Федеральное государственное бюджетное образовательное учреждение
высшего образования
"Тверской государственный университет"

На правах рукописи

Сорокина Ирина Владимировна

Методы оценки параметров возможностных распределений и их применение для прогнозирования неисправностей электрооборудования

Специальность 05.13.01 — «Системный анализ, управление
и обработка информации (в промышленности)»

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор физико-математических наук,
профессор Язенин А.В.

Тверь, 2018

Оглавление

Введение **5**

1. Исчисление возможностей и агрегирование информации на основе t-норм **13**

1.1. Возможностное пространство и его свойства 13
1.2. Нечёткие величины и их исчисление 15
1.2.1. Преобразования 18
1.2.2. Операции 19
1.2.3. Нечёткие числа (L, R)-типа 20
1.3. Агрегирование нечёткой информации на основе t-норм 21
1.3.1. Определения, основные свойства и разновидности t-норм 21
1.3.2. Понятие аддитивного генератора 24
1.4. Подходы к интерпретации понятия функции распределения (принадлежности) и способам измерения её значений 27
1.4.1. Интерпретации понятия принадлежности 27
1.4.2. Подход, основанный на правдоподобии 29
1.4.3. Подход на основе случайных множеств 30
1.4.4. Интерпретация принадлежности как меры сходства объектов 31
1.4.5. Ценностный подход 32
1.4.6. Подход на основе теории меры 33
1.5. Выводы по первой главе 34

2. Оценка параметров возможностных распределений в одномерном случае **37**

2.1. Постановка задачи 37
2.2. Существующие методы оценивания параметров 39
2.2.1. Оценка параметров возможностных распределений в одномерном случае с использованием информации об уровне риска её ошибочного определения 39
2.2.2. Максиминная μ/E оценка параметров одномерных возможностных распределений 41
2.3. Сравнительный анализ рассмотренных методов 43
2.4. Выводы по второй главе 45

3. Оценка параметров многомерных возможностных распределений при различных t-нормах **47**

3.1. Оценка параметров многомерных распределений минисвязанных возможностных величин 47

3.1.1. Метод нахождение оценки параметров многомерных возможностных распределений при заданном уровне возможности 47

3.1.2. Максиминная μ/E оценка параметров многомерных возможностных распределений 51

3.1.3. Свойства μ/E оценки 55

3.2. Оценка параметров многомерных возможностных распределений в случае Архимедовой t-нормы 59

3.2.1. Нахождение оценки параметра **c** при фиксированной матрице A 59

3.2.2. Оценка параметра A 64

3.3. Выводы по третьей главе 66

4. Разработка и тестирование алгоритмов построения нечётких моделей с применением разработанного математического аппарата **69**

4.1. Применение методов оценки параметров возможностных распределений в задачах обучения машин нечёткого вывода 69

4.1.1. Понятие нечёткого правила 69

4.1.2. Описание системы нечёткого вывода 70

4.1.3. Алгоритм проектирования систем нечёткого вывода . . . 71

4.1.4. Аппроксимация нелинейной функции одной переменной 73

4.1.5. Задача прогнозирования значений нестационарного временного ряда 75

4.1.6. Усовершенствование алгоритма обучения машин нечёткого вывода 75

4.2. Двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе 80

4.3. Слияние алгоритма обучения машин нечёткого вывода с алгоритмом AdaBoost 83

4.4. Выводы по четвёртой главе 87

5. Программный комплекс анализа данных на основе возможностного моделирования **89**

5.1. Реализация алгоритмов обучения нечёткого контроллера в среде Matlab 89

5.2. Реализация алгоритмов обучения нечёткого контроллера на языке программирования Python 90

5.2.1. Прогнозирование неисправностей вагонного электрооборудования на железнодорожном транспорте 90

5.2.1.1. Постановка задачи 90

5.2.1.2. Проведение первичного анализа данных 93

5.2.1.3. Апробация алгоритма 94

5.2.2. Идентификация состояния воздушного пространства . . 97

5.2.3. Применение алгоритма Adaboost совместно с μ/E алгоритмом обучения машин нечёткого вывода к задаче классификации и регрессии 98

5.2.3.1. Апробация на типовой задаче классификации . 98

5.2.3.2. Апробация работы слияния алгоритма μ/E и Adaboost на задаче регрессии 99

5.3. Выводы по пятой главе 101

Заключение **103**

Список литературы **106**

Приложения **116**

А Одно из собственных чисел матрицы A 116

Б Востановленное двумерное нормальное распределение 122

В Восстановленное распределение, порождённое линейной функцией 123

Г Библиотека программ для прогнозирования 124

Д Акты о внедрении результатов диссертационной работы 133

Е Свидетельства о государственной регистрации программ для ЭВМ 136

Введение

Актуальность

В задачи, поставленные ОАО “РЖД” в “Стратегии развития железнодорожного машиностроения до 2030 года”, входит создание комплексных систем диагностики и прогнозирования технического состояния инфраструктуры подвижного состава.

Внедрение автоматизированных систем выявления и прогнозирования неисправностей позволяет одновременно снизить расходы на обслуживание и ремонт электрооборудования и увеличить его безопасность. Раннее выявление повреждённых частей позволяет сократить их воздействие на другие механизмы.

Разработка новых подходов к прогнозированию неисправностей позволит повысить надёжность производимого электрооборудования и оперативность реагирования.

Проанализировав литературу, посвящённую решению данной проблемы, можно сделать вывод о необходимости разработки такой системы с применением современных методов обработки информации.

Для создания систем подобного рода могут быть использованы машины нечёткого вывода, обучение которых может производиться на основе аппарата теории неопределённостей. Данный инструмент отлично зарекомендовал себя для решения задач интеллектуального анализа данных.

Тем не менее, можно отметить, что распространённые в настоящее время алгоритмы обучения машин нечёткого вывода опираются на интуитивный подход при определении параметров нечётких распределений. Теоретической основной для решения этой задачи должны быть математически обоснованные методы оценки параметров нечётких распределений.

Однако, к настоящему времени такие методы были разработаны только для одномерных распределений, что существенно ограничивает возможность использования этих результатов при решении практических задач, в том числе при обучении машин нечёткого вывода.

Таким образом, можно ожидать, что разработка методов оценки параметров многомерных возможностых распределений и их внедрение в алгорит-

мы обучения позволит повысить качество работы машин нечёткого вывода и расширить круг практических задач, решаемых с их помощью, в том числе задачи прогнозирования неисправностей электрооборудования.

Этим определяется актуальность темы диссертации.

Обзор литературы

Теория нечётких множеств и связанная с ней теория возможностей активно развивалась профессором Калифорнийского университета Беркли Лотфи Заде. В своей работе "Fuzzy Sets" [97] (1965 год) он расширил классическое понятие множества, допустив, что характеристическая функция может принимать значения не только из множества $\{0, 1\}$, но и из отрезка $[0, 1]$, таким образом, определяя "степень принадлежности" элемента нечёткому множеству (fuzzy set).

Дальнейшее развитие теория получила в 1978 году, когда S. Nahmias в работе "Fuzzy Variables" [71] предложил "аксиоматическую базу, которая может являться основой для построения строгой теории возможностей". В данной работе было введно понятие нечёткой переменной (возможностной (нечёткой) величины) и несвязанности (unrelatedness) нечётких величин, также были аксиоматически определены бинарные операции над несвязанными нечёткими величинами. Таким образом, была заложена теореческая основа современного исчисления возможностей.

В 1981 году M.B. Rao и A. Rashed в статье "Some comments on fuzzy variables" заметили, что введённую Nahmias несвязанность корректнее называть минисвязанностью (min-relatedness). Эти работы можно считать основополагающими в теории возможностей.

Дальнейшие развитие теории нечетких множеств и ее применение для моделирования неопределенности привело к появлению и применению систем нечёткого вывода, которые анализируют входные значения в терминах лингвистических переменных, осуществляют их отображение во множество выходных значений на основе базы нечётких правил. Эта база может быть построена с применением методов оценки параметров возможностных распределений.

Данная тема получила развитие в статье W. Näther "On possibilistic inference" (1990) [72], в которой рассматриваются некоторые идеи получения оценок параметров. Одновременно Zhen-Yuan W. и Shou-Mei L. в статье

"Fuzzy linear regression analysis of fuzzy-valued variables" [102] предложили μ/E метод оценки параметров возможностных распределений, который далее получил развитие в работах W. Xizhao и H. Minghu [89], [90]. W. Xizhao и H. Minghu помимо нахождения оценки для одномерных возможностных распределений в случае T-нормы минимум одни из немногих определили свойства полученной оценки.

Badard R. в [27] рассматривал "минимальную оценку нечеткости" ("minimum fuzziness estimator") для нечётких моделей и изучал сходимость этих оценок. Dishkant H. в статье "About membership functions estimation" [39] изучал оценку параметров нечетких переменных в значении Л. Заде. Cai K.-Y. [36] занимался изучением оценки параметров для нормальных нечётких переменных, основаваясь на понятиях S. Nahmias. На основе его работ D.H.Hong в работе "Parameter estimations of mutually T-related fuzzy variables" [56] оценивал параметры обобщённых T-связанных нечётких величин.

Проанализировав существующую литературу по теме оценки параметров возможностных распределений можно сделать вывод о том, что тема активно развивается, получены оценки в одномерном случае, но ряд вопросов остаётся открытым. Одним из таких вопросов является исследование методов оценки параметров в многомерном случае.

Что касается задачи прогнозирования неисправностей электрооборудования применительно к железнодорожным вагонам, можно отметить, что большинство исследований и существующих систем направлено на анализ состояния рельсов, колёсных пар и подвески железнодорожных вагонов [52], [67], [87]. Это не удивительно, так как данные системы оказывают первостепенное влияние на безопасность железнодорожных перевозок. В таких системах ведётся обработка сравнительно небольшого числа аналоговых сигналов высокой частоты, собранных с помощью дополнительных датчиков. Анализ данных осуществляется с использованием Фурье [86] или вейвлет-преобразований [59] для начальной обработки данных.

Характер данных, доступных в современных системах управления электрооборудованием пассажирскими вагонами, существенно отличается. Число сигналов достигает нескольких сотен, большинство из них является бинарными и редко изменяется. В таких условиях применение методов спектрального анализа не является оправданным. Поэтому, прямой перенос имеющегося

опыта на другие системы вагонов невозможен. Необходима разработка новых методов поиска и выявления признаков нештатных режимов работы оборудования.

Этим и определяются цель и задачи диссертационного исследования.

Цель работы

Целью данной работы повышение надёжности работы электрооборудования на основе применения методов оценки параметров возможностных распределений.

Данная цель предполагает разработку методов определения параметров многомерных распределений, ориентированных на использование для обучения машин нечёткого вывода, которые могут быть применены на практике при решении задач прогнозирования неисправностей электрооборудования.

Основные задачи

Для достижения поставленной цели необходимо решить следующие задачи:

- рассмотреть существующие методы построения оценки параметров возможностных распределений в одномерном случае;
- разработать методы оценивания параметров возможностных распределений в многомерном случае для сильнейшей и Архимедовых t-норм;
- определить свойства полученных оценок;
- разработать алгоритм обучения машины нечёткого вывода на основе метода оценки параметров многомерных возможностных распределений;
- реализовать библиотеку программ для восстановления параметров многомерных возможностных распределений;
- провести сравнение разработанного алгоритма с существующими аналогами с помощью численных экспериментов;
- разработать алгоритмы структурно-параметрической идентификации сложных систем на основе объединения нечёткого вывода с эффективными методами машинного обучения;

- на основе разработанных алгоритмов реализовать прототипы систем прогнозирования, ориентированные на применение в промышленности.

Методология и методы исследования

В работе используются методы теории возможностей, системного анализа, общей теории систем, моделирования систем, теории управления, нейронные сети, методы оптимизации, а также технологии сбора информации, проведение численных экспериментов, разработки и испытания опытного образца. Программный комплекс реализован на языке программирования Python.

Полученные реультаты диссертационной работы проверяются с помощью проведения вычислительных экспериментов и применения результатов исследований в условиях реальной проектной деятельности.

Научная новизна

Научная новизна заключается в разработке комплекса методов оценки параметров многомерных возможностных распределений:

1. Оценка параметров многомерных распределений минисвязанных возможностных величин:

   1.1. Метод нахождение оценки параметров многомерных возможностных распределений при заданном уровне возможности;

   1.2. Максиминная оценка параметров многомерных возможностных распределений;

2. Оценка параметров многомерных возможностных распределений в случае Архимедовых t-норм.

Также научную новизну составляет разработка алгоритмов обучения машин нечёткого вывода, алгоритма идентификации сложных систем, основанного на нейронных сетях и нечётком выводе, и алгоритма идентификации сложных систем, использующего бустинг нечётких контроллеров.

Практическая значимость

Предложенный алгоритм обучения машин нечёткого вывода обеспечивает более высокую точность по сравнению с существующими аналогами, что создаёт предпосылки для повышения эффективности функционирования различных систем управления и анализа данных, в том числе для повышения надёжности работы электрооборудования на предприятиях вагоностроения и электронной промышленности.

На основе предложенных алгоритмов структурно-параметрической идентификации в диссертации разработаны:

- система анализа данных для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте;
- система прогнозирования параметров прозрачных проводниковых материалов.

Основные положения, выносимые на защиту

На защиту выносятся следующие результаты диссертационной работы:

1. Методы оценки параметров многомерных возможностных распределений.
2. Алгоритм обучения машин нечёткого вывода.
3. Библиотека программ для восстановления параметров многомерных нечётких распределений.
4. Алгоритм идентификации сложных систем, основанный на нейронных сетях и нечётком выводе.
5. Алгоритм идентификации сложных систем, использующий бустинг нечётких контроллеров.
6. Система прогнозирования свойств прозрачных проводниковых материалов, использующая алгоритм Adaboost совместно с разработанным алгоритмом обучения машин нечёткого вывода.
7. Двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте.

Достоверность результатов работы

Достоверность полученных в диссертационной работе результатов подтверждается строгостью проводимых математических обоснований при формулировании и доказательстве теорем, результатами численных расчётов, сравнительным анализом полученных в ходе модельных экспериментов результатов с известными.

По результатам диссертационной работы разработан и внедрен программный комплекс прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте в рамках совместных проектов РФФИ и РЖД № 12-07-13117-офи_м_РЖД и №13-07-13160-офи_м_РЖД.

Методы оценки параметров многомерных возможностных распределений использованы в ЗАО «РТИС ВКО» для разработки программного обеспечения по заказу Министерства обороны РФ.

Результаты диссертационной работы внедрены в учебный процесс на факультете прикладной математики и кибернетики Тверского годарственного университета в курсе "Анализ нечётких информационных систем".

Апробация результатов

Основные результаты работы докладывались на 6 научных конференциях:

- студенческая научно-практическая конференция факультета прикладной математики и кибернетики;
- интегрированные модели и мягкие вычисления в искусственном интеллекте. VIII-я Международная научно-техническая конференция;
- 20th East West Fuzzy Colloquium 2013. 20th Zittau Fuzzy Colloquium;
- комплексное использование и охрана подземных пространств: Междунар. науч.-практ. конф., посвящ. 100-летнему юбилею науч. и туристско-экскурсионной деятельности в Кунгурской Ледяной пещере и 100-летию со дня рожд. В.С. Лукина;
- гибридные и синергетические интеллектуальные системы: II международный Поспеловский симпозиум;

- ICFLS 2015: XIII International Conference on Fuzzy Logic Systems, Barcelona, Spain.

Личный вклад автора

Все основные научные результаты, разработка и написание программного обеспечения принадлежат лично автору.

Публикации

Основные результаты диссертации опубликованы в 11 печатных изданиях, 5 из которых в журналах, из перечня ВАК РФ, 6 — в тезисах докладов конференций. Получено 2 свидетельства о регистрации программных продуктов для ЭВМ.

Структура работы и ее содержание

Диссертация состоит из введения, пяти глав, заключения, списка литературы и 6 приложений. Общий объем диссертации, не включая приложения, 115 страниц текста с 30 рисунками и 7 таблицами. Список литературы содержит 104 наименования.

1. Исчисление возможностей и агрегирование информации на основе t-норм

1.1. Возможностное пространство и его свойства

В диссертационной работе используется ряд терминов, связанных с теорией возможностей. Для полносты изложения и понимания введём определение такого базового понятия, как мера неопределённости, а далее определим другие понятия теории возможностей и рассмотрим некоторые теоретические аспекты, которые потребуются нам в дальнейшем, следуя [40], [71], [100], [22] и др.

Пусть Γ — произвольное множество с элементами $\gamma \in \Gamma$. Пусть $C(\Gamma)$ — некоторое непустое множество его подмножеств, так что $\varnothing, \Gamma \in C(\Gamma)$. $P(\Gamma)$ — множество всех подмножеств множества Γ.

Определение 1. *Мера неопределённости есть функция множеств*

$$\sigma : C(\Gamma) \to [0, 1],$$

обладающая следующими свойствами:

$$\sigma\{\varnothing\} = 0, \sigma\{\Gamma\} = 1, \tag{1.1}$$

$$(\forall A, B \in C(\Gamma), A \subseteq B) \Rightarrow \sigma\{A\} \leqslant \sigma\{B\}. \tag{1.2}$$

Будем говорить, что мера неопределённости σ непрерывна снизу, если

$$\sigma\{\bigcup_{i=1}^{\infty} A_i\} = \lim_{i\to\infty} \sigma\{A_i\}$$

для любой возрастающей последовательности множеств

$$A_1 \subseteq A_2 \subseteq \ldots, A_i, \bigcup_{i=1}^{\infty} A_i \in C(\Gamma).$$

Будем называть меру неопределённости непрерывной, если она непрерывна сверху и снизу.

Рассмотрим некоторые свойства мер неопределённости.

Из аксиом (1.1), (1.2) следует ограниченность меры неопределённости

$$0 \leqslant \sigma\{A\} \leqslant 1, \forall A \in C(\Gamma).$$

Пусть множество $C(\Gamma)$ замкнуто относительно конечных объединений и пересечений. Тогда

$$\sigma\{A \cup B\} \geqslant \max\{\sigma\{A\}, \sigma\{B\}\},$$

$$\sigma\{A \cap B\} \leqslant \min\{\sigma\{A\}, \sigma\{B\}\}. \tag{1.3}$$

Предполагая выполнение равенства в (1.3), получим предельные случаи мер неопределённости, выделив два вида специальных мер.

Определение 2. *Мерой возможности называется функция*

$$\pi : P(\Gamma) \to [0, 1],$$

обладающая следующими свойствами:

$$\pi\{\varnothing\} = 0, \pi\{\Gamma\} = 1, \tag{1.4}$$

$$\pi\{\bigcup_{i \in I} A_i\} = \sup_{i \in I} \pi\{A_i\}, \forall A_i \in P(\Gamma), \forall I. \tag{1.5}$$

Определение 3. *Мерой необходимости называется функция*

$$\nu : P(\Gamma) \to [0, 1],$$

обладающая следующими свойствами:

$$\nu\{\varnothing\} = 0, \nu\{\Gamma\} = 1, \tag{1.6}$$

$$\nu\{\bigcap_{i \in I} A_i\} = \inf_{i \in I} \nu\{A_i\}, \forall A_i \in P(\Gamma), \forall I. \tag{1.7}$$

Определение 4. *Триплет* $(\Gamma, P(\Gamma), \pi)$ *называется возможностным пространством.*

Отметим важные свойства мер возможности и необходимости:

1. π и ν есть меры неопределённости.

2. $\nu = 1 - \pi\{A^c\}$, где A^c есть дополнение A.

   Данное свойство, с одной стороны, демонстрирует отношение двойственности между мерами необходимости и возможности, а с другой стороны, представляет собой численное выражение двойственности отношений таких понятий, как "необходимость"и "возможность": "нечто A является безусловно необходимым, когда противоположное ему A^c невозможно".

3. $\max\{\pi\{A\}, \pi\{A^c\}\} = 1$, $\min\{\nu\{A\}, \nu\{A^c\}\} = 0$.

4. $\pi\{A\} + \pi\{A^c\} \geqslant 1$, $\nu\{A\} + \nu\{A^c\} \leqslant 1$.

5. $\pi\{A\} \geqslant \nu\{A\} : \nu\{A\} \geqslant 0 \Rightarrow \pi\{A\} = 1, \pi\{A\} \leqslant 1 \Rightarrow \nu\{A\} = 0$.

6. Мера возможности непрерывна снизу, мера необходимости непрерывна сверху.

Очевидно, что мера возможности непрерывна, если она непрерывна сверху, мера необходимости непрерывна, если она непрерывна снизу.

1.2. Нечёткие величины и их исчисление

Определим понятие возможностной переменной, придерживаясь схемы, предложенной Stephen Nahmias [71] :

Определение 5. *Возможностной величиной (переменной) называется функция* $Z : \Gamma \to \mathbb{R}^1$. *Функция* $\mu_Z : \mathbb{R}^1 \to [0, 1]$ *называется распределением возможностной величины* Z *и определяется следующим образом:*

$$\mu_Z(z) = \pi\{\gamma \in \Gamma | Z(\gamma) = z\}, \forall z \in \mathbb{R}^1.$$

Здесь и далее $\mathbb{R}^1, \mathbb{R}^n$ *— множества действительных чисел соответствующих размерностей.*

Значение $\mu_Z(z)$ интерпретируется как “возможность того, что переменная Z может принять значение z”.

В двумерном случае возможностные переменные будут определяться согласно [8]. Данное определение также можно расширить и на случай многомерных возможностных величин.

Нетрудно показать, что

1. $0 \leqslant \mu_Z(z) \leqslant 1, \forall z \in \mathbb{R}^1$,

2. $\sup\limits_{z \in \mathbb{R}^1} \mu_Z(z) = 1$.

Следующая теорема показывает, что полунепрерывность сверху возможностных распределений задается аксиоматикой.

Теорема 1. *Пусть Z — непрерывная возможностная величина, определенная на возможностном пространстве $(\Gamma, P(\Gamma), \pi)$ с непрерывной мерой π. Тогда распределение μ_Z есть полунепрерывная сверху функция на $\mathbb{R}^1$ [95].*

Определение 6. *Носителем возможностной величины Z называется множество*

$$supp(Z) = \{z \in \mathbb{R}^1 | \mu_Z(z) > 0\}.$$

Определение 7. *Множеством α-уровня возможностной величины Z называется множество*

$$[Z]^\alpha = \{z \in \mathbb{R}^1 | \mu_Z(z) \geqslant \alpha\}, \alpha \in (0, 1].$$

Рассмотрим некоторые свойства возможностных величин, которые определяются их распределениями.

Определение 8. *Возможностная величина Z называется выпуклой, если ее функция распределения является квазивогнутой, то есть*

$$\mu_Z(\lambda z_1 + (1 - \lambda) z_2) \geqslant \min\{\mu_Z(z_1), \mu_Z(z_2)\} \forall z_1, z_2 \in \mathbb{R}^1, \forall \lambda \in [0, 1].$$

Определение 9. *Точка $a \in \mathbb{R}^1$ называется модальным значением возможностной величины Z, если $\mu_Z(a) = 1$.*

Определение 10. *Возможностная величина Z называется нормальной, если Z – выпуклая и существует хотя бы одно модальное значение Z.*

Определение 11. *Возможностная величина Z называется унимодальной, если Z – выпуклая и существует единственная точка, являющаяся модальным значением Z.*

Определение 12. *Под нечётким числом N_μ с функцией распределения $\mu(x)$ понимают выпуклую возможностную величину, функция распределения которой удовлетворяет следующим условиям:*

1. $\mu_\alpha = \{x : \mu(x) \geqslant \alpha\}$ *ограничена на замкнутом интервале для всех* $\alpha \in (0,1)$

2. $\mu_1 = \{x : \mu(x) = 1\} \neq \varnothing$

Лемма 1. *[95] Пусть* Z *— возможностная величина, определенная на возможностном пространстве* $(\Gamma, P(\Gamma), \pi)$*, распределение* μ_Z *квазивогнутое, имеющее конечный носитель,* π *непрерывна,* c *— константа,* $c > 0$*. Тогда*

$$[cZ]^\alpha = [cZ^-(\alpha), cZ^+(\alpha)],$$

где

$$Z^-(\alpha) = \inf_{z \in Z(\alpha)} z, Z^+(\alpha) = \sup_{z \in Z(\alpha)} z.$$

Лемма 2. *[95] Пусть* Z *— возможностная величина, определенная на возможностном пространстве* $(\Gamma, P(\Gamma), \pi)$*,* μ_Z *— неубывающая функция,* π *непрерывна,*
$|Z^-(\alpha)| < \infty$*,* c *— константа,* $c > 0$*. Тогда*

$$[cZ]^\alpha = [cZ^-(\alpha), +\infty).$$

Лемма 3. *Пусть* Z *— возможностная величина, определенная на возможностном пространстве* $(\Gamma, P(\Gamma), \pi)$*,* μ_Z *— невозрастающая функция,* π *непрерывна,* $|Z^+(\alpha)| < \infty$*,* c *— константа,* $c > 0$*. Тогда*

$$[cZ]^\alpha = (-\infty, cZ^+(\alpha)].$$

Введем понятие минисвязанности нечётких множеств. Пусть $A, B \in P(\Gamma)$. Следуя [76], будем называть множества A и B минисвязанными (относительно возможностной меры π), если:

$$\pi\{A \cap B\} = \min\{\pi\{A\}, \pi\{B\}\}.$$

Далее для того, чтобы ввести понятие минисвязанности для возможностных переменных, определим понятие совместной функции распределения n нечётких величин.

Определение 13. *Совместная функция распределения совокупности*

нечётких величин $Z_1, \ldots, Z_n$ *определяется следующим образом:*

$$\mu_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) = \pi\{\gamma \in \Gamma | Z_1(\gamma) = z_1, \ldots, Z_n(\gamma) = z_n\}, (z_1, \ldots, z_n) \in \mathbb{R}^n.$$

Определение 14. *Возможностные величины* $Z_1, \ldots, Z_n$ *называются минисвязанными, если для любого подмножества* $\{i_1, \ldots, i_k\}$ *множества* $\{1, \ldots, n\}$

$$\mu_{Z_{i_1},\ldots,Z_{i_k}}(z_1, \ldots, z_k) = \min\{\mu_{Z_{i_1}}(z_1), \ldots, \mu_{Z_{i_k}}(z_k)\}, \forall (z_1, \ldots, z_k) \in \mathbb{R}^k,$$

где $\mu_{Z_{i_s}}$ *— одномерные функции распределения возможностей.*

1.2.1. Преобразования

Приведём формулы преобразования для нахождения функций распределения возможностей от возможностных переменных.

Пусть $g : \mathbb{R}^1 \to \mathbb{R}^1$ и существует g^{-1}, Z — возможностная переменная, определенная на возможностном пространстве $(\Gamma, P(\Gamma), \pi)$. Тогда $Y = g(Z)$ также является возможностной переменной с распределением возможностей:

$$\mu_{g(Z)}(z) = \pi\{\gamma \in \Gamma | g(Z(\gamma)) = z\}.$$

Справедлива следующая формула [71]:

$$\mu_{g(Z)}(z) = \sup_{u:g(u)=z} \mu_Z(u), \forall z \in \mathbb{R}^1.$$

Ввиду того, что функция g имеет обратную, данная формула принимает вид:

$$\mu_{g(Z)}(z) = \mu_Z(g^{-1}(z)), \forall z \in \mathbb{R}^1. \tag{1.8}$$

Если существует $t \in \mathbb{R}^1$ такое, что $g^{-1}(t) = \varnothing$, то полагаем $\mu_{g(Z)}(t) = 0$.

С помощью формулы (1.8) определим следующие величины:

$-Z$ — противоположная величина. $\mu_{-Z}(z) = \mu_Z(-z)$.

λZ — преобразование растяжения/сжатия. $\mu_{\lambda Z}(z) = \mu_Z(z/\lambda), \lambda \neq 0$.

$1/Z$ — обратная величина. $\mu_{1/Z}(z) = \mu_Z(1/z)$.

Z^p — степень. $\mu_{Z^p}(z) = \mu_Z(z^{1/p}), p \neq 0$.

e^Z — экспонента. $\mu_{e^Z}(z) = \mu_Z(\log(z)), z \geqslant 0$.

Теперь пусть $g : \mathbb{R}^n \to \mathbb{R}^1, Z_1, \ldots, Z_n$ — возможностные переменные. Тогда $Z = g(Z_1, \ldots, Z_n)$ также является возможностной переменной с распределением возможностей

$$\mu_{g(Z_1,\ldots,Z_n)}(z) = \pi\{\gamma \in \Gamma | g(Z_1(\gamma), \ldots, Z_n(\gamma)) = z\}, z \in \mathbb{R}^1.$$

В случае, если переменные $Z_1, \ldots, Z_n$ являются минисвязанными, будет верна следующая формула [96]:

$$\mu_{g(Z_1,\ldots,Z_n)}(z) = \sup_{u_1,\ldots,u_n : g(u_1,\ldots,u_n)=z} \min_{1 \leqslant i \leqslant n} \mu_{Z_i}(u_i).$$

Теорема 2. *Пусть $Z_1, \ldots, Z_n$ — минисвязанные возможностные переменные, определенные на возможностном пространстве $(\Gamma, P(\Gamma), \pi)$, их распределения $\mu_{Z_1}, \ldots, \mu_{Z_n}$ квазивогнутые, $g : \mathbb{R}^n \to \mathbb{R}^1$ непрерывна. Тогда $Z = g(Z_1, \ldots, Z_n)$ есть возможностная переменная и для любого $\alpha \in (0, 1]$*

$$[Z]^\alpha = g([Z_1]^\alpha, \ldots, [Z_n]^\alpha).$$

1.2.2. Операции

Теорема 3. *[71] Пусть Σ — множество арифметических операций: $\Sigma = \{+, -, \times\}$, Z_1, Z_2 — минисвязанные возможностные переменные, заданные на возможностном пространстве $(\Gamma, P(\Gamma), \pi)$. Тогда возможностная величина $Z = Z_1 \star Z_2$, где $\star \in \Sigma$, определяется функцией распределения:*

$$\mu_Z(z) = \bigvee_{z_1 \star z_2 = z} (\mu_{Z_1}(z_1) \wedge \mu_{Z_2}(z_2)),$$

где $\vee, \wedge$ есть, соответственно, взятие минимума и максимума на отрезке $[0, 1]$.

Опираясь на данную теорему, можно получить следующие равенства:

1. $\mu_{Z_1+Z_2}(z) = \vee_{z_1}(\mu_{Z_1}(z_1) \wedge \mu_{Z_2}(z - z_1))$,
2. $\mu_{Z_1-Z_2}(z) = \vee_{z_1}(\mu_{Z_1}(z_1) \wedge \mu_{Z_2}(z_1 - z))$,
3. $\mu_{Z_1 \times Z_2}(z) = \vee_{z_1}(\mu_{Z_1}(z_1) \wedge \mu_{Z_2}(z/z_1)), \mathrm{supp}(Z) \not\ni 0$,

4. $\mu_{Z_1/Z_2}(z) = \vee_{z_1}(\mu_{Z_1}(z_1) \wedge \mu_{Z_2}(z \times z_1))$.

Следующая теорема может быть применена для нахождения модального значения суммы минисвязанных нечётких величин:

Теорема 4. *Пусть Z_1 и Z_2 — минисвязанные нечёткие величины с модальными значениями $M[Z_1]$ и $M[Z_2]$ соответственно. Тогда $M[Z_1] + M[Z_2]$ есть модальное значение нечеткой величины $Z_1 + Z_2$.*

Следует заметить, что в общем случае вычисление функции распределения возможностной величины $Z = Z_1 \star Z_2$ представляет определенные трудности.

1.2.3. Нечёткие числа (L, R)-типа

На практике зачастую применяются специальные классы нечетких переменных, одним из которых является класс нечетких чисел (L, R)-типа [10]:

Определение 15. *(L, R)-функциями (функциями представления формы) называют невозрастающие полунепрерывные сверху, определённые на неотрицательной части числовой прямой функции, обладающие следующими свойствами:*

1) $L(0) = R(0) = 1$,

2) $L(t), R(t) < 1$, $\forall t \in \mathbb{R}^1_+$,

3) $\lim\limits_{t \to +\infty} L(t) = \lim\limits_{t \to +\infty} R(t) = 0$,

где $\mathbb{R}^1_+ = \{t \in \mathbb{R}^1 : t \geqslant 0\}$.

Определение 16. *Возможностная переменная Z является возможностной величиной (L, R)-типа, если ее распределение имеет вид:*

$$\mu_Z(t) = \begin{cases} L\Big(\dfrac{m^- - t}{d^-}\Big), & t \leqslant m^-, \\ 1, & m^- \leqslant t \leqslant m^+, \\ R\Big(\dfrac{t - m^+}{d^+}\Big), & t \geqslant m^+, \end{cases}$$

где m^-, m^+ — границы интервала толерантности, $m^-, m^+ \in \mathbb{R}^1$, d^-, d^+ — коэффициенты нечёткости, $d^-, d^+ \in \mathbb{R}^1_>$, $\mathbb{R}^1_> = \{t \in \mathbb{R}^1 : t > 0\}$.

Коэффициенты нечёткости позволяют управлять "размытостью" возможностной величины. Нетрудно видеть, что при увеличении их значений нечёткость величины Z возрастёт, и наоборот.

Для унимодальных нечётких величин (L, R)– типа $m^- = m^+ = m$, где m — модальное значение.

Среди распределений (L, R)-типа наибольший практический интерес представляют следующие классы распределений:

1. Триангулярные распределения $m^+ = m^-, L(t) = R(t) = \max\{0, 1 - t\}$, $t \geqslant 0$,

2. Трапециевидные распределения $m^+ \neq m^-, L(t) = R(t) = \max\{0, 1 - t\}$, $t \geqslant 0$,

3. Нормальные распределения $m^+ = m^-, L(t) = R(t) = e^{-t^2}$, $t \geqslant 0$.

Следует заметить, что все классы возможностных величин, характеризующихся распределениями (L, R)-типа, замкнуты относительно операций сложения, вычитания и умножения на положительное число.

1.3. Агрегирование нечёткой информации на основе t-норм

1.3.1. Определения, основные свойства и разновидности t-норм

Агрегирование нечёткой информации в диссертационной работе основывается на использовании треугольных норм (t-норм) [60], [61]. Данный математический инструмент позволяет обобщить понятие t-связанности на совокупность возможностных величин. На основе t-норм строится соответствующее исчисление возможностей [76]. Рассмотрим основные понятия, связанные с t-нормами [28].

Определение 17. *Отображение* $T : [0, 1] \times [0, 1] \to [0, 1]$ *называется треугольной нормой (или t-нормой), если для любого* $x \in [0, 1]$ *оно обладает следующими свойствами:*

1) ограниченность: $T(1, x) = x$;

2) симметричность: $T(x, y) = T(y, x)$;

3) ассоциативность: $T(x, T(y, z)) = T(T(x, y), z)$;

4) монотонность: $T(w, y) \leqslant T(x, z)$, *если* $w \leqslant x$, $y \leqslant z$.

