

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

СПИСОК-2022

МАТЕРИАЛЫ 9-Й ВСЕРОССИЙСКОЙ НАУЧНОЙ КОНФЕРЕНЦИИ ПО ПРОБЛЕМАМ ИНФОРМАТИКИ

27–29 апреля 2022 г.
Санкт-Петербург

Санкт-Петербург

2022

УДК 004 (063)
C72

*Печатается по рекомендации
кафедры системного программирования
Санкт-Петербургского государственного университета*

Материалы 8-й всероссийской научной конференции по проблемам информатики СПИСОК-2022. 27–29 апреля 2022 г. Санкт-Петербург — СПб.: ВВМ, 2022. — 504 с.

ISSN 2310-4724, 2310-4732

C72 Тематика сборника затрагивает широкий круг актуальных проблем теоретической и прикладной математики и информатики. Слово «СПИСОК», ставшее названием конференции, — это не только обозначение фундаментальной структуры данных, но и сокращение от названий трех направлений исследований: Системное Программирование, Интеллектуальные Системы, Обеспечение Качества. результаты исследований в этих областях являются значительной частью того знания, которое в настоящее время можно назвать словом «информатика».

Для студентов и аспирантов естественно-научных специальностей.

ISSN 2310-4724 (инпринт)

ISSN 2310-4732 (онлайн)

© Авторы, 2022
© Издательство «ВВМ», 2022



Вообще-то идея этой конференции принадлежит не мне, она была придумана в Екатеринбурге много лет назад. И потом авторы этой идеи перебрались в Санкт-Петербург и забрали идею конференции с собой. СПИСОК – это аббревиатура, из которой я помню значение первых двух букв: Системное Программирование. Преподаватели нашей кафедры считают, что студенты должны заниматься научной работой и уметь рассказывать о своих результатах. Лучшие результаты мы публикуем в трудах конференции. Этот год не станет исключением.

С уважением,  А. Н. Терехов
д.ф.-м.н., профессор, заведующий кафедрой
системного программирования СПбГУ
Президент ООО «Ланит-Терком»

Системное программирование



Терехов Андрей Николаевич

д.ф.-м.н., профессор, заведующий Кафедрой системного
программирования СПбГУ, президент ООО «Ланит-Терком»

Эмулятор ЕМ1.

Задачи, возможности и результаты

Баклановский М.В., старший преподаватель кафедры системного программирования СпбГУ, baklanovsky@mail.ru

Лагунов Л.Л., старший разработчик ПО ООО «Софтком», leonid.lagunov@softcom.su

Сибиряков А.Е., старший преподаватель департамента математики, механики и компьютерных наук УрФУ, a.sibiryakov@urfu.ru

Ханов А.Р. , старший преподаватель кафедры системного программирования СпбГУ, st036451@student.spbu.ru

Аннотация

Рассмотрены основные подходы, использованные при разработке эмулятора ЕМ1. Предложены способы эмуляции внешних устройств и многопроцессорных систем. Описана методика тестирования эмуляторов и процессоров на соответствие с документацией.

Ключевые слова

Внешние устройства, симуляция, тестирование эмуляторов, эмуляторы, эмуляция процессоров.

Введение

Эмулятором называют программу, которая имитирует поведение реального вычислителя. Эмулируются наборы API-вызовов, инструкции целевых процессоров, микроархитектура вычислителей либо отдельные логические элементы [1,2].

Известно большое количество эмуляторов для популярных процессоров [3,4,5,6]. Практически все они (даже популярный эмулятор Qemu) обладают рядом серьезных недостатков:

- добавление новых процессоров требует существенных

изменений/добавлений в коде и не всегда возможно в разумные сроки даже при наличии в открытом доступе исходных текстов и соответствующей лицензии;

- скорость работы эмулируемых программ обычно значительно ниже чем на реальных вычислителях (иногда до сотен раз);
- многие разработки не являются отечественными и могут содержать недокументированные фрагменты кода.

Разработчики редких вычислителей, таких как DSP-ядра, процессоры для микроконтроллеров, иногда включают эмуляторы в коммерческие пакеты поставки. В отсутствие таковых приходится писать новый эмулятор "с нуля". При этом такая программа может не обладать всеми необходимыми возможностями по трассировке и отладке и потребовать существенных трудовых затрат. Неэффективная реализация (обычно на ЯВУ) не позволяет выполнять эмулируемые программы с достаточной скоростью. Немаловажную роль здесь играет и точность эмуляции - эмулятор может не соответствовать документации, а также не достаточно точно эмулировать микроархитектуру (например такты процессора), что может быть существенным при решении некоторых задач, работающих с аппаратурой. Известны попытки предложить универсальные решения данных проблем.

Система WInter [6] предназначена для разработки эмуляторов целых вычислительных систем. В ней поддержано несколько известных процессоров для микроконтроллеров, а также процессор общего назначения Intel 8086. Однако все возможности отладки были реализованы внутри собственной графической оболочки, что не позволяет подключать сторонние программы. Кроме того, проект устарел и не поддерживается с 2003 года.

Еще одной такой попыткой можно считать известный эмулятор qemu. Однако его архитектура не подходит для тонкой эмуляции микроархитектуры: работы конвейеров, таймеров, шин обмена данными. Поскольку реализация кодов инструкций производится на языке C++, такой эмулятор будет медленным.

Большая часть современных работ на данную тему посвящена эмуляции вычислительных систем на FPGA, что лишает их возможности работы на компьютерах общего назначения

Многие эмуляторы разрабатываются для уже существующих гостевых аппаратных платформ и не содержат встроенных механизмов низкоуровневого тестирования на соответствие между документацией, эмулятором и целевой вычислительной системой.

Постановка задачи

Важной [7] прикладной задачей представляется разработка мультиэмулатора, обладающего следующими характеристиками:

- быстрая (2-3 месяца) разработка эмуляторов для еще не выпущенных процессоров и включение их в мультиэмулатор;
- возможность начать полноценную разработку прикладного ПО параллельно с производством процессоров, значительно уменьшив при этом время готовности решений;
- использование дополнительных возможностей, предоставляемых эмуляцией, при отладке и оптимизации разрабатываемого ПО;
- высокая скорость выполнения эмулируемых программ;
- максимально точная эмуляция времени выполнения;
- использование внешних по отношению к пространству вычислений средств отладки и трассировки;
- встроенный механизм тестирования эмулятора;
- простое решение для эмуляции внешних устройств;
- API для интеграции в различные графические среды;
- эмуляция многопроцессорных систем, состоящих в том числе и из процессоров разных архитектур;
- собственная хорошо документированная разработка, позволяющая вносить существенные изменения и дополнения в алгоритм работы эмуляторов.

Основные решения

Создание нового эмулятора каждый раз "с нуля" предполагает серьезные затраты на разработку, поэтому при создании очередного эмулятора для EM1 используются готовые, заранее разработанные модули:

- парсеры команд и мнемоник;
- поддержка протокола GDB;
- API для работы с консольной или графической оболочкой;
- инструменты интеграции, отладки и трассировки.

Описание каждого эмулируемого процессора и правила разбора опкодов инструкций сохраняются в специализированной базе данных [8]. Обработчики инструкций разрабатываются на языке ассемблера хост-компьютера (Intel x64) в ручном или полуавтоматическом режиме. Исполняемый файл эмулятора автоматически генерируется с использованием данных из БД, модулей и обработчиков инструкций. Для

увеличения скорости работы парсеров при разработке эмулятора для каждого очередного процессора выполняется тестирование различных вариантов разбора, в финальную сборку попадает самый скоростной вариант.

EM1 содержит встроенный GDB-сервер. В качестве клиентов могут использоваться популярные средства отладки, поддерживающие протокол GDB, например - Visual Studio Code, IDA Pro, GDB-Multiarch, PEDA, Angr, Radare2. Подключение к эмулятору выполняется по сети с любого компьютера в Интернете.

Предлагаемая система тестирования эмуляторов и микропроцессоров (ТЭМП) содержит тестовую базу, разрабатываемую в соответствии с документацией на процессор. ТЭМП использует специализированный язык описания тестов, позволяющий выполнять тесты в пакетном режиме. Эмулятор перед сдачей в эксплуатацию проходит полное тестирование. Позднее такое же тестирование пройдут образцы процессора. Это даст нам уверенность в том, что программы на процессоре будут выполняться точно так же, как на эмуляторе.

EM1 эмулирует не только алгоритмы исполнения инструкций, но и времена их выполнения в тактах процессора. До момента готовности образцов процессора время выполнения каждой инструкции вычисляется в соответствии с документацией. В дальнейшем в ходе тестирования процессоров выполняется дополнительное изучение времен выполнения инструкций и их комбинаций, исследуется влияние конвейеризации и вносятся изменения в соответствии с полученными результатами.

Достигнутая точность эмуляции позволяет выполнять на эмуляторе разработки практически любой сложности, включая, например, компиляторы и другое сложное ПО. Используемый при разработке ПО язык программирования большого значения не имеет, достаточно применять кросс-компилятор, максимально поддерживающий возможности процессора.

Эмуляция внешних устройств является одной из сложнейших задач при создании эмуляторов. Основная причина состоит в том,, что разработчики процессоров при разработке SoC чаще всего пользуются готовыми устройствами внешних производителей. Документация на устройства не всегда доступна в полном объеме и часто не содержит информации, достаточной для разработки полноценных драйверов. В случае разработки в таких условиях драйвер устройства для эмулятора с большой вероятностью не сможет в дальнейшем использоваться при работе с реальной аппаратурой и подобная эмуляция устройств становится

неэффективной.

Вместо эмуляции работы устройств предлагаются выполнять эмуляцию API-функций, работающих с устройствами. Основные преимущества такого подхода:

- разработка прикладного ПО может выполняться при отсутствии ОС;
- функции реализуются один раз для каждого конкретного устройства и в дальнейшем переиспользуются;
- в случае если эти функции по логике работы совпадают с API той ОС, которая будет использована, затраты на доработку ПО под реальный вычислитель могут быть сведены к нулю;

К эмулятору прилагаются библиотеки для специальной сборки исполняемых файлов. Библиотеки содержат в том числе и необходимые для работы ПО функции (например printf и malloc). Собранный с такими библиотеками исполняемый файл работает одинаково как под эмулятором так и на "голом" железе.

Один запущенный эмулятор процессора (ЭП) эмулирует один процессор. На одном или нескольких компьютерах в сети могут быть запущены несколько эмуляторов одного и того же или разных процессоров. Отдельный эмулятор может быть запущен в режиме эмуляции многопроцессорного вычислителя. Такой эмулятор многопроцессорной системы (ЭМПС) связывает между собой эмуляторы различных (в т.ч. и по архитектуре) процессоров и обеспечивает обмен информацией между ними используя общую память. Каждый ЭП имеет несколько интерфейсов - Init, Cons и Link - обеспечивающих возможности подключения. Схема сборки многопроцессорного эмулятора приведена на Рис.1.

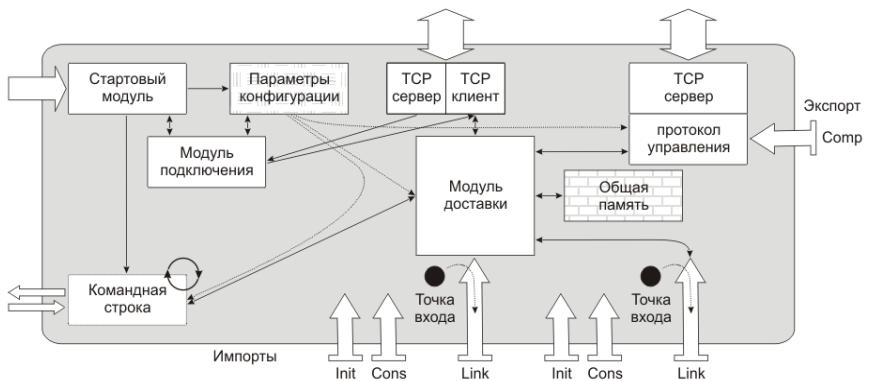


Рис. 1. Схема сборки многопроцессорного эмулятора.

ЭМПС позволяет отлаживать и оптимизировать одновременную (параллельную) работу ПО на разных процессорах. При этом каждый ЭП может быть запущен в отдельном потоке на отдельном ядре или на отдельном компьютере.

Внешняя по отношению к вычислительному пространству гостевого процессора эмуляция предоставляет уникальные возможности для отладки и трассировки, отсутствующие в реальных процессорах, например:

- точки останова без внесения изменений в секции кода (в т.ч. и в ПЗУ);
- управление отладкой в зависимости от состояния или изменения областей данных (в любом количестве);
- наблюдение за служебными регистрами эмулируемого процессора даже если они не доступны пользователю в user mode.

Разработанный механизм наблюдения за процессом вычислений позволяет не только отлаживать код, работающий в привилегированных режимах, но и отлаживать программы-отладчики.

Заключение

К настоящему времени с использованием предложенной технологии разработаны эмуляторы процессоров из семейств ARM, STM32, MIPS32, КОМДИВ64, DSP TMS320C6x.

Скорость работы эмуляторов настолько высока что, например, известный тест LINPACK в эмуляторе MIPS32 на компьютере с далеко не

современным процессором Intel PentiumG 4го поколения выполняется примерно в 6 раз быстрее чем на реальном микрокомпьютере Onion Omega2+.

EM1 работает в среде ОС Linux (Ubuntu, Arch, Debian, Astra и др.) или Windows 64 (в т.ч. под WSL).

Режимы внешней трассировки и отладки в сочетании с ЭМПС позволяют разрабатывать и оптимизировать программные решения со сложной синхронизацией.

Реализована концепция “одного бинарника”, когда любой исполняемый файл одинаково работает как на эмуляторе, так и на процессоре без ОС.

Все разработки выполнены "с нуля" без использования готовых (чужих) программ или их фрагментов. Основной язык программирования — ассемблер.

Литература

1. A. Svensson, "Software primitives for emulation of multiprocessor architectures," Twenty-Third Annual Hawaii International Conference on System Sciences, 1990, pp. 48-56 vol.1, doi: 10.1109/HICSS.1990.205098.
2. Речистов Г.С. и др., <https://atakua.org/w/images/simulation-lectures-latest.pdf> [дата просмотра 25.04.2022].
3. https://en.wikipedia.org/wiki/List_of_emulators [дата просмотра 25.04.2022].
4. https://en.wikipedia.org/wiki/List_of_computer_system_emulators [дата просмотра 25.04.2022].
5. QEMU's documentation, <https://www.qemu.org/docs/master/> [дата просмотра 25.04.2022].
6. Интегрированная среда разработки программного обеспечения встроенных систем Winter, <https://newit.gsu.by/ru/winter/> [дата просмотра 25.04.2022].
7. Баклановский М.В., Оносовский В.В., Терехов А.Н., Тимохин Д.В., Мартынов В.И., Милюченко С.Г., Нестеренко А.Н., Сгонников А.С., Сгонников Д.С., “Инструментальная среда разработки программно-аппаратных комплексов”, XLV академические чтения по космонавтике, Москва, 30 марта – 02 апреля 2021 года, стр. 523-526.
8. Баклановский М.В., Сибиряков А.Е., Нестеренко А.Н., Дмитриев В.С., Сгонников Д.С., “Эмуляторы многопроцессорных вычислительных систем и отладка на уровне системы”, XLV академические чтения по космонавтике, Москва, 30 марта – 02 апреля 2021 года, стр. 517-519.

Фундаментальная информатика



Косовская Татьяна Матвеевна

д.ф.-м.н., профессор, заведующая Кафедрой информатики

ИНСТРУМЕНТАРИЙ ДЛЯ РАБОТЫ С ДИНАМИЧЕСКИМИ ОТНОШЕНИЯМИ

Федорченко Л. Н., к. ф.-м. н., доцент кафедры информатики СПбГУ,
LNF@iias.spb.su

Криворучко Д. И., студент группы 19.Б10мм, кафедра системного
программирования СПбГУ, denis.krvrchk@gmail.com

Аннотация

Динамическое отношение — это отношение, множество элементов которого изменяется в зависимости от времени. В докладе дается краткое описание прототипа программного инструмента для работы с динамическими отношениями, описывающими практические ситуации. Представлены примеры представления динамических отношений.

Введение

На сегодняшний день работа во многих прикладных областях знаний предполагает взаимодействие с данными, которые имеют динамическую структуру и изменяются в зависимости от времени. Для обработки информации подобного рода часто используют алгебраические модели, благодаря которым можно выразить динамику предметной области, например, динамические отношения, то есть отношения на множестве объектов, зависящие от времени. Для анализа ситуаций на практике и принятия решений, удобно визуализировать динамические отношения.

Описываемый в докладе инструмент предоставляет возможность в автоматическом режиме проводить операции над заданными динамическими отношениями, визуализировать и анализировать их и получать результат в виде программного модуля на языке высокого уровня для его дальнейшего использования в практической задаче.

Представление динамического отношения в виде циклограммы

Программная система реализована на языке Python3 с использованием библиотек PyQt5 [2] для пользовательского интерфейса и Matplotlib [3] для отображения динамических отношений в виде циклограмм (набора импульсов), графиков. Алгебра динамических отношений описана в работе [1]. На текущий момент реализованы 3 бинарные операции (объединение, пересечение и вычитание циклограмм).

Пример отображения бинарного динамического отношения представлен на рисунке 1. По оси OX располагаются моменты времени. По оси OY — элементы отношения. Закрашенные прямоугольники соответствуют моментам времени, когда элементы принадлежат отношению (имеют значение "истина").

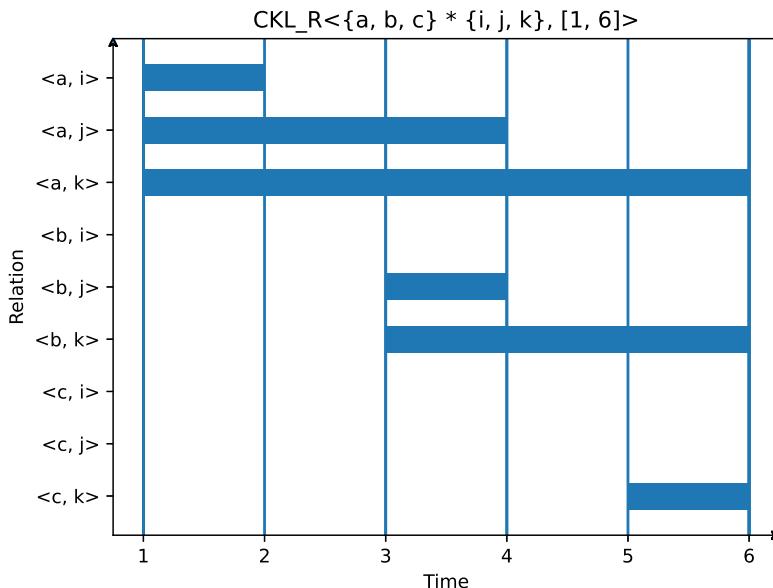


Рис. 1: Пример отображения динамического отношения

Графическая часть приложения содержит два окна. При запуске программы, на экране появляется главное окно, в котором можно открыть циклограммы из файла, сохранить циклограммы в файл, поме-

нять местами циклограммы, а также отменить или восстановить действие и применить операции из списка. Данная функциональность реализована в виде выпадающего меню. Главное окно представлено на рисунке 2. Результатом реализованных операций является циклограмма на рисунке 3. Она отображается в отдельном окне, которое поддерживает функции сохранения в файл или открытия циклограммы в главном окне.

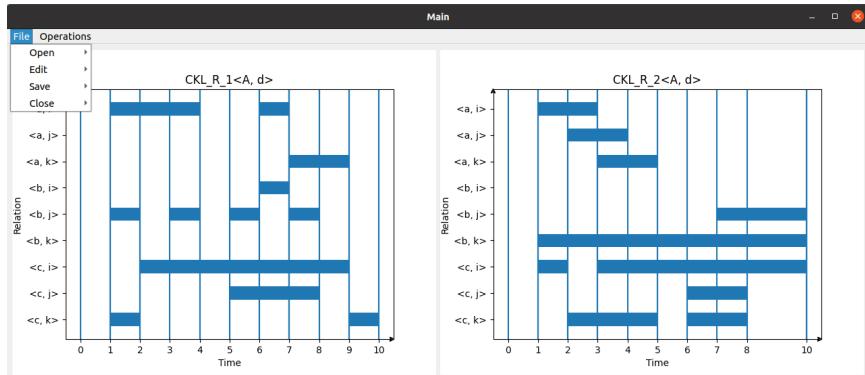


Рис. 2: Основное окно

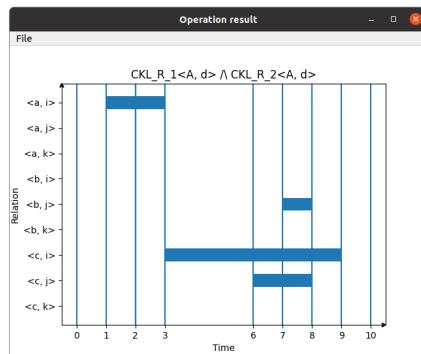


Рис. 3: Окно результата применения операции (пересечения)

Заключение

Полученный инструмент содержит простые операции над динамическими отношениями, что позволяет решать некоторый класс прикладных задач обработки динамических знаний с помощью представления их в виде системы динамических отношений. Однако, гибкая архитектура позволяет добавлять реализацию новых операций и других возможностей.

Список литературы

- [1] Динамические отношения в задачах обработки знаний. URL: <https://rureportal.spbu.ru/ru/publications/динамические-отношения-в-задачах-обработки-знаний> (дата обращения 01.02.2022).
- [2] Документация PyQt5. URL: <https://doc.qt.io/qtforpython/> (дата обращения 14.02.2022).
- [3] Документация matplotlib. URL: <https://matplotlib.org/> (дата обращения 23.02.2022).

СТЕПЕНИ ДВОЙКИ И ВЕЩЕСТВЕННО-ЦЕЛОЧИСЛЕННАЯ ЛИНЕЙНАЯ ЭЛИМИНАЦИЯ КВАНТОРОВ

И.Д Мишурев, студент СПбГУ, ilucha556@gmail.com

Аннотация

В докладе сделан обзор известных алгоритмов и подходов для элиминации кванторов, явного алгоритма элиминации кванторов для арифметики целых чисел с предикатом возведения двойки в степень, а так же доказаны теоремы о неразрешимости двух расширений арифметики Семёнова. Исследование нацелено на поиск алгоритмов элиминации кванторов, которые могут стать расширением пакета RedLog.

Введение

В современном мире арифметика Пресбургера нашла применение не только в математике, но и во многих прикладных областях [2]. Арифметика Пресбургера представляет собой элементарную теорию целых чисел с нулём, единицей, сложением и отношением порядка.

Уже долгое время ведутся исследования как самой арифметики, так и вопросов, связанных с различными ее расширениями. Большая часть исследований была посвящена проблемам сложности и разрешимости этой теории, а в дальнейшем и различных расширений этой теории.

Есть большое количество результатов, в которых доказывается и явно приводится алгоритм элиминации кванторов для различных структур. Обработка на компьютере формул, содержащих кванторы, требует больших затрат времени, так как каждый квантор рассматривается как цикл по всем значениям переменных, принимающих значения из предметной области. В связи с этим задача элиминации (удаления) кванторов той или иной теории является весьма актуальной в теоретической информатике. Множество из этих алгоритмов реализовано в системе Redlog [1].

Важной задачей является нахождение достаточно выразительных структур (в частности, расширений арифметики Пресбургера), для которых существует алгоритм элиминации кванторов. Такие алгоритмы затем могут быть внедрены в систему RedLog для отдельных классов формул, для которых применение элиминации кванторов оказывается

достаточно эффективным. Такого рода чисто теоретические вопросы представляют неменьший интерес.

В статье [3] Ф.Пуан привела явный алгоритм элиминации кванторов для расширения арифметики Пресбургера $\langle \mathbb{N}, \leq, +, 2^x \rangle$. В этой же статье Ф.Пуан было показано что любая формула вида $\exists x\theta(x, \bar{y})$, где $\theta(x, \bar{y})$ – это конъюнкция атомарных формул, эквивалентна некоторой открытой формуле. Также, было показано что можно ограничить переменную x термом в \bar{y} .

Основные определения

В этом разделе будут даны основные определения и обозначения, а в дальнейшем будут получены некоторые утверждения о неразрешимости.

Определение 1. Через L_σ обозначается язык первого порядка сигнатуры σ , а формула языка L_σ называется L -формулой.

Определение 2. Через $\exists L_\sigma$ будем обозначать экзистенциальные формулы, а через $\forall L_\sigma$ - универсальные L_σ -формулы.

Определение 3. Множество всех замкнутых L -формул, истинных в структуре $\langle M; \sigma \rangle$, называется L -теорией структуры $\langle M; \sigma \rangle$ и обозначается $L - Th\langle M; \sigma \rangle$.

Определение 4. $\exists L_\sigma$ -выразимые в структуре $\langle M; \sigma \rangle$ отношения называются экзистенциально выражимыми в структуре $\langle M; \sigma \rangle$.

Определение 5. $Th\langle M; \sigma \rangle$ есть элементарная теория структуры $\langle M; \sigma \rangle$, а $\exists Th\langle M; \sigma \rangle$ экзистенциальная теория структуры $\langle M; \sigma \rangle$.

Определение 6. Алгоритмом элиминации кванторов для языка L_σ в структуре $\langle M; \sigma \rangle$ называется алгоритм, который по всякой L_σ -формуле вида $\exists x\phi(x, y_1, \dots, y_n)$, где $\phi(x, y_1, \dots, y_n)$ бескванторная L_σ -формула, строит эквивалентную ей в этой структуре бескванторную L_σ -формулу $\psi(y_1, \dots, y_n)$.

Алгоритм элиминации кванторов позволяет построить по всякой L_σ -формуле эквивалентную в соответствующей структуре бескванторную L_σ -формулу.

Неразрешимость некоторых расширений арифметики Семёнова

С прикладной точки зрения полезно иметь достаточно выражительные структуры с алгоритмом элиминации кванторов. Однако, в действительности, не для всех интересующих нас теорий такой алгоритм существует. В рамках данной работы были доказаны две теоремы о неразрешимости экзистенциальных теорий структур $\langle \mathbb{N}, 0, 1, +, 2^x, | \rangle$ и $\langle \mathbb{R}, 0, 1, +, -, 2^x, =, <, Z \rangle$.

Напомним, что $|$ есть двухместный предикат делимости целых чисел, а Z соответствует свойству "быть целым числом".

Расширение предикатом делимости

Определение 7. Заданный на натуральных числах двухместный предикат T является предикатом степенного роста, если существуют положительные рациональные числа c, d, c_1, d_1 , такие что $d_1 > 1$ и

1. Каковы бы ни были натуральные числа x, y , если имеет место $T(x, y)$ и $x > 0$, то $y \leq cx^d$
2. для всякого натурального числа x существует натуральное число y , такое, что имеют место $y \geq c_1 x^{d_1}$ и $T(x, y)$.

Теорема 1 ([4]). Пусть есть структура $\langle \mathbb{N}; 0, 1, +, |, T \rangle$, где T - предикат степенного роста. Тогда в этой структуре выразимо умножение.

Теорема 2. Экзистенциальная теория структуры $\langle \mathbb{N}; 0, 1, +, 2^x, | \rangle$ неразрешима.

Доказательство. Для доказательства данного утверждения воспользуемся результатами из работы Н.К.Косовского [4].

В данной структуре выразим предикат равенства, а именно $y = x \iff y|x \wedge x|y$, полагая что верно $0|0$.

Выразим в структуре предикат, обозначающий длину записи числа, обычно определяемый как:

$$\|x\| = \begin{cases} 1 & \text{если } x = 0 \\ [\log x] + 1 & \text{если } x > 0 \end{cases}$$

В структуре $\langle \mathbb{N}; 0, 1, +, 2^x, | \rangle$ выразим формально данный предикат $||y|| \leq 3||x||$:

1. $y \leq x \iff \exists z(x = y + z)$
2. $y < x \iff y \leq x \wedge \neg(y = x)$
3. $y = ||x|| \iff \exists z(2^y > x \geq 2^z \wedge z + 1 = y)$
4. $||y|| \leq 3||x|| \iff \exists t_1 \exists t_2(t_1 = ||y|| \wedge t_2 = ||x|| \wedge t_1 \leq 3t_2)$

По определению, предикат $||y|| \leq 3||x||$ является предикатом степенного роста. Для доказательства этого достаточно взять в определении 7 значения $c = 2^3, d = 1, c_1 = 2^{-3}, d_1 = 1$. Из теоремы 1 тогда следует что данная структура содержит в себе арифметику натуральных чисел со сложением и умножением. Из неразрешимости десятой проблемы Гильберта [5] следует что такая структура неразрешима.

Расширение предикатом быть целым

Рассмотрим ещё одно расширение арифметики Семёнова. Несложно показать, что если дополнить структуру из теоремы В.Вайспеннигга функцией 2^x , то это позволит выразить в ней умножение с помощью экзистенциальной формулы, что приведёт к неразрешимости уже экзистенциальной теории этой структуры.

Теорема 3. Экзистенциальная теория структуры $\langle \mathbb{R}, Z, 0, 1, +, -, 2^x, =, < \rangle$, где Z — одноместный предикатный символ, который интерпретируется с помощью свойства «быть целым числом», неразрешима.

Доказательство. Для доказательства утверждения сначала заметим, что для всякого положительного целого y имеет место $\exists x(y = 2^x)$, где x является вещественной переменной. Теперь несложно видеть, что можно выразить умножение натуральных чисел следующим образом

$$\begin{aligned} z = x \cdot y \wedge x > 0 \wedge y > 0 \wedge Z(z) \wedge Z(x) \wedge Z(y) \Leftrightarrow \\ \exists u \exists v (z = 2^{u+v} \wedge x = 2^u \wedge y = 2^v) \wedge Z(z) \wedge Z(x) \wedge Z(y) \end{aligned}$$

Таким образом, ввиду того, что в сигнатуре имеется унарный минус, в данной структуре экзистенциально выразим график функции умножения. Для этого достаточно рассмотреть различные знаки переменных x, y, z . Тогда данная структура содержит в себе арифметику натуральных чисел со сложением и умножением. Из неразрешимости десятой проблемы Гильберта [5] следует что такая структура неразрешима.

Заключение

В рамках данной работы были доказаны две теоремы о неразрешимости экзистенциальных теорий структур $\langle \mathbb{N}, 0, 1, +, 2^x, | \rangle$ и $\langle \mathbb{R}, 0, 1, +, -, 2^x, =, <, Z \rangle$. Следовательно, для данных структур не существует алгоритмов элиминации кванторов. Эти результаты приводят к следующему вопросу. Пусть $2^{[l]}$ — это функция возведения 2 в степень целой части числа, можно ли тогда построить алгоритм элиминации кванторов для $\langle \mathbb{R}, 0, 1, +, -, 2^{[x]}, =, <, Z \rangle$. Если ожидать положительного решения этого вопроса, то первой задачей является преобразование алгоритма Ф.Пуан [3] так, чтобы новый алгоритм элиминации кванторов работал со структурой $\langle \mathbb{Z}, =, +, 2^x \rangle$. Дальнейшие исследования планируется продолжать с решения этой задачи.

Список литературы

- [1] Andreas Dolzmann, Thomas Sturm. Redlog: Computer algebra meets computer logic // SIGSAM Bull. — 1997. — Т. 31 С. 2–9
<https://doi.org/10.1145/261320.261324>
- [2] Christoph Haase. A survival guide to presburger arithmetic // ACM SIGLOG News. — 2018. — Т. 5 С. 67–82
<https://doi.org/10.1145/3242953.3242964>
- [3] Francoise Poin. On the expansion $(n, +, 2x)$ of presburger arithmetic // — 2007. — <https://webusers.imj-prg.fr/~francoise.point/papiers/Pres.pdf>
- [4] Н. К. Косовский. О решении систем, состоящих одновременно из уравнений в словах и неравенств в длинах слов // Исследования по конструктивной математике и математической логике. VI, Зап. научн. сем. ЛОМИ. — 1974. — Т. 40 С. 24–29
<http://mi.mathnet.ru/znsl2678>
- [5] Ю. В. Матиясевич. Диофантовость перечислимых множеств // Докл. АН СССР. — 1970. — Т. 191 С. 279–282
<http://mi.mathnet.ru/dan35274>

Параллельные алгоритмы и вэйвлетная обработка числовых потоков



Демьянович Юрий Казимирович

д.ф.-м.н., профессор, заведующий кафедрой параллельных алгоритмов
СПбГУ

Об устойчивости всплесковых разложений

Демьянович Ю.К., профессор кафедры параллельных алгоритмов СПбГУ,
y.demjanovich@spbu.ru

Иванцова О. Н., доцент кафедры параллельных алгоритмов СПбГУ,
o.ivancova@spbu.ru

Аннотация

В работе устанавливаются критерии устойчивости сплайн-всплесковых разложений. Полученные критерии применимы к всплесковым разложениям нулевого, первого и второго порядков.

Введение

Всплесковые (вэйвлетные) разложения широко используются при обработке числовых информационных потоков; объемы таких потоков постоянно возрастают, и это является стимулом к дальнейшему развитию теории всплесков ([1], [2]). Используемый в данной работе подход к построению всплесков основывается на применении аппроксимационных соотношений, так что автоматически обеспечивается эффективная аппроксимация (чаще всего, она асимптотически оптимальна по N -поперечнику стандартных компактов). В противоположность классическим вэйвлетам упомянутый подход позволяет без дополнительных сложных исследований использовать неравномерную сетку (как конечную, так и бесконечную), что весьма важно для экономии компьютерных ресурсов в случае появления сингулярных изменений рассматриваемых потоков. Кроме того, при построении всплесковых разложений в многомерном случае (и даже на произвольном дифференцируемом многообразии) могут применяться известные конечно-элементные аппроксимации (Куранта, Зламала и т.п.), что существенно расширяет возможности использования всплесковых разложений. В классическом случае большую трудность представляет построение всплескового (вэйвлетного) базиса в том или ином функциональном пространстве (часто в $L_2(\mathbb{R}^1)$); используемый здесь подход не требует предварительного построения всплескового базиса (при желании этот базис может быть получен после проведения основных исследований). С другой стороны, знание всплескового базиса позволяет достичь существенной экономии компьютерных и сетевых ресурсов. Заметим, что для получения упомянутой экономии не нужен всплесковый базис в пространстве функций с континуальной областью определения, достаточно лишь получить подходящий базис для пространства всплесковых

числовых потоков, но для этого необходимо все построения выполнять без использования функций с континуальной областью определения.

Цель данной работы состоит в том, чтобы в рамках упомянутого подхода рассмотреть условия устойчивости сплайн-всплескового разложение относительно возмущения исходных данных. Для формул декомпозиции исходными данными является, так называемый, исходный поток. При рассмотрении формул реконструкции исходными данными являются основной и всплесковый потоки.

Предварительные сведения

Рассмотрим линейное пространство \mathbb{L} функций, определенных на интервале $(\alpha, \beta) \subset \mathbb{R}^1$. На интервале (α, β) введем сетку

$$X : \dots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 \dots, \quad (1.1)$$

$$\lim_{j \rightarrow -\infty} x_j = \alpha, \quad \lim_{j \rightarrow +\infty} x_j = \beta. \quad (1.2)$$

Пусть $\varphi(t) \stackrel{\text{def}}{=} (\varphi_0(t), \varphi_1(t), \dots, \varphi_m(t))$ — вектор-функция с компонентами, принадлежащими пространству \mathbb{L} : $\varphi_i \in \mathbb{L}, i \in \{0, 1, \dots, m\}$.

Рассмотрим множество G линейных функционалов $g^{(s)} \in \mathbb{L}^*$, $G \stackrel{\text{def}}{=} \{g^{(s)}\}_{s \in \mathbb{Z}}$, со свойством

$$\text{supp } g^{(s)} \subset (x_s, x_{s+1}) \quad \forall s \in \mathbb{Z}. \quad (1.3)$$

Результат действия функционала $g^{(s)}$ на функцию $u \in \mathbb{L}$ обозначим остройми скобками $\langle g^{(s)}, u \rangle$. Упомянутые скобки будем использовать также для вектор-столбца с числовыми компонентами

$$\langle g^{(s)}, \varphi \rangle \stackrel{\text{def}}{=} (\langle g^{(s)}, \varphi_0 \rangle, \langle g^{(s)}, \varphi_1 \rangle, \dots, \langle g^{(s)}, \varphi_m \rangle)^T.$$

Условие невырожденности. Полная цепочка векторов

Предположим, что выполнено условие

$$\det(\langle g^{(s)}, \varphi \rangle, \langle g^{(s+1)}, \varphi \rangle, \dots, \langle g^{(s+m)}, \varphi \rangle) \neq 0 \quad \forall s \in \mathbb{Z}. \quad (2.1)$$

Условие (2.1) называется *условием невырожденности*.

Рассмотрим цепочку $m + 1$ -мерных векторов $\dots, \mathbf{a}_{-2}, \mathbf{a}_{-1}, \mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \dots$. Если при каждом фиксированном $s \in \mathbb{Z}$ система векторов $\{\mathbf{a}_s, \mathbf{a}_{s+1}, \dots, \mathbf{a}_{s+m+1}\}$ линейно независимая, то цепочка векторов $\dots, \mathbf{a}_{-2}, \mathbf{a}_{-1}, \mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \dots$ называется *полной*.

