

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (НОВОСИБИРСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ, НГУ)

Механико-математический факультет

Кафедра математического моделирования
Направление подготовки «Прикладная математика и информатика»

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
МАГИСТРА**

Звягина Максима Алексеевича

**Восстановление функциональных зависимостей вида
полиномов из экспонент по интервальным данным**

«К защите допущен»

Заведующий кафедрой

д.т.н., профессор

Барахнин В. Б. / _____

(Фамилия, И.О.) / (подпись, МП)

«_____» _____ 20____ г.

«Научный руководитель»

Д.ф.-м.н., профессор

кафедры мат. мод.

ММФ НГУ

Шарый С.П. / _____

(Фамилия, И.О.) / (подпись, МП)

«_____» _____ 20____ г.

Дата защиты: «_____» _____ 20____ г.

**Новосибирск
2024**

Реферат

Тема работы: Восстановление функциональных зависимостей вида полиномов из экспонент по интервальным данным.

Объем работы составляет 45 страниц, список использованной литературы включает 15 источников, в работе приводятся 9 таблиц и 9 рисунков.

Ключевые слова: интервал, интервальный анализ данных, задача восстановления зависимости, регрессия, полиномы из экспонент, оптимизация.

Настоящая работа посвящена решению задачи восстановления зависимостей вида полиномов из экспонент по интервальным выборкам. Цель работы построить, описать, реализовать и протестировать алгоритм нахождения заданной функциональной зависимости. Алгоритм построения зависимости является обобщением метода максимума совместности на заданный класс функций. Ключевая идея состоит в ведении меры совместности данных и поиска максимума этой меры. Таким образом задача восстановления зависимостей сводится к задаче негладкой невогнутой условной максимизации функции называемой распознающим функционалом. Для её решения использовались субградиентные методы, метод штрафных функций и технология мультистарта.

В ходе работы построен явный вид задачи оптимизации, сделана программная реализация, построено несколько примеров демонстрирующих работоспособность технологии. Алгоритм обобщен на более широкий класс функциональных зависимостей. Предложен новый подход к анализу результатов восстановления функциональных зависимостей по интервальным данным. Результаты работы могут быть использованы в прикладных исследованиях для решения задачи восстановления зависимостей, когда неопределенность данных может быть описана с помощью интервалов и требуется гарантированный характер оценок.

Результатами настоящего исследования являются описание и реализация нового подхода к решению задачи восстановления зависимостей.

Содержание

1	Введение	5
2	Постановка задачи восстановления зависимостей и сведение её к задаче оптимизации	8
3	Распознающий функционал и метод максимума совместности	12
4	Анализ задачи максимизации распознающего функционала	14
4.1	Невыпуклость распознающего функционала	14
4.2	Многоэкстремальность	16
4.3	Липшицевость	17
5	Обзор методов оптимизации	19
5.1	r - алгоритмы	19
5.2	Работа с ограничениями	21
6	Относительная совместность	23
7	Примеры	27
7.1	Регулярные данные	27
7.2	Несовместность данных и кривой	28
7.3	Данные с перекрытиями	29

7.4	Пример Ланцоша	31 ⁴
8	Усовершенствованные постановки	34
8.1	Отрицательные коэффициенты при экспонентах .	34
8.2	Дополнительные линейные слагаемые	34
9	Примеры	37
9.1	Отрицательные коэффициенты	37
9.2	Степенные слагаемые	38
9.3	Функция с точками перегиба	40
9.4	Немонотонная функция	41
10	Заключение	42
	Список литературы	44

1. Введение

*Не следует надеяться, что
какой-либо иной математический
прием мог бы привести к лучшим
результатам, ибо трудность
заключается не в способе
вычисления, а в исключительной
чувствительности показательных
функций и амплитуд к весьма
малым изменениям в данных.*

– Корнелий Ланцош

Цитата, которая была использована в качестве эпиграфа, принадлежит Корнелию Ланцошу. В своей классической монографии «Практические методы прикладного анализа» [3], он исследует различные подходы к решению многих важных практических задач. Одна из таких задач — *выделение показательных функций по эмпирическим данным*. Рассмотрев примеры, Ланцош приходит к пессимистическому выводу, который стал эпиграфом к данной работе. Хотя наша работа не полностью преодолевает все трудности, связанные с особенностями задачи и, в частности, со свойствами сумм экспонент, нам удалось получить продвижение в сравнение с предшественниками в этой сложной задаче и разработать алгоритм восстановления функциональных зависимостей.

В настоящее время методы интервального анализа все чаще используются в робототехнике [4], автоматике и киберфизических системах [5], где неопределенность значений параметров системы может быть ограничена интервалами. Задача аппроксимации данных полиномами из экспонент является достаточно распространенной в обработке физических экспериментов [9]. Например, при изучении плотности земной породы с помощью методов мюонной плотнометрии [6] возникает необходимость восстановить функцио-

нальную зависимость интенсивности потока космических частиц от глубина. Эту зависимость в силу физического смысла рассматривают, как полином из экспонент. При решении подобных задач возникает несколько проблем, с которыми столкнулись практики и которые мотивировали их обратиться к методам интервального анализа. Первая заключается в том, что каждое измерение занимает достаточно долгий промежуток времени, а значит данных нельзя сделать достаточно много для применения статистических методов, основанных на теоретико-вероятностных идеях. Вторая проблема — это размерность измерительного прибора (около метра) и скважины в которой проводится измерение (около 10–15 метров), а значит считать измерения точечными по глубине — очень грубое упрощение которое может привести к существенным ошибкам, а значит к плохой точности всей процедуры измерения плотности.

Интервальные методы позволяют преодолеть обе эти проблемы. Во-первых, так как интервальные методы не опираются на теоретико-вероятностные законы, мы можем использовать их на малых выборках. Во-вторых, у нас есть возможность учитывать сколь угодно большие погрешности за счет выбора ширины интервалов по входным или выходным переменным. Если у нас есть теоретические представления о возможных отклонениях в измерениях и ошибках, мы можем использовать эти знания при выборе размеров интервалов. Например, если в процессе расчетов мы не смогли получить ожидаемую зависимость, которая бы проходила через все измерения, это может означать, что наши предположения об отклонениях были слишком оптимистичны. Если же, наоборот, мы получили стабильное решение, мы можем попробовать сузить интервалы, уточняя тем самым теоретическую оценку ошибок. Если наши интервальные оценки являются накрывающими, то есть они содержат истинные значения измеряемых функций, то интервальные подходы, основанные на методах распознающих функционалов позволят нам найти такую функциональную зависимость, которая будет

проходить через все измерения. Подробнее об накрывающих и ненакрывающих измерениях можно прочитать в [1].

В традиционных («точечных») методах восстановления зависимостей, как правило, невозможно добиться, чтобы функциональная зависимость проходила через все измерения. Поэтому целью точечных методов восстановления зависимостей часто становится минимизация отклонения кривой от данных. Для данных с интервальной неопределенностью мы нередко можем построить кривую, проходящую через все интервальные измерения, что представляется более предпочтительным, чем минимизация отклонения от этих интервалов. Это особенно важно с учетом того, что измеряемые интервалы могут представлять собой гарантированные верхние и нижние границы истинного значения функции. Поэтому построенная функциональная зависимость должна также находиться в пределах этих границ.

В ходе проведения численных расчетов была осознана необходимость ввести новую характеристику для оценки результатов восстановления зависимостей. Характеристика \mathcal{RC} — *относительная совместность* позволяет нам количественно оценить совместность конкретной модели и данных, а также оценить насколько эта оценка теоретически могла бы быть улучшена.

2. Постановка задачи восстановления зависимостей и сведение её к задаче оптимизации

Рассматривается задача восстановления зависимостей (регрессии) вида

$$y = \sum_{j=1}^m a_j e^{-b_j x},$$

где $a_j, b_j \geq 0$, по набору измерений

$$(x_i, y_i), \quad i = 1, \dots, n,$$

где $x_i = [\underline{x}_i, \overline{x}_i]$, $y_i = [\underline{y}_i, \overline{y}_i]$ — вещественные интервалы.

Дадим краткое описание интервальной арифметики. Подробности можно найти в [2]. Интервалом будем называть замкнутый отрезок вещественной оси. Для двух интервалов x, y определим арифметические операции руководствуясь принципом

$$x \star y = \{x \star y : x \in x, y \in y\}.$$

Таким образом, интервальный аналог операции определяет область значений операции на основе точек из заданных интервалов. Интервальные арифметические операции, введённые таким образом, прежде всего предназначены для вычисления областей значений разнообразных выражений.

Для арифметических операций с интервалами $x = [\underline{x}, \overline{x}]$, $y = [\underline{y}, \overline{y}]$ получаются следующие конструктивные формулы:

$$x + y = [\underline{x} + \underline{y}, \overline{x} + \overline{y}],$$

$$x - y = [\underline{x} - \overline{y}, \overline{x} - \underline{y}],$$

$$x \cdot y = \left[\min\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\} \right],$$

$$x/y = x \cdot [1/\overline{y}, 1/\underline{y}] \quad \text{для } 0 \notin y.$$

Множество вещественных интервалов с вышеопределёнными операциями называется *классической интервальной арифметикой* и обозначается \mathbb{IR} . Так-

же нам понадобятся две операции над интервалами, такие как взятие радиуса и середины интервала:

$$\text{rad } \mathbf{x} = \frac{1}{2}(\overline{\mathbf{x}} - \underline{\mathbf{x}}),$$

$$\text{mid } \mathbf{x} = \frac{1}{2}(\overline{\mathbf{x}} + \underline{\mathbf{x}})$$

Напомним важное понятие — *интервальное расширение функции*.

Определение 1. Будем говорить, что интервальная функция $\mathbf{f} : \mathbb{IR}^n \rightarrow \mathbb{IR}^m$ является интервальным продолжением точечной функции $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ на множестве $D \subset \mathbb{R}^n$, если $\mathbf{f}(x) = f(x)$ для всех точечных аргументов $x \in D$

Определение 2. Интервальная функция $\mathbf{f} : \mathbb{IR}^n \rightarrow \mathbb{IR}^m$ называется интервальным расширением точечной функции $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ на множестве $D \subset \mathbb{R}^n$, если

1. является интервальным продолжением f на D ,
2. монотонна по включению на $\mathbb{I}D$, т. е. для любых $\mathbf{x}, \mathbf{y} \in \mathbb{I}D$ имеет место импликация $\mathbf{x} \subseteq \mathbf{y} \implies \mathbf{f}(\mathbf{x}) \subseteq \mathbf{f}(\mathbf{y})$.