Среди наиболее используемых t-норм отметим следующие (рис. 1.1):

1. Операция взятия минимума: $T_M(x, y) = \min\{x, y\}$. Данную t-норму также называют сильнейшей.

2. Алгебраическое произведение: $T_P(x, y) = xy$.

3. T-норма Лукасевича: $T_L(x, y) = \max\{x + y - 1, 0\}$.

4. Слабейшая t-норма: $T_W(x, y) = \begin{cases} \min\{x, y\}, & \text{если } \max\{x, y\} = 1, \\ 0, & \text{иначе.} \end{cases}$

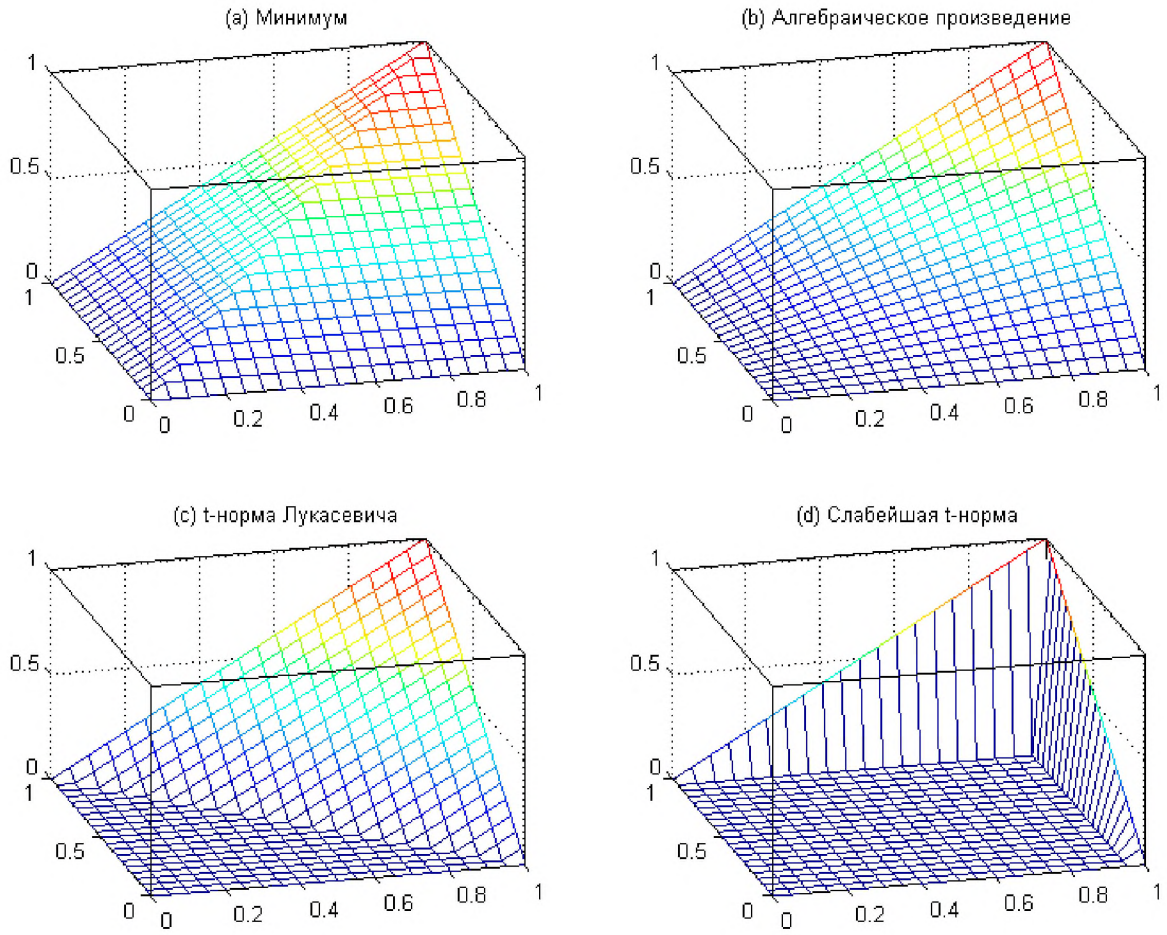


Рис. 1.1. Основные t-нормы

Определение 18. *T-норма* T_1 *слабее (сильнее)* t*-нормы* T_2, *если* $\forall x, y \in [0, 1]$

$$T_1(x, y) \leqslant T_2(x, y) \quad \big(T_1(x, y) \geqslant T_2(x, y)\big)$$

и существуют такие $x, y \in [0,1]$, *что*

$$T_1(x,y) < T_2(x,y) \quad \big(T_1(x,y) > T_2(x,y)\big).$$

Следующая теорема поясняет, почему сильнейшая и слабейшая t-нормы имеют такие называния.

Теорема 5. *Если* T *есть* t*-норма, то* $\forall x, y \in [0,1]$

$$T_W(x,y) \leqslant T(x,y) \leqslant T_M(x,y).$$

Другими словами, треугольные нормы T_W и T_M являются экстремальными. И не существует более слабой t-нормы, чем T_W, и более сильной, чем T_M. Также для приведённых t-норм справедливо следующее соотношение:

$$T_W(x,y) \leqslant T_L(x,y) \leqslant T_P(x,y) \leqslant T_M(x,y).$$

Замечание 1. *В случае, когда число аргументов больше двух,* t*-нормы минимума и алгебраического умножения не меняются. Т-норма Лукасевича и слабейшая* t*-норма будут выглядеть следующим образом:*

$$T_L(x_1,..,x_n) = \max\Big(\sum_{i=1}^{n} x_i - (n-1), 0\Big),$$

$$T_W(x_1,..,x_n) = \begin{cases} x_i, & \text{если } x_j = 1 \text{ для всех } j \neq i, \\ 0, & \text{иначе.} \end{cases}$$

Определение 19. *Т-норма* T *называется строгой, если она непрерывная и строго монотонна на* $(0,1] \times (0,1]$, *т.е.* $T(t,u) < T(t,v)$ *когда* $t > 0$ *и* $u < v$. *Т-норма* T *называется нульпотентной, если она непрерывная и существует*
$n \in \{1,2,...\}$ *такое, что* $T(\underbrace{a,..,a}_{n}) = 0$ *для любых* $a \in (0,1)$.

При освещении понятия t-нормы нельзя не рассмотреть класс Архимедовых t-норм [60], [28].

Определение 20. *Т-норма называется Архимедовой, если для каждой пары*
$(a,b) \in [0,1] \times [0,1]$ *выполняется:* $T(\underbrace{a,..,a}_{n}) < b$ *при* $n \in \{1,2,..,\}$.

Существует и другое, эквивалентное определение.

Определение 21. *T-норма называется Архимедовой, если* $\lim\limits_{n\to\infty} T_n(\underbrace{t,..,t}_{n}) = 0$ *при* $t \in (0,1)$.

Большинство Архимедовых t-норм являются непрерывными, существует всего несколько примеров, когда это свойство не выполняется. И учитывая, что свойство непрерывности играет важную роль, в дальнейшем будем рассматривать только непрерывные Архимедовы t-нормы.

Теорема 6. *Непрерывная t-норма* T *является Архимедовой тогда и только тогда, когда она строго возрастает на подмножестве* $\{(x,y) \in [0,1] \times [0,1] \mid T(x,y) > 0\}$.

Данное свойство называют свойством условного сокращения в [60]:

$$\forall x, y, z : T(x,y) = T(x,z) > 0, \text{ при } y = z.$$

Замечание 2. *Треугольные нормы* T_P *и* T_L *являются Архимедовыми t-нормами.*

Есть несколько причин, почему Архимедовы t-нормы представляют интерес:
a) они формируют плотное подмножество в множестве всех непрерывных t-норм;
b) они могут быть представлены с помощью аддитивных генераторов, понятие которых будет рассмотрено далее.

1.3.2. Понятие аддитивного генератора

Одним из отличительных свойств Архимедовых t-норм является то, что они могут быть представлены с помощью непрерывных аддитивных генераторов, использование которых позволяет упростить и свести вычисление t-нормы (как многомерной функции) к вычислению значения одномерного генератора.

Рассмотрим непрерывную строго убывающую функцию $g : [0,1] \to [0,\infty]$ с $g(1) = 0$. Обозначим через g^{-1} функцию, обратную к g. Заметим, что при $0 < a < \infty$ область допустимых значений $g^{-1} = [0,a]$ (рис. 1.2).

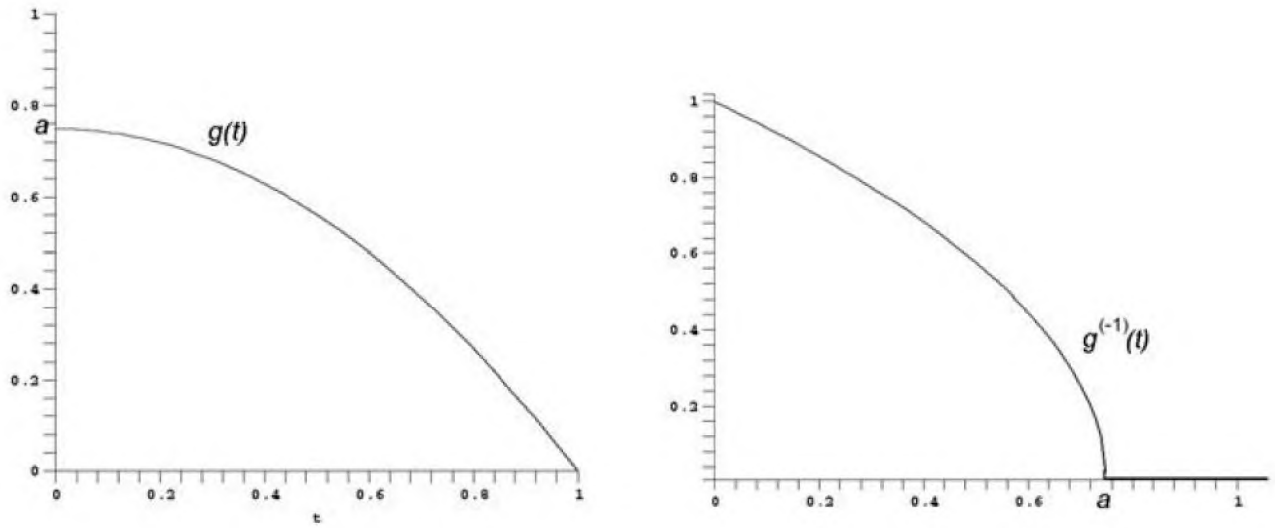


Рис. 1.2. Пример аддитивного генератора нульпотентной t-нормы(слева) и её псевдообратной функции(справа)

Для того, чтобы g^{-1} можно было подавать на вход любой неотрицательный аргумент, введём псевдообратную функцию $g^{[-1]}$:

$$g^{[-1]}(t) = \begin{cases} g^{-1}(t), & \text{если } t \in [0, a], a \in \mathbb{R}^1 \text{ и } g(0) = a, \\ 0, & \text{иначе.} \end{cases} \tag{1.9}$$

Замечание 3. *Псевдообратную функцию также можно определить следующим образом:*

$$g^{[-1]}(t) = sup\{z \in [0,1] \mid g(z) > t\} \text{ для всех } t \in [0, \infty].$$

Данный вид позволяет включить оба случая: $a < \infty$ *и* $a = \infty$.

Причина рассмотрения именно такой функции состоит в том, что любая непрерывная Архимедова t-норма может быть представлена с помощью непрерывного аддитивного генератора (рис. 1.3).

Теорема 7. *Пусть* T *– непрерывная Архимедова t-норма. Тогда она может быть представлена в следующем виде:*

$$T(x, y) = g^{[-1]}(g(x) + g(y)),$$

где $g : [0,1] \to [0, \infty]$, $g(1) = 0$ *является непрерывной строго убывающей функцией, называемой аддитивным генератором для* T *[28].*

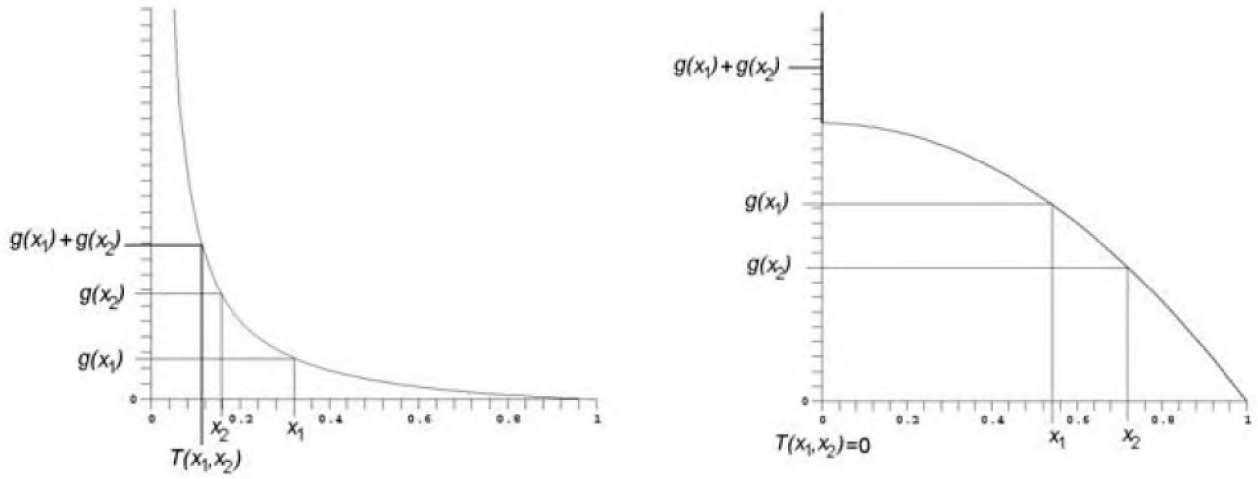


Рис. 1.3. Построение строгой(слева) и нульпотентной t-норм, используя аддитивный генератор

Заметим, что данное утверждение работает в две стороны – если t-норму можно представить через аддитивный генератор, то она является Архимедовой.

Замечание 4. *В случае большого числа аргументов (> 2), имеем:*

$$T(\mathbf{x}) = g^{[-1]}(g(x_1) + g(x_2) + ... + g(x_n)),$$

где $\mathbf{x} = (x_1, ..., x_n)$.

Замечание 5. *Некоторые t-нормы можно представить с помощью следующих аддитивных генераторов:*

1. *Для t-нормы Лукасевича* $g(t) = 1 - t$;

2. *Для t-нормы произведения* $g(t) = -ln(t)$.

Заметим также, что t-норма минимум не имеет аддитивного генератора (она не является Архимедовой).

Теорема 8. *Если* $g(t)$ – *аддитивный генератор для* T, *то* $cg(t), c > 0$ – *также является аддитивным генератором для* T.

Таким образом, одна t-норма может быть выражена через несколько аддитивных генераторов. Например, с помощью $g(t) = -2ln(t) = -ln(t^2)$ и $g(t) = -ln(t)$ можно представить t-норму произведения.

1.4. Подходы к интерпретации понятия функции распределения (принадлежности) и способам измерения её значений

После введения в 1965 году Л.Заде [97] понятия нечетких множеств одной из основных проблем стала интерпретация понятия функции распределения (принадлежности) и её измерения.

Существует большое количество точек зрения на источники появления нечёткости (размытости). Это приводит к различным определениям понятия функции распределения. Для более детального понимания ситуации рассмотрим существующиие походы к интерпретации и построению теоретических моделей для различных понятий функции распределения [33].

1.4.1. Интерпретации понятия принадлежности

При знакомстве с теорией нечётких множеств идея измерения принадлежности выглядит достаточно логичной, так как она является обобщением хорошо известной концепции: бинарная принадлежность множеству расширяется до степени принадлежности этому множеству. Однако, после размышлений возникает вопрос: “Как можно измерить степень принадлежности?” Этот вопрос рассматривался исследователями из разных дисциплин, включая многозначные логики [99], [49], [47]. Более 2000 лет назад Аристотель писал о “неопределенном значение истины” [26], но интерес к формальным аспектам многозначной логики действительно возрос только в начале 1900-х. Тем не менее, интерпретация понятия нескольких значений истинности остаётся открытым вопросом.

Как бы то ни было, для того, чтобы использовать нечёткие множества, необходимо ответить на следующие вопросы:

1. Что означает “измерить принадлежность”?

2. Как она измеряется?

3. Какие операции применимы к этому понятию?

Для ответа на первый вопрос необходимо выбрать одну из точек зрения на понятие нечёткости. Существует два основных направления мысли (рис. 1.4):

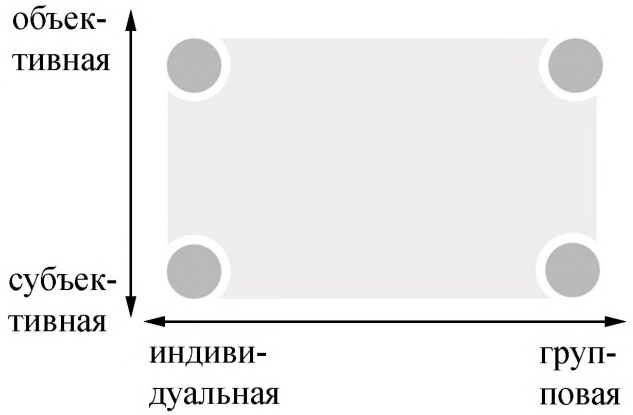


Рис. 1.4. Основные интерпретации понятия нечёткости

1. Нечёткость может рассматриваться как субъективное понятие или как объективное.

2. Понятие нечёткости может приписываться как отдельному индивиду, так и группе людей (сенсоров, информационных систем и т.д.).

Рассмотрим разные точки зрения на интерпретацию понятия нечёткости на примере. Пусть есть нечёткое утверждение "Вася (x) высокий (T)", которое можно представить как число $\mu_T(x)$ из единичного интервала. Существует несколько ответов на вопрос "Что означает, что $\mu_T(x) = 0.7$?"

1. Подход, основанный на правдоподобии: 70% опрошенных утверждают, что Вася высокий.

2. Подход на основе случайных множеств: 70% опрошенных описывают слово "высокий" с помощью интервала, содержащего рост Васи.

3. Интерпретация принадлежности как меры сходства объектов: рост Васи отличается от типового объекта, который можно назвать "высоким", на 0.3 (нормализованное расстояние).

4. Ценностный подход: 0.7 – это то, насколько информация о высоком росте Васи полезна.

5. Подход на основе теории меры: Вася выше некоторых людей и существует шкала, на которой этот факт можно представить в числовом виде как 0.7.

Подходы, основанные на правдоподобии и на случайных множествах, предполагают, что существует более чем один оценивающий субъект или

же эксперименты повторяются. Следовательно, если считать, что функции принадлежности используются для оображения смысла, а смысл является объективным, то получается, что нечёткость порождается несоответствием или ошибками. С другой стороны, следует сказать, что на начальных этапах развития теории нечётких множеств было повсеместно принято, что функции распределения носят субъективный характер и контекстнозависимы [97], [99]. Подход, основанный на сходстве объектов, и ценностный подход отличаются от остальных тем, что поддерживают субъективную интерпретацию. Теория мер связывает субъективные и объективные подходы, так как её аппарат может быть использован наблюдателем, который придерживается любой точки зрения. Сравнения могут быть результатами субъективных оценок или результатами "точных" (или идеализированных) измерений. Рассмотрим более подробно каждый из подходов.

1.4.2. Подход, основанный на правдоподобии

Этот подход к концепции измерения функции распределения был представлен Hisdal [54], [55]. Её модель подразумевает три источника нечёткости:

1. Ошибки в измерениях, которым присуща случайность по своей природе.
2. Неполнота информации.
3. Межличностные противоречия.

Кроме того, Hisdal утверждает, что для единичного объекта, информация о котором является исчерпывающей (как рост Васи, например), не существует нечёткости как таковой. В данном случае рост Васи можно охарактеризовать как "высокий", если его он больше определённого порогового значения, и нельзя, если меньше. Hisdal предлалает три способа определения функции распределения:

1. Эксперимент на выбор класса: Вася высокий, где "высокий" $\in$ {очень высокий, высокий, низкий}.
2. Бинарный эксперимент: Вася высокий? (да/нет)
3. Эксперимент на степень принадлежности: с какой степенью Вася принадлежит к множеству высоких людей?

Hisdal постулирует, что степень принадлежности объекта – это модификация её (чётких) ответов в экспериментах выбора класса или бинарных экспериментах, скорректированных на её оценку кривой ошибок.

Отвечая на вопрос о степени принадлежности, субъект, имеющий неточные знания о росте Васи (он считает, что рост находится в интервале 180 ± 5 см), строит кривую ошибок, оценивая возможную ошибку в окрестности 180 см. Далее он размышляет следующим образом: “В 70% случаев рост находится внутри моего интервала для понятия “высокий". Следовательно:

$$\mu_T(x) = P(\text{высокий}|x = 180\text{см}) = 0.7,$$

где $P(\text{высокий}|x = 180\text{см})$ обозначает правдоподобие того, что метка “высокий"может быть приписана объекту с ростом 180см.

Похожих взглядов придерживаются Mabuchi [69], Thomas [82].

1.4.3. Подход на основе случайных множеств

Для более подробного изучения этого подхода рассмотрим ещё один способ представления функции распределения – горизонтальный, при котором возможностная величина T представлена в терминах альфа-уровней [98]:

$$\mu_T = sup\{\alpha \in (0, 1] : t \in [T]^\alpha\},$$

где $[T]^\alpha(t) = \{t \in \mathbb{R}^1 | \mu_T(t) \geqslant \alpha\}, \alpha \in (0, 1]$.

Данное представление даёт возможность рассматривать функцию распределения (принадлежности) как семейство вложенных альфа-уровней (в том смысле, что если $\alpha'' \geqslant \alpha'$, то $[T]^{\alpha''} \subseteq [T]^{\alpha'}$). Тогда функция распределения может быть представлена в виде интеграла [42]:

$$\mu_T(t) = \int_0^1 [T]^\alpha(t) d\alpha, \text{ где } [T]^\alpha(t) = \begin{cases} 1, & \text{если} \quad t \in [T]^\alpha, \\ 0, & \text{иначе.} \end{cases}$$

Таким образом, с данной точки зрения на функцию распределения она может быть представлена как равномерно распределённое случайное множество, состоящее из меры Лебега на $[0, 1]$ и отображения $[T]^\alpha : (0, 1] \to X$, значениями которого являются множества.

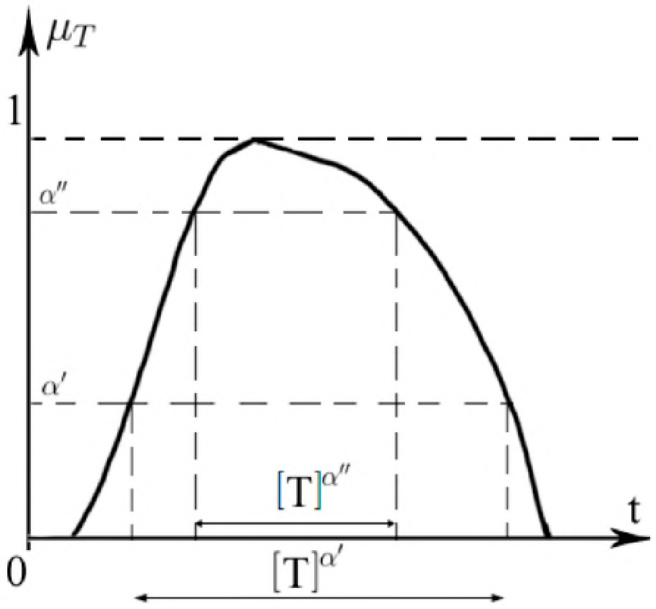


Рис. 1.5. Горизонтальное представление функции распределения

Следовательно, выражение $\mu_T(x) = 0.7$ означает, что 70% населения определяют T как интервал на множестве X, включающий x (где x – рост Васи). Оставшиеся 30% определяют T как интервал, в который x не входит.

1.4.4. Интерпретация принадлежности как меры сходства объектов

Эта точка зрения на функцию распределения возникает в теории прототипов, где принадлежность означает меру сходства с представителем категории [78], [64]. Преполагается, что существует идеальный элемент множества (или категория), который принадлежит этому множеству в полной мере. Принадлежность других элементов определяется расстоянием от них до идеального элемента.

Zysno рассматривает индивида в качестве измерительного устройства [104] и предполагает, что нечёткость берёт начало в ограниченности познавательных способностей человека, который сталкивается с задачей "сравнения объекта с прототипом или воображаемой идеей". Такой подход приводит к понятию расстояния, которое Zysno ассоциирует с функцией распределения:

$$\mu(x) = \frac{1}{1 + f(d_x)},$$

где d_x – расстояние между объектом x и идеалом, а $f(d_x)$ – некоторая функция расстояния.

Zimmermann и Zysno дали теоретическое описание этого метода [101], в котором математическая модель функции распределения (принадлежности) построена на действительной прямой, при этом принадлежность рассматривается как функция расстояния между рассматриваемым и идеальным объектами.

Говоря о расстояниях, многие исследователи неявно предполагают введение метрического пространства. Прежде чем его использовать, важно определить свойства этого пространства и его соотношение с психологическими законами.

Zwick, Carlstein и Budescu изучали различные меры сходства [103], используемые для введения понятия нечёткости. Они рассмотрели меры подобия, определённые как в метрических пространствах, так и с использованием теоретико-множественного подхода, и пришли к заключению, что при грубой категоризации все меры приводят к схожим результатам. Однако, меры, введённые на метрических пространствах, позволяют точнее определить степень сходства.

1.4.5. Ценностный подход

Giles предложил интерпретацию функции распределения в ключе теории решений [48]. С его точки зрения понятие функции распределения следует рассматривать в контексте задачи принятия решений в условиях неопределённости.

Он определяет нечёткое утверждение как “предложение, которому мы приписываем некоторую степень доверия”, являющееся функцией от возможных состояний мира (или сознания, природы и т.д.). Он выделяет предложения, которые утверждают, а не просто произносят. К примеру, когда кто-то утверждает “Вася высокий”, Giles предполагает, что существует функция ценности, связанная с этим утверждением. И чем ближе утверждение к правде, тем большую ценность привносит функция (человек, говорящий правду, пользуется большим доверием, чем врун). И это предположение позволяет ввести функцию распределения в контексте теории ценностей. Giles пишет “Ценность утверждений определяет смысл утверждений”[48].

Такой подход приводит к определению функции принадлежности на интервальной шкале (уникальной с точностью до положительных линейных

преобразований) [45].

Giles избавляется от влияния контекста на значение функции принадлежности (ребёнок ростом 160см – это высокий ребёнок, но не высокий человек), утверждая, что необходимо попытаться представить среднее значение предположения в обычном обществе.

Триангулярные нормы и конормы не подходят для вычисления степени истинности утверждений, поэтому не могут быть использованы в качестве операций в теории нечётких множеств, определённой в работах Giles.

1.4.6. Подход на основе теории меры

Зачастую математические структуры изобретаются для нужд математики без оглядки на практическое применение. Однако, теория возможностей является в большой степени эмпирической наукой – утверждается, что она описывает естественное явление, которое мы наблюдаем каждый день: нечёткость. А это означает, что объекты теории нечётких множеств должны иметь соответствующие прообразы в нашей повседневной жизни.

Подход на основе теории меры является связующим звеном, отображая явления, которые могут быть смоделированы в виде алгебраических структур (например, нечёткость), в числовые структуры, тем самым выражая сущность процесса измерения. В такой теории можно рассматривать представления качественной структуры и значимость этого представления.

Рассматривая понятие возможности в рамках теории меры, можно выделить две основные задачи: измерение принадлежности и ранжирование свойств. И хотя первая из них более популярна в научной литературе, вторая тесно связана с вопросом “какие операторы использовать?”.

Задача измерения принадлежности состоит в измерении степени принадлежности конечного числа объектов, агентов, элементов (A) одному нечёткому множеству (F). Ставится вопрос: “С какой степенью агент из A принадлежит нечёткому множеству F?”. Основой для определения степени принадлежности является утверждение вида “агент A_1 более F, чем агент A_2”, где F – нечёткий терм (как правило, прилагательное). Полученное представление измеряет степень, с которой каждый агент принадлежит нечёткому множеству F. Важно заметить, что это представление сравнивает несколько агентов по отношению к одному нечёткому терму.

Во второй задаче рассматривается отношение одного агента к нескольким нечётким термам. В этом случае измерения основываются на отношениях вида “Агент A более F_i , чем он F_j” (Вася лучше описывается с помощью термина "высокий чем "старый").

Такая постановка задачи позволяет обосновать выбор конкретных операторов на роли объединения и пересечения для нечётких множеств. Bellman и Giertz были первыми, кто дал аксиоматическое подтверждение использования операций min и max [29]. Yager также рассматривал эту проблему и предложил аксиомы [91], согласно которым min и max являются единственными решениями. Введение триангулярных норм, конорм и параметризованных семейств распределений в качестве возможных кандидатов для отражение связей послужило толчком к большому количеству исследований: [92], [40], [84], [25] и др.

В работах Norwich и Türkşen [74], [75] и Bilgiç и Türkşen [31], [32] задачи измерения функций принадлежности формализуются с помощью введения соответствующих алгебраических структур. Рассматриваются аксиомы, которые позволяют гарантировать существование численных представлений функции принадлежности на шкалах разного типа (порядковой, относительной, абсолютной) и определяют, какие арифметические операции могут быть использованы.

1.5. Выводы по первой главе

В первой главе рассмотрены понятия и термины теории возможностей, которые будут использоваться далее по ходу диссертации.

1. В параграфе 1.1 введено понятие возможностного пространства, определены его свойства, введены понятия мер неопределённости, возможности и необходимости.

2. Определена возможностная величина(переменная), её носитель, множество α-уровня, рассмотрены более подробно выпуклые возможностные величины. Приведены формулы преобразования для нахождения функций распределения возможностей от возможностных переменных. Рассмотрены операции над минисвязанными возможностными величинами. Определён специальный класс возможностных переменных – нечёткие числа (L, R)-типа, описаны свойства (L, R) – функций.

3. Параграф 1.3 посвящён агрегированию нечёткой информации на основе t-норм. Вначале дано понятие треугольной нормы (t-нормы), описаны её основные свойства, особое внимание уделено сильнейшей и слабейшей t-нормам. Показан вид t-нормы для большого числа аргументов.

   Рассмотрен класс Архимедовых t-норм и его взаимосвязь с остальными треугольными нормами.

   Далее введено понятие аддитивного генератора, обозначена важность данного математического инструмента. Показана его связь с Архимедовыми t-нормами.

4. В параграфе 1.4 рассмотрены существующие подходы к объяснению понятия функции распределения (принадлежности) и способам измерения её значений. Вначале с помощью каждого из подходов дана интерпретация предложенного нечёткого утверждения, а далее подробно рассмотрен каждый из них.

   Первым описан подход, основанный на правдоподобии, и введённый Hisdal, далее на рассмотрение вынесен подход, основанный на случайных множествах, в котором вводится горизонтальный способ представления функции распределения, дающий возможность рассматривать её как семейство вложенных альфа-уровней.

   Следующим описан подход, исходя из которого принадлежность интерпретируется как мера сходства объектов и преполагается, что существует идеальный элемент множества, который принадлежит этому множеству в полной мере, а принадлежность других элементов определяется расстоянием от них до идеального элемента.

   Далее рассмотрен ценностный подход, основная идея которого заключается в том, что существует связь между утверждением и функцией ценности. И чем ближе утверждение к правде, тем большую ценность привносит функция. Это предположение позволяет ввести функцию распределения в контексте теории ценностей.

   В заключение описан подход на основе теории меры, который является связующим звеном, отображая явления, которые могут быть смоделированы в виде алгебраических структур в числовые структуры. В рамках

этого подхода выделены две основные задачи: измерение принадлежности и ранжирование свойств. Далее рассмотрена каждая из задач.

После рассмотрения существующих способов измерения значений функции распределения в качестве основополагающего для данной работы был выбран подход, основанный на интерпретации принадлежности как меры сходства объектов. В исследованиях, которым посвящана данная работа, будем далее придерживаться именно этого подхода.

2. Оценка параметров возможностных распределений в одномерном случае

Хотя нечёткие переменнные, введённые S.Nahmias [71], служат базой для различных теоретических приложений и им уделяется большое внимание, связанная с ними тема оценки параметров возможностных распределений развита слабо. Данный вопрос освещён в научной литературе для одномерного случая [36], [56], [90].

В данной главе будет рассмотрено несколько существующих подходов к оценке параметров, на основе которых в последующих главах будут построены методы для оценки параметров в многомерном случае.

2.1. Постановка задачи

Здесь и далее будем рассматривать возможностные распределения, зависящие от одного или нескольких параметров. Введём некоторые обозначения [90].

Определение 22. *Множество функций распределения одного вида, зависящие от параметров, которые варьируются в одном допустимом интервале, называют семейством функций распределения.*

Определение 23. *Пусть* $\xi_1, \xi_2, ..., \xi_n$ — *n t-связанных, одинаково распределённых возможностных величин, тогда* $(\xi_1, \xi_2, ..., \xi_n)$ *называют выборкой семейства, а* $x_1, x_2, ..., x_n$ — *наблюдаемыми значениями выборки* $(\xi_1, \xi_2, ..., \xi_n)$.

Определение 24. *Любая функция от выборки* $\xi_1, \xi_2, ..., \xi_n$, *которая не включает неизвестные параметры, называется статистикой.*

Заметим, что статистика является возможностной переменной.

Пусть семейство представляет собой множество распределений, зависящих от вектора параметров Θ:

$$\{\mu_\xi(x, \boldsymbol{\theta}) | \boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k) \in \Theta\}, x \in \mathbb{R}^1,$$

где Θ – пространство параметров распределения $\Theta \subseteq \mathbb{R}^k$.

Задача оценивания параметров возможностного распределения состоит в том, чтобы найти подходящий вектор параметров $\boldsymbol{\theta}$ на основе значений $(x_1, x_2, ..., x_n)$ выборки $(\xi_1, \xi_2, ..., \xi_n)$ после проведения n наблюдений.

Для того, чтобы определить качество оценки параметров, используют различные критерии. Основываясь на [90], введём определения используемых далее критериев.

Определение 25. *Пусть F представляет собой семейство функций распределения, $\bar{\xi} = (\xi_1, \xi_2, ..., \xi_n)$ — выборка, и пусть $T(\bar{\xi})$ - статистика, функция распределения которой принадлежит семейству F. Будем говорить, что $T(\bar{\xi})$ — достаточна по отношению к F, если совокупная функция распределения от $(T(\bar{\xi}), \bar{\xi})$ не зависит от x.*

Каждая выборка предоставляет определённую информации о семействе. И данное определение показывает, что достаточная статистика предоставляет тот же объём информации о семействе, что и вся выборка. Из чего следует, что такая достаточная статистика может быть использована для уменьшения объёма выборочных данных без потери информации.