По определению положим

$$\mathbf{a}_s = \langle g^{(s)}, \varphi \rangle. \quad (2.2)$$

Цепочка векторов (2.2) — полная, поскольку из предположения (1.3), (2.1) следует линейная независимость системы векторов $\mathbf{a}_s, \mathbf{a}_{s+1}, \dots, \mathbf{a}_{s+m+1}$ при любом фиксированном целом s .

Замечание. Если \mathbb{L} является пространством непрерывных функций на интервале (α, β) , и $\langle g^{(s)}, u \rangle \stackrel{\text{def}}{=} u(x_s), s \in \mathbb{Z}$, а система функций $\{\varphi_0, \varphi_1, \dots, \varphi_m\}$ является системой Чебышева на (α, β) , то свойство (2.1) выполнено. Заметим, что примером системы Чебышева на любом интервале $(\alpha, \beta) \subset \mathbb{R}^1$ является система $\{1, t, \dots, t_m\}$.

Апроксимационные соотношения и пространство минимальных сплайнов

Теперь рассмотрим соотношения

$$\sum_{i=k-m}^k \mathbf{a}_i \omega_i(t) = \varphi(t) \quad \forall t \in (x_k, x_{k+1}) \quad \forall k \in \mathbb{Z}, \quad (3.1)$$

$$\text{supp } \omega_s \subset [x_s, x_{s+m+1}] \quad \forall s \in \mathbb{Z}. \quad (3.2)$$

Соотношения (3.1) – (3.2) называются *аппроксимационными соотношениями*. Ввиду предположения (2.1) из аппроксимационных соотношений однозначно определяются функции $\omega_i(t)$. При предположениях (2.1) эти функции образуют линейно независимую систему. Они называются координатными сплайнами.

Предположим, что $\omega_s \in \mathbb{L}$. Из формул (1.3), (2.1) – (2.2), (3.1) – (3.2) следуют соотношения биортогональности системы функционалов $\{g^{(j)}\}$ по отношению к системе координатных сплайнов $\{\omega_s\}_{s \in \mathbb{Z}}$:

$$\langle g^{(j)}, \omega_s \rangle = \delta_{j,s} \quad \forall j, s \in \mathbb{Z}.$$

Рассмотрим линейное пространство \mathbb{S} , состоящее из линейных комбинаций координатных сплайнов, $\mathbb{S} = \mathbb{S}(X, \varphi) \stackrel{\text{def}}{=} \mathcal{L}\{\omega_s\}_{s \in \mathbb{Z}}$, где через \mathcal{L} обозначена линейная оболочка функций, находящихся в фигурных скобках. Пространство (X, φ) называется *пространством минимальных сплайнов*.

Вложенное пространство минимальных сплайнов

Пусть \tilde{X} — подмножество сетки X такое, что

$$X : \dots < \tilde{x}_{-2} < \tilde{x}_{-1} < \tilde{x}_0 < \tilde{x}_1 < \tilde{x}_2 < \dots,$$

$$\lim_{j \rightarrow -\infty} \tilde{x}_j = \alpha, \quad \lim_{j \rightarrow +\infty} \tilde{x}_j = \beta, \quad \tilde{X} \subset X.$$

Предположим, что $\{\tilde{g}^{(i)}\}_{i \in \mathbb{Z}}$ — система линейных функционалов со свойствами

$$\text{supp } \tilde{g}^{(i)} \subset (\tilde{x}_i, \tilde{x}_{i+1}) \quad \det(\langle \tilde{g}^{(i)}, \varphi \rangle) \neq 0 \quad \forall i \in \mathbb{Z}. \quad (4.1)$$

Рассмотрим функции $\tilde{\omega}_i$, связанные с новой сеткой \tilde{X} аналогично предыдущему, и введем линейное пространство минимальных сплайнов $\tilde{\mathbb{S}}$, ассоциированное с новой сеткой $\tilde{\mathbb{S}} = \mathbb{S}(\tilde{X}, \varphi)$.

Предположим, что пространство $\tilde{\mathbb{S}}$ является подпространством пространства \mathbb{S} , так что

$$\tilde{\mathbb{S}} \subset \mathbb{S} \subset \mathbb{L}. \quad (4.2)$$

Операция проектирования на вложенное пространство

Из соотношений (4.1) следует, что система функционалов $\{\tilde{g}^{(s)}\}_{s \in \mathbb{Z}}$ биортогональна системе функций $\{\tilde{\omega}_i\}_{i \in \mathbb{Z}}$,

$$\langle \tilde{g}^{(s)}, \tilde{\omega}_i \rangle = \delta_{s,i} \quad \forall s, i \in \mathbb{Z}. \quad (5.1)$$

Рассмотрим операцию P проектирования пространства \mathbb{S} на подпространство $\tilde{\mathbb{S}}$, которая действует по формуле

$$Pu \stackrel{\text{def}}{=} \sum_{s \in \mathbb{Z}} \langle \tilde{g}^{(s)}, u \rangle \omega_s, \quad \forall u \in \mathbb{S}(X, \varphi). \quad (5.2)$$

Ясно, что если $t \in (\tilde{x}_k, \tilde{x}_{k+1})$ фиксировано, то правая часть формулы (5.2) имеет не более $m+1$ слагаемого:

$$Pu(t) \stackrel{\text{def}}{=} \sum_{s=k-m}^k \langle \tilde{g}^{(s)}, u \rangle \omega_s(t), \quad \forall t \in (\tilde{x}_k, \tilde{x}_{k+1}). \quad (5.3)$$

Операция проектирования P определяет прямую сумму

$$\mathbb{S} = \tilde{\mathbb{S}} \dot{+} \mathbb{W}. \quad (5.4)$$

Декомпозиция и реконструкция

Пусть $\mathbf{c} = (\dots, c_{-2}, c_{-1}, c_0, c_1, c_2, \dots)$ — исходный поток числовой информации. Рассмотрим функцию

$$u(t) \stackrel{\text{def}}{=} \sum_j c_j \omega_j(t). \quad (6.1)$$

Ее проекция $\tilde{u} \stackrel{\text{def}}{=} P u$ на пространство $\tilde{\mathbb{S}}$ может быть представлена в форме

$$\tilde{u} = \sum_i a_i \tilde{\omega}_i. \quad (6.2)$$

Итак, имеем так называемый основной поток

$$\mathbf{a} \stackrel{\text{def}}{=} (\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots),$$

который соответствует укрупнению \tilde{X} сетки X , а также всплесковый поток $\mathbf{b} \stackrel{\text{def}}{=} (\dots, b_{-2}, b_{-1}, b_0, b_1, b_2, \dots)$, который определяется разложением разности $w = u - \tilde{u}$ по базису пространства \mathbb{S} : $w = \sum_s b_s \omega_s$ (см. формулы (5.3) – (5.4), (6.1) – (6.2)).

Переход от исходного потока \mathbf{c} к потокам \mathbf{a} и \mathbf{b} называется *декомпозицией*, а обратный переход называется *реконструкцией*.

Формулы декомпозиции могут быть представлены в форме $\mathbf{a} = \mathfrak{Q}\mathbf{c}$, $\mathbf{b} = \mathbf{c} - \mathfrak{P}^T \mathfrak{Q}\mathbf{c}$, а формулы реконструкции — в форме $\mathbf{c} = \mathbf{b} + \mathfrak{P}^T \mathbf{a}$. Здесь \mathfrak{P} и \mathfrak{Q} — матрицы, которые называются матрицами сужения и продолжения соответственно.

В частном случае для $m = 1$, $\varphi(t) = (1, t)^T$, $X \setminus \tilde{X} = \{x_{k+1}\}$ формулы декомпозиции имеют вид

$$a_i = c_i \quad \text{при } i \leq k-1, \quad a_i = c_{i+1} \quad \text{при } i \geq k, \quad (6.3)$$

$$b_j = 0 \quad \text{при } j \neq k, \quad b_k = -\frac{x_{k+2} - x_{k+1}}{x_{k+2} - x_k} \cdot c_{k-1} + c_k - \frac{x_{k+1} - x_k}{x_{k+2} - x_k} \cdot c_{k+1}, \quad (6.4)$$

а формулы реконструкции могут быть представлены в форме

$$c_j = a_j + b_j \quad \text{при } j \leq k-1, \quad (6.5)$$

$$c_k = \frac{x_{k+2} - x_{k+1}}{x_{k+2} - x_k} \cdot a_{k-1} + \frac{x_{k+1} - x_k}{x_{k+2} - x_k} \cdot a_k + b_k, \quad (6.6)$$

$$c_j = a_{j-1} + b_j \quad \text{при } j \geq k+1. \quad (6.7)$$

Заметим, что если отрезок $[a, b]$ содержится в интервале (α, β) , то все предыдущие построения справедливы для сужения рассматриваемых функций на этот отрезок; при этом рассматривающиеся сетки, а также исходный, основной и всплесковый потоки оказываются конечными.

О численной устойчивости декомпозиции и реконструкции

Определение. Сетка $X \stackrel{\text{def}}{=} \{x_i\}_{i \in \mathbb{Z}}$ вида (1.1) – (1.2) называется локально квазиравномерной, если существует число $K_0 \geq 1$ такое, что справедливо соотношение

$$K_0^{-1} \leq \frac{x_{s+1} - x_s}{x_s - x_{s-1}} \leq K_0 \quad \forall s \in \mathbb{Z}. \quad (7.1)$$

Множество сеток, удовлетворяющих неравенствам (7.1), называется классом локально квазиравномерных сеток типа K_0 . Этот класс сеток обозначается $\mathcal{X}(K_0)$. Множество всех сеток обозначаем $\mathcal{X}(\infty)$.

Заметим, что $\mathcal{X}(1)$ — множество равномерных сеток (при этом, очевидно, должно быть $\alpha = -\infty, \beta = +\infty$). Дальше считаем, что $K_0 \geq 1$.

Лемма 1. Если сетка (1.1) – (1.2) является локально квазиравномерной сеткой типа K_0 , то

$$(1 + K_0)^{-1} \leq \frac{x_{k+1} - x_k}{x_{k+2} - x_k} \leq (1 + K_0^{-1})^{-1}, \quad (7.2)$$

$$(1 + K_0)^{-1} \leq \frac{x_{k+2} - x_{k+1}}{x_{k+2} - x_k} \leq (1 + K_0^{-1})^{-1}. \quad (7.3)$$

Доказательство. Используя соотношение (7.1), имеем

$$K_0^{-1} + 1 \leq \frac{x_{k+2} - x_k}{x_{k+1} - x_k} = \frac{x_{k+2} - x_{k+1}}{x_{k+1} - x_k} + 1 \leq K_0 + 1.$$

Учитывая неотрицательность всех частей этого двойного неравенства, перейдем к обратным величинам. В результате получим соотношение (7.2).

Неравенство (7.3) доказывается аналогично доказательству соотношения (7.2). ■

Лемма 2. Если сетка (1.1) – (1.2) является локально квазиравномерной сеткой типа K_0 , то для формул декомпозиции справедливы соотношения

$$|a_i| = |c_i| \quad \text{при } i \leq k-1, \quad |a_i| = |c_{i+1}| \quad \text{при } i \geq k, \quad (7.4)$$

$$|b_k| \leq (1 + K_0^{-1})^{-1} \cdot (|c_{k-1}| + |c_{k+1}|) + |c_k|. \quad (7.5)$$

Доказательство. Соотношения (7.4) вытекают из формул (6.3), а соотношения (7.5) получаются из формулы (6.4) и неравенств (7.2) – (7.3). ■

Лемма 3. Если сетка (1.1) – (1.2) является локально квазиравномерной сеткой типа K_0 , то для формул реконструкции верны неравенства

$$|c_j| \leq |a_j| + |b_j| \quad \text{при } j \leq k-1, \quad (7.6)$$

$$|c_k| \leq (1 + K_0^{-1})^{-1} \cdot (|a_{k-1}| + |a_k|) + |b_k| \quad (7.7)$$

$$|c_j| \leq |a_{j-1}| + |b_j| \quad \text{при } j \geq k+1. \quad (7.8)$$

Доказательство. Соотношения (7.6) и (7.8) вытекают из формул (6.5) и (6.7), а соотношение (7.7) следует из формулы (6.6) и неравенств (7.2) – (7.3). ■

Введем обозначения

$$\|\mathbf{a}\| \stackrel{\text{def}}{=} \sup_{i \in \mathbb{Z}} |a_i|, \quad \|\mathbf{b}\| \stackrel{\text{def}}{=} \sup_{i \in \mathbb{Z}} |b_i|, \quad \|\mathbf{c}\| \stackrel{\text{def}}{=} \sup_{i \in \mathbb{Z}} |c_i|.$$

Теорема 1. (Об устойчивости декомпозиции) Если X — локально квазиравномерная сетка типа K_0 , то справедливы неравенства

$$\|\mathbf{a}\| \leq \|\mathbf{c}\|, \quad (7.9)$$

$$\|\mathbf{b}\| \leq (2(1 + K_0^{-1})^{-1} + 1)\|\mathbf{c}\|. \quad (7.10)$$

Доказательство. Из (7.4) имеем

$$|a_i| = |c_i| \leq \|\mathbf{c}\| \quad \text{при } i \leq k-1, \quad |a_i| = |c_{i+1}| \leq \|\mathbf{c}\| \quad \text{при } i \geq k,$$

откуда вытекает неравенство (7.9). Из (7.5) выводим

$$|b_k| \leq 2(1 + K_0^{-1})^{-1} \cdot \|\mathbf{c}\| + \|\mathbf{c}\|.$$

Таким образом, неравенство (7.10) доказано. ■

Теорема 2. (Об устойчивости реконструкции) Если X — локально квазиравномерная сетка типа K_0 , то справедливы неравенства

$$\|\mathbf{c}\| \leq \max\{2(1 + K_0^{-1})^{-1}, 1\} \cdot \|\mathbf{a}\| + \|\mathbf{b}\|. \quad (7.11)$$

Доказательство. Из неравенств (7.6) и (7.6) имеем

$$|c_j| \leq \|\mathbf{a}\| + \|\mathbf{b}\| \quad \text{при } j \in \mathbb{Z} \setminus \{k\},$$

а из соотношения (7.7) имеем

$$|c_k| \leq 2(1 + K_0^{-1})^{-1} \cdot \|\mathbf{a}\| + \|\mathbf{b}\|.$$

Из последних двух неравенств следует соотношение (7.11). ■

Заключение

В работе установлены критерии устойчивости сплайн-всплесковых разложений. Доказаны теоремы об устойчивости декомпозиции и реконструкции.

Список литературы

- [1] Демьянович Ю. К. Всплески & минимальные сплайны. СПб., 2003.
- [2] Демьянович Ю. К. Теория сплайн-всплесков. СПб., 2013.

Распараллеливание в OpenMP и сплайновые аппроксимации



Бурова Ирина Герасимовна

д.ф.-м.н., профессор, профессор Кафедры вычислительной математики

ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ПАРАЛЛЕЛЬНОГО ПРОГРАММИРОВАНИЯ ПРИ РЕШЕНИИ УРАВНЕНИЯ ФРЕДГОЛЬМА ВТОРОГО РОДА

Бурова И. Г., профессор кафедры вычислительной математики
СПбГУ, burovaig@mail.ru,

Алцыбеев Г. О., аспирант кафедры вычислительной математики
СПбГУ, gleb.alcybeev@gmail.com

Аннотация

В работе рассматривается применение технологии параллельных вычислений OpenMP для решения интегрального уравнения Фредгольма второго рода с помощью локальных интерполяционных сплайнов второго порядка аппроксимации.

Введение

В настоящее время при разработке ресурсозатратного программного обеспечения много внимания уделяется различного рода оптимизациям программного кода для повышения скорости вычислений. В качестве одного из основных способов можно отметить параллельные вычисления. Параллельные вычисления — это вид вычислений, при которых сразу несколько вычислительных процессов выполняются одновременно в течение одного и того же периода времени. Пионерами в области параллельных вычислений являются Эдсгер Дейкстра [1], Пер Бринч Хансен [2], [3] и Хоар К. А. Р. [4]. Большой вклад в области параллельных вычислений внесли Воеводин В. В. и Воводин В. В. [5], Гергель В. П. [6], Антонов А. С. [7], Корнеев В. Д. [8], и Немногин С. А. [9].

В качестве одного из инструментов параллельных вычислений можно выделить технологию OpenMP. OpenMP — это интерфейс прикладного программирования, который поддерживает многоплатформенное многопроцессорное программирование с общей памятью на языках C, C++ и Fortran.

Решение уравнения Фредгольма и сплайновые аппроксимации

Пусть $a, b \in \mathbb{R}$. Рассмотрим линейное интегральное уравнение Фредгольма второго рода

$$y(x) - \int_a^b K(x, s) y(s) ds = f(x), \quad x \in [a, b],$$

где $f(x)$ такая, что $f \in C[a, b]$ — правая часть, $K(x, s)$ — ядро, определенное в квадрате $\Pi = \{(x, s) | a \leq x \leq b, a \leq s \leq b\}$, полагаем, что ядро $K(x, s)$ непрерывно в квадрате Π , а $y(x)$ — искомая непрерывная функция, $x \in [a, b]$.

На промежутке $[a, b]$ задаем узлы сетки $a = x_1 < x_2 < \dots < x_n = b$. Приближенные значения $\tilde{y}(s)$ функции $y(s)$ на промежутке $[x_j, x_{j+1}]$ определяем по правилу

$$\tilde{y}(s) = y(x_j) \omega_j(s) + y(x_{j+1}) \omega_{j+1}(s), \quad s \in [x_j, x_{j+1}], \quad (1)$$

где $\omega_j(s)$ и $\omega_{j+1}(s)$ — базисные полиномиальные сплайны

$$\omega_j(s) = \frac{s - x_{j+1}}{x_j - x_{j+1}}, \quad \omega_{j+1}(s) = \frac{s - x_j}{x_{j+1} - x_j}, \quad s \in [x_j, x_{j+1}],$$

отметим, что удобно представление $\omega_{j+1}(s)$: $\omega_{j+1}(s) = 1 - \omega_j(s)$.

Обозначим норму

$$\|y''\|_{[x_j, x_{j+1}]} = \max_{[x_j, x_{j+1}]} |y''(x)|.$$

Пусть $h = x_{j+1} - x_j$. Можно показать, что в случае полиномиальных сплайнов справедливо неравенство

$$|\tilde{y}(s) - y(s)| \leq 0,25h^2 \|y''\|_{[x_j, x_{j+1}]}, \quad s \in [x_j, x_{j+1}].$$

Также можно использовать неполиномиальные базисные функции (см. [11], [12]) с погрешностью аппроксимации порядка $O(h^2)$.

Нетрудно видеть, что

$$\int_a^b K(x, s) y(s) ds \approx \sum_{k=1}^{n-1} \int_{x_j}^{x_{j+1}} K(x, s) \tilde{y}(s) ds, \quad x \in [a, b],$$

где $\tilde{y}(s)$ имеет вид (1). В результате применения сплайновых аппроксимаций, получаем систему линейных алгебраических уравнений (СЛАУ)

$$\tilde{y}(x_k) + \sum_{j=1}^{n-1} W_j(x_k) = f(x_k), \quad k = 1, 2, \dots, n, \quad (2)$$

где

$$\begin{aligned} W_j(x_k) &= \tilde{y}(x_j) \int_{x_j}^{x_{j+1}} K(x_k, s) \omega_j(s) ds + \\ &+ \tilde{y}(x_{j+1}) \int_{x_j}^{x_{j+1}} K(x_k, s) (1 - \omega_j(s)) ds. \end{aligned} \quad (3)$$

Равенство (3) можно упростить, таким образом оно принимает следующий вид

$$\begin{aligned} W_j(x_k) &= (\tilde{y}(x_j) - \tilde{y}(x_{j+1})) \int_{x_j}^{x_{j+1}} K(x_k, s) \omega_j(s) ds + \\ &+ \tilde{y}(x_{j+1}) \int_{x_j}^{x_{j+1}} K(x_k, s) ds, \end{aligned}$$

кроме этого, в случае, если интеграл трудно вычислять, можно использовать следующую форму записи

$$\begin{aligned} W_j(x_k) &= \left(\frac{\tilde{y}(x_j) - \tilde{y}(x_{j+1})}{x_j - x_{j+1}} \right) \int_{x_j}^{x_{j+1}} K(x_k, s) s ds + \\ &+ \left(\tilde{y}(x_{j+1}) - \frac{x_{j+1}}{x_j - x_{j+1}} \right) \int_{x_j}^{x_{j+1}} K(x_k, s) ds. \end{aligned}$$

В большинстве случаев интегралы вычисляются в конечном виде, либо можно применить соответствующие квадратурные формулы.

Приведем СЛАУ (2) к виду $A\tilde{y} = b$. Выпишем случай, когда $n = 3$, тогда $\tilde{y} = (\tilde{y}(x_1), \tilde{y}(x_2), \tilde{y}(x_3))^T$, а $b = (f(x_1), f(x_2), f(x_3))^T$.

С учетом того, что $n = 3$ имеем систему

$$\begin{cases} \tilde{y}(x_1) - \tilde{y}(x_1)K_{00} - \tilde{y}(x_2)K_{10} - \tilde{y}(x_2)K_{20} - \tilde{y}(x_3)K_{30} = f(x_1), \\ \tilde{y}(x_2) - \tilde{y}(x_1)K_{01} - \tilde{y}(x_2)K_{11} - \tilde{y}(x_2)K_{21} - \tilde{y}(x_3)K_{31} = f(x_2), \\ \tilde{y}(x_3) - \tilde{y}(x_1)K_{02} - \tilde{y}(x_2)K_{12} - \tilde{y}(x_2)K_{22} - \tilde{y}(x_3)K_{32} = f(x_3), \end{cases} \quad (4)$$

в которой приняты следующие обозначения

$$K_{00} = \int_{x_1}^{x_2} K(x_1, s) \omega_1(s) ds, \quad K_{10} = \int_{x_2}^{x_3} K(x_1, s) \omega_2(s) ds,$$

$$\begin{aligned}
 K_{20} &= \int_{x_1}^{x_2} K(x_1, s) \omega_2(s) ds, & K_{30} &= \int_{x_2}^{x_3} K(x_1, s) \omega_3(s) ds, \\
 K_{01} &= \int_{x_1}^{x_2} K(x_2, s) \omega_1(s) ds, & K_{11} &= \int_{x_2}^{x_3} K(x_2, s) \omega_2(s) ds, \\
 K_{21} &= \int_{x_1}^{x_2} K(x_2, s) \omega_2(s) ds, & K_{31} &= \int_{x_2}^{x_3} K(x_2, s) \omega_3(s) ds, \\
 K_{02} &= \int_{x_1}^{x_2} K(x_3, s) \omega_1(s) ds, & K_{12} &= \int_{x_2}^{x_3} K(x_3, s) \omega_2(s) ds, \\
 K_{22} &= \int_{x_1}^{x_2} K(x_3, s) \omega_2(s) ds, & K_{32} &= \int_{x_2}^{x_3} K(x_3, s) \omega_3(s) ds.
 \end{aligned}$$

Нетрудно заметить, что система (4) преобразуется к системе вида

$$\begin{cases}
 \underbrace{\tilde{y}(x_1)(1 - K_{00})}_{a_{11}} - \underbrace{\tilde{y}(x_2)(K_{10} + K_{20})}_{a_{12}} - \underbrace{\tilde{y}(x_3)K_{30}}_{a_{13}} = f(x_1), \\
 -\underbrace{\tilde{y}(x_1)K_{01}}_{a_{21}} + \underbrace{\tilde{y}(x_2)(1 - K_{11} - K_{21})}_{a_{22}} - \underbrace{\tilde{y}(x_3)K_{31}}_{a_{23}} = f(x_2), \\
 -\underbrace{\tilde{y}(x_1)K_{02}}_{a_{31}} - \underbrace{\tilde{y}(x_2)(K_{12} + K_{22})}_{a_{32}} + \underbrace{\tilde{y}(x_3)(1 - K_{32})}_{a_{33}} = f(x_3),
 \end{cases}$$

которую можно записать в матричном виде

$$\begin{pmatrix} a_{11} & -a_{12} & -a_{13} \\ -a_{21} & a_{22} & -a_{23} \\ -a_{31} & -a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \tilde{y}(x_1) \\ \tilde{y}(x_2) \\ \tilde{y}(x_3) \end{pmatrix} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \end{pmatrix}.$$

Можно показать, что матрица системы уравнений имеет диагональное преобладание при произвольном n , поэтому она — неособенная. Полученную СЛАУ можно решать численными методами, например методом Гаусса.

В ходе работы был разработан программный комплекс на языке C++ для работы с интегральными уравнениями, а также библиотеки **AS::LinearAlgebra** и **AS::MathAnalysis** для выполнения сопутствующих операций. Информация об этом в следующем разделе.

Распараллеливание вычислений в методе Гаусса

Полученную СЛАУ можно решать различными численными методами, например в данном случае хорошо подходит метод Гаусса. В ходе работы была разработана библиотека **AS::LinearAlgebra** на языке

C++ для работы с матрицами и векторами. В библиотеку вошло большинство стандартных операций для решения задач линейной алгебры, в том числе различные методы для решения СЛАУ. Рассмотрим реализацию процедуры **AS::LinearSolve** на примере метода Гаусса. Параметры процедуры: A — экземпляр класса **AS::Matrix** из библиотеки **AS::LinearAlgebra**, b — экземпляр класса **AS::Vector** (может использоваться также класс **AS::Matrix**) из библиотеки **AS::LinearAlgebra**, столбец свободных членов.

Рассмотрим процедуру метода Гаусса в общем виде и внесем в нее некоторые модификации с помощью технологии OpenMP. Программы тестировались на СЛАУ из 3000-4000 уравнений. Обозначим $y_i = \tilde{y}(x_i)$ и $b_i = f(x_i)$. Данна СЛАУ порядка n

$$\begin{cases} a_{11}^{(0)}y_1 + a_{12}^{(0)}y_2 + \dots + a_{1n}^{(0)}y_n = b_1^{(0)} \\ a_{21}^{(0)}y_1 + a_{22}^{(0)}y_2 + \dots + a_{2n}^{(0)}y_n = b_2^{(0)} \\ \dots \\ a_{n1}^{(0)}y_1 + a_{n2}^{(0)}y_2 + \dots + a_{nn}^{(0)}y_n = b_n^{(0)} \end{cases}.$$

Разделим первое уравнение системы на $a_{11}^{(0)}$, тогда получим

$$y_1 + a_{12}^{(1)}y_2 + a_{13}^{(1)}y_3 + \dots + a_{1n}^{(1)}y_n = b_1^{(1)}, \quad (5)$$

где $a_{1j}^{(1)} = a_{1j}^{(0)}/a_{11}^{(0)}$, $j = 2, 3, \dots, n$, $b_1^{(1)} = b_1^{(0)}/a_{11}^{(0)}$. Предположим, что система уравнений такова, что $n > 3000$. Вычисления в цикле деления элементов можно распараллелить используя директивы OpenMP **parallel** и **for**. В результате имеем конструкцию представленную на Листинге 2.

```
#pragma omp parallel
{
    #pragma omp for
    for (int i = 0; i < Ab.GetColSize(); i++) {
        Matrix_Ab_c[k][i] = Matrix_Ab_c[k][i] / Ab[k][k];
    }
}
```

Листинг 2. Участок кода с циклом деления элементов в процедуре **LinearSolve** с использованием директив OpenMP

Далее исключаем неизвестную y_1 из каждого уравнения системы, начиная со второго. Это делается вычитанием уравнения (5), умно-

женного на коэффициент при переменной y_1 в соответствующем уравнении. Преобразованные уравнения имеют вид

$$\left\{ \begin{array}{l} y_1 + a_{12}^{(1)}y_2 + a_{13}^{(1)}y_3 + \dots + a_{1n}^{(1)}y_n = b_1^{(1)} \\ a_{22}^{(1)}y_2 + a_{23}^{(1)}y_3 + \dots + a_{2n}^{(1)}y_n = b_2^{(1)} \\ \dots \\ a_{n2}^{(1)}y_2 + a_{n3}^{(1)}y_3 + \dots + a_{nn}^{(1)}y_n = b_n^{(1)} \end{array} \right., \quad (6)$$

где $a_{ij}^{(1)} = a_{ij}^{(0)} - a_{1j}^{(1)} \cdot a_{i1}^{(0)}$, $j = 2, 3, \dots, n$, $b_i^{(1)} = b_i^{(1)} \cdot a_{i1}^{(0)}$, $i = 2, 3, \dots, n$. Поступаем аналогично со следующим уравнением из преобразованной системы. В конечном итоге приводим исходную систему к эквивалентной системе с треугольной матрицей

$$\left\{ \begin{array}{l} y_1 + a_{12}^{(1)}y_2 + a_{13}^{(1)}y_3 + \dots + a_{1n}^{(1)}y_n = b_1^{(1)} \\ y_2 + a_{23}^{(2)}y_3 + \dots + a_{2n}^{(2)}y_n = b_2^{(2)} \\ \dots \\ a_{nn}^{(n)}y_n = b_n^{(n)} \end{array} \right.. \quad (7)$$

Цикл с исключением неизвестной y_i из каждого уравнения системы также распараллеливается с применением директив **parallel** и **for**. В результате имеем конструкцию представленную на Листинге 3.

```
#pragma omp parallel
{
    #pragma omp for
    for (int i = k + 1; i < Ab.GetRowSize(); i++) {
        double K = Matrix_Ab_c[i][k] / Matrix_Ab_c[k][k];
        for (int j = 0; j < Ab.GetColSize(); j++) {
            Matrix_Ab_c[i][j] = Matrix_Ab_c[i][j] - Matrix_Ab_c[
                k][j] * K;
        }
    }
}
```

Листинг 3. Участок кода с циклом исключения неизвестной y_i в процедуре **LinearSolve** с использованием директив OpenMP

Далее обратным ходом из системы (7) находим неизвестные y_1, y_2, \dots, y_n . Соответствующие циклы распараллеливаются аналогично. Кроме этого, аналогичная конструкция была построена для цикла записи ответа, однако данные действия привели к незначительным ускорениям программы.

Заключение

В работе было рассмотрено применение технологий параллельный вычислений для решения интегрального уравнения Фредгольма второго рода с помощью локальных интерполяционных сплайнов второго порядка аппроксимации.

Список литературы

- [1] Edsger W. Dijkstra. Selected Writings on Computing: A Personal Perspective. Monographs in Computer Science. New York: Springer Science & Business Media, 2012. 362 p.
- [2] Per Brinch Hansen. Design Principles. New Jersey: Prentice Hall, Englewood Cliffs, 1977. 314 p.
- [3] Per Brinch Hansen. Experience with Modular Concurrent Programming // IEEE Transactions on Software Engineering. 1977. No. 3 (2). P. 156–159.
- [4] Hoare C. A. R., Communicating Sequential Processes. New Jersey: Prentice Hall, 1985. 260 p.
- [5] Воеводин В. В., Воеводин В. В. Параллельные вычисления. СПб.: БХВ, 2002. 608 с.
- [6] Гергель В. П. Высокопроизводительные вычисления для многоядерных многопроцессорных систем. Нижний Новгород: Изд-во Нижегородского госуниверситета, 2010. 421 с.
- [7] Антонов А. С. Технологии параллельного программирования MPI и OpenMP. М.: Изд-во МГУ, 2012. 344 с.
- [8] Корнеев В. Д. Параллельное программирование в MPI. Новосибирск: Изд-во ИВМИМГ СО РАН, 2002. 215 с.
- [9] Немнюгин С. А., Стесик О. Л. Параллельное программирование для многопроцессорных вычислительных систем. СПб.: БХВ-Петербург, 2002. 396 с.
- [10] Алицыбеев Г. О., Бурова И. Г. Газотурбинный двигатель и сплайновые приближения // Процессы управления и устойчивость. 2021. Т. 8 (24). С. 101–107.

- [11] Burova I. G. On left integro-differential splines and Cauchy problem // International Journal of Mathematical Models and Methods in Applied Sciences. 2015. Vol. 9. P. 683–690.
- [12] Burova I. G., Alcybeev G. O. Application of Splines of the Second Order Approximation to Volterra Integral Equations of the Second Kind. Applications in Systems Theory and Dynamical Systems // International Journal of Circuits, Systems and Signal Processing. 2021. Vol. 15. P. 63–71.

Вычислительная стохастика и статистические модели



Ермаков Сергей Михайлович

д.ф.-м.н., профессор, заведующий кафедрой статистического моделирования



Голяндина Нина Эдуардовна

к.ф.-м.н., доцент кафедры статистического моделирования

О ПРОВЕРКЕ ГИПОТЕЗ В ДИСПЕРСИОННОМ АНАЛИЗЕ ПОВТОРЯЮЩИХСЯ НЕПОЛНЫХ НАБЛЮДЕНИЙ¹

Алексеева Н.П., доцент кафедры статистического моделирования
СПбГУ, nina.alekseeva@spbu.ru

Федорченко С.А., консультант в Insilico Medicine,
fedorchenko.ser.eja@yandex.ru

Аннотация

В модели дисперсионного анализа для повторяющихся наблюдений (ANOVA Repeated Measures) при условии неполных данных во временных точках кроме первой решается задача централизации модели, приводящей к коррелированности ошибок. В соответствии со структурой инцидентности наблюдений вычисляется ковариационная матрица ошибок. Статистическая модель с коррелированными ошибками используется для проверки значимости отклонения от нуля дифференциальных эффектов. В аналогичной модели со случайными эффектами для факторов повторения и их взаимодействия вычисляются математические ожидания статистик, необходимых для построения отношения Фишера. Показано, что для проверки значимости эффекта взаимодействия используется одна и та же статистика вне зависимости от дифференциального или случайного характера данного эффекта. Для проверки значимости случайного фактора повторения получен дополнительный критерий в виде равенства, справедливого, в частности, при условии одинакового числа наблюдений в ячейке.

Введение

Дисперсионный анализ повторяющихся наблюдений, в литературе известный под названием ANOVA Repeated Measures (AVRM) – один из самых востребованных методов анализа лонгитюдных данных [1], позволяющий в рамках одной статистической модели осуществлять проверку значимости динамики наблюдений, главных эффектов и эффек-

¹ Работа поддержана грантом РФФИ: 20-01-00096.

тов их взаимодействия с фактором времени. Проблема возникает, если данные неполные. Не умаляя достоинств методов, предлагающих заполнение пропущенных данных, остановимся на эргодическом методе, предложенном автором в [2], примененном для анализа неполных кардиологических наблюдений в [3], представленном на международной конференции [4]. Идея метода заключается в централизации модели через введение индивидуальной поправки и дальнейшем пересчете ковариационной матрицы ошибок. В [5] было предложено усовершенствование модели за счет введения групповой поправки, необходимой в случае неравномерности распределения пропущенных данных в отдельных группах. Несмотря на то, что удалось довести исследование до числа и получить адекватные результаты, осталось несколько нераскрытых тем. Так ранее не обсуждалось то, на чем основана идея построения поправок. Кроме того не рассмотрены возможности применения случайных эффектов, и актуальным остается тестирование работы полученных критериев на модельных данных.

Модель AVRM с дифференциальными эффектами

Модель дисперсионного анализа для повторяемых наблюдений AVRM с дифференциальными эффектами можно представить в виде

$$x_{ijk} = \mu + \alpha_i + \delta_{ij} + \beta_k + \gamma_{ik} + e_{ijk}, \quad (1)$$

где x_{ijk} наблюдение j -го индивида из группы i в момент времени t , μ параметр генерального среднего, α_i , β_k , γ_{ik} дифференциальные эффекты, соответствующие группе $i = 1, 2, \dots, I$, временной точке $k = 1, 2, \dots, T$ и эффектам взаимодействия факторов времени и группы. Ошибки δ_{ij} , e_{ijk} предполагаются независимыми нормально распределенными величинами с нулевыми средними и соответственно дисперсиями σ_1^2 и σ^2 . Обозначим через M_{it} номера индивидов из группы i , имеющие наблюдения в момент времени t , а через m_{it} их количество, $m_i = \sum_{t=1}^T m_{it}$, $m_{..} = \sum_{i=1}^I m_{it}$, $m_{..} = \sum_{t=1}^T m_{..}$. Пусть N_{ij} множество временных точек индивида j из группы i , и n_{ij} их количество. Ограничения на параметры выбираем в соответствии с частичным планом $\sum_{i=1}^I \frac{\alpha_i m_{..}}{m_{..}} = 0$, $\sum_{t=1}^T \frac{\beta_t m_{..}}{m_{..}} = 0$, $\sum_{i=1}^I \frac{\gamma_{it} m_{it}}{m_{..}} = 0$, $\sum_{t=1}^T \frac{\gamma_{it} m_{it}}{m_{..}} = 0$ [6]. Оценки параметров модели (1) с учетом этих ограничений $\hat{\alpha}_i = x_{i..} - x_{...}$, $\hat{\beta}_t = x_{..t} - x_{...}$, $\hat{\gamma}_{it} = x_{i..t} - x_{i..} - x_{..t} + x_{...}$ выражаются через разнообразные формы усреднения наблюдений: индивидуальное

среднее для j -го индивида из i -й группы $x_{ij\cdot}$, среднее $x_{..t}$ в заданный момент времени t , внутригрупповое среднее $x_{i..}$, общее среднее $x_{...}$. В случае полных данных проверку гипотез для параметров модели (1) можно свести к проверке гипотез для двух моделей с математическими ожиданиями $\mathbb{E}x_{ij\cdot} = \mu + \alpha_i$, $\mathbb{E}(x_{ijk} - x_{ij\cdot}) = \beta_k + \gamma_{ik}$. В случае неполных данных эти модели становятся смещенными. В этом можно убедиться, доказав следующее утверждение.