Для дальнейшего использования нам необходимо определить интервальное расширение экспоненты. Положим по определению

$$\exp \mathbf{x} = \left[\min_{x \in \underline{\mathbf{x}}} e^x, \max_{x \in \overline{\mathbf{x}}} e^x \right] = \left[e^{\underline{\mathbf{x}}}, e^{\overline{\mathbf{x}}} \right].$$

Легко видеть, что определенная таким образом интервальная функция является интервальным расширением вещественной экспоненты. В дальнейшем будем называть её *интервальной экспонентой* и обозначать её для краткости $e^{\mathbf{x}}$.

Существуют разные способы понимать решение задачи восстановления зависимостей по интервальным данным. Под *совместностью данных и кривой* мы далее будем понимать ту ситуацию, при восстанавливаемая кривая

проходит через все измерения. Но в силу того, что в задаче восстановления зависимостей входные и выходные переменные часто принципиально неравноправны существует разделение на виды совместности [8].

Зафиксировав смысл, в котором мы будем понимать решение задачи восстановления зависимостей, мы можем построить систему формальных математических условий, которым должна соответствовать модель. Эти условия могут представлять собой уравнения или неравенств.

Определение 3. Информационное множество задачи восстановления функциональной зависимости по интервальным данным — это множество решений интервальной системы уравнений (неравенств и т. п. соотношений), которая построена по виду восстанавливаемой функциональной зависимости, интервальным данным задачи и другим условиям, вытекающим из постановки задачи.

Подробнее об информационном множестве можно прочитать в [1]. Любую точку из информационного множества можно использовать в качестве решения задачи восстановления зависимостей. Сформулируем решение задачи восстановления зависимости в терминах нашей постановки.

Определение 4. Решением задачи восстановления зависимости вида полиномов из экспонент будем называть такой набор параметров $(a, b) \in \mathbb{R}_+^m \times \mathbb{R}_+^m$, что верно

$$\sum_{j=1}^m a_j e^{-b_j x_i} \subseteq y_i, \quad i = 1, \dots, n,$$

Множество параметров модели, для которых это верно, по определению совпадает с допусковым множеством решений интервальной системы уравнений и неравенств:

$$\sum_{j=1}^m a_j e^{-b_j x_i} = y_i, \quad i = 1, \dots, n, \quad (2.1)$$

$$a_j, b_j \geq 0, \quad j = 1, \dots, m, \quad (2.2)$$

Дадим определение допусковому множеству решений в терминах нашей задачи.

Определение 5. Допусковым множеством решений системы уравнений (2.1) при дополнительных условиях (2.2) называется множество:

$$\Xi_{tol}(\mathbf{x}, \mathbf{y}) = \left\{ (a, b) \in \mathbb{R}_+^m \times \mathbb{R}_+^m \mid (\forall x \in \mathbf{x}) (\exists y \in \mathbf{y}) \left(\sum_{j=1}^m a_j e^{-b_j x_i} = y_i, \quad i = 1, \dots, n \right) \right\}$$

Обобщая идеи, описанные, например, в [8] и [10], сведем задачу отыскания точек допускового множества решений (или проверки, что их нет) к задаче оптимизации.

3. Распознающий функционал и метод максимума совместности

Известен факт (см. [2]), что для интервалов \mathbf{p}, \mathbf{q} выполнено

$$\mathbf{p} \subset \mathbf{q} \iff \text{rad } \mathbf{q} - \text{rad } \mathbf{p} \geq |\text{mid } \mathbf{q} - \text{mid } \mathbf{p}|.$$

Тогда

$$\sum_{j=1}^m a_j e^{-b_j \mathbf{x}_i} \subset \mathbf{y}_i \iff \text{rad } \mathbf{y}_i - \sum_{j=1}^m a_j \text{rad } e^{-b_j \mathbf{x}_i} - \left| \text{mid } \mathbf{y}_i - \sum_{j=1}^m a_j \text{mid } e^{-b_j \mathbf{x}_i} \right| \geq 0,$$

где

$$\text{rad } e^{\mathbf{x}} = \frac{1}{2}(e^{\bar{\mathbf{x}}} - e^{\underline{\mathbf{x}}}), \quad \text{mid } e^{\mathbf{x}} = \frac{1}{2}(e^{\bar{\mathbf{x}}} + e^{\underline{\mathbf{x}}}).$$

Эти условия должны быть выполнены для всех i , тогда это эквивалентно

$$\min_{i=1, \dots, n} \left\{ \text{rad } \mathbf{y}_i - \sum_{j=1}^m a_j \text{rad } e^{-b_j \mathbf{x}_i} - \left| \text{mid } \mathbf{y}_i - \sum_{j=1}^m a_j \text{mid } e^{-b_j \mathbf{x}_i} \right| \right\} \geq 0.$$

Так как нас интересуют решения, в которых $b_j \geq 0$, преобразуем радиусы

$$\min_{i=1, \dots, n} \left\{ \text{rad } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_i} - e^{-b_j \bar{\mathbf{x}}_i}) - \left| \text{mid } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_i} + e^{-b_j \bar{\mathbf{x}}_i}) \right| \right\} \geq 0.$$

Появляется необходимость в следующем определении.

Определение 6. Распознающим функционалом задачи (2.1), (2.2) будем называть функцию $\text{Tol}(\mathbf{x}, \mathbf{y}, a, b) : X \times Y \times Q \rightarrow \mathbb{R}$, где $X, Y \subset \mathbb{IR}^n$ — множества интервальных значений входной и выходной (зависимой) переменных и множество параметров модели $Q = \{(a, b) : a \in \mathbb{R}_+^m, b \in \mathbb{R}_+^m\}$, определяемую следующим образом:

$$\begin{aligned} \text{Tol}(\mathbf{x}, \mathbf{y}, a, b) = \min_{i=1, \dots, n} \left\{ \text{rad } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_i} - e^{-b_j \bar{\mathbf{x}}_i}) \right. \\ \left. - \left| \text{mid } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_i} + e^{-b_j \bar{\mathbf{x}}_i}) \right| \right\} \end{aligned} \quad (3.1)$$

Из описанного выше следует следующее утверждение.

Утверждение 1. Пусть $(a, b) \in Q$, тогда

$$(a, b) \in \Xi_{\text{tol}}(\mathbf{x}, \mathbf{y}) \iff \text{Tol}(\mathbf{x}, \mathbf{y}, a, b) \geq 0.$$

Таким образом, для ответа на вопрос о существовании и поиска решения задачи (2.1), (2.2) достаточно решить задачу негладкой условной оптимизации

$$\max_{(a,b) \in Q} \text{Tol}(\boldsymbol{x}, \boldsymbol{y}, a, b).$$

Этот метод отыскания функциональной зависимости по интервальным данным называется методом максимума совместности [10], так как распознающий функционал показывает меру совместности интервальных данных и зависимости.

4. Анализ задачи максимизации распознающего функционала

Как было описано в предыдущем разделе, мы можем свести задачу восстановления зависимостей к поиску экстремума распознающего функционала. Методы решения оптимизационных задач напрямую зависят от свойств целевой функции и ограничений, накладываемых на переменные. Как было сказано, наша задача является задачей негладкой условной оптимизации. В силу негладкости из возможного арсенала методов исключаются методы высоких порядков, но остаются субградиентные методы и методы нулевого порядка. С точки зрения ограничений наша задача является одной из самых базовых среди задач условной оптимизации. Условия типа $a_j, b_j \geq 0$ часто возникают на практике, что породило много эффективных подходов обработке таких ограничений.

Используемые нами подходы описаны в разделе 5. Далее обсудим наличие или отсутствие часто используемых в практике свойств целевой функции (*распознающего функционала*).

4.1. Невыпуклость распознающего функционала

Важнейшим классом задач оптимизации являются задачи выпуклой оптимизации, которые в общем виде представляют собой задачу типа

$$\begin{cases} \min f_0(x) \\ s.t. f_i(x) \leq 0, \quad i = 1, \dots, m, \\ Ax = b, \end{cases}$$

где функции $f_i(x)$ для всех $i = 0, \dots, m$ являются выпуклыми. Отсутствие вогнутости (выпуклости) в общем случае может приводить к экспоненциальному росту числа шагов метода, необходимых для нахождения решения с заданной точностью, с ростом размерности x .

В дальнейшем мы рассмотрим примеры и увидим на графиках линий уровня, что распознающий функционал не является вогнутой или даже квазивогнутой функцией. Существуют области седлового характера. Это приводит нас к очевидным трудностям, связанным с тем фактом, что многие методы негладкой оптимизации, используют субградиенты или суперградиенты, определенные только для локально выпуклых или вогнутых функций соответственно.

Для формализации описания оптимизационных алгоритмов будем использовать язык *почти дифференцируемых функций* (см. [13]).

Определение 7. Функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ называется почти дифференцируемой, если она удовлетворяет следующим условиям:

1. в любой ограниченной области удовлетворяет условию Липшица;
2. почти везде дифференцируема;
3. ее градиент непрерывен на том множестве M , где он определен.

Можно показать, что в определении почти дифференцируемости условие 1 влечет 2.

Определение 8. Почти-градиентом функции f в точке x_0 называется вектор, являющийся предельной точкой некоторой последовательности градиентов $g(x_1), g(x_2), \dots, g(x_k), \dots$, где $\{x_k\}_{k=1}^\infty$ — последовательность точек, сходящихся к x_0 , и такая, что во всех точках этой последовательности f дифференцируема.

Теорема 1. Множество $G(x)$ почти-градиентов в любой точке $x \in \mathbb{R}^n$ почти дифференцируемой функции является непустым, ограниченным и замкнутым.

Теорема 2. Произвольная выпуклая функция $f(x)$, определенная в \mathbb{R}^n , является почти дифференцируемой, а её любой почти-градиент в точке x совпадает с некоторым субградиентом.

Доказательства теорем можно найти в [13].

Пример 1. Для функции $f(x, y) = |x| - |y|$ в точке $(0, 0)$ множество почти-градиентов представляет собой набор из 4 векторов $\{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$, т.к. для любой точки дифференцируемости имеем $\nabla f(x, y) = (\text{sgn}(x), -\text{sgn}(y))$.