Определение 26. *Пусть F - семейство функций распределения, представленных следующим образом: $F = \{F(x, \boldsymbol{\theta}) | \boldsymbol{\theta} \in \Theta, \Theta \subseteq \mathbb{R}^k\}, \bar{\xi} = (\xi_1, \xi_2, ..., \xi_n)$ — выборка и $T(\bar{\xi})$ — статистика. Тогда:*
1)$T(\bar{\xi})$ — достаточная оценка вектора параметров $\boldsymbol{\theta}$, если $T(\bar{\xi})$ — достаточна по отношению к F;
2)$T(\bar{\xi})$ — состоятельная оценка вектора параметров $\boldsymbol{\theta}$, если $\Pi(\{T(\bar{\xi}) = \boldsymbol{\theta}\}) = 1$;
3)$T(\bar{\xi})$ — максимально правдоподобная оценка вектора параметров $\boldsymbol{\theta}$, если:

$$D(x_1, ..., x_n, T(\bar{\xi})) = \max_{\boldsymbol{\theta} \in \Theta} D(x_1, ..., x_n, \boldsymbol{\theta}),$$

где $D(x_1, ..., x_n, \boldsymbol{\theta}) = \bigwedge_{j=1}^{n} F(x_j, \boldsymbol{\theta})$.

Состоятельность отражает возможность того, что полученная оценка параметров представляет собой действительные значения параметров. Максимально правдоподобные оценки строятся по методу максимального правдоподобия, откуда и берут своё название, и максимизируют функцию правдоподобия при фиксированной реализации выборки [13].

2.2. Существующие методы оценивания параметров

В [36], [56], [90] и др. предлагаются различные варианты методов оценивания параметров возможностных распределений. В данном параграфе будут рассмотрены два наиболее интересных из них: метод Cai Kai-Yuan, который далее обобщил Dug Hun Hong на случай оценивания параметров взаимно t-связанных нечётких величин, и метод максиминной μ/E оценки параметров одномерных возможностных распределений, разработанный Wang Xizhao и Ha Minghu. Вначале подробно рассмотрим первый из них.

2.2.1. Оценка параметров возможностных распределений в одномерном случае с использованием информации об уровне риска её ошибочного определения

Cai Kai-Yuan разработал метод нахождения оценки параметров нормальных возможностных переменных [36] в случае t = min. В своей работе он рассмотрел методы точечного оценивания и максимального правдоподобия для нахождения значения параметра a при известном b, где a представляет собой модальное значение, а b – коэффициент нечёткости. Им были предложены следующие оценки для a:

$$\hat{a} = \frac{x_1 + ... + x_n}{n}, \qquad \hat{a} = \frac{1}{2}\Big(\max_{1 \leqslant i \leqslant n} x_i + \min_{1 \leqslant i \leqslant n} x_i\Big).$$

где $x_1, ..., x_n$ – наблюдаемые значения нечётких переменных $\xi_1, ..., \xi_n$.

Также он получил оценку параметра b при известном a, используя метод интервального оценивания:

$$\hat{b} = \frac{\max\limits_{1 \leqslant i \leqslant n} |x_i - a|}{\varepsilon_\alpha},$$

где ε_α удовлетворяет условию: $\sigma(|(\xi_i - a)/b| > \varepsilon_\alpha, i = 1, ..n) = \alpha$.

В заключение Cai приводит метод оценивания a и b, в случае, когда оба параметра неизвестны:

$$\hat{a} = \frac{1}{2}\Big(\max_{1 \leqslant i \leqslant n} x_i + \min_{1 \leqslant i \leqslant n} x_i\Big), \quad \hat{b} = \frac{\max\limits_{1 \leqslant i \leqslant n} x_i - \min\limits_{1 \leqslant i \leqslant n} x_i}{2\varepsilon_\alpha}.$$

Dug Hun Hong обобщил результаты Cai и в [56] представил методы оце-

нивания параметров взаимно T-связанных нечётких величин.

Наиболее интересным из рассмотренных им случаев является получение оценки параметров распределения возможностных переменных $\xi_1, ..., \xi_n$ в одномерном случае при неизвестных a и b и заданном α – уровне риска ошибочного определения оценки.

Для нахождения оценки в одномерном случае Hong решал следующую задачу [56]:

$$\xi_\alpha^+ - \xi_\alpha^- \to min; \tag{2.1}$$

$$\mu_{\xi_1,...\xi_n}(x_1, .., x_n; a, b) = \alpha, \tag{2.2}$$

где α определяет уровень риска того, что оценка a не находится в интервале $(\xi_\alpha^-; \xi_\alpha^+)$, то есть α задаёт минимальный уровень возможности для точек выборки, $0 < \alpha \leqslant 1$; ξ_α^- и ξ_α^+ – границы α-уровневых множеств.

При t-норме $T_m(x, y) = min(x, y)$ и заданном значении α им были получены следующие результаты для одномерного случая:

$$\hat{b} = \frac{\max\limits_{1\leqslant i\leqslant n} x_i - \min\limits_{1\leqslant i\leqslant n} x_i}{H_+^{-1}(\alpha) - H_-^{-1}(\alpha)}, \tag{2.3}$$

$$\hat{a} = \frac{1}{2}\Big(\max_{1\leqslant i\leqslant n} x_i + \min_{1\leqslant i\leqslant n} x_i\Big) + \frac{\hat{b}}{2}(H_-^{-1}(\alpha) + H_+^{-1}(\alpha)), \tag{2.4}$$

где $H(x)$ – функция представления формы, такая что для нечёткой переменной X типа H функция распределения $\mu_X(x) = H(\frac{x-a}{b})$, $H(H_-^{-1}(\alpha)) = \alpha = H(H_+^{-1}(\alpha))$, $H : \mathbb{R}^1 \to [0, 1]$ непрерывна, строго возрастает на $\{H > 0\} \cap (-\infty, 0)$ и строго убывает на $\{H > 0\} \cap (0, \infty)$, $H(0) = 1$, $H(x) \neq 1$ для $\forall x \neq 0$ и $\lim\limits_{x\to\pm\infty} H(x) = 0$, a, $b > 0$.

Следует заметить, что ни Cai Kai-Yuan, ни Dug Hun Hong не представили свойства полученных ими оценок.

В статье [90] Wang Xizhao и Ha Minghu, опираясь на [72], [89], [102], разработали метод нахождения μ/E оценки параметров a, b одномерных возможностных распределений. Они исследовали семейство нечётких чисел, которое задаётся функцией μ, определяющей форму распределения с параметрами

$a = 0$ и $b = 1$. Ими были получены следующие результаты:

$$\hat{a} = \frac{1}{2}\Big(\max_{1\leqslant i\leqslant n} x_i + \min_{1\leqslant i\leqslant n} x_i\Big), \quad \hat{b} = \frac{\max\limits_{1\leqslant i\leqslant n} x_i - \min\limits_{1\leqslant i\leqslant n} x_i}{2c},$$

где $c = \arg\max\limits_{t\geqslant 0}\{t\mu(t)\}$.

В следующем параграфе рассмотрим подробнее, как были получены эти результаты.

2.2.2. Максиминная μ/E оценка параметров одномерных возможностных распределений

Представим метод оценки параметров распределения минисвязанных одномерных возможностных величин, предложенный Wang Xizhao и Ha Minghu в [90].

Рассмотрим следующее семейство распределений:

$$F(\mu) = \{Q_\mu(a,b) | Q_\mu(a,b)(x) = \mu\Big(\frac{x-a}{b}\Big), a \in \mathbb{R}^1, b > 0\}, \tag{2.5}$$

где N_μ – симметричное в нуле нечёткое число, т.е. $\mu(-x) = \mu(x),\ \forall x \in \mathbb{R}^1$. Согласно [90] получаем:

$$\sum_{i=1}^{n} c_i Q_\mu(a_i, b_i) = Q_\mu\Big(\sum_{i=1}^{n} c_i a_i, \sum_{i=1}^{n} |c_i| b_i\Big) \in F(\mu),$$

где $c_i \in \mathbb{R}^1$ и $Q_\mu(a_i, b_i) \in F(\mu), i = 1, 2, ..., n$.

Обозначим через $E(\xi) = \int\limits_R \eta_\xi(t)dt$ меру нечёткости распределения η_ξ.

При условии, что N_μ — симметричное в нуле нечёткое число, $Q_\mu(a,b)$ является функцией распределения для возможностной величины ξ и $\int\limits_R \mu(t)dt = l$, легко показать, что $E(\xi) = lb$ [90].

Совместное распределение возможности наблюдения значений $(x_1, ..., x_n)$ имеет вид:

$$\mu(x_1, ..., x_n, \boldsymbol{\theta}) = \bigwedge_{j=1}^{n} \mu_\xi(x_j, \boldsymbol{\theta}), \tag{2.6}$$

где $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k)$. Принимая во внимание тот факт, что любые изменения параметров $\theta_1, \theta_2, ..., \theta_k$ непосредственно повлияют как на возможность

появления тех или иных наблюдаемых значений, так и на значение $E(\xi)$, мы должны выбрать такой вектор $\boldsymbol{\theta}$, чтобы максимизировать возможность (2.6) и при этом получить минимальное $E(\xi)$.

Введём следующее обозначение:

$$L(\boldsymbol{\theta}) = \bigwedge_{j=1}^{n} \mu_{\xi}(x_j, \boldsymbol{\theta})/E(\xi), \forall \boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_k). \tag{2.7}$$

Определение 27. *Если найдётся такая оценка* $\hat{\boldsymbol{\theta}} \in \Theta$, *что* $L(\hat{\boldsymbol{\theta}}) = \max\limits_{\boldsymbol{\theta}\in\Theta} L(\boldsymbol{\theta})$,
$L(\boldsymbol{\theta})$ *получена из* (2.7), *то* $\hat{\boldsymbol{\theta}}$ *называют максиминной* μ/E *оценкой вектора параметров* $\boldsymbol{\theta}$.

Теорема 9. *[90] Пусть* N_μ *— симметричное в нуле нечёткое число.* $F(\mu) = \{Q_\mu(a,b)|a \in \mathbb{R}^1, b > 0\}$ *— семейство распределений* (2.5). *Тогда максиминная оценка вектора параметров* $\boldsymbol{\theta} = (a,b)$ *может быть найдена следующим образом:*

$$(\hat{a}, \hat{b}) = \big((x^{(1)} + x^{(n)})/2, (x^{(n)} - x^{(1)})/2c\big), \tag{2.8}$$

где $x^{(1)} \leqslant x^{(2)} \leqslant, ..., \leqslant x^{(n)}$ *– упорядоченные наблюдаемые значения* $x_1, x_2, ..., x_n$ *возможностной величины, принадлежащей семейству распределений,* $n \geqslant 2$, c *— действительное число, при котором функция* $g(t) = t\mu(t) (t \geqslant 0)$ *достигает своего максимума.*

Для того, чтобы показать, что $(\hat{a}, \hat{b})$ является действительно качественной оценкой параметров распределения, в [90] была доказана следующая теорема:

Теорема 10. *Пусть* $F(\mu)$ *задаёт семейство распределений,* $(\xi_1, \xi_2, ..., \xi_n)$ *— выборка и* $(x_1, x_2, ..., x_n)$ *— её наблюдаемые значения. Пусть* $m = \min(\xi_1, \xi_2, ..., \xi_n)$, $M = \max(\xi_1, \xi_2, ..., \xi_n)$, $\bar{X} = \big(M + m\big)/2$, $S = \big(M - m\big)/2c$, *где* c *— значение, при котором функция* $t\mu(t)$ *достигает своей максимума. Тогда:*
1. $\bar{X}$ *— состоятельная оценка параметра* a *и максимально правдоподобная оценка параметра* a, *когда* b *известно.*
2. $(\bar{X}, S)$ *— достаточная оценка вектора параметров* (a,b).

Таким образом, получаем, что $(\hat{a}, \hat{b})$ является состоятельной, достаточной и максимально правдоподобной оценкой параметров в одномерном случае для семейства возможностных симметричных унимодальных распределений и её можно использовать для определения параметров распределения без потери информации.

2.3. Сравнительный анализ рассмотренных методов

Рассмотрим взаимосвязь между двумя методами нахождения оценки параметров: μ/E и метод Dug Hun Hong.

Можно видеть, что метод μ/E оценки является частным случаем метода Dug Hun Hong.

Действительно, если произвести оценку параметров μ/E методом, то мы получим некоторые параметры распределения, и при этом значение функции распределения на крайних точках будет равно некоторому значению α. Если выполнить оценкку параметров по методу Dug Hun Hong с уровнем риска, равным α, то получим те же значения оцениваемых параметров распределения: параметр a, отвечающий за положение центра распределения, определяется по одинаковым формулам (2.4 для симмметричных распределений сводится к 2.8). Оценки параметра, отвечающего за коэффициент нечёткости, также буду совпадать ввиду способа построения оценки Dug Hun Hong: оценка строится так, что минимальные значения равны уровню риска α.

При использовании метода Dug Hun Hong исследователь задаёт подходящий уровень риска и получает параметры распределения. При этом от значения уровня риска зависит только коэффициент нечёткости b, определяющий ширину распределения. Таким образом, b является функцией от уровня риска α. Построим график этой зависимости (рис.2.1) на основе числовых данных от Complex Systems Summer School (Santa Fe) [46] о работе инфракрасного лазера.

Метод μ/E оценки не требует задания параметров и выдаёт одну точку с этого графика. Чтобы понять, где расположена эта точка, воспользуемся графическим способом, аналогичным применяемому для решения задач линейного программирования.

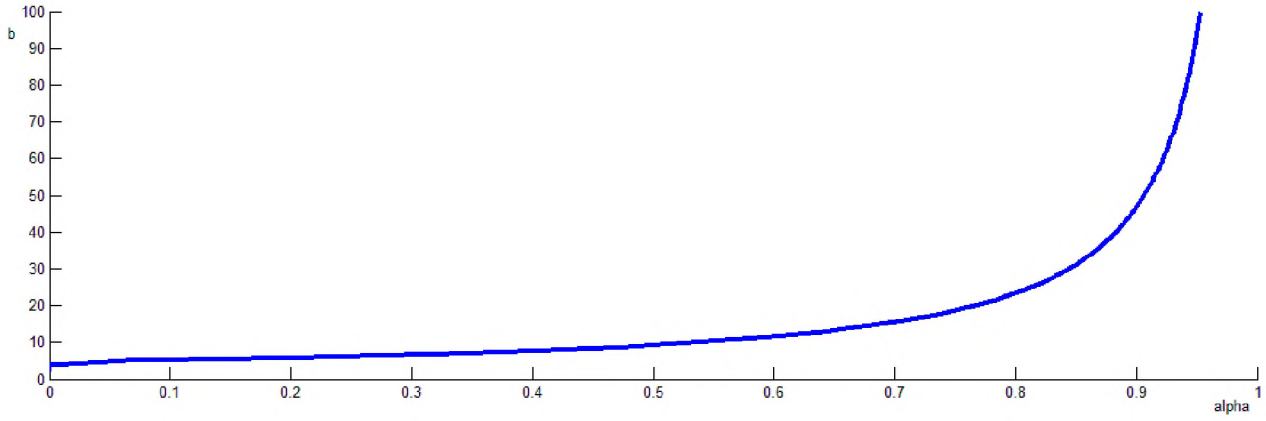


Рис. 2.1. Значения параметра b при различных уровнях α

Рассмотрим множество точек, для которых

$$\mu/E = c, c = const.$$

Согласно [90] $E = lb$. Из 2.6 следует, что $\mu(x_1, ..., x_n, \boldsymbol{\theta}) = \alpha$, т.к. среди значений μ_ξ выбираются минимальные, которые достигаются на краях распределения.

Соответственно, если $\mu/E = c$ и $c' = c/l$, то $\alpha/b = c'$ – это уравнение описывает прямую L, проходящую через начало координат (рис 2.2).

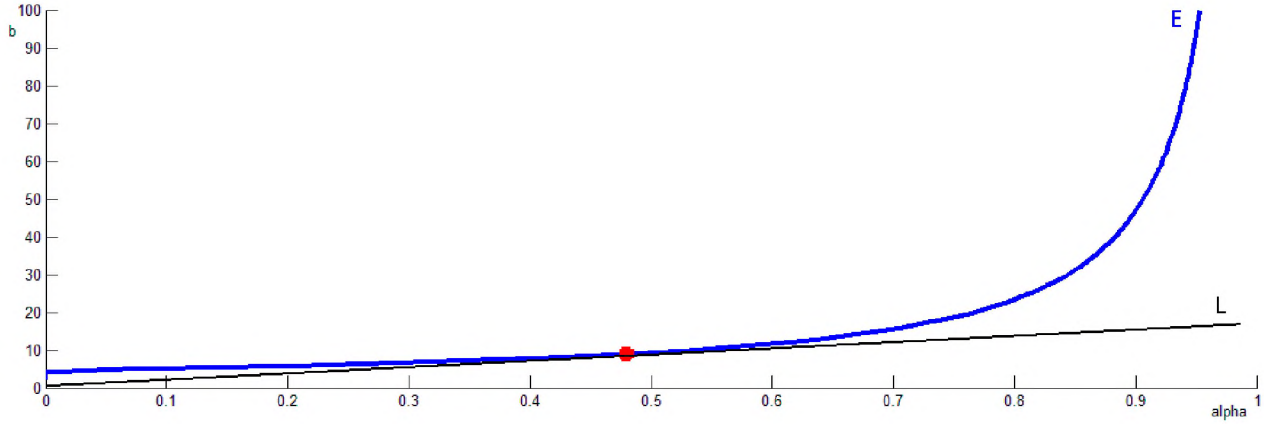


Рис. 2.2. Касание прямой L кривой, отражающей возможные значения параметра b

Увеличение значения отношения μ/E приводит к повороту прямой L по часовой стрелке. При этом все допустимы знаения пар α/b лежат на линии, задаваемой методом Dug Hun Hong.

Оптимальное с точки зрения метода μ/E значение соответствует последнему положению этой прямой из начала координат, при котором она ещё

касается кривой E.

2.4. Выводы по второй главе

Во второй главе были рассмотрены существующие подходы к оценке параметров возможностных распределений в одномерном случае.

1. В параграфе 2.1 произведена постановка задачи, определены свойства, которым должна отвечать полученная оценка.

2. В параграфе 2.2 сделан обзор двух существующих методов получения оценки в одномерном случае: в параграфе 2.2.1 метод Cai Kai-Yuan и его расширение, сделанное Dug Hun Hong, и в параграфе 2.2.2. подробно описан метод Wang Xizhao и Ha Minghu для нахождения максиминной оценки параметров.

   Оценка, которую получил Hong, выглядит следующим образом (при сильнейшей t-норме и заданном значении α):

$$\hat{b} = \frac{\max\limits_{1 \leqslant i \leqslant n} x_i - \min\limits_{1 \leqslant i \leqslant n} x_i}{H_+^{-1}(\alpha) - H_-^{-1}(\alpha)}, \tag{2.9}$$

$$\hat{a} = \frac{1}{2}\Big(\max_{1 \leqslant i \leqslant n} x_i + \min_{1 \leqslant i \leqslant n} x_i\Big) + \frac{\hat{b}}{2}(H_-^{-1}(\alpha) + H_+^{-1}(\alpha)), \tag{2.10}$$

   где $H(x)$ – функция представления формы, такая что для нечёткой переменной X типа H функция распределения $\mu_X(x) = H(\frac{x-a}{b})$. $H(H_-^{-1}(\alpha)) = \alpha = H(H_+^{-1}(\alpha))$, $H : \mathbb{R}^1 \to [0,1]$ непрерывна, строго возрастает на $\{H > 0\} \cap (-\infty, 0)$ и строго убывает на $\{H > 0\} \cap (0, \infty)$, $H(0) = 1$, $H(x) \neq 1$ для $\forall x \neq 0$ и $\lim\limits_{x \to \pm\infty} H(x) = 0$, a, $b \in \mathbb{R}^1$, $b > 0$.

   Wang Xizhao и Ha Minghu предлагают следующую оценку:

$$\hat{a} = \frac{1}{2}\Big(\max_{1 \leqslant i \leqslant n} x_i + \min_{1 \leqslant i \leqslant n} x_i\Big), \quad \hat{b} = \frac{\max\limits_{1 \leqslant i \leqslant n} x_i - \min\limits_{1 \leqslant i \leqslant n} x_i}{2c},$$

   где $c = \arg\max\limits_{t \geqslant 0}\{t\mu(t)\}$.

   В отличие от Hong, они доказали, что их оценка является состоятельной, достаточной и максимально правдоподобной.

3. В параграфе 2.3 произведён сравнительный анализ рассмотренных методов.

   Если произвести оценку параметров предложенным Wang Xizhao и Ha Minghu методом, то получим некоторые параметры распределения, и при этом значение функции распределения на крайних точках будет равно некоторому значению α. Если выполнить оценкку параметров по методу Dug Hun Hong с уровнем риска, равным α, то получим те же значения оцениваемых параметров распределения: параметр a, отвечающий за положение центра распределения, определяется по одинаковым формулам. Действительно, следующая формула

$$\hat{a} = \frac{1}{2}\Big(\max_{1\leqslant i\leqslant n} x_i + \min_{1\leqslant i\leqslant n} x_i\Big) + \frac{\hat{b}}{2}(H_-^{-1}(\alpha) + H_+^{-1}(\alpha)),$$

   для симмметричных распределений сводится к формуле:

$$(\hat{a}, \hat{b}) = \big((x^{(1)} + x^{(n)})/2, (x^{(n)} - x^{(1)})/2c\big),$$

   Оценки параметра, отвечающего за коэффициент нечёткости, также будут совпадать ввиду способа построения оценки Dug Hun Hong: оценка строится так, что минимальные значения равны уровню риска α.

   При использовании метода Dug Hun Hong исследователь задаёт подходящий уровень риска и получает параметры распределения. При этом от значения уровня риска зависит только коэффициент нечёткости b, определяющий ширину распределения. Таким образом, b является функцией от уровня риска α.

   После проведённого анализа можно утверждать, что метод μ/E оценки является частным случаем метода Dug Hun Hong.

3. Оценка параметров многомерных возможностных распределений при различных t-нормах

В предыдущей главе были рассмотрены существующие методы оценки параметров одномерных возможностных распределений. Насколько нам известно, вопрос оценки параметров в многомерном случае никем не рассматривался. Этот факт находит своё отражение в ограниченном использовании данного математического инструмента в решении практических задач. Для ликвидации этого пробела в текущей главе предлагается несколько методов оценки параметров многомерных возможностных распределений.

3.1. Оценка параметров многомерных распределений минисвязанных возможностных величин

3.1.1. Метод нахождение оценки параметров многомерных возможностных распределений при заданном уровне возможности

В данной главе будет разработан метод нахождения оценки параметров многомерных минисвязанных нечётких величин, основанный на подходе Dug Hun Hong [56].

Для нахождения оценки параметров в многомерном случае можно расширить критерии (2.1) и (2.2) следующим образом: критерий (2.2) может быть непосредственно применён в многомерном случае, естественным расширением критерия (2.1) на многомерный случай является

$$\int\limits_{\mu_{\boldsymbol{\xi}}(\mathbf{x},\boldsymbol{\theta})\geqslant\alpha} \cdots \int 1 d\mathbf{x} \to min,$$

где $x \in \mathbb{R}^m, \boldsymbol{\theta} \in \Theta, \Theta \subseteq \mathbb{R}^k, 0 < \alpha \leqslant 1$. Данный критерий минимизирует объём α–уровневого множества распределения $\mu_{\boldsymbol{\xi}}(\mathbf{x})$.

Таким образом, можно ввести следующее определение оценки:

Определение 28. $\boldsymbol{\theta}_\alpha = (\mathbf{c}, A)$ *назовём оценкой параметров возможностного распределения, если она является решением следующей задачи оптимизации:*

$$\int\limits_{\mu_{\boldsymbol{\xi}}(\mathbf{x},\boldsymbol{\theta})\geqslant\alpha} \cdots \int 1 d\mathbf{x} \to min; \tag{3.1}$$

$$\mu_{\boldsymbol{\xi}}(\mathbf{x}_1, ..., \mathbf{x}_n, \boldsymbol{\theta}) = \alpha, \tag{3.2}$$

где α *задаёт минимальный уровень возможности для точек выборки,* $0 < \alpha \leqslant 1$, $\mathbf{x}_i \in \mathbb{R}^m$.

Используя критерии (3.1) и (3.2), перенесём подход Dug Hun Hong на многомерный случай. Способ определения оценки параметров $\mathbf{c}, A$ даёт следующая теорема.

Теорема 11. *Пусть:*

1. $f : \mathbb{R}^1 \to \mathbb{R}^1$ *— строго убывающая на луче* $[0, +\infty)$ *функция,* $f(0) = 1$.

2. $Q = \{\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A)\}$ *— семейство распределений, где* $\mathbf{c} \in \mathbb{R}^m$, A *– положительно определённая симметричная матрица* $m \times m$, $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$.

3. $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ *— наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ *минисвязанных возможностных величин, принадлежащих семейству распределений* Q, $\mathbf{x}_i \in \mathbb{R}^m, i = 1, ..., n$.

4. W *— многомерный эллипсоид минимальноого объёма, описанный вокруг множества* X, *заданный параметрами* $(\breve{\mathbf{c}}, \breve{A})$, *где* $\breve{\mathbf{c}}$ *– центр эллипсоида* W, $\breve{A}$ *– матрица, задающая ориентацию осей эллипсоида и длину его радиусов.*

Тогда оценкой параметра $\boldsymbol{\theta}_\alpha = (\mathbf{c}, A)$ *является* $\hat{\boldsymbol{\theta}}_\alpha = (\hat{\mathbf{c}}, \hat{A}) = (\breve{\mathbf{c}}, k\breve{A})$, *где* $k = (f^{-1}(\alpha))^2, 0 < \alpha \leqslant 1$.

Доказательство:

Очевидно, что границами α–уровневых множеств для данного вида распределений являются эллипсоиды. Рассмотрим минимальное α–уровневое множество, включающее все точки из X. Его границей является эллипсоид, который обозначим как V. Тогда его объём равен $\int\limits_{\mu_{\boldsymbol{\xi}}(\mathbf{x})\geqslant\alpha} \cdots \int 1 d\mathbf{x}$. Исходя из

критерия (3.1) следует подбирать параметры таким образом, чтобы объём эллипсоида V был минимальным.

Пусть вектор параметров распределения $\boldsymbol{\theta}_\alpha$ выбирается таким образом, чтобы одной из линий равного уровня являлся эллипсоид минимального объёма W, описанный вокруг множества X. Тогда $\boldsymbol{\theta}_\alpha = (\breve{\mathbf{c}}, k\breve{A})$, $k \in \mathbb{R}^1$.

Рассмотрим подмножество $Z = \{\mathbf{x}_{z_1}, ..., \mathbf{x}_{z_k}\} = X \cap W$ – наблюдаемые значения, которые лежат на W. Обозначим через μ_Z значение функции распределения на точках из подмножества Z. Тогда $\mu_Z = \min\limits_{\mathbf{x}\in X} \mu_{\boldsymbol{\xi}}(\mathbf{x}, \boldsymbol{\theta}) = \mu_{\boldsymbol{\xi}}(\mathbf{x}_1, ..., \mathbf{x}_n, \boldsymbol{\theta})$.

Согласно (3.2) $\mu_{\boldsymbol{\xi}}(\mathbf{x}_1, ..., \mathbf{x}_n, \boldsymbol{\theta}) = \alpha$, следовательно $\mu_Z = \alpha$.

По условию теоремы $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$, следовательно $f\big(\sqrt{(\mathbf{x}-\breve{\mathbf{c}})^T k\breve{A}(\mathbf{x}-\breve{\mathbf{c}})}\,\big) = f\big(\sqrt{k}\sqrt{(\mathbf{x}-\breve{\mathbf{c}})^T \breve{A}(\mathbf{x}-\breve{\mathbf{c}})}\,\big) = \alpha$ для всех $\mathbf{x} \in Z$. Заметим, что для эллипсоида W выполняется соотношение $(\mathbf{x}-\breve{\mathbf{c}})^T \breve{A}(\mathbf{x}-\breve{\mathbf{c}}) = 1$. Следовательно, $f(\sqrt{k}) = \alpha$ и $k = (f^{-1}(\alpha))^2$.

Таким образом, оценкой параметра $\boldsymbol{\theta}_\alpha = (\mathbf{c}, A)$ является $\hat{\boldsymbol{\theta}}_\alpha = (\hat{\mathbf{c}}, \hat{A}) = (\breve{\mathbf{c}}, k\breve{A})$, где $k = (f^{-1}(\alpha))^2$. ■

Пример 1. Рассмотрим пример нахождения оценки параметров распределения в двумерном случае для функции $f(t) = e^{-t^2}$.

Для того, чтобы воспользоваться доказанной ранее теоремой, необходимо задать уровень α. Рассмотрим 3 случая:

1. $\alpha = 0.1$;
2. $\alpha = 0.4$;
3. $\alpha = 0.7$.

Построив вокруг множества X эллипс минимальной площади W и получив параметры этого эллипса, найдём значения параметров нормального распределения согласно теореме 11:

$$k = -ln(\alpha),\ \hat{A} = -ln(\alpha)\breve{A},\ \hat{\mathbf{c}} = \breve{\mathbf{c}},$$

где $(\hat{\mathbf{c}}, \hat{A})$ – оценки параметров распределения, $\breve{\mathbf{c}}$ – центр эллипса W, $\breve{A}$ – матрица, задающая ориентацию осей эллипса и длину его радиусов.

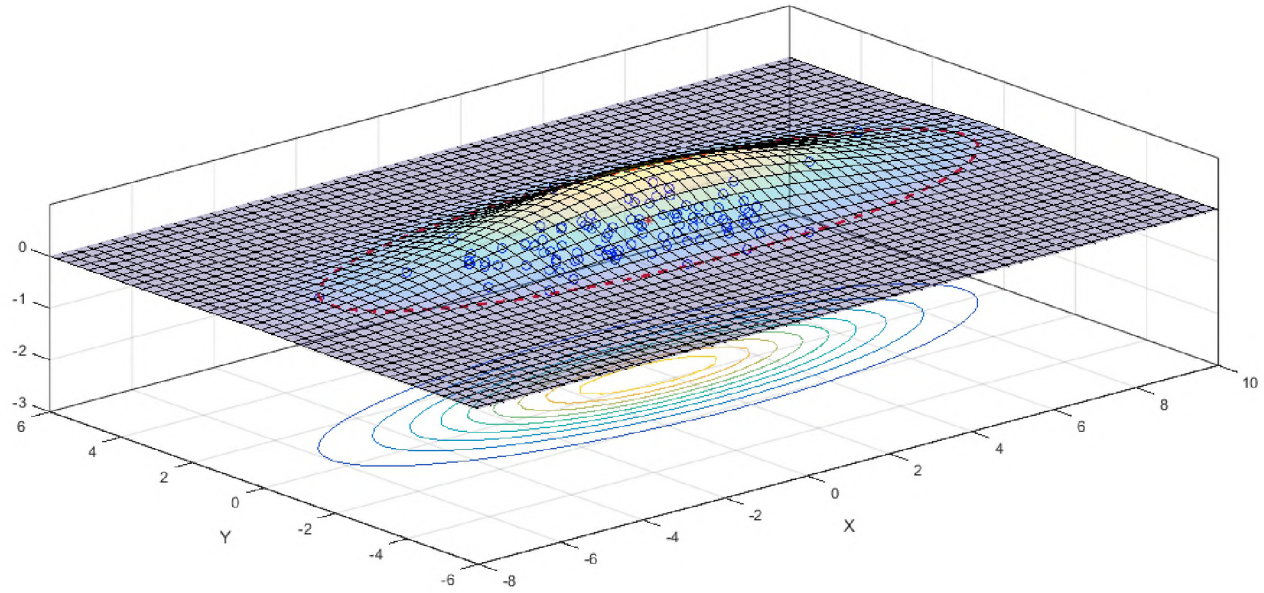


Рис. 3.1. Восстановление параметров двумерного нормального распределения в случае $\alpha = 0.1$

На рисунках 3.1, 3.2 и 3.3 изображены графики функций нормального распределения $\mu_{\boldsymbol{\xi}}(\hat{\mathbf{c}}, \hat{A})$ при $\alpha = 0.1, 0.4, 0.7$ соответственно, где: $\circ$ – наблюдаемые значения возможностной величины; —— – описанный эллипс минимальной площади W; $\bullet$ – центр W; эллипсы, расположенные в нижней части рисунков, отражают границы α–уровневых множеств функции распределения μ.