Лемма 0.1 *Если существует индивид j из группы i такой, что $n_{ij} \neq T$, то для индивидуального среднего имеет место смещение $\mathbb{E}x_{ij\cdot} = \mu + \alpha_i + W_{ij}$, где $W_{ij} = \frac{1}{n_{ij}} \sum_{t \in N_{ij}} (\beta_t + \gamma_{it})$.*

Из оценок параметров получим $\hat{\beta}_t + \hat{\gamma}_{it} = x_{i..t} - x_{i..}$ и оценим смещение

$$\begin{aligned}\hat{W}_{ij} &= \frac{1}{n_{ij}} \sum_{t \in N_{ij}} (\hat{\beta}_t + \hat{\gamma}_{it}) = \frac{1}{n_{ij}} \sum_{t \in N_{ij}} \frac{1}{m_{it}} \sum_{l \in M_{it}} (x_{ilt} - x_{il\cdot}), \\ \mathbb{E}(x_{ij\cdot} - \hat{W}_{ij}) &= \mu + \alpha_i + \frac{1}{n_{ij}} \sum_{t \in N_{ij}} \frac{1}{m_{it}} \sum_{l \in M_{it}} W_{ij}.\end{aligned}$$

Можно повторить еще раз эту процедуру, рассмотрев математическое ожидание разности $\mathbb{E} \left(x_{ij\cdot} - \hat{W}_{ij} - \frac{1}{n_{ij}} \sum_{t \in N_{ij}} \frac{1}{m_{it}} \sum_{l \in M_{it}} \hat{W}_{ij} \right)$. Опять получим какое-то смещение меньшее предыдущего, и так далее. В общем виде, обозначим через $A_{ij}(0) = \frac{1}{n_{ij}} \sum_{t \in N_{ij}} \frac{1}{m_{it}} \sum_{l \in M_{it}} (x_{ilt} - x_{il\cdot})$ и определим рекуррентным образом последовательность $A_{ij}(k+1) = \frac{1}{n_{ij}} \sum_{t \in N_{ij}} \frac{1}{m_{it}} \sum_{l \in M_{it}} A_{ij}(k)$. В качестве индивидуального смещения будем рассматривать $H_{ij} = \sum_{k=1}^{\infty} A_{ij}(k)$. В [3] доказано, что H_{ij} конечно при условии полноты наблюдений хотя бы в один момент времени, вычислено его математическое ожидание. Это позволило использовать H_{ij} в качестве индивидуальной поправки при небольшом ограничении, связанном с равномерностью пропущенных наблюдений по группам, которое в [5] было снято посредством введения дополнительной поправки.

Поправки модели при неполноте данных

Оценку параметров общей модели (1) можно получить посредством введения двух несмешанных моделей, для которых

$$\mathbb{E}X_{ij} = \mu + \alpha_i \text{ и } \mathbb{E}(x_{ijk} - X_{ij}) = \beta_k + \gamma_{ik}. \quad (2)$$

В случае полных данных используется $X_{ij} = x_{ij.}$, а в случае неполных данных [5] нужно использовать поправки. Пусть J^i матрица инцидентности в i -й группе, Λ_{iT} и Λ_{ν_i} диагональные матрицы с векторами на главной диагонали вида $(m_{it})_{t=1}^T$ и $(n_{ij})_{j=1}^{\nu_i}$, $R_i = \Lambda_{\nu_i} J^i$, $P_i = R_i \Lambda_{iT} (J^i)^T$ – стохастическая матрица с ненулевыми компонентами при условии наличия полных данных хотя бы в одной точке, $P_i^\infty = \lim_{k \rightarrow \infty} P_i^k$, $Q_i = (I - P_i + P_i^\infty)^{-1}$, векторы $V_i = \{x_{ij.}\}_{j=1}^{\nu_i}$, $U_i = \{x_{i..t}\}_{t=1}^T$, $L = \{x_{..t} - x_{...}\}_{t=1}^T$, $K = \{x_{i..} - x_{...}\}_{i=1}^I$ и матрицы

$$M = \left\{ \frac{m_{i1}}{m_{i.}}, \frac{m_{i2}}{m_{i.}}, \dots, \frac{m_{iT}}{m_{i.}} \right\}_{i=1}^I, \quad N = \left\{ \frac{m_{1t}}{m_{.t}}, \dots, \frac{m_{it}}{m_{.t}} \right\}_{t=1}^T. \quad (3)$$

Тогда для i -й группы вектор индивидуальной поправки H_i с компонентами $\{H_{ij}\}_{i=1, j=1}^{I, \nu_i}$ вводится как

$$H_i = (I - C_i)V_i + D_iU_i, \quad \text{где } C_i = P_i^\infty + Q_i, D_i = Q_iR_i. \quad (4)$$

Групповые поправки G_i определяются как компоненты вектора $G = \sum_{i=0}^{\infty} (MN)^i (ML - MNK)$.

Теорема 0.1 [5] Пусть M, N матрицы из (3) и $P_0 = MN$ стохастическая матрица, $Q_0 = (I - P_0 + P_0^\infty)^{-1}$, $A = P_0Q_0$, $B = Q_0M$, a_{ik} и b_{it} соответственно элементы этих матриц и $d_{j\tau}^i$ элементы матриц C_i, D_i из (4). Будем рассматривать два вида наблюдений

$$X_{ij} = x_{ij.} - (H_{ij} + G_i) \quad u \quad y_{ijt} = x_{ijt} - X_{ij}. \quad (5)$$

Пусть $\Delta_{ij} = e_{ij.} - \mathcal{E}_{ij} - \varepsilon_i$, $\mathcal{E}_{ij} = H_{ij} - \mathbb{E}H_{ij}$, $\varepsilon_i = G_i - \mathbb{E}G_i$. Тогда

$$\mathcal{E}_{ij} = e_{ij.} - \sum_{l=1}^{\nu_i} c_{jl}^i e_{il.} + \sum_{\tau=1}^T d_{j\tau}^i e_{i.\tau}, \quad \varepsilon_i = \sum_{k=1}^I \sum_{t=1}^T m_{kt} \left(\frac{b_{it}}{m_{.t}} - \frac{a_{ik}}{m_{k.}} \right) e_{i..t},$$

и имеют место несмешанные модели вида

$$X_{ij} = \mu + \alpha_i + e_{ij}^1 + \Delta_{ij} \quad u \quad y_{ijt} = \beta_t + \gamma_{it} + e_{ijt} - \Delta_{ij}. \quad (6)$$

Ковариационная матрица $\sigma^2 \Lambda$ ошибок $\varepsilon_{ijt} = e_{ijt} - \Delta_{ij}$ получена в [5].

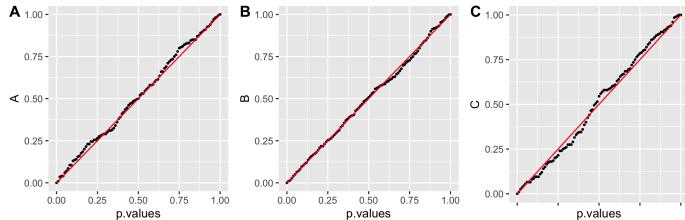


Рис. 1: Распределение p -значений для проверки значимости дифференциальных эффектов факторов: A группы, B времени, C взаимодействия. Объем выборки $m_1 = 72$, три временные точки, две группы, 6% процентов пропущенных данных.

Модель с дифференциальными эффектами

Рассмотрим элементы $y_{ijt} = x_{ijt} - X_{ij}$ из второго уравнения (5) $y_{ijt} = \beta_t + \gamma_{it} + \varepsilon_{ijt}$, для каждой пары i, t соберем m_{it} элементов y_{ijt} в векторе Y_{it} и обозначим через Y вектор $(Y_{11}, Y_{12}, \dots, Y_{1T}, Y_{21}, Y_{22}, \dots, Y_{2T}, \dots, Y_{I1}, Y_{I2}, \dots, Y_{IT})$. Согласно (5), вектор $Y = \{y_{ijk}\}$ состоит из $m_{..}$ компонент, из которых линейно независимые $n = m_{..} - m_1$, поэтому ранг матрицы Λ равен n .

Матрицу частичного плана обозначим через $X = [X_1 | X_2]$, где X_1, X_2 соответственно матрицы плана усеченных моделей, вектор ошибок ε_{ijk} через ε , вектор параметров через $\Theta = (\beta_1, \dots, \beta_{T-1}, \gamma_{1,1}, \dots, \gamma_{1,T-1}, \dots, \gamma_{I-1,1}, \dots, \gamma_{I-1,T-1})$. В матричном виде модель $y_{ijt} = \beta_t + \gamma_{it} + \varepsilon_{ijt}$ имеет вид $Y = X\Theta + \varepsilon$, где $\mathbb{E}\varepsilon\varepsilon^T = \sigma^2\Lambda$. При помощи ортогонального преобразования преобразуем эту модель к виду $V = H\Theta + \delta$, где δ независимые ошибки. Минимизируя выражение $(V - H\Theta)^T(V - H\Theta)$, получаем остаточные суммы квадратов

$$\begin{aligned} R_0 &= (V - H\hat{\Theta})^T(V - H\hat{\Theta}) = \delta^T(\mathbb{I} - H(H^TH)^{-1}H^T)\delta = \delta^TA\delta, \\ R_1 &= (V - H_1\hat{\beta})^T(V - H_1\hat{\beta}) = (H_2\gamma + \delta)^TA_1(H_2\gamma + \delta), \\ R_2 &= (V - H_2\hat{\gamma})^T(V - H_2\hat{\gamma}) = (H_1\beta + \delta)^TA_2(H_1\beta + \delta), \end{aligned} \quad (7)$$

рассматривая $A_i = \mathbb{I}_n - H_1(H_1^TH_1)^{-1}H_1^T$, $i = 0, 1, 2$, $A_0 = A$. Кроме того $R_2 - R_0 = (H_1\beta + H_2\gamma + \delta)^T(\mathbb{I} - A_1)(H_1\beta + H_2\gamma + \delta)$.

Используя обратное ортогональное преобразование, можно выразить компоненты R_0 , R_1 , R_2 через матрицы плана $X = X_0$, X_1 , X_2 и матрицу Λ . В результате имеем $R_i = Y^TB_iY$, где $i = 0, 1, 2$, $B_i = \Lambda^{-1} - \Lambda^{-1}X_i(X_i^T\Lambda^{-1}X_i)^{-1}X_i^T\Lambda^{-1}$. Далее применяются стандартное отношение Фишера. На рис.1 представлены согласованные с равно-

мерным распределением функции распределения p -значений, полученных при проверке значимости отклонения от нуля дифференциальных эффектов на модельных данных с числом итераций 400.

Проверка значимости случайных эффектов

В отличии от модели (1) эффекты факторов времени и взаимодействия будем считать случайными, то есть

$$x_{ijk} = \mu + \alpha_i + \delta_{ij} + b_k + g_{ik} + e_{ijk}, \quad (8)$$

где b_k , g_{ik} независимые нормально распределенные центрированные случайные величины с дисперсиями σ_b^2 и σ_g^2 соответственно. Вычислим математические ожидания статистик (7), полученных в анализе модели с дифференциальными эффектами,

$$\begin{aligned} \mathbb{E}R_0 &= \sigma^2 \text{Tr}(A) = \sigma^2(n - I(T - 1)), \\ \mathbb{E}R_1 &= \mathbb{E}(H_2 g + \delta)^T A_1 (H_2 g + \delta) = \mathbb{E}g^T H_2^T A_1 H_2 g + \mathbb{E}\delta^T A_1 \delta = \\ &= \text{Tr}(H_2^T A_1 H_2) \sigma_g^2 + \text{Tr}(A_1) \sigma^2, \text{ где } \text{Tr}A_1 = n - (T - 1). \\ \mathbb{E}(R_1 - R_0) &= \text{Tr}(H_2^T A_1 H_2) \sigma_g^2 + (I - 1)(T - 1)\sigma^2. \end{aligned}$$

Отсюда видно, для проверки гипотезы $H_0 : \sigma_g^2 = 0$ можно применить стандартный критерий, а в случае $H_0 : \sigma_b = 0$ для построения статистики Фишера нужно выполнение равенства $(I - 1)\text{Tr}(H_2^T(\mathbb{I} - A_1)H_2) = \text{Tr}(H_2^T A_1 H_2)$, которое справедливо в случае одинакового числа наблюдений $m_{it} = m$ в каждой точке и в каждой группе. Если это равенство не выполняется, то нужно исследовать возможность выравнивания коэффициентов через оценки дисперсий σ_g^2 и σ^2 , но этот вопрос является предметом уже дальнейшего исследования.

Заключение

В большинстве статистических моделей изучения влияния факторов на зависимую переменную метрического типа является закономерным требование одинакового числа наблюдений в ячейке. В анализе повторяющихся наблюдений это требование удалось сузить до более приемлемого требования наличия полных данных хотя бы в одной точке за счет корректировки индивидуального среднего по времени в зависимости от числа имеющихся наблюдений. Вычисление корреляционной

матрицы ошибок позволило построить статистики для проверки значимости отклонения от нуля дифференциальных эффектов. Моделирование свидетельствует об адекватности предлагаемого метода проверки гипотез.

Предпринята попытка применить данный подход в модели со случайными эффектами. Как и ожидалось, статистика для проверки значимости случайного эффекта взаимодействия совпала с соответствующей статистикой в модели с дифференциальными эффектами. Для проверки значимости случайного эффекта фактора времени приведено равенство, при выполнении которого допустимо построение отношения Фишера. В дальнейшем предполагается исследовать область допустимых отклонений от этого равенства для получения приближенных результатов и удостовериться в адекватности данного подхода при увеличении числа дополнительных факторов.

Литература

- [1] Longitudinal Data Analysis // Ed. by G. Fitzmaurice, M. Davidian, G. Verbeke, G. Molenberghs. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Boca Raton: Chapman and Hall/CRC, 2008.
- [2] Alexeyeva N.P., Tatarinova A.A., Bondarenko B.B. et al. (2011). Analysis of repeated cardiological incomplete data based on ergodic centralization of model. Bulleten of Almazov Center, 3(8):59–63.
- [3] Alexeyeva N. (2013). Analysis of biomedical systems. Reciprocity. Ergodicity. Synonymy. Publishing of the Saint-Petersburg State University, Saint-Petersburg.
- [4] Ufland A., Alexeyeva N. (2014). The dependence of the ergodicity on the time effect in the repeated measures anova with missing data based on the unbiasedness recovery. In Topics in Statistical Simulation, Springer Proceedings in Mathematics and Statistics 114, pages 517–527. Springer Science + Business Media New York 2014.
- [5] Alexeyeva N. (2017) Dual balance correction in repeated measures ANOVA with missing data January Electronic Journal of Applied Statistical Analysis 10(1):146-159 DOI:10.1285/i20705948v10n1p146
- [6] Scheffe H. (1999). The Analysis of Variance. John Wiley and Sons, Canada.

L-ОПТИМАЛЬНЫЕ ПЛАНЫ ДЛЯ ТРИГОНОМЕТРИЧЕСКОЙ РЕГРЕССИИ БЕЗ СВОБОДНОГО ЧЛЕНА.¹

Шпилев П.В., доцент кафедры статистического моделирования
СПбГУ, p.shpilev@spbu.ru

Аннотация

Работа посвящена исследованию задачи построения L -оптимальных планов для тригонометрической модели без свободного члена. Для заданной диагональной матрицы L получены аналитические результаты для тригонометрических моделей без свободных членов сколь угодно большого порядка кратного 3 ($m = 2k$, k -натуральное число).

Введение

Тригонометрические регрессионные модели Фурье широко используются на практике для описания периодических процессов и до настоящего времени внимание исследователей было ограничено классической моделью в которой свободный член предполагается не равным нулю. Вместе с тем, на практике возникают ситуации, когда нулевой отклик, то есть начальное положение объекта экспериментирования, уже известен или эта информация не важна. В этих случаях целесообразнее использовать модели без свободного члена. Подобные модели до настоящего времени еще мало изучены. Для полиномиальной регрессионной модели без свободного члена недавно был получен ряд результатов для c - и e_k - критериев ([1], [2], [3]).

Для тригонометрической модели без свободного члена в статье [4] автором (совместно с Меласом В.Б.) были получены аналитические результаты, определяющие свойства информационных матриц для планов специального вида. В настоящей работе, данные результаты использованы для построения L -оптимальных планов для моделей порядка $m = 3k$, $k = 1, 2, 3, \dots$.

¹ Работа частично поддержана грантом РФФИ: 20-01-00096-а.

Постановка задачи

Рассмотрим регрессионную модель Фурье без свободного члена:

$$y = \sum_{i=1}^m \theta_{2i-1} \sin(ix) + \sum_{i=1}^m \theta_{2i} \cos(ix) + \epsilon = \theta^T f(x) + \epsilon, \quad t \in \chi, \quad (1)$$

где $f(x) = (\sin(x), \cos(x), \dots, \sin(mx), \cos(mx))^T$ вектор регрессионных функций, $\theta = (\theta_1, \dots, \theta_{2m})^T$ - вектор неизвестных параметров, ϵ - случайная величина, характеризующая ошибки наблюдения, $\chi = [-\pi, \pi]$ - интервал планирования. Результаты различных экспериментов предполагаются независимыми. Под (непрерывным) планом эксперимента мы будем понимать вероятностную меру ξ с конечным носителем на интервале планирования $[-\pi, \pi]$.

Информационной матрицей Фишера плана ξ называется матрица

$$M(\xi) = \left(\int_{-\pi}^{\pi} f(t) f(t)^T d\xi(t) \right) \in R^{2m \times 2m}$$

Вырожденным планом называется план, информационная матрица которого вырожденная.

Введем несколько специальных классов, которые будем использовать в дальнейшем, а также дадим определение L -оптимального плана в общем случае:

Определение 1 Для заданной матрицы

$$L = \sum_{i=1}^k l_i l_i^T$$

с векторами $l_i \in R^{2m}$, **класс** Ξ_L определим, как множество всех непрерывных планов эксперимента для которых линейная комбинация параметров $l_i^T \beta, i = 0, \dots, k$ оцениваема, то есть вектор l_i принадлежит пространству столбцов матрицы $M(\xi) : l_i \in \mathcal{R}(M(\xi)) i = 0, \dots, k$.

Определение 2 Будем говорить, что непрерывный план η принадлежит **классу** Ξ_L^* , если $\eta \in \Xi_L$ и для любого непрерывного плана ξ существует предел

$$\lim_{\alpha \rightarrow 0} f^T(t) M^+(\xi_\alpha) L M^+(\xi_\alpha) f(t) = f^T(t) M^+(\eta) L M^+(\eta) f(t), \quad (2)$$

где $\xi_\alpha = (1 - \alpha)\eta + \alpha\xi$, $\alpha \in [0, 1]$.

Определение 3 План ξ^* будем называть *L-оптимальным*, если $\xi^* \in \Xi_L$

$$\xi^* = \arg \min_{\xi \in \Xi_L} \text{tr} LM(\xi)^+,$$

где матрица L фиксированная неотрицательно определенная и $M(\xi)^+$ - обобщенно-обратная в смысле Мура-Пенроуза матрица для $M(\xi)$.

Определение 4 Будем говорить, что непрерывный план η является **абсолютно симметричным** (принадлежит классу абсолютно симметричных планов), если данный план имеет следующий вид:

$$\begin{aligned} \eta &= \begin{pmatrix} \bar{t}_1 & \dots & \bar{t}_{4n+2} \\ \bar{\omega}_1 & \dots & \bar{\omega}_{4n+2} \end{pmatrix} = \begin{pmatrix} -\eta_{\frac{n}{2}} & \eta_{\frac{n}{2}} \end{pmatrix}, \quad \text{где} \\ -(-\eta_{\frac{n}{2}}) &= \eta_{\frac{n}{2}} = \begin{pmatrix} t_1 & \dots & t_n & \frac{\pi}{2} & \pi - t_n & \dots & \pi - t_1 \\ \omega_1 & \dots & \omega_n & \omega_{n+1} & \omega_n & \dots & \omega_1 \end{pmatrix}, \end{aligned}$$

точки $t_i \in [0, \frac{\pi}{2}]$, а веса удовлетворяют равенству: $2 \sum_{i=1}^n \omega_i + \omega_{n+1} = \frac{1}{2}$.

Удобным инструментом проверки плана на *L-оптимальность*, является следующая теорема (см. [5]):

Теорема 1 Пусть матрица $L \in R^{2m \times 2m}$ фиксированная, неотрицательно определенная матрица. Имеют место следующие утверждения:

(a) План $\xi \in \Xi_L$ тогда и только тогда, когда

$$l_i^T M(\xi)^- M(\xi) = l_i^T, \quad i = 0, \dots, 2m,$$

где $M(\xi)^-$ - обобщено-обратная для $M(\xi)$ матрица.

(b) План $\xi^* \in \Xi_L$ является *L-оптимальным* тогда и только тогда, когда существует такая матрица $M^-(\xi^*)$, что

$$\max_{t \in \chi} \varphi(t, \xi^*) = \text{tr} LM^+(\xi^*),$$

где $\varphi(t, \xi) = f^T(t) M(\xi)^- L M(\xi)^- f(t)$. При этом в точках $t_i \in \text{supp}(\xi^*)$ имеет место равенство

(c)

$$\varphi(t_i) = \text{tr} LM(\xi^*)^+.$$

Доказательство данной теоремы повторяет стандартные рассуждения для невырожденного случая и здесь приводиться не будет.

Замечание Отметим, что в случае, когда оптимальный план ξ^* является невырожденным, матрица $M^-(\xi^*)$ из условия (b) совпадает с $M^{-1}(\xi^*)$, а в случае, когда он вырожден и принадлежит классу Ξ_L^* , эта матрица совпадает с $M^+(\xi^*)$. Наиболее сложным для исследования является случай, когда оптимальный план $\xi^* \in \Xi_L$ и $\notin \Xi_L^*$. Для этого случая некоторые методы построения экстремального полинома $\varphi(t, \xi^*)$ рассмотрены в работе [5].

L-оптимальные планы для моделей порядка $m = 3k$

Теоретические результаты, описывающие свойства информационной матрицы абсолютно симметричного плана, получены в работе [4] (Теоремы 3.1, 3.2). Нижеприведенная теорема является прямым следствием данных результатов:

Теорема 2 Рассмотрим абсолютно симметричный план $\eta^{(1)}$ следующего вида:

$$\eta^{(1)} = \begin{pmatrix} -\eta_{\frac{\pi}{2}}^{(1)} & \eta_{\frac{\pi}{2}}^{(1)} \end{pmatrix}, \quad \text{где}$$

$$-\left(-\eta_{\frac{\pi}{2}}^{(1)}\right) = \eta_{\frac{\pi}{2}}^{(1)} = \begin{pmatrix} x_1 & x_2 & \pi/2 & \pi - x_2 & \pi - x_1 \\ \omega_1 & \omega_2 & 1/2 - 2\omega_2 - 2\omega_1 & \omega_2 & \omega_1 \end{pmatrix}.$$

За $\eta_{\frac{\pi}{2}+t}^{(k)}$ обозначим следующую меру (не нормированную):

$$\eta_{\frac{\pi}{2}+t}^{(k)} = \begin{pmatrix} \frac{x_1+t}{k} & \frac{x_2+t}{k} & \frac{\pi/2+t}{k} & \frac{\pi-x_2+t}{k} & \frac{\pi-x_1+t}{k} \\ \frac{\omega_1}{k} & \frac{\omega_2}{k} & \frac{1/2-2\omega_2-2\omega_1}{k} & \frac{\omega_2}{k} & \frac{\omega_1}{k} \end{pmatrix}.$$

Для тригонометрической модели (1) третьего порядка ($m = 3$) план $\eta^{(1)}$ является L-оптимальным планом для диагональной матрицы $L^{(1)}$ с произвольным сочетанием нулей и единиц на главной диагонали (значения точек x_1, x_2 и весов ω_1, ω_2 найденных, как решения соответствующей оптимизационной задачи, для всевозможных вариантов матрицы L , представлены в Таблицах 1,2,3). Для тригонометрической модели (1) произвольного порядка кратного 3 ($m = 3k, k = 1, 2, \dots$) следующий план (для матрицы $L^{(k)}$ у которой все элементы нулевые, кроме диагональных элементов, стоящих на позициях кратных k (и совпадающих с соответствующими элементами матрицы $L^{(1)}$)) является L-оптимальным:

$$\eta^{(k)} = \begin{pmatrix} -\eta_{\frac{\pi}{2}+(k-1)\pi}^{(k)} & \dots & -\eta_{\frac{\pi}{2}}^{(k)} & \eta_{\frac{\pi}{2}}^{(k)} & \dots & \eta_{\frac{\pi}{2}+(k-1)\pi}^{(k)} \end{pmatrix}$$

Доказательство Теоремы 2: В том, что для тригонометрической модели(1) третьего порядка и диагональной матрицы $L^{(1)}$ (с произвольным сочетанием нулей и единиц на главной диагонали) L -оптимальный план принадлежит классу абсолютно симметричных планов (т.е. может быть представлен в виде $\eta^{(1)}$) легко убедиться непосредственно, решив соответствующую оптимизационную задачу и проверив, что для найденных точек и весов (см. Таблицы 1,2,3) выполнены условия Теоремы 1 (т.е. что найденное решение является глобальным минимумом). То, что для заданной матрицы $L^{(k)}$ (для тригонометрической модели порядка кратного 3 ($m = 3k$, $k = 1, 2, \dots$)) оптимальным является план $\eta^{(k)}$, следует из результатов, полученных в работе [4] (Теоремы 3.1, 3.2).

Замечание Отметим, что в таблицу не включены 2 случая:

- 1) диагональная матрица L содержит одну единицу (остальные нули);
- 2) матрица L является единичной. В первом случае L -оптимальный план называется e_k -оптимальным и для классической регрессионной модели порядка m в явном виде найден в работе [6] (Theorem 1,2). Несложно показать, найденные планы будут e_k -оптимальными и для модели без свободного члена. Во втором случае, как показано в работе [4] (Теорема 3.3), L -оптимальный план совпадает с D -оптимальным (вид которого хорошо известен).

Матрица L	ω_1	ω_2	x_1	x_2	$trLM^+$
diag{0,1,1,1,1,1}	0.0531	0.1242	0.1340	0.8354	9.8647
diag{1,0,1,1,1,1}	0.0576	0.1091	0.5234	0.5237	9.0000
diag{1,1,0,1,1,1}	0.0616	0.1334	0.1353	0.8013	9.8990
diag{1,1,1,0,1,1}	0.1423	0.1049	0.3733	1.1367	9.8647
diag{1,1,1,1,0,1}	0.0833	0.1667	0.0000	1.0472	9.0000
diag{1,1,1,1,1,0}	0.0850	0.1002	0.2079	0.8500	9.8990

Таблица 1: Значения весов и точек L -оптимальных планов из Теоремы 2 для вектора переупорядоченных регрессионных функций: $\tilde{f}(t) = (\cos(t), \cos(3t), \dots, \cos(2t), \cos(4t), \dots, \sin(t), \sin(3t), \dots, \sin(2t), \sin(4t), \dots)^T$.

Литература

- [1] H. Dette, V. B. Melas, and P. V. Shpilev, “Optimal designs for estimating individual coefficients in polynomial regression with no intercept.,” *Statistics and Probability Letters*, vol. 158, p. 108636, 2020.

Матрица L	ω_1	ω_2	x_1	x_2	$trLM^+$
diag{0,0,1,1,1,1}	0.0000	0.1378	0.0819	0.6404	6.7062
diag{0,1,0,1,1,1}	0.0476	0.1402	0.1663	0.8433	7.6990
diag{0,1,1,0,1,1}	0.0858	0.1077	0.2550	0.8712	8.0000
diag{0,1,1,1,0,1}	0.0635	0.1865	0.0052	1.0472	6.8777
diag{0,1,1,1,1,0}	0.1064	0.1274	0.3870	1.1969	7.7692
diag{1,0,0,1,1,1}	0.0000	0.1792	0.5587	0.5236	6.9220
diag{1,0,1,0,1,1}	0.0108	0.1865	1.5704	0.5236	6.8780
diag{1,0,1,1,0,1}	0.1258	0.0822	0.4075	1.0695	8.0000
diag{1,0,1,1,1,0}	0.0601	0.0912	0.5238	0.5235	6.9102
diag{1,1,0,0,1,1}	0.1053	0.1108	0.3105	0.8685	7.6990
diag{1,1,0,1,0,1}	0.0708	0.1792	0.0000	1.0472	6.9220
diag{1,1,0,1,1,0}	0.0832	0.1095	0.2417	0.8600	8.0000
diag{1,1,1,0,0,1}	0.1122	0.1378	0.0000	0.9304	6.7062
diag{1,1,1,0,1,0}	0.1354	0.0870	0.3237	1.0707	7.7692
diag{1,1,1,1,0,0}	0.0988	0.1512	0.0007	1.0472	6.9138

Таблица 2: Значения весов и точек L -оптимальных планов из Теоремы 2 для вектора переупорядоченных регрессионных функций: $\tilde{f}(t) = (\cos(t), \cos(3t), \dots, \cos(2t), \cos(4t), \dots, \sin(t), \sin(3t), \dots, \sin(2t), \sin(4t), \dots)^T$.

- [2] V. B. Melas and P. Shpilev, “Constructing c-optimal designs for polynomial regression without an intercept,” *Vestnik St. Petersburg University: Mathematics*, vol. 53, pp. 223–231, 2020.
- [3] H. Dette, V. B. Melas, and P. V. Shpilev, “A note on optimal designs for estimating the slope of a polynomial regression.,” *Statistics and Probability Letters*, vol. 170, p. 108992, 2021.
- [4] Мелас В. Б., и Шпилев П. В., “ L -оптимальные планы для регрессионной модели Фурье без свободного члена,” *Вестник СПбГУ. Математика. Астрономия.*, Том 9(1), р. 64–75, 2022.
- [5] P. V. Shpilev, “Equivalence theorem for singular l -optimal designs,” *Vestnik St. Petersburg University. Mathematics*, vol. 48, no. 1, p. 29–34, 2015.
- [6] H. Dette, V. B. Melas, and P. V. Shpilev, “Optimal designs for estimating the coefficients of the lower frequencies in trigonometric regression models,” *Ann. Inst. Statist. Math.*, vol. 59, pp. 655–673, 2007.

Матрица L	ω_1	ω_2	x_1	x_2	$trLM^+$
diag{1,1,1,0,0,0}	0.1333	0.1167	0.0000	1.0169	4.4736
diag{1,1,0,1,0,0}	0.0833	0.1667	0.0005	1.0472	5.0000
diag{1,1,0,0,1,0}	0.0749	0.1271	0.0657	0.7553	5.8634
diag{1,1,0,0,0,1}	0.0851	0.1649	0.0000	0.8597	4.3935
diag{1,0,1,1,0,0}	0.1137	0.0847	0.3106	1.0365	5.8284
diag{1,0,1,0,1,0}	0.1667	0.0000	0.5236	1.4149	5.0000
diag{1,0,1,0,0,1}	0.1176	0.1324	0.0000	0.7575	5.4861
diag{1,0,0,1,1,0}	0.0302	0.1365	0.5245	0.5234	5.0000
diag{1,0,0,1,0,1}	0.0978	0.1146	0.3622	0.9434	5.8284
diag{1,0,0,0,1,1}	0.0104	0.1999	0.5225	0.5237	4.3981
diag{0,1,1,1,0,0}	0.0833	0.1667	0.0053	1.0472	5.0000
diag{0,1,1,0,1,0}	0.1186	0.0909	0.3281	1.1002	5.8284
diag{0,1,1,0,0,1}	0.0833	0.1667	0.0000	1.0472	5.0000
diag{0,1,0,1,1,0}	0.0915	0.1094	0.3764	1.0304	5.8634
diag{0,1,0,1,0,1}	0.0397	0.2103	0.0007	1.0472	4.3981
diag{0,1,0,0,1,1}	0.0518	0.1464	0.0150	0.7856	5.8284
diag{0,0,1,1,1,0}	0.0000	0.1167	0.4086	0.5539	4.4736
diag{0,0,1,1,0,1}	0.1225	0.0099	0.8133	0.8133	5.4861
diag{0,0,1,0,1,1}	0.1667	0.0000	0.5236	1.3961	5.0000
diag{0,0,0,1,1,1}	0.0000	0.1650	0.5226	0.7105	4.3935
diag{1,1,0,0,0,0}	0.1100	0.1400	0.0000	0.9242	2.7044
diag{1,0,1,0,0,0}	0.1563	0.0937	0.0001	0.9553	3.2500
diag{1,0,0,1,0,0}	0.0830	0.1174	0.2077	0.8988	4.0000
diag{1,0,0,0,1,0}	0.1147	0.0853	0.5234	0.5238	2.7778
diag{1,0,0,0,0,1}	0.2457	0.0043	0.5236	0.5207	2.6667
diag{0,1,1,0,0,0}	0.1022	0.1478	0.0000	1.1668	2.6180
diag{0,1,0,1,0,0}	0.0500	0.2000	0.0004	1.0472	2.7778
diag{0,1,0,0,1,0}	0.0626	0.1249	0.0164	0.7857	4.0000
diag{0,1,0,0,0,1}	0.1910	0.0590	0.9425	0.0000	2.6180
diag{0,0,1,1,0,0}	0.0857	0.0080	0.6155	0.6155	3.2500
diag{0,0,1,0,1,0}	0.1371	0.0108	0.4040	0.4040	2.6180
diag{0,0,1,0,0,1}	0.0032	0.1252	0.1224	0.8149	4.0000
diag{0,0,0,1,1,0}	0.1401	0.0056	0.6476	1.5708	2.7043
diag{0,0,0,1,0,1}	0.1896	0.0604	1.0471	1.0475	2.6667
diag{0,0,0,0,1,1}	0.1547	0.0363	0.6283	0.6283	2.6180

Таблица 3: Значения весов и точек L -оптимальных планов из Теоремы 2 для вектора переупорядоченных регрессионных функций: $\tilde{f}(t) = (\cos(t), \cos(3t), \dots, \cos(2t), \cos(4t), \dots, \sin(t), \sin(3t), \dots, \sin(2t), \sin(4t), \dots)^T$.

О равномерной состоятельности непараметрического критерия Неймана

Ермаков М. С., д. ф.-м. н., профессор кафедры статистического моделирования СПбГУ, erm2512@mail.ru¹

Капаца Д. Ю., студент-магистр кафедры статистического моделирования
СПбГУ, david@kapatsa.com

Аннотация

Для задачи проверки гипотезы согласия в работе исследуется равномерная состоятельность критериев против непараметрических множеств альтернатив, когда тестовой статистикой является бесконечная линейная комбинация квадратов оценок коэффициентов Фурье разложения плотности распределения в ряд Фурье.

Получены условия равномерной состоятельности, близкие к необходимым. Доказана асимптотическая нормальность тестовой статистики при гипотезе или альтернативах.

Введение

Состоительность непараметрических критериев при параметрическом задании альтернатив довольно хорошо изучена ([2], [15], [16], [18]). С конца восьмидесятых годов прошлого века началось интенсивное исследование задач непараметрической проверки гипотез при непараметрических множествах альтернатив ([12], [9], [10], [11], [4]). До этих публикаций данная проблематика исследовалась только в отдельных работах ([17], [13], [14]).

Изучению скорости сходимости классических непараметрических критериев при этом уделялось немного внимания. В то же время, на тот момент уже были получены определённые результаты о состоятельности критериев против непараметрических множеств альтернатив ([17], [16], [15]).

Следует отметить, что исследования в конце восьмидесятых годов прошлого века начались под влиянием крайне интенсивного в этот период развития теории непараметрического оценивания. В свою очередь, развитие теории непараметрического оценивания базировалась на уже созданные к тому времени теории аппроксимации и в какой-то мере на теории информации. В результате исследований теория непараметрического оценивания в каком-то смысле стала аналогом теории аппроксимации, но с введённой случайной

¹Исследование поддержано грантом РФФИ 20-01-00273

ошибкой. Как следствие, теория непараметрической проверки гипотез — использующая схожую технику — есть не что иное, как теория проверки гипотез в функциональных пространствах.