Пример 2. Представим распознающий функционал в виде

$$\begin{aligned} \text{Tol}(\mathbf{x}, \mathbf{y}, a, b) &= \min_{i=1, \dots, n} g_i, \\ g_i &= \text{rad } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j \left(e^{-b_j \mathbf{x}_i} - e^{-b_j \bar{\mathbf{x}}_i} \right) - \left| \text{mid } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j \left(e^{-b_j \mathbf{x}_i} + e^{-b_j \bar{\mathbf{x}}_i} \right) \right|, \\ &\quad i = 1, \dots, n. \end{aligned}$$

Тогда в точках дифференцируемости

$$\begin{aligned} \frac{\partial \text{Tol}(\mathbf{x}, \mathbf{y}, a, b)}{\partial a_k} &= \frac{\partial g_l}{\partial a_k} = \\ &\quad \text{rad } e^{-b_k \mathbf{x}_l} + \text{sgn} \left(\text{mid } \mathbf{y}_l - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \mathbf{x}_l} + e^{-b_j \bar{\mathbf{x}}_l}) \right) \text{mid } e_l^{-b_k \mathbf{x}}, \\ \frac{\partial \text{Tol}(\mathbf{x}, \mathbf{y}, a, b)}{\partial b_k} &= \frac{\partial g_l}{\partial b_k} = \frac{1}{2} a_k \left(\mathbf{x}_l e^{-b_k \mathbf{x}_l} - \bar{\mathbf{x}}_l e^{-b_k \bar{\mathbf{x}}_l} \right) \\ &\quad - \frac{1}{2} a_k \text{sgn} \left(\text{mid } \mathbf{y}_l - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \mathbf{x}_l} + e^{-b_j \bar{\mathbf{x}}_l}) \right) \left(\mathbf{x}_l e^{-b_k \mathbf{x}_l} + \bar{\mathbf{x}}_l e^{-b_k \bar{\mathbf{x}}_l} \right), \end{aligned}$$

где $g_l = \min_{i=1, \dots, n} g_i$. Вектор из этих выражений мы и будем использовать в качестве почти-градиента в наших дальнейших расчетах. Для нашей функции, за исключением множеств меры 0, это выражение совпадает с суперградиентом. Поэтому для решения задачи поиска экстремума мы будем применять супер(суб)градиентные методы.

4.2. Многоэкстремальность

Заметим, что отсутствие выпуклости также следует из того факта, что если набор параметров $a = (a_1, \dots, a_m)$, $b = (b_1, \dots, b_m)$ является оптимумом $\text{Tol}(\mathbf{x}, \mathbf{y}, a, b)$, то для любой перестановки $\sigma \in S_m$ (где S_m — множество всех

перестановок на m элементах) набор параметров $\tilde{a} = (a_{\sigma(1)}, \dots, a_{\sigma(m)})$, $\tilde{b} = (b_{\sigma(1)}, \dots, b_{\sigma(m)})$ также является оптимумом. Однако в силу того, что значения распознающего функционала во всех таких точках совпадают, нас не пугает перспектива попасть в такой локальный максимум.

4.3. Липшицевость

Определение 9. Функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ называется L -липшицевой на множестве $X \subseteq \mathbb{R}^n$, для заданной нормы $\|\cdot\|$, если существует число $L \geq 0$ такое, что верно

$$|f(x) - f(y)| \leq L\|x - y\|$$

для любых x, y из X . Константа L в таких случаях называется константой Липшица (функции f на множестве X)

Лемма 1. Пусть функции $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ являются L_1 и L_2 -липшицевыми на X соответственно, тогда $\phi(x) = \min\{f_1(x), f_2(x)\}$ является липшицевой с константой $\max\{L_1, L_2\}$.

Доказательство. Необходимо оценить выражение $|\phi(x) - \phi(y)|$. Заметим, что если в обоих точках x и y одна из функций меньше другой (например $f_1 \leq f_2$), то имеем

$$|\phi(x) - \phi(y)| = |f_1(x) - f_1(y)| \leq L_1\|x - y\| \leq \max\{L_1, L_2\}\|x - y\|.$$

Осталось рассмотреть ситуацию, когда это не так. Пусть (для определенности)

$$\begin{aligned}\phi(x) &= \min\{f_1(x), f_2(x)\} = f_1(x), \\ \phi(y) &= \max\{f_1(y), f_2(y)\} = f_2(y).\end{aligned}$$

Тогда

$$\begin{aligned}\phi(x) - \phi(y) &= f_1(x) - f_2(y) \leq f_2(x) - f_2(y) \leq \\ &L_2\|x - y\| \leq \max\{L_1, L_2\}\|x - y\|,\end{aligned}$$

аналогично

$$\begin{aligned}\phi(y) - \phi(x) &= f_2(y) - f_1(x) \leq f_1(y) - f_1(x) \leq \\ &L_1\|x - y\| \leq \max\{L_1, L_2\}\|x - y\|,\end{aligned}$$

откуда

$$|\phi(x) - \phi(y)| \leq \max\{L_1, L_2\}\|x - y\|.$$

■

Замечание 1. Пусть функции $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, k$, являются L_i - липшицевыми на X соответственно, тогда функции $\phi(x) = \min_{i=1, \dots, k} \{f_i(x)\}$ и $\psi(x) = \max_{i=1, \dots, k} \{f_i(x)\}$ являются липшицевыми с константами $\max_{i=1, \dots, k} \{L_i\}$.

Утверждение 2. Функционал $\text{Tol}(\mathbf{x}, \mathbf{y}, a, b)$ — липшицева функция на любом ограниченном множестве.

Доказательство. Используя доказанное выше и тот факт, что $|\mathbf{x}| = \max\{-\underline{\mathbf{x}}, \overline{\mathbf{x}}\}$, имеем, что липшицевость функции $\text{Tol}(\mathbf{x}, \mathbf{y}, a, b)$ сводиться к липшицевости функций $\phi_i = \sum_{j=1}^m a_j e^{-b_j \tilde{x}_i}$, где $\tilde{x}_i \in \{\underline{x}_i, \overline{x}_i\}$. Последние же липшицевы на любом ограниченном множестве в силу своей непрерывной дифференцируемости.

■

5. Обзор методов оптимизации

Классические методы минимизации гладких функций, такие как градиентные спуски и модификации метода Ньютона, могут быть получены, как результат сведения исходной задачи к задаче локальной минимизации линейной или квадратичной аппроксимации исходной функции с помощью формулы Тэйлора. Но в случае негладких целевых функций эти методы оказываются не применимы. Однако задачи негладкой оптимизации возникают довольно часто. Это привело к бурной разработке множества эффективных методов различными научными школами. На данный момент мы используем описанный далее алгоритм, но планируется расширение «парка» методов и поиск оптимальных.

5.1. r - алгоритмы

Один из методов преодоления возникших трудностей, предложенный Шором, заключается в прямой замене градиента в градиентном спуске на субградиент. Тогда

$$x^{(k+1)} = x^{(k)} - \gamma_k \nabla f(x^{(k)}), \quad k = 1, 2, \dots$$

превращается в

$$x^{(k+1)} = x^{(k)} - \gamma_k \partial f(x^{(k)}), \quad k = 1, 2, \dots$$

Такой метод получил название *субградиентный спуск*. На практике такие методы в исходном виде оказались не очень применимы в силу медленной сходимости. Одним из способов ускорения сходимости субградиентных методов является применение специальных неортогональных преобразований пространства. Опишем общую схему таких преобразований.

При переходе с шага k на $k + 1$ происходит преобразование координат $x = B_k y$ где B_k матрица линейного неособого(невырожденного) преобразования. Тогда $\phi(y) = f(B_k x)$ и $\partial \phi(y) = B_k^T \partial f(x)$. И теперь решается задача

минимизации функции $\phi(y)$ исходным методом.

$$y^{(k+1)} = y^{(k)} - h_k \frac{\partial \phi(y^{(k)})}{\|\partial \phi(y^{(k)})\|}, \quad k = 1, 2, \dots$$

Если записать эту схемы в исходных координатах x , то имеем

$$x^{(k+1)} = x^{(k)} - h_k B_k \frac{B_k^T \partial f(x^{(k)})}{\|B_k^T \partial f(x^{(k)})\|}, \quad k = 1, 2, \dots$$

Если дополнить это рекуррентным пересчетом матрицы $B_{k+1} = T B_k$, то получаем *субградиентный спуск с растяжением пространства*.

Есть два основных подхода к выбору направления растяжения:

1. направление субградиента
2. направление разности двух последовательных субградиентов

Вариации последнего подхода получили название *r - алгоритмы*.

Введем оператор растяжения пространства $R_\alpha(\xi)$ по формуле

$$R_\beta(\xi) = I + (\beta - 1)\xi\xi^\top.$$

Тогда в мы можем полностью описать *r-алгоритм* следующим образом.

r - Алгоритм минимизации функции $f(x)$ называется процедура итеративного нахождения последовательности n -мерных векторов $\{x^{(k)}\}_{k=0}^\infty$ и последовательности матриц $\{B_k\}_{k=0}^\infty$ по следующему правилу:

$$x^{(k+1)} = x^{(k)} - h_k B_k \xi^{(k)}, \quad B_{k+1} = B_k R_{\beta_k}(\eta^{(k)}), \quad k = 0, 1, 2, \dots$$

где

$$\xi^{(k)} = \frac{B_k^T \partial f(x^{(k)})}{\|B_k^T \partial f(x^{(k)})\|}, \quad h_k \geq h_k^* = \arg \min f(x^{(k)} - h B_k \xi^{(k)}),$$

$$\eta^{(k)} = \frac{B_k^T r^{(k)}}{\|B_k^T r^{(k)}\|}, \quad r^{(k)} = \partial f(x^{(k)}) - \partial f(x^{(k-1)}), \quad \beta_k < 1.$$

Если $\partial f(x^{(k)}) = 0$ то $x^{(k)}$ — решение задачи и процесс останавливается. В качестве B_0 можно брать единичную матрицу, $x^{(0)}$ — стартовая точка.

5.2. Работа с ограничениями

Существует множество различных подходов к решению задач условной оптимизации. В данной работе мы используем *метод штрафных функций*, так как он прост в реализации и хорошо адаптируется к нашей задаче.

Метод основан на идее сведения задачи условной минимизации к решению серии задач безусловной оптимизации с скорректированной целевой функцией. Пусть решается задача

$$\begin{cases} \min f(x), \\ x \in X. \end{cases}$$

Тогда её можно свести к серии задач

$$\begin{cases} \min \Phi_k(x), & k = 1, 2, \dots \\ x \in X_0, \end{cases}$$

где $X \subset X_0$. Эти функции Φ_k подбираются так, чтобы их значение было близко к значению исходной функции на множестве X . На множестве $X \setminus X_0$ последовательность функций должна быстро возрастать, за счет чего минимум едва ли будет оказываться за пределами множества X при больших номерах k .