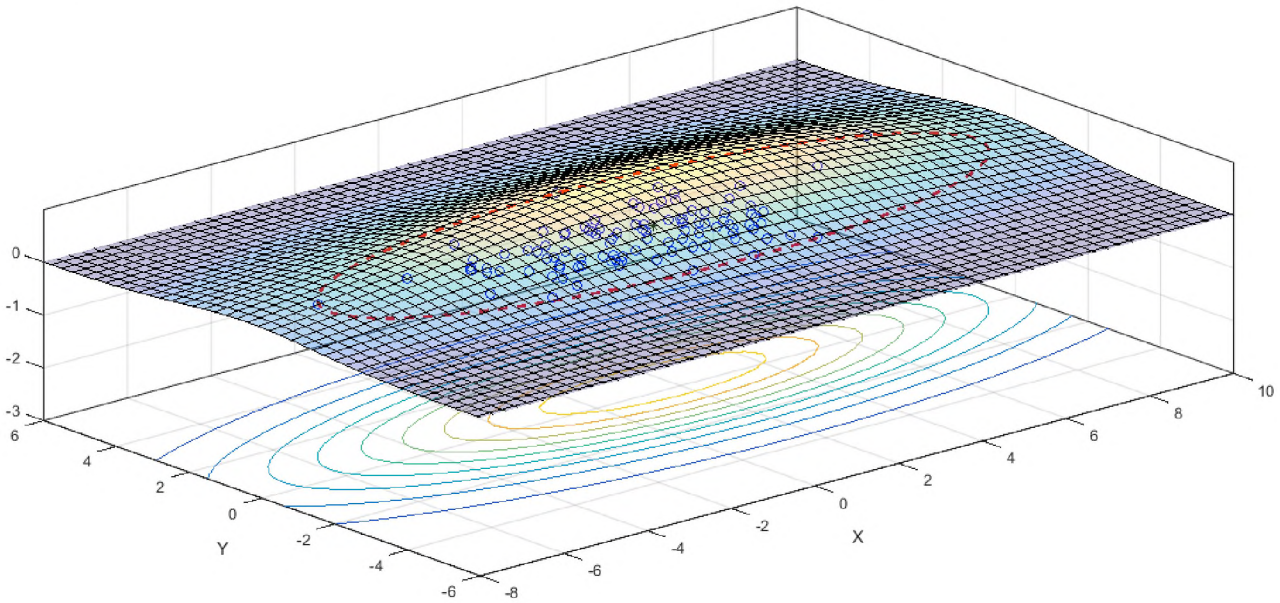


Рис. 3.2. Восстановление параметров двумерного нормального распределения в случае $\alpha = 0.4$

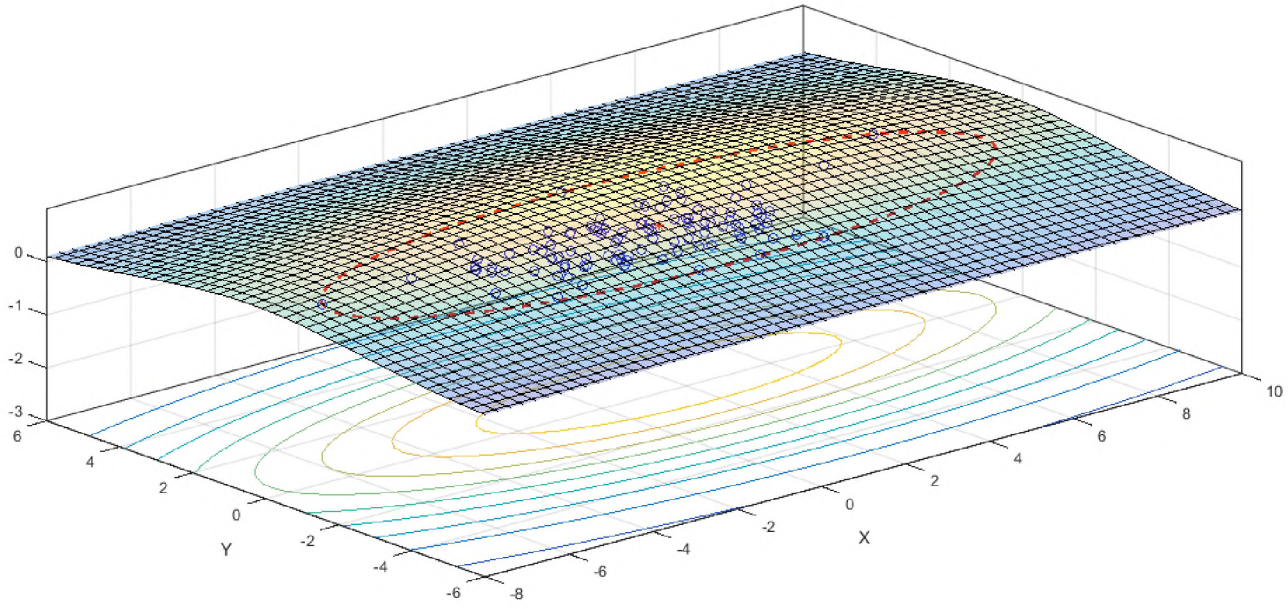


Рис. 3.3. Восстановление параметров двумерного нормального распределения в случае $\alpha = 0.7$

Из рисунков видно, что параметр α влияет на форму распределения: чем больше значение параметра, тем более пологим становится распределение.

3.1.2. Максиминная μ/E оценка параметров многомерных возможностных распределений

Подход Wang Xizhao и Ha Minghu [88] в отличие от подхода, рассмотренного в предыдущей главе, не использует дополнительную информацию о значении уровня риска ошибочного определения оценки. В данной главе их идея будет расширена на многомерный случай.

Рассмотрим многомерное нечёткое число, имеющее симметричную, строго убывающую при увеличении расстояния от нуля функцию распределения μ_0:

$$\mu_0(\mathbf{x}) = \bar{f}(\|\mathbf{x}\|) = \max\{0, f(\|\mathbf{x}\|)\}. \tag{3.3}$$

где $\mathbf{x} \in \mathbb{R}^m$, f — строго убывающая на луче $[0, +\infty)$ функция, $f(0) = 1$.

Рассмотрим Q — семейство возможностных распределений, которые могут быть получены из данного нечёткого числа с помощью аффинных преобразований. Функции распределения, входящие в это семейство, могут быть представлены в виде

$$\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}, \tag{3.4}$$

где $\mathbf{c} \in \mathbb{R}^m$, A – положительно определённая симметричная матрица $m \times m$. Параметр $\mathbf{c}$ определяет расположение модального значения нечёткого числа и задаёт сдвиг функции распределения μ_0, $\mu(\mathbf{c}, \mathbf{c}, A) = 1$. Поворот и масштабирование задаётся матрицей A, при этом углы поворота распределения определяются ориентацией собственных векторов матрицы A, а изменение масштаба вдоль осей обратно пропорционально квадратному корню из её собственных чисел.

Заметим, что поверхностями равных уровней данного распределения являются эллипсоиды.

Обозначим через $E(\boldsymbol{\xi}) = \int_R \cdots \int \mu_{\boldsymbol{\xi}}(\mathbf{t}, \boldsymbol{\theta}) d\mathbf{t}$ меру нечёткости распределения $\mu_{\boldsymbol{\xi}}$ [90], $\boldsymbol{\theta}$ – параметр распределения, $\boldsymbol{\theta} = (\mathbf{c}, A)$, $\boldsymbol{\theta} \in \Theta$, $\mathbf{t} \in \mathbb{R}^m$.

Допустим, что известны значения $(\mathbf{x}_1, ..., \mathbf{x}_n)$ многомерной выборки $(\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_n)$ после проведения n наблюдений. В силу взаимной минисвязанности совместное распределение возможностей наблюдения значений $(\mathbf{x}_1, ..., \mathbf{x}_n)$ имеет вид:

$$\mu((\mathbf{x}_1, ..., \mathbf{x}_n), \boldsymbol{\theta}) = \bigwedge_{j=1}^{n} \mu_{\boldsymbol{\xi}}(\mathbf{x}_j, \boldsymbol{\theta}), \tag{3.5}$$

где $\mathbf{x_j} = (x_j^1, ..., x_j^m)$.

Принимая во внимание тот факт, что любые изменения параметров $\boldsymbol{\theta}$ непосредственно повлияют как на возможность появления тех или иных наблюдаемых значений, так и на значение $E(\boldsymbol{\xi})$, мы должны выбрать такое $\boldsymbol{\theta}$, чтобы максимизировать возможность (3.5) и при этом получить минимальное $E(\boldsymbol{\xi})$.

Введём следующее обозначение:

$$L(\boldsymbol{\theta}) = \bigwedge_{j=1}^{n} \mu_{\boldsymbol{\xi}}(\mathbf{x}_j, \boldsymbol{\theta}) / E(\boldsymbol{\xi}). \tag{3.6}$$

Определение 29. *Если найдётся такая оценка* $\hat{\boldsymbol{\theta}} \in \Theta$, *что* $L(\hat{\boldsymbol{\theta}}) = \max\limits_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta})$, *то* $\hat{\boldsymbol{\theta}}$ *называется максиминной* μ/E *оценкой параметра* $\boldsymbol{\theta}$.

Способ определения оценки параметров $\mathbf{c}, A$ даёт следующая теорема.

Теорема 12. *Пусть:*

1. $f : R \to R$ — *строго убывающая на луче* $[0, +\infty)$ *функция,* $f(0) = 1$.

2. $Q = \{\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A)\}$ — *семейство распределений, где* $\mathbf{c} \in \mathbb{R}^m$, A – *положительно определённая симметричная матрица,* $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$.

3. $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ — *наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ *минисвязанных возможностных величин, принадлежащих семейству распределений* Q, $\mathbf{x}_i \in \mathbb{R}^m, i \in [1, ..., n]$.

4. W — *многомерный эллипсоид минимальноого объёма, описанный вокруг множества* X.

Тогда максиминной оценкой параметра $\boldsymbol{\theta} = (\mathbf{c}, A)$ *является* $(\hat{\mathbf{c}}, \hat{A})$, *такие, что эллипсоид* W *задаётся уравнением* $(\mathbf{x} - \hat{\mathbf{c}})^T \hat{A}(\mathbf{x} - \hat{\mathbf{c}}) = 1/q^2$, $q = \arg\max\limits_{t \geqslant 0} t^m \bar{f}(t)$.

Доказательство:

Рассмотрим распределение $\mu_{\boldsymbol{\xi}}$ семейства Q с параметрами $\mathbf{c}, A$ и $\tilde{A} = p^2 A$, $p \in \mathbb{R}^1$. Пусть $\int\limits_R \cdots \int \mu_0(t^1, ..., t^m) dt^1 ... dt^m = I$. Нетрудно видеть, что тогда $E(\boldsymbol{\xi}) = Ib^1 \cdot ... \cdot b^m/(p)^m$, где $b^i = 1/\sqrt{v^i}$, v^i – i-ое собственное число матрицы $\tilde{A}$.

Согласно (3.6), имеем:

$$L(\mathbf{c}, A) = \bigwedge_{j=1}^{n} \frac{(p)^m \bar{f}(\sqrt{\mathbf{u}}\,)}{Ib^1 \cdot ... \cdot b^m}, \tag{3.7}$$

где $\mathbf{u} = (\mathbf{x} - \mathbf{c})^T A(\mathbf{x} - \mathbf{c})$.

Рассмотрим множество $Z = \{\mathbf{x}_{z_1}, ..., \mathbf{x}_{z_k}\}$ – наблюдаемые значения, на которых достигается минимум (3.7). Для них выполняется следующее соотношение: $\mu_{\boldsymbol{\xi}}(\mathbf{x}_{z_i}) = \mu_Z = \min\limits_{\mathbf{x} \in X} \mu_{\boldsymbol{\xi}}(\mathbf{x})$, $i \in [1, .., k]$. Тогда:

$$L(\mathbf{c}, A) = \bigwedge_{j=1}^{n} \frac{(p)^m \bar{f}(\sqrt{\mathbf{u}}\,)}{Ib^1 \cdot ... \cdot b^m} = \frac{(p)^m \mu_Z}{Ib^1 \cdot ... \cdot b^m} \tag{3.8}$$

Заметим, что поверхностями равного уровня для распределения, определённого согласно (3.4), являются эллипсоиды. Соответственно, все точки из множества Z находятся на одном из этих эллипсоидов (эллипсоид U), а

остальные точки – внутри него. Пусть $\tilde{\mathbf{c}} = (\tilde{c}^1, ..., \tilde{c}^m)$ и матрица $\tilde{A}$ определяют эллипсоид U.

Для того, чтобы найти μ_Z, возьмём точку $\mathbf{H} = (h^1, ..., h^m)$ на конце длинной оси эллипсоида U (рис. 3.4). Получим: $\mu_Z = \mu_{\boldsymbol{\xi}}(\mathbf{H}) = \bar{f}(p)$. Тогда $L(\tilde{\mathbf{c}}, \tilde{A}) = (p)^m \bar{f}(p)/Ib^1 \cdot ... \cdot b^m$.

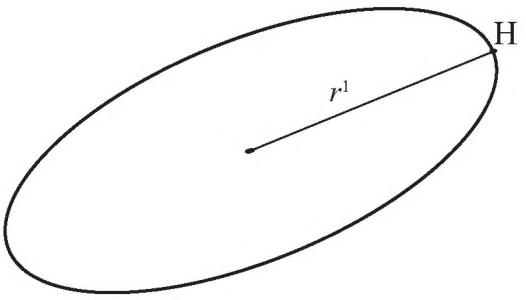


Рис. 3.4. Эллипс U. Двумерный случай

Учитывая, что объём эллипсоида U равен [51]

$$V_U = \frac{2\pi^{\frac{m}{2}}}{m\Gamma(\frac{m}{2})} b^1 \cdot ... \cdot b^m,$$

где Γ – гамма-функция.

Тогда

$$L(\tilde{\mathbf{c}}, \tilde{A}) = \frac{2\pi^{\frac{m}{2}}(p)^m \bar{f}(p)}{m\Gamma(\frac{m}{2}) I V_U}.$$

Так как необходимо найти оценку, максимизирующую $L(\mathbf{c}, A)$, возьмём $p = q$, максимизирующее функцию $(p)^m \bar{f}(p)$, и эллипсоид минимального объёма W в качестве эллипсоида U.

Тогда эллипсоид W задаётся уравнением $(\mathbf{x}-\hat{\mathbf{c}})^T \tilde{A}(\mathbf{x}-\hat{\mathbf{c}}) = 1$. Учитывая, что $A = \frac{1}{p^2}\tilde{A}$, получаем $(\mathbf{x} - \hat{\mathbf{c}})^T \hat{A}(\mathbf{x} - \hat{\mathbf{c}}) = 1/q^2$. ■

Пример 2. Рассмотрим μ/E оценку параметров двумерного распределения семейства $Q_\mu(\mathbf{c}, A)$, задаваемого следующими функциями:

1. Экспоненциальное распределение порождается функцией $f(t) = e^{-t^2}$, в этом случае $q = \arg\max\limits_{t \geqslant 0} t^2 \bar{f}(t) = 1$.

2. Линейное распределение задаётся функцией $f(t) = 1 - t$, здесь $q = \frac{2}{3}$.

В приложениях Б и В, соответственно, изображены графики функций распределения $\mu(\hat{\mathbf{c}}, \hat{A})$ для случаев 1 и 2, где: ∘ – наблюдаемые значения возможностной величины; —— – W, эллипс минимальной площади, описанный вокруг наблюдаемых значений; • – центр W; эллипсы, расположенные в нижней части рисунка, отражают линии равного уровня функции распределения μ.

3.1.3. Свойства μ/E оценки

Для того, чтобы определить качество оценки параметров, используют различные критерии. Основываясь на [90], введём определения используемых далее критериев.

Определение 30. *Пусть F представляет собой семейство функций распределения, $\bar{\boldsymbol{\xi}} = (\bar{\boldsymbol{\xi}}_1, \bar{\boldsymbol{\xi}}_2, ..., \bar{\boldsymbol{\xi}}_n)$ — выборка, $\bar{\boldsymbol{\xi}}_i = (\bar{\xi_i^1}, ..., \bar{\xi_i^m})$, $i = \overline{1,n}$; и пусть $T(\bar{\boldsymbol{\xi}})$ — статистика, функция распределения которой принадлежит семейству F. Будем говорить, что $T(\bar{\boldsymbol{\xi}})$ — достаточна по отношению к F, если совокупная функция распределения от $(T(\bar{\boldsymbol{\xi}}), \bar{\boldsymbol{\xi}})$ не зависит от $\mathbf{x}$.*

Каждая выборка предоставляет определённую информации о семействе. И данное определение показывает, что достаточная статистика предоставляет тот же объём информации о семействе, что и вся выборка. Из чего следует, что такая достаточная статистика может быть использована для уменьшения объёма выборочных данных без потери информации.

Определение 31. *Пусть F — семейство функций распределения, представленных следующим образом: $F = \{F(\mathbf{x}, \boldsymbol{\theta}) | \mathbf{x} = (x^1, ..., x^m)$, $\boldsymbol{\theta} \in \Theta, \Theta \subseteq \mathbb{R}^k\}, \bar{\boldsymbol{\xi}} = (\bar{\boldsymbol{\xi}}_1, \bar{\boldsymbol{\xi}}_2, ..., \bar{\boldsymbol{\xi}}_n)$ — выборка и $T(\bar{\boldsymbol{\xi}})$ — статистика. Тогда:*
1)$T(\bar{\boldsymbol{\xi}})$ — достаточная оценка вектора параметров $\boldsymbol{\theta}$, если $T(\bar{\boldsymbol{\xi}})$ достаточна по отношению к F;
2)$T(\bar{\boldsymbol{\xi}})$ — состоятельная оценка вектора параметров $\boldsymbol{\theta}$, если $\Pi(\{T(\bar{\boldsymbol{\xi}}) = \boldsymbol{\theta}\}) = 1$.
3)$T(\bar{\boldsymbol{\xi}})$ — максимально правдоподобная оценка вектора параметров $\boldsymbol{\theta}$, если:

$$M(\mathbf{x}_1, ..., \mathbf{x}_n, T(\bar{\boldsymbol{\xi}})) = \max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} M(\mathbf{x}_1, ..., \mathbf{x}_n, \boldsymbol{\theta}),$$

где $M(\mathbf{x}_1, ..., \mathbf{x}_n, \boldsymbol{\theta}) = \bigwedge_{j=1}^{n} F(x_j^1, ..., x_j^m, \boldsymbol{\theta})$, $\mathbf{x}_i = (x_i^1, ..., x_i^m)$, $i \in [1, n]$.

Состоятельность отражает возможность того, что полученная оценка параметров представляет собой действительные значения параметров. Максимально правдоподобные оценки строятся по методу максимального правдоподобия, откуда и берут своё название, и максимизируют функцию правдоподобия при фиксированной реализации выборки.

Докажем вспомогательную теорему:

Теорема 13. *Пусть* $\mu(\mathbf{x})$ *— унимодальная функция распределения вида* (3.3)*;* $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ *— наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ *минисвязанных возможностных величин с функцией распределения* $\mu(\mathbf{x})$*,* $\mathbf{x}_i \in \mathbb{R}^m$*,* $i \in [1, ..., n]$*;* $\mathbf{S}$ *— статистика, соответствующая* X *и удовлетворяющая условию:* $\mathbf{S} \subseteq W$*, где* W *— эллипсоид минимального объёма, описанный вокруг множества* X*.*

Тогда $\mu_{\mathbf{S}}(\mathbf{t}) = \sup\limits_{\mathbf{S}(\mathbf{x})=\mathbf{t}} \mu(\mathbf{x}; \mathbf{a})$*, т.е.* $\mathbf{S}$ *имеет то же распределение, что и любое из наблюдений* $\mathbf{x}$*.*

Доказательство.

Необходимость:

Так как наблюдения X являются минисвязанными, то согласно принципу обобщения Л. Заде [11], [21]:

$$\mu_{\mathbf{S}}(\mathbf{t}) = \sup_{\mathbf{S}(\mathbf{x_1},...,\mathbf{x_n})=\mathbf{t}} \min(\mu(\mathbf{x_1}), ..., \mu(\mathbf{x_n})).$$

По условию теоремы $\mathbf{S} \subseteq W$, следовательно, $\forall n > 1\ \forall \mathbf{x} \in \mathbb{R}^m$:

$$\mathbf{S}(\underbrace{\mathbf{x}, ..., \mathbf{x}}_{\text{n раз}}) = \mathbf{S}(\mathbf{x}).$$

Тогда $\forall X\ \exists \mathbf{x} \in W :\ \mathbf{S}(\mathbf{x_1}, ..., \mathbf{x_n}) = \mathbf{S}(\mathbf{x})$. Следовательно

$$\begin{aligned}\mu_{\mathbf{S}}(\mathbf{t}) &= \sup_{\mathbf{S}(\mathbf{x_1},...,\mathbf{x_n})=\mathbf{t}} \min(\mu(\mathbf{x_1}), ..., \mu(\mathbf{x_n})) \\ &\geqslant \sup_{\mathbf{S}(\mathbf{x},...,\mathbf{x})=\mathbf{S}(\mathbf{x})=\mathbf{t}} \min(\mu(\mathbf{x}), ..., \mu(\mathbf{x})) = \sup_{\mathbf{S}(\mathbf{x})=t} \mu(\mathbf{x}).\end{aligned}$$

Достаточность:

Пусть $\mathbf{S}(\mathbf{x}) = \mathbf{t}$ и $\mathbf{S}(\mathbf{x_1}, ..., \mathbf{x_n}) = \mathbf{t}$. Так как $\mathbf{S} \subseteq W$, то

$$\mu(\mathbf{x}) \geqslant \min(\mu(\mathbf{x_1}), ..., \mu(\mathbf{x_n})).$$

Тогда

$$\sup_{\mathbf{S}(\mathbf{x})=\mathbf{t}} \mu(\mathbf{x}) \geqslant \sup_{\mathbf{S}(\mathbf{x_1},...,\mathbf{x_n})=\mathbf{t}} \min(\mu(\mathbf{x_1}), ..., \mu(\mathbf{x_n})) = \mu_{\mathbf{S}}(\mathbf{t}).\ \blacksquare$$

Используя теорему 13, докажем, что оценка, полученная по теореме 12, обладает свойствами достаточности и состоятельности при фиксированной

матрице A.

Теорема 14. *Оценка* $\hat{\boldsymbol{\theta}} = (\hat{\mathbf{c}}, \hat{A})$, *полученная согласно теореме 12, является достаточной оценкой параметров* $\mathbf{c}, A$ *и состоятельной оценкой* $\mathbf{c}$ *при фиксированной матрице* A.

Доказательство:

а) Докажем достаточность оценки. Пусть $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ — наблюдаемые значения выборки $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ минисвязанных возможностных величин с функцией распределения возможностей $\mu(\mathbf{x})$, $\mathbf{x}_i \in \mathbb{R}^m$, $i \in [1, ..., n]$. Рассмотрим W — эллипсоид минимального объёма, описанный вокруг множества X, и точки, расположенные на концах его полуосей (рис. 3.5): $H_1 = (\hat{\mathbf{c}} + \mathbf{r}^1)$, $h_1 = (\hat{\mathbf{c}} - \mathbf{r}^1)$,...,$H_m = (\hat{\mathbf{c}} + \mathbf{r}^m)$, $h_m = (\hat{\mathbf{c}} - \mathbf{r}^m)$. Заметим, что для них выполняются следующие условия:

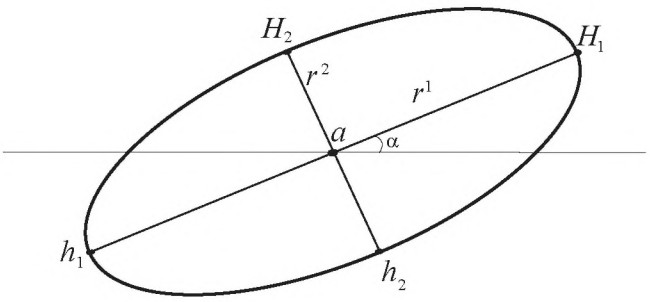


Рис. 3.5. Эллипс W. Двумерный случай

1. Согласно теореме 13 H_1, h_1,...,H_m, h_m являются минисвязанными возможностными величинами с распределением $\mu_{\boldsymbol{\xi}}$;

2. $\forall i = 1, .., n : \mu(H_1) \leqslant \mu(\mathbf{x_i})$, $\mu(h_1) \leqslant \mu(\mathbf{x_i})$,...,$\mu(H_m) \leqslant \mu(\mathbf{x_i})$, $\mu(h_m) \leqslant \mu(\mathbf{x_i})$.

Тогда:

$$\Pi(\{\boldsymbol{\xi_1} = \mathbf{x_1}, ..., \boldsymbol{\xi_n} = \mathbf{x_n}, \hat{\boldsymbol{\theta}} = (\hat{\mathbf{c}}, \hat{A})\}) =$$
$$= \Pi(\{\boldsymbol{\xi_1} = \mathbf{x_1}, ..., \boldsymbol{\xi_n} = \mathbf{x_n}, H_1 = (\hat{\mathbf{c}} + \mathbf{r}^1), h_1 = (\hat{\mathbf{c}} - \mathbf{r}^1), ..., H_m = (\hat{\mathbf{c}} + \mathbf{r}^m),$$
$$h_m = (\hat{\mathbf{c}} - \mathbf{r}^m)\}) \overset{1}{=} \mu(H_1, \hat{\boldsymbol{\theta}}) \wedge \mu(h_1, \hat{\boldsymbol{\theta}}) \wedge ... \wedge \mu(H_m, \hat{\boldsymbol{\theta}}) \wedge \mu(h_m, \hat{\boldsymbol{\theta}}) \wedge \bigwedge_{i=1}^{n} \mu(\mathbf{x_i}, \hat{\boldsymbol{\theta}}) \overset{2}{=}$$
$$\overset{2}{=} \mu(H_1, \hat{\boldsymbol{\theta}}) \wedge \mu(h_1, \hat{\boldsymbol{\theta}}) \wedge ... \wedge \mu(H_m, \hat{\boldsymbol{\theta}}) \wedge \mu(h_m, \hat{\boldsymbol{\theta}}) =$$
$$= \Pi(\{H_1 = (\hat{\mathbf{c}} + \mathbf{r}^1)\}) \wedge \Pi(\{h_1 = (\hat{\mathbf{c}} - \mathbf{r}^1)\}) \wedge ...$$
$$... \wedge \Pi(\{H_m = (\hat{\mathbf{c}} + \mathbf{r}^m)\}) \wedge \Pi(\{h_m = (\hat{\mathbf{c}} - \mathbf{r}^m)\}) \overset{1}{=}$$
$$\overset{1}{=} \Pi(\{H_1 = (\hat{\mathbf{c}} + \mathbf{r}^1), h_1 = (\hat{\mathbf{c}} - \mathbf{r}^1), ..., H_m = (\hat{\mathbf{c}} + \mathbf{r}^m), h_m = (\hat{\mathbf{c}} - \mathbf{r}^m)\}) =$$
$$= \Pi(\{\hat{\boldsymbol{\theta}} = (\hat{\mathbf{c}}, \hat{A})\}).$$

Следовательно, оценка $\hat{\boldsymbol{\theta}} = (\hat{\mathbf{c}}, \hat{A})$ является достаточной оценкой параметров распределения.

b) Докажем состоятельность оценки.

Из теоремы 13 следует, что $\hat{\mathbf{c}}$ имеет то же распределение, что и $\mathbf{x} \in X$. Тогда $\Pi(\{\hat{\mathbf{c}}(X) = \hat{\mathbf{c}}\}) = \mu(\hat{\mathbf{c}}; \hat{\mathbf{c}}, \hat{A}) = 1$. Следовательно, оценка параметра $\hat{\mathbf{c}}$ является состоятельной оценкой при фиксированной матрице A.

Теорема 15. *Оценка $\hat{\boldsymbol{\theta}} = (\hat{\mathbf{c}}, \hat{A})$, полученная согласно теореме 12, является максимально правдоподобной оценкой параметра* $\mathbf{c}$ *при фиксированных собственных числах матрицы A.*

Доказательство:

Рассмотрим множество $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$, где $\mathbf{x}_i \in \mathbb{R}^m, i \in [1, ..., n]$, X — наблюдаемые значения возможностной величины, для которого была построена оценка $\hat{\boldsymbol{\theta}} = (\hat{\mathbf{c}}, \hat{A})$. Пусть $\widehat{W}$ – соответствующий оценке эллипсоид минимального объёма, описанный вокруг множества X.

Методом от противного докажем, что данная оценка является максимально правдоподобной среди прочих оценок набора параметров при фиксированных собственных числах матрицы $\hat{A}$. Предположим, что существует другая оценка $\tilde{\boldsymbol{\theta}} = (\tilde{\mathbf{a}}, \tilde{A})$, собственные числа матриц $\tilde{A}$ и $\hat{A}$ совпадают и

$$M(\mathbf{x}_1, ..., \mathbf{x}_n, \tilde{\boldsymbol{\theta}}) > M(\mathbf{x}_1, ..., \mathbf{x}_n, \hat{\boldsymbol{\theta}}).$$

Пусть $\mathbf{x}_{\hat{\mathbf{k}}} = \operatorname{argmin}\mu(\mathbf{x}_j, \hat{\boldsymbol{\theta}})$ и $\mathbf{x}_{\tilde{\mathbf{k}}} = \operatorname{argmin}\mu(\mathbf{x}_j, \tilde{\boldsymbol{\theta}})$, $j \in [1,..,n]$. Согласно предположению $\mu(\mathbf{x}_{\tilde{k}}, \tilde{\boldsymbol{\theta}}) > \mu(\mathbf{x}_{\hat{k}}, \hat{\boldsymbol{\theta}})$. Так как $\mu(\mathbf{x}, \mathbf{c}, A) = \bar{f}\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)$, то

$$\bar{f}\big(\sqrt{(\mathbf{x}_{\tilde{k}}-\tilde{\mathbf{c}})^T \tilde{A}(\mathbf{x}_{\tilde{k}}-\tilde{\mathbf{c}})}\,\big) > \bar{f}\big(\sqrt{(\mathbf{x}_{\hat{k}}-\hat{\mathbf{c}})^T \hat{A}(\mathbf{x}_{\hat{k}}-\hat{\mathbf{c}})}\,\big).$$

По определению (3.3) f — строго убывающая на луче $[0, +\infty)$ функция, следовательно

$$\sqrt{(\mathbf{x}_{\tilde{k}}-\tilde{\mathbf{c}})^T \tilde{A}(\mathbf{x}_{\tilde{k}}-\tilde{\mathbf{c}})} < \sqrt{(\mathbf{x}_{\hat{k}}-\hat{\mathbf{c}})^T \hat{A}(\mathbf{x}_{\hat{k}}-\hat{\mathbf{c}})}. \tag{3.9}$$

Так как согласно (3.4) A является положительно определённой симметричной матрицей, то она может быть представлена через разложение Холецкого в виде: $A = P^T P$, где P — верхняя треугольная матрица со строго положительными элементами на диагонали. Тогда $\sqrt{x^T A x} = \sqrt{(Px)^T Px}$.

Рассмотрим преобразования $\tilde{y}(\mathbf{x}) = \tilde{P}(\mathbf{x}-\tilde{\mathbf{c}})$ и $\hat{y}(\mathbf{x}) = \hat{P}(\mathbf{x}-\hat{\mathbf{c}})$, где $\tilde{A} = \tilde{P}^T\tilde{P}$ и $\hat{A} = \hat{P}^T\hat{P}$. Заметим, что $\mu(\mathbf{x}, \tilde{\boldsymbol{\theta}}) = \bar{f}(\|\tilde{y}(\mathbf{x})\|)$ и $\mu(\mathbf{x}, \hat{\boldsymbol{\theta}}) = \bar{f}(\|\hat{y}(\mathbf{x})\|)$.

Пусть $\widetilde{Y}$ и $\widehat{Y}$ — гиперсферы с радиусами $\tilde{y}(\mathbf{x}_{\tilde{k}})$ и $\hat{y}(\mathbf{x}_{\hat{k}})$ и центрами в начале координат. Причём, $\tilde{y}(X) \in \widetilde{Y}$, $\hat{y}(X) \in \widehat{Y}$ и $\widehat{Y} = \hat{y}(\widehat{W})$. Из (3.9) следует, что $\tilde{y}(\mathbf{x}_{\tilde{k}}) < \hat{y}(\mathbf{x}_{\hat{k}})$, тогда $V(\widetilde{Y}) < V(\widehat{Y})$, где $V(\cdot)$ — объём.

Рассмотрим эллипсоид $\widetilde{W} = \tilde{y}^{-1}(\widetilde{Y})$. Так как $\tilde{y}(X) \in \widetilde{Y}$, то $X \in \widetilde{W}$, то есть $\widetilde{W}$ – описанный эллипсоид множества X. Так как собственные числа матриц $\tilde{A}$ и $\hat{A}$ совпадают, то $V(\widetilde{Y})/V(\widehat{Y}) = V(\widetilde{W})/V(\widehat{W})$. Тогда $V(\widetilde{W}) < V(\widehat{W})$, что невозможно, т.к. эллипсоид минимального объёма $\widehat{W}$ является единственным [51]. Данное противоречие доказывает теорему. ■

3.2. Оценка параметров многомерных возможностных распределений в случае Архимедовой t-нормы

3.2.1. Нахождение оценки параметра c при фиксированной матрице A

Как уже было отмечено выше, любая непрерывная Архимедова t-норма может быть представлена с помощью непрерывного аддитивного генератора g:

$$T(\mu_1, ..., \mu_n) = g^{[-1]}(g(\mu_1) + ... + g(\mu_n)),$$

где $g : [0,1] \to [0,\infty]$, $g(1) = 0$, g непрерывная, строго убывающая функция.

Используя данное представление, найдём оценку параметра $\mathbf{c}$ в случае Архимедовой t-нормы.

Теорема 16. *Пусть:*

1. $f : \mathbb{R}^1 \to \mathbb{R}^1$ — *строго убывающая на луче* $[0, +\infty)$ *функция,* $f(0) = 1$.

2. $Q = \{\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A)\}$ — *семейство распределений, где* $\mathbf{c} \in \mathbb{R}^m$, A – *положительно определённая симметричная матрица* $m \times m$, $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \bar{f}\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$.

3. $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ — *наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ *t-связанных возможностных величин, принадлежащих семейству распределений* Q, *где t- Архимедова t-норма,* $\mathbf{x}_i \in \mathbb{R}^m, i = 1, ..., n$.

4. $\exists i, j : \mathbf{x}_i \neq \mathbf{x}_j$.

Тогда оценкой параметра $\mathbf{c}$ *является* $\mathbf{c} = \bar{\mathbf{x}}$, *где* $\bar{\mathbf{x}} = \frac{\sum_{i=1}^n \mathbf{x}_i}{n}$.

Доказательство:

Применим аддитивный генератор к функции совместного распределения возможностей наблюдения значений $(\mathbf{x}_1, ..., \mathbf{x}_n)$, где каждая функция распределения имеет вид (3.4). Получаем:

$$\mu((\mathbf{x}_1, ..., \mathbf{x}_n), \boldsymbol{\theta}) = g^{[-1]}\Big(\sum_{i=1}^n g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i - \mathbf{c})^T A(\mathbf{x}_i - \mathbf{c})}\big)\big)\Big)$$

Тогда функция $L(\boldsymbol{\theta})$ примет вид:

$$L(\boldsymbol{\theta}) = g^{[-1]}\Big(\sum_{i=1}^n g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i - \mathbf{c})^T A(\mathbf{x}_i - \mathbf{c})}\big)\big)\Big)/E(\mu_{\boldsymbol{\xi}}), \tag{3.10}$$

где $E(\mu_{\boldsymbol{\xi}}) = \int\limits_R \cdots \int \mu_{\boldsymbol{\xi}}(t^1, ..t^m, \boldsymbol{\theta}) dt^1 ... dt^m$ – мера нечёткости распределения $\mu_{\boldsymbol{\xi}}$.

Найдём максимум функции $L(\boldsymbol{\theta})$ для определения оценки параметра $\boldsymbol{\theta} = (\mathbf{c}, A)$:

$$\frac{g^{[-1]}\Big(\sum_{i=1}^n g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i - \mathbf{c})^T A(\mathbf{x}_i - \mathbf{c})}\big)\big)\Big)}{Ib^1 \cdot ... \cdot b^m} \to \max_{\mathbf{c} \in \mathbb{R}^m}. \tag{3.11}$$

Сделаем замену: $\mathbf{d} = \bar{\mathbf{x}} - \mathbf{c}$. При этом $\mathbf{x}_i - \mathbf{c} = \mathbf{x}_i - \bar{\mathbf{x}} + \bar{\mathbf{x}} - \mathbf{c} = \mathbf{x}_i - \bar{\mathbf{x}} + \mathbf{d}$. И $(\mathbf{x}_i - \bar{\mathbf{x}} + \mathbf{d})^T A(\mathbf{x}_i - \bar{\mathbf{x}} + \mathbf{d}) = (\mathbf{x}_i - \bar{\mathbf{x}})^T A(\mathbf{x}_i - \bar{\mathbf{x}}) + \mathbf{d}^T A\mathbf{d}$.