Принятие гипотезы носит рекомендательный характер и говорит лишь о том, что данные не сильно противоречат модели. Понятно, что взяв большое количество данных, мы всегда сможем опровергнуть простую статистическую модель. Поэтому важным фактором в выборе критерия является различимость гипотезы и альтернативы в асимптотическом сценарии. Описание наибольших множеств альтернатив — способных быть «различимыми» непараметрическими критериями — получило развитие в работах ([3], [6], [8], [1]). В них авторы, по сути, показывали состоятельность критерииев против альтернатив, изучаемых в рамках метода расстояний (distance method), на основе которого построены базовые статистические непараметрические критерии (в том числе построенный в данной работе).

Метод расстояний Опишем метод расстояний в контексте проверки гипотезы согласия. Пусть дана выборка независимых одинаково распределённых случайных наблюдений $X^{(n)} = X_1, X_2, \dots, X_n$, имеющих функцию распределения $F(x)$, $x \in (0, 1)$. Мы проверяем простую гипотезу

$$\mathbb{H}_0 : F(x) = F_0(x)$$

против альтернатив

$$\mathbb{H}_n : F(x) \neq F_0(x).$$

Для проверки гипотезы мы задаём тестовые статистики $T_n(\hat{F}_n)$, где

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i < x\}}$$

— эмпирическая функция распределения. Здесь $T_n(F) = \bar{T}_n(F - F_0)$ — функционал, заданный на множестве всех функций распределения \mathcal{F} . В качестве $\bar{T}_n(F - F_0)$ обычно берётся некоторая псевдометрика, заданная на множестве $\mathcal{F} - \mathcal{F}$ разностей функций распределения.

Например, для критерия Колмогорова

$$T_n(X^{(n)}) = \sup_x |F_0(x) - \hat{F}_n(x)|,$$

критерия фон Мизеса

$$T_n(X^{(n)}) = \int_{-\infty}^{+\infty} (\hat{F}_n(x) - F_0(x))^2 dF_0(x).$$

Равномерная состоятельность Ясно, что мы не можем состоятельно проверить гипотезу \mathbb{H}_0 для всех возможных функциональных альтернатив. В работах [5], [7], [8] показывалось, что возможно состоятельно проверить гипотезу для таких альтернатив F , для которых $T(F) > \rho_n$, где $\rho_n \rightarrow 0$ специальным образом при $n \rightarrow \infty$.

Равномерная состоятельность в теории проверки гипотез при этом рассматривается обычно в двух наиболее распространённых смыслах. Пусть стоит задача проверки гипотезы

$$\mathbb{H}_0 : F = F_0 \quad (1)$$

против альтернатив

$$\mathbb{H}_n : F_n \in \mathcal{F}_n \subset \mathcal{F}. \quad (2)$$

Для критерия K_n проверки гипотезы обозначим $\alpha(K_n) := \mathbf{E}_{F_0} K_n$ — вероятность ошибки первого рода и $\beta(K_n, F_n) := \mathbf{E}_{F_0}(1 - K_n)$ — вероятность ошибки второго рода при альтернативе $F_n \in \mathcal{F}_n$.

Скажем, что последовательность критериев K_n , где $0 < \alpha(K_n) = \alpha < 1$, *слабо равномерно состоятельна*, если

$$\lim_{n \rightarrow \infty} \sup_{F \in \mathcal{F}_n} \beta(K_n, F) = 0,$$

и *равномерно состоятельна*, если

$$\lim_{n \rightarrow \infty} \sup_{F \in \mathcal{F}_n} \beta(K_n, F) < 1 - \alpha. \quad (3)$$

Слабая равномерная состоятельность непараметрических критериев довольно хорошо изучена ([16], [10], [11]). Равномерная состоятельность (3) изучается по настоящее время; её изучению для случая непараметрического критерия Неймана посвящена настоящая работа.

Приведём одно следствие из равномерной состоятельности. В работе Ле Кама [14] утверждается, что если критерии равномерно состоятельны против множества альтернатив $\mathcal{F}_0 \subset \mathcal{F}$ в следующем смысле: если существует n_0 и критерий K_{n_0} , $0 < \alpha(K_{n_0}) = \alpha < 1$ такой, что

$$\sup_{F \in \mathcal{F}_0} \beta(K_{n_0}, F) < 1 - \alpha,$$

то существуют $c, c_1 > 0$ и последовательность критериев K_n такая, что

$$\sup_{F \in \mathcal{F}_n} \beta(K_n, F) < c_1 \exp\{-Cn\}.$$

Также, если непараметрические критерии равномерно состоятельны для $r_n \rightarrow 0$ при $n \rightarrow \infty$, то они слабо равномерно состоятельны против множеств альтернатив

$$\mathcal{F}_{1n} = \{F : \bar{T}(F - F_0) > r_n > 0, F \in \mathcal{F}\},$$

где $\rho_n/r_n \rightarrow \infty$ при $n \rightarrow \infty$.

Постановка задачи

Далее предполагается, что рассматриваемые в задаче функции распределения абсолютно непрерывны, а соответствующие им плотности являются ограниченными в пространстве L_2 . Поэтому можем рассматривать задачу в терминах плотностей распределения.

Пусть X_1, \dots, X_n — независимые одинаково распределённые с.в., имеющие неизвестную плотность распределения $p(x)$ в $L_2(\nu)$. Предполагается, что плотность представима в виде

$$p(x) = 1 + f(x), \quad f(x) = \sum_{j=1}^{\infty} \theta_j \varphi_j(x) \quad (4)$$

где $p(x)$ — плотность распределения, $\{\varphi_j(x)\}_0^{\infty}$ — ортонормированная система функций в $L_2(\nu)$, ν — вероятностная мера, $\varphi_0(x) = 1$ для всех $x \in \Omega$. Функция $f(x)$ предполагается принадлежащей некоторому компакту \mathcal{U} из L_2 .

Требуется проверить гипотезу о равномерности распределения наблюдаемых случайных величин

$$\mathbb{H}_0 : p(x) = 1, \quad x \in [0, 1], \quad (5)$$

т.е. $\theta_j = 0$ для всех j против альтернатив

$$\mathbb{H}_n : f \in \Psi_n(R_n, c), \quad \Psi_n = \{f : R_n(f) > c, f \in \mathcal{U}\}, \quad (6)$$

где

$$R_n(f) = A_n(\boldsymbol{\theta}) = n^2 \sum_{j=1}^{\infty} \varkappa_{nj}^2 \theta_j^2, \quad (7)$$

где \varkappa_{nj}^2 — некоторая убывающая последовательность весов, которая удовлетворяет условиям В1–В3 (см. далее). Соответствующие функциям $f \in \Psi_n(R_n, c)$ бесконечномерные векторы $\boldsymbol{\theta}$ условимся относить к множеству Q_n .

Критерии типа Неймана

Последовательность весов \varkappa_{nj} выбирается из условий В1–В3. Введём

$$R_n(f) = A_n(\theta) = n^2 \sum_{j=1}^{\infty} \varkappa_{nj}^2 \theta_j^2, \quad A_n = n^2 \sum_{j=1}^{\infty} \varkappa_{nj}^4.$$

$f(x) = \sum_{j=1}^{\infty} \theta_j \varphi_j(x)$ — разложение функции $f(x)$ по ортонормированному базису. Определим тестовые статистики

$$T_n(X^{(n)}) = \sum_{j=1}^{\infty} \varkappa_{nj}^2 \left(\sum_{s=1}^{\infty} \varphi_j(X_s) \right)^2. \quad (8)$$

Данная статистика, помимо условий на веса \varkappa_{nj}^2 , отличается от предложенной в [5] тем, что соответствующие суммы по \varkappa_{nj}^2 имеют бесконечное число слагаемых.

Условия на последовательности $\{\varkappa_j^2\}_{j=1}^{\infty}$

B1. Для любого n последовательность \varkappa_{nj}^2 является убывающей.

B2. Существуют константы C_1 и C_2 такие, что для любого n

$$C_1 < A_n = n^2 \sum_{j=1}^{\infty} \varkappa_{nj}^4 < C_2. \quad (9)$$

B3. Существуют C_1 и $\lambda > 1$, что для любого $\delta > 0$ и n имеет место

$$\varkappa_{[n,(1+\delta)k_n]}^2 < C_1(1+\delta)^{-\lambda} \varkappa_n^2,$$

где $k_n = \sup \left\{ k : \sum_{j < k} \varkappa_{nj}^2 \leq \frac{1}{2} \sum_{j=1}^{\infty} \varkappa_{nj}^2 < \infty \right\}$. Также имеет место конечность сумм $\sum_{j=1}^{\infty} \varkappa_{nj} < C < \infty$.

Для любого события D обозначим за $\chi(D)$ индикатор этого события. Зададим критерии

$$K_n(X^{(n)}) = \chi \left(T_n(X^{(n)}) - n \sum_{j=1}^k \varkappa_j^2 > (2A_n)^{1/2} x_{\alpha_n} \right).$$

Задачей исследования является обобщение результатов работы [5] на большее множество критериев² и на более широкий класс альтернатив. Таким образом, главным итогом работы является доказательство следующей теоремы.

Теорема 1 (О равномерной состоятельности критериев типа Неймана)

Пусть выполнены условия В1–В3, а также условие ограниченности системы ортонормированных функций $\{\varphi_j\}_{j=1}^{\infty}$. Тогда для последовательности критериев K_n имеют место $\alpha(K_n) = \alpha + o(1)$ и

$$\beta(K_n, f_n) = \Phi(x_{\alpha} - R_n(\boldsymbol{\theta}_n)(2A_n)^{-1/2})(1 + o(1))$$

равномерно для всех последовательностей $\boldsymbol{\theta}_n$, для которых выполнено условие ограниченности f .

Из данного результата сразу следует результат о равномерной состоятельности последовательности заданных критериев.

Результаты

Для доказательства теоремы 1 достаточно доказать две следующих леммы.

Лемма 1 (Об асимптотической нормальности) Пусть $\boldsymbol{\theta} = \boldsymbol{\theta}^{(n)} \in \Psi_n$ и $\mathbf{D}_{\boldsymbol{\theta}}T_n = O(A_n)$, $A_n(\boldsymbol{\theta}) \asymp A_n$. $A_n(\boldsymbol{\theta}) := n^2 \sum_{j=1}^{\infty} \theta_j^2 \varkappa_{nj}^2$. Тогда $\mathrm{P}_{\boldsymbol{\theta}}$ -распределения тестовых статистик

$$\left(T_n(X^{(n)}) - n \sum_{j=1}^k \varkappa_{nj}^2 - A_n(\boldsymbol{\theta}) \right) (2A_n)^{-1/2}$$

асимптотически нормальны.

Доказательство леммы 1 основано на проверке условий теоремы из работы [19] и на оценках, схожих с построенными в следующей лемме.

Лемма 2 Пусть $\boldsymbol{\theta} = \boldsymbol{\theta}^{(n)} \in \Psi_n$. Тогда

$$\begin{aligned} \mathbf{E}_{\boldsymbol{\theta}}T_n &= A_n(\boldsymbol{\theta}) + o(A_n(\boldsymbol{\theta})), \\ \mathbf{D}_{\boldsymbol{\theta}}T_n &= 2A_n + o(A_n^2(\boldsymbol{\theta})). \end{aligned}$$

²за счёт неявного задания последовательностей весов \varkappa_{nj}^2

Заключение

По результатам работы для задачи проверки гипотезы о распределении построены асимптотические оценки для математического ожидания и дисперсии статистики типа Неймана в случае гипотезы и альтернативы, а также проведено доказательство асимптотической нормальности данной статистики при заданных условиях на рассматриваемое пространство альтернатив.

Таким образом, построено новое семейство непараметрических критериев о распределении, обладающее свойством равномерной состоятельности против широких множеств непараметрических альтернатив.

Список литературы

- [1] A. R. Barron, Uniformly powerful goodness of fit tests.— Ann. Statist., v.17 (1989), 107-124.
- [2] J. Durbin, Distribution Theory for Tests Based on the Sample Distribution function.— Regional Conference Series in Applied Mathematics, v.9 (1973), SIAM, Philadelphia.
- [3] J. Horowitz, V. Spokoiny, An Adaptive, Rate-Optimal Test of a Parametric Mean-Regression Model Against a Nonparametric Alternative.— Econometrica, v.69(3) (2001), 599-631.
- [4] М. С. Ермаков, Минимаксное обнаружение сигнала в гауссовском белом шуме.— Теория вероятн. и ее примен., в.35(4) (1990), 704-715.
- [5] М. С. Ермаков, Минимаксная проверка гипотез о плотности распределения.— Теория вероятн. и ее примен., в.39(5) (1994), 488-512.
- [6] М. С. Ермаков, Асимптотическая минимаксность критериев хи-квадрат.— Теория вероятн. и ее примен., в.42(4) (1997), 668-695.
- [7] M. S. Ermakov, On asymptotic minimaxity of kernel-based tests.— ESAIM Probab. Stat., v.7 (2003), 279-312.
- [8] М. С. Ермаков, Минимаксное обнаружение сигнала в весовом гауссовском белом шуме.— Зап. научн. сем. ПОМИ, в.320 (2004), 54-68.
- [9] И. А. Ибрагимов, Р. З. Хасьминский, Об оценке бесконечномерного параметра в гауссовом белом шуме.— Докл. АН СССР, в.236(5) (1977), 1053-1055.

- [10] Ю. И. Ингстер, О сравнении минимаксных свойств тестов Колмогорова, ω^2 и χ^2 .— Теория вероятн. и ее примен., в.32(2) (1987), 374-378.
- [11] Yu. I. Ingster, I. A. Suslina, Nonparametric Goodness-of-fit Testing under Gaussian Models.— Lecture Notes in Statistics, 169, Springer: N.Y., 2002.
- [12] E. Gine, R. Nickl, Mathematical Foundation of Infinite-Dimensional Statistical Models.— Cambridge University Press, Cambridge, 2015.
- [13] L. Le Cam, L. Schwartz, A necessary and sufficient conditions for the existence of consistent estimates.— Ann. Math. Statist., v.31 (1960), 140-150.
- [14] L. Le Cam, Convergence of estimates under dimensionality restrictions.— Ann. Statist., v.1 (1973), 38-53.
- [15] M. Кендалл, Ф. Стьюарт, Статистические выводы и связи.— М. : Наука, 1973, 810с.
- [16] E. L. Lehmann, Testing Statistical Hypotheses.— Wiley: N.Y., 1986, 604p.
- [17] H. Mann, A Wald, On the Choice of the Number of Class Intervals in the Application of Chi-Square Test.— Annals of Mathematical Statistics, v.13 (1942), 306-317.
- [18] G. R. Shorack, J. A. Wellner, Empirical Processes with Application to Statistics.— Wiley: N.Y, 1986.
- [19] P. Hall, Central limit theorem for integrated square error of multivariate nonparametric density estimators.— J. Multivar. Anal., v.14(1) (1984), 1-16.

Модели, методы и приложения тропической математики



Кривулин Николай Кимович

д.ф.-м.н., профессор кафедры статистического моделирования СПбГУ

РЕШЕНИЕ ЗАДАЧИ О РАЗМЕЩЕНИИ ДВУХ ОБЪЕКТОВ В ПРОСТРАНСТВЕ С МЕТРИКОЙ ЧЕБЫШЕВА

Кривулин Н. К., д. ф.-м. н., профессор кафедры статистического моделирования СПбГУ, nkk@math.spbu.ru

Брюшинин М. А., студент кафедры статистического моделирования
СПбГУ, st076630@student.spbu.ru

Аннотация

Рассматривается минимаксная задача размещения двух объектов в многомерном пространстве с метрикой Чебышева. Задача формулируется в терминах тропической математики, которая изучает алгебраические системы с идемпотентными операциями. Предлагается прямое аналитическое решение задачи на основе методов и результатов тропической оптимизации.

Введение

Решение задач размещения часто осложняется нелинейностью и негладкостью целевой функции и ограничений. В некоторых случаях решение подобных задач можно упростить путем их представления на языке тропической математики и использования ее результатов. Тропическая (идемпотентная) математика охватывает область, связанную с изучением теории полуполей с идемпотентным сложением и ее приложениями. Решение задач размещения в пространствах с чебышевской и прямоугольной (манхэттенской) метрикой с помощью тропической математики представлены в [2, 3, 5, 9]. В настоящей статье рассматривается минимаксная задача размещения двух объектов в многомерном пространстве с метрикой Чебышева. Сначала представлен краткий обзор основных понятий и результатов тропической математики, необходимых для последующего решения задачи. Задача размещения формулируется в терминах тропической математики, а затем предлагается прямое аналитическое решение задачи на основе методов и результатов тропической оптимизации.

Элементы тропической математики

В этом разделе представлены базовые понятия и результаты тропической математики [1, 2, 4], на основе которых проведено решение.

Рассмотрим непустое множество \mathbb{X} , на котором определены операции сложения \oplus и умножения \otimes . По сложению \mathbb{X} является идемпотентным коммутативным моноидом с нейтральным элементом \emptyset . По умножению множество $\mathbb{X} \setminus \{\emptyset\}$ образует абелеву группу с нейтральным элементом $\mathbb{1}$. В итоге имеем набор $\langle \mathbb{X}, \emptyset, \mathbb{1}, \oplus, \otimes \rangle$.

На \mathbb{X} задан частичный порядок, индуцированный идемпотентостью сложения: $x \leq y$ тогда и только тогда, когда $x \oplus y = y$. Отсюда следует равносильность неравенства $x \oplus y \leq z$ неравенствам $x \leq z$ и $y \leq z$. Векторные неравенства рассматриваются как покомпонентные.

Справедлив тропический аналог неравенства между арифметическим и геометрическим средними:

$$(x_1 \oplus x_2 \oplus \cdots \oplus x_n)n^{-1} \geq (x_1 \otimes x_2 \otimes \cdots \otimes x_n)^{1/n}.$$

Операции \oplus и \otimes монотонны по каждому аргументу относительно указанного порядка. Умножение дистрибутивно относительно сложения, и для любого $x \in \mathbb{X} \setminus \{\emptyset\}$ существует обратный по умножению x^{-1} такой, что $x \otimes x^{-1} = \mathbb{1}$. Для тех же иксов и натуральных p определены степени $x^0 = \mathbb{1}$, $\emptyset^p = \emptyset$, $x^p = x^{p-1} \otimes x$, $x^{-p} = (x^{-1})^p$. Далее для простоты записи знак умножения \otimes опущен.

Множество \mathbb{X} не является группой по сложению, так как для его произвольного элемента противоположного относительно сложения не существует. Подобную структуру называют идемпотентным полуполем.

Примером служит $\mathbb{R}_{max,+} = \langle \mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, + \rangle$. В данном вещественном полуполе сложение определено как \max , умножение как $+$, нейтральный элемент относительно сложения есть $-\infty$, а относительно умножения — 0 .

Операции сложения и умножения матриц выполняются по обычным правилам с заменой соответствующих покомпонентных операций на \oplus и \otimes .

Мультипликативно сопряженным транспонированием вектора $\mathbf{x} = (x_j)$ называют преобразование, при котором \mathbf{x} переводится в вектор-строку $\mathbf{x}^- = (x_j^-)$ с элементами $x_j^- = x_j^{-1}$, если $x_j \neq 0$, и $x_j^- = 0$ иначе. Так, для ненулевого вектора \mathbf{x} справедливо равенство $\mathbf{x}^- \mathbf{x} = \mathbb{1}$. Спектральным радиусом матрицы \mathbf{A} называется её максимальное собствен-

ное число. Спектральный радиус может быть вычислен по формуле

$$\lambda = \bigoplus_{m=1}^n \text{tr}^{1/m}(\mathbf{A}^m).$$

Задача оптимизации

Задача размещения — частный случай задачи оптимизации. Методы решения задач оптимизации, сформулированных в терминах тропической математики, получены в [6, 7, 10]. В данном тексте используется следующий результат:

Теорема 1. Пусть \mathbf{A} — матрица со спектральным радиусом $\lambda > 0$, а \mathbf{q} — регулярный вектор. Тогда минимум в задаче

$$\min \mathbf{x}^\top \mathbf{Ax} + \mathbf{q}^\top \mathbf{x} + \mathbf{x}^\top \mathbf{p} + r$$

равен

$$\mu = \lambda \oplus \bigoplus_{m=1}^n (\mathbf{q}^\top \mathbf{A}^{m-1} \mathbf{p})^{1/(m+1)} \oplus r,$$

а все регулярные решения задачи имеют вид

$$\mathbf{x} = (\mu^{-1} \mathbf{A})^* \mathbf{u}, \quad \mu^{-1} \mathbf{p} \leq \mathbf{u} \leq \mu \left(\mathbf{q}^\top (\mu^{-1} \mathbf{A})^* \right)^-.$$

Доказательство представлено в [7, 8, 10].

Метрика Чебышева в терминах полуполя $\mathbb{R}_{\max,+}$

Определение. Метрикой Чебышёва называют максимум модулей разности соответствующих координат векторов:

$$d(\mathbf{r}, \mathbf{s}) = \max_{1 \leq i \leq n} |r_i - s_i|,$$

где $\mathbf{r} = (r_1, \dots, r_n)^T$ и $\mathbf{s} = (s_1, \dots, s_n)^T$ — произвольные векторы из \mathbb{R}^n .

В терминах полуполя $\mathbb{R}_{\max,+}$ метрика принимает следующий вид:

$$d(\mathbf{r}, \mathbf{s}) = \bigoplus_{i=1}^n (r_i s_i^{-1} \oplus r_i^{-1} s_i) = \mathbf{s}^\top \mathbf{r} \oplus \mathbf{r}^\top \mathbf{s}.$$

Задача размещения двух объектов в терминах полуполя $\mathbb{R}_{\max,+}$

Даны k векторов $\mathbf{r}_1, \dots, \mathbf{r}_k \in \mathbb{R}^n$ и l векторов $\mathbf{s}_1, \dots, \mathbf{s}_l \in \mathbb{R}^n$. Требуется найти векторы $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, которые решают задачу

$$\min_{\mathbf{x}, \mathbf{y}} \quad \max \left\{ \max_{1 \leq j \leq k} d(\mathbf{x}, \mathbf{r}_j), \max_{1 \leq j \leq l} d(\mathbf{y}, \mathbf{s}_j), d(\mathbf{x}, \mathbf{y}) \right\}.$$

В терминах полуполя $\mathbb{R}_{\max,+}$ задача записывается в следующей форме:

$$\min_{\mathbf{x}, \mathbf{y}} \quad \left\{ \bigoplus_{j=1}^k d(\mathbf{x}, \mathbf{r}_j) \oplus \bigoplus_{j=1}^l d(\mathbf{y}, \mathbf{s}_j) \oplus d(\mathbf{x}, \mathbf{y}) \right\}. \quad (1)$$

Лемма 1. Минимум в задаче (1) равен

$$\mu = (\mathbf{q}_1^- \mathbf{p}_1 \oplus \mathbf{q}_2^- \mathbf{p}_2)^{1/2} \oplus (\mathbf{q}_1^- \mathbf{p}_2 \oplus \mathbf{q}_2^- \mathbf{p}_1)^{1/3},$$

а все регулярные решения задачи имеют вид

$$\begin{aligned} \mathbf{x} &= \mathbf{u}_1 \oplus \mu^{-1} \mathbf{u}_2, \quad \mathbf{y} = \mu^{-1} \mathbf{u}_1 \oplus \mathbf{u}_2, \\ \mu^{-1} \mathbf{p}_1 &\leq \mathbf{u}_1 \leq \mu (\mathbf{q}_1^- \oplus \mu^{-1} \mathbf{q}_2^-)^-, \quad \mu^{-1} \mathbf{p}_2 \leq \mathbf{u}_2 \leq \mu (\mu^{-1} \mathbf{q}_1^- \oplus \mathbf{q}_2^-)^-, \end{aligned}$$

т.е.

$$\mathbf{p}_1 = \bigoplus_{j=1}^k \mathbf{r}_j, \quad \mathbf{p}_2 = \bigoplus_{j=1}^l \mathbf{s}_j, \quad \mathbf{q}_1^- = \bigoplus_{j=1}^k \mathbf{r}_j^-, \quad \mathbf{q}_2^- = \bigoplus_{j=1}^l \mathbf{s}_j^-.$$

Доказательство. Приведем целевую функцию к более компактному виду. Для начала обозначим:

$$\varphi(\mathbf{x}, \mathbf{y}) = \bigoplus_{j=1}^k d(\mathbf{x}, \mathbf{r}_j) \oplus \bigoplus_{j=1}^l d(\mathbf{y}, \mathbf{s}_j) \oplus d(\mathbf{x}, \mathbf{y}).$$

После подстановки выражений для функций расстояния и группировки слагаемых получим

$$\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^- \bigoplus_{j=1}^k \mathbf{r}_j \oplus \bigoplus_{j=1}^k \mathbf{r}_j^- \mathbf{x} \oplus \mathbf{y}^- \bigoplus_{j=1}^l \mathbf{s}_j \oplus \bigoplus_{j=1}^l \mathbf{s}_j^- \mathbf{y} \oplus \mathbf{y}^- \mathbf{x} \oplus \mathbf{x}^- \mathbf{y}.$$

Для последующего упрощения формулы обозначим:

$$\mathbf{z} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}, \quad \mathbf{q} = \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0 & \mathbf{I}_n \\ \mathbf{I}_n & 0 \end{pmatrix}.$$

С использованием введенных обозначений, получим следующий вид целевой функции:

$$\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{z}^\top \mathbf{A} \mathbf{z} \oplus \mathbf{q}^\top \mathbf{z} \oplus \mathbf{z}^\top \mathbf{p}.$$

Из теоремы 1 известно, что минимум для целевой функции равен

$$\mu = \lambda \oplus \bigoplus_{m=1}^{2n} (\mathbf{q}^\top \mathbf{A}^{m-1} \mathbf{p})^{1/(m+1)},$$

где λ — спектральный радиус, вычисляемый по формуле

$$\lambda = \bigoplus_{m=1}^{2n} \text{tr}^{1/m}(\mathbf{A}^m) = \bigoplus_{i=1}^n \text{tr}^{1/(2i-1)}(\mathbf{A}^{2i-1}) \oplus \bigoplus_{i=1}^n \text{tr}^{1/2i}(\mathbf{A}^{2i}).$$

Ниже будем использовать обозначение $\mathbf{J}_{2n} = \mathbf{A}$. Так как для всех целых $i \geq 0$ выполняется $\mathbf{A}^{2i} = \mathbf{I}_{2n}$, $\mathbf{A}^{2i-1} = \mathbf{J}_{2n}$, то $\text{tr } \mathbf{J}_{2n} = 0$, $\text{tr } \mathbf{I}_{2n} = 1$ и $\lambda = 1$. Также по теореме 1 μ примет вид:

$$\mu = \lambda \oplus \bigoplus_{i=1}^n (\mathbf{q}^\top \mathbf{I}_{2n} \mathbf{p})^{1/2i} \oplus \bigoplus_{i=1}^n (\mathbf{q}^\top \mathbf{J}_{2n} \mathbf{p})^{1/(2i+1)}.$$

Преобразуем слагаемые приведенных выше сумм для их явного представления:

$$\mathbf{q}^\top \mathbf{I}_{2n} \mathbf{p} = \mathbf{q}_1^\top \mathbf{p}_1 \oplus \mathbf{q}_2^\top \mathbf{p}_2, \quad \mathbf{q}^\top \mathbf{J}_{2n} \mathbf{p} = \mathbf{q}_1^\top \mathbf{p}_2 \oplus \mathbf{q}_2^\top \mathbf{p}_1.$$

Тогда μ принимает следующий вид:

$$\mu = \lambda \oplus \bigoplus_{i=1}^n (\mathbf{q}_1^\top \mathbf{p}_1 \oplus \mathbf{q}_2^\top \mathbf{p}_2)^{1/2i} \oplus \bigoplus_{i=1}^n (\mathbf{q}_1^\top \mathbf{p}_2 \oplus \mathbf{q}_2^\top \mathbf{p}_1)^{1/(2i+1)}.$$

Из определения сопряженного транспонирования следует, что

$$\mathbf{q}_1^\top \mathbf{p}_1 = \left(\bigoplus_{j=1}^k \mathbf{r}_j^\top \right) \left(\bigoplus_{j=1}^k \mathbf{r}_j \right) \geq 1,$$

аналогично $\mathbf{q}_2^- \mathbf{p}_2 \geq \mathbb{1}$.

Тогда результат суммирования будет равен первому слагаемому:

$$\bigoplus_{i=1}^n (\mathbf{q}_1^- \mathbf{p}_1 \oplus \mathbf{q}_2^- \mathbf{p}_2)^{1/2i} = (\mathbf{q}_1^- \mathbf{p}_1 \oplus \mathbf{q}_2^- \mathbf{p}_2)^{1/2}.$$

В свою очередь, из условия и тропического аналога неравенства между геометрическим и арифметическим средними имеем неравенство

$$\mathbf{q}_1^- \mathbf{p}_2 \oplus \mathbf{q}_2^- \mathbf{p}_1 \geq (\mathbf{q}_1^- \mathbf{p}_2 \mathbf{q}_2^- \mathbf{p}_1)^{1/2} \geq \mathbb{1}.$$

Следовательно,

$$\bigoplus_{i=1}^n (\mathbf{q}_1^- \mathbf{p}_2 \oplus \mathbf{q}_2^- \mathbf{p}_1)^{1/(2i+1)} = (\mathbf{q}_1^- \mathbf{p}_2 \oplus \mathbf{q}_2^- \mathbf{p}_1)^{1/3}.$$

Получаем:

$$\mu = (\mathbf{q}_1^- \mathbf{p}_1 \oplus \mathbf{q}_2^- \mathbf{p}_2)^{1/2} \oplus (\mathbf{q}_1^- \mathbf{p}_2 \oplus \mathbf{q}_2^- \mathbf{p}_1)^{1/3}.$$

Минимум μ целевой функции $\varphi(\mathbf{x}, \mathbf{y})$ по теореме 1 доставляют следующие регулярные векторы

$$\mathbf{z} = (\mu^{-1} \mathbf{A})^* \mathbf{u}, \quad \mu^{-1} \mathbf{p} \leq \mathbf{u} \leq \mu \left(\mathbf{q}^- (\mu^{-1} \mathbf{A})^* \right)^-$$

Приведём $(\mu^{-1} \mathbf{A})^*$ к явному виду:

$$(\mu^{-1} \mathbf{A})^* = \bigoplus_{i=0}^{2n-1} (\mu^{-i} \mathbf{A}^i) = \left(\bigoplus_{i=0}^{n-1} \mu^{-2i} \right) \mathbf{I}_{2n} \oplus \left(\bigoplus_{i=0}^{n-1} \mu^{-(2i+1)} \right) \mathbf{J}_{2n}.$$

Так как $\mu \geq \mathbb{1}$, то $\bigoplus_{i=0}^{n-1} \mu^{-2i} = \mu^0 = \mathbb{1}$, также $\bigoplus_{i=0}^{n-1} \mu^{-(2i+1)} = \mu^{-1}$, таким образом,

$$(\mu^{-1} \mathbf{A})^* = \mathbf{I}_{2n} \oplus \mu^{-1} \mathbf{J}_{2n} = \begin{pmatrix} \mathbf{I}_n & \mu^{-1} \mathbf{I}_n \\ \mu^{-1} \mathbf{I}_n & \mathbf{I}_n \end{pmatrix}.$$

Обозначим блоки вектора \mathbf{u} :

$$\mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}.$$

Тогда имеем

$$\mathbf{x} = \mathbf{u}_1 \oplus \mu^{-1} \mathbf{u}_2, \quad \mathbf{y} = \mu^{-1} \mathbf{u}_1 \oplus \mathbf{u}_2,$$

$$\mu^{-1} \mathbf{p}_1 \leq \mathbf{u}_1 \leq \mu (\mathbf{q}_1^- \oplus \mu^{-1} \mathbf{q}_2^-)^-, \quad \mu^{-1} \mathbf{p}_2 \leq \mathbf{u}_2 \leq \mu (\mu^{-1} \mathbf{q}_1^- \oplus \mathbf{q}_2^-)^-.$$

□

Заключение

В работе с помощью методов тропической алгебры была решена задача размещения двух объектов в чебышёвской метрике. В дальнейшем планируется обобщить решение на случаи задачи с числом объектов больше двух, задачи с весами.

Список литературы

- [1] Маслов В. П., Колокольцев В. Н. Идемпотентный анализ и его применение в оптимальном управлении. М. : Физматлит, 1994.
- [2] Krivulin N. Algebraic solution to a constrained rectilinear minimax location problem on the plane // 2011 International Conference on Multimedia Technology (ICMT), P. 6212–6220 <https://arxiv.org/pdf/1212.6089.pdf>
- [3] Krivulin N. An algebraic approach to multidimensional minimax location problems with Chebyshev distance // WSEAS Transactions on Mathematics, 2011. Vol. 10. N6. P. 191–200 <http://www.wseas.us/journal/pdf/mathematics/2012/54-882.pdf>
- [4] Кривулин Н. К. Экстремальное свойство собственного значения неразложимых матриц в идемпотентной алгебре и решение задачи размещения Ролса // Вестник СПбГУ., Сер. 1., 2011. Вып. 4 С. 42–51 <https://elibrary.ru/item.asp?id=17048464>
- [5] Krivulin N. A new algebraic solution to multidimensional minimax location problems with Chebyshev distance // WSEAS Transactions on Mathematics. 2012. Vol. 11 N7. P. 605–614 <http://www.wseas.us/journal/pdf/mathematics/2012/54-882.pdf>
- [6] Krivulin N., Zimmermann K. Direct solutions to tropical optimization problems with nonlinear objective functions and boundary constraints // WSEAS Press. 2013. P. 86–91
- [7] Krivulin N. Extremal properties of tropical eigenvalues and solutions to tropical optimization problems // Linear Algebra Appl. 2015. Vol. 468. P. 211–232.
- [8] Кривулин Н., Сорокин В. Решение задачи тропической оптимизации с линейными ограничениями // Вестник СПбГУ., Сер.

- 1., Т. 2 (60). 2015. Вып. 4 С. 541–552 <https://math-mech-astr-journal.spbu.ru/article/view/11191/7873>
- [9] Krivulin N. Using tropical optimization to solve constrained minimax single-facility location problems with rectilinear distance, Computational Management Science // Comput. Manag. Sci. 2017. Vol. 14, N4. P. 493–518 <http://link.springer.com/article/10.1007/s10287-017-0289-2>
- [10] Krivulin N. Direct solution to constrained tropical optimization problems with application to project scheduling // Comput. Manag. Sci. 2017. Vol. 14, N1. P. 91–113.

АЛГЕБРАИЧЕСКОЕ РЕШЕНИЕ ЗАДАЧ СОСТАВЛЕНИЯ ОПТИМАЛЬНОГО ГРАФИКА ВЫПОЛНЕНИЯ ПРОЕКТА

Кривулин Н. К., д.ф.-м.н., профессор кафедры статистического моделирования СПбГУ, nkk@math.spbu.ru

Губанов С. А., СПбФ КБ “Луч” инженер-программист,
segubanov@mail.ru

Аннотация

Предлагаются прямые решения на основе методов тропической оптимизации для задач оптимального планирования графика выполнения работ проекта при различных ограничениях на время выполнения работ. В качестве критериев оптимальности плана используются минимум максимального разброса времени начала всех работ и минимум общей продолжительности проекта.

Введение

Одной из основных проблем управления проектами является задача составления оптимального графика (календарного плана) выполнения работ проекта [1, 2]. Для решения задач календарного планирования находят применение модели и методы тропической математики, которая изучает полукольца и полуполя с идемпотентным сложением [3, 4, 5]. Задачи планирования сводятся к задачам оптимизации, которые формулируются и решаются в терминах тропической математики (задачи тропической оптимизации) [6, 7, 8, 9]. В настоящей работе рассматриваются обобщения задач минимизации максимального разброса времени начала работ проекта и общей продолжительности проекта при заданных временных ограничениях, которые изучались в работах [7, 8, 9], и предлагаются новые решения указанных задач.

Задачи оптимального календарного планирования

Изучаются две задачи, которые возникают при составлении оптимальных графиков выполнения работ проекта при необходимости синхронизировать время выполнения работ во времени. В первой задаче требуется минимизировать максимальный разброс времени начала всех

работ, во второй – минимизировать общую продолжительность проекта. Заданы ограничения в форме минимальных допустимых временных интервалов между временем начала и завершения различных работ, а также границы для времени начала и завершения каждой работы.

Рассмотрим проект, который заключается в параллельном выполнении n работ при условии временных ограничений в виде отношений предшествования “старт-старт”, “старт-финиш” и “финиш-старт”, а также в виде границ для наиболее раннего и наиболее позднего времени начала и наиболее позднего времени завершения работы.