В качестве функции Φ_k как правило берется сумма исходной функции f и вспомогательной *штрафной функции* P_k определяемой условием:

$$\lim_{k \rightarrow \infty} P_k(x) = \begin{cases} 0, & \forall x \in X, \\ \infty, & \forall x \in X \setminus X_0 \end{cases}$$

Если $P_k(x) = A_k \psi(x)$, где A_k — некоторый скаляр и $\psi(x)$ — специально подобранная функция, то параметр A_k называется *штрафным параметром*. Для нашей задачи максимизации мы используем в качестве штрафной функции

$$P_k(x) = \alpha_k \sum_{j=0}^m \min\{a_j, 0\} + \beta_k \sum_{j=0}^m \min\{b_j, 0\},$$

где $\alpha_k, \beta_k \geq 0$ Легко заметить, что при положительных коэффициентах a_j, b_j она равно 0 и становится отрицательной иначе. В нашей задаче мы вместо постепенного роста штрафного параметра сразу берем достаточно большие значение и увеличиваем их, если решение не находится. По сути своей это применение *метода точных штрафных функций*, более формальное описание которого можно найти, например, в [15].

Таким образом, для решения задачи восстановления зависимостей мы будем применять метод максимума совместности, который сводит задачу восстановления зависимостей к задаче оптимизации. Задача оптимизации оказалась невыпуклой, негладкой и имеет ограничения вида неравенств. Ограничения вида неравенств мы учитываем, корректируя оптимизационную функцию путём введения штрафную функции. Для решения этой оптимизационной задачи мы будем применять технологию мултистарта из случайных точек области поиска решения. Из каждой точки мы будем запускать r - алгоритм для поиска локального экстремума, а потом среди всех экстремумов будет выбран наилучший.

6. Относительная совместность

Прежде чем рассматривать примеры нахождения функциональных зависимостей, введем полезную характеристику, на которую будем обращать внимание в при анализе решения задачи. Общий вид распознающего функционала в случае восстановления произвольной функционально зависимости, часто имеет вид минимума разностей:

$$\text{Tol} = \min_{i=1,\dots,n} \{\text{rad } \mathbf{y}_i - \psi_i\}, \quad (6.1)$$

где ψ_i это положительные функции от данных и неизвестных параметров (3.1). Из (6.1) следует ограниченность сверху соответствующих распознающих функционалов. В линейном случае, в связке с вогнутостью, это гарантирует существование максимума и его глобальность. В случае произвольной нелинейной функциональной зависимости у нас могут отсутствовать свойства вогнутости или унимодальности. А значит, для нахождения глобального оптимума необходимо решение задачи невыпуклой оптимизации, которая в общем случае требует экспоненциального времени. Если для нахождения глобального оптимума мы используем локальные методы, то у нас нет никаких гарантий, что найденный нами максимум является глобальным. Поэтому нами была осознанна необходимость оценивать, насколько найденный нами максимум может отличаться от глобального максимума.

Вторая идея, приводящая нас к мысли о необходимости ввести специальную количественную характеристику совместности, заключается в желании сравнивать качество моделей. Если у нас есть две модели с разными типами функциональных зависимостей и мы хотим выбрать более предпочтительную, то в случае накрывающих данных (подробное обсуждение этого понятия см. [1]) имеет смысл отдать предпочтение модели, для которой значение распознающего функционала наибольшее, так как распознающий функционал по сути и является мерой совместности данных и функциональной зависимости.

Но нам хочется уметь говорить о качестве моделей без привязки к данным. То есть, получить аналог коэффициента детерминации в случае интервальных данных. Это позволит говорить о моделях, приближающих данные, без оглядки на размеры данных. Например, значение распознающего функционала 100 или 5, или 1 ничего нам не говорит, также как средне-квадратичное отклонение 100. Но значение коэффициента детерминации 0.8 сразу дает понять, что модель достаточно неплохо согласуется с данными.

Определение 10. Пусть имеется множество функциональных зависимостей (моделей) M , а функция f какая-то выбранная функция из M . Тогда точной относительной совместностью модели f относительно множества M будем называть

$$\mathcal{RC}^* = \frac{\text{Tol}(\mathbf{X}, \mathbf{y}, f)}{\max_{g \in M} \text{Tol}(\mathbf{X}, \mathbf{y}, g)} \quad (6.2)$$

Для случая, когда M это линейные модели с заданным количеством входных переменных, мы умеем с помощью метода максимума совместности точно находить максимум знаменателя. А значит мы можем добиться $\mathcal{RC}^* = 1$. Но в случае произвольного множества M это становится значительно сложнее. Мы уже выяснили, что в случае зависимости вида полинома из экспонент, для нахождения этого максимума нам необходимо решать задачу невыпуклой оптимизации. Поэтому вычисление \mathcal{RC}^* в общем случае является крайне сложной задачей. Но мы точно знаем, что $\mathcal{RC}^* \leq 1$. Заметим, что если распознающий функционал имеет вид 6.1, то верно

$$\max_{g \in M} \text{Tol}(\mathbf{X}, \mathbf{y}, g) \leq \min_{i=1, \dots, n} \text{rad } \mathbf{y}_i \quad (6.3)$$

Это значит, что:

$$\mathcal{RC}^* = \frac{\text{Tol}(\mathbf{X}, \mathbf{y}, f)}{\max_{g \in M} \text{Tol}(\mathbf{X}, \mathbf{y}, g)} \geq \frac{\text{Tol}(\mathbf{X}, \mathbf{y}, f)}{\min_{i=1, \dots, n} \text{rad } \mathbf{y}_i} = \mathcal{RC} \quad (6.4)$$

Это приводит нас к следующей конструкции. В качестве достаточного условия совместности произвольной модели f и данных \mathbf{X}, \mathbf{y} предлагается

брать характеристику

$$\mathcal{RC} = \frac{\text{Tot}(\mathbf{X}, \mathbf{y}, f)}{\min_{i=1, \dots, n} \text{rad } \mathbf{y}_i},$$

которую будем называть *отношением совместности*. Таким образом мы получили оценку снизу для совместности «лучшего приближения» данных. Если в случае расчетов мы получили модель f для которой $\mathcal{RC} = 0.9$, это значит, что мы можем быть уверенны, что «лучшая модель», которую мы можем найти, «не очень лучше» найденной. Имея только значение распознающего функционала, мы не смогли бы сделать такой вывод.

Легко заметить, что значение характеристики \mathcal{RC} принадлежит полуинтервалу $] - \infty, 1]$. Эта характеристика, как и распознающий функционал, может применяться для выбора лучшей из нескольких моделей. За счет деления на радиус зависимой переменной мы достигаем меньшей зависимости от масштаба данных.

Если при расчетах \mathcal{RC} оказалась близка к 1, то мы можем считать, что найденный нами экстремум распознающего функционала или является глобальным или отличается от него не очень сильно.

При работе с этой характеристикой, может возникнуть проблема деления на нуль. Но $\min \text{rad } \mathbf{y}_j = 0$, соответствует ситуации, когда у нас есть точечные данные по \mathbf{y} , в этом случае $\text{Tot} \leq 0$ и совместности данных и кривой практически (за исключением случая $\text{Tot} = 0$) нет.

Возможно, имеет смысл рассмотреть модификации характеристики \mathcal{RC} , в которых в знаменателе стоит другое выражение зависящее от $\text{rad } \mathbf{y}$, чтобы избавиться от проблемы деления на нуль.

Рассмотрим несколько примеров вычисления \mathcal{RC} для разных наборов данных и моделей. Первым примером покажем, что данное условие не является необходимым. Близость к нулю не означает отсутствие совместности данных. На рисунке №. 1 пример построения линейной зависимости методом максимума совместности для представленных данных. Интуитивно кажется,

что такие данные очень совместны с линейной функцией, однако $\mathcal{RC} = 0$. Это связано с тем, что максимальное значение распознающего функционала для этих данных равно нулю и существует единственная прямая отвечающая условию сильной совместности. В данном случае близость \mathcal{RC} к нулю скорее демонстрирует отсутствие устойчивости восстановленной зависимости.

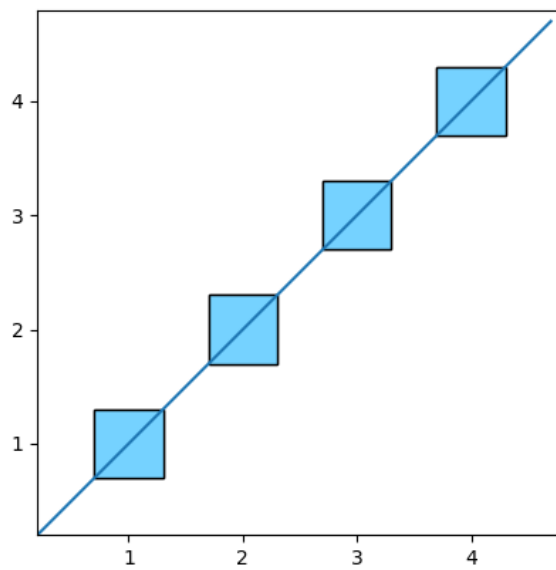


Рис. 1: Пример неустойчивой совместности

7. Примеры

В этом разделе рассматриваются несколько примеров применения метода максимума совместности для интервальных данных.

7.1. Регулярные данные

Рассмотрим первый набор данных из таблицы 1

Таблица 1: Наборы данных № 1

#	1	2	3
\mathbf{x}_i	$[-0.3, 0.5]$	$[1.6, 2.4]$	$[3.6, 4.4]$
\mathbf{y}_i	$[1.59, 2.79]$	$[0.88, 2.08]$	$[0.88, 2.08]$
#	1	2	3
\mathbf{x}_i	$[5.6, 6.4]$	$[7.6, 8.4]$	$[9.6, 10.4]$
\mathbf{y}_i	$[0.38, 1.58]$	$[0.25, 1.45]$	$[0.13, 1.33]$

По этому набору данных попытаемся восстановить зависимость вида:

$$y = a_1 e^{-b_1 x} + a_2 e^{-b_2 x}, \quad (7.1)$$

где коэффициенты a_i , b_i , $i = 1, 2$ — неизвестные параметры модели. Используя описанную выше технологию, было найдено решение $a^* = (1.76, 0.46)$, $b^* = (2.47 \cdot 10^{-1}, 2.93 \cdot 10^{-10})$ со значением распознающего функционала $\text{Tol}(\mathbf{x}, \mathbf{y}, a^*, b^*) = 0.43$. Значение $\mathcal{RC} = 0.72$. Полученное решение изображено на рисунке 2.