Получаем задачу следующего вида:

$$\frac{g^{[-1]}\Big(\sum_{i=1}^{n} g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\big)\big)\Big)}{Ib^1\cdot\ldots\cdot b^m}\to\max_{\mathbf{d}\in\mathbb{R}^m}. \tag{3.12}$$

Выражение $Ib^1\cdot\ldots\cdot b^m$ не зависит от $\mathbf{d}$, поэтому (3.12) можно упросить следующим образом:

$$g^{[-1]}\Big(\sum_{i=1}^{n} g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\big)\big)\Big)\to\max_{\mathbf{d}\in\mathbb{R}^m}. \tag{3.13}$$

Функция $g^{[-1]}(t)$ является псевдообратной функцией для $g(t)$ и согласно (1.9) строго убывает при $t\in[0,a], a\in\mathbb{R}^1$ и $g(0)=a$ и принимает значение 0 при $t>a$.

Рассмотрим случай $t>a$. Если

$$\sum_{i=1}^{n} g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\big)\big)>a,$$

то $\mu((\mathbf{x}_1,\ldots,\mathbf{x}_n),\boldsymbol{\theta})=0$. Тогда $L(\boldsymbol{\theta})=0$ и на луче $t>a$ точки максимума функции $L(\boldsymbol{\theta})$ отсутствуют.

Таким образом, следует рассматривать только случай, когда $t\in[0,a]$. На этом интервале $g^{[-1]}$ строго убывает и задача (3.13) эквивалентна задаче

$$\sum_{i=1}^{n} g\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\big)\big)\to\min_{\mathbf{d}\in\mathbb{R}^m}.$$

Для нахождения значения $\mathbf{d}$, при котором достигается максимум, найдём производную и приравняем к нулю полученное выражение:

$$\sum_{i=1}^{n}\Big[g'_f\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\big)\big)\cdot\bar{f}'_d\Big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\Big)\cdot$$
$$\cdot\Big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\Big)'_d\Big] \;=\; 0.$$

$$\sum_{i=1}^{n}\Big[g'_f\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\big)\big)\cdot\bar{f}'_d\Big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\,\Big)\cdot$$
$$\cdot\big((\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}\big)'_d\cdot\big((\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}\big)^{-\frac{1}{2}}\Big]=0.$$

$$\sum_{i=1}^{n}\Big[g'_f\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\big)\big)\cdot\bar{f}'_d\Big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}}+\mathbf{d})^T A(\mathbf{x}_i-\bar{\mathbf{x}}+\mathbf{d})}\Big)\cdot$$
$$\cdot\,(\mathbf{d}^T(A^T+A))\cdot\big((\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}\big)^{-\frac{1}{2}}\Big]\;=\;0.$$

Заметим, что выражение $\mathbf{d}^T(A^T+A)$ не зависит от $\mathbf{x}$, его можно вынести за знак суммы:

$$(\mathbf{d}^T(A^T+A))\cdot\sum_{i=1}^{n}\Big[g'_f\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\big)\big)\cdot$$
$$\cdot\bar{f}'_d\Big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}}+\mathbf{d})^T A(\mathbf{x}_i-\bar{\mathbf{x}}+\mathbf{d})}\Big)\cdot\big((\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}\big)^{-\frac{1}{2}}\Big]=0.$$

Данное выражение сожержит несколько сомножителей. Убедимся, что максимум функции $L(\boldsymbol{\theta})$ достигается только в случае, когда $\mathbf{d}^T(A^T+A)=0$.

Рассмотрим выражение

$$\sum_{i=1}^{n}\Big[g'_f\big(\bar{f}\big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}}\big)\big)\cdot\bar{f}'_d\Big(\sqrt{(\mathbf{x}_i-\bar{\mathbf{x}}+\mathbf{d})^T A(\mathbf{x}_i-\bar{\mathbf{x}}+\mathbf{d})}\Big)\cdot$$
$$\cdot\,\big((\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}\big)^{-\frac{1}{2}}\Big]. \qquad (3.14)$$

В силу строгой монотонности $g'_f(t)\neq 0$ и $f'_d(t)\neq 0$.

Матрица A является положительно определённой, следовательно, $(\mathbf{x}_i-\bar{\mathbf{x}})^T A(\mathbf{x}_i-\bar{\mathbf{x}})+\mathbf{d}^T A\mathbf{d}=0$ только в случае, когда $\mathbf{x}_i=\bar{\mathbf{x}}$. Однако, по условию 4 должно существовать хотя бы одно $i:\mathbf{x}_i\neq\bar{\mathbf{x}}$. Следовательно, рассматриваемая сумма (3.14) должна быть строго положительной.

Тогда можно утверждать, что

$$\mathbf{d}^T(A^T+A)=0.$$

Так как A –положительно определённая матрица, то $(A^T+A)\neq 0$, тогда $\mathbf{d}^T=0$ и $\mathbf{c}=\bar{\mathbf{x}}$. ■

Заметим, что в случае Архимедовой t-нормы центр распределения находится через среднее арифметическое наблюдаемых значений выборки, а при t-норме минимум – как центр эллипсоида минимального объёма, описанного вокруг наблюдаемых значений [20]. То есть центры совпадут только при симметричном распределении наблюдаемых значений выборки.

Пример 3. Рассмотрим пример нахождения оценки параметра центра распределения семейства $Q_\mu(\mathbf{c}, A)$ в двумерном случае для выборки на основе нормального распределения (рис. 3.6). Произведём оценку для случая t-нормы минимум и t-нормы произведение.

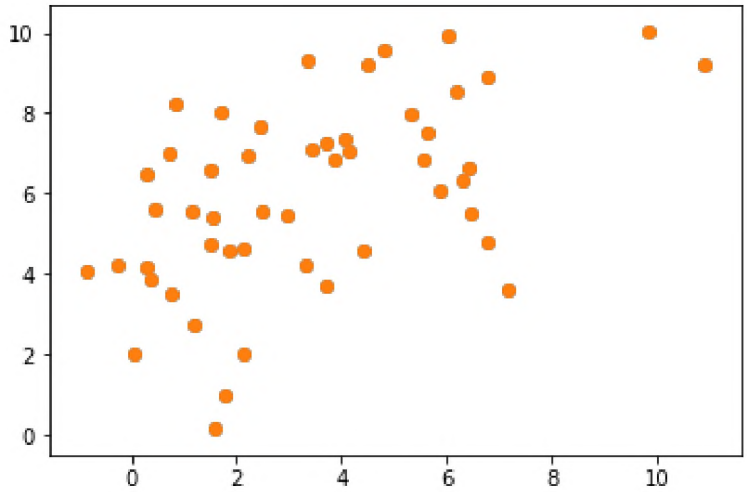


Рис. 3.6. Двумерное нормальное распредедение

В случае сильнейшей t-нормы согласно [20] требуется найти эллипс минимальной площади (в двумерном пространстве), описанный вокруг точек распределения. Для t-нормы произведения будем пользоваться тем, что она входит в класс Архимедовых t-норм, следовательно, можем применить доказанную ранее теорему 19. Рисунок 3.6 демонстрирует полученный результат, где ★ – центр распределения при t-норме минимум, ▲ – центр при t-норме произведение.

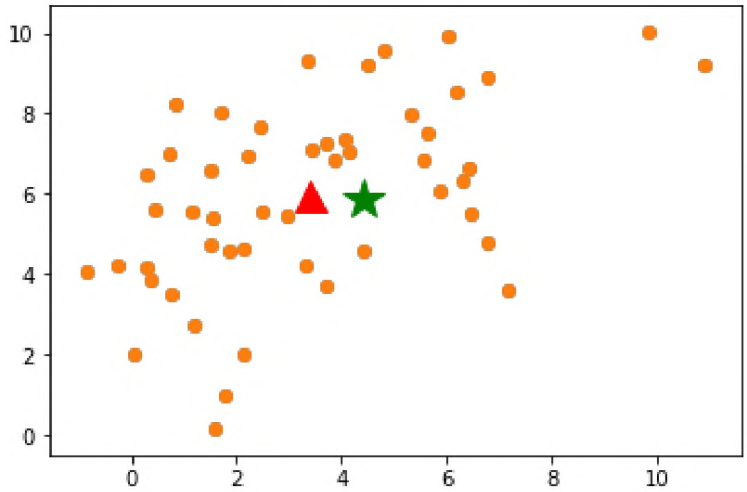


Рис. 3.7. Восстановление параметров двумерного нормального распределения

Как можно видеть из рисунка 3.7, оценки центров распределений не совпадают. Это является следствием различия применяемых t-норм. И демонстрирует важность выбора типа t-нормы под конкретные данные для получения результата, наилучшим образом отражающего действительность.

3.2.2. Оценка параметра A

Для нахождения оценки параметра A при известном $\mathbf{c}$ требуется решить задачу (3.11) относительно матрицы A. Аналитическое решение этой задачи приводит к громоздким выражениям. Например, в приложении А приведено выражение, выведенное аналитически, для нахождения одного из собственных чисел матрицы A для двумерной выборки объёмом 5 точек и t-нормы произведения. Заметим, что с ростом количества точек, растёт объём выражения и сложность его вывода.

Поэтому для решения этой задачи целесообразно использовать численные методы оптимизации. Автором был разработан и реализован следующий алгоритм:

1. Найти значение параметра $\mathbf{c}$ через вектор средних значений выборки по каждому измерению согласно теореме 19.

2. Задать произвольные начальные значения для собственных чисел $\mathbf{q}$ и собственных векторов $\mathbf{v}$ матрицы A и представить её в виде разложения: $A = VQV^T$, где Q – диагональная матрица собственных чисел, а V – матрица собственных векторов.

3. Подставить все компоненты в (3.11), выбрав в качестве $g^{[-1]}(t)$ и $g(t)$ конкретные представления (к примеру, для t-нормы произведения $g(t) = -ln(t)$ и $g^{[-1]}(t) = e^{-t}$).

4. Оптимизировать полученную функцию численными методами.

Для численного нахождения параметров была написана программа на языке программирования Python, использовались библиотеки numpy, scipy и sympy. В качестве функции оптимизации использовалась функция fmin() модуля scipy.optimize.optimize, реализующая метод Нелдера — Мида (метод деформируемого многогранника) [80].

Рассмотрим результаты работы программы в случае двух разных наборов данных.

В качестве первого варианта данных за основу взята выборка из 25 двумерных точек, равномерно распределённых по окружности (рис. 3.8).

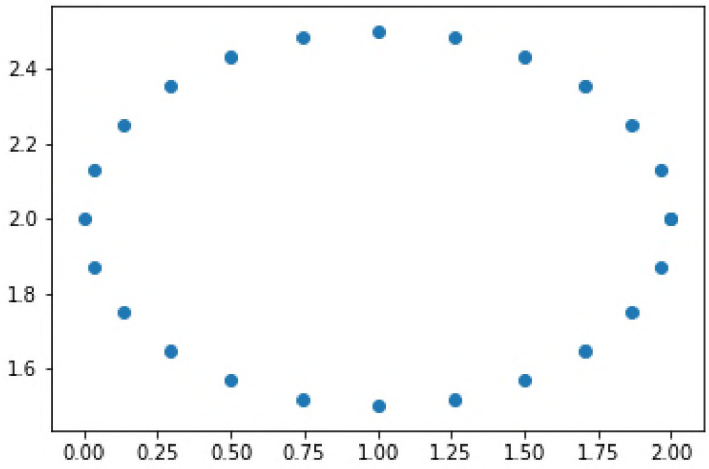


Рис. 3.8. Выборка 1

В качестве второй выборки использовалось нормальное распределение (рис. 3.9) с центром в (4; 4) и дисперсией (2.8; 1.9).

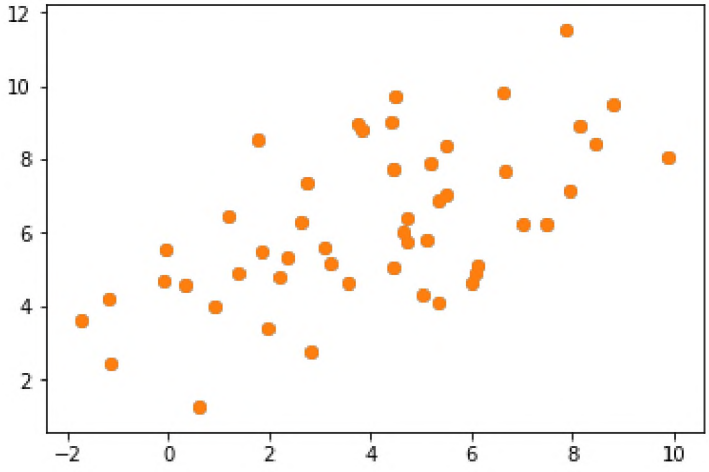


Рис. 3.9. Выборка 2

Вычисления производились на основе t-нормы произведения, входящей в класс Архимедовых и представленной с помощью аддитивного генератора $g(\mathbf{t}) = -ln(\mathbf{t})$.

В качестве оценки параметра $\mathbf{c}$ выступал вектор средних значений выборки по каждому измерению согласно 19. В результате работы программы были получены следующие значения:

1 случай:

$$A = \begin{pmatrix} 1.964 \cdot 10^{-1} & -2.285 \cdot 10^{-9} \\ -2.285 \cdot 10^{-9} & 4.082 \cdot 10^{-1} \end{pmatrix},$$

$$c = (1.040, 1.999).$$

2 случай:

$$A = \begin{pmatrix} 0.042 & -0.016 \\ -0.016 & 0.057 \end{pmatrix},$$

$$c = (4.041, 6.225).$$

3.3. Выводы по третьей главе

Третья глава посвящена методам нахождения оценки параметров многомерных возможностных распределений.

В параграфе 3.1 рассматривается случай минисвязанных возможностных величин.

Параграф 3.1.1 посвящён оценке параметров в случае заданного уровня возможности и описывает метод оценивания, основанный и расширяющий подход Dug Hun Hong.

Была доказана следующая теорема:

Теорема 17. *Пусть:*

1. $f : \mathbb{R}^1 \to \mathbb{R}^1$ — *строго убывающая на луче* $[0, +\infty)$ *функция,* $f(0) = 1$.

2. $Q = \{\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A)\}$ — *семейство распределений, где* $\mathbf{c} \in \mathbb{R}^m$, A – *положительно определённая симметричная матрица* $m \times m$, $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$.

3. $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ — *наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ *минисвязанных возможностных величин, принадлежащих семейству распределений* Q, $\mathbf{x}_i \in \mathbb{R}^m, i = 1, ..., n$.

4. W — *многомерный эллипсоид минимальноого объёма, описанный вокруг множества* X, *заданный параметрами* $(\breve{\mathbf{c}}, \breve{A})$, *где* $\breve{\mathbf{c}}$ – *центр эллипсоида* W, $\breve{A}$ – *матрица, задающая ориентацию осей эллипсоида и длину его радиусов.*

Тогда оценкой параметра $\boldsymbol{\theta}_\alpha = (\mathbf{c}, A)$ *является* $\hat{\boldsymbol{\theta}}_\alpha = (\hat{\mathbf{c}}, \hat{A}) = (\breve{\mathbf{c}}, k\breve{A})$, *где* $k = (f^{-1}(\alpha))^2$.

Также рассмотрен пример нахождения оценки параметров распределения в двумерном случае для функции $f(t) = e^{-t^2}$.

На основе проведённого исследования и рассмотренного примера сделан вывод, что параметр α влияет на форму распределения: чем больше значение параметра, тем более пологим становится распределение.

Параграф 3.1.2 описывает метод нахождения максиминной μ/E оценки параметров многомерных возможностных распределений и является расширением подхода Wang Xizhao и Ha Minghu на многомерный случай.

В ходе исследования была доказана теорема, описывающая способ определения оценки параметров $\mathbf{c}, A$:

Теорема 18. *Пусть:*

1. $f : R \to R$ — *строго убывающая на луче* $[0, +\infty)$ *функция,* $f(0) = 1$.

2. $Q = \{\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A)\}$ — *семейство распределений, где* $\mathbf{c} \in \mathbb{R}^m$, A – *положительно определённая симметричная матрица,* $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$.

3. $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ — *наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ *минисвязанных возможностных величин, принадлежащих семейству распределений* Q, $\mathbf{x}_i \in \mathbb{R}^m, i \in [1, ..., n]$.

4. W — *многомерный эллипсоид минимального объёма, описанный вокруг множества* X.

Тогда максиминной оценкой параметра $\boldsymbol{\theta} = (\mathbf{c}, A)$ *является* $(\hat{\mathbf{c}}, \hat{A})$, *такие, что эллипсоид* W *задаётся уравнением* $(\mathbf{x} - \hat{\mathbf{c}})^T \hat{A}(\mathbf{x} - \hat{\mathbf{c}}) = 1/q^2$, $q = \arg\max\limits_{t \geqslant 0} t^m \bar{f}(t)$.

Рассмотрен пример нахождения параметров двумерных распределений, задаваемых функциями:

1. Экспоненциальное распределение порождается функцией $f(t) = e^{-t^2}$, в этом случае $q = \arg\max\limits_{t \geqslant 0} t^2 \bar{f}(t) = 1$.

2. Линейное распределение задаётся функцией $f(t) = 1 - t$, здесь $q = \frac{2}{3}$.

В параграфе 3.1.3 доказывается, что полученная таким способом оценка является достаточной оценкой параметров $\mathbf{c}, A$, состоятельной оценкой $\mathbf{c}$ при фиксированной матрице A и максимально правдоподобной оценкой параметра $\mathbf{c}$ при фиксированных собственных числах матрицы A.

В параграфе 3.2 производится нахождение оценки параметра c, доказывается следующая теорема:

Теорема 19. *Пусть:*

1. $f : \mathbb{R}^1 \to \mathbb{R}^1$ — *строго убывающая на луче* $[0, +\infty)$ *функция,* $f(0) = 1$.

2. $Q = \{\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A)\}$ — *семейство распределений, где* $\mathbf{c} \in \mathbb{R}^m$, A – *положительно определённая симметричная матрица* $m \times m$, $\mu_{\boldsymbol{\xi}}(\mathbf{x}, \mathbf{c}, A) = \bar{f}\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big) = \max\{0, f\big(\sqrt{(\mathbf{x}-\mathbf{c})^T A(\mathbf{x}-\mathbf{c})}\,\big)\}$.

3. $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ — *наблюдаемые значения выборки* $(\boldsymbol{\xi}_1, .., \boldsymbol{\xi}_n)$ t-*связанных возможностных величин, принадлежащих семейству распределений* Q, *где* t- *Архимедова* t-*норма,* $\mathbf{x}_i \in \mathbb{R}^m, i = 1, ..., n$.

4. $\exists i, j : \mathbf{x}_i \neq \mathbf{x}_j$.

Тогда оценкой параметра $\mathbf{c}$ *является* $\mathbf{c} = \bar{\mathbf{x}}$, *где* $\bar{\mathbf{x}} = \frac{\sum_{i=1}^{n} \mathbf{x}_i}{n}$.

Также представлен алгоритм численного нахождения параметра A, отвечающего за поворот и масштабирование.

Продемонстрирована работа программы, позволяющей находить оценку параметров многомерных возможностных распределений в случае Архимедовой t-нормы.

4. Разработка и тестирование алгоритмов построения нечётких моделей с применением разработанного математического аппарата

4.1. Применение методов оценки параметров возможностных распределений в задачах обучения машин нечёткого вывода

В предыдущей главе были рассмотрены методы получения оценки параметров многомерных возможностных распределений. Одним из практических применений полученных результатов является обучение машины нечёткого вывода (или нечёткого регулятора) — системы, которая основанна на нечёткой логике, т.е. она анализирует входные значения в терминах лингвистических переменных, осуществляет их отображение во множество выходных значений на основе базы нечётких правил.

Системы нечёткого контроля применяются в различных областях:
– в составе комбинированных систем оптимального управления, у которых в прямом контуре задействованы обычные регуляторы, а в дополнительном контуре используются нечёткие контроллеры, подстраивающие коэффициенты усиления контроллера прямого контура в зависимости от изменяющихся условий;
– при решении задач алгоритмической обработки информации от изучаемого объекта (задачи фильтрации);
– в помехозащищённых адаптивных системах автоматического управления;
– в системах с нечёткой процедурой проверки статистических гипотез [14].

4.1.1. Понятие нечёткого правила

Нечёткое правило – это выражение вида:

$$If\ \mathbf{x}\ is\ A\ then\ \mathbf{y}\ is\ B,$$

где A и B – нечёткие термы, характеризующиеся соответствующими функциями распределения. Благодаря их краткости нечёткие правила часто используются для фиксации нечётких моделей рассуждений, которые играют

важную роль в человеческой способности принимать решения в условиях неопределённости.

В качестве примера приведём следующее выражение:

$$\textit{if давление is высокое, then объём is маленький},$$

где *давление* и *объём* – лингвистические переменные [11], а *высокий* и *маленький* – нечёткие термы, являющиеся названиями соответствующих возможностных распределений.

В 1983 году Takagi и Sugeno [81] предложили другой вид нечётких правил, в которых нечёткие множества используются только в предпосылке правила. В качестве примера данного подхода приведём правило, описывающее силу сопротивления, действующую на движущийся объект:

$$\textit{If скорость is высокая, then сила} = k * (\textit{скорость})^2,$$

где *высокая* – лингвистический терм, характеризующийся функцией распределения. Однако выходная часть правила представляет собой классическую функцию входной переменной *скорость*, которая обычно задаётся полиномом. Модель, в которой используются полиномы нулевого порядка, называют моделями нечёткого вывода Сугено 0-го порядка, а модели, включающие линейные функции – моделями Сугено 1-го порядка.

4.1.2. Описание системы нечёткого вывода

Системы нечёткого вывода также известны как нечёткие регуляторы или нечёткие системы. Они состоят из 5 функциональных блоков (рис. 4.1) [58]:

- база правил, содержащая нечёткие правила;
- база термов, определяющая функции распределения нечётких термов для входных и выходных переменных;
- блок нечёткого вывода, осуществляющий нечёткий логический вывод;
- блок фаззификации, который преобразует чёткие входные значения в лингвистическое описание;
- блок дефаззификации, отображающий нечёткие результаты в чёткие выходные значения.

Обычно базу правил и базу термов в совокупности называют базой знаний. Отметим, что в моделях Сугено блок дефаззификации не требуется, т.к. результатом работы нечёткого вывода является чёткое значение.

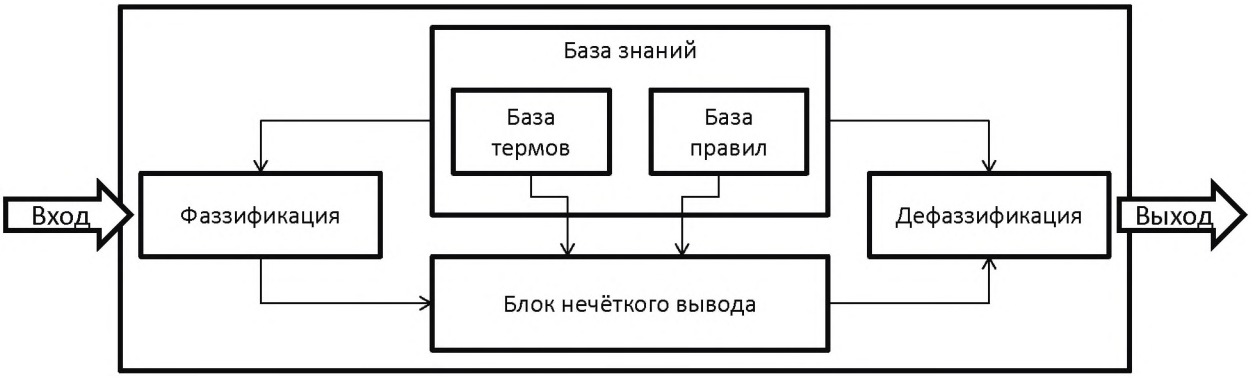


Рис. 4.1. Система нечёткого вывода

Нечёткий вывод включает в себя следующие шаги:

1. Сопоставление входных значений с функциями распределения нечётких термов с целью получения соответствующих значений принадлежности.

2. Объединение значений функций распределения, входящих в предпосылку одного правила, с помощью t-нормы для получения уровня активации правила.

3. Вычисление взвешенного выходного значения для каждого правила с учётом его уровня активации.

4. Объединение взвешенных выходов правил и вычисление чёткого выходного значения.

Существуют разные типы нечёткого вывода, различающиеся видом используемых правил, процедурой дефаззификации и др. параметрами.

Известно, что машина нечёткого вывода может аппроксимировать любую непрерывную функцию с заданной точностью [63].

4.1.3. Алгоритм проектирования систем нечёткого вывода

Рассмотрим алгоритм проектирования нечёткого контроллера Cluster Estimation, который был предложен Stepohen L. Chiu в 1994 году [37]. Он включает следующие этапы:

1. Проведение кластеризации данных при помощи модифицированного алгоритма Mountain method. Данная модификация основана на алгоритме, предложенном Yager R. R. и Filev D. P. в [93]. Для каждой точки данных вычисляется потенциал, который сильно зависит от числа ближайших соседей. Далее выбирается точка с максимальным потенциалом – она является центром первого кластера, потенциал соседних к ней точек уменьшается и заново вычисляется потенциал точек. Так происходит, пока максимальный потенциал не уменьшится до определённого уровня. Параметром алгоритма является радиус r, определяющий степень соседства точек.

2. Для каждого кластера создаётся правило, для которого функция распределения нечёткого терма определяется как нечёткое число с модальным значением a в центре кластера (входная часть), а коэффициент нечёткости b определяется параметром алгоритма.

3. Начальное значение выходной части правил для метода Сугено 0 приравнивается выходной части координат центров кластеров.

4. Выходная часть правил оптимизируется, используя метод наименьших квадратов.

Отметим, что в пункте 2 коэффициент нечёткости b задаётся произвольно, что является недостатком подхода Stepohen L. Chiu. Попробуем улучшить алгоритм с помощью метода, разработанного в главе 3. Будем действовать по следующему алгоритму (μ/E алгоритм 1):

1. Кластеризовать исходное множество входных примеров с помощью алгоритма k-средних. На каждой итерации перевычисляется центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике. Алгоритм завершается, когда на какой-либо итерации не происходит изменения центров кластеров.

2. Для каждого кластера создать правило, для которого функция распределения нечёткого терма определяется посредствам использования μ/E оценки для множества точек кластера согласно теореме 12.

3. Начальное значение выходной части правил для алгоритма Сугено 0 положить равным выходной части координат центров кластеров.

4. Оптимизировать выходную часть правил, используя метод наименьших квадратов.

Чтобы оценить качество работы μ/E алгоритма 1, рассмотрим далее его работу на тестовых данных, которые были использованы для тестирования алгоритма Cluster Estimation в работе [37].

4.1.4. Аппроксимация нелинейной функции одной переменной

В качестве первого теста в работе [37] рассматривалась задача аппроксимации нелинейной функции одной переменной $z = \sin y/y$. Обучение и тестирование нечёткого регулятора проводилось на одном и том же наборе данных из 100 значений равномерно распределённых на отрезке [-10,10]. В качестве функции распределения для нечётких термов использовались экспоненциальная функция $f(t) = e^{-t^2}$, параметры которой подбирались экспериментально, и функция распределения, используемая алгоритмом нечётких k-средних [37]. В качестве критерия определения качества аппроксимации использовалось среднеквадратическое отклонение (RMSE).

Была проведена аппроксимация с использованием линейных и экспоненциальных функций распределения нечётких термов. Следуя [37], число кластеров для алгоритма k-средних равнялось 7. Соответственно, полученная модель содержала 7 нечётких правил.

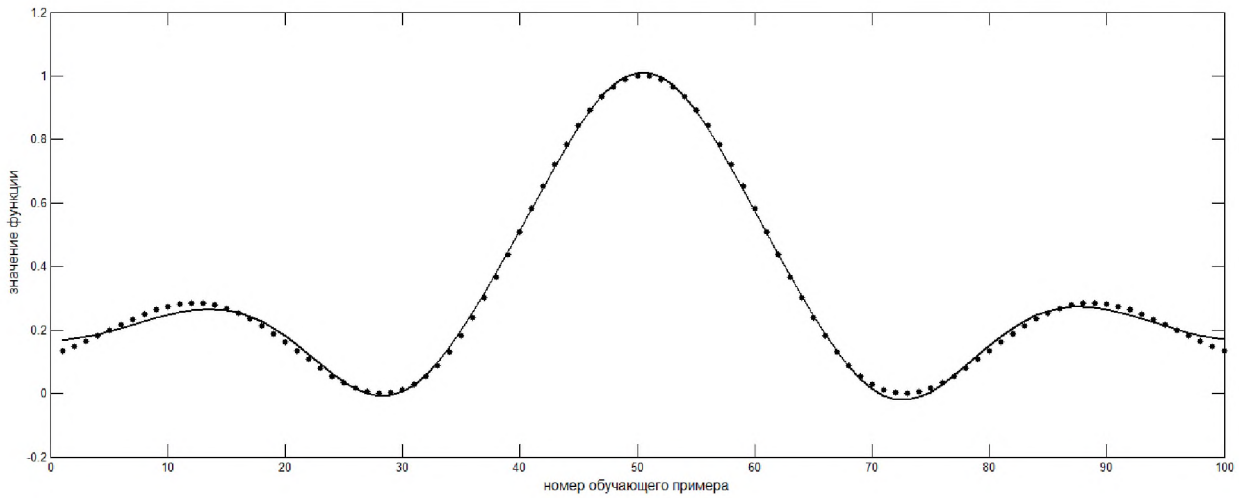


Рис. 4.2. Аппроксимация функции $z = \sin y/y$

На рисунке 4.2 представлен один из полученных результатов аппроксимации. Символом • показаны исходные обучающие данные, непрерывной линией – результат аппроксимации. На рисунке 4.3 приведён график среднеквадратического отклонения.

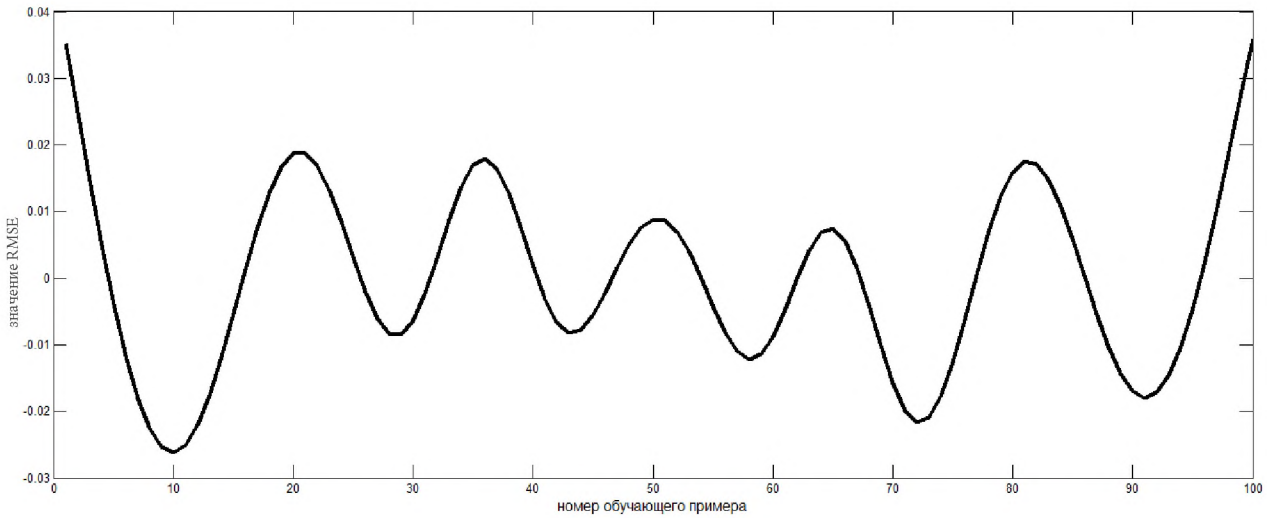


Рис. 4.3. Среднеквадратическое отклонение при аппроксимации функции $z = \sin y/y$

Эксперименты с работой μ/E алгоритма 1 показали, что качество аппроксимации существенно зависит от результата кластеризации, который, в свою очередь, зависит от изначального расположения центров.

Для получения более точных результатов моделирования было проведено 500 итераций и получены средние значения ошибок. В таблице 4.1 отражены результаты моделирования. Во 2 и 3 колонках представлены результаты работы алгоритма Cluster Estimation согласно [37] и в 4 и 5 – полученные автором.

Таблица 4.1

Сравнение среднеквадратичного отклонения для различных способов оптимизации и в зависимости от вида функции μ

Модель нечёткого вывода	Алгоритм Cluster Estimation		μ/E алгоритм 1	
	$\mu = e^{-t^2}$	нечёткое k-средних μ	$\mu = e^{-t^2}$	$\mu = 1 - t$
без оптимизации	0.1800	0.1190	0.2433	0.0543
Сугено 0	0.0540	0.0980	0.0202	0.0328
Сугено 1	0.0100	0.0150	0.0042	0.0224

Анализируя данные таблицы 4.1, можно сказать, что, действуя согласно μ/E алгоритму 1, наиболее хорошие результаты получаются при использовании модели нечёткого вывода Сугено 1 и экспоненциальной функции распределения, и в целом авторский алгоритм даёт меньшую ошибку в сравнении с алгоритмом Cluster Estimation.

4.1.5. Задача прогнозирования значений нестационарного временного ряда

Более сложной задачей является прогнозирование временного ряда, порождаемого дифференциальным уравнением Mackey-Glass [68].:

$$\dot{x}(t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t).$$

Задача прогнозирования состоит в том, чтобы, используя известные предыдущие значения x, за время t определить будущие значения x на момент времени $t + \triangle t$. Для наиболее точного сравнения результатов ставится задача нахождения значения $x(t+6)$ на основе $x = \{x(t-18), x(t-12), x(t-6), x(t)\}$. Множество данных содержит 1000 точек наблюдения, начиная с $t = 118$ и заканчивая $t = 1117$. Первые 500 точек используются для обучения и следующие 500 – для проверки способности модели к обобщению. В роли критерия качества прогнозирования применяется среднеквадратичное отклонение на контрольных данных, поделенное на стандартное отклонение временного ряда.

4.1.6. Усовершенствование алгоритма обучения машин нечёткого вывода

Эксперименты с μ/E алгоритмом 1 показали, что он является достаточно эффективным, но, тем не менее, иногда уступает алгоритму Cluster Estimation. Значения ошибки, полученные в работе [37], удавалось достигнуть только при наиболее удачных вариантах кластеризации, получаемых при помощи алгоритма k-средних.

Чтобы добиться улучшения результатов и уменьшить ошибку, модифицируем μ/E алгоритм 1 и поменяем метод кластеризации. Вместо алгоритма

k-средних будем использовать модифицированный алгоритм кластеризации Mountain method, реализованный командой *subclust* в среде Matlab. Особенностью данного алгоритма является то, что он вычисляет только центры кластеров. Для того, чтобы получить набор точек, принадлежащих кластерам, будем относить каждую из них к кластеру, расстояние до центра которого минимально. После этого, используя максиминную оценку параметров, получаем функции распределения для нечётких термов. Полученный таким образом алгоритм будем называть μ/E алгоритм 2.