Для каждой работы $i = 1, \dots, n$ обозначим время начала и завершения через x_i и y_i . Пусть заданы величины g_i и h_i , которые определяют самое раннее и самое позднее время начала, а также f_i , которая определяет наиболее позднее время завершения работы. Эти величины задают границы для времени начала и завершения в виде неравенств

$$g_i \leq x_i \leq h_i, \quad y_i \leq f_i.$$

Ограничения типа “старт-старт” для работы i определены в форме неравенств $b_{ij} + x_j \leq x_i$ для всех $j = 1, \dots, n$, где b_{ij} обозначает минимально допустимый интервал времени между началом работы i и началом работы j . Считаем $b_{ij} = -\infty$, если величина b_{ij} не задана. Объединение неравенств по всем j дает эквивалентное неравенство

$$\max_{1 \leq j \leq n} (b_{ij} + x_j) \leq x_i.$$

Обозначим минимальный допустимый интервал между временем начала работы i и завершением работы j как c_{ij} ($c_{ij} = -\infty$, если интервал не задан) и запишем ограничение “старт-финиш” в виде неравенства $c_{ij} + x_j \leq y_i$. После объединения неравенств по всем j получим

$$\max_{1 \leq j \leq n} (c_{ij} + x_j) \leq y_i.$$

Будем предполагать, что работа завершается немедленно при выполнении заданных для нее ограничений “старт-финиш”, а тогда хотя бы для одного j выполняется равенство $c_{ij} + x_j = y_i$. В этом случае предыдущее неравенство следует заменить равенством

$$\max_{1 \leq j \leq n} (c_{ij} + x_j) = y_i.$$

Пусть d_{ij} обозначает минимальный допустимый интервал между временем завершения работы i и начала j ($d_{ij} = -\infty$, если интервал не

задан). Ограничения “финиш-старт” записываются в виде неравенств $d_{ij} + y_j \leq x_i$, объединение которых дает неравенство

$$\max_{1 \leq j \leq n} (d_{ij} + y_j) \leq x_i.$$

Определим критерий оптимальности, который требуется минимизировать, в виде максимального разброса времени начала работ

$$\max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i = \max_{1 \leq i \leq n} x_i + \max_{1 \leq i \leq n} (-x_i).$$

В качестве другого критерия будем рассматривать общую продолжительность проекта, которая определяется по формуле

$$\max_{1 \leq i \leq n} y_i - \min_{1 \leq i \leq n} x_i = \max_{1 \leq i \leq n} y_i + \max_{1 \leq i \leq n} (-x_i).$$

Сформулируем задачу минимизации максимального разброса времени начала при заданных ограничениях, как задачу определения для всех $i = 1, \dots, n$ значений x_i и y_i , которые обеспечивают минимум

$$\begin{aligned} \min_{x_i, y_i} & \quad \max_{1 \leq i \leq n} x_i + \max_{1 \leq i \leq n} (-x_i); \\ & \quad \max_{1 \leq j \leq n} (b_{ij} + x_j) \leq x_i, \quad \max_{1 \leq j \leq n} (c_{ij} + x_j) = y_i, \\ & \quad \max_{1 \leq j \leq n} (d_{ij} + y_j) \leq x_i, \quad g_i \leq x_i \leq h_i, \quad y_i \leq f_i, \quad i = 1, \dots, n. \end{aligned} \tag{1}$$

Задачу составления оптимального плана работ в соответствии с критерием минимума общей продолжительности всех работ проекта представим в виде

$$\begin{aligned} \min_{x_i, y_i} & \quad \max_{1 \leq i \leq n} y_i + \max_{1 \leq i \leq n} (-x_i); \\ & \quad \max_{1 \leq j \leq n} (b_{ij} + x_j) \leq x_i, \quad \max_{1 \leq j \leq n} (c_{ij} + x_j) = y_i, \\ & \quad \max_{1 \leq j \leq n} (d_{ij} + y_j) \leq x_i, \quad g_i \leq x_i \leq h_i, \quad y_i \leq f_i, \quad i = 1, \dots, n. \end{aligned} \tag{2}$$

Ниже эти задачи будут сформулированы в терминах тропической математики и решены с помощью методов тропической оптимизации.

Элементы тропической математики

Приведем обзор основных определений и результатов тропической (идемпотентной) математики [7, 8, 4, 5], необходимых для описания и решения задач тропической оптимизации в следующем разделе.

Пусть множество \mathbb{X} замкнуто относительно ассоциативных и коммутативных операций сложения \oplus и умножения \otimes , и содержит их нейтральные элементы ноль 0 и единицу 1 . Сложение является идемпотентным (для каждого $x \in \mathbb{X}$ выполняется равенство $x \oplus x = x$), а умножение дистрибутивным относительно сложения и обратимым (для любого ненулевого $x \neq 0$ существует элемент x^{-1} такой, что $x \otimes x^{-1} = 1$). Алгебраическую систему $\langle \mathbb{X}, 0, 1, \oplus, \otimes \rangle$ обычно называют идемпотентным полуполем. Знак \otimes операции умножения далее будет опускаться.

Идемпотентное сложение задает частичный порядок: $x \leq y$ тогда и только тогда, когда $x \oplus y = y$. Будем предполагать, что указанный частичный порядок дополнен на \mathbb{X} до линейного порядка.

Для любого $x \neq 0$ и целого $p > 0$ обычным путем определена целая степень: $x^0 = 1$, $x^p = x^{p-1}x$, $x^{-p} = (x^{-1})^p$, $0^p = 0$. Считается, что операция возвведения в рациональную степень также определена.

Примером алгебраической системы $\langle \mathbb{X}, 0, 1, \oplus, \otimes \rangle$ является вещественное полуполе $\mathbb{R}_{\max,+} = \langle \mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, + \rangle$, для которого $0 = -\infty$, $1 = 0$, $\oplus = \max$ и $\otimes = +$.

Множество матриц, состоящих из m строк и n столбцов с элементами из \mathbb{X} , обозначается через $\mathbb{X}^{m \times n}$. Операции сложения и умножения согласованных по размеру матриц $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$ и $\mathbf{C} = (c_{ij})$, а также умножение на скаляр x определяются формулами

$$\{\mathbf{A} \oplus \mathbf{B}\}_{ij} = a_{ij} \oplus b_{ij}, \quad \{\mathbf{BC}\}_{ij} = \bigoplus_k b_{ik} c_{kj}, \quad \{x\mathbf{A}\}_{ij} = x a_{ij}.$$

Заданное выше отношение порядка обобщается на матрицы и понимается покомпонентно.

Рассмотрим квадратные матрицы в $\mathbb{X}^{n \times n}$. Матрица \mathbf{I} с элементами, равными 1 на главной диагонали и 0 вне ее, является единичной.

Для любой квадратной матрицы $\mathbf{A} = (a_{ij})$ и целого $p > 0$ определена степень: $\mathbf{A}^0 = \mathbf{I}$, $\mathbf{A}^p = \mathbf{A}^{p-1}\mathbf{A}$, а также следующие функции:

$$\text{tr } \mathbf{A} = \bigoplus_{i=1}^n a_{ii}, \quad \text{Tr}(\mathbf{A}) = \bigoplus_{k=1}^n \text{tr } \mathbf{A}^k,$$

Если $\text{Tr}(\mathbf{A}) \leq 1$, то определена матрица Клини в форме

$$\mathbf{A}^* = \bigoplus_{k=0}^{n-1} \mathbf{A}^k.$$

Множество векторов-столбцов, состоящих из n элементов обозначается через \mathbb{X}^n . Вектор без нулевых элементов называется регулярным. Вектор, состоящий из единиц, обозначается как $\mathbf{1} = (1, \dots, 1)^T$.

Для вектора $\mathbf{x} = (x_i) \in \mathbb{X}^n$ транспонированный вектор обозначается как \mathbf{x}^T . Мультиликативно сопряженный вектор для \mathbf{x} – это вектор-строка $\mathbf{x}^- = (x_i^-)$, где $x_i^- = x_i^{-1}$, если $x_i \neq 0$, и $x_i^- = 0$ – иначе.

Введем тропические аналоги нормы вектора и матрицы. Для любых вектора $\mathbf{x} = (x_i) \in \mathbb{X}^n$ и матрицы $\mathbf{A} = (a_{ij}) \in \mathbb{X}^{m \times n}$ имеем

$$\|\mathbf{x}\| = \bigoplus_{i=1}^n x_i, \quad \|\mathbf{A}\| = \bigoplus_{i=1}^m \bigoplus_{j=1}^n a_{ij}.$$

Решение задач оптимального планирования

Рассмотрим задачи минимизации максимального разброса времени начала работ (1) и минимизации общей продолжительности проекта (2), в которых заданы ограничения вида “старт-старт”, “старт-финиш”, “финиш-старт” и границы для самого раннего и самого позднего допустимого времени начала работ. Представим эти задачи в векторной форме как задачи тропической оптимизации, а затем приведем результаты, которые описывают полное решение задач.

Сначала сформулируем задачу (1) в терминах идемпотентного полуполя $\mathbb{R}_{\max,+}$. Для этого введем следующие матрицы и векторы:

$$\begin{aligned} \mathbf{B} &= (b_{ij}), & \mathbf{C} &= (c_{ij}), & \mathbf{D} &= (d_{ij}), \\ \mathbf{x} &= (x_i), & \mathbf{f} &= (f_i), & \mathbf{g} &= (g_i), & \mathbf{h} &= (h_i). \end{aligned}$$

После замены арифметических операций на операции полуполя $\mathbb{R}_{\max,+}$ задачу можно представить в векторном виде в форме

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^- \mathbf{1} \mathbf{1}^T \mathbf{x}; \\ \text{при } \quad & \mathbf{Bx} \leq \mathbf{x}, \quad \mathbf{Cx} = \mathbf{y}, \\ & \mathbf{Dy} \leq \mathbf{x}, \quad \mathbf{g} \leq \mathbf{x} \leq \mathbf{h}, \quad \mathbf{y} \leq \mathbf{f}. \end{aligned} \tag{3}$$

На основе применения методов тропической оптимизации получен следующий результат.

Лемма 1 Пусть \mathbf{B} и \mathbf{D} – матрицы, а \mathbf{C} – регулярная по столбцам матрица, такие, что для матрицы $\mathbf{R} = \mathbf{B} \oplus \mathbf{DC}$ выполняется условие $\text{Tr}(\mathbf{R}) \leq 1$. Пусть \mathbf{g} – вектор, а \mathbf{f} и \mathbf{h} – регулярные векторы такие, что для вектора $\mathbf{s}^T = \mathbf{f}^- \mathbf{C} \oplus \mathbf{h}^-$ выполняется условие $\mathbf{s}^T \mathbf{R}^* \mathbf{g} \leq 1$.

Тогда минимальное значение целевой функции в задаче (3) равно

$$\theta = \|\mathbf{R}^*\| \oplus \bigoplus_{0 \leq i+j \leq n-2} \|s^T \mathbf{R}^i\| \|\mathbf{R}^j \mathbf{g}\|,$$

а все регулярные решения имеют вид

$$\mathbf{x} = \mathbf{G}\mathbf{u}, \quad \mathbf{y} = \mathbf{C}\mathbf{G}\mathbf{u}, \quad \mathbf{G} = \mathbf{R}^* \oplus \bigoplus_{0 \leq i+j \leq n-2} \theta^{-1} \mathbf{R}^i \mathbf{1} \mathbf{1}^T \mathbf{R}^j,$$

где \mathbf{u} – любой регулярный вектор, который удовлетворяет условиям

$$\mathbf{g} \leq \mathbf{u} \leq (s^T \mathbf{G})^-.$$

Задача минимизации общей продолжительности проекта (2) в векторной форме имеет вид

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{1} \mathbf{1}^T \mathbf{y}; \\ \text{при } \quad & \mathbf{Bx} \leq \mathbf{x}, \quad \mathbf{Cx} = \mathbf{y}, \\ & \mathbf{Dy} \leq \mathbf{x}, \quad \mathbf{g} \leq \mathbf{x} \leq \mathbf{h}, \quad \mathbf{y} \leq \mathbf{f}. \end{aligned} \tag{4}$$

Решение задачи описывает следующее утверждение.

Лемма 2 Пусть \mathbf{B} и \mathbf{D} – матрицы, \mathbf{C} – регулярная по столбцам матрица, такие, что матрица $\mathbf{R} = \mathbf{B} \oplus \mathbf{DC}$ удовлетворяет условию $\text{Tr}(\mathbf{R}) \leq 1$. Пусть \mathbf{g} – вектор, а \mathbf{f} и \mathbf{h} – регулярные векторы такие, что вектор $s^T = \mathbf{f}^T \mathbf{C} \oplus \mathbf{h}^T$ удовлетворяет условию $s^T \mathbf{R}^* \mathbf{g} \leq 1$.

Тогда минимальное значение целевой функции в задаче (4) равно

$$\theta = \|\mathbf{CR}^*\| \oplus \bigoplus_{0 \leq i+j \leq n-2} \|s^T \mathbf{R}^i\| \|\mathbf{CR}^j \mathbf{g}\|,$$

а все регулярные решения имеют вид

$$\mathbf{x} = \mathbf{Gu}, \quad \mathbf{y} = \mathbf{CGu}, \quad \mathbf{G} = \mathbf{R}^* \oplus \bigoplus_{0 \leq i+j \leq n-2} \theta^{-1} \mathbf{R}^i \mathbf{1} \mathbf{1}^T \mathbf{CR}^j,$$

где \mathbf{u} – любой регулярный вектор, который удовлетворяет условиям

$$\mathbf{g} \leq \mathbf{u} \leq (s^T \mathbf{G})^-.$$

Заключение

Рассмотрены задачи составления оптимального графика выполнения работ, которые заключаются в минимизации максимального разброса времени начала работ и минимизации общей продолжительности проекта при заданных ограничениях вида “старт-старт”, “старт-финиш”, “финиш-старт” и границах для самого раннего и самого позднего допустимого времени начала работ. Получены прямые решения задач, которые могут быть использованы для их формального анализа и непосредственных вычислений в практических приложениях.

Список литературы

- [1] T'kindt V. and Billaut J.-C. , Multicriteria Scheduling. 2 ed. Berlin: Springer, 2006.
- [2] Kerzner H. Project Management. 10 ed. Hoboken: Wiley, 2010. 1094 p.
- [3] Маслов В. П., Колокольцев В. Н. Идемпотентный анализ и его применение в оптимальном управлении. М.: Физматлит, 1994.
- [4] Кривулин Н. К. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб.: Изд-во С.-Петерб. ун-та., 2009.
- [5] Butkovič P., Max-linear Systems: Theory and Algorithms. Springer Monographs in Mathematics. London: Springer, 2010.
- [6] Krivulin N. Direct solution to constrained tropical optimization problems with application to project scheduling // Computational Management Science. 2017. Vol. 14. N 1. P. 91-113.
- [7] Krivulin N., Tropical optimization problems with application to project scheduling with minimum makespan // Annals of Operations Research. 2017. Vol. 256, N 1. P. 75-92.
- [8] Krivulin N., Tropical optimization problems in time-constrained project scheduling. // Optimization. 2017. Vol. 66, N 2. P. 205-224.
- [9] Кривулин Н. К., Губанов С. А. Алгебраическое решение задачи оптимального планирования сроков проекта в управлении проектами // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2021. Т. 8. Вып. 1. С. 73-87.

РЕШЕНИЕ ЗАДАЧИ ВЫБОРА ПРОЕКТА ИНФОРМАЦИОННОЙ СИСТЕМЫ МЕТОДАМИ ТРОПИЧЕСКОЙ МАТЕМАТИКИ

Приньков А.С., аспирант кафедры статистического моделирования
СПбГУ, aprinkov@yahoo.com

Руководитель: Кривулин Н.К., д.ф.-м.н. профессор кафедры
статистического моделирования СПбГУ, nkk@math.spbu.ru

Аннотация

Рассматривается многокритериальная задача оценки рейтингов альтернатив на основе парных сравнений при принятии решения по выбору проекта информационной системы. Для решения задачи используется подход, основанный на взвешенной минимаксной log-чебышевской аппроксимации и применении методов тропической математики. Полученное решение сравнивается с известным решением с помощью метода анализа иерархий.

Введение

Многокритериальные задачи оценки альтернатив на основе парных сравнений составляют важный класс задач принятия решений, которые распространены во многих сферах. В многокритериальных задачах выбор оптимального решения осуществляется в соответствии с несколькими критериями. Основная сложность таких задач — отсутствие в общем случае решения, которое является наилучшими по всем критериям сразу. Исходными условиями такой задачи выступает набор m альтернатив и их попарное сравнение по n критериям. Такие сравнения удобно представлять в виде матриц парных сравнений A_k , где $1 \leq k \leq n$. Критерии также сравниваются между собой попарно, а результаты сравнений записываются в матрицу сравнений критериев C . Решением такой задачи является вектор абсолютных рейтингов, по которому определяются ранги альтернатив. Один из подходов к решению опирается на log-чебышевскую аппроксимацию матриц парных сравнений согласованными матрицами (обратно симметрическими матрицами единичного ранга) и сводится к решению задач оптимизации в терминах тропической математики. С помощью методов тропической математики эти задачи могут быть решены аналитически.

Элементы тропической математики

Тропическая (идемпотентная) математика изучает теорию и приложения алгебраических систем с идемпотентными операциями [1, 2, 3]. Операцию называют идемпотентной, если при ее применении к одинаковым аргументам она дает в результате этот аргумент. Например, операция максимум является идемпотентной: $\max(x, x) = x$. Задачи оптимизации, сформулированные в терминах идемпотентных алгебраических систем, могут быть решены методами тропической оптимизации. Это относится к классическим и вновь сформулированным задачам. Многие из этих задач могут быть решены в явной аналитической форме, другие же имеют только алгоритмическое решение.

В работе используется тах-алгебра — алгебраическая система, которая представляет собой множество неотрицательных вещественных чисел $\mathbb{R}_+ = \{x \in \mathbb{R} | x \leq 0\}$ с операциями сложения и умножения. В такой алгебраической системе сложение определено как максимум и обозначается знаком \oplus . Умножение определено и обозначается как обычно. Векторные и матричные операции выполняются по стандартным правилам с заменой арифметического сложения на операцию \oplus . Единичная матрица обозначается символом \mathbf{I} и имеет обычный вид. Целая неотрицательная степень квадратной матрицы \mathbf{A} обозначает результат произведения матрицы на себя и определена для всех натуральных p так, что $\mathbf{A}^0 = \mathbf{I}, \mathbf{A}^p = \mathbf{A}^{p-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{p-1}$. След матрицы $\mathbf{A} = (a_{ij})$ порядка n вычисляется по формуле $\text{tr } \mathbf{A} = a_{11} \oplus \cdots \oplus a_{nn}$.

Линейной комбинацией векторов $\mathbf{a}_1, \dots, \mathbf{a}_n$ с коэффициентами $x_1, \dots, x_n \in \mathbb{R}_+$ называется выражение $x_1\mathbf{a}_1 \oplus \cdots \oplus x_n\mathbf{a}_n$. Вектор \mathbf{b} линейно зависит от векторов $\mathbf{a}_1, \dots, \mathbf{a}_n$, если существуют числа $x_1, \dots, x_n \in \mathbb{R}_+$ такие, что выполняется равенство $\mathbf{b} = x_1\mathbf{a}_1 \oplus \cdots \oplus x_n\mathbf{a}_n$. Коллинеарность двух векторов имеет обычный смысл: векторы \mathbf{a} и \mathbf{b} являются коллинеарными, если $\mathbf{b} = x\mathbf{a}$ для некоторого $x \in \mathbb{R}_+$.

Спектральным радиусом матрицы \mathbf{A} называется число, которое вычисляется согласно формуле

$$\lambda = \text{tr } \mathbf{A} \oplus \cdots \oplus \text{tr}^{1/n}(\mathbf{A}^n) = \bigoplus_{i=1}^n \text{tr}^{1/i}(\mathbf{A}^i).$$

Если $\lambda < 1$, то для матрицы \mathbf{A} определен оператор Клини

$$\mathbf{A}^* = \mathbf{I} \oplus \mathbf{A} \oplus \cdots \oplus \mathbf{A}^{n-1} = \bigoplus_{i=0}^{n-1} \mathbf{A}^i.$$

Более подробные сведения по теории, методам и приложениям тропической математики можно найти, например, в работах [1, 2, 3].

Решение задачи выбора проекта информационной системы

Рассмотрим задачу выбора проекта информационной системы, описанную в работе [4]. Эта задача состоит в том, чтобы по совокупности критериев выбрать наиболее приоритетный проект информационной системы для реализации. В работе [4] приводится решение с помощью метода анализа иерархий Т. Саати [5], а также обсуждается преимущество такого решения над классическими экспертными методами маркетинга при сравнении альтернатив по нескольким критериям. При использовании нескольких критериев выбора проекта информационной системы возникает ряд сложностей, основные из которых — присутствие нематериальных выгод и невозможность установить относительную важность разных типов затрат и прибыли. Таким образом, без учета относительной важности критериев можно ошибочно выделить проекты с высокой материальной прибылью, игнорируя проекты с высокой нематериальной прибылью.

В рассматриваемой задаче используется шкала оценок от 1 до 9. Всего рассматривается $m = 6$ конкурирующих альтернативных проектов информационных систем. Альтернативы сравниваются по $n = 4$ критериям: увеличение точности канцелярских операций (accuracy), эффективность обработки информации (efficiency), продвижение обучения в организации (organisational learning) и стоимость реализации (implementation costs).

Далее приведем решение многокритериальной задачи по выбору информационного проекта на основе минимаксной log-чебышевской аппроксимации. Такое решение в терминах тропической оптимизации может быть получено в аналитической форме [6, 7].

В задаче задана матрица парных сравнений критериев:

$$C = \begin{pmatrix} 1 & 1/9 & 1/7 & 1/5 \\ 9 & 1 & 2 & 5 \\ 7 & 1/2 & 1 & 3 \\ 5 & 1/5 & 1/3 & 1 \end{pmatrix}.$$

Матрицы парных сравнений альтернатив по каждому критерию:

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 1/3 & 1/6 & 1/6 & 1/3 & 1/9 \\ 3 & 1 & 1/3 & 1/3 & 1 & 1/8 \\ 6 & 3 & 1 & 1 & 3 & 1/8 \\ 6 & 3 & 1 & 1 & 3 & 1/8 \\ 3 & 1 & 1/3 & 1/3 & 1 & 1/8 \\ 9 & 8 & 8 & 8 & 8 & 1 \end{pmatrix}; \quad \mathbf{A}_2 = \begin{pmatrix} 1 & 4 & 3 & 1 & 3 & 4 \\ 1/4 & 1 & 7 & 3 & 1/5 & 1 \\ 1/3 & 1/7 & 1 & 1/5 & 1/5 & 1/6 \\ 1 & 1/3 & 5 & 1 & 1 & 1/3 \\ 1/3 & 5 & 5 & 1 & 1 & 3 \\ 1/4 & 1 & 5 & 3 & 1/3 & 1 \end{pmatrix};$$

$$\mathbf{A}_3 = \begin{pmatrix} 1 & 1/5 & 2 & 1/3 & 1/2 & 2 \\ 5 & 1 & 7 & 2 & 3 & 7 \\ 1/2 & 1/7 & 1 & 1/5 & 1/2 & 1 \\ 3 & 1/2 & 5 & 1 & 2 & 5 \\ 2 & 1/3 & 2 & 1/2 & 1 & 3 \\ 1/2 & 1/7 & 1 & 1/5 & 1/3 & 1 \end{pmatrix}; \quad \mathbf{A}_4 = \begin{pmatrix} 1 & 5 & 4 & 2 & 3 & 1/3 \\ 1/5 & 1 & 1/2 & 1/4 & 1/3 & 1/8 \\ 1/4 & 2 & 1 & 1/3 & 1/2 & 1/6 \\ 1/2 & 4 & 3 & 1 & 2 & 1/4 \\ 1/3 & 3 & 2 & 1/2 & 1 & 1/5 \\ 3 & 8 & 6 & 4 & 5 & 1 \end{pmatrix}.$$

Для решения задачи с помощью минимаксной log-чебышевской аппроксимации будем производить аналитические вычисления в терминах тропического полуполя с операцией максимума в роли сложения, которое называют тах-алгеброй.

Для определения весов критерииев необходимо вычислить $(\lambda^{-1}\mathbf{C})^*$.

Сначала для матрицы \mathbf{C} определим спектральный радиус:

$$\lambda = \text{tr } \mathbf{C} \oplus \dots \oplus \text{tr}^{1/4}(\mathbf{C}^4) = (25/9)^{1/3} \approx 1.4057.$$

Вычисление матрицы Клини, столбцы которой генерируют все оптимальные векторы весов критерииев, дает следующий результат:

$$\mathbf{D} = (\lambda^{-1}\mathbf{C})^* = \begin{pmatrix} 1 & 1/9\lambda & 2\lambda/25 & \lambda/5 \\ 9\lambda & 1 & 2/\lambda & 5/\lambda \\ 27\lambda/5 & 3/5 & 1 & 3/\lambda \\ 5/\lambda & \lambda/5 & 2/5 & 1 \end{pmatrix}.$$

Такая матрица может генерировать не единственный (с точностью до положительного множителя) вектор весов, а множество векторов \mathcal{S} . В случае неединственного решения в качестве «наилучшего» и «наихудшего» векторов весов возьмем такие векторы из множества \mathcal{S} , которые максимизируют и минимизируют отношение между максимальным и минимальным элементами вектора весов. Эти решения являются наилучшим и наихудшим дифференцирующими векторами весов.

Нормируем столбцы матрицы \mathbf{D} относительно максимального элемента, тогда максимальные элементы всегда будут равны 1. В этом случае наилучшим решениям будут соответствовать столбцы, у которых минимальные элементы являются наименьшими среди всех столбцов, а наихудшие — у которых минимальные элементы наибольшие.

Нормированная матрица \mathbf{D} равна

$$\begin{pmatrix} 1/9\lambda & 1/9\lambda & 1/9\lambda & 1/9\lambda \\ 1 & 1 & 1 & 1 \\ 3/5 & 3/5 & \lambda/2 & 3/5 \\ \lambda/5 & \lambda/5 & \lambda/5 & \lambda/5 \end{pmatrix}.$$

Первый, второй и четвертый столбцы коллинеарны. Для всех столбцов минимальным элементом является $1/9\lambda \approx 0.0790$.

Наилучшими дифференцирующим вектором весов является вектор

$$\mathbf{v} = (1/9\lambda, 1, 3/5, \lambda/5)^T,$$

а наихудшим - вектор

$$\mathbf{w} = (1/9\lambda, 1, \lambda/2, \lambda/5)^T.$$

Возьмем наилучший дифференцирующий вектор весов \mathbf{v} и составим взвешенную сумму

$$\begin{aligned} \mathbf{P} &= 1/9\lambda \mathbf{A}_1 + \mathbf{A}_2 + 3/5 \mathbf{A}_3 + \lambda/5 \mathbf{A}_4 = \\ &= \begin{pmatrix} 1 & 4 & 3 & 1 & 3 & 4 \\ 3 & 1 & 7 & 3 & 9/5 & 21/5 \\ 2/3\lambda & 2\lambda/5 & 1 & 1/5 & 3/10 & 3/5 \\ 9/5 & 4\lambda/5 & 5 & 1 & 6/5 & 3 \\ 6/5 & 5 & 5 & 1 & 1 & 3 \\ 3\lambda/5 & 8\lambda/5 & 6 & 3 & \lambda & 1 \end{pmatrix}. \end{aligned}$$

После вычисления следов степеней матрицы \mathbf{P} определим ее спектральный радиус:

$$\mu = \text{tr } \mathbf{P} \oplus \dots \oplus \text{tr}^{1/6}(\mathbf{P}^6) = 45^{1/3} \approx 3.5569.$$

Матрица Клини ($\mu^{-1} \mathbf{P}$)* равна:

$$\begin{pmatrix} 1 & 675/\mu^5 & 378/\mu^4 & 189/\mu^4 & 3/\mu & 7/5 \\ 3/\mu & 1 & 1134/\mu^5 & 567/\mu^5 & 405/\mu^5 & 21/5\mu \\ 2/15 & 2\lambda/5\mu & 1 & 126/\lambda\mu^5 & 90/\lambda\mu^5 & 42/\lambda\mu^4 \\ 81/\mu^4 & 3/5 & 18/\mu^2 & 1 & 243/\mu^5 & 3/\mu \\ 675/\mu^5 & 225/\mu^4 & 14/5 & 7/5 & 1 & 945/\mu^5 \\ 8/15 & 8\lambda/5\mu & 6/\mu & 3/\mu & 72\lambda/5\mu^3 & 1 \end{pmatrix}.$$

С помощью генерирующей матрицы $(\mu^{-1} \mathbf{P})^*$ вычислим наилучший дифференцирующий вектор рейтингов альтернатив:

$$\begin{aligned}\mathbf{x} &= (3/\mu, 405/\mu^5, 90/\lambda\mu^5, 243/\mu^5, 1, 72\lambda/5\mu^3)^T \approx \\ &\approx (0.8434, 0.7114, 0.1125, 0.4268, 1.0000, 0.4498)^T.\end{aligned}$$

Вычисленный вектор задает порядок $A_5 > A_1 > A_2 > A_6 > A_4 > A_3$.

Возьмем наихудший дифференцирующий вектор весов \mathbf{w} и составим взвешенную сумму

$$\begin{aligned}\mathbf{R} &= 1/9\lambda \mathbf{A}_1 + \mathbf{A}_2 + \lambda/2 \mathbf{A}_3 + \lambda/5 \mathbf{A}_4 = \\ &= \begin{pmatrix} 1 & 4 & 3 & 1 & 3 & 4 \\ 5\lambda/2 & 1 & 7 & 3 & 3\lambda/2 & 7\lambda/2 \\ 2/3\lambda & 2\lambda/5 & 1 & 1/5 & \lambda/4 & \lambda/2 \\ 3\lambda/2 & 4\lambda/5 & 5 & 1 & \lambda & 5\lambda/2 \\ \lambda & 5 & 5 & 1 & 1 & 3 \\ 3\lambda/5 & 8\lambda/5 & 6 & 3 & \lambda & 1 \end{pmatrix}.\end{aligned}$$

После вычисления следов степеней матрицы \mathbf{R} определим ее спектральный радиус:

$$\mu = \text{tr } \mathbf{R} \oplus \dots \oplus \text{tr}^{1/6}(\mathbf{R}^6) = (75\lambda/2)^{1/3} \approx 3.7495$$

Матрица Клини $(\mu^{-1} \mathbf{R})^*$ равна

$$\begin{pmatrix} 1 & 1125\lambda/2\mu^5 & 315\lambda/\mu^4 & 315\lambda/2\mu^4 & 225\lambda/2\mu^4 & 7/5 \\ 375\lambda^2/4\mu^4 & 1 & 1575\lambda^2/2\mu^5 & 1575\lambda^2/4\mu^5 & 1125\lambda^2/4\mu^5 & 525\lambda^2/4\mu^4 \\ 75\lambda^3/2\mu^5 & 15\lambda^2/\mu^4 & 1 & 21\lambda^2/5\mu^3 & 3\lambda^2/\mu^3 & 105\lambda^3/2\mu^5 \\ 225\lambda^2/4\mu^4 & 3/5 & 15\lambda/\mu^2 & 1 & 675\lambda^2/4\mu^5 & 5\lambda/2\mu \\ 1875\lambda^2/4\mu^5 & 375\lambda/2\mu^4 & 14/5 & 7/5 & 1 & 2625\lambda^2/4\mu^5 \\ 150\lambda^3/\mu^5 & 60\lambda^2/\mu^4 & 6/\mu & 3/\mu & 12\lambda^2/\mu^3 & 1 \end{pmatrix}.$$

Вычислим наихудший дифференцирующий вектор рейтингов альтернатив для генерирующей матрицы оптимальных альтернатив:

$$\begin{aligned}\mathbf{y} &= (4\mu^5/1875\lambda^2, 2\mu^4/375\lambda, 5/14, 5/7, 1, 4\mu^5/2625\lambda^2)^T \approx \\ &\approx (0.8001, 0.7499, 0.3571, 0.7143, 1, 0.5715)^T.\end{aligned}$$

Вычисленный вектор задает порядок $A_5 > A_1 > A_2 > A_4 > A_6 > A_3$.

В работе [4], где была сформулирована эта задача, порядок альтернатив при использовании метода анализа иерархий получился $A_1 >$

$A_2 > A_5 > A_6 > A_4 > A_3$. Если сравнивать с наилучшим лог-чебышевским решением, то видно, что разница в решениях есть в первых трех рангах, последние три ранга полностью совпадают. В случае сравнения с результатами наихудшей лог-чебышевской аппроксимации тройка наиболее приоритетных и тройка наименее приоритетных альтернатив совпадают, но внутри троек разный порядок. В случае существенного различия результатов, выбор одного из них не вполне очевиден. В рассматриваемой задаче все методы дали по существу схожие результаты, поэтому различные решения могут дополнять информацию при окончательной приоритезации.

Список литературы

- [1] Маслов В. П., Колокольцев В. Н. Идемпотентный анализ и его применение в оптимальном управлении. М.: Физматлит, 1994
- [2] Heidergott B., Olsder G. J., van der Woude J. Max Plus at Work. Princeton Series in Applied Mathematics. Princeton: Princeton Univ. Press, 2006. 226 p.
- [3] Кривулин Н. К. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб.: Изд-во С.-Петерб. ун-та, 2009. 255 с.
- [4] Muralidhar K., Santhanam R., Wilson R. L. Using the analytic hierarchy process for information system project selection // Information and Management. 1990. Vol. 18, N 2. P. 87–95.
- [5] Саати Т. Принятие решений. Метод анализа иерархий. — М.:Радио, 1993 — 278 с.
- [6] Кривулин Н. К., Агеев В. А. Методы тропической оптимизации в многокритериальных задачах оценки альтернатив на основе парных сравнений // Вестн. С.-Петерб. ун-та. Прикладная математика. 2019. Т. 15. Вып. 4. С. 472–488.
- [7] Krivulin N. Methods of tropical optimization in rating alternatives based on pairwise comparisons // Operations Research Proceedings 2016 / Ed. by A. Fink, A. Fugenschuh, M. J. Geiger. Cham: Springer, 2018. P. 85–91.

РАЗРАБОТКА ПРОГРАММНЫХ СРЕДСТВ ВИЗУАЛИЗАЦИИ ТРОПИЧЕСКИХ РАЦИОНАЛЬНЫХ ФУНКЦИЙ

Шкурат Д. Е., студент магистратуры СПбГУ,
st097097@student.spbu.ru

Аннотация

Тропическая математика — раздел математики, активно использующийся для решения практических задач в разных областях науки: в физике, биологии, экономике и оптимизации. Работа посвящена созданию программных средств визуализации тропических многочленов и тропических рациональных функций.

Введение

Тропическая математика — относительно новый раздел математики, который появился в 1960-х годах и быстро развивается в последние десятилетия. Изначально эта область разрабатывалась и применялась в контексте дискретной математики и оптимизации, но в последнее время тропическая математика находит также применение в информатике [1], экономике [2], биологии [3] и в других науках.

Быстрое развитие тропической математики объясняется тем, что многие нелинейные в обычном смысле задачи, представленные в терминах тропической математики, становятся линейными и сводятся к решению линейных уравнений, нахождению собственных векторов и другим алгебраическим задачам.

Рассмотрим тропическую алгебраическую систему (идемпотентное полуполе) $\mathbb{R}_{(\max,+)}$, которую называют $(\max,+)$ -алгеброй. Эта система задана на множестве $\mathbb{X} = \mathbb{R} \cup \{-\infty\}$, где \mathbb{R} обозначает множество вещественных, на котором определены операции сложения: $x \oplus y = \max(x, y)$ и умножения: $x \otimes y = x + y$. Заметим, что сложение *идемпотентно*, т.е. для любого x из \mathbb{X} верно, что $x \oplus x = x$.

Определим операцию возведения в степень x^y как $x \cdot y$. Пусть a_i коэффициенты из \mathbb{X} , тогда можно определить тропический многочлен следующим образом:

$$P(x) = \bigoplus_i a_i x^i.$$

Тропические полиномы часто используются в решении задач обработки изображений [4], криптографии [5] и других.

Тропическая рациональная функция — это тропическое частное двух тропических полиномов:

$$R(x) = \frac{P(x)}{Q(x)}, \quad P(x) = \bigoplus_i a_i x^i, \quad Q(x) = \bigoplus_j b_j x^j.$$

Рациональные функции могут использоваться, например, при обучении нейронных сетей [6].

Известно, что рациональные функции не являются выпуклыми, поэтому задача нахождения экстремумов не является тривиальной. Для решения таких задач могли бы оказаться полезными средства визуализации тропических полиномов и рациональных функций, поэтому разработка приложения, которое обрабатывает многочлены, записанными на языке тропической математики, а также отображает их на графиках, представляется достаточно актуальным.

Тропическая математика

Пусть \mathcal{X} — числовое множество, на котором заданы операции сложения \oplus и умножения \otimes . Будем предполагать, что $\langle \mathcal{X}, \oplus, \otimes \rangle$ является коммутативным полукольцом с нулём и единицей, в котором сложение идемпотентно, а каждый ненулевой элемент имеет обратный по умножению. Обозначим нулевой и единичный элементы полукольца символами 0 и 1 .

Свойства операций

Множество \mathcal{X} образует относительно операции сложения идемпотентную коммутативную полугруппу с нейтральным элементом, т.е. удовлетворяет условиям:

1. $x \oplus (y \oplus z) = (x \oplus y) \oplus z$ (ассоциативность),
2. $x \oplus y = y \oplus x$ (коммутативность),
3. $x \oplus 0 = x$ (существование нуля),
4. $x \oplus x = x$ (идемпотентность).