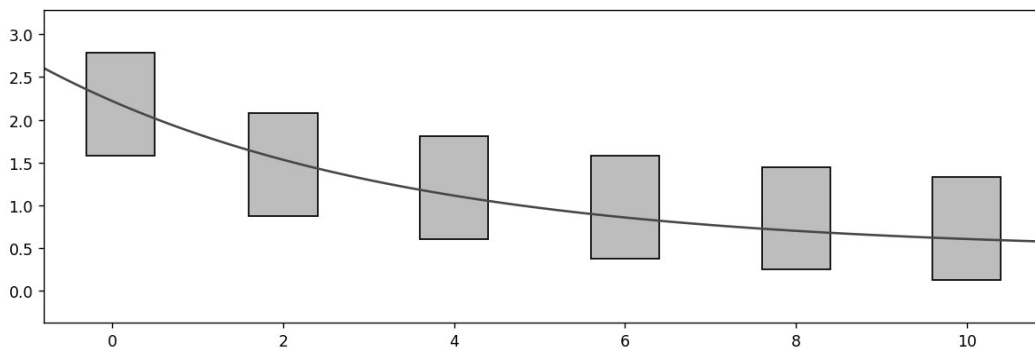


Рис. 2: Функциональная зависимость для набора данных № 1.

При увеличении количества экспонент в нашем примере, значительный рост значения максимума распознающего функционала не наблюдается (разница порядка 10^{-10}).

Продemonстрируем невогнутость распознающего функционала. На рисунке 3 представлен график распознающего функционала как функции b при фиксированном $a = (1.6, 0.4)$.

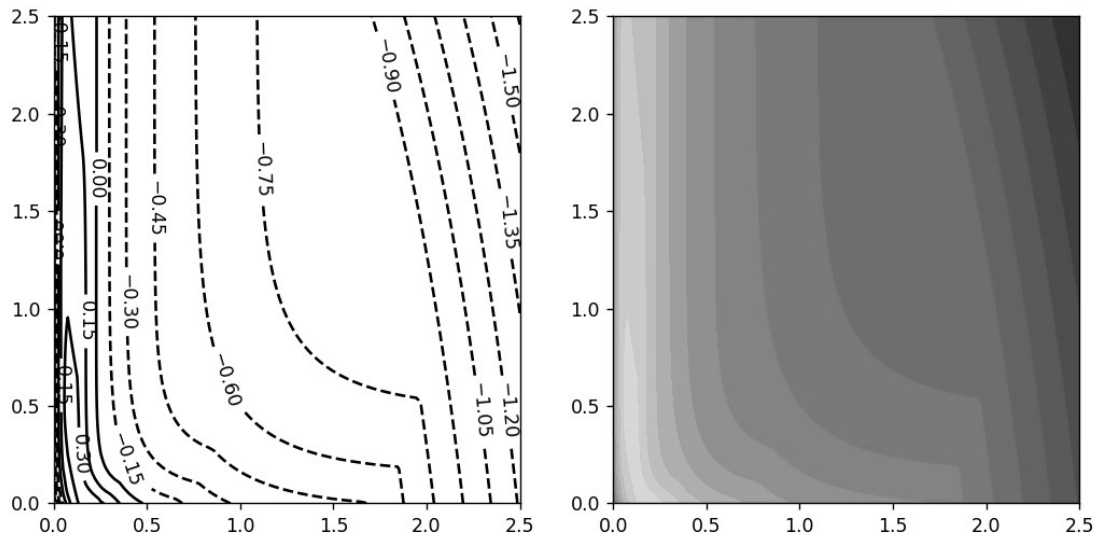


Рис. 3: Линии уровня распознающего функционала как функции b для набора данных №1.

7.2. Несовместность данных и кривой

Интересно рассмотреть случай, когда кривой, проходящей через все данные в нужном нам смысле, не существует. Набор данных, демонстрирующий эту ситуацию, приведен в Таблице 2. Если мы начнем равномерно увеличи-

Таблица 2: Набор данных № 2

#	1	2	3	4
x_i	[1.9, 2.1]	[3.82, 4.18]	[5.86, 6.14]	[7.89, 8.11]
y_i	[0.80, 2.0]	[0.36, 1.44]	[0.46, 1.3]	[1.17, 1.53]
#	5	6	7	
x_i	[9.88, 10.12]	[11.87, 12.13]	[13.86, 14.14]	
y_i	[0.23, 0.71]	[0.16, 0.521]	[-0.088, 0.752]	

вать радиусы всех выходных интервалов, то функция, полученная методом

максимальной совместимости, будет первой кривой из заданного параметрического семейства, что коснется всех данных. Результаты продемонстрированы на рисунке №. 4.

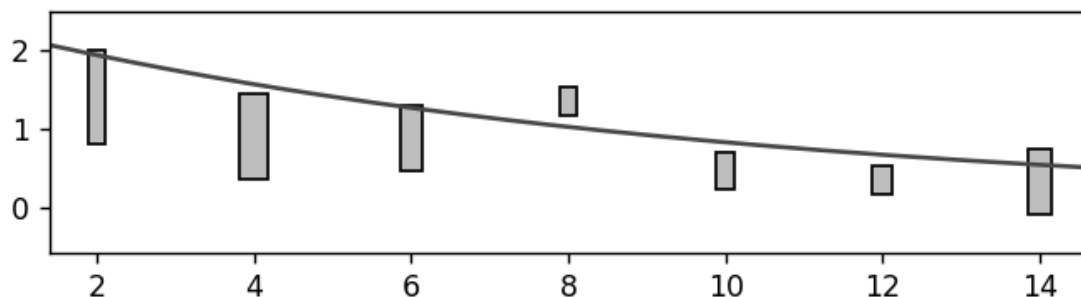


Рис. 4: Пример несовместимости данных и кривой.

В данном примере решение: $a^* = (2.94, 2.40)$, $b^* = (12.47, 0.11)$, и $\text{Tol}(x, y, a^*, b^*) = -0.15$. Как и ранее увеличение количества экспонент не приводит к значительному изменению значения максимума распознающего функционала. Например, в равномерной норме на отрезке аппроксимации разница между решениями для 2 и 10 экспонент равна $8.53 \cdot 10^{-8}$.

7.3. Данные с перекрытиями

Далее, мы исследуем работоспособность метода на данных, где есть пересечения интервалов (что означает, что интервальные данные по существу широки). Два примера таких данных приведены в Таблице №. 3.

Данные и результаты расчетов представлены на Рис. 5. В этих примерах мы снова замечаем, что рост значения распознающего функционала останавливается при количестве экспонент большем чем 2.

Эти примеры показывают, что описанный выше метод хорошо работает с пересекающимися интервальными данными.

Таблица 3: Наборы данных № 3, 4

#	\mathbf{x}_i^3	\mathbf{y}_i^3	\mathbf{x}_i^4	\mathbf{y}_i^4
1	[0.47, 1.23]	[1.21, 1.88]	[0.83, 1.29]	[0.90, 1.82]
2	[1.34, 1.80]	[0.70, 1.46]	[1.04, 1.89]	[0.95, 1.90]
3	[1.46, 2.24]	[1.09, 1.68]	[1.47, 2.14]	[0.66, 1.97]
4	[1.95, 2.41]	[0.64, 1.59]	[1.69, 2.55]	[1.31, 1.80]
5	[1.96, 2.90]	[0.21, 0.92]	[2.33, 3.22]	[0.28, 1.22]
6	[2.85, 3.39]	[0.96, 1.43]	[2.58, 3.31]	[0.15, 0.78]
7	[3.20, 3.89]	[0.32, 1.47]	[3.03, 3.57]	[0.00, 0.99]
8	[3.36, 4.07]	[0.16, 0.99]	[3.38, 3.86]	[0.21, 1.55]
9	[3.73, 4.67]	[0.54, 1.86]	[3.75, 4.51]	[0.23, 1.21]
10	[4.40, 5.12]	[0.28, 1.67]	[4.18, 5.15]	[0.31, 0.85]
11	[4.61, 5.18]	[0.21, 0.97]	[4.71, 5.30]	[-0.04, 0.58]
12	[4.94, 5.82]	[0.42, 0.92]	[5.12, 5.77]	[-0.08, 1.09]
13	[5.32, 6.25]	[0.63, 1.54]	[5.14, 6.14]	[-0.12, 0.50]
14	[5.76, 6.44]	[-0.03, 0.51]	[5.86, 6.32]	[0.55, 1.42]
15	[6.46, 6.93]	[0.22, 1.10]	[6.13, 7.06]	[0.35, 0.94]
16	[6.62, 7.61]	[0.28, 0.91]	[6.68, 7.54]	[-0.11, 1.11]
17	[7.08, 7.96]	[-0.31, 1.05]	[6.84, 7.71]	[0.43, 1.04]
18	[7.19, 8.12]	[-0.12, 0.46]	[7.42, 7.85]	[0.17, 0.64]
19	[7.97, 8.67]	[-0.16, 0.75]	[7.71, 8.59]	[0.21, 1.48]
20	[8.10, 8.92]	[0.57, 1.13]	[8.31, 9.27]	[0.24, 1.32]