Сначала оценим качество работы μ/E алгоритма 2 на первой задаче – аппроксимации нелинейной функции $z = \sin y/y$. В таблице 4.2 отражена зависимость количества правил и RMSE от значений радиуса r.

Таблица 4.2

Сравнение значений среднеквадратичного отклонения для различных моделей нечёткого вывода и в зависимости от значений радиуса

Радиус	Количество правил	RMSE при Сугено 0	RMSE при Сугено 1
0.178	15	0.0210	0.0072
0.20	10	0.0432	0.0065
0.25	9	0.0449	0.0037
0.3	7	0.0362	0.0067
0.35	4	0.0647	0.544

Наилучшие значения были достигнуты при r = 0.25, 9 кластерах и модели нечёткого вывода Сугено 1-го порядка, достигнутая ошибка составила 0.0037, что превосходит лучшие значения, полученные согласно μ/E алгоритму 1.

Вернёмся к задаче прогнозирования значений нестационарного временного ряда. В таблице 4.3 представлены результаты работы μ/E алгоритма 2. Использовалась модель Сугено 1-го порядка.

Таблица 4.3

Сравнение ошибок прогнозирования значений временного ряда в зависимости от радиуса

Радиус	Количество правил	Ошибка	
		обучающие данные	тестовые данные
0.18	53	0.0085	0.0102
0.24	47	0.0106	0.0123
0.25	30	0.0161	0.0152
0.27	27	0.0225	0.0225
0.3	26	0.0238	0.0233
0.33	21	0.0259	0.0255
0.35	20	0.0267	0.0263
0.4	15	0.0339	0.0335

Наилучшие результаты были достигнуты при r = 0.18 и 53 кластерах, ошибка составила: на обучающих данных = 0.0085, на тестовых = 0.0101. При увеличении радиуса ошибка возрастает (рис. 4.4), а дальнейшее уменьшение радиуса вызывало получение вырожденных матриц, используемых при построении эллипсоидов минимального объёма, описанных вокруг каждого кластера из множества входных данных для получения оценок параметров возможностных распределений.

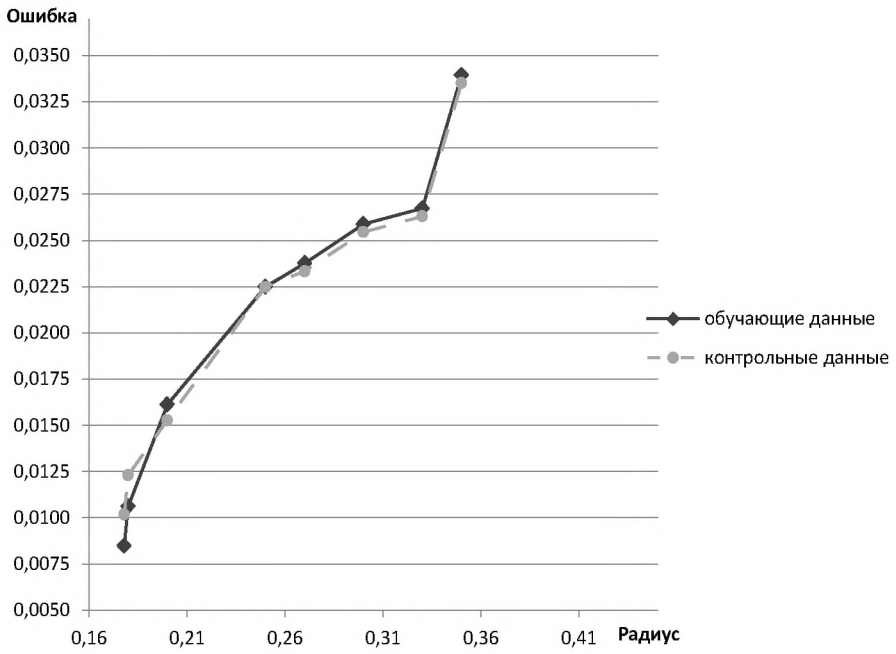


Рис. 4.4. Ошибка прогнозирования временного ряда, порождённого уравнением Mackey-Glass

На рисунке 4.5 представлен прогноз значений на контрольных данных.

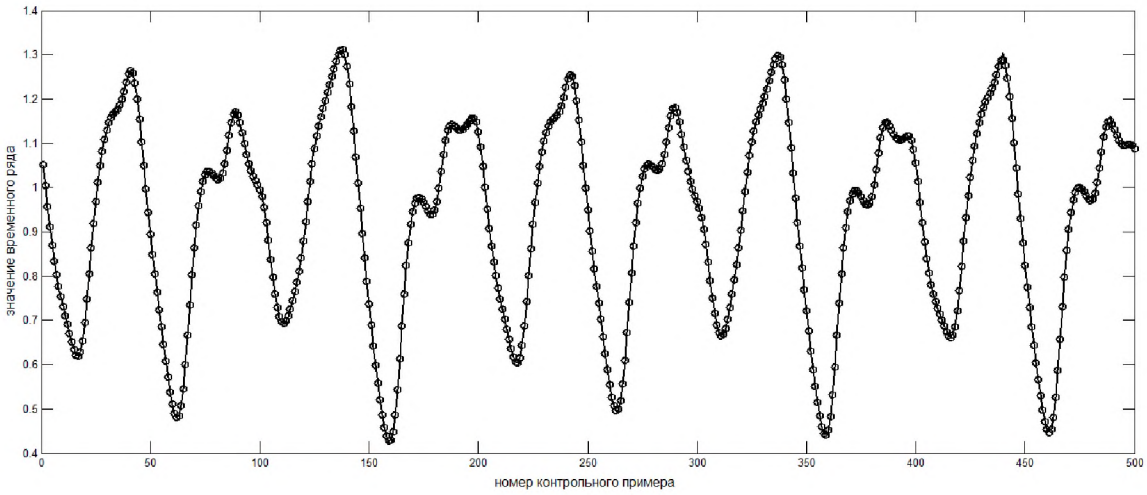


Рис. 4.5. Прогноз значений временного ряда, порождённого уравнением Mackey-Glass

Как можно видеть, модель практически соответствует временному ряду в масштабе изображения. На рисунке 4.6 можно видеть график изменения значений ошибок прогнозирования.

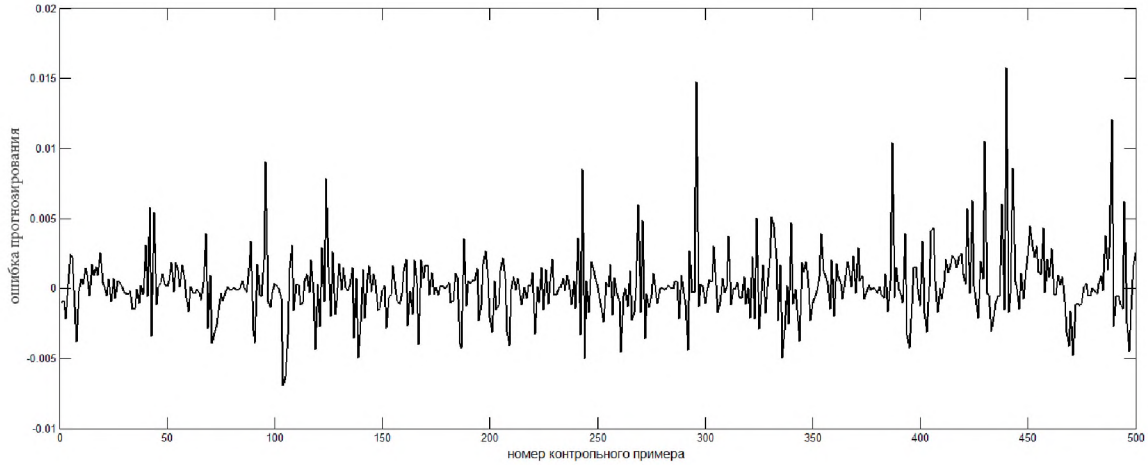


Рис. 4.6. Ошибка прогнозирования временного ряда, порождённого уравнением Mackey-Glass

Заметим, что во время обучения не наблюдалось явления переобучения – значения ошибок постепенно уменьшались с уменьшением радиуса.

В таблице 4.4 приведены значения ошибок, полученных при использовании различных методов на примере прогнозирования значений временного ряда. Лидирующее положение занимает алгоритм ANFIS [58]. Его работа заключается в последовательном применении двух шагов:

1. Нахождение оптимальных значений параметров для заключительных частей правил с помощью метода наименьших квадратов при фиксированных параметрах функции распределения нечётких термов.

2. Уточнение параметров функций распределения нечётких термов с помощью алгоритма обратного распространения ошибки при фиксированных заключительных частях правил.

И эти шаги повторяются до тех пор, пока повышается точность аппроксимации, обеспечиваемая нечёткой системой.

Отметим, что алгоритм ANFIS работает только с одномерными нечёткими термами, и, соответственно, его база правил порождается декартовым произведением терм-множеств для всех входных переменных. А разработанные нами алгоритмы и алгоритм Claster Estimation работают с многомерными термами.

Таблица 4.4

Сравнение значений ошибки моделирования
при использовании различных методов

Метод	Ошибка	Источник
ANFIS	0.007	J.-S. R. Jang, [58]
μ/E алгоритм 2	0.010	
Cluster Estimation	0.014	S. L. Chiu, [37]
μ/E алгоритм 1	0.016	
Back-Prop NN	0.02	R. S. Crowder, [38]
6-th order polynomial	0.04	R. S. Crowder, [38]
Cascaded-Correlation NN	0.06	R. S. Crowder, [38]
Auto-regressive Model	0.19	J.-S. R. Jang, [58]
Lineal Predictive Method	0.55	R. S. Crowder, [38]

Исходя из данных, представленных в таблице 4.4, видно, что при использовании μ/E алгоритма 2 был получен результат, уступающий лишь методу ANFIS .

4.2. Двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе

Известно, что обучение больших нейронных сетей с помощью классического алгоритма градиентного спуска наталкивается на значительные трудности. Это обусловлено проблемами затухания градиента по мере его распространения вглубь сети [77], [30] и «застреванием» в локальных минимумах[66].

В 2006 году Hinton и Salakhutdinov [79] предложили эффективный алгоритм послойного обучения нейронных сетей в режиме без учителя с помощью ограниченных машин Больцмана (Restricted Bolzman Machine, RBM). Опыт практического применения подобных сетей показывает, что с её помощью можно значительно сократить размерность входных данных. Например, в работе [66] показано, что такая сеть может сокращать размерность многомерных массивов данных вплоть до двух измерений, при этом сохраняя исходные взаимосвязи между элементами.

Ограниченная машина Больцмана – это ненаправленная вероятностная графическая модель, то есть случайное Марковское поле. В RBM чаще всего используются элементы с распределением Бернулли [24], в некоторых случаях – с экспоненциальным [85]. Структура RBM представляет собой полностью связанный двудольный граф (рис. 4.7), в котором одна группа узлов (видимые узлы v) моделирует данные, другая группа (скрытые узлы h) – скрытые переменные.

Так как RBM является особым случаем Марковского случайного поля, то совместное распределение по всем узлам является распределением Больцмана, которое описывается функцией энергии $E(v, h)$. Чаще всего в качестве такой функции используется линейная функция состояний видимых и скрытых узлов [53]:

$$E(v,h) = -\sum_{i,j} W_{ij} v_i h_j - \sum_i b_i v_i - \sum_j c_j h_j, \tag{4.1}$$

где $W_{i,j}$ – вес ненаправленных связей между

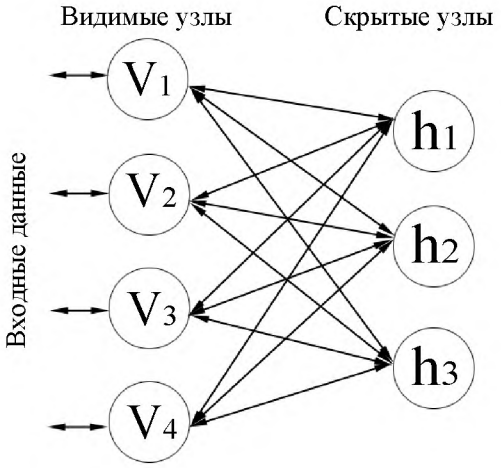


Рис. 4.7. Ограниченная машина Больцмана

узлами v_i и h_j, b_i – смещение узла v_i, c_j – смещение узла h_j.

Следует заметить, что состояния видимых узлов являются независимыми при заданных состояниях скрытых узлов, и наоборот. Тогда, при условии использования линейной функции энергии, условные вероятности $P(v_i = 1|h)$ и $P(h_j = 1|v)$ можно определить следующим образом [79]:

$$P(v_i = 1|h) = \frac{1}{1 + e^{\sum\limits_j W_{ij} h_j - b_i}},$$

$$P(h_j = 1|v) = \frac{1}{1 + e^{\sum\limits_i W_{ij} v_i - c_j}}.$$

Задача обучения RBM состоит в определении таких весов W и смещений b и c, что маргинальное распределение наблюдаемых узлов в модели (P_{model}) будет близко к наблюдаемому распределению (P_{data}). Для этого минимизируется расстояние Кульбака-Лейблера, градиент которого по весам W_{ij} определяется выражением [79]:

$$\frac{\partial KL(P_{data}||P_{model})}{\partial W_{ij}} = E[v_i h_j]_{P_{data}} - E[v_i h_j]_{P_{model}},$$

где $E[\cdot]_{P_{model}}$ – ожидаемое значение для модельного распределения, $E[\cdot]_{P_{data}}$ – ожидаемое значение для распределения данных.

Сложность обучения RBM заключается в том, что значение $E[v_i h_j]_{P_{model}}$ не может быть получено аналитически, а его численное вычисление занимает много времени. В работе [79] был предложен алгоритм обучения RBM, в котором используется измененная целевая функция. Ее градиент задается выражением

$$E[v_i h_j]_{P_{data}} - E[v_i h_j]_{P_1},$$

где $E[\cdot]_{P_1}$ – ожидаемое значение после одного шага работы модели.

После обучения RBM может функционировать в режиме ассоциативной памяти. При подаче на видимые узлы значений, описывающих текущую ситуацию, на скрытых узлах будет получено её описание в терминах скрытых переменных, выделенных сетью во время обучения. Данное описание устойчиво к шумам и отсутствию части входных данных: для заданного входа сеть находит ближайший пример из обучающего набора. При этом уровень энергии сети (4.1) может быть использован для определения степени новизны

ситуации: энергия будет низкой для состояний, близких к использованным при обучении, и высоким для принципиально новых.

Характерной особенностью обучения RBM является то, что оно проходит в режиме, известном как «обучение без учителя»: для обучения сети предъявляются только образцы данных для видимых узлов (соответствующие значениям признаков G_i, описывающих ситуацию), а значения скрытых переменных сеть находит самостоятельно. Таким образом, в ходе обучения сеть автоматически выделяет некоторые значимые характеристики, которые могут быть использованы для описания данных.

Это свойство позволяет использовать стек RBM для обучения глубоких нейронных сетей: построенное одной машиной описание может быть использовано следующей для выявления более высокоуровневых скрытых переменных и построения описания в их терминах и так далее. Как показывает практика, сформированные таким образом высокоуровневые описания можно эффективно использовать для решения задач классификации, что говорит о том, что они действительно отражают существенные высокоуровневые понятия.

Создание глубокой нейронной сети осуществляется следующим образом:

1. Обучается RBM, соответствующая первому слою нейронной сети.

2. Для каждой точки данных вычисляются наиболее вероятные значения скрытых элементов h_j.

3. Полученные на предыдущем шаге значения используются в качестве входных для обучения RBM, соответствующей второму слою.

4. Процесс продолжается до тех пор, пока не будут обучены RBM для всех слоев нейронной сети.

Стек обученных RBM используется для построения глубокой нейронной сети. При этом отбрасываются смещения видимых элементов b_i, а матрицы ненаправленных связей W_{ij} используются как матрицы связей между нейронами соответствующих слоев.

На рисунке 4.8 схематично представлена архитектура нейронной сети, которая позволит сократить размерность [83], [70].

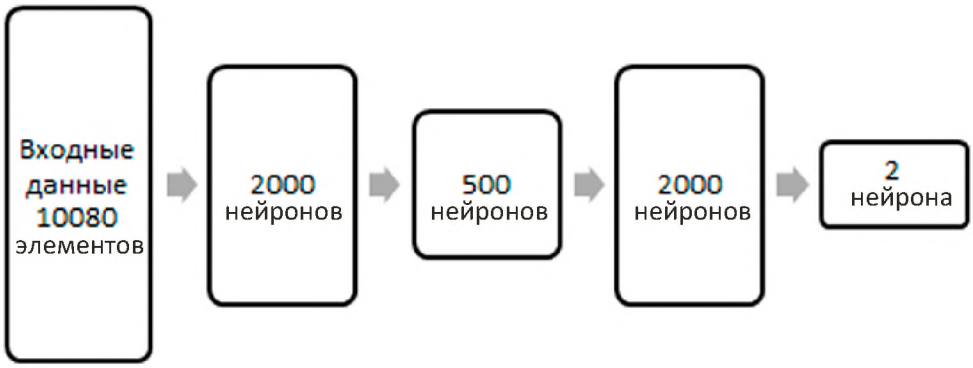


Рис. 4.8. Архитектура нейронной сети

Первый слой нейронной сети состоит из 2000 нейронов. Ему на вход подаётся вектор, состоящий из 10080 элементов. За первым слоем находятся скрытые слои размером 500 и 2000 нейронов соответственно. В выходном слое сети число нейронов равно требуемой выходной размерности, например, 2 или 3.

Для выделения областей, соответствующих различным ситуациям в пространстве выходов нейронной сети, может быть использована машина нечеткого вывода (или нечеткий регулятор). Для её обучения в части 5.2.1.3. применяется μ/E алгоритм 2, предложенный в параграфе 4.1.6. и статье [20].

4.3. Слияние алгоритма обучения машин нечёткого вывода с алгоритмом AdaBoost

В настоящий момент конкуренцию нейронным сетям составляет такой метод машинного обучения как бустинг.

Бустинг (англ. boosting — улучшение) — это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов. [5]

Бустинг представляет собой жадный алгоритм построения композиции алгоритмов.

Основные причины его популярности состоят в простоте, универсальности, гибкости (возможность построения различных модификаций) и высокой обобщающей способности.

Суть алгоритмов бустинга сводится к следующим шагам. В начале каждой точке из совокупности обучающих данных присваивается вес (одинако-

вый для всех точек). Далее применяется алгоритм обучения и в зависимости от вычисленных результатов обновляются веса точек обучающей выборки, причём больший вес присваивается тем наблюдениям, которые сложнее всего классифицировать. Таким образом, классификатору сообщаются "проблемные"точки, на которые необходимо обратить внимание.

Далее алгоритм работает уже со взвешенными данными и на основе результатов происходит обновление весов обучающих данных. И так далее, пока не будет достигнуто заданное количество итераций.

Одной из разновидностей бустинга является алгоритм AdaBoost (сокр. от adaptive boosting) — алгоритм машинного обучения, предложенный Йоавом Фройндом (Yoav Freund) и Робертом Шапиром (Robert Schapire) в 1996 году [44]. Является мета-алгоритмом, в процессе обучения строит композицию из базовых алгоритмов обучения для улучшения их эффективности. AdaBoost является алгоритмом адаптивного бустинга в том смысле, что каждый следующий классификатор строится по объектам, которые плохо классифицируются предыдущими классификаторами [6].

Рассмотрим задачу классификации на два класса, $Y = \{-1, +1\}$. Допустим, что базовые алгоритмы $b_1, \ldots, b_T$ также возвращают только два ответа -1 и $+1$. $W^l = (w_1, \ldots w_l)$ — вектор весов объектов.

$Q(b, W^l) = \sum_{i=1}^{l} w_i[y_i b(x_i) < 0]$ — стандартный функционал качества алгоритма классификации b.

Задачу оптимизации параметра α_t можно решать аналитически, аппроксимируя пороговую функцию потерь $[z < 0]$ с помощью экспоненты $E(z) = \exp^{-z}$.

Алогоритм AdaBoost для классификации на 2 класса будет выглядеть следующим образом [6]:
дано: X^l- обучающая выборка;

$b_1, \ldots, b_T$ - базовые алгоритмы классификации;

1. Инициализация весов объектов: $w_i = 1/l, i = 1, \ldots, l$;
2. Для всех $t = 1, \ldots, T$, пока не выполнен критерий останова.
2.1 Находим классификатор $b_t : X \to \{-1, +1\}$ который минимизирует взвешенную ошибку классификации;

$$b_t = \arg\min_b Q(b, W^l);$$

2.2 Пересчитываем кооэффициент взвешенного голосования для алгоритма классификации b_t:

$$\alpha_t = \frac{1}{2} \ln \frac{1 - Q(b, W^l)}{Q(b, W^l)};$$

2.3 Пересчет весов объектов:$w_i = w_i \exp\left(-\alpha_t y_i b_t(x_i)\right), i = 1, \ldots, l;$

2.4 Нормировка весов объектов: $w_0 = \sum_{j=1}^{l} w_j; w_i = w_i / w_0, i = 1, \ldots, l;$

4. Возвращаем:$a(x) = sign\left(\sum_{i=1}^{T} \alpha_i b_i(x)\right).$

Данный алгоритм также имеет расширения для классификации на несколько классов [44]. В главе 5.2.3.1. описывается применение алгоритма AdaBoost к типовой задаче классификации. В качестве базовых алгоритмов классификации выступают алгоритмы, основанные на методе нахождения μ/E-оценки (4.1.6.).

Помимо задачи классификации алгоритм AdaBoost с успехом применяется в задачах регрессии [44]. Данная технология была опробована авторами в задаче прогнозирования свойств проводников и описана в параграфе 5.2.3.2.. Но перед тем, как использовать метод на сложных данных, рассмотрим его применение к задаче аппроксимации нелинейной функции одной переменной $z = sin(y)/y$.

Для обучения сгенерировано 130 точек выборки, наложен белый шум, проверочная выборка содержит 480 значений функциональной зависимости без шума.

Исходные данные представлены на рис. 4.9.

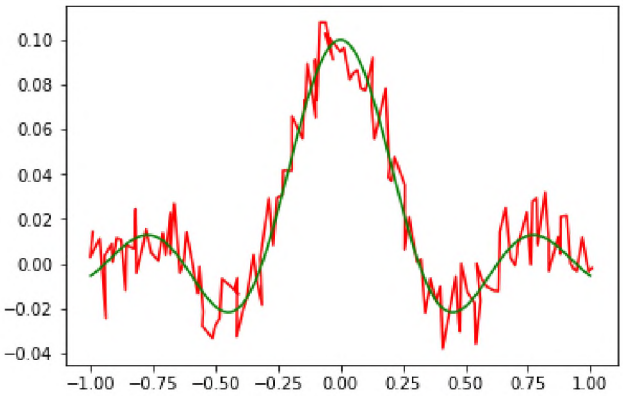


Рис. 4.9. Функция $z = \sin y/y$ с наложенным на неё шумом

Работа алгоритма была проверена при различных значениях радиусов, использованных при кластеризации. В качестве критерия качества был взят коэффициент детерминации R^2, то есть наилучшее значение, которые может быть получено – 1.0, а в случае, если модель работает плохо, то коэффициент может быть отрицательным.

Результаты работы отражены на рис. 4.10. Верхняя линия представляет коэффициент детерминации на тестовых данных, нижняя – на обучающих.

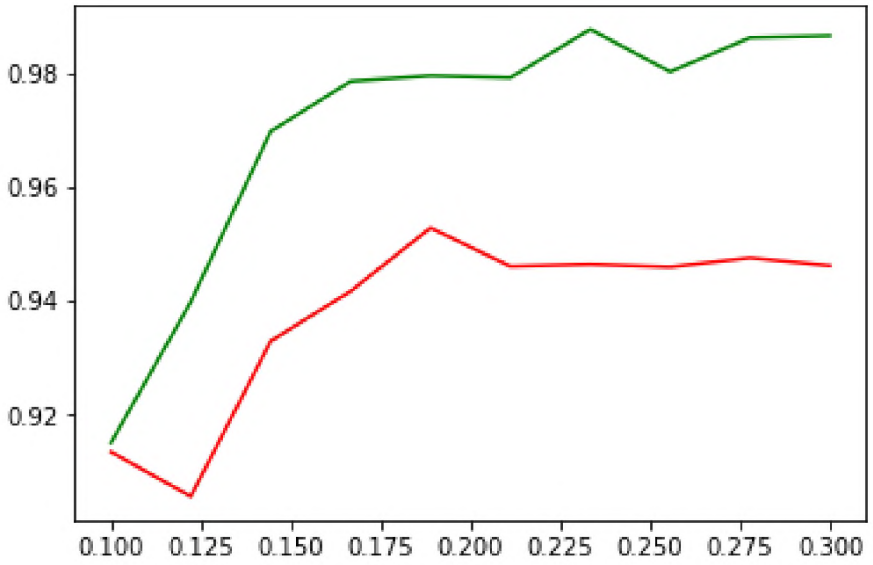


Рис. 4.10. Прогнозирование значений функции $z = \sin y/y$ при различных радиусах

Проанализировав данные рисунка 4.10 и таблицы 4.5 можно видеть, что ошибка на тестовых данных ниже, чем на обучающих и максимальное значение R^2 достигается при радиусах, больших, чем 0.23.

Отметим, что показатели на тестовых данных лучше, чем на обучающих.

Такой эффект является результатом успешного обобщения, в процессе которого алгоритмом был скомпенсирован шум в обучающих данных.

Также заметим, что при некоторых значениях радиуса были достигнуты значения коэффициента детерминации, равные 0.99, что является почти максимально возможным.

Все эти факты говорят о высоком качестве работы представленного алгоритма.

Таблица 4.5

Значения коэффициента детерминации при различных значениях радиуса

Радиус	R^2 на обучающей выборке	R^2 на тестовой выборке
0.1	0.91	0.91
0.12	0.91	0.94
0.14	0.93	0.97
0.17	0.94	0.98
0.19	0.95	0.98
0.21	0.95	0.98
0.23	0.95	0.99
0.26	0.95	0.98
0.28	0.95	0.99
0.3	0.95	0.99

4.4. Выводы по четвёртой главе

В четвёртой главе в параграфе 4.1.1 дано понятие нечёткого правила, в параграфе 4.1.2 произведено описание системы нечёткого вывода. Перечислены составные элементы таких систем.

Параграф 4.1.3 посвящён разработке алгоритма обучения машин нечёткого вывода, основанного на идее алгоритма Cluster Estimation, который был предложен Stepohen L. Chiu.

Для проверки работоспособности разработанного алгоритма в параграфе 4.1.4 рассмотрена задача аппроксимации нелинейной функции одной переменной $z = sin(y)/y$. После чего произведено сравнение результатов с результатами алгоритма Cluster Estimation. Показано, что действуя согласно разработанному алгоритму, наиболее качественные результаты получаются при использовании модели нечёткого вывода Сугено 1 и экспоненциальной функции распределения, и в целом авторский алгоритм даёт меньшую ошибку в сравнении с алгоритмом Cluster Estimation.

В параграфе 4.1.5 рассматривается задача прогнозирования значений нестационарного временного ряда, порождаемого дифференциальным уравнением Mackey-Glass:

$$\dot{x}(t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t).$$

Результаты работы над этой задачей с использованием разработанного алго-

ритма показали, что он является достаточно эффективным, но, тем не менее, иногда уступает алгоритму Cluster Estimation.

Для улучшения результатов в параграфе 4.1.6 была произведена модификация алгоритма путём замены метода кластеризации на Mountain method.

По результатам сравнения работы нового алгоритма с предыдущим разработанным алгоритмом на первой задаче – аппроксимации нелинейной функции $z = \sin y/y$ выяснилось, что модифицированный алгоритм работает значительно лучше.

После применения нового алгоритма к задаче прогнозирования значений нестационарного временного ряда был получен результат, уступающий лишь методу ANFIS.

В параграфе 4.2 описывается разработанная двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе. Описывается процесс создания глубокой нейронной сети. В рассматриваемой системе роль глубокой нейронной сети состоит в сокращении размерности исходных данных. Применение ограниченных машин Больцмана для построения нейронной сети позволяет выявлять существенные закономерности в данных и переходить от пространства входов размерностью в несколько сотен или тысяч измерений к пространству скрытых переменных небольшой размерности, например, 2 или 3. Далее для выделения областей, соответствующих различным классам (при решении задач классификации), в пространстве выходов нейронной сети используется машина нечеткого вывода. Для её обучения применяется разработанный алгоритм обучения машин нечёткого вывода.

В параграфе 4.3 рассматривается процедура бустинга и алгоритм Adaboost. Была изучена возможность использования алгоритма бустинга Adaboost для построения композиций систем нечёткого вывода, обучающихся разработанным алгоритмом. Работоспособность метода продемонстрирована на примере прогнозирования значений функции $z = \sin y/y$. Были достигнуты максимально возможные значения коэффициента детерминации, равные 0.99, что является показателем высокого качества работы представленного алгоритма.

5. Программный комплекс анализа данных на основе возможностного моделирования

5.1. Реализация алгоритмов обучения нечёткого контроллера в среде Matlab

Для проверки разработанных алгоритмов была реализована программа в среде Matlab (рис. 5.1).

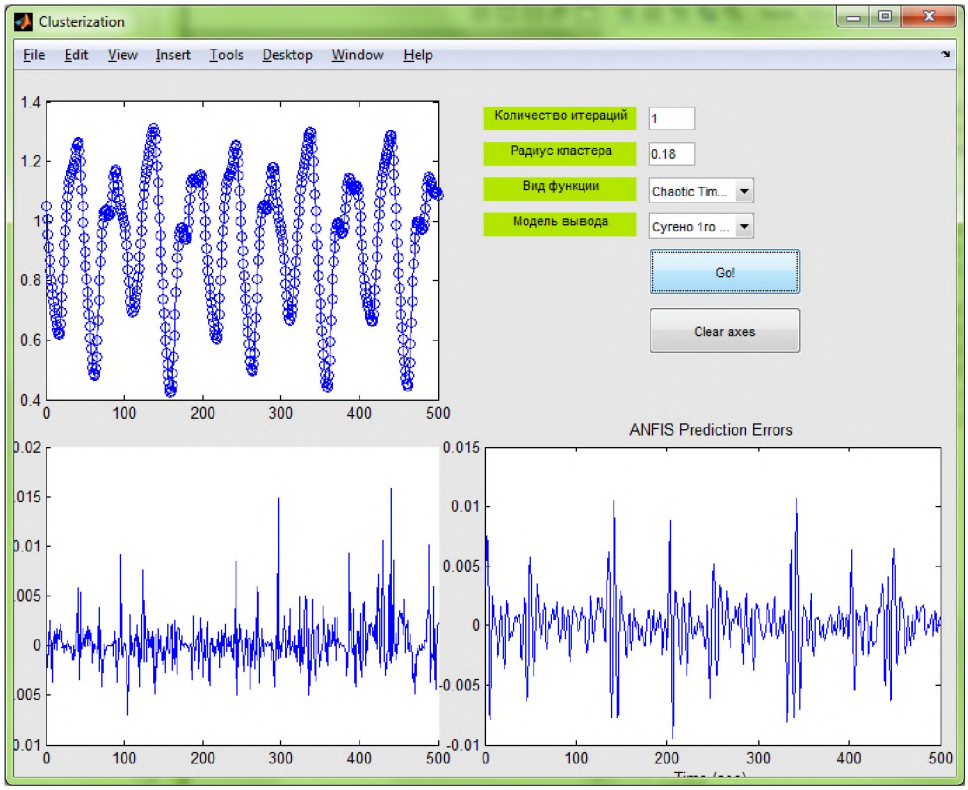


Рис. 5.1. Интерфейс программы

Matlab был выбран в силу того, что он является интуитивно понятным языком программирования и в нём имеется большое число функций, реализующих известные алгоритмы: метод наименьших квадратов, поиск эллипсоида минимального объёма, кластеризация различными методами, ANFIS и другие. Использование матричной формы записи позволяет в компактной форме записать алгоритмы нечёткого вывода и обучения.

Разработанная программа поддерживает модели нечёткого вывода Сугено 0-го и 1-го порядка. Пользователь может выбрать вид аппроксимируемой функции и параметр алгоритма. Возможно проведение серии испытаний с

вычислением статистики результатов. Строятся графики исходной функции и её аппроксимации, ошибки аппроксимации. Для сравнения строится график ошибки аппроксимации по методу ANFIS.

5.2. Реализация алгоритмов обучения нечёткого контроллера на языке программирования Python

5.2.1. Прогнозирование неисправностей вагонного электрооборудования на железнодорожном транспорте

В данном параграфе будут рассмотрен вариант применения разработанного математического аппарата для решения проблем интеллектуального анализа данных.

Одной из таких задач является прогнозирование нештатных ситуаций в электрооборудовании вагонов на железнодорожном транспорте.

5.2.1.1. Постановка задачи

Среди задач “Стратегии развития железнодорожного машиностроения до 2030 года” [18] можно выделить проблему повышения безопасности железнодорожных перевозок. Для этого необходимо определить и ликвидировать причины возникновения аварийных ситуаций, в состав которых, в том числе, входит неисправность вагонного оборудования. Таким образом, ставится задача заблаговременного выявления возможных нештатных ситуаций в подвижном составе крупнейшего оператора сети российских железных дорог ОАО «РЖД» для вагонов, выпускаемых ОАО “ТВЗ".

Большинство существующих систем и исследований направлено на анализ состояния железнодорожных вагонов, так как они оказывают первостепенное влияние на безопасность железнодорожных перевозок. В этих системах ведётся обработка сравнительно небольшого числа аналоговых сигналов высокой частоты, собранных с помощью дополнительных датчиков. Анализ данных осуществляется с использованием преобразования Фурье или вейвлет-преобразований для начальной обработки данных.

Немаловажным является тот факт, что характер данных, доступных в современных системах управления электрооборудованием пассажирских вагонов, существенно отличается от используемых в системах обработки информации о состоянии рельсов, колёсных пар и подвески – число сигналов

достигает нескольких сотен, большинство из них является бинарными и редко изменяется. Следовательно, применение методов спектрального анализа невозможно и прямой перенос имеющегося опыта на другие системы вагонов заструднителен.

Таким образом, необходима разработка новых методов поиска и выявления признаков нештатных режимов работы оборудования. Верхние слои систем анализа данных хорошо зарекомендовали себя в раличных отраслях, в том числе железнодорожной, и их можно использовать при разработке систем прогнозирования неисправностей.

В современных вагонах, выпускаемых ОАО «ТВЗ», используется модульная информационная система контроля, диагностики и управления (СКДУ). Она представляет собой сеть взаимодействующих устройств, которые управляют различными системами вагона. Одним из таких устройств, входящих в состав СКДУ, является контроллер управления электрооборудованием вагона (КУЭВ).