Относительно умножения имеем коммутативную группу, в которой выполняется:

5. $x \otimes (y \otimes z) = (x \otimes y) \otimes z$ (ассоциативность),
6. $x \otimes y = y \otimes x$ (коммутативность),
7. $x \otimes 1 = x$ (существование единицы),
8. если $x \neq 0$, то $\exists x^{-1} : x \otimes x^{-1} = 1$ (существование обратного).

Сложение и умножение связаны свойствами:

9. $x \otimes (y \oplus z) = x \otimes y \oplus x \otimes z$ (дистрибутивность),
10. $x \otimes 0 = 0$ (закон поглощения).

С учётом групповых свойств умножения полукольцо $\langle \mathcal{X}, \oplus, \otimes \rangle$ часто называют полуядром.

Примеры идемпотентных полуколец

Следующие полукольца являются идемпотентными:

- $\mathbb{R}_{\max,+} = \langle \mathbb{R} \cup \{-\infty\}, \max, + \rangle$,
- $\mathbb{R}_{\max,\times} = \langle \mathbb{R}_+ \cup \{0\}, \max, \times \rangle$,
- $\mathbb{R}_{\min,+} = \langle \mathbb{R} \cup \{+\infty\}, \min, + \rangle$,
- $\mathbb{R}_{\min,\times} = \langle \mathbb{R}_+ \cup \{+\infty\}, \min, \times \rangle$.

где \mathbb{R} — множество всех вещественных чисел, а $\mathbb{R}_+ = \{x \in \mathbb{R} | x > 0\}$.

Рассмотрим подробнее первый пример. В $\mathbb{R}_{\max,+}$ нулевым элементом является $-\infty$, а единичным — число 0.

Для всех x из \mathbb{R} существует обратный по умножению элемент x^{-1} , равный $-x$ в обычной арифметике.

Для всех x, y из \mathbb{R} определена степень x^y , значение которой соответствует арифметическому произведению xy .

Полиномы и рациональные функции

Рассмотрим тропический полином, заданный в терминах $(\max, +)$ -алгебры в форме:

$$P(x) = \bigoplus_i a_i x^i.$$

При использовании обычных обозначений это выражение можно записать в виде:

$$\max_i(i \cdot x + a_i).$$

Показатель степени может быть неотрицательным целым числом, но также можно рассматривать и целые показатели (получится тропический многочлен Лорана), и рациональные показатели (тропический многочлен Пюизё), и действительные (обобщённый тропический многочлен Пюизё).

Тропическая рациональная функция — это тропическое частное двух полиномов:

$$R(x) = \frac{P(x)}{Q(x)}, \quad P(x) = \bigoplus_i a_i x^i, \quad Q(x) = \bigoplus_j b_j x^j.$$

что в обычной алгебре будет выглядеть так:

$$\max_i(i \cdot x + a_i) - \max_j(j \cdot x + b_j).$$

Тропический полином образует выпуклую кусочно-линейную функцию, а тропическая линейная функция образует невыпуклую кусочно-линейную функцию. Далее будет приведён пример, иллюстрирующий форму графика рациональной функции

Разработка программного обеспечения для визуализации

Цели и задачи приложения

- Интерпретировать полиномы, записанные на языке тропической математики;
- вычислять значения рациональной функции;
- Строить графики тропических полиномов;
- Строить графики обыкновенных функций;
- Строить график функции ошибки;
- Строить график тропической рациональной функции;

Выбор языка и средств разработки

Для реализации данного проекта был выбран язык C# 8.0 на платформе .NET 5.0. Целевая платформа — Windows 10, конфигурация — AnyCPU. Приложение должно работать на всех операционных системах Windows, которые поддерживают версию .NET 5.0 и выше. Выбор рабочей среды .NET (а не .Net Framework) должен обеспечить возможность сборки проекта под другими операционными системами, например, Linux.

В качестве интегрированной среды разработки была выбрана Microsoft Visual Studio 2022, поскольку она является наиболее удобной средой разработки на языке C#, имеет встроенный диспетчер пакетов NuGet, а также позволяет видеть изменения с системы контроля версий GitHub.

Принцип работы приложения

Для изображения графиков используется пакет ScottPlot, позволяющий отображать несколько графиков на одной плоскости. Чтобы нарисовать график функции достаточно передать два массива одинакового размера — абсцисс и ординат.

В качестве синтаксического анализатора используется система компьютерной алгебры с открытым исходным кодом — AngouriMath. Она работает с объектами-выражениями, постепенно разбивая их на более мелкие объекты-выражения. Она не имеет встроенной поддержки тропической алгебры, поэтому требуется реализовать средства, позволяющие обрабатывать выражения, записанные на языке тропической алгебры.

Продемонстрируем, как реализуется $(\max, +)$ -алгебра в нашем приложении:

```
double MaxPlus(Entity expr)
=> expr switch {
    Number.Real r => (double)r,
    Sumf(var a, var b) => (MaxPlus(a) > MaxPlus(b))
        ? MaxPlus(a) : MaxPlus(b),
    Powf(var a, var b) => MaxPlus(a) * (double)b.
        EvalNumerical().RealPart,
    Mulf(var a, var b) => MaxPlus(a) + MaxPlus(b),
    Divf(var a, var b) => MaxPlus(a) - MaxPlus(b),
};
```

В данном случае алгебра реализуется с помощью функции, которая возвращает double, принимает на вход объект-выражение Entity, кото-

рый в последствии рекурсивно разбирает. Обратим ваше внимание, что операнд b не в Powf не подвергается рекурсивному разбору, поскольку показатель степени должен вычисляться в обычновенной алгебре.

Все источники выложены на GitHub и могут быть найдены по ссылке: <https://github.com/dragondangun/Tropical-Palm>.

Пример работы с приложением

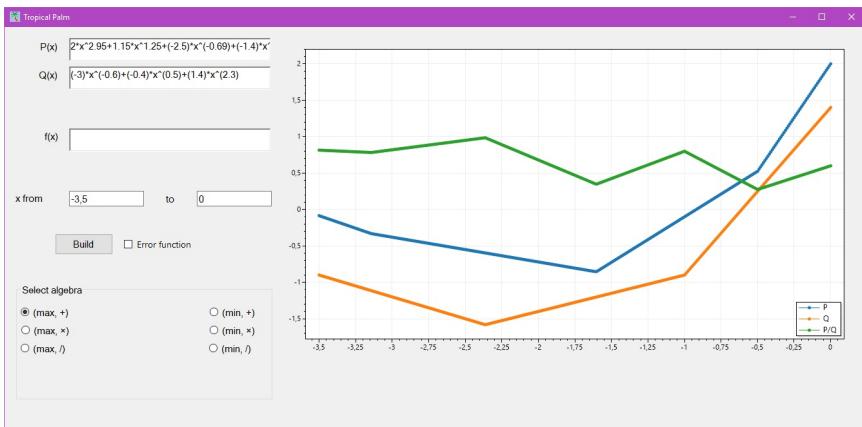


Рис. 1: Приложение во время работы

На рис. 1 показано, что сначала требуется задать отрезок, в котором будут строиться графики функций: $[-3.5; 0]$. Затем, в поле многочлена $P(x)$ вводится полином:

$$2 * x^{2.95} + 1.15 * x^{1.25} + (-2.5) * x^{-0.69} + (-1.4) * x^{-0.34} + (-1),$$

а в $Q(x)$ полином:

$$(-3) * x^{-0.6} + (-0.4) * x^{0.5} + (1.4) * x^{2.3}.$$

После нажатия на кнопку Build строятся и отображаются графики тропических полиномов и тропической рациональной функции. На графике видно, что рациональная функция не является выпуклой. Дополнительно, может быть задана произвольная функция, которая записывается в терминах обычной алгебры.

При желании можно сменить алгебру и/или включить построение графика функции ошибки.

Заключение

Разработано приложение, которое позволяет строить графики тропических полиномов и дробно-рациональных функций. Дальнейшая работа над приложением предусматривает расширение его функциональных возможностей и усовершенствование пользовательского интерфейса.

Список литературы

- [1] Litvinov, G.L., Idempotent and tropical mathematics; complexity of algorithms and interval analysis // Computers & Mathematics with Applications, Volume 65, Issue 10, 2013
- [2] Baldwin, E. and Klempner, P. Understanding Preferences: “Demand Types”, and the Existence of Equilibrium With Indivisibilities. // Econometrica 2019, 87: 867-932. <https://doi.org/10.3982/ECTA13693>
- [3] Radulescu, O. Tropical Geometry of Biological Systems. // Computer Algebra in Scientific Computing: 22nd International Workshop, CASC 2020, Linz, Austria, September 14–18, 2020, Proceedings. Sep 2020. Pages 1–13. https://doi.org/10.1007/978-3-030-60026-6_1
- [4] Li, D. Morphological template decomposition with max-polynomials. // J Math Imaging Vis 1, 215–221, 1992. <https://doi.org/10.1007/BF00129876>
- [5] Grigoriev, D., Shpilrain, V. Tropical Cryptography // Communications in Algebra, 42:6, 2624-2632, 2014 DOI: 10.1080/00927872.2013.766827
- [6] Zhang, L., Naizat, G., Lim, L.-H. Tropical Geometry of Deep Neural Networks // Proceedings of the 35th International Conference on Machine Learning, PMLR 80:5824-5832, 2018, <http://proceedings.mlr.press/v80/zhang18i.html>
- [7] Кривулин, Н.К. Методы идепotentной алгебры в задачах моделирования и анализа сложных систем. изд. С.-Петерб. ун-та, 2009.

ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ СРОКОВ НАЧАЛА ПРОЕКТА ПРИ ПОМОЩИ МЕТОДОВ ТРОПИЧЕСКОЙ ОПТИМИЗАЦИИ

Губанов С. А., СПбФ КБ “Луч” инженер-программист,
segubanov@mail.ru

Аннотация

Исследована задача управления проектами, которая состоит в минимизации максимального отклонения от директивных сроков начала работ при заданных ограничениях. Представлено прямое аналитическое решение задачи, основанное на методах тропической оптимизации.

Введение

Задачи оптимального временного планирования (составления календарных графиков) выполнения работ являются важной проблемой, которая возникает при управлении проектами [1, 2]. Эффективным подходом к решению таких задач является использование методов тропической математики, изучающей различные аспекты теории и применения алгебраических систем с идемпотентными операциями [3, 4, 5]. Применение методов тропической математики в задачах планирования исследуется, например, в статьях [6, 7, 8, 9]. В настоящей работе рассматривается задача минимизации максимального отклонения от директивных сроков начала работ при различных ограничениях на время начала и завершения работ. Задача сводится к задаче тропической оптимизации, при решении которой используется результат работы [10].

Задачи оптимального планирования

При составлении календарных графиков выполнения работ проекта возникают задачи, связанные с необходимостью обеспечить директивные сроки начала или завершения выполнения работ. Изучим подобную задачу, в которой требуется минимизировать максимальное отклонение времени начала работ от директивных сроков. В задаче имеются строгие ограничения, заданные в форме минимальных допустимых временных интервалов между началом и завершением различных работ (ограничения “старт-старт”, “старт-финиш” и “финиш-старт”), а также в виде границ для времени начала и завершения каждой работы.

Для каждой работы $i = 1, \dots, n$ обозначим время начала и завершения через x_i и y_i . Самое раннее и самое позднее допустимое время начала работы обозначим через g_i и h_i , а наиболее позднее время завершения — через f_i . Эти величины устанавливают следующие временные границы для начала и завершения работ:

$$g_i \leq x_i \leq h_i, \quad y_i \leq f_i.$$

Обозначим через b_{ij} минимально допустимый интервал времени между началом работы i и началом работы j . Если величина b_{ij} не задана, считаем $b_{ij} = -\infty$. Определим ограничения вида “старт-старт” в форме неравенств $b_{ij} + x_j \leq x_i$. После объединения неравенств по всем $j = 1, \dots, n$ получим

$$\max_{1 \leq j \leq n} (b_{ij} + x_j) \leq x_i.$$

Пусть c_{ij} — минимальный допустимый интервал между временем начала работы i и временем завершения работы j ($c_{ij} = -\infty$, если интервал не задан). Запишем ограничения “старт-финиш” в виде неравенств $c_{ij} + x_j \leq y_i$. Объединив неравенства по всем j , имеем неравенство

$$\max_{1 \leq j \leq n} (c_{ij} + x_j) \leq y_i.$$

Предполагается, что работа завершается как только для нее выполняются заданные ограничения “старт-финиш”. Из этого следует, что хотя бы для одного j выполняется равенство $c_{ij} + x_j = y_i$, а предыдущее неравенство можно заменить равенством

$$\max_{1 \leq j \leq n} (c_{ij} + x_j) = y_i.$$

Для представления ограничений “финиш-старт” обозначим минимальный допустимый интервал между временем завершения работы i и временем начала j — через d_{ij} (если интервал не задан, будем считать $d_{ij} = -\infty$) и запишем ограничения в виде неравенства

$$\max_{1 \leq j \leq n} (d_{ij} + y_j) \leq x_i.$$

Пусть для каждой работы i определен директивный срок (установочное время) начала p_i . Будем считать, что в задаче требуется по возможности минимизировать максимальное отклонение от директивного времени начала снизу $p_i - x_i$ и сверху $x_i - p_i$ по всем работам

$i = 1, \dots, n$. Тогда критерий оптимальности плана, который требуется минимизировать, записывается в виде

$$\max \left(\max_{1 \leq i \leq n} (p_i - x_i), \max_{1 \leq i \leq n} (x_i - p_i) \right).$$

Сформулируем задачу минимизации максимального отклонения от директивных сроков начала:

$$\begin{aligned} \min_{x_i, y_i} \quad & \max \left(\max_{1 \leq i \leq n} (p_i - x_i), \max_{1 \leq i \leq n} (x_i - p_i) \right); \\ & \max_{1 \leq j \leq n} (b_{ij} + x_j) \leq x_i, \quad \max_{1 \leq j \leq n} (c_{ij} + x_j) = y_i, \\ & \max_{1 \leq j \leq n} (d_{ij} + y_j) \leq x_i, \quad g_i \leq x_i \leq h_i, \\ & y_i \leq f_i, \quad i = 1, \dots, n. \end{aligned} \quad (1)$$

В следующих разделах эта задача и ее решение будут представлены в терминах тропической математики в компактной векторной форме.

Определения и обозначения

Для описания и решения задачи тропической оптимизации, к которой в дальнейшем сводится рассматриваемая задача составления календарного графика, потребуется обзор используемых определений и обозначений тропической (идемпотентной) математики [3, 4, 5].

Идемпотентным полуполем называется алгебраическая система $\langle \mathbb{X}, \emptyset, \mathbb{1}, \oplus, \otimes \rangle$, где \mathbb{X} – множество, замкнутое относительно ассоциативных и коммутативных операций сложения \oplus и умножения \otimes . Множество \mathbb{X} содержит нейтральные элементы ноль \emptyset и единицу $\mathbb{1}$. Сложение обладает свойством идемпотентности, т.е. для любого $x \in \mathbb{X}$ выполняется равенство $x \oplus x = x$, а умножение дистрибутивно относительно сложения и обратимо, т.е. для каждого $x \neq \emptyset$ существует обратный элемент x^{-1} такой, что $x \otimes x^{-1} = \mathbb{1}$. Знак операции умножения \otimes для простоты далее будет опускаться.

При помощи идемпотентного сложения задан следующий частичный порядок: $x \leq y$ тогда и только тогда, когда $x \oplus y = y$. Считается, что этот частичный порядок дополнен до линейного порядка на \mathbb{X} .

Целая степень определяется как обычно: $x^0 = \mathbb{1}$, $x^p = x^{p-1}x$, $x^{-p} = (x^{-1})^p$, $\emptyset^p = \emptyset$ для каждого $x \neq \emptyset$ и целого $p > 0$. Предполагается, что операция возведения в рациональную степень также определена.

Вещественное полуполе $\mathbb{R}_{\max,+} = \langle \mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, + \rangle$, где $0 = -\infty$, $1 = 0$, $\oplus = \max$ и $\otimes = +$ является примером идемпотентного полуполя и будет далее использовано для решения исследуемой задачи.

Обозначим через $\mathbb{X}^{m \times n}$ множество матриц, состоящих из m строк и n столбцов с элементами из \mathbb{X} , а через \mathbb{X}^n – множество векторов-столбцов из n элементов. Матрица называется регулярной по столбцам, если она не имеет нулевых столбцов.

Для согласованных по размеру матриц и векторов сложение, умножение и умножение на скаляр выполняются по обычным формулам с заменой арифметических операций $+$ и \times на \oplus и \otimes . Отношение порядка, которое было определено выше, обобщается на матрицы и векторы, и понимается покомпонентно.

Для любого вектора $x \in \mathbb{X}^n$ обозначим транспонированный вектор как x^T . Вектор, все элементы которого равны 0 , является нулевым. Вектор называется регулярным, если он не имеет нулевых элементов.

Для ненулевого вектора $x = (x_i)$ определен мультипликативно со-пряженный вектор-строка $x^- = (x_i^-)$ с элементами $x_i^- = x_i^{-1}$ при условии, что $x_i \neq 0$, и $x_i^- = 0$ в противном случае.

Для любой квадратной матрицы A определена целая степень $p > 0$ следующим образом: $A^0 = I$, $A^p = A^{p-1}A$, где I – единичная матрица с элементами, равными 1 на главной диагонали и 0 – вне ее.

Для квадратной матрицы $A = (a_{ij})$ порядка n след вычисляется по формуле

$$\text{tr } A = a_{11} \oplus \cdots \oplus a_{nn} = \bigoplus_{i=1}^n a_{ii}.$$

Для матрицы A введем функцию

$$\text{Tr}(A) = \text{tr } A \oplus \cdots \oplus \text{tr } A^n = \bigoplus_{k=1}^n \text{tr } A^k.$$

При выполнении условия $\text{Tr}(A) \leq 1$ определена матрица Клини

$$A^* = I \oplus A \oplus \cdots \oplus A^{n-1} = \bigoplus_{k=0}^{n-1} A^k.$$

Решение задачи оптимального планирования

Представим задачу минимизации максимального отклонения от директивных сроков начала работ (1) в терминах идемпотентного полу-

поля $\mathbb{R}_{\max,+}$. Сперва введем следующие матрицы и векторы:

$$\begin{aligned} \mathbf{B} &= (b_{ij}), \quad \mathbf{C} = (c_{ij}), \quad \mathbf{D} = (d_{ij}), \\ \mathbf{x} &= (x_i), \quad \mathbf{y} = (y_i), \quad \mathbf{f} = (f_i), \quad \mathbf{g} = (g_i), \quad \mathbf{h} = (h_i), \quad \mathbf{p} = (p_i). \end{aligned}$$

Путем замены арифметических операций на операции полуполя $\mathbb{R}_{\max,+}$ представим задачу в векторном виде в форме

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{y}} \quad & \mathbf{x}^-\mathbf{p} \oplus \mathbf{p}^-\mathbf{x}, \\ \mathbf{B}\mathbf{x} \leq \mathbf{x}, \quad & \mathbf{C}\mathbf{x} = \mathbf{y}, \\ \mathbf{D}\mathbf{y} \leq \mathbf{x}, \quad & \mathbf{g} \leq \mathbf{x} \leq \mathbf{h}, \\ \mathbf{y} \leq \mathbf{f}. \end{aligned} \tag{2}$$

На основе применения результатов работы [10] получим решение задачи оптимизации в виде следующего утверждения.

Лемма 1 Пусть \mathbf{B} и \mathbf{D} – матрицы, \mathbf{C} – регулярная по столбцам матрица такие, что матрица $\mathbf{R} = \mathbf{B} \oplus \mathbf{DC}$ удовлетворяет условию $\text{Tr}(\mathbf{R}) \leq 1$. Пусть \mathbf{g} – вектор, а \mathbf{f} и \mathbf{h} – регулярные векторы такие, что вектор $\mathbf{s}^T = \mathbf{f}^-\mathbf{C} \oplus \mathbf{h}^-$ удовлетворяет условию $\mathbf{s}^T \mathbf{R}^* \mathbf{g} \leq 1$.

Тогда минимальное значение целевой функции в задаче (2) равно

$$\theta = (\mathbf{p}^-\mathbf{R}^*\mathbf{p})^{1/2} \oplus \mathbf{s}^T \mathbf{R}^* \mathbf{p} \oplus \mathbf{p}^-\mathbf{R}^*\mathbf{g},$$

а все регулярные решения имеют вид

$$\mathbf{x} = \mathbf{R}^* \mathbf{u}, \quad \mathbf{y} = \mathbf{C} \mathbf{R}^* \mathbf{u},$$

где \mathbf{u} – любой регулярный вектор, который удовлетворяет условиям

$$\mathbf{g} \oplus \theta^{-1} \mathbf{p} \leq \mathbf{u} \leq ((\mathbf{s}^T \oplus \theta^{-1} \mathbf{p}^-) \mathbf{R}^*)^-.$$

Заключение

Изучена задача минимизации максимального отклонения времени начала работ от директивных сроков при ограничениях вида “старт-старт”, “старт-финиш”, “финиш-старт” и допустимых границах для времени начала и времени окончания работ. Предложено аналитическое решение задачи в компактной векторной форме на основе использования методов и результатов тропической оптимизации.

Список литературы

- [1] T'kindt V. and Billaut J.-C. Multicriteria Scheduling. 2 ed. Berlin: Springer, 2006.
- [2] Kerzner H. Project Management. 10 ed. Hoboken: Wiley, 2010. 1094 p.
- [3] Маслов В. П., Колокольцев В. Н. Идемпотентный анализ и его применение в оптимальном управлении. М.: Физматлит, 1994.
- [4] Кривулин Н. К. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб.: Изд-во С.-Петерб. ун-та., 2009.
- [5] Butkovič P., Max-linear Systems: Theory and Algorithms. Springer Monographs in Mathematics. London: Springer, 2010.
- [6] Krivulin N. Direct solution to constrained tropical optimization problems with application to project scheduling // Computational Management Science. 2017. Vol. 14. N 1. P. 91-113.
- [7] Krivulin N., Tropical optimization problems with application to project scheduling with minimum makespan // Annals of Operations Research. 2017. Vol. 256, N 1. P. 75-92.
- [8] Krivulin N., Tropical optimization problems in time-constrained project scheduling. // Optimization. 2017. Vol. 66, N 2. P. 205-224.
- [9] Кривулин Н. К., Губанов С. А. Алгебраическое решение задачи оптимального планирования сроков проекта в управлении проектами // Вестник Санкт-Петербургского университета. Математика. Механика. Астрономия. 2021. Т. 8. Вып. 1. С. 73-87.
- [10] Krivulin N. Complete solution of a constrained tropical optimization problem with application to location analysis // Relational and Algebraic Methods in Computer Science. Cham: Springer, 2014. P. 362-378. (Lecture Notes in Computer Science, Vol. 8428.)

ОЦЕНКА АЛЬТЕРНАТИВ НА ОСНОВЕ ПАРНЫХ СРАВНЕНИЙ В ЗАДАЧЕ О ВЫБОРЕ СТРОИТЕЛЬНОГО ПОДРЯДЧИКА

Григорьев Д. А., студент 4-го курса бакалавриата СПбГУ,
 dmitry.hpbgrigorev@gmail.com

Руководитель: Кривулин Н. К., д.ф.-м.н., профессор кафедры статистического моделирования СПбГУ, nkk@math.spbu.ru

Аннотация

Рассматривается многокритериальная задача оценки альтернатив на основе парных сравнений, которая возникает при выборе строительного подрядчика. Для решения задачи используются методы анализа иерархий, взвешенных геометрических средних и log-чебышевской аппроксимации. Сопоставление полученных решений показывает близость результатов применения указанных методов.

Введение

Рассматривается многокритериальная задача оценки рейтингов альтернатив на основе парных сравнений по нескольким неравноценным критериям на примере задачи о выборе строительного подрядчика. Существует ряд методов, которые позволяют найти решение для таких задач, включая метод анализа иерархий Т. Саати [1, 2], метод взвешенных геометрических средних [3], а также метод log-чебышевской аппроксимации [4, 5], использующий инструменты тропической математики. Известно [2, 6], что решение, полученное одним методом, может существенно отличаться от результатов другого метода. В этих условиях для получения более обоснованного решения целесообразно сопоставлять результаты разных методов: если они мало различаются или совпадают, это может служить дополнительным обоснованием выбора любого из решений в качестве близкого к оптимальному.

Постановка задачи

В задаче о выборе строительного подрядчика [7] исследуются 5 альтернатив выбора строительной компании (A, B, C, D, E), которые сравниваются попарно по 6 критериями (опыт работы, финансовая стабиль-

ность, качество работы, обеспечение персоналом, обеспечение оборудованием, текущая рабочая загрузка). Результаты парных сравнений альтернатив по каждому критерию описываются матрицами:

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 1/3 & 1/2 & 1/6 & 2 \\ 3 & 1 & 2 & 1/2 & 4 \\ 2 & 1/2 & 1 & 1/3 & 3 \\ 6 & 2 & 3 & 1 & 7 \\ 1/2 & 1/4 & 1/3 & 1/7 & 1 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 1 & 6 & 3 & 2 & 7 \\ 1/6 & 1 & 1/4 & 1/2 & 3 \\ 1/3 & 4 & 1 & 1/3 & 5 \\ 1/2 & 2 & 3 & 1 & 7 \\ 1/7 & 1/3 & 1/5 & 1/7 & 1 \end{pmatrix},$$

$$\mathbf{A}_3 = \begin{pmatrix} 1 & 7 & 1/3 & 2 & 8 \\ 1/7 & 1 & 1/5 & 1/4 & 4 \\ 3 & 5 & 1 & 4 & 9 \\ 1/2 & 4 & 1/4 & 1 & 6 \\ 1/8 & 1/4 & 1/9 & 1/6 & 1 \end{pmatrix}, \quad \mathbf{A}_4 = \begin{pmatrix} 1 & 1/2 & 1/4 & 2 & 5 \\ 2 & 1 & 1/3 & 5 & 7 \\ 4 & 3 & 1 & 4 & 6 \\ 1/2 & 1/5 & 1/4 & 1 & 2 \\ 1/5 & 1/7 & 1/6 & 1/2 & 1 \end{pmatrix},$$

$$\mathbf{A}_5 = \begin{pmatrix} 1 & 1/6 & 1/8 & 2 & 3 \\ 6 & 1 & 1/4 & 5 & 7 \\ 8 & 4 & 1 & 9 & 9 \\ 1/2 & 1/5 & 1/9 & 1 & 2 \\ 1/3 & 1/7 & 1/9 & 1/2 & 1 \end{pmatrix}, \quad \mathbf{A}_6 = \begin{pmatrix} 1 & 1/5 & 1/3 & 3 & 3 \\ 5 & 1 & 5 & 6 & 6 \\ 3 & 1/5 & 1 & 2 & 2 \\ 1/3 & 1/6 & 1/2 & 1 & 2 \\ 1/3 & 1/6 & 1/2 & 1/2 & 1 \end{pmatrix},$$

а результаты парных сравнений критериев — матрицей:

$$\mathbf{C} = \begin{pmatrix} 1 & 2 & 3 & 6 & 6 & 5 \\ 1/2 & 1 & 3 & 6 & 6 & 5 \\ 1/3 & 1/3 & 1 & 4 & 4 & 3 \\ 1/6 & 1/6 & 1/4 & 1 & 2 & 1/2 \\ 1/6 & 1/6 & 1/4 & 1/2 & 1 & 1/4 \\ 1/5 & 1/5 & 1/3 & 2 & 4 & 1 \end{pmatrix}.$$

Требуется найти вектор рейтингов альтернатив \mathbf{x} , который задаёт порядок на множестве альтернатив. Для решения используются три метода: метод анализа иерархий, метод взвешенных геометрических средних и метод лог-чебышевской аппроксимации.

Метод анализа иерархий

Метод анализа иерархий (Analytic Hierarchy Process – АНР) использует главные собственные векторы матриц парных сравнений, которые соответствуют их максимальным собственным числам. Собственные векторы нормируются относительно суммы элементов, а затем вектор рейтингов альтернатив вычисляется как взвешенная сумма собственных векторов $\mathbf{a}_1, \dots, \mathbf{a}_6$ матриц $\mathbf{A}_1, \dots, \mathbf{A}_6$, в которой в качестве весов берутся элементы собственного вектора $\mathbf{w} = (w_i)$ матрицы \mathbf{C} .

После вычисления нормированных собственных векторов получим вектор весов критериев

$$\mathbf{w} \approx (0.377, 0.299, 0.155, 0.050, 0.037, 0.082)^T,$$

а также векторы

$$\mathbf{a}_1 \approx (0.086, 0.250, 0.152, 0.458, 0.055)^T,$$

$$\mathbf{a}_2 \approx (0.428, 0.085, 0.176, 0.273, 0.038)^T,$$

$$\mathbf{a}_3 \approx (0.271, 0.067, 0.472, 0.160, 0.030)^T,$$

$$\mathbf{a}_4 \approx (0.147, 0.269, 0.463, 0.078, 0.043)^T,$$

$$\mathbf{a}_5 \approx (0.077, 0.261, 0.574, 0.053, 0.036)^T,$$

$$\mathbf{a}_6 \approx (0.134, 0.547, 0.180, 0.079, 0.060)^T.$$

Вычисление вектора рейтингов альтернатив дает результат

$$\sum_{i=1}^6 w_i \mathbf{a}_i \approx (0.223, 0.198, 0.242, 0.291, 0.045)^T.$$

После нормировки относительно максимального элемента получим

$$\mathbf{x}_{АНР} \approx (0.766, 0.680, 0.832, 1, 0.155)^T.$$

Метод взвешенных геометрических средних

Метод взвешенных геометрических средних (Weighted Geometric Means – WGM) опирается на вычисление геометрических средних строк матриц парных сравнений. Сначала определяется нормированный относительно суммы элементов вектор $\mathbf{w} = (w_i)$ геометрических средних матрицы парных сравнений критериев, а также векторы $\mathbf{a}_1, \dots, \mathbf{a}_6$ геометрических средних матриц парных сравнений альтернатив. Элементы каждого вектора \mathbf{a}_i возводятся в степень w_i , а затем результат перемножения соответствующих элементов этих векторов берется в качестве элемента x_i вектора рейтингов $\mathbf{x} = (x_i)$.

Нормированный вектор геометрических средних для матрицы \mathbf{C} :

$$\mathbf{w} \approx (0.377, 0.299, 0.156, 0.051, 0.036, 0.081)^T.$$

Векторы геометрических средних для матриц $\mathbf{A}_1, \dots, \mathbf{A}_6$:

$$\mathbf{a}_1 \approx (0.561, 1.644, 1.0, 3.022, 0.359)^T,$$

$$\mathbf{a}_2 \approx (3.022, 0.574, 1.173, 1.838, 0.267)^T,$$

$$\mathbf{a}_3 \approx (2.063, 0.491, 3.519, 1.246, 0.225)^T,$$

$$\mathbf{a}_4 \approx (1.046, 1.878, 3.104, 0.549, 0.299)^T,$$

$$\mathbf{a}_5 \approx (0.660, 2.208, 4.817, 0.467, 0.305)^T,$$

$$\mathbf{a}_6 \approx (0.903, 3.898, 1.191, 0.561, 0.425)^T.$$

Тогда вектор рейтингов альтернатив определяется выражением

$$\prod_{i=1}^6 (\mathbf{a}_i)^{w_i} \approx (1.227, 1.086, 1.452, 1.695, 0.305)^T,$$

где операции над векторами выполняются покомпонентно.

Нормированный относительно максимума вектор имеет вид

$$\mathbf{x}_{WGM} \approx (0.725, 0.640, 0.857, 1, 0.180)^T.$$

Метод log-чебышевской аппроксимации

Решение задачи оценки рейтингов альтернатив этим методом состоит в аппроксимации матриц парных сравнений обратно симметрическими матрицами единичного ранга в метрике Чебышева в логарифмической шкале. Применение метода log-чебышевской аппроксимации приводит к ряду задач оптимизации, которые формулируются и решаются в терминах тропической математики [8, 9]. Процедура оценки рейтингов альтернатив на основе парных сравнений с использованием методов тропической оптимизации представлена в работах [4, 5].

Тропическая математика изучает теорию и приложения алгебраических систем с идемпотентными операциями. Примером системы является max-алгебра, которая определена на множестве неотрицательных вещественных чисел с операциями сложения \oplus и умножения \times . Сложение определено как $x \oplus y = \max(x, y)$ и является идемпотентным, так как $x \oplus x = \max(x, x) = x$, а умножение определено как обычно.

Векторные и матричные операции выполняются по стандартным правилам с заменой операции арифметического сложения $+$ на тропическое сложение \oplus . Свойство коллинеарности векторов имеет обычный смысл. Понятия линейной комбинации и линейной зависимости векторов определены как обычно с заменой векторной операции $+$ на \oplus .

След квадратной матрицы $\mathbf{A} = (a_{ij})$ размера n вычисляется по формуле $\text{tr } \mathbf{A} = a_{11} \oplus \dots \oplus a_{nn}$. Спектральным радиусом матрицы \mathbf{A} называется число

$$\lambda = \text{tr } \mathbf{A} \oplus (\text{tr } \mathbf{A}^2)^{1/2} \oplus \dots \oplus (\text{tr } \mathbf{A}^n)^{1/n}.$$

Если $\lambda \leq 1$, то для матрицы \mathbf{A} определен оператор Клини в виде

$$\mathbf{A}^* = \mathbf{I} \oplus \mathbf{A} \oplus \mathbf{A}^2 \oplus \dots \mathbf{A}^{n-1}.$$

Нахождение весов критериев

Сначала для матрицы \mathbf{C} сравнений критериев следует найти спектральный радиус λ и матрицу Клини $(\lambda^{-1}\mathbf{C})^*$. В результате получаем

$$\lambda = 12^{1/5}, \quad (\lambda^{-1}\mathbf{C})^* = \begin{pmatrix} 1 & 2/\lambda & 6/\lambda^2 & 24/\lambda^3 & 72/\lambda^4 & 18/\lambda^3 \\ 6/\lambda^4 & 1 & 3/\lambda & 12/\lambda^2 & 36/\lambda^3 & 9/\lambda^2 \\ 2/\lambda^3 & 4/\lambda^4 & 1 & 4/\lambda & 12/\lambda^2 & 3/\lambda \\ 1/3\lambda^2 & 2/3\lambda^3 & 2/\lambda^4 & 1 & 2/\lambda & 1/2 \\ 1/6\lambda & 1/3\lambda^2 & 1/\lambda^3 & 4/\lambda^4 & 1 & 3/\lambda^4 \\ 2/3\lambda^2 & 4/3\lambda^3 & 4/\lambda^4 & 4/3 & 4/\lambda & 1 \end{pmatrix}.$$

Столбцы этой матрицы определяют один вектор весов критериев, если все столбцы коллинеарны, или несколько векторов весов в противном случае. Если среди векторов-столбцов матрицы $(\lambda^{-1}\mathbf{C})^*$ имеется несколько линейно независимых векторов, то среди них выбираются в некотором смысле наилучший и наихудший вектор весов.

В качестве наилучшего вектора берется столбец, для которого отношение между наибольшим и наименьшим элементами максимальное, т.е. столбец, который наилучшим образом дифференцирует (различает) критерии с максимальным и минимальным весами. В роли наихудшего вектора берется столбец, для которого это отношение минимально, т.е. столбец, который хуже всего дифференцирует критерии.