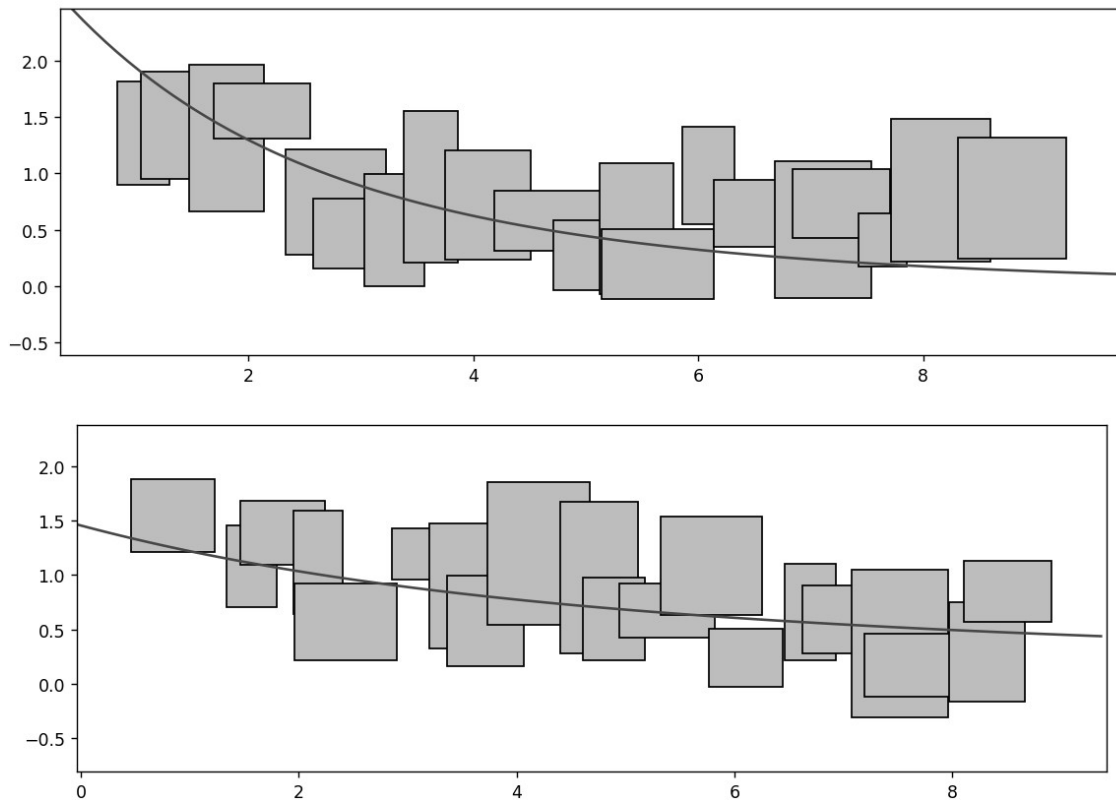


Рис. 5: Примеры данных с перекрытиями.

7.4. Пример Ланцоша

Рассмотрим численный пример, взятый из классической монографии [3]. В её шестой главе разбирается вопрос выделения показательных функций из эмпирических данных. Мы не будем описывать метод, предложенный для решения этой задачи в [3], а лишь используем набор данных и сравним с результатами, полученными в книге. Сами данные приведены в Таблице 4.

Эти данные считаются абсолютно точными по входным переменным (что в терминах интервального анализа означает $\text{rad } x_i = 0$) и имеют погрешность в половину второго десятичного знака по выходным переменным (т. е. $\text{rad } y_i = 0.005$).

Эти данные были получены из кривой, заданной уравнением

$$f(x) = 0.0951e^{-x} + 0.8607e^{-3x} + 1.5576e^{-5x},$$

с последующим округлением до второго знака после запятой. В ходе решения задачи отыскания показательных функций, описывающих эти данные,

Таблица 4: Наборы данных для примера Ланцоша

#	x_i	y_i	#	x_i	y_i	#	x_i	y_i	#	x_i	y_i
1	0.0	2.51	7	0.3	0.77	13	0.6	0.27	19	0.9	0.11
2	0.05	2.04	8	0.35	0.64	14	0.65	0.23	20	0.95	0.1
3	0.1	1.67	9	0.4	0.53	15	0.7	0.2	21	1.0	0.09
4	0.15	1.37	10	0.45	0.45	16	0.75	0.17	22	1.05	0.08
5	0.2	1.12	11	0.5	0.38	17	0.8	0.15	23	1.1	0.07
6	0.25	0.93	12	0.55	0.32	18	0.85	0.13	24	1.15	0.06

автором книги было показано, что решение задачи для 3 экспонент требует проводить вычисления с числами, значительно меньшими погрешности измерений. Для двух экспонент было найдено следующее решение:

$$f_1(x) = 2.202e^{-4.45x} + 0.305e^{-1.58x}.$$

Используя данные таблицы 4, мы провели серию расчетов для 2, 3 и 4 экспонент. Результаты этих расчетов:

$$f_2(x) = 2.084e^{-4.61x} + 0.426e^{-1.85x},$$

$$f_3(x) = 2.084e^{-4.61x} + 0.425e^{-1.85x} + 4.571 \cdot 10^{-12} \cdot e^{-0.52x},$$

$$f_4(x) = 2.084e^{-4.61x} + 0.426e^{-1.85x} + 6.685 \cdot 10^{-11} \cdot e^{-3.93x} + 1.752 \cdot 10^{-6} \cdot e^{-3.8 \cdot 10^3 \cdot x}.$$

Отметим, что в трех этих случаях первые два слагаемых практически совпадают, а оставшиеся вносят малый вклад в значение функции, который находится далеко за пределами погрешности измерений. Например,

$$\max_{[0,1.15]} \{f_2(x) - f_3(x)\} \sim 10^{-12}, \quad (7.2)$$

$$\max_{[0,1.15]} \{f_2(x) - f_4(x)\} \sim 10^{-6}. \quad (7.3)$$

Значение распознающего функционала для f_2, f_3, f_4 практически не отличается. В дальнейшем, из трех полученных решений будем рассматривать f_2 . Для того чтобы сравнить решения, будем рассматривать 3 характеристики:

1. Максимальное абсолютное отклонение от данных

$$E_{\infty}(f) = \max_{i=1,\dots,n} \{|f(x_i) - y_i|\}, \quad (7.4)$$

2. Стандартное отклонение от данных

$$RMSE(f) = \left(\frac{1}{n} \sum_{i=1}^n (f(x_i) - y_i)^2 \right)^{\frac{1}{2}}, \quad (7.5)$$

3. Значение распознающего функционала Tol .

Результаты представлены в Таблице 5:

Из результатов расчетов видим, что решение полученное методом макси-

Таблица 5: Сравнение функциональных зависимостей для примера Ланцоша

#	E_{∞}	$RMSE$	Tol	\mathcal{RC}
f	0.0043	0.0024	0.00066	0.132
f_L	0.0066	0.0025	-0.00162	—
f_2	0.0033	0.0023	0.00168	0.336

мум совместности обладает наименьшими ошибками. Решение, полученное Ланцошом, отвечает отрицательному значению распознающего функционала. Это говорит о том, что аппроксимирующая кривая не проходит через интервальные данные. Резюмируя, можно сделать вывод, что на данном модельном примере метод, предложенный в текущей работе, предпочтительнее.

Стоит отметить, что значение характеристики $\mathcal{RC} = 0.336$ для найденного решения. Это демонстрирует нам, что это решение может оказаться не самым оптимальным из возможных.

8. Усовершенствованные постановки

В ходе расчета реальных данных часто возникает потребность в модификации восстанавливаемой зависимости. В данном разделе будут рассмотрено несколько таких обобщений.

8.1. Отрицательные коэффициенты при экспонентах

В практике применения технологии нам понадобилось приближать немотонную функцию, обладающую одним максимумом и эквивалентную полиному из экспонент на бесконечности. Для такой ситуации была проведена попытка отказаться от требования положительности коэффициентов a_j . В рамках предложенного нами подхода это эквивалентно отказу от применения штрафной функции для коэффициентов a_j .

8.2. Дополнительные линейные слагаемые

Рассматривается задача восстановления функциональной зависимости вида

$$y = \sum_{j=1}^m a_j e^{-b_j x} + \sum_{j=1}^p c_j f_j(x),$$

где a_j , b_j , c_j — вещественные константы и $b_j \geq 0$, по набору измерений

$$(\mathbf{x}_i, \mathbf{y}_i), \quad i = 1, \dots, n,$$

где $\mathbf{x}_i = [\underline{\mathbf{x}}_i, \overline{\mathbf{x}}_i]$, $\mathbf{y}_i = [\underline{\mathbf{y}}_i, \overline{\mathbf{y}}_i]$ — вещественные интервалы.

Определение 11. Решением задачи восстановления зависимости вида полиномов из экспонент с линейной добавкой будем называть такой набор параметров $(a, b, c) \in \mathbb{R}^m \times \mathbb{R}_+^m \times \mathbb{R}^p$, что верно

$$\sum_{j=1}^m a_j e^{-b_j \mathbf{x}_i} + \sum_{j=1}^p c_j f_j(\mathbf{x}_i) \subseteq \mathbf{y}_i, \quad i = 1, \dots, n,$$

где f_j — заданные интервальные расширения функций, не зависящие от параметров.

Тогда

$$\begin{aligned} \sum_{j=1}^m a_j e^{-b_j \mathbf{x}_i} + \sum_{j=1}^p c_j \mathbf{f}_j(\mathbf{x}_i) &\subset \mathbf{y}_i \\ \Updownarrow \\ \text{rad } \mathbf{y}_i - \sum_{j=1}^m a_j \text{rad } e^{-b_j \mathbf{x}_i} - \sum_{j=1}^p c_j \text{rad } \mathbf{f}_j(\mathbf{x}_i) \\ - \left| \text{mid } \mathbf{y}_i - \sum_{j=1}^m a_j \text{mid } e^{-b_j \mathbf{x}_i} - \sum_{j=1}^p c_j \text{mid } \mathbf{f}_j(\mathbf{x}_i) \right| &\geq 0, \end{aligned}$$

Для случая монотонной функции f вид распознающего функционала можно явно выписать следующим образом. В качестве интервального расширения функции f возьмем интервальную функцию $\mathbf{f}(\mathbf{x}) = [\min\{f(\underline{\mathbf{x}}), f(\overline{\mathbf{x}})\}, \max\{f(\underline{\mathbf{x}}), f(\overline{\mathbf{x}})\}]$.

$$\text{rad } \mathbf{f}(x_i) = \begin{cases} \frac{1}{2}(f(\overline{\mathbf{x}}_i) - f(\underline{\mathbf{x}}_i)), & \text{для монотонно возрастающих функций} \\ \frac{1}{2}(f(\underline{\mathbf{x}}_i) - f(\overline{\mathbf{x}}_i)), & \text{для монотонно убывающих функций} \end{cases}$$

или что тоже самое:

$$\text{rad } \mathbf{f}(x_i) = \frac{1}{2} \left| f(\overline{\mathbf{x}}_i) - f(\underline{\mathbf{x}}_i) \right|.$$

Из этого следует что вид распознающего функционала в случае монотонной функции f имеет вид

$$\begin{aligned} \text{Tol}(\mathbf{x}, \mathbf{y}, a, b, c) &= \\ &= \min_{i=1, \dots, n} \left\{ \text{rad } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_i} - e^{-b_j \overline{\mathbf{x}}_i}) - \frac{1}{2} \sum_{j=1}^p c_j \left| f_j(\overline{\mathbf{x}}_i) - f_j(\underline{\mathbf{x}}_i) \right| \right. \\ &\quad \left. - \left| \text{mid } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_i} + e^{-b_j \overline{\mathbf{x}}_i}) - \frac{1}{2} \sum_{j=1}^p c_j (f_j(\overline{\mathbf{x}}_i) + f_j(\underline{\mathbf{x}}_i)) \right| \right\} \end{aligned}$$