В функции КУЭВ входит приём аналогово-цифровых и дискретных сигналов обратной связи с вагонного оборудования, присвоение этим сигналам логических имён и фиксация соответствующих значений. КУЭВ хранит данные об актуальных значениях своих переменных, которые периодически фиксируются в базе данных срезов значений параметров КУЭВ. Схематично это можно представить следующим образом (рис. 5.2):

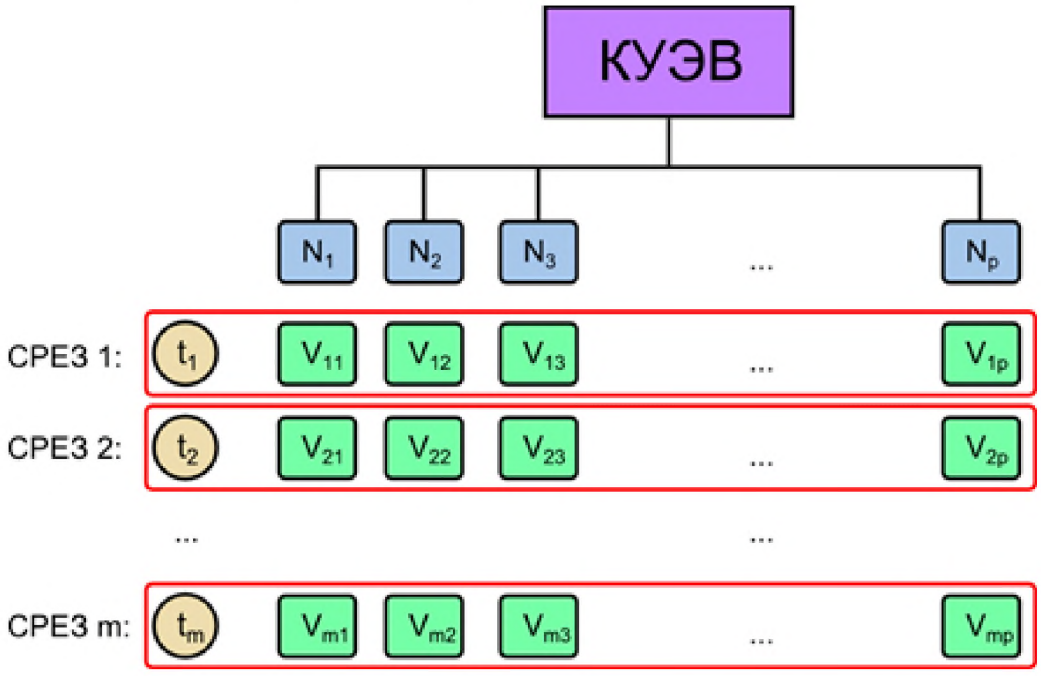


Рис. 5.2. База данных срезов значений параметров КУЭВ

В момент времени t_i переменная с именем N_j имеет значение V_{ij}, $i \in 1,..,m$, $j \in 1,..,p$.

Срезом системы в момент времени t_i называют набор данных $[t_i, V_{i1}, .., V_{ip}]$. Срезы добавляются в базу данных каждые 5 секунд.

Количество параметров, фиксируемых КУЭВ, варьируется от 147 до 237 в зависимости от модификации вагона. Большинство из них является бинарными.

Значения параметров могут храниться в структурах одного из двух типов:

1. $\{V, S\}$, где $V \in \{ON, OFF\}$, $S \in \{NORMAL, KZ, NA, FAULT\}$.
2. $\{N, S\}$, где $N \in \mathbb{R}$, $S \in \{NORMAL, KZ, NA, FAULT\}$.

Для анализа будут использоваться параметры 1-го типа, которые могут принимать бинарные значения.

При возникновении неисправности в вагоне поезда её регистрирует СКДУ. В описание неисправности входит код неисправности, приоритет, системная дата и время возникновения, тип устройства. Приоритет отражает степень важности неисправности, которая может быть 4 типов (по возрастанию важности):

- неисправности с приоритетом 0 не влияют на ход эксплуатации оборудования; пример такого события – “включен приточный вентилятор”;
- приоритет 1 имеют неисправности, которые могут повлиять на ход эксплуатации оборудования; пример – “питание кондиционера производится от АКБ”;
- приоритет 2 имеют неисправности, которые могут повлечь за собой невыполнение основных функций вагонного оборудования или потенциально привести к нештатным ситуациям на вагоне; например “Бак накопитель заполнен на 95%”;
- неисправности с приоритетом 3 представляют непосредственную угрозу жизням и здоровью пассажиров; например – “Пожар в купе №2”.

Под неисправностью будем понимать зафиксированное в определённый момент времени событие.

5.2.1.2. Проведение первичного анализа данных

Первичный анализ данных включает следующие стадии:

1. подготовка данных;
2. выбор значимых параметров КУЭВ;
3. формирование обучающей и контрольной выборок.

На этапе подготовки производится предварительная обработка данных, собранных СКДУ вагона. Осуществляются операции фильтрации ошибочных данных и объединения журналов.

При выборе значимых параметров КУЭВ из дальнейшей обработки исключим параметры, не несущие существенной информации для прогнозирования возникновения неисправности. Для этого будем использовать алгоритмы построения деревьев решения и случайных лесов. Сначала построим деревья регрессии [34] для каждого набора данных и все параметры, задействованные в узлах данных деревьев, будем считать значимыми и оставим для дальнейшего анализа. Затем, с помощью алгоритма построения случайных лесов [35], оценим такие показатели, как среднее снижение точности и среднее снижение коэффициента Gini. Выберем те параметры, у которых хотя бы один из показателей больше 0.

На этапе формирования выборок данные преобразуются к виду, пригодному для подачи на входы нейронной сети. Для этого из параметров КУЭВ формируются два вида срезов блоков данных информационной системы:

1. Приводящие через заданное пользователем время к возникновению неисправности;
2. Не приводящие через заданное пользователем время к возникновению неисправности.

При их формировании сначала выделяются блоки, приводящие к возникновению неисправности. Затем случайным образом выбирается такое же колчество блоков, не приводящих к неисправности. Параметры, входящие в каждый блок, разворачиваются из последовательности срезов в вектор.

5.2.1.3. Апробация алгоритма

В качестве целевой выберем неисправность устройства HEATEQ с кодом 17721 (далее – HEATEQ__17721). Эта неисправность свидетельствует о том, что был выключен приточный вентилятор и управление электрокалорифером в автоматическом режиме становится невозможным. Она была выбрана вследствие большой частоты возникновения и, соответственно, большого объёма данных для анализа. Будем считать, что КУЭВ содержит 147 параметров, а количество различных неисправностей в системе равно 570.

База данных срезов значений параметров КУЭВ содержит 241 920 срезов, полученных путём снятия данных с периодичностью в 5 секунд в течение 14 дней, включая случаи временной недоступности системы вследствие отсутствия электроснабжения вагона.

Прогнозным горизонтом, т.е. крайним сроком, для которого прогноз действителен с заданной точностью, будем считать промежуток времени в 15 минут, а в качестве периода основания прогноза, т.е. промежутка времени, в рамках которого выбирается информация для разработки прогноза, возьмём 30 минут. Таким образом, прогноз строится на основании срезов, снятых в течение 30 минут (360 срезов), и взятых за 15 минут до возникновения рассматриваемой неисправности. После применения алгоритмов построения деревьев решения и случайных лесов было отобрано 28 значимых параметров КУЭВ.

Блоки данных случайным образом были разделены на обучающие и контрольный наборы: обучающий набор содержал 24 блока, контрольный – 16 блоков. Далее они были использованы для обучения глубокой нейронной сети, с архитектурой представленной на рисунке 4.8. Затем блоки проецировались в двумерное пространство.

Так как алгоритм обучения является стохастическим, при каждом его запуске получаются различные сети и соответствующие им проекции. На основании проведённых испытаний путём многократной визуализации данных можно утверждать, что получаемые проекции не имеют существенных отличий друг от друга. Пример проекции представлен на рисунке 5.3, где $\circ$ – точки, соответствующие предаварийным ситуациям, $\times$ – нормальным.

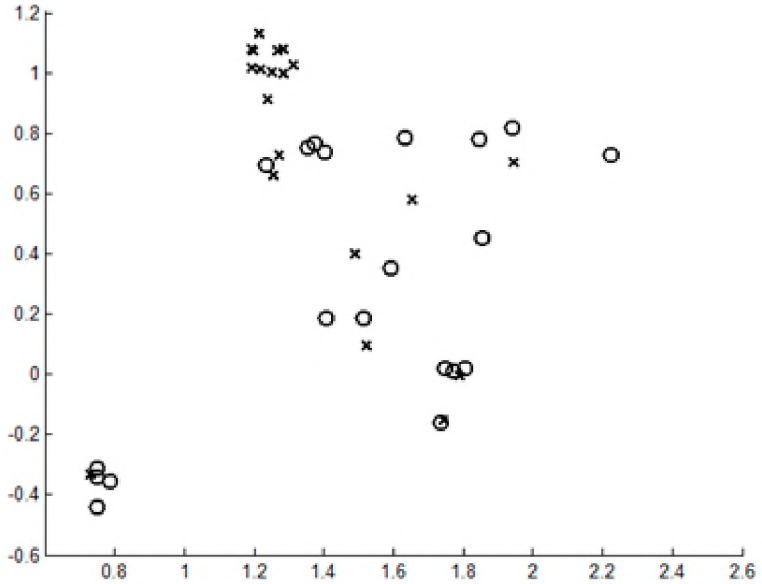


Рис. 5.3. Пример проекции исходных данных в двумерное пространство выходов нейронной сети

Для того, чтобы проверить работу μ/E алгоритма 2, сравним результаты его работы с двумя зарекомендовавшими себя алгоритмами обучения машин нечёткого вывода: ANFIS [58] и Claster Estimation [37]. Все перечисленные алгоритмы работают с моделями нечёткого вывода Сугено.

Алгоритмы Claster Estimation и ANFIS были подробно рассмотрены в параграфах 4.1.3. и 4.1.6..

На рисунке 5.4 представлен пример работы системы нечёткого вывода, созданной на основе алгоритма μ/E-оценки. Выходное значение больше 0.5 будем рассматривать как прогноз аварийной ситуации, меньше 0.5 – нормальной. В обучающей выборке эти ситуации представлялись значениями 1 и 0 соответственно. Прогнозные значения обозначены с помощью $*$, $\circ$ – значение из контрольной выборки.

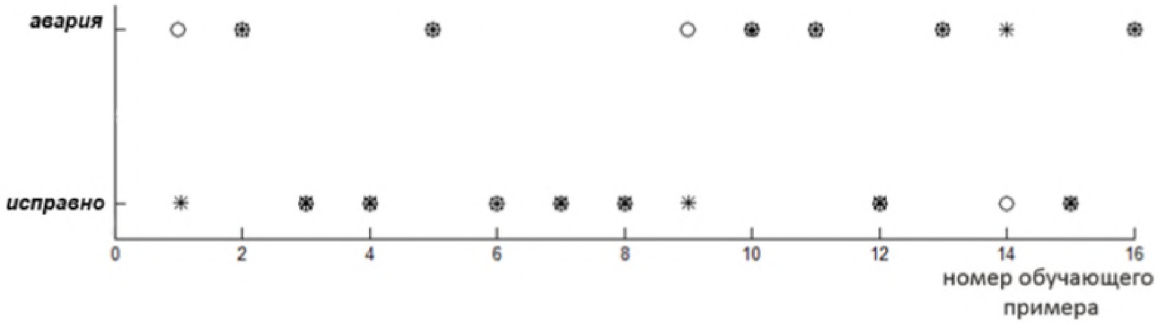


Рис. 5.4. Пример работы спроектированной системы нечёткого вывода

После проведения 1000 экспериментов были получены следующие ре-

зультаты:

Таблица 5.1

Результаты тестирования алгоритмов обучения машин нечёткого вывода

Алгоритм	Среднее число ошибочных классификаций	
	Без нормализации данных	С нормализацией данных
ANFIS	5.96 (37%)	5.40 (34%)
Claster Estimation	4.23 (26%)	3.48 (22%)
μ/E алгоритм 2	4.03 (25%)	4.16 (26%)

В качестве критерия оценивания работы алгоритмов использовалось среднее число ошибок при классификации точек. Обучение алгоритма ANFIS велось с параметрами по умолчанию, для двух других алгоритмов проводился поиск оптимального значения радиуса кластеров, обеспечивающего наименьшую ошибку на контрольном множестве.

Проанализировав результаты, представленные в таблице 5.1, можно сделать вывод о том, что алгоритм ANFIS хуже всех справился с поставленной задачей. Возможно, причиной является структура данных, которая лучше описывается с помощью многомерных термов, чем с помощью декартова произведения одномерных термов.

Также отметим, что после нормализации данных алгоритмы ANFIS и Claster Estimation улучшили свои результаты, тогда как алгоритм, основанный на μ/E-оценке, стал работать хуже. Авторы считают, что это может быть вызвано тем, что в случае μ/E-оценки могут использоваться не симметричные относительно осей координат распределения, в то время в алгоритме Claster Estimation распределения ограничены симметричными относительно нуля функциями. А после проведения нормализации данных распределение принимает более равномерный характер и авторский алгоритм работает хуже.

Из отрицательных сторон алгоритма μ/E-оценки отметим вычислительную неустройчивость для некоторых наборов данных: в связи с небольшим числом точек обучающего множества в ряде случаев размер получаемых кластеров слишком мал для того, чтобы провести оценку параметров соответствующей функции распределения. Данный недостаток можно исправить, например, путём увеличения размера обучающей выборки.

Анализ работы алгоритмов показал, что в случае ненормализованных данных использование метода μ/E-оценки параметров функций распределения нечётких термов позволяет более точно описать структуру данных, чем с помощью задаваемого пользователем параметра – радиуса R.

В целом можно сказать, что разработанный алгоритм μ/E-оценки достойно справился с поставленной задачей и в случае ненормализованных данных работает даже лучше, чем Claster Estimation.

Полученная архитектура может являться основой системы прогнозирования возникновения неисправностей в электрооборудовании вагона. Как показали предварительные расчёты, проведенные специалистами ОАО «ТВЗ», ожидаемое сокращение затрат на техническое обслуживание вагонов от внедрения подобной системы может достигать 25 тысяч рублей в год на каждый вагон.

5.2.2. Идентификация состояния воздушного пространства

Задача оценки состояния воздушного пространства может быть рассмотрена в контексте решения проблем классификации, распознавания образов и диагностики [16], [4]. Ее решение требует привлечения технических средств «добывания» информации (радиолокационные станции загоризонтного обнаружения воздушных целей (РЛС ЗГО) и др.), а также экспертной информации, которая, в основном, носит качественный характер.

В такой ситуации адекватным средством моделирования информации является современная теория возможностей, а решающие правила, позволяющие классифицировать оперативную ситуацию, могут быть построены на основе технологий мягких вычислений [71], [3], [17].

Данная задача решалась Арефьевым В.И., Петровым М.О., Талалаевым А. Б., Язениным А.В. и Сорокиным С.В. и представлена в статье [1]. В процессе решения был применён разработанный в диссертации математический аппарат.

Возможность применения нечетких классификаторов для решения этой задачи показана на примере успешной реализации нечёткого классификатора оперативных ситуаций.

5.2.3. Применение алгоритма Adaboost совместно с μ/E алгоритмом обучения машин нечёткого вывода к задаче классификации и регрессии

5.2.3.1. Апробация на типовой задаче классификации

В качестве задачи классификации для проверки работы алгоритма бустинга Adaboost с базовыми классификаторами, основанными на μ/E алгоритме обучения нечёткого контроллера, была выбрана задача определение вида ириса из набора данных ирисы Фишера.

Ирисы Фишера — это набор данных для задачи классификации, на примере которого Рональд Фишер в 1936 году продемонстрировал работу разработанного им метода дискриминантного анализа. Этот набор данных является классическим и часто используется для иллюстрации работы различных статистических алгоритмов [57].

Ирисы Фишера состоят из данных о 150 экземплярах ириса, по 50 экземпляров из трёх видов — Ирис щетинистый (Iris setosa), Ирис виргинский (Iris virginica) и Ирис разноцветный (Iris versicolor). Для каждого экземпляра существует четыре характеристики (в сантиметрах):

- длина наружной доли околоцветника;
- ширина наружной доли околоцветника;
- длина внутренней доли околоцветника;
- ширина внутренней доли околоцветника.

На основании этого набора данных требуется построить правило классификации, определяющее вид растения по данным измерений. Это задача многоклассовой классификации, так как имеется три класса — три вида ириса.

Для решения этой задачи было произведено обучение на 50 примерах (33% выборки), выборанных случайным образом с помощью функции model_selection.train_test_split библиотеки sklearn.

Для классификации с помощью бустинга была использована функция ensemble.AdaBoostClassifier библиотеки sklearn, на вход которой подаётся ба-

зовый классификатор. В данном случае в качестве базового выступал классификатор, основанный на μ/E алгоритме. Были получены следующие результаты: ошибка на обучающих данных - 0.0%, на тестовых – 2.0%.

Для сравнения был проведён идентичный численный эксперимент, но с применением в качестве базового классификатора стандартного DecisionTreeClassifier – классификация с помощью деревьев решений. В этом случае ошибки составили: 0.0% на обучающих данных и 6.0% на тестовых.

Таким образом, очевидно, что в данном случае разработаный метод даёт сравнительно более хорошие результаты, чем стандартный классификатор. Следовательно, можно применить полученный комплекс для решения конкретных практически важных задач.

5.2.3.2. Апробация работы слияния алгоритма μ/E и Adaboost на задаче регрессии

В настоящее время большое значение для радиоэлектронной промышленности имеет поиск материалов, которые могут быть использованы для изготовления прозрачных проводников.

Для того, чтобы отобрать наиболее многообещающие материалы для проведения реальных экспериментов, можно применить прогнозирование для оценки интересующих свойств материала на основе информации о его составе.

Такие технологии позволяют ускороить процесс разработки прозрачных проводников.

Рассмотрим задачу определения параметров прозрачных проводников. Каждый проводник описывается 11 свойствами, среди которых:

- метка, идентифицирующая симметрию материала;
- общее количество атомов Al, Ga, In и O;
- относительные составы Al, Ga и In;
- решетчатые векторы и углы.

Задача состоит в прогнозировании значения такого ключевого свойства проводников как энергии образования – важный показатель устойчивости материала.

Для оценки качества алгоритма был использован набор данных, включающий 3000 примеров.

Данная задача относится к разделу задач регрессии. Для её решения был применён подход, описанный в 4.3..

Выяснилось, что у разработанной системы, проверенной на простой задаче, появляются трудности на текущих данных. Проблемы возникают на этапе определения параметров многомерных возможностных распределений при использовании алгоритма поиска минимальных описанных эллипсоидов из работы [51].

Было выявлено две проблемы:

1. Вычислительная неустойчивость, которая может приводить к ошибкам расчёта параметров эллипса и низкому качеству работы полученной машины нечёткого вывода на сложных данных.

2. Низкая скорость работы.

Вследствие чего было решено отказаться от использования библиотеки для расчёта минимального эллипсоида. В качестве приближения к минимальному эллипсоиду было предложено использовать построение эллипсоида, ориентированного по осям, которые определяются с помощью метода главных компонент.

В таблице 5.2 отражены полученные результаты. Можно сделать вывод о том, что наилучший результат был достигнут при радиусе, равном 0.7, в композиции из 10 алгоритмов и его коэффициент детерминации составил 0.736.

Данная задача также была решена с помощью одного из самых популярных в настоящее время методов анализа данных – деревьев решений с применением бустинга. Были получены следующие результаты: при количестве алгоритмов в композиции равном 10, коэффициент детерминации составил 0.6610, при количестве алгоритмов в композиции равном 15 – 0.6296.

Таким образом, можно сделать вывод о том, что разработанный алгоритм превосходит используемые в настоящее время решения для задач такого рода.

Таблица 5.2

Значения коэффициента детерминации при разном количестве алгоритмов в композиции бустинга и разных значениях радиуса на обучающих и тестовых выборках

Радиус	10 алгоритмов		15 алгоритмов	
	R^2 на об. выб.	R^2 на тест. выб.	R^2 на об. выб.	R^2 на тест. выб.
0.1	0.6919	0.665	0.6501	0.6466
0.2	0.7424	0.7076	0.7117	0.6652
0.3	0.7532	0.6832	0.7026	0.6723
0.4	0.7164	0.6735	0.7164	0.6735
0.5	0.7364	0.6851	0.7065	0.6793
0.6	0.7277	0.6765	0.7140	0.6922
0.7	0.7360	0.6993	0.7434	0.7336
0.8	0.7257	0.6861	0.7484	0.6926
0.9	0.7115	0.7115	0.7099	0.7053

5.3. Выводы по пятой главе

В пятой главе сделано описание библиотек программ и прототипов систем прогнозирования, разработанных в ходе работы над диссертацией.

В параграфе 5.1 описывается реализация алгоритма обучения нечёткого контроллера в среде Matlab. Разработанная программа поддерживает модели нечёткого вывода Сугено 0-го и 1-го порядка. Пользователь может выбрать вид аппроксимируемой функции и параметры алгоритма. Возможно проведение серии испытаний с вычислением статистики результатов. Строятся графики исходной функции и её аппроксимации, ошибки аппроксимации. Для сравнения строится график ошибки аппроксимации по методу ANFIS.

В параграфе 5.2 производится реализация алгоритма обучения нечёткого контроллера на языке программирования Python (Приложение Г). Алгоритм реализован в соответствии с соглашениями библиотеки scikit-learn, что позволяет интегрировать его с современными системами анализа данных.

Параграф 5.2.1 посвящён решению проблемы прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте с применением разработанных технологий. Производится поставновка задачи и первичный анализ данных. Далее происходит апробация двухуровневой системы анализа данных, приведённой в параграфе 4.2 и оценка её работы

по сравнению с алгоритмами ANFIS и Claster Estimation.

Анализ работы алгоритмов показал, что в случае ненормализованных данных использование разработаннно метода оценки параметров функций распределения нечётких термов позволяет более точно описать структуру данных, чем с помощью задаваемого пользователем параметра – радиуса R для алгоритма Claster Estimation и ANFIS.

В целом можно сказать, что разработанная двухуровневая система анализа данных справляется с задачей прогнозирования лучше, чем алгоритм Claster Estimation и ANFIS.

В параграфе 5.2.2 описывается применение алгоритма Adaboost совместно с разработанным алгоритмом обучения машин нечёткого вывода к типовой задаче классификации (параграф 5.2.2.1) и к задаче определения параметров прозрачных проводниковых материалов (параграф 5.2.2.2). По результатам работы можно сделать вывод о том, что разработанный алгоритм превосходит используемое в настоящее время решения для задач такого рода.

По результатам работы была написана библиотека программ для прогнозирования и

Заключение

В работе были рассмотрены существующие методы построения оценки параметров возможностных распределений в одномерном случае.

На основе работ Wang Xizhao и Ha Minghu автором была получена максиминная оценка параметров для многомерных возможностных распределений, доказано, что оценка является достаточной и состоятельной при фиксированной матрице A. Также была получена оценка параметров многомерных возможностных распределений при заданом уровне возможности – расширение подхода Dug Hun Hong.

Для случая Архимедовых t-норм была получена оценка параметра, отвечающего за расположение центра распределения, при фиксированной матрице A и разработан алгоритм получаения оценки матрицы A при фиксированном параметре, отвечающем за расположение центра распределения.

В ходе исследования был разработан алгоритм обучения машин нечёткого вывода, на основе метода оценки параметров многомерных возможностных распределений. Было проведено сравнение результатов с результатами алгоритма Cluster Estimation на базовом примере. Далее сравнение было проведено на более сложной задаче прогнозирования значений нестационарного временного ряда, порождаемого дифференциальным уравнением Mackey-Glass. И на основе проведённых исследований была произведена модифиция алгоритма путём замены метода кластеризации k-means на Mountain method.

Была рассмотрена суть процедуры бустинга и алгоритм Adaboost и исследована возможность использования алгоритма Adaboost для построения композиций систем нечёткого вывода, обучающихся разработанным алгоритмом. Работоспособность метода продемонстрирована на примере задачи регрессии. Было достигнуто значение коэффициента детерминации, равное 0.99, что является показателем высокого качества работы представленного алгоритма.

Был разработана двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе. В данной системе роль глубокой нейронной сети состоит в сокращении размерности исходных данных. Применение ограниченных машин Больцмана для построения нейронной сети позволяет

выявить существенные закономерности в данных и сократить размерность. Для обучения машин нечёткого вывода было предложено применять разработанный алгоритм обучения.

Разботанная двухуровневая система анализа данных была использована для решения задачи прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте. В процессе работы системы для прогнозирования используются данные, доступные в современных системах управления электрооборудованием пассажирских вагонов, что позволило внедрить систему без установки дополнительного оборудования.

Была произведена апробация и оценка её работы по сравнению с алгоритмами ANFIS и Claster Estimation. Проведённые вычислительные эксперименты показали, что использование для обучения машины нечёткого вывода предложенного в диссертации алгоритма позволяет достичь более высокой точности прогнозирования по сравнению с распространёнными алгоритмами обучения. Это создаёт предпосылки для повышения эффективности функционирования различных систем прогнозирования, управления и анализа данных, в том числе для повышения надёжности работы электрооборудования на предприятиях вагоностроения.

В ходе работы над диссертацией были разработаны библиотеки программ и прототипов систем прогнозирования:

1. Реализация алгоритма обучения нечёткого контроллера в среде Matlab.

   Разработанная программа поддерживает модели нечёткого вывода Сугено 0-го и 1-го порядка. Пользователь может выбрать вид аппроксимируемой функции и параметры алгоритма. Возможно проведение серии испытаний с вычислением статистики результатов. Строятся графики исходной функции и её аппроксимации, ошибки аппроксимации. Для сравнения строится график ошибки аппроксимации по методу ANFIS.

2. Реализация алгоритма обучения нечёткого контроллера на языке программирования Python. Алгоритм реализован в соответствии с соглашениями библиотеки scikit-learn, что позволяет интегрировать его с современными системами анализа данных.

Реализация на языке программирования Python совместно с алгоритмом Adaboost была применена для решения типовой задачи классификации

и для решения задачи прогнозирования параметров прозрачных проводниковых материалов.

По результатам работы можно сделать вывод о том, что разработанный алгоритм превосходит используемое в настоящее время решение для задач такого рода.

Проведённые исследования показали, что в случае использования модели, состоящей из одной системы нечёткого вывода, в зависимости от характера исследуемой системы, возможны два варианта:

- в части случаев наиболее эффективными оказываются системы, построенные на одномерных термах с помощью алгоритма ANFIS;
- в других случаях более эффективным является использование многомерных возможностных распределений, и в этом случае предложенный в работе алгоритм позволяет достичь более высоких результатов по сравнению с существующими аналогами.

При решении сложных практических задач применение разработанных в ходе диссертации гибридных систем структурно-параметрической идентификации позволяет достичь более высокой точности, чем при применении систем, использующих одну машину нечёткого вывода.

Список литературы

1. Арефьев В.И., Петров М.О., Талалаев А. Б., Язенин А.В., Сорокин С.В. Классификация состояния системы на основе технологий мягких вычислений // Нечёткие системы и мягкие вычисления. 2016. Т. 11. №1. С. 33-56.

2. Базовые понятия и операции обработки экспериментальных данных. — Электрон. дан. — Режим доступа: http://dvo.sut.ru/libr/opds/i130hodo_part1/2.htm. — Загл. с экрана.

3. Батыршин И.З., Недосекин А.О., Стецко А.А., Тарасов В.Б., Язенин А.В., Ярушкина Н.Г. Нечеткие гибридные системы. Теория и практика / Под ред. Н.Г. Ярушкиной. - , М.: ФИЗМАТЛИТ, 2007. - 208с.

4. Биргер И.А. Техническая диагностика. М., 1978.

5. Бустинг .– Электрон. дан. – Режим доступа: www.machinelearning.ru/wiki/index.php?title=Бустинг#cite_note-0. – Загл. с экрана.

6. Бустинг.– Электрон. дан. – Режим доступа: www.machinelearning.ru/wiki/index.php?title=Алгоритм_AdaBoost.

7. Высшая математика: Учеб. для вузов: В 3 т. / Я. С. Бугров, С. М. Никольский; Под ред. В. А. Садовничего. — 6-е изд., стереотип. — М.: Дрофа, 2004. — (Высшее образование: Современный учебник). Т. 2: Дифференциальное и интегральное исчисление. — 512 с.

8. Гордеев Р.Н., Язенин А.В. Метод решения одной задачи возможностного программирования // Изв. РАН. Теор. и сист. упр. 2006. №3. С. 112-119.

9. Демьянов В.Ф., Малоземов В.Н. Введение в минимакс. М.: Наука, 1972.

10. Дюбуа Д., Прад А. Теория возможностей. Приложения к представлению знаний в информатике. М.: Радио и связь, 1990.

11. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений.-М.:Мир, 1976.-165 с.

12. Лекции по алгоритмическим композициям.– Электрон. дан. – Режим доступа: http://www.machinelearning.ru/wiki/images/0/0d/Voron-ML-Compositions.pdf. – Загл. с экрана.

13. Метод максимального правдоподобия. — Электрон. дан. — Режим доступа: ru.wikipedia.org/wiki/Метод_максимального_правдоподобия. — Загл. с экрана.

14. Методы робастного, нейро-нечёткого и адаптивного управления: Учебник / Под ред. Н.Д. Егупова. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. - 744 с.

15. Окружность, описанная около треугольника. Полные уроки. — Электрон. дан. — Режим доступа: school.xvatit.com/index.php?title=Окружность,_описанная около треугольника. Полные_уроки. — Загл. с экрана.

16. Патрик Э. Основы теории распознавания образов. М., 1980.

17. Пегат А. Нечеткое моделирование и управление. М., 2013.

18. Распоряжение Правительства РФ от 17.06.2008 №877-р “О Стратегии развития железнодорожного транспорта в Российской Федерации до 2030 года".– Электрон. дан. – Режим доступа: http://doc.rzd.ru/doc/public/ru?STRUCTURE ID=704&layer id=5104&id=3997

19. Саати Т.Л. Взаимодействия в иерархических системах // Техническая кибернетика. 1979. №1. С. 68-84.

20. Сорокина И.В., Сорокин С.В. К задаче оценки параметров многомерных возможностных распределений // Нечеткие системы и мягкие вычисления. 2013. Том 8. №2. С. 101-111.

21. Язенин А.В. К задаче максимизации возможности достижения нечеткой цели // Изв. РАН. Теория и системы управления. 1999. №4. С.120-123.

22. Язенин А.В. Методы оптимизации и принятия решений при нечетких данных. Автореферат диссертации на соискание ученой степени доктора физико-математических наук. Тверь. 1995.

23. Язенин А.В. Основные понятия теории возможностей: математический аппарат для принятия решений в условиях гибридной неопределённости. - М.: ФИЗМАТЛИТ, 2016. - 144с.

24. Ackley D.H., Hinton G.E., Sejnowski T.J. A learning algorithm for Boltzmann machines // Cognitive science. Elsevier. 1985. №9(1). С.147–169.

25. Alsina C. On a family of connectives for fuzzy sets // Fuzzy Sets and Systems. 1985. №16. C.231-235.

26. Aristotle's Categories and De Interpretatione. Tr. J. L. Ackrill, Oxford University Press, 1963. 162 p.

27. Badard R. The law of large numbers for fuzzy processes and the estimation problem // Information Sciences. 1982. № 28. С.161–178.

28. Beliakov G., Pradera A., Calvo T., Aggregation Functions: A Guide for Practitioners. Studies in Fuzziness and Soft Computing, Springer-Verlag, Berlin Heidelberg. 2007. Vol.221. 361 p.

29. Bellman R., Giertz M. On analytic formalism of the theory of fuzzy sets // Information Sciences. 1973. Vol. 5. Pp. 149-156.

30. Bengio Y., Simard P., Frasconi P. Learning long-term dependencies with gradient descent is difficult // Neural Networks. IEEE Transactions on 5.2. 1994. Pp. 157-166.

31. Bilgiç T. & Türkşen I. B. Measurement-theoretic justification of fuzzy set connectives // Fuzzy Sets and Systems. 1995. Vol. 76(3). Pp. 289-308.

32. Bilgiç T. & Türkşen I. B. Measurement-theoretic frameworks for fuzzy set theory // Working notes of the IJCAI-95 workshop on "Fuzzy Logic in Artificial Intelligence". 14th International Joint Conference on Artificial Intelligence Montréal, Canada. 1995. Pp. 55-65.

33. Bilgiç T. & Türkşen I. B. Measurement of Membership Functions: Theoretical and Empirical Work. Chapter 3 in D. Dubois and H. Prade (eds) Handbook of Fuzzy Sets and Systems Vol. 1, Fundamentals of Fuzzy Sets, Kluwer. 1999. Pp. 195-232.

34. Breiman L., Friedman J.H., Olshen R.A., Stone C.J. Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software. 1984.

35. Breiman L. Random Forests // Machine Learning. October 2011. Vol. 45(1). Pp 5-32.

36. Cai K.-Y. Parameter estimations of normal fuzzy variables // Fuzzy Sets and Systems. 1993. Vol. 5. Pp. 179-185.

37. Chiu Stephen L. Fuzzy model identification based on cluster estimation // Fuzzy Sets and Systems. 1994. Vol. 3. Pp. 267-278.

38. Crowder R. S. Predicting the Mackey-Glass timeseries with cascade-correlation learning, in Proc. 1990 Connectionist Models Summer School,Carnegie Mellon University. 1990. Pp. 117-123.

39. Dishkant H. About membership functions estimation // Fuzzy Sets and Systems. 1981. Vol. 5. Pp. 141-147.

40. Dubois D., Prade H. Fuzzy Sets and Systems: Theory and Applications. New York:Academic Press, 1980.

41. Dubois D. and Prade H., Possibility Theory. Plenum Press, New York, 1988.

42. Dubois D. & Prade H. Fuzzy sets, probability and measurement // European Journal of Operational Research. 1989. Vol. 40. Pp. 135-154.

43. Freund Y., Schapire R. E. Experiments with a new boosting algorithm // International Conference on Machine Learning. 1996. Pp. 148–156.

44. Freund Y., Schapire R. E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting // Journal of computer and system sciences. 1997. Vol. 55. Article № SS971504. Pp. 119-139.

45. Fishburn P. C. Utility theory for decision making // Operations Research Society of America. Publications in operations research. 1970. Vol. 18.

46. Gershenfeld N. A., Weigend A. S. The Future of Time Series: Learning and Understanding. Times Series Prediction: Forecasting the Future and Understanding the Past, eds. Reading, MA: Addison-Wesley, 1993.

47. Giles R. Lukasiewicz logic and fuzzy set theory // International Journal of Man-Machine Studies. 1976. Vol. 8(3). Pp. 312-327.