Определение наихудшего дифференцирующего вектора весов для рассматриваемой матрицы $(\lambda^{-1}\mathbf{C})^*$ дает следующий результат:

$$\mathbf{w}_1 = (1, 6/\lambda^4, 2/\lambda^3, 1/2\lambda^2, 1/6\lambda, 2/3\lambda^2)^T.$$

Имеются два наилучших вектора весов, которые после нормировки относительно максимального элемента записываются в виде

$$\mathbf{w}_2 = (1, 6/\lambda^4, 2/\lambda^3, \lambda^3/36, 1/6\lambda, 2/3\lambda^2)^T,$$

$$\mathbf{w}_3 = (1, 6/\lambda^4, 2/\lambda^3, \lambda^3/24, 1/6\lambda, 2/3\lambda^2)^T.$$

Определение рейтингов альтернатив

Для определения рейтингов альтернатив необходимо исследовать взвешенные тропические суммы (максимумы) матриц парных сравнений альтернатив $\mathbf{A}_1, \dots, \mathbf{A}_6$, где векторы весов определены на основе матрицы \mathbf{C} парных сравнений критериев, как указано выше. Результаты построения взвешенных сумм показывают, что они одинаковы для

всех векторов $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$. Пусть $\mathbf{w} = (w_i)$ обозначает любой из этих векторов. Тогда взвешенная сумма представляет собой матрицу

$$\mathbf{B} = \bigoplus_{k=1}^6 w_k \mathbf{A}_k = \begin{pmatrix} 1 & 3\lambda & 3\lambda/2 & \lambda & 7\lambda/2 \\ 3 & 1 & 2 & 4/\lambda^2 & 4 \\ 2 & 2\lambda & 1 & 8/\lambda^3 & 5\lambda/2 \\ 6 & 2 & 3 & 1 & 7 \\ 1/2 & \lambda/6 & 1/3 & 1/7 & 1 \end{pmatrix}.$$

Теперь для матрицы \mathbf{B} требуется найти спектральный радиус μ и матрицу Клини $(\mu^{-1}\mathbf{B})^*$, которые записываются в виде

$$\mu = 3\lambda^{1/2}, \quad (\mu^{-1}\mathbf{B})^* = \begin{pmatrix} 1 & 27\lambda^2/\mu^3 & 2/3 & 4/3\lambda^2 & 63\lambda^2/(2\mu^3) \\ 27\lambda/\mu^3 & 1 & 18\lambda/\mu^3 & 4/\lambda^2\mu & 7/6 \\ 36\lambda^3/\mu^4 & 12\lambda^3/\mu^3 & 1 & 8/(\lambda^3\mu) & 14\lambda^3/\mu^3 \\ 54\lambda/\mu^3 & 2 & 36\lambda/\mu^3 & 1 & 7/3 \\ 9\lambda/(2\mu^3) & 1/6 & 3\lambda/\mu^3 & 6/\lambda\mu^3 & 1 \end{pmatrix}.$$

Столбцы матрицы $(\mu^{-1}\mathbf{B})^*$ используются для определения наихудшего и наилучшего дифференцирующих векторов рейтингов альтернатив. В результате находим наихудший дифференцирующий вектор

$$\mathbf{x}_1 = (\lambda^{1/2}/2, 1/2, 3\lambda^{1/2}/4, 1, 3/7)^T \approx (0.641, 0.5, 0.962, 1, 0.429)^T.$$

Наилучший дифференцирующий вектор рейтингов имеет вид

$$\begin{aligned} \mathbf{x}_2 &= (4/3\lambda^2, 4/3\lambda^{5/2}, 8/3\lambda^{7/2}, 1, 2/9\lambda^{5/2})^T \approx \\ &\approx (0.493, 0.385, 0.468, 1, 0.064)^T. \end{aligned}$$

Заключение

Для задачи о выборе строительного подрядчика на основе парных сравнений альтернатив A, B, C, D, E по 6 критериям получены векторы рейтингов альтернатив с помощью метода анализа иерархий $\mathbf{x}_{АНР}$, метода взвешенных геометрических средних \mathbf{x}_{WGM} , а также наихудший \mathbf{x}_1 и наилучший \mathbf{x}_2 дифференцирующие векторы рейтингов по методу лог-чебышевской аппроксимации.

В соответствии с величиной рейтингов векторы $\mathbf{x}_{АНР}$, \mathbf{x}_{WGM} и \mathbf{x}_1 устанавливают один и тот же порядок предпочтения альтернатив

$$D \succ C \succ A \succ B \succ E,$$

который естественно рассматривать как близкий к оптимальному.

Вектор рейтингов \mathbf{x}_2 определяет несколько иной порядок

$$D \succ A \succ C \succ B \succ E,$$

в котором альтернативы A и C меняются местами.

Объединение результатов можно представить в виде

$$D \succ C \parallel A \succ B \succ E,$$

где взаимный порядок альтернатив A и C считается не определенным.

Список литературы

- [1] Саати Т. Принятие решений. Метод анализа иерархий / пер. с англ. Р. Г. Вачнадзе. М.: Радио и связь, 1993. 315 с.
- [2] Saaty T. L., Vargas L. G. Comparison of eigenvalue, logarithmic least squares and least squares methods in estimating ratios // Math. Modelling. 1984. Vol. 5, N 5, P. 309–324.
- [3] Crawford G., Williams C. A note on the analysis of subjective judgment matrices // J. Math. Psych. 1985. Vol. 29, N 4. P. 387–405.
- [4] Кривулин Н. К., Агеев В. А. Методы тропической оптимизации в многокритериальных задачах оценки альтернатив на основе парных сравнений // Вестн. С.-Петерб. ун-та. Прикл. матем. 2019. Т. 15, №4. С. 472–488.
- [5] Krivulin N., Sergeev S. Tropical implementation of the Analytical Hierarchy Process decision method // Fuzzy Sets and Systems. 2019. Vol. 377. P. 31–51.
- [6] Tran N. M. Pairwise ranking: Choice of method can produce arbitrarily different rank order // Linear Algebra Appl. 2013. Vol. 438, N 3. P. 1012–1024.
- [7] Al-Harbi K. M. Application of the AHP in project management // Int. J. Project Manag. 2001. Vol. 19, N 1. P. 19–27.
- [8] Маслов В. П., Колокольцев В. Н. Идемпотентный анализ и его применение в оптимальном управлении. М.: Физматлит, 1994.
- [9] Кривулин Н. К. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб.: Изд-во С.-Петерб. ун-та., 2009.

**Вероятностные графические модели,
нечеткие системы, мягкие вычисления и
социокомпьютинг**



Тулупьев Александр Львович

д.ф.-м.н., профессор кафедры информатики СПбГУ, главный научный
сотрудник лаб. ТиМПИ СПб ФИЦ РАН

Методы предобработки текстовых постов

Эйрих М., студент кафедры информатики СПбГУ,
michael.eirich@mail.ru

Олисеенко В.Д., ассистент кафедры информатики СПбГУ,
vdo@dscs.pro

Абрамов М. В., к.т.н., доц. кафедры информатики СПбГУ,
mva@dscs.pro

Аннотация

Статья описывает методы предобработки текстовых данных на русском языке, которые необходимы для использования в машинном обучении. Методы включают удаление стоп-слов, токенизацию, нормализацию, стемминг и лемматизацию, и помогают улучшить качество анализа и классификации данных. В статье представляются примеры использования морфологических анализаторов и библиотек, для решения различных задач в области NLP [1].

Введение

В настоящее время предобработка данных остается одной из основных проблем при работе с большими объемами информации, что требует разработки более эффективных подходов к созданию и преобразованию данных. Это особенно важно при создании приложений для работы с данными, которые должны быть способны обрабатывать и анализировать большие объемы информации.

Предварительная обработка данных является наиболее важным этапом в анализе данных, поскольку, если этот этап пропустить, то аналитические алгоритмы, используемые для дальнейшего анализа, не будут работать правильно. Это может привести к некорректным или неэффективным результатам, что известно как принцип GIGO — garbage in, garbage out¹.

В статье рассмотрены различные методы предварительной обработки текстовых данных на русском языке, необходимые для эффективного использования алгоритмов машинного обучения. Также подробно описаны методы, применяемые в разработанном приложении, которые могут помочь улучшить качество обработки текста и повысить точность аналитических моделей.

¹ GIGO — <https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.1456>

Методология обработки данных

Стоп-слова

Одной из основных форм предварительной обработки является фильтрация ненужных данных [2]. В обработке естественного языка "бесполезные слова" называются стоп-словами [3]. Стоп-слова – это часто используемые слова, например "и", "или", "а", "в" и т.д., которые поисковая система должна игнорировать как при индексировании записей для поиска, так и при получении их в результате поискового запроса, за исключением строгого поиска по определенной фразе. При анализе текстовых данных и построении моделей NLP эти слова не придают документу особой важности. В случаях, когда необходимо классифицировать текст, например, при фильтрации спама [4] или генерации заголовков (комментариев) [5] к изображению или тегам, можно использовать техники удаления стоп-слов. Напротив, в таких задачах, как машинный перевод, языковое моделирование, краткое резюме текста, рекомендуется оставлять стоп-слова, так как они имеют большое значение (листинг 1).

```

import nltk
from nltk.corpus import stopwords

nltk.download("stopwords")
stop_words = stopwords.words("russian")
```

Листинг 1. Сбор стоп-слов с помощью NLTK

Токенизация

После фильтрации данных необходимо удалить знаки препинания и служебные символы. Это можно сделать вручную, написав код для проверки и удаления ненужных символов, либо использовать инструменты из NLTK (Natural Language Toolkit)². Кроме того, этот процесс может быть объединен с токенизацией предложений на слова - выделением отдельных токенов каждого предложения. Токенизация предложений должна учитывать особенности языка, с которым вы работаете [6]. Например, при использовании инструментов NLTK для русского языка мы можем получить неожиданные результаты. Например, слово "Санкт-Петербург" может быть разделено на два отдельных слова "Санкт" и "Петербург", что

² NLTK — <https://www.nltk.org>

может привести к потере смысла в некоторых случаях. Поэтому на этом этапе необходимо быть особенно внимательным (листинг 2).

Не ветер, а какой-то ураган!

Листинг 2. Исходный текст до токенизации и удаления стоп-слов

```
[ 'Не', 'ветер', ',', 'какой-то', 'ураган', '!' ]
```

Листинг 3. Текст после токенизации и удаления стоп-слов

Нормализация слов

В любом естественном языке слова могут быть записаны в более чем в одной форме, в зависимости от ситуации.

Например:

- «Я был на конференции вчера».
- «Я буду выступать на конференции 28 апреля».
- «Я иногда бываю на конференциях».

Во всех этих предложениях мы видим, что слово «быть» употребляется в нескольких различных формах. Для нас, людей, это легко понять, что «быть» представляет собой некую деятельность, независимо от формы, в которой мы видим это слово — «был», «бываем», «был», «бываем» и т.д. Однако для компьютеров все эти формы являются разными словами, поэтому необходима их нормализация к корневому слову «быть». Нормализация — это процесс приведения слова к единой канонической форме [7], который может быть выполнен двумя способами — стеммингом [8] и лемматизацией [9].

Стемминг и Лемматизация

Варианты применения алгоритмов стемминга и лемматизации зависят от цели и задач, которые перед ними ставятся.

Стемминг часто используется в поисковых системах для уменьшения количества форм слова и упрощения поиска по ключевым словам. Однако этот метод не учитывает контекст и может приводить к ошибкам, так как не всегда корневая форма совпадает с исходным словом.

Лемматизация используется для более точной обработки текста, так как она учитывает контекст и грамматические отношения слов в предложении. Она часто применяется в задачах обработки естественного языка, таких как машинный перевод, анализ тональности, категоризация текста и т.д. Однако этот метод требует больше вычислительных ресурсов и времени на обработку, чем стемминг.

Также есть комбинированные методы, которые используют как стемминг, так и лемматизацию в зависимости от цели обработки текста и типа задачи.

Для неанглийских слов можно использовать стеммер Snowball³. На самом деле, Snowball — это язык для создания стеммеров и был добавлен в NLTK версии 2.0b9 в виде отдельного класса SnowballStemmer. Этот стеммер поддерживает следующие языки: датский, английский, финский, немецкий, испанский, шведский, финский, самый важный для нас русский и некоторые другие языки. Так как NLTK имеет мало возможностей для русского языка, мы рассмотрим только стемминг с помощью SnowballStemmer. Рассмотрим исходный текст и текст полученный после стемминга. Как мы видим, после выполнения обработки слов с помощью стемминга, большинство слов были обрезаны (листинг 4).

```
from nltk.stem import SnowballStemmer
snowball = SnowballStemmer(language="russian")

# Функция обработки
def snowball_text(text):

    # Приводим весь текст к нижнему регистру
    text = text.lower()
    # Разбиваем текст на слова
    words = text.split()
    # Получаем токены
    tokens = [snowball.stem(word) for word in words]

    # Объединяем токены в текст
    text = ' '.join(tokens)

    # Возвращаем лемматизированный текст
    return text

# Проверка работоспособности функции лемматизации на основе SnowballStemmer
text = "Никто так не нуждается в отпуске, как человек, только что вернувшийся из отпуска (с)"
snowball_text(text)

'никт так не нужда в отпуске, как человек, тольк что вернувш из отпуск (с)'
```

Листинг 4. Стемминг с использованием SnowballStemmer из NLTK

Когда доступны два варианта, лемматизация всегда будет лучшим

³Snowball — https://www.nltk.org/_modules/nltk/stem/snowball.html

вариантом, чем стемминг [10]. Стемминг алгоритмы являются оптимизированным способом идентификации родственных слов с помощью относительно короткого алгоритма и без необходимости в словарных данных для каждого языка. Недостатком является то, что он не всегда точен: иногда он соединяет родственными отношениями слова, которые не происходят от одного и того же слова, но, с другой стороны, не идентифицирует родственные формы конкретного слова. В свою очередь, лемматизация всегда даст лучший результат, потому что лемматизаторы полагаются на правильные языковые данные (словари) для идентификации слова с его леммой. Кроме того, результатом всегда будет другой элемент словаря (инфinitивы, формы единственного числа и т.д.), а не "основа", с которой иногда могут возникнуть трудности.

Рассмотрим лемматизацию с помощью морфологического анализатора pymorphy2⁴ (листинг 5).

```
import pymorphy2
morph = pymorphy2.MorphAnalyzer()

# Функция обработки
def pymorphy2_text(text):

    # Приводим весь текст к нижнему регистру
    text = text.lower()
    # Разбиваем текст на слова
    words = text.split()
    # Получаем токены
    tokens = [morph.parse(word)[0].normal_form for word in words]

    # Объединяем токены в текст
    text = ' '.join(tokens)

    # Возвращаем лемматизированный текст
    return text

# Проверка работоспособности функции лемматизации на основе pymorphy2
text = "Никто так не нуждается в отпуске, как человек, только что вернувшийся из отпуска (с)"
pymorphy2_text(text)

'никто так не нуждается в отпуске, как человек, только что вернуться из отпуска (с)'
```

Листинг 5. Лемматизация с помощью PyMorphy2

Заключение

В данной работе были рассмотрены основные методы предварительной обработки данных, необходимые для их эффективного использования в задачах машинного обучения. Были

⁴Pymorphy2 — <https://pymorphy2.readthedocs.io/en/stable/>

подробно описаны такие методы, как токенизация, удаление стоп-слов и нормализация. Кроме того, были представлены различные подходы к реализации этих методов с использованием библиотек NLTK, SnowballStemmer и ruymorph2. Однако, следует помнить, что выбор конкретного метода зависит от задачи и свойств исходных данных.

Литература

1. Ofer D., Brandes N., Linial M. The language of proteins: NLP, machine learning & protein sequences //Computational and Structural Biotechnology Journal. – 2021. – Т. 19. – С. 1750-1758.
2. Maharana K., Mondal S., Nemade B. A review: Data pre-processing and data augmentation techniques //Global Transitions Proceedings. – 2022.
3. Shelke N. et al. An efficient way of text-based emotion analysis from social media using LRA-DNN //Neuroscience Informatics. – 2022. – С. 100048.
4. Dada E. G. et al. Machine learning for email spam filtering: review, approaches and open research problems //Heliyon. – 2019. – Т. 5. – №. 6. – С. e01802.
5. Huang Y. et al. Towards automatically generating block comments for code snippets //Information and Software Technology. – 2020. – Т. 127. – С. 106373.
6. Aso M. et al. Acoustic model-based subword tokenization and prosodic-context extraction without language knowledge for text-to-speech synthesis //Speech Communication. – 2020. – Т. 125. – С. 53-60.
7. Mehmmood K. et al. An unsupervised lexical normalization for Roman Hindi and Urdu sentiment analysis //Information Processing & Management. – 2020. – Т. 57. – №. 6. – С. 102368.
8. Singh J., Gupta V. A novel unsupervised corpus-based stemming technique using lexicon and corpus statistics //Knowledge-Based Systems. – 2019. – Т. 180. – С. 147-162.
9. Freihat A. A. et al. Towards an optimal solution to lemmatization in Arabic //Procedia computer science. – 2018. – Т. 142. – С. 132-140.

10. Moon S., Chi S., Im S. B. Automated detection of contractual risk clauses from construction specifications using bidirectional encoder representations from transformers (BERT) //Automation in Construction. – 2022. – Т. 142. – С. 104465.

Автоматизация сбора данных для оценки выраженности личностных особенностей пользователей социальных сетей

Ляпин Н.Е., студент СПбГУ, st075646@student.spbu.ru
Корепанова А.А., м.н.с. СПб ФИЦ РАН, aak@dscs.pro

Аннотация

В данной статье описывается реализация инструмента для автоматизации сбора данных пользователей социальной сети ВКонтакте для последующего поиска взаимосвязей между личностными особенностями пользователей и публикуемым ими контентом, с целью применения найденных фактов в сфере информационной безопасности и других смежных областях.

Введение

Повседневность человека в 2022 году невозможно представить без взаимодействия с самыми разными информационными системами, переносящими в цифровую среду множество различных видов деятельности [1]. Общение также подвержено цифровизации, и это особенно заметно в последние годы [2]. Важной площадкой для данного процесса служат социальные сети, количество новых пользователей которых с каждым годом лишь увеличивается [3]. Вместе с ростом вовлечённости в них, растёт и объём данных, которые пользователи оставляют о себе в сети.

Личностные особенности и психологические черты миллионов людей отражены на их страницах в виде записей, фотографий, лайков и т.д. Это огромный массив информации, который может быть потенциально использован для различных исследований в областях, изучающих человеческое поведение или аспекты, связанные с ним [4-5].

Данный проект посвящён разработке веб-приложения для упрощения и ускорения сбора данных для оценки личностных особенностей пользователей социальных сетей. Найденные с его помощью зависимости могут быть применены во множестве областей, вроде информационной безопасности, психологии и др.

Концепция приложения

Для удобства сбора данных было решено реализовать мини-приложение с психологическими и иными тестами [6]. Оно легко интегрируется в социальную сеть ВКонтакте и позволяет получать информацию, представляющую собой результаты тестирований по любым специально разработанным методикам (в том числе, для исследования личностных особенностей респондентов).

В концепции приложения были реализованы следующие подходы:

- Каждый тест конструируется из нескольких блоков, взаимодействующих друг с другом. Если подходящих блоков среди ранее реализованных нет – их можно реализовать и внедрить. Благодаря данному подходу получается достичь большей гибкости за счёт того, что в приложение возможно добавить тестирования любой конфигурации и сложности.
- Документно-ориентированная база данных. Использование такой БД позволяет легко внедрять новые функциональности на уровне, связывающем серверную часть и БД, без необходимости постоянного исправления SQL-скриптов [7].

Использованные технологии

1. Интеграция:
 - 1.1. VK Mini Apps — платформа социальной сети ВКонтакте для мини-приложений;
2. Клиентская часть:
 - 2.1. React — JavaScript-библиотека для простой разработки пользовательских интерфейсов;
 - 2.2. VKUI, VK Bridge, VK Icons — библиотеки команды ВКонтакте, предназначенный для удобного создания React-приложений на платформе VK Mini Apps в общей для социальной сети стилистике;
3. Серверная часть:
 - 3.1. Node.js — программная платформа, выполняющая роль веб-сервера, предоставляющая возможность программирования на языке JavaScript;
 - 3.2. Express — фреймворк-обёртка для веб-приложений на

Node.js, позволяющий более простым образом использовать серверные функции;

3.3. Mongoose — библиотека для взаимодействия веб-приложения на Express с документно-ориентированной базой данных MongoDB;

4. База данных:

4.1. MongoDB — документно-ориентированная система управления базами данных, не требующая описания схемы таблиц, использующая JSON-подобные документы и схему базы данных;

4.2. MongoDB Compass — графический интерфейс для администрирования и просмотра данных MongoDB;

Архитектура приложения

Клиентская часть

Структура клиентской части веб-приложения была спроектирована через предварительную декомпозицию проекта на атомарные функциональные компоненты. Компоненты были объединены в следующие логические группы:

- Основные компоненты, которые описывают главные панели, содержащие все функциональности клиентской части.
- Панель с тестами, которая содержит компоненты с внешним оформлением тестов и функциональности, реализующие их логику.
- Панель с профилем пользователя, которая содержит компоненты с оформлением профиля и функциональности, отвечающие за показ результатов тестирований.
- Панель с настройками, которая содержит компоненты с внешним оформлением настроек и функциональности, отвечающие за их изменение внутри приложения.

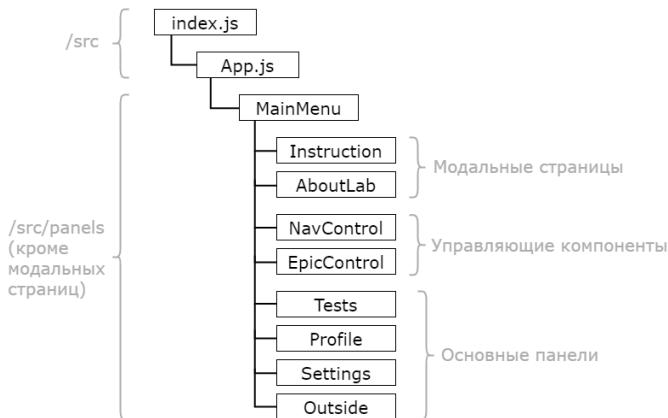


Рисунок 1: Иерархия основных компонентов

Серверная часть

Архитектура серверной части представлена в виде API, позволяющего взаимодействовать на клиенте с базой данных. Функции API были объединены в следующие логические группы:

- tests – функции, отвечающие за взаимодействие с тестами (например, получение конкретного теста для его прохождения).
- users – функции, отвечающие за взаимодействие с данными пользователя (например, инициализация стандартных пользовательских настроек при первом входе в приложение).
- passings – функции, отвечающие за прохождения тестов (например, сохранения ответов пользователя).
- results – функции, отвечающие за формирование и сохранение результатов тестирований (например, сохранение результатов пользователя по завершению прохождения).

База данных

Для хранения и управления данными в приложении используется документно-ориентированная NoSQL СУБД MongoDB. Для приложения были спроектированы следующие три коллекции, содержащиеся в БД:

- Tests – содержит информацию обо всех имеющихся тестах, каждый документ в ней представляет собой описание

конкретного теста.

- Users – содержит информацию обо всех пользователях, каждый документ представляет собой описание конкретного пользователя.
- Passings – содержит информацию о прохождениях тестов, каждый документ представляет собой описание конкретного прохождения.

Заключение

В статье было представлено описание веб-приложения, являющегося полноценным, масштабируемым инструментом для автоматизации сбора данных пользователей социальной сети ВКонтакте с целью последующей оценки связи между публикуемым пользователями контентом и их личностными особенностями. Факты, установленные с помощью этих данных, можно будет применить во множестве областей. Например, для более полного изучения и моделирования распространения социоинженерных атак и иных аспектов информационной безопасности, связанных с человеческим фактором, а также в социологии, психологии, маркетинговых исследованиях и т.д. [8–11].

Литература

1. Forbes: «Record 3.8 Trillion Hours Spent On Mobile Apps During 2021 In Another Blockbuster Year For Digital Economy». <https://www.forbes.com/sites/roberthart/2022/01/12/record-38-trillion-hours-spent-on-mobile-apps-during-2021-in-another-blockbuster-year-for-digital-economy/?sh=76ca663b2a42> [дата просмотра: 18.04.2022].
2. Nguyen, Minh Hao; Gruber, Jonathan; Fuchs, Jaelle (2020). Changes in digital communication during the COVID-19 global pandemic: implications for digital inequality and future research. Social Media and Society.
3. We Are Social: «DIGITAL 2022: ANOTHER YEAR OF BUMPER GROWTH». <https://wearesocial.com/uk/blog/2022/01/digital-2022-another-year-of-bumper-growth-2/> [дата просмотра: 18.04.2022].
4. Open Access Government «Social media data for social and behavioural research». <https://www.openaccessgovernment.org/social-media-data/113292/> [дата просмотра: 18.04.2022].
5. Cerruto, F., Cirillo, S., Desiato, D. et al. Social network data analysis to highlight privacy threats in sharing data. J Big Data 9, 19 (2022).
6. Мини-приложение ВКонтакте «Психологические тесты». <https://vk.com/app7794698> [дата просмотра: 19.04.2022].
7. Эрик Редмонд, Джим. Р. Уилсон. Семь баз данных за семь недель.

- Введение в современные базы данных и идеологию NoSQL = MongoDB in Action. — ДМК Пресс, 2013. — 384 с.
8. Тулупьева Т.В., Абрамов М.В., Тулупьев А.Л. Социоинженерные атаки: социальные сети и оценки защищенности пользователей // Монография.: Санкт-Петербург, 2018.
 9. Корепанова А.А., Олисеенко В.Д., Абрамов М.В. Применимость коэффициентов сходства в задаче сравнения социального окружения // Международная конференция по мягким вычислениям и измерениям. 2020. Т. 1. С. 39-42.
 10. Корепанова А.А., Тулупьева Т.В. Идентификация аккаунтов пользователя в различных социальных сетях по социальному окружению // В сборнике: ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ РЕГИОНОВ РОССИИ (ИБРР-2019). материалы конференции. 2019. С. 442-443.
 11. Арамов М.В., Азаров А.А., Фильченков А.А. Распространение социоинженерной атаки злоумышленника на пользователей информационной системы, представленных в виде графа социальных связей // Международная конференция по мягким вычислениям и измерениям. 2015. Т. 1. С. 329-331.

ИСТОЧНИКИ НЕОПРЕДЕЛЕННОСТИ И СПОСОБЫ ИХ ОБРАБОТКИ В ЗАДАЧЕ ОЦЕНКИ СВОДНЫХ ЧИСЛОВЫХ ХАРАКТЕРИСТИК ПОВЕДЕНИЯ ПО ДАННЫМ ИЗ САМООТЧЕТОВ ИНДИВИДОВ О ЕГО ПОСЛЕДНИХ ЭПИЗОДАХ¹

Столярова В.Ф., м.н.с. лаборатории теоретических и
междисциплинарных проблем информатики СПб ФИЦ РАН,
svf@dscs.pro

Аннотация

Работа посвящена задаче учета неопределенности ситуации сбора самоотчетов при оценивании кумулятивных характеристик эпизодического поведения индивидов. В работе представлена классификация типов неопределенности, которые возникают в ситуации сбора и анализа данных самоотчетов. Классическая модель эпизодического поведения представляет собой случайный точечный процесс. В работе описаны как классический подход к учету неопределенности в оценке кумулятивных характеристик поведения посредством регрессионного анализа, так и байесовский подход на основе байесовских сетей доверия, которые позволяют строить гибкие модели с учетом особенностей каждой ситуации сбора самоотчетов.

Введение

Сбор и анализ данных о поведении человека является важной частью решения задач в социоориентированных областях знаний. Такие данные используются для описания популяции или выборки, с которой работает специалист, но и для построения прогностических моделей и систем поддержки принятия решений. Подобные системы часто направлены на оценку риска, который связан с исследуемым поведением, и на последующий анализ экономической составляющей различных вариантов развития событий. Например, при решении задачи оценки уровня защищенности пользователей от социоинженерных атакующих

¹ Работа выполнена в рамках проекта по государственному заданию СПб ФИЦ РАН СПИИРАН № FFZF-2022-0003.

воздействий может быть важно выявить степень приверженности индивида рискованным действиям, связанным с информационной системой, как предоставление пароля третьим лицам[1, 2, 3]. Другим примером является исследование обыденного поведения индивидов, которые может быть связано с риском развития хронических заболеваний, как физическая активность [4].

Для извлечения информации о подобном поведении используются как объективные методы, включающие различные мониторы, анализы или прямое наблюдение [4], или же косвенные методы, которые описаются на различные опросы и интервью. Если первый класс методов является ресурсозатратным и часто нереализуемым при исследовании девиантного поведения, то одномоментные самоотчеты о поведении часто недооценивают истинные показатели, в отличие от дневниковых методов, анкет из нескольких пунктов или сбора данных в режиме реального времени[6]. Золотым стандартом при анкетировании или интервьюировании респондентов об их поведении является календарные методы [7, 8]. С развитием индивидуальных информационных технологий, а также доступности интернета, появился еще один способ сбора информации о повседневном поведении человека: EMA (ecological momentary assessment) [5], который подразумевает внесение индивидом данных об эпизодах своего поведения в момент их реализации посредством технических средств. Такой подход комбинирует в себе черты классического опроса и прямого наблюдения, при этом позволяя собирать дополнительную информацию о внешних факторах, сопутствующих поведению.

Однако поведение индивида является многогранным объектом, включающим множество различных факторов и характеристик. При этом при сборе самоотчетов исследователь сталкивается с рядом неопределенностей, связанных с когнитивными особенностями респондента. Поэтому при построении оценок кумулятивных характеристик эпизодического рискованного поведения возникает задача выявления, а также моделирования возникающей неопределенности. Целью исследования является классификация источников неопределенности в ситуации сбора самоотчетов индивидов об эпизодическом поведении, а также способы обработки неопределенности при построении оценок кумулятивных характеристик такого поведения. Новизна исследования в подходе (на основе вероятностных графических моделей) к учету неопределенности. Теоретическая значимость состоит в системном подходе к анализу рисков в ситуации сбора самоотчетов с целью построения оценок характеристик поведения, соответствующие вероятностные графи-



Рис. 1: Математическая постановка задачи оценки кумулятивной характеристики поведения по данным о нескольких последних эпизодах

ческие модели составляют практическую значимость.

Математическая модель эпизодического поведения индивидов

В работах [9, 10, ?] была предложена и разработана математическая модель эпизодического поведения индивида, которая легла в основу метода построения оценки кумулятивных характеристик такого поведения по неполным и неточным данным об эпизодах. Для построения оценки в модели используются данные о нескольких последовательных эпизодах поведения. Математическим ядром для этой модели служит теория анализа повторяющихся событий [12]. На рисунке 1 представлена схематическое изображение предложенного подхода.

Предположим, что исследуется какое-то эпизодическое поведение. Пусть для каждого индивида i из выборки n извлекается информация о нескольких последовательных эпизодах поведения с помощью некоторого метода $0 < t_{m_i}^i < \dots < t_2^i < t_1^i < t_0^i < I$ (I - это момент опроса или сбора данных). Это может быть как метод ЕМА или же стандартное интервью. Обозначим как $N^i(t)$ количество эпизодов в интервале $(0, t]$, а как $H(t) = \{N(s), 0 \leq s \leq t\}$ — историю точечного процесса до времени $t > 0$. Такой точечный процесс может задаваться посредством функции интенсивности, которая отражает мгновенную вероятность реализации эпизода.

$$\lambda(t|H(t)) = \lim_{\Delta t \rightarrow 0} \frac{Pr\{\Delta N(t) = 1 | H(t)\}}{\Delta t}.$$

Такая формализация процесса эпизодического поведения позволяет использовать регрессионный анализ для оценки интенсивности процесса,

которая в свою очередь служит основой для вычисления различных кумулятивных характеристик процесса реализации эпизодов [12].

Однако для практических приложений была предложена модель оценивания искомых характеристик процесса на основе байесовской сети доверия[13, 14]. Такой подход является достаточно гибким, позволяющим моделировать неопределенность из различных источников.

Классификация видов неопределенности и подходы к их обработке

В целом ситуация сбора самоотчетов позволяет выявить следующие источники неопределенности:

1. Неопределенность внешних факторов, возникающая в силу возможной неизвестности факторов, оказывающих влияние на эпизоды поведения, которые могут быть взаимозависимыми.
2. Неопределенность, связанная с неоднородностью поведения в популяции: различные респонденты имеют различную склонность к поведению, а также могут предоставлять информацию о различном числе эпизодов.
3. Неопределенность информации о дате эпизода поведения, которая может поступатьискаженной в силу особенностей мышления респондента.

Далее приведен сравнительный анализ двух подходов к построению оценки кумулятивных характеристик поведения с точки зрения учета выявленных неопределенностей.

Классический подход к анализу повторяющихся событий позволяет учесть неопределенность, связанную с неоднородностью выборки, путем включения случайного множителя u_i , отражающего индивидуальную склонность к поведению, в функции интенсивности процесса. В этом случае функция интенсивности будет иметь вид:

$$\lambda(t|u_i) = u_i \rho_i(t),$$

где $\rho_i(t)$ — функция, отражающая зависимость от времени и фиксированных ковариат, влияющих на процесс генерации эпизода.

Наиболее распространенной и простой с вычислительной точки зрения является выбор гамма-распределения для u_i . Кроме того, число

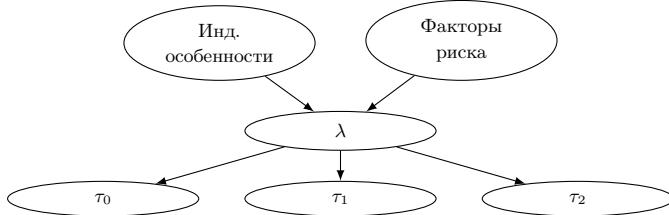


Рис. 2: Расширенная для учета возникающих типов неопределенности структура байесовской сети доверия

эпизодов, наблюдаемое для каждого индивида, может быть различным при квантификации функции интенсивность по статистическим данным посредством регрессии (регрессии Кокса [12]). Такая модель широко используется в различных приложениях [12], включая моделирование поведения человека [9, 15]. Внешние факторы, влияющие на поведение, могут быть включены в модель только в форме регрессии [11], что ограничивает возможности моделирования первого типа неопределенности, связанной со структурой внешних факторов. При этом отметим, что учет неточности в самих самоотчетах, т.е. неточности данных о времени реализации эпизода, с помощью регрессионной модели практически невозможен.

Для учета перечисленных типов неопределенности байесовская сеть доверия, предложенная в работе [13], может быть дополнена соответствующими узлами (см. рисунок 2). В ситуации малого числа наблюдений, байесовские сети доверия позволяют использовать экспертную информацию как для задания структуры сети, что может быть полезно при обработке неопределенности внешних факторов и их структуры зависимости, так и для квантификации модели. Обратим внимание также, что существуют подходы к учету в байесовских сетях доверия неопределенности самих данных (тип 3) путем добавления скрытых переменных

Существуют подходы к численному заданию байесовских сетей доверия, которые позволяют использовать непрерывные случаный элементы для байесовского вывода напрямую, минуя этап дискретизации [16].

Заключение

В работе представлена классификация типов неопределенности, которые возникают в ситуации сбора и анализа данных самоотчетов. Классическая модель эпизодического поведения представляет собой случайный точечный процесс. В работе описан классический подход к учету неопределенности в оценке кумулятивных характеристик поведения посредством регрессионного анализа. Однако область применимости этого метода ограничена в силу жесткости задания регрессионной модели. Более гибким подходом оказываются байесовские сети доверия. В работе предложена структура байесовской сети доверия для оценивания кумулятивных характеристик поведения в рамках гамма-пуассоновской модели эпизодического поведения индивидов с учетом возникающих источников неопределенности.

Список литературы

- [1] Абрамов М.В., Тулупьева Т.В., Тулупьев А.Л. Социоинженерные атаки: социальные сети и оценки защищенности пользователей. СПб.: ГУАП, 2018. 266 с.
- [2] Khlobystova A., Abramov M. The models separation of access rights of users to critical documents of information system as factor of reduce impact of successful social engineering attacks //CEUR Workshop Proceedings. 2020. Т 2782. С. 264-268.
- [3] Abramov M. V., Tulupyev A. L. Soft Estimates of User Protection from Social Engineering Attacks // Conference on Artificial Intelligence and Natural Language. – Springer, Cham, 2019. – С. 47-58.
- [4] Prince S.A., Adamo K.B., Hamel M.E., Hardt J., Gorber S.C., Tremblay MA comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review //International journal of behavioral nutrition and physical activity. – 2008. – Т. 5. №. 1. С. 1-24.
- [5] Shiffman S., Stone A. A., Hufford M. R. Ecological momentary assessment //Annu. Rev. Clin. Psychol. – 2008. – Т. 4. – С. 1-32.
- [6] Prince S. A., Cardilli L., Reed J. L., Saunders T. J., Kite C., Douillette K., Fournier K., Buckley J. P. A comparison of self-reported and

- device measured sedentary behaviour in adults: a systematic review and meta-analysis //International Journal of Behavioral Nutrition and Physical Activity. – 2020. – Т. 17. – №. 1. – С. 1-17.
- [7] Sobell L. C., Sobell M. B. Timeline follow-back //Measuring alcohol consumption. – Humana Press, Totowa, NJ, 1992. – С. 41-72.
- [8] Liu W., Li R., Zimmerman M. A., Walton M. A., Cunningham R. M., Buu A. Liu W. et al. Statistical methods for evaluating the correlation between timeline follow-back data and daily process data with applications to research on alcohol and marijuana use // Addictive behaviors. – 2019. – Т. 94. – С. 147-155.
- [9] Пашченко А. Е., Тулупьев А. Л., Николенко С. И. Моделирование заражения ВИЧ-инфекцией на основе данных о последних эпизодах рискованного поведения //Известия высших учебных заведений. Приборостроение. – 2006. – Т. 49. – №. 11. – С. 33-34.
- [10] Tulupiyev A., Suvorova A., Sousa J., Zelterman D. Beta prime regression with application to risky behavior frequency screening //Statistics in medicine. – 2013. – Т. 32. – №. 23. – С. 4044-4056.
- [11] Stolarova V., Tulupiyev A. L. Cox regression in the problem of risky behavior parameter estimation based on the last episodes' data //St. Petersburg State Polytechnical University Journal. Physics and Mathematics. – 2021. – Т. 14. – №. 4. – С. 202.
- [12] Cook R. J. et al. The statistical analysis of recurrent events. – New York : Springer, 2007. – С. 128-133.
- [13] Суворова А. В., Тулупьев А. Л., Сироткин А. В. Байесовские сети доверия в задачах оценивания интенсивности рискованного поведения //Нечеткие системы и мягкие вычисления. – 2014. – Т. 9. – №. 2. – С. 115-129.
- [14] Тулупьев А. Л., Николенко С. И., Сироткин А. В. Основы теории байесовских сетей доверия. – 2019. 399 стр.
- [15] Lin T. H., Tsai M. H. Solving unobserved heterogeneity with latent class inflated Poisson regression model //Journal of Applied Statistics. – 2022. – Т. 49. – №. 11. – С. 2953-2963.
- [16] Stolarova V., Tulupiyev A. Probabilistic Graphical Models with Continuous Variables for the Decision Making About Risky Episodic

Behavior in the Framework of Gamma Poisson Model with Application to Public Posting Data //International Conference on Intelligent Information Technologies for Industry. – Springer, Cham, 2023. – C. 465-474.