А почти-градиенты в точках непрерывности имеют вид:

$$\begin{aligned}\frac{\partial \text{Tol}(\mathbf{x}, \mathbf{y}, a, b, c)}{\partial a_k} &= \text{rad } e^{-b_k \mathbf{x}_l} + \Psi \text{mid } e^{-b_k \mathbf{x}_l}, \\ \frac{\partial \text{Tol}(\mathbf{x}, \mathbf{y}, a, b, c)}{\partial b_k} &= \frac{1}{2} a_k (\underline{\mathbf{x}}_l e^{-b_k \underline{\mathbf{x}}_l} - \overline{\mathbf{x}}_l e^{-b_k \overline{\mathbf{x}}_l}) - \frac{1}{2} a_k \Psi (\underline{\mathbf{x}}_l e^{-b_k \underline{\mathbf{x}}_l} + \overline{\mathbf{x}}_l e^{-b_k \overline{\mathbf{x}}_l}), \\ \frac{\partial \text{Tol}(\mathbf{x}, \mathbf{y}, a, b, c)}{\partial c_k} &= \frac{1}{2} |f_k(\overline{\mathbf{x}}_l) - f_k(\underline{\mathbf{x}}_l)| + \frac{1}{2} \Psi (f_k(\overline{\mathbf{x}}_l) + f_k(\underline{\mathbf{x}}_l)),\end{aligned}$$

где l определяется условием $g_l = \min_{i=1, \dots, n} g_i$.

$$\Psi = \text{sgn} \left(\text{mid } \mathbf{y}_i - \frac{1}{2} \sum_{j=1}^m a_j (e^{-b_j \underline{\mathbf{x}}_l} + e^{-b_j \overline{\mathbf{x}}_l}) - \frac{1}{2} \sum_{j=1}^p c_j (f_j(\overline{\mathbf{x}}_l) + f_j(\underline{\mathbf{x}}_l)) \right)$$

Таким образом, для того чтобы расширить класс восстанавливаемых зависимостей, нам необходимо немного изменить расчет значения распознающего функционала и расчет его почти-градиентов. Сама технология поиска экстремума остается неизменной.

9. Примеры

В данном разделе рассматривается несколько примеров восстановления функциональных зависимостей вида полиномов из экспонент в усовершенствованной постановке.

В следующих нескольких примерах будем тестировать работу на модельных данных для известных функций. В каждом примере для фиксированной функции сгенерируем набор данных следующим образом:

1. зафиксируем набор входных переменных x
2. входные переменные считаем точными
3. посчитаем значения $\text{mid } y_i = f(x_i)$
4. округлим до 4 знака после запятой
5. возьмем $\text{rad } y_i = 0.07$ для всех измерений

9.1. Отрицательные коэффициенты

Попытаемся восстановить функцию

$$f(x) = 2e^{-x} - 1.5e^{-3x} \quad (9.1)$$

по набору данных:

Таблица 6: Данные для степенного слагаемого

x	0.2	0.35	0.9	1.2	1.9	2.4
\underline{y}	0.7442	0.8145	0.6423	0.4914	0.2241	0.1103
\overline{y}	0.8842	0.9545	0.7823	0.6314	0.3641	0.2503

В ходе расчетов было найдено решение:

$$f^*(x) = 1.9993e^{-x} - 1.4995e^{-3.0014x}$$

Этому решению соответствует $\text{Tol} = 0.06998$ и $\mathcal{RC} = 0.9997$. Настолько близкое значение \mathcal{RC} к 1 демонстрирует нам, решение даже теоретически не может быть улучшено. Это согласуется с тем, что найденное нами решение отличается от исходной функции четвертым знаком.

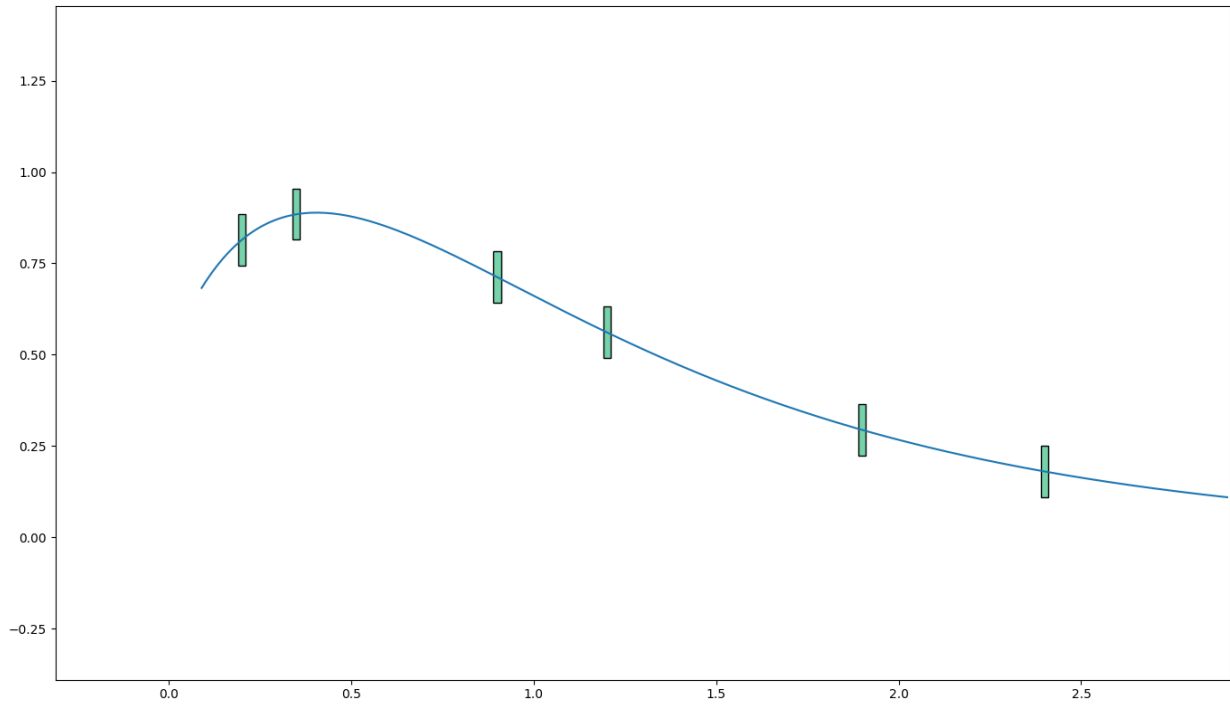


Рис. 6: Восстановленная зависимость с отрицательными a_j

9.2. Степенные слагаемые

Рассмотрим, как метод максимума совместности помогает отделить степенную функцию от показательной. Для этого рассмотрим следующую функцию:

$$f(x) = 0.01x^{-2.5} + 0.8e^{-2x}$$

Способом описанным выше получим набор данных на Таблице 7:

Таблица 7: Данные для степенного слагаемого

x	0.12	0.2	0.6	0.7	0.9	1.4	1.6	1.9	2.4
y	2.564	1.0253	0.2068	0.1517	0.0753	-0.017	-0.0343	-0.0501	-0.0623
\bar{y}	2.704	1.1653	0.3468	0.2917	0.2153	0.123	0.1057	0.0899	0.0777

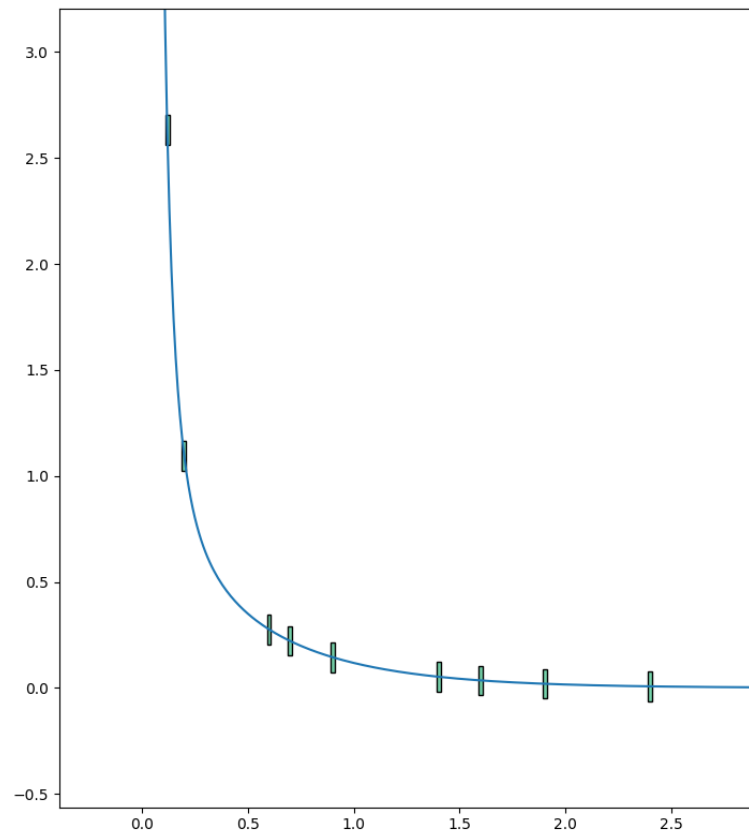


Рис. 7: Результат работы алгоритма на данных с дополнительным степенным слагаемым

Для степенной функции показатель степени, равный -2.5 , считаем известным. Он реально дается в геофизических задачах из соображений асимптотической эквивалентности. То есть, ищем функцию в виде:

$$f^*(x) = ae^{-bx} + cx^{-2.5}$$

Используя метод максимума совместности, получаем решение:

$$f^*(x) = 0.7972e^{-1.9967x} + 0.01x^{-2.5}$$

Этому решению соответствует значение распознающего функционала $\text{Tol} = 0.0696$, значение $\mathcal{RC} = 0.994$. Это значит, что найденное нами решение можно считать оптимальным. График найденного решения и данные представлены на Рис. 7.