АГРЕГАЦИЯ И АНАЛИЗ СВЕДЕНИЙ ЛОГИСТИЧЕСКИХ КОМПАНИЙ ДЛЯ ПОСТРОЕНИЯ СЛОЖНОГО МАРШРУТА ПЕРЕВОЗКИ ГРУЗА

Сабреков А. А., м.н.с. лаб. ТиМПИ СПб ФИЦ РАН, aas@dscs.pro

Есин М. С., студент СПбГУ, mse@dscs.pro

Аннотация

В статье рассматривается концепция калькуляторов по расчету стоимости доставки грузов и их применение в агрегаторах предложений логистических компаний. Исследуются возможности по улучшению итоговой стоимости полученных предложений и увеличению покрытия обслуживаемых городов путем построения мультимодальных маршрутов. Проводится обзор динамических и графовых алгоритмов по построению оптимальных многозвездных путей, а также существенные препятствия к их применению. Описываются эвристические алгоритмы построения сложного маршрута, основанные на анализе логистических сетей России. Также аргументируется избыточность и невыгодность маршрутов с четырьмя и более звенями.

Введение

Логистика не только играет большую роль в развитии экономики, но и прочно входит в повседневную жизнь: грузоперевозками пользуются не только крупные промышленные компании, маркетплейсы и интернет-магазины, но и обычные люди, например при переезде или доставке важных бумаг.

Для повышения лояльности клиентов логистические компании создают специальные калькуляторы по расчету стоимости доставки груза: с их помощью клиент может самостоятельно рассчитать примерную стоимость перевозки своего груза, не разбираясь в сложных тарифах.

Данные калькуляторы по параметрам груза, начальному и конечному пункту вычисляют примерную стоимость и время доставки, за

которые компания готова перевезти груз. Подробная концепция калькуляторов была описана в [1].

Агрегатор калькуляторов на сайте Cargotime.ru, консолидирующий данные от разных компаний, способен выбрать самое выгодное с точки зрения стоимости предложение. Однако, прямой маршрут не всегда самый выгодны, поэтому возникает задача построения мультимодального маршрута, исполняемого несколькими перевозчиками, и проходящего через промежуточные пункты, в которых осуществляется перевалка (передача груза).

Задача построения сложных маршрутов уже решалась в рамках улучшения механизма принятия логистических решений для крупной китайской транспортной 4PL-компании. В статье [2] для построения оптимальных маршрутов используется алгоритм Дейкстры, а также применяются нечеткие комплексные оценки стоимости и сроков доставки. Помимо графового подхода к построению сложного маршрута, существует ничем не уступающий ему в оптимальности результата динамический подход, как, например, в статье [3].

В решениях выше для построения используется готовый набор данных о предложениях перевозчиков, образующий насыщенный взвешенный граф. В контексте работы с калькуляторами главным препятствием к применению графовых и динамических алгоритмов является невозможность оперативного получения подобного графа из-за низкой пропускной способности сайтов компаний.

В масштабах статьи планируется рассмотреть подходы к решению задачи построения сложного маршрута, тем самым расширив возможности агрегатора по расчету стоимости доставки.

Построение сложных маршрутов

Анализ логистических сетей

Ввиду затрудненности использования стандартных графовых и динамических оптимизаций, возникает необходимость создать некоторый эвристический алгоритм, который опирается в основном на логические предположения об оптимальности, нежели на полный набор данных.

Чтобы принять во внимание слабые стороны существующей транспортной системы, приведенные в литературном обзоре, в рамках решения задачи построения сложных маршрутов для начала стоит понять, какие города чаще всего входят в транспортные сети перевозчиков, чтобы эмулировать их механизмы принятия логистических решений.

Степень вовлеченности города в цепи поставок непрерывно связана с количеством грузовых терминалов логистических компаний. В предположении, что объем складских помещений в городе коррелирует с количеством грузовых терминалов, можно выделить порядка 18 городов России, объем складских помещений которых больше 150 тыс. кубометров [4].

Стоит принять во внимание главные транзитные железнодорожные линии России: преимущество городов, лежащих вдоль этих путей, в том, что перевозка грузов по железным дорогам – самый дешевый из способов транспортировки грузов.

Также стоит отобрать крупные города Сибири и Дальнего Востока, не подходящие ни под один из критериев выше, потому что в локальных областях они являются единственными крупными как по населению и площаади, так и по количеству терминалов компаний.

Стратегии построения сложных маршрутов

Назовем улучшением маршрута замену одного из звеньев в маршруте на пару смежных звеньев, которые соединяют начало и конец исходного звена, причем их суммарная стоимость меньше стоимости исходного звена. Тогда сложный маршрут с любым количеством звеньев можно получить композицией улучшений.

При каждом улучшении количество звеньев в маршруте будет только расти, что может оказаться на практической пользе таких маршрутов: увеличиваются расходы на транспортировку между компаниями, при этом повышается риск потери или порчи груза при передаче. Поэтому в реальных условиях выгодными нужно считать только двух- и трехзвенные маршруты.

Первое рассматриваемое улучшение заменяет звено на пару звеньев, которые соединяются в крупном городе, ближайшем к пункту отправки или доставки. Данное улучшение опирается на предположение о том, что компании, специализирующиеся на перевозках в пределах пары федеральных округов, должны:

- доставлять грузы до локальных населенных пунктов, которые вообще не обслуживаются остальными компаниями
- предлагать более низкие цены по сравнению с остальными компаниями

В таком случае, часть маршрута будет пройдена по тарифам местных компаний, что, возможно, будет дешевле.

Другое улучшение ориентируется на минимизацию пройденного расстояния: чем меньше километраж – тем дешевле. В качестве промежуточного города выбирается крупный город с минимальным суммарным расстоянием до концов исходного звена.

Главным вектором дальнейшего развития проекта является оценка сопутствующих расходов на транспортировку между компаниями и временное хранение, потому что эти расходы являются самым неизученным аспектом сложных маршрутов, так как калькуляторы компаний при расчете никак не учитывают то, что впоследствии будет построен сложный маршрут.

Заключение

В статье была описана концепция построения сложных грузовых маршрутов в рамках развития сервиса агрегации данных о стоимости и сроках перевозок и протестирован тестовый механизм. Были обозначены основные векторы развития проекта с целью стать уникальным сервисом на российском рынке логистики.

Список литературы

- [1] Абрамов М. В., Есин М. С. Агрегация сведений и оценка параметров грузовых маршрутов в условиях информационного дефицита – Материалы конференции «ИБРР-2021». С. 328–330. – 2021.
- [2] Wei-Qiong F. Fourth-party logistics optimization decision-making Based on graph model with multi-dimensions. – 2014.
- [3] Zavalishchin D. S., Popkov A. Dynamic programming method for optimization problem of multi-modal transportation. – 2017.
- [4] Рынок складской недвижимости России в 2018 году, обзорная статья агентства Knight Frank Research – 2019.

Моделирование управления общественными процессами посредством массовых рассылок

Зубов С. В., ст. преп. кафедры диагностики функциональных систем
СПбГУ, s.zybov@spbu.ru

Аннотация

Рассматривается один из возможных способов управления большими социальными системами при использовании массового распространения специальной информации через электронную почту, через социальные сети и через другие системы массовой коммуникации. В качестве аналога приводится одна задача из математической теории устойчивости.

Введение

Толчком к проведению настоящего исследования было осмысление автором социальных процессов, связанных с так называемыми "цветными революциями" в странах Ближнего Востока и с миграционными процессами в Европе. Несомненно, что эти события были запланированы и в значительной степени реализованы при использовании информационных средств массового воздействия на целевые социальные группы. То есть говорить о том, что события происходили сами собой, были спонтанными не приходится. Если внешне события и выглядели хаотическими, то, бесспорно, это был "управляемый хаос". Кавычки здесь поставлены потому, что существует даже научная теория управления социальными системами посредством организации событий, внешне выглядящими как случайные (хаотические). Есть данные [1-3] о том, что в настоящее время такие цифровые гиганты, как Google, Facebook и им подобные являются "глобальными квазиорисдикциями", осуществляя некоторые функции глобального управления, которые ранее были присущи только государствам. Возможно также появление на мировой арене глобальных игроков какого-либо иного типа.

Постановка задачи и основные подходы к ее решению

Мы будем рассматривать проблему в следующей постановке. Пусть в некоторой социальной системе структура управления иерархическая, т.е. принятие управленческого решения на каком-либо уровне управления влечет за собой безоговорочное принятие этого решения на более низких уровнях управления. Еще этот принцип называют централизмом. Пусть имеется конкретное лицо, принимающее решения (ЛПР), которое подлежит

нашему управлению. То есть мы хотим каким-либо образом воздействовать на ЛПР с той целью, чтобы ЛПР приняло нужное нам управленческое решение. Естественно, рассматривается случай, когда ЛПР принуждается к нужному решению, а не идет на него добровольно (иначе не было бы и проблемы). Как говорят практические психологи, переубедить оппонента малореально, но выиграть бой можно. Таким образом, предлагается рассмотреть своего рода информационное сражение с вышеуказанным ЛПР.

Существуют три уровня воздействия на ЛПР: 1) личное воздействие, т.е. воздействие на личностные ценности ЛПР, а также работа с его внутренним миром;

1) личное воздействие, т.е. воздействие на личностные ценности ЛПР, а также работа с его внутренним миром;

2) воздействие на ближайшее окружение ЛПР, на его референтную группу;

3) воздействие на окружение ЛПР в широком смысле слова, т.е. на лиц, группы или слои населения, непосредственно не связанные с ЛПР.

Первый уровень относится к личностным установкам ЛПР, его жизненным целям как руководителя. Если известны слабые стороны личности ЛПР (а они есть почти у каждого), то при воздействии на эти слабые стороны можно сразу же достичь цели.

Простой пример такого воздействия. В случае, когда ЛПР считает себя великим, хочет достичь величия посредством своих управленческих решений или же намерен при их помощи оставить свой след в истории - тогда для воздействия на ЛПР достаточно простой личной беседы (или личного письма), где осуществляется давление на указанные "болевые точки".

В строгом смысле слова массовое информационное воздействие на ЛПР осуществляется на втором и третьем уровнях. На втором уровне (наиболее эффективном по воздействию на ЛПР) возникают две задачи - техническая и идеологическая. Технической задачей является поиск точек доступа к референтной группе ЛПР - найти адреса электронной почты членов этой группы, их аккаунты в социальных сетях, позволяющие написать им сообщения. Идеологической задачей является собственно определение состава этой референтной группы, а также структуры отношений между ЛПР и членами этой группы. Именно от структуры отношений между ЛПР и членами его референтной группы зависит содержание писем и сообщений, отправляемых нами этим членам с той целью, чтобы они включили свои каналы воздействия на ЛПР и реально воздействовали на него в наших интересах. Указанная идеологическая задача даже в простых случаях может быть неимоверно сложна для

неподготовленного исследователя. Однако опыт показывает, что с годами нарабатываются навыки и способности к решению подобного рода задач, которые, по сути дела, являются своего рода искусством, а овладевший этим искусством специалист может быть признан как обладающий талантом в указанной сфере деятельности.

На третьем уровне воздействия на ЛПР вышеописанные техническая и идеологическая задачи также возникают, но их содержание является более простым по сравнению со вторым уровнем. Здесь дело определяется количеством адресатов для рассылки соответствующей информации - их может быть на несколько порядков больше, чем членов референтной группы ЛПР на втором уровне.

Простой практический пример. Рассылка осуществляется рядовым работникам предприятия, группы предприятий, холдинга, министерства и т.п., где работает интересующее нас ЛПР, а число адресатов в этом случае может исчисляться десятками тысяч. Здесь техническая задача дополняется проблемой организации именно массовых рассылок, а также, возможно, создания специальных аккаунтов в социальных сетях и специализированных сайтов.

Идеологическая же задача сильно упрощается по сравнению со вторым уровнем, т.к. для наших целей уже не так важна структура отношений между адресатами наших рассылок и ЛПР (может быть, что до получения нашей информации эти адресаты даже и не знали о самом существовании указанного ЛПР). Таким образом, третий уровень по своей организации аналогичен обычным СМИ – газетам, радио, телевидению.

Резюмируя вышеизложенное, мы приходим к выводу, что на всех трех рассмотренных уровнях информационного воздействия на ЛПР принцип наших действий является общим - мы привлекаем внешние ресурсы для достижения цели. На первом уровне этими ресурсами являются вышеуказанные свойства личности самого ЛПР, которые могут быть привлечены для достижения наших целей. На втором и третьем уровне этими ресурсами являются соответственно референтная группа ЛПР и (более или менее) широкие слои общественности. Как тут не вспомнить популярное в свое время высказывание: "Идеи становятся материальной силой, когда они овладевают массами", принадлежащее известным политическим деятелям прошлого. Однако, многое зависит от того, кого или что нужно иметь ввиду под понятием "массы". В нашем случае - это вышеуказанные внешние по отношению к ЛПР ресурсы.

Аналогия с задачами теории устойчивости

Рассмотрим автономное дифференциальное уравнение первого порядка с малым параметром при производной следующего вида:

$$\varepsilon \frac{dx}{dt} = f(x), t \geq 0, \quad (1)$$

где $\varepsilon > 0$ - малый параметр, x - скалярная переменная, а скалярная функция $f(x)$ задана для всех $x, a \leq x \leq b$, непрерывна по x для всех $x, a \leq x \leq b$, и удовлетворяет там условию Липшица по аргументу x , здесь a, b - константы. Малый параметр ε в уравнении (1) моделирует наличие как раз той "массы", о которой шла речь во второй части статьи. В самом деле, стоит поделить в (1) обе части на ε , как становится ясно, что в правой части возникает "большая" величина, связанная с наличием множителя $1/\varepsilon$. А это как раз и соответствует многократному усилению воздействия на ЛПР, о чём и шла речь выше. Предположим, что существует такое значение x_0 , $a < x_0 < b$, что $f(x_0) = 0$, т.е. x_0 является корнем функции $f(x)$. Тогда в уравнении (1) имеется стационарное решение $x(t, x_0) \equiv x_0, t \geq 0$ (точка покоя). Указанная величина x_0 может быть поставлена в соответствие желательному для нас управлению решению ЛПР, о котором шла речь во второй части статьи. Все остальные значения аргумента x будут соответствовать каким-то другим управлением решениям ЛПР, нежелательным для нас. Здесь и далее через $x(t, x_1)$ обозначено решение уравнения (1), удовлетворяющее начальному условию $x(0, x_1) = x_1$, где x_1 - любая величина, подчиняющаяся неравенствам $a < x_1 < b$. Упомянутое стационарное решение асимптотически устойчиво по Ляпунову, если функция $f(x)$ при переходе через корень x_0 меняет знак с плюса на минус. Иными словами, все другие близкие решения $x(t, \dot{x}_0)$, начинающиеся в некоторой окрестности точки x_0 , асимптотически приближаются к этому решению $x(t, x_0) \equiv x_0$ при неограниченном возрастании времени, т.е. $x(t, \dot{x}_0) \rightarrow x_0$ при $t \rightarrow +\infty$. Указанная окрестность (область притяжения решения $x(t, x_0) \equiv x_0$) включает в себя все те значения x , для которых выполняются условия $f(x) > 0, a < x < x_0$, а также условия $f(x) < 0, x_0 < x < b$. То есть указанные значения образуют область асимптотической устойчивости решения $x(t, x_0) \equiv x_0$. Заметим, что здесь имеется ввиду односвязная область.

Замечание. Поскольку дифференциальное уравнение (1) автономно (правая

часть не зависит от аргумента t), можно перейти к новому времени τ по формуле $\tau = t/\varepsilon$. Получаем дифференциальное уравнение

$$\frac{dx}{d\tau} = f(x), \tau \geq 0, \quad (2)$$

где τ - новое «быстрое» время, и здесь коэффициент $1/\varepsilon$ моделирует уже длительность воздействия на ЛПР вышеуказанных внешних факторов. То есть воздействующие на ЛПР внешние факторы могут быть не интенсивными, но длительными, и эффект их воздействия на ЛПР в соответствии с рассматриваемой моделью будет такой же.

Заключение

В докладе предлагается подход к управлению социальными процессами в современном обществе. Предлагается разнообразное использование массовых рассылок (e-mail, социальные сети).

Список литературы

1. Буров В. Трансформация глобального управления / Буров В. // Материалы научной интернет-конференции по проблемам глобальных вызовов XXI века «Футуратон 2017». — 2017. <https://www.youtube.com/watch?v=ioBUTtot4Yk> [дата просмотра: 25.04.2022], <https://www.facebook.com/futurathon/> [дата просмотра: 25.04.2022].
2. <https://www.youtube.com/c/tcehlive/featured> [дата просмотра: 25.04.2022].
3. https://t.me/tceh_com [дата просмотра: 25.04.2022].

Проверка ацикличности алгебраической байесовской сети с применением третичной структуры

Вяткин А. А., студент математико-механического факультета СПбГУ,
vyatkin.artex@gmail.com

Харитонов Н. А., аспирант СПбГУ, nak@dscs.pro

Тулупьев А. Л., д.ф.-м.н., профессор кафедры информатики СПбГУ,
 главный научный сотрудник лаб. ТиМПИ СПб ФИЦ РАН, alt@dscs.pro

Аннотация

В теории алгебраических байесовских сетей (АБС) существует понятие ацикличности алгебраической байесовской сети. Это свойство имеет место быть, когда АБС представима в виде дерева смежности. В данной работе приведен способ проверки непротиворечивости АБС с применением третичной структуры, что может быть полезно при независимом использовании третичной структуры, например, при глобальном апостериорном выводе.

Введение

Зачастую необходимо опираться на знания, собранные в данной конкретной области, для того, чтобы составлять планы будущей деятельности, находить решения, размышлять в терминах этой предметной области. Подобные знания большей своей частью заключают в себе неопределенность, анализ и обработка которых является одной из сторон современной информатики [6]. Такие знания могут рассматриваться в виде совокупности утверждений,ключающих между собой логические и стохастические связи. Данные соотношения между утверждениями способны охарактеризовать эксперты рассматриваемой области. В то же время суждения экспертов характеризуют связи между небольшим набором сущностей из предметной области. В итоге комплекс знаний экспертов разделяется на отдельные компоненты, фрагменты знаний, которые в общей сложности образуют базу фрагментов знаний [5].

Для описания подобной экспертной системы необходима математическая модель, в качестве которой могут выступать вероятностные графические модели (ВГМ). Вероятностные графические модели находят свое применение в разных задачах, таких, как, например, исследование влияния окружающей среды и генов на получаемые заболевания [3], анализ социоинженерных атак [2]. В класс вероятностных графических моделей также входят

алгебраические байесовские сети. В качестве математической модели фрагмента знаний в теории алгебраических байесовских сетей могут выступать идеалы конъюнктов, определенных над алфавитом из набора пропозициональных переменных. Каждому элементу идеала приписывается либо точечная, либо интервальная оценка вероятности истинности. В совокупности все фрагменты знаний составляют алгебраическую байесовскую сеть.

Алгебраическая байесовская сеть может рассматриваться с использованием различных типов структур. Первичная структура представляет из себя обычное множество фрагментов знаний, без указания связей между ними. Вторичная структура, как и третичная, представляются в виде специального рода графов, использующих фрагменты знаний и их пересечения в качестве нагрузок ребер и вершин. Вторичная структура является графом смежности, при этом для одной заданной первичной структуры может быть несколько различных вторичных структур. Важным понятием является непротиворечивость алгебраической байесовской сети — согласованность оценок вероятности истинности элементов фрагментов знаний. Еще одним важным свойством первичной структуры и алгебраической байесовской сети в целом служит их ацикличность — возможность построить вторичную структуру, являющуюся деревом смежности. Наличие ацикличности у алгебраической байесовской сети позволит уменьшить вычислительную сложность проверки и поддержания непротиворечивости. Также именно для ациклических байесовских сетей доказана корректность применения алгоритма глобального апостериорного вывода, основанного на распространении виртуального свидетельства и заключающегося в переоценке вероятности истинности элементов фрагментов знаний на основе новых поступивших знаний — свидетельств. При этом для пропагации свидетельств может быть использована только третичная структура, применяющаяся сейчас для построения вторичных структур, используемых в дальнейшем для распространения свидетельств. Таким образом, проверка ацикличности с применением третичной структуры будет полезна при ее обосновленном от вторичной структуры использовании и данная работа посвящена решению этого вопроса.

Теоретическая основа

В данной главе опишем систему терминов и ряд алгоритмов, используемых в работе и приведенных в [4, 7, 8].

Основные определения

Прежде всего рассмотрим объекты, которые будут соответствовать переменным, заключающим утверждения. Они образуют *алфавит* — множество, состоящее из атомарных пропозициональных формул (которые могут называться атомами). $A = \{x_1, \dots, x_n\}$ определяет алфавит из n атомов. Для оценки самих атомарных пропозиций, а также связей между ними определим *идеал конъюнктов*, построенный над алфавитом $A = \{x_1, \dots, x_n\}$ — множество формул вида $\{x_{i_1} x_{i_2} \dots x_{i_k} \mid 0 \leq i_1 < \dots < i_k \leq n - 1, k \leq n\}$, где $x_{i_1} x_{i_2} \dots x_{i_k}$ представляет конъюнкцию соответствующих переменных.

Фрагментом знаний (математической моделью фрагмента знаний), который построен над алфавитом A , назовем пару (C, p) , где C — идеал конъюнктов над соответствующим алфавитом, p — интервальные или скалярные (точечные) оценки вероятностей для каждого конъюнкта из идеала C .

Для дальнейшей работы с фрагментами знаний и их наборами удобно определить вес, который соответствует каждому ФЗ — нагрузкой или весом фрагмента знаний $W(C, p)$ назовем подалфавит алфавита, над которым задан фрагмент знаний, $W(C, p) = \{x_i \mid x_i \in C, x_i \in A\}$.

Назовем *набором максимальных фрагментов знаний* (набор МФЗ, первичная структура алгебраической байесовской сети) такой набор фрагментов знаний, что никакая нагрузка фрагмента знаний не содержится полностью в нагрузке другого фрагмента знаний из представленного набора. То есть $\forall i \neq j$ выполнено: $W(V_i) \not\subseteq W(V_j)$ и $W(V_j) \not\subseteq W(V_i)$.

Вторичная структура АБС

Сепаратором двух МФЗ, V_i и V_j , назовем подалфавит, который является пересечением нагрузок этих ФЗ: $W(V_i, V_j) = W(V_i) \cap W(V_j)$, $i \neq j$. Пара МФЗ называются сочлененными, если их сепаратор непуст.

Граф максимальных фрагментов знаний — ненаправленный граф, вершинам которого сопоставлены МФЗ, вошедшие в АБС и ребра возможны только между сочлененными ФЗ. *Нагрузкой* $W(\{V_i, V_j\})$ *ребра* $\{V_i, V_j\} \in E(G)$ графа G назовем сепаратор его концов: $W(\{V_i, V_j\}) = W(V_i) \cap W(V_j)$. Определим и *нагрузку* $W(H)$ *подграфа* $H \subseteq G$ — наибольший по включенияния подалфавит, входящий в нагрузку всех вершин подграфа: $W(H) = \bigcap_{V \in H} W(V)$.

Магистральный путь между сочлененными вершинами V_i и V_j — такой путь между этими вершинами, что нагрузка каждой вершины пути содержит сепаратор концов этого пути. Далее граф будет *магистрально связан*, если между каждой из его сочлененных вершин существует магистральный путь.

Граф смежности — магистрально связный граф МФЗ. *Дерево смежности* — граф смежности, представимый в виде дерева.

В результате, помимо первичной структуры АБС, можно дать определение *вторичной*. Такой структурой будет являться некоторый граф смежности АБС.

Так же существует понятие *максимального графа смежности* G_{max} — наибольшего по числу ребер графа смежности. Для заданного множества вершин существует единственный максимальный граф смежности, то есть тот, в котором между вершинами существует ребро только тогда, когда они сочлененные.

Дополнительно предположим, что первичная структура *связна*, то есть связан максимальный граф смежности, построенный над этой структурой. В противном случае можно рассматривать наборы вершин из каждой компоненты связности как отдельные АБС.

Третичная структура АБС

Сужением $G \downarrow U$ графа G на нагрузку U назовем граф, в который входят те и только те ребра и вершины исходного графа G , нагрузки которых равны или содержат U . *Значимое сужение* — сужение на нагрузку, которая является сепаратором для некоторой пары МФЗ. На сужение можно наложить дополнительные ограничения, тогда получим *сильное сужение* $G \downarrow U$ — сужение $G \downarrow U$, из которого удалили все ребра нагрузки U . После сильного сужения граф $G \downarrow U$ разбивается на компоненты связности, после сужения же $G \downarrow U$ граф остается связным.

Одним из основных объектов в новоопределенной структуре будет *значимая нагрузка* U — непустой сепаратор некоторой пары ФЗ первичной структуры. *Замкнутым же снизу множеством нагрузок* назовем объединение множества значимых нагрузок с множеством нагрузок вершин МФЗ. *Замкнутое множество нагрузок* — объединение замкнутого снизу множества нагрузок с одноэлементным множеством, содержащим пустое множество.

При этом на множестве нагрузок существует частичный порядок, являющийся отношением включения. Таким образом, *родительским графом* (третичной структурой АБС) назовем диаграмму Хассе замкнутого множества нагрузок. Диаграмму Хассе можно рассматривать как транзитивное сокращение, поэтому родительский граф единственный при заданной первичной структуре АБС [1].

Проверка ацикличности

Алгоритм проверки того, что вторичная структура представима в виде дерева смежности основывается на следующей теореме, описанной в [9]:

Теорема 1. *Связная первичная структура АБС циклична тогда и только тогда, когда не выполняется соотношение:*

$$|MKP| = \sum_{U \in Sep} Conn(G_{max} \downarrow U) - |Sep| + 1$$

где MKP — первичная структура АБС, набор $\Phi 3$, $Conn(G_{max} \downarrow U)$ — число компонент связности графа $G_{max} \downarrow U$, Sep — множество непустых сепараторов.

Все слагаемые из выражения данной теоремы можно подсчитать, используя только первичную и вторичную структуры АБС, при этом наибольшую сложность здесь представляет расчет числа компонент связности, количество сепараторов же равно количеству вершин родительского графа, за исключением верхней, пустой вершины, и листьев-фрагментов знаний. Поэтому далее рассмотрим, как можно подсчитать сумму компонент связности графов сильных сужений с использованием третичной структуры. Для этого докажем теорему:

Теорема 2. *Две вершины с нагрузками в виде фрагментов знаний k_{p1} и k_{p2} лежат в одной компоненте связности C_u графа $G_{max} \downarrow u$, полученной после сильного сужения на значимую нагрузку u , тогда и только тогда, когда существует последовательность таких вершин с нагрузками в виде фрагментов знаний и с k_{p1} и k_{p2} как крайними элементами, что для каждого из двух соседних вершин в этой последовательности существует нагрузка, которая является предком по отношению к этим вершинам и потомком по отношению к вершине с нагрузкой u в родительском графе. Формально:*

$$k_{p1}, k_{p2} \in C_u \Leftrightarrow \exists v_1 = k_{p1}, \dots, v_n = k_{p2} :$$

$$\forall i = 1, \dots, n-1 : (\exists w : v_i, v_{i+1} \in \text{descendants}(w) \& w \in \text{descendants}(u))$$

Доказательство. Действительно, если k_{p1} и k_{p2} лежат в одной компоненте связности, то между ними существует путь, связанный ребрами, нагрузки которых включают, но не равны нагрузке сильного сужения u . С другой стороны, такая последовательность и будет путем, связывающим k_{p1} и k_{p2} в графе $G_{max} \downarrow u$, ведь наличие общей вершины-предка между двумя соседними элементами последовательности означает то, что ребро между этими

элементами будет содержать вес, включающий вес вершины-предка. Но так как в последовательности подобные вершины-предки являются потомками по отношению к u , то вес связующих ребер будет включать, но не равняться u , что и означает наличие пути в $G_{max} \downarrow u$. ■

Замечание 1. Отметим, что если в родительском графе между двумя вершинами существует общая вершина-потомок, то все фрагменты знаний, являющиеся потомками по отношению к первым двум вершинам будут лежать в одной компоненте связности, полученной после сильного сужения на общую для этих двух вершин нагрузку вершины-предка. То есть, если $\exists w_1, w_2 : kp_1 \in \text{descendants}(w_1) \& kp_2 \in \text{descendants}(w_2) \& w_1, w_2 \in \text{descendants}(u)$, а также $\exists w_3 : w_3 \in \text{descendants}(w_1) \cap \text{descendants}(w_2)$, то $kp_1, kp_2 \in C_u$.

Основываясь на доказанной теореме, предложим следующий алгоритм подсчета числа компонент связности. Пусть нам необходимо найти количество компонент связности для вершины u . Тогда распространим по каждой дочерней вершине u различные маркеры, назовем их *цветами*. Затем от каждой такой вершины будем распространять по дочерним узлам цвет, полученный ранее. Если в одну вершину поступило несколько разных цветов, то признаем эти цвета одинаковыми. В итоге, количество цветов, оставшихся после распространения их до листьев и будет совпадать с количеством компонент связности $G_{max} \downarrow u$.

Алгоритм работает корректно. Если две вершины с фрагментами знаний в качестве нагрузок получили один цвет, то, по замечанию 1, они будут лежать в одной компоненте связности. С другой стороны, предположим, что два листа kp_1 и kp_2 в родительском графе получили разные цвета, но лежат в одной компоненте. Тогда, по утверждению 2, будут существовать последовательности из вершин v_1, \dots, v_n и w_1, \dots, w_{n-1} , которые по действию алгоритма должны быть окрашены в один цвет. Но в таком случае $v_1 = kp_1$ и $v_n = kp_n$ будут окрашены в единый цвет — противоречие.

Заключение

В данной работе были представлен алгоритм, позволяющий применять только третичную структуру АБС для проверки ацикличности, а также доказана его корректность. Применение этого алгоритма может быть полезно при обосновленном использовании третичной структуры, например при ее использовании для глобального апостериорного вывода.

Список литературы

- [1] Aho A., Garey M., Ullman J. The Transitive Reduction of a Directed Graph // SIAM Journal on Computing. 1972. Vol. 1, No. 2. P. 131–137.
- [2] Khlobystova A. O., Abramov M. V., Tulupyev A.L. An approach to estimating of criticality of social engineering attacks traces // Studies in Systems, Decision and Control. 2019. Vol. 199. P. 446–456.
- [3] Su C., Andrew A., Karagas M.R. et al. Using Bayesian networks to discover relations between genes, environment, and disease // BioData Mining. 2013. Vol. 6, No. 6.
- [4] Сироткин А. В., Тулупьев А. Л. Моделирование знаний и рассуждений в условиях неопределенности: матрично-векторная формализация локального синтеза согласованных оценок истинности // Труды СПИИРАН. 2011. Вып. 18. С. 108–135.
- [5] Тулупьев А. Л. Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. // СПб.: ООО Издательство «Анатолия», 2007. 80 с.
- [6] Тулупьев А. Л., Николенко С. И., Сироткин А. В. Байесовские сети: логико-вероятностный подход. // СПб.: Наука, 2006. 607 с.
- [7] Тулупьев А. Л., Сироткин А. В. Локальный апостериорный вывод в алгебраических байесовских сетях как система матрично-векторных операций // Интегрированные модели и мягкие вычисления в искусственном интеллекте. V-я Международная научно-практическая конференция, 9 сентября — 12 сентября 2009 г. Сборник научных трудов. В 2-х т. Т. 1. СПб.: Наука, 2009. С. 425–434.
- [8] Фильченков А. А., Тулупьев А. Л. Третичная структура алгебраической байесовской сети // Труды СПИИРАН. 2011. Вып. 18. С. 164–187.
- [9] Фильченков А. А., Тулупьев А.Л. Связность и ацикличность первичной структуры алгебраической байесовской сети // Вестник Санкт-Петербургского государственного университета. Серия 1. Математика. Механика. Астрономия. 2013. Вып. 1. С. 110–119.

Вспомогательный ресурс для помощи поступающим в выборе образовательной программы СПбГУ¹

А.А. Чекалев

Санкт-Петербургский государственный университет
Университетская наб., 7, Санкт-Петербург, 199034, Россия
st087200@student.spbu.ru

А. О. Хлобыстова

Федеральное государственное учреждение науки “Санкт-Петербургский Федеральный исследовательский центр Российской академии наук”,
14 линия В.О., д. 39, Санкт-Петербург, 199178, Россия
Санкт-Петербургский государственный университет
Университетская наб., 7, Санкт-Петербург, 199034, Россия
aok@dscs.pro

Аннотация

В статье была рассмотрена актуальная проблема выбора абитуриентами образовательной программы университета. Было предложено решение в виде Telegram-бота, который бы способствовал улучшению процесса выбора абитуриентом образовательной программы, за счёт построения рекомендаций на основе его предрасположенности, интересов и навыков. Практическая значимость полученных результатов заключается в создании инструмента, который расширяет перечень существующих информационных ресурсов для абитуриентов СПбГУ, а также строит индивидуальные рекомендации.

Введение

Около 70% выпускников испытывают проблемы с выбором будущей профессией [1]. При этом неверный выбор образовательной программы при поступлении в ВУЗ ведёт к неудовлетворённости своей профессией в будущем [2]. Согласно статистике последних лет [3, 4], с каждым годом в Санкт-Петербургский Государственный Университет (СПбГУ) поступает

¹ Работа выполнена в рамках проекта по государственному заданию СПб ФИЦ РАН № FFZF-2022-0003, поддержанная Санкт-Петербургским государственным университетом, проект № 75254082 и РФФИ грант № 20-07-00839

Содержание

Системное программирование	5
Баклановский М.В., Лагунов Л.Л., Сибиряков А.Е., Ханов А.Р. Эмулятор ЕМ1. Задачи, возможности и результаты	7
Фундаментальная информатика	15
Федорченко Л. Н., Криворучко Д. И. Инструментарий для работы с динамическими отношениями	17
Мироненко И. Д. Степени двойки и вещественно-целочисленная линейная элиминация квантоворов	21
Параллельные алгоритмы и вэйвлетная обработка числовых потоков	27
Демьянович Ю.К., Иванцова О. Н. Об устойчивости всплесковых разложений	29
Распараллеливание в OpenMP и сплайновые аппроксимации	37
Бурова И. Г., Алтыбеев Г. О. Применение технологии параллельного программирования при решении уравнения Фредгольма второго рода	39
Вычислительная стохастика и статистические модели	47
Алексеева Н.П., Федорченко С.А. О проверке гипотез в дисперсионном анализе повторяющихся неполных наблюдений 1	49
Шпилев П.В. L-оптимальные планы для тригонометрической регрессии без свободного члена. 1	56
Ермаков М. С., Капаца Д. Ю. О равномерной состоятельности непараметрического критерия Неймана	63
Модели, методы и приложения тропической математики	71
Кривулин Н. К., Брюшинин М. А. Решение задачи о размещении двух объектов в пространстве с метрикой Чебышева	73
Кривулин Н. К., Губанов С. А. Алгебраическое решение задач составления оптимального графика выполнения проекта	81
Приньков А.С., Кривулин Н.К. Решение задачи выбора проекта информационной системы методами тропической математики . . .	88
Шкурат Д. Е. Разработка программных средств визуализации тропических рациональных функций	95
Губанов С. А. Оптимальное планирование сроков начала проекта при помощи методов тропической оптимизации	102
Григорьев Д. А., Кривулин Н. К. Оценка альтернатив на основе парных сравнений в задаче о выборе строительного подрядчика	108

Организаторы конференции СПИСОК-2022 благодарят Санкт-Петербургский государственный университет за неоценимую ресурсную и организационную помощь в подготовке и проведении конференции.

Научное издание

МАТЕРИАЛЫ
9-Й ВСЕРОССИЙСКОЙ НАУЧНОЙ
КОНФЕРЕНЦИИ ПО ПРОБЛЕМАМ
ИНФОРМАТИКИ СПИСОК-2022

*27–29 апреля 2022 г.
Санкт-Петербург*

Компьютерная верстка: Д.В. Луцив

Издательство «ВВМ»
190000, Санкт-Петербург,
ул. Декабристов, 6, лит. А, пом. 10н
E-mail: vvmpub@yandex.ru

Подписано в печать ___. Формат 60 × 84 $\frac{1}{16}$.
Бумага офсетная. Гарнитура Times. Печать цифровая.
Усл. печ. л. ___. Тираж 100 экз. Заказ №___.

Отпечатано в Издательстве ВВМ.
198095, Санкт-Петербург, ул. Швецова, 41.