9.3. Функция с точками перегиба

В следующем примере в качестве модельной функции, взята функция

$$f(x) = 0.8e^{-(x-1)^2} + 1.8e^{-2x}.$$

По ней были получены следующие данные:

Таблица 8: Данные с изменением вогнутости

x	0.12	0.2	0.6	0.7	0.9	1.4	1.6	1.9	2.4
y	1.7147	1.5584	1.1539	1.105	1.0196	0.7212	0.5615	0.3262	0.0575
\bar{y}	1.8547	1.6984	1.2939	1.245	1.1596	0.8612	0.7015	0.4662	0.1975

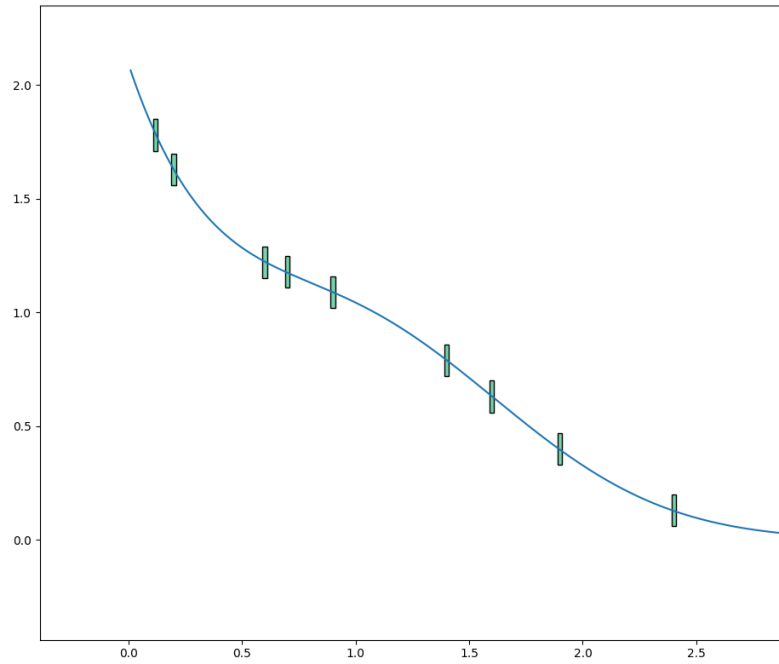


Рис. 8: Результат работы алгоритма на данных с меняющийся вогнутостью

Решение со значением распознающего функционала 0.067 и $\mathcal{RC} = 0.957$ имеет вид:

$$f^*(x) = 0.8037e^{-(x-1)^2} + 1.8024e^{-2.0153x}$$

Данные и результаты работы алгоритма представлены на Рис. 8

9.4. Немонотонная функция

В следующем примере модельной является следующая функция:

$$f(x) = 10x^{0.7}e^{-x} + 1.8e^{-2x} + 0.7e^{-3x}$$

Таблица 9: Данные с немонотонной функцией

x	0.12	0.2	0.35	0.9	1.2	1.9	2.4
\underline{y}	3.8448	4.1745	4.4482	4.0512	3.5344	2.3167	1.6197
\overline{y}	3.9848	4.3145	4.5882	4.1912	3.6744	2.4567	1.7597

Ищем решения в виде:

$$f(x) = a_1e^{-b_1x} + a_2e^{-b_2x} + cx^{0.7}e^{-x},$$

приводит нас к решению:

$$f^*(x) = 10.0089x^{0.7}e^{-x} + 2.0054e^{-2.0677x} + 0.4920e^{-3.1452x}$$

Этому решению соответствует значение распознающего функционала 0.0696 и относительной совместностью 0.994.

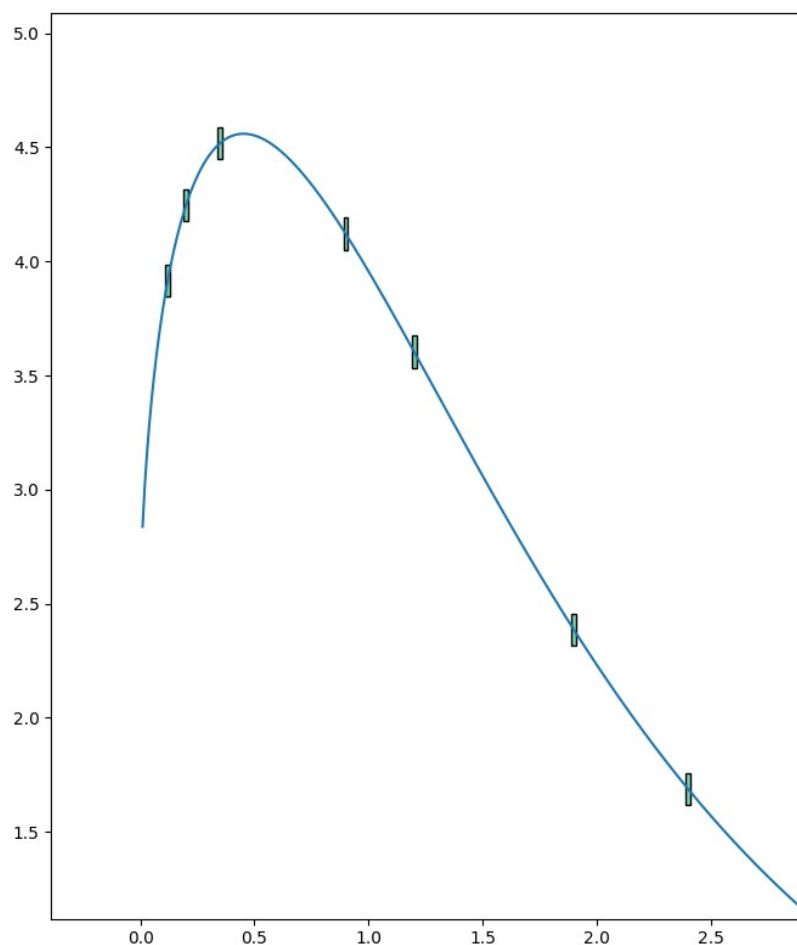


Рис. 9: Результат работы алгоритма на данных, имеющих локальный максимум

10. Заключение

В данной работе описан подход к решению задачи восстановления зависимостей вида полиномов из экспонент с возможностью учитывать дополнительные линейные слагаемые по интервальным данным. Продемонстрированы результаты численных экспериментов, подтверждающих работоспособность реализации программного обеспечения. Мы обнаружили «эффект отсутствия роста точности» по мере роста сложности. Появилась гипотеза, что любой монотонный полином из экспонент можно аппроксимировать небольшим числом экспонент с достаточно малой ошибкой.

С реализацией Python можно ознакомиться по адресу [14]. После более тщательного тестирования, программы будут включены в библиотеку интер-

вальных вычислений IntvalPy [14]. По теоретическим аспектам предлагаемого подхода опубликована статья [7].

В работе была введена новая характеристика *отношение совместности*. Она позволяет оценивать совместность модели и данных, а также дает способ валидировать нахождение наилучшего решения.

Важной проблемой остается разработка и реализация алгоритмов для построения внутренних и внешних оценок допускового множества решений системы уравнений и неравенств (2.1), (2.2). Это позволит нам находить не только кривую, наиболее совместную с данными, но и множество совместных с данными кривых. Это задача видится вычислительно сложной.

Тот факт, что нет существенного улучшения в совместности оценки при росте количества экспонент, может значительно упростить и ускорить решение нашей задачи.

Результаты, полученные в работе, могут быть использованы для решения практических задач, связанных с обработкой экспериментальных данных для соответствующих математических моделей или обобщены на более широкие классы задач.

Автор благодарен Владимиру Ивановичу Ерохину за внимание к работе и ценные комментарии.

6. Список литературы

- [1] **Баженов А. Н., Жилин С. И., Кумков С. И., Шарый С. П.** Обработка и анализ интервальных данных. Ижевск–Москва: Издательство «ИКИ»; 2024: 355.
- [2] **Шарый С. П.** Конечномерный интервальный анализ. Новосибирск: Издательство «XYZ», 2022. 653 с. URL: <http://www.nsc.ru/interval/Library/InteBooks/SharyBook.pdf>
- [3] **Ланцош К.** Практические методы прикладного анализа. Москва: Государственное издательство физико-математической литературы, 1961. 524 с.
- [4] **Жолен Л., Кифер М., Дидри О., Вальтер Э.** Прикладной интервальный анализ. – Москва-Ижевск: Издательство «РХД», 2005
- [5] International Online Seminar on Interval Methods in Control Engineering
Доступен на: <https://www.interval-methods.de/seminars>
- [6] **Зюбин В. Е., Петухов А. Д., Розов А. С.** Автоматизация измерений мюонным скважинным плотномером // Промышленные АСУ и контроллеры, 2018, № 11, С. 35–41.
- [7] **Zvyagin M. A., Shary S. P.** Curve Fitting for Exponential Polynomials from Interval Data // IEEE 24th International Conference of Young Professionals in Electron Devices and Materials, June 2023, pp. 1880–1883
- [8] **Shary S. P.** Weak and strong compatibility in data fitting problems under interval uncertainty // Advances in Data Science and Adaptive Analysis. 2020. Vol. 12, No. 01. 2050002.
- [9] **Леонтьев А.Ф.** Последовательности полиномов из экспонент. Москва: Издательство «Наука», 1980. 386 с.

- [10] **С. П. Шарый.** Восстановление функциональных зависимостей по данным с интервальной неопределенностью // Информатика и системы управления, 2022, № 3. С. 130-142.
- [11] **Стеценко С. И. Нурминский Е. А.** Негладкий штраф и субградиентные алгоритмы для решения задачи проекции на политоп // Кибернетика и системный анализ. 2010. № 1. С. 59-63
- [12] **Шор Н. З., Стеценко С. И.** Квадратичные экстремальные задачи и недифференцируемая оптимизация — Институт кибернетики им. В.М. Глушкова. Киев: Наук. думка., 1989.
- [13] **Шор Н. З.** Методы минимизации недифференцируемых функций и их приложения — Институт кибернетики им. В.М. Глушкова. Киев: Наук. думка., 1979.
- [14] **Андросов А.С.** IntvalPy. Библиотека для языка Python Открытый программный репозиторий. Доступен на: <https://github.com/AndrosovAS/intvalpy> и [https://github.com/ZvyaginMA/ Curve-Fitting-for-Exponential-Polynomials-from-Interval-Data](https://github.com/ZvyaginMA/Curve-Fitting-for-Exponential-Polynomials-from-Interval-Data)
- [15] **Гилл Ф. Мюррей У. Райт М.** Практическая оптимизация — Москва: «Мир», 1985. 509 с.