48. Giles R., The concept of grade of mambership // Fuzzy Sets and Systems. 1988. Vol.25. Pp. 297-323.

49. Goguen J. The logic of inexact concepts // Synthese. 1969. Vol. 19. Pp. 325-373.

50. Gupta M.M., Qi J. Theory of T-norms and fuzzy inference methods // Fuzzy Sets and Systems. 1991. Vol. 40. Pp. 431-450.

51. Hans-Peter Schröcker, Uniqueness Results for Minimal Enclosing Ellipsoids // Computer Aided Geometric Design. December 2008. Vol. 25(9). Pp. 756-762.

52. Hayashi Y., Tsunashima H., Marumo Y. Fault Detection Of Railway Vehicle Suspension System Using Muliple-Model Approach // Journal of Mechanical Systems for Transportation and Logistics. 2008. Vol. 1(1).

53. Hinton G. Training products of experts by minimizing contrastive divergence // Neural Computation. 2002. Vol. 14(8). Pp. 1771–1800.

54. Hisdal E. Reconciliation of the Yes-No versus grade of membership dualism in human thinking, in M.M.Gupta, A.Kandel, W.Bandler&J.B.Kiszka(eds), Approximate Reasoning in Expert Systems. North-Holland, Amsterdam. 1985. Pp. 33-46.

55. Hisdal E. Are grades of membership probabilities? // Fuzzy Sets and Systems. 1988. Vol.25. Pp. 325-348.

56. Hong D.H. Parameter estimations of mutually T-related fuzzy variables // Fuzzy Sets and Systems. 2001. Vol. 123. Pp. 63–71.

57. Iris Data Set.– Электрон. дан. – Режим доступа:https://archive.ics.uci.edu/ml/datasets/iris. – Загл. с экрана.

58. Jang J.-S. R., ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man, and Cybernetics. 1993. Vol. 23(3). Pp. 665-685.

59. Jia S., Dhanasekar M. Detection of Rail Wheel Flats Using Wavelet Approaches // In Structural Health Monitoring. 2007. Vol. 6(2). Pp. 121-131.

60. Klement E.P., Mesiar R., Pap E. Triangular norms. I. Basic analytical and algebraic properties // Fuzzy Sets and Systems. 2004. Vol. 143. Pp. 5-26.

61. Klement E.P., Mesiar R., Pap E. Triangular norms. Position paper II: general constructions and parameterized families // Fuzzy Sets and Systems. 2004. Vol. 145. Pp. 411-438.

62. Kosko B. Fuzziness vs. probability // General Systems. 1990. Vol 17. Pp.211-240.

63. Kosko B., Fuzzy Systems as Universal Approximators. IEEE Transactions on Computers. 1994. , Vol. 43 (11). Pp. 1329-1333.

64. Lakoff G. Women, Fire and Dangerous Things: What categories reveal about the mind, The Universuty of Chicago Press, Chicago. 1987.

65. Lapedes. A., Farber. R.: Nonlinear signal processing using neural networks: Prediction and system modelling. Tech. Rep. LA-UR-87-2662, Los Alamos Nat. Lab., Los Alamos, NM, 1987.

66. Laurens van der Maaten. Learning a Parametric Embedding by Preserving Local Structure // Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AI-STATS), JMLR W&CP 5:384-391, 2009. Pp.384-391.

67. Li P., Goodal R.M. Model Based Approach to Railway Vehicle Fault Detection and Isolation [Электронный ресурс]. – 2003. – Режим доступа: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.65.6317&rep=rep1&type=pdf.

68. Mackey M.C., Glass L. Oscillation and chaos in physiological system // Science. 1977. Pp. 287 - 289.

69. Mabuchi S. An interpretation of membership functions and the properties of general probabilistic operators as fuzzy set operators. Part I: Case of type 1 fuzzy sets // Fuzzy Sets and Systems. 1992. Vol. 49(3). Pp. 271-283.

70. Marsland S. Novelity Detection in Learning Systems // Neural Computing Surveys. 2002. Vol. 3. Pp. 1-39.

71. Nahmias S. Fuzzy variables // Fuzzy Sets and Systems. 1978. Vol.1. Pp. 97-110.

72. Näther W., On possibilistic inference // Fuzzy Sets and Systems. 1990. Vol. 36. Pp. 327-337.

73. Näther W. and Albrecht M., Fuzzy model fitting based on the truth of the model. Freiberger Forschungshefte. 1987. 187p.

74. Norwich A. M. & Türkşen I. B. The fundamental measurement of fuzziness, in R. R. Yager (ed.) // Fuzzy Sets and Possibility Theory: Resent Developments. Pergamon Press, New York. 1982. Pp. 49-60.

75. Norwich A. M. & Türkşen I. B. A model for the measurement of membership and the consequences of its emperical implementation // Fuzzy Sets and Systems. 1984. Vol. 12. Pp. 1-25.

76. Rao M.B., Rashed A. Some comments on fuzzy variables // Fuzzy Sets and System. 1981.Vol. 6. Pp. 285-292.

77. Raul Rojas. Neural Networks. A systematic introduction. Springer-Verlag, Berlin, 1996.

78. Rosch E. & Mervis C.B. Family resemblances: Studies in the internal structure of categories // Cagnitive Pschology. 1975. Vol. 7. Pp. 573-605.

79. Salakhutdinov R., Hinton G. An Efficient Learning Procedure for Deep Boltzmann Machines // Neural Computation. 2012. Vol. 24(8). Pp. 1967 - 2006.

80. Scipy.optimize.fmin. — Электрон. дан. — Режим доступа: https://docs.scipy.org/doc/scipy-0.16.0/reference/generated/scipy.optimize.fmin.html. — Загл. с экрана.

81. Takagi T., Sugeno M., Fuzzy identification of systems and its applications to modeling and control. IEEE Trans. Systems, Man, Cybernetics 1985 SMC-15, (1), 116–132.

82. Thomas S.F. Fuzziness and Probability, AGG Press, Wichita Kanzas USA, 1995.

83. Tsunashima H., Naganuma Y., Matsumoto A., Mizuma T., Mori H. Condition Monitoring of Railway Track Using In-Service Vehicle, Reliability and Safety in Railway, Dr.Xavier Perpinya (Ed.), ISBN: 978-953-51-0451-3, InTech, 2012

84. Weber S. A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms // Fuzzy Sets and Systems. 1983. Vol. 11. Pp. 115-134.

85. Welling M., Rosen-Zvi M., Hinton G. Exponential family harmoniums with an application to information retrieval // Advances in Neural Information Processing Systems. 2004. Vol. 17. Pp. 1481–1488.

86. Wolfs P., Bleakley S., Senini S., Thomas P. A Distributed Low Cost Device for the Remote Observation of Track and Vehicle Interactions, CORE 2006, 30 April – 3 May 2006, Melbourne, Australia, Railway Technical Society of Australasia, Melbourne.

87. Xiukun W., Limin J., Hai L. A comparative study on fault detection methods of rail vehicle suspension systems based on acceleration measurements // Vehicle System Dynamics. 2013. Vol. 51(5). Pp. 700-720.

88. Xizhao W., Minghu H. Possibilistic statistics, Proc. Sino Japan Joint Meeting on Fuzzy Sets and Systems (International Academic Publishers, 1990).

89. Xizhao W., Minghu H. Fuzzy linear regression analysis // Fuzzy Sets and Systems. 1992. Vol. 52. Pp. 179 -188.

90. Xizhao W., Minghu H. Note on maxmin μ/E estimation // Fuzzy Sets and Systems. 1998. Vol. 94. Pp. 71-75.

91. Yager R.R. A measurement-informational discussion of fuzzy union and intersection // International Journal of Man-Machine Studies. 1979. Vol. 11. Pp. 189-200.

92. Yager R.R. On a general class of fuzzy connectives // Fuzzy Sets and Systems. 1980. Vol. 4. Pp. 235-242.

93. Yager R. R., Filev D. P. Approximate clustering via the mountain method. IEEE Transactions on Systems, Man and Cybernetics. 1994. Vol. 24. Pp. 1279-1284.

94. Yazenin A.V. Fuzzy and stochastic programming // Fuzzy Sets and Systems. 1987. Vol. 22. Pp. 171-180.

95. Yazenin A.V. On the problem of possibilistic optimization // Fuzzy Sets and Systems. 1996. Vol.81. Pp. 133-140.

96. Yazenin A., Wagenknecht M. Possibilistic optimization. A measure-based approach. BUTC-UW, 1996. Vol.6.

97. Zadeh L.A. Fuzzy Sets // Information and Control. 1965. Vol. 8. Pp. 338-353.

98. Zadeh L.A. Similarity relations and fuzzy orderings // Information Sciences. 1971. Vol. 3. Pp. 177-200.

99. Zadeh L.A. Fuzzy logic and approximate reasoning // Synthese. 1975. Vol. 30. Pp. 407-428.

100. Zadeh L.A., Fuzzy sets as a basis for a theory of possibility // Fuzzy Sets and Systems. 1978. Vol. 1. Pp. 3-28.

101. Zimmermann H. J. & Zysno P. Quantifying vagueness in decision models // European Journal of Operational Research. 1985. Vol. 22. Pp. 148-154.

102. Zhen-Yuan W., Shou-Mei L. Fuzzy linear regression analysis of fuzzy-valued variables // Fuzzy Sets and Systems. 1990. Vol. 36. Pp. 125 - 136.

103. Zwick R., Carlstein E. & Budescu D. V., Measures if similarity among fuzzy concepts: A comparative analysis // International Journal of Approximate Reasoning. 1987. Vol. 1(2). Pp. 221-242.

104. Zysno P. Modeling membership functions in B.B.Rieger(ed.), Empirical Semantics I, Vol.1 of Quantitative Semantics, Studienverlag Brockmeyer, Bochum, 1981, Vol. 12. Pp. 350-375.

Приложение А
Одно из собственных чисел матрицы A

(для двумерной выборки объёмом 5 точек и t-нормы произведения)

$$(-(2b_1(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5)cos(\alpha)^2 - 2b_1(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5)sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5) - (-2b_1(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5)sin(\alpha)cos(\alpha) + 2b_1(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5)sin(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5)) \cdot exp - ((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5) \cdot exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5))(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))/(b_1b_2) + (-(2b_1(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5)cos(\alpha)^2 -$$

$$2b_1(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5)sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5) - (-2b_1(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5)sin(\alpha)cos(\alpha) + 2b_1(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5)sin(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2cos(\alpha)^2 + b_2^2sin(\alpha)^2)(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2sin(\alpha)cos(\alpha) + b_2^2sin(\alpha)cos(\alpha))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))/(b_1b_2) + (-(2b_1(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5)cos(\alpha)^2 - 2b_1(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5)sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5) - (-2b_1(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5)sin(\alpha)cos(\alpha) + 2b_1(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5)sin(\alpha)^2)(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5))exp(-((b_1^2sin(\alpha)^2 + b_2^2cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 +$$

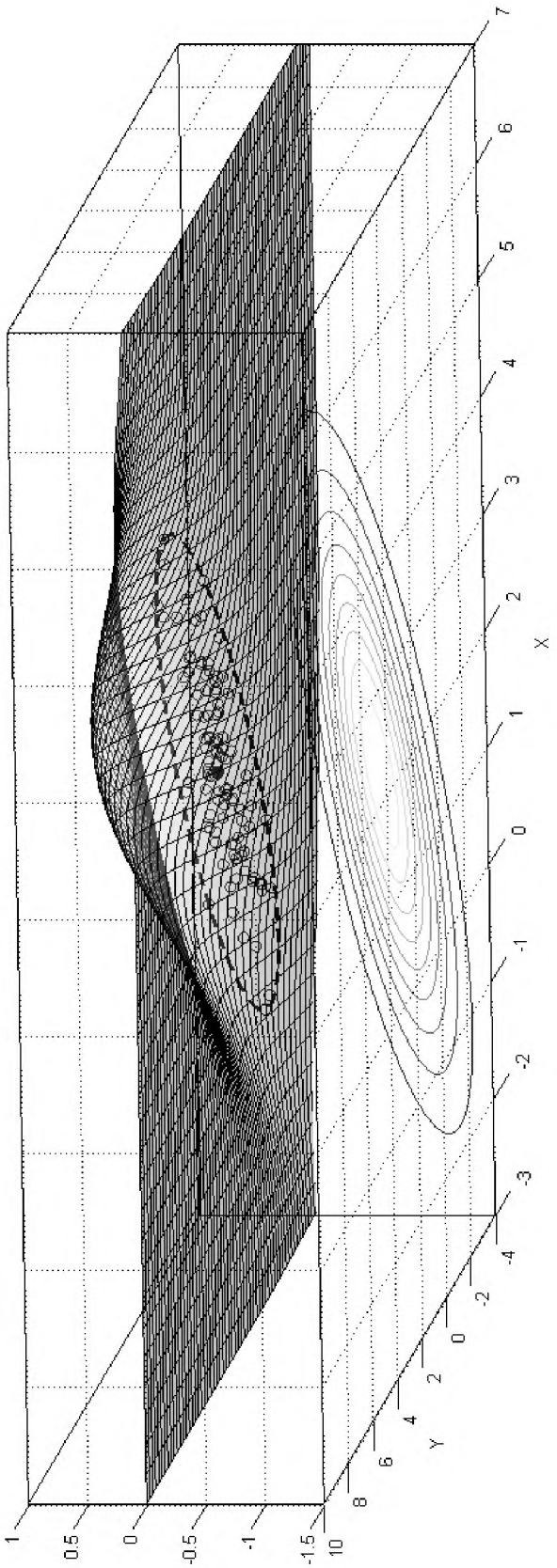
$4x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))exp(-((b_1^2 sin(\alpha)^2 + b_2^2 cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 + 4x_{24}/5 - x_{25}/5))(-x_{11}/5 - x_{12}/5 - x_{13}/5 + 4x_{14}/5 - x_{15}/5))exp(-((b_1^2 sin(\alpha)^2 + b_2^2 cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 + 4x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 - x_{12}/5 + 4x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2 sin(\alpha)^2 + b_2^2 cos(\alpha)^2)(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2 sin(\alpha)^2 + b_2^2 cos(\alpha)^2)(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))/(b_1 b_2) + (-(2b_1(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5)cos(\alpha)^2 - 2b_1(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5)sin(\alpha)cos(\alpha))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) - (-2b_1(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5)sin(\alpha)cos(\alpha) + 2b_1(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5)sin(\alpha)^2)(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))exp(-((b_1^2 sin(\alpha)^2 + b_2^2 cos(\alpha)^2)(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(-x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 + 4x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(-x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 + 4x_{25}/5))(-x_{11}/5 - x_{12}/5 -$

$$x_{13}/5-x_{14}/5+4x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5))(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5))(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5))(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5))(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5))(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5))(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))/(b_1b_2)+(-(2b_1(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)cos(\alpha)^2-2b_1(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)sin(\alpha)cos(\alpha))(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)-(-2b_1(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)sin(\alpha)cos(\alpha)+2b_1(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)sin(\alpha)^2)(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5+4x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5+4x_{15}/5))(-x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5+4x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5+4x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5+4x_{25}/5))(-x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5+4x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5))(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5))(-x_{11}/5-x_{12}/5-$$

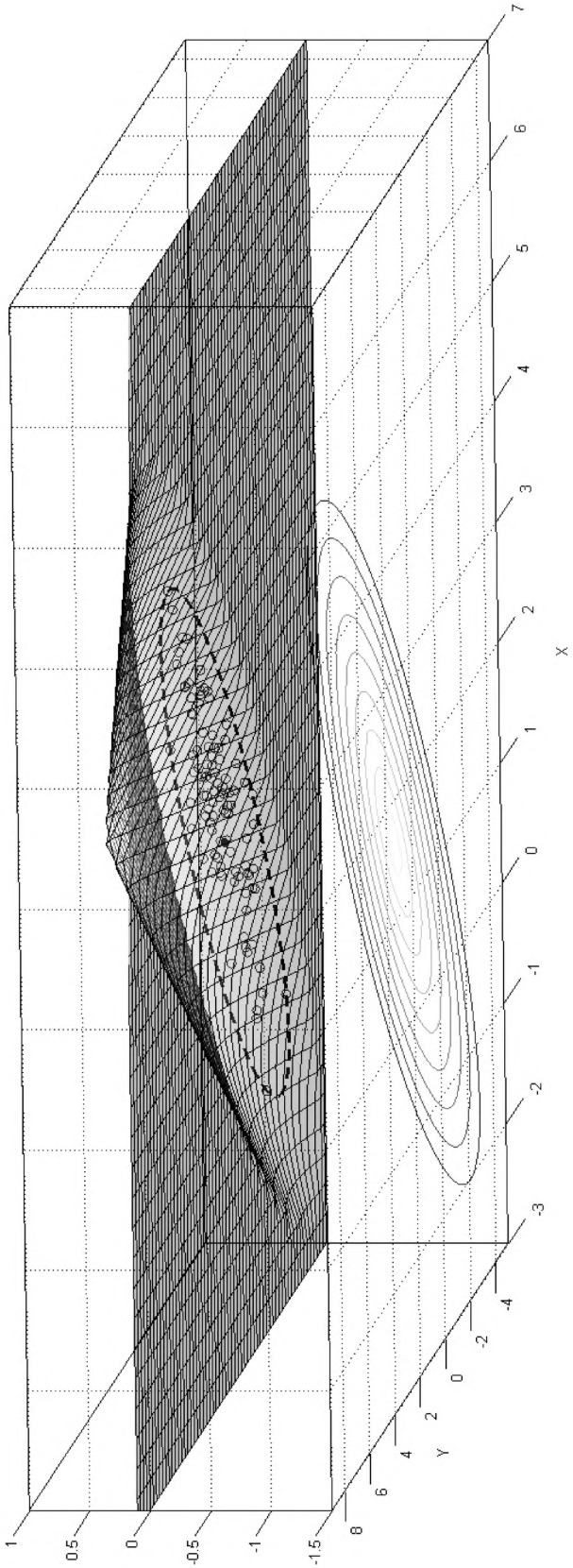
$$x_{13}/5+4x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5))(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5))(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5))(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(4x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5))(4x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))/(b_1b_2)-exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5+4x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5+4x_{15}/5))(-x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5+4x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5+4x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5-x_{23}/5-x_{24}/5+4x_{25}/5))(-x_{11}/5-x_{12}/5-x_{13}/5-x_{14}/5+4x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5))(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5-x_{23}/5+4x_{24}/5-x_{25}/5))(-x_{11}/5-x_{12}/5-x_{13}/5+4x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5))(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{21}/5-x_{22}/5+4x_{23}/5-x_{24}/5-x_{25}/5))(-x_{11}/5-x_{12}/5+4x_{13}/5-x_{14}/5-x_{15}/5))exp(-((b_1^2sin(\alpha)^2+b_2^2cos(\alpha)^2)(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)+(-b_1^2sin(\alpha)cos(\alpha)+b_2^2sin(\alpha)cos(\alpha))(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5))(-x_{21}/5+4x_{22}/5-x_{23}/5-x_{24}/5-x_{25}/5)-((b_1^2cos(\alpha)^2+b_2^2sin(\alpha)^2)(-x_{11}/5+4x_{12}/5-x_{13}/5-x_{14}/5-x_{15}/5)+(-b_1^2sin(\alpha)cos(\alpha)+$$

$b_2^2 sin(\alpha)cos(\alpha))(-x_{21}/5 + 4x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(-x_{11}/5 + 4x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))exp(-((b_1^2 sin(\alpha)^2 + b_2^2 cos(\alpha)^2)(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5) - ((b_1^2 cos(\alpha)^2 + b_2^2 sin(\alpha)^2)(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5) + (-b_1^2 sin(\alpha)cos(\alpha) + b_2^2 sin(\alpha)cos(\alpha))(4x_{21}/5 - x_{22}/5 - x_{23}/5 - x_{24}/5 - x_{25}/5))(4x_{11}/5 - x_{12}/5 - x_{13}/5 - x_{14}/5 - x_{15}/5))/(b_1^2 b_2)$

Приложение Б
Востановленное двумерное нормальное распределение



Приложение B
Восстановленное распределение, порождённое линейной функцией



Приложение Г
Библиотека программ для прогнозирования

Класс Predictor

```
import pandas as pd
import numpy as np
import scipy.sparse as sp
import sys
from sklearn.base import BaseEstimator, ClusterMixin,
   ↪ RegressorMixin, ClassifierMixin
from sklearn import linear_model
from sklearn import cross_validation
from sklearn import ensemble
from sklearn import metrics
import matplotlib.pyplot as plt
from matplotlib import cm
from sklearn import datasets
from itertools import count
from sklearn import preprocessing
from Clusterer import Clusterer
from sklearn import tree

class Predictor(BaseEstimator, RegressorMixin):

    def __init__(self, sigma = 0.18, radius = 0.75):
        self._sigma = sigma
        self._radius = radius
        self._estimator_type = "regressor"

    def _findParam(self, P):
        tolerance = 0.0001;
        d = P.shape[0]
        N = P.shape[1]
```

```
        Q = np.zeros((d+1,N))
        Q[0:d,:] = P
        Q[d,:] = np.ones(N)
        count = 1;
        err = 1;
        u = (1.0/N) * np.ones(N)
        while (err > tolerance):
                        X = np.dot((np.dot(Q, np.diag(u))), Q.T)
            M = np.diag(np.dot(np.dot(Q.T, np.linalg.inv(X)), Q))
            maximum = np.max(M)
            j = np.argmax(M)
            step_size = (maximum - d - 1)/((d + 1)*(maximum - 1))
            new_u = (1 - step_size)*u
            new_u[j] = new_u[j] + step_size
            count = count + 1
            err = np.linalg.norm(new_u - u)
            u = new_u
        U = np.diag(u)
        A = (1/d) * np.linalg.inv(np.dot(np.dot(P, U), P.T) - np.
           ↪ dot(P, u)*((np.dot(P, u)).reshape((-1, 1), order = 'F
           ↪ ')))
        C = np.dot(P, u)
        vv = np.linalg.eigvals(A)
        return (A, C)

    def _findNormalizedLevels(self, enters, C, A):
        N = enters.shape[0];
        nOfDimens = enters.shape[1];
        klast = C.shape[1];
        mu = np.zeros((N, klast));
        q = np.sqrt(nOfDimens/2);
        for k in range(0, klast):
            for i in range (0, N):
                XminusC = (enters[i, :] - C[:, k])
```

```
            XA = np.dot(XminusC, A[:,:,k])
            temp = np.dot(XA, XminusC)
            if (temp <= 0):
                temp = 0.01
            sqr = np.sqrt(temp)
            mu[i, k] = max(0, np.exp(-np.power(q*sqr, 2)))
    p = np.zeros((N,klast))
    for i in range(0, N):
        for k in range (0, klast):
            p[i,k] = mu[i,k]/(mu[i,:].sum())
    return p


def _splitClusters(self, enters, labels):
   clusters, counts = np.unique(labels, return_counts = True)
   clustToDelete = []
   for cluster, c in zip(clusters, counts):
        if c < 2:
            point = enters[labels == cluster][0]
            distance = np.sum((enters - point)**2, axis = 1)
            minVal = np.min(distance[distance > 0])
            labels[labels == cluster] = labels[distance ==
               ↪ minVal][0]
            clustToDelete.append(cluster)
   return [enters[labels == cl,:] for cl in clusters if cl not
      ↪  in clustToDelete]

def fit(self, enters, output, sample_weight = None):
    if (sample_weight is None):
        print("!Make␣new␣weights!")
        sample_weight = np.ones(enters.shape[0])
    if (len(output.shape) == 1):
        output = output[:, np.newaxis]
    trn_data = np.hstack((enters, output))
```

```
        g = Clusterer(radius = self._radius)
        labels = g.fit_predict(trn_data, sample_weight)
        labels[labels > self._clustNum] = self._clustNum
        clusters = self._splitClusters(enters, labels)
        klast = len(clusters)
        A = np.zeros((enters.shape[1], enters.shape[1], klast))
        C = np.zeros((enters.shape[1], klast))
        for i,p in enumerate(clusters):
            (A[:, :, i], C[:, i]) = self._findParam(p.T)
        self._A_ = A
        self._C_ = C
        for i in range(A.shape[2]):
            eigvals = np.linalg.eigvals(A[:,:,i])
        N = trn_data.shape[0];
        nOfDimens = trn_data.shape[1];
        p = self._findNormalizedLevels(enters, C, A)
        pPlusEnters = np.zeros((N,klast*nOfDimens))
        for i in range(0, N):
            for k in range(0, klast):
                pPlusEnters[i,nOfDimens*k] = p[i,k]
                for j in range(0, nOfDimens-1):
                    pPlusEnters[i, nOfDimens*k+j+1] = p[i, k]*
                        ↪ trn_data[i,j]
        pPlusEntersW = pPlusEnters * np.sqrt(sample_weight).
            ↪ reshape(-1, 1)
        outputW = (output.T * np.sqrt(sample_weight)).T
        h = np.linalg.lstsq(pPlusEntersW, outputW)
        self._H_ = h[0].reshape((nOfDimens, klast), order='F')
        return self


    def predict(self, enters):
        Nrul = enters.shape[0];
        Nvar = enters.shape[1];
```

```
        w = np.zeros(Nrul);
        klast = self._H_.shape[1];
        p = self._findNormalizedLevels(enters);
        sys.stdout.flush();
        for i in range(0, Nrul):
            for k in range(0, klast):
                outSum = 0;
                for j in range(0, Nvar):
                     outSum = outSum + enters[i, j]*self._H_[j+1,k];
                w[i] = w[i] + p[i, k]*(self._H_[0, k]+outSum)
        return w

    def score (self, X, y, sample_weight=None):
        w = self.predict(X)
        k = r2_score(y, w)
        return k
```

Класс Clusterer

```
import numpy as np
import scipy.sparse as sp
import sys
from sklearn.base import BaseEstimator, ClusterMixin
from sklearn import ensemble
import matplotlib.pyplot as plt
from sklearn import datasets
from itertools import count
from sklearn import preprocessing
from matplotlib import cm

class Clusterer(BaseEstimator, ClusterMixin):
    def __init__(self, radius = 0.25):
        self.set_params({"radius" : radius})

    def _subclust(self, X, radii1, sample_weight):
```

```
X = X.copy();
(numPoints, numParams) = X.shape;
options = [1.25, 0.5, 0.15, 1];
xBounds = [];
radii = np.ones(numParams)
radii *= radii1
sqshFactor = options[0]
acceptRatio = options[1]
rejectRatio = options[2]
verbose = options[3]
accumMultp = 1.0 / radii;
sqshMultp = 1.0 / (sqshFactor * radii);
if np.size(xBounds) == 0:
    minX = np.amin(X, axis=0);
    maxX = np.amax(X, axis=0);
    index = np.nonzero(maxX == minX);
    minX[index] = minX[index] - 0.0001*(1 + abs(minX[index
       ↪ ]));
    maxX[index] = maxX[index] + 0.0001*(1 + abs(maxX[index
       ↪ ]));
else:
    minX = xBounds[0,:];
    maxX = xBounds[1,:];
for id in range(0, numParams):
    X[:,id] = (X[:,id] - minX[id]) / (maxX[id] - minX[id]);
X = np.minimum(np.maximum(X,0),1);
potVals = np.zeros(numPoints);
new_accumMultp = accumMultp.reshape((1, accumMultp.shape
   ↪ [0]), order = 'F'
                                    )[np.zeros(numPoints,
                                       ↪ dtype = int), :]
for i in range(0, numPoints):
    thePoint = X[i,:];
    thePoint = thePoint.reshape((1, -1), order = 'F')[np.
```

```
        ↪ zeros(numPoints, dtype = int), :];
    dx = (thePoint - X) * new_accumMultp;
    if (numParams == 1):
        potVals[i] = np.sum(np.exp(-4*dx**2)*sample_weight[
            ↪ i])
    else:
        potVals[i] = np.sum(np.exp(-4*np.sum(dx**2,axis =
            ↪ 1))*sample_weight[i])
refPotVal = np.max(potVals);
maxPotIndex = np.argmax(potVals);
        maxPotVal = refPotVal;
centers = [];
numClusters = 0;
findMore = 1;
while (findMore and maxPotVal):
    findMore = 0;
    maxPoint = X[maxPotIndex,:];
    maxPotRatio = maxPotVal/refPotVal;
    if (maxPotRatio > acceptRatio):
        findMore = 1;
    elif maxPotRatio > rejectRatio:
        minDistSq = -1;
        for i in range(0, numClusters):
            dx = (maxPoint - centers[i,:]) * accumMultp;
            dxSq = np.dot(dx, dx.T);
            if (minDistSq < 0 or dxSq < minDistSq):
                minDistSq = dxSq;
        minDist = np.sqrt(minDistSq);
        if ((maxPotRatio + minDist) >= 1):
             findMore = 1
        else:
            findMore = 2
        if (findMore == 1):
        if (len(centers) != 0):
```

```
                centers = np.vstack((centers, maxPoint))
            else:
                centers = maxPoint.copy().reshape((1, -1), order
                    ↪ = 'F')
            numClusters = numClusters + 1
            new_sqshMultp = sqshMultp.reshape((1, -1), order =
                ↪ 'F')[np.zeros(numPoints, dtype = int),:]
            tmp = maxPoint.reshape((1, -1), order = 'F')[np.
                ↪ zeros(numPoints, dtype = int), :]
            dx = (tmp - X) * new_sqshMultp;
            if (numParams == 1):
                deduct = maxPotVal*np.exp(-4*dx**2)
            else:
                deduct = maxPotVal*np.exp(-4*np.sum(dx**2,1))
            potVals = potVals - deduct.T
            potVals[potVals < 0] = 0
            maxPotVal = np.max(potVals)
            maxPotIndex = np.argmax(potVals)
        else:
            if (findMore == 2):
                potVals[maxPotIndex] = 0
                maxPotVal = np.max(potVals)
                maxPotIndex = np.argmax(potVals)
    for i in range(0, numParams):
        centers[:,i] = np.dot(centers[:,i] , maxX[i] - minX[i])
            ↪  + minX[i]
    sigmas = (radii * (maxX - minX)) / np.sqrt(8.0);
    return centers

def transform(self, trn_data):
    N = trn_data.shape[0]
    k = self._centers.shape[0]
    d = np.zeros((N, k))
    for i in range (0, N):
```

```
            for j in range (0, k):
                d[i, j] = np.sum(np.power(trn_data[i,:] - self.
                    ↪ _centers[j, :], 2))
        return d

    def fit(self, trn_data, sample_weight = None):
        if (sample_weight is None):
            sample_weight = np.ones(trn_data.shape[0])
        self._centers = self._subclust(trn_data, self._radius,
            ↪ sample_weight)
        self._labels = self.predict(trn_data)
        return self

    def fit_predict(self, trn_data, sample_weight = None):
        self.fit(trn_data, sample_weight)
        return self._labels

    def fit_transform(self, trn_data, sample_weight):
        self.fit(trn_data, sample_weight)
        return self.transform(trn_data)

    def predict(self, newX):
        distances = self.transform(newX)
        labels = np.argmin(distances, axis = 1)
        return labels

    def set_params(self, params):
        self._radius = params["radius"]

    def get_params(self):
        return {"radius": self._radius}
```

Приложение Д
Акты о внедрении результатов диссертационной работы

УТВЕРЖДАЮ
Генеральный директор
ЗАО «РТИС-ВКО»
кандидат технических наук,
старший научный сотрудник
А.А. Храмичев
«28» 08 2018 г.

АКТ
внедрения результатов кандидатской диссертационной работы Сорокиной Ирины Владимировны
«Методы оценки параметров возможностных распределений и их применение для прогнозирования неисправностей электрооборудования»,
Представленной на соискание учёной степени кандидата технических наук по специальности 05.13.01 - «Системный анализ, управление и обработка информации (в промышленности)»

Комиссия в составе:

председателя – заместителя начальника отдела Арефьева В.И.

членов комиссии: начальника сектора Абрамова С.В., ведущего научного сотрудника Тихонова В.В.

составила настоящий акт о том, что материалы, разработанные в кандидатской диссертации Сорокиной И.В. использованы в работах, проводимых ЗАО «РТИС-ВКО» и опубликованы в книге «Разработка материалов в эскизный проект системы» (шифр «Сирена ВКО-РТИС ЭП») инв. № Р30/171с 2014г.

Председатель комиссии: к.т.н., снс В.И.Арефьев
Члены комиссии: к.т.н. С.В. Абрамов
к.т.н., профессор В.В.Тихонов

«28» августа 2018 г.

АКТ

внедрения результатов кандидатской диссертационной работы Сорокиной Ирины Владимировны

«Методы оценки параметров возможностных распределений и их применение для прогнозирования неисправностей электрооборудования», представленной на соискание учёной степени кандидата технических наук по специальности 05.13.01 - «Системный анализ, управление и обработка информации (в промышленности)»

Материалы, разработанные в диссертационной работе Сорокиной И.В. были применены ОАО «ТВЗ» для прогнозирования неисправностей вагонного электрооборудования на железнодорожном транспорте.

Разработанная в рамках диссертации Сорокиной И.В. двухуровневая система анализа данных, основанная на нейронных сетях и нечётком выводе, была успешно внедрена в рамках совместных проектов РФФИ и РЖД № 12-07-13117-офи_м_РЖД и №13-07-13160-офи_м_РЖД.

Начальник УИТ ОАО «ТВЗ» Сивров А.В.

28 августа 2018 г.

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное
образовательное учреждение
высшего образования
«Тверской государственный университет»
ТвГУ

ул. Желябова, д. 33, Тверь, 170100
Телефоны: 34-24-52, 32-15-50
Факс: (4822)32-12-74 E-mail: TverSU@tversu.ru
ОГРН 1026900577109
ИНН/КПП 6905000791/695001001

30.08.2018 № 57-02-81

На_____ от _____

СПРАВКА

о внедрении результатов диссертационной работы Сорокиной Ирины Владимировны в учебный процесс на факультете прикладной математики и кибернетики ФГБОУ ВО «Тверской государственный университет»

Результаты, полученные Сорокиной Ириной Владимировной в диссертации «Методы оценки параметров возможностных распределений и их применение для прогнозирования неисправностей электрооборудования»:

- методы оценки параметров многомерных возможностных распределений;
- алгоритм обучения машины нечёткого вывода,

внедрены в учебный процесс на факультете Прикладной математики и кибернетики Тверского государственного университета и используются в курсе "Анализ нечётких информационных систем", читаемом кафедрой Информационные технологии для магистров 2 курса по направлению подготовки 02.04.02 – «Фундаментальная информатика и информационные технологии».

Зав. кафедрой ИТ,
д. ф.-м. н., профессор _____ А.В. Язенин

Проректор по УВР,
д.г.н., профессор _____ Н.Е. Сердитова

013672

Приложение Е

Свидетельства о государственной регистрации программ для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2014662971

u/E-анализатор данных эксперимента ATLAS

Правообладатель: ***федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Тверской государственный университет» (RU)***

Авторы: ***Сорокина Ирина Владимировна (RU), Сорокин Сергей Владимирович (RU)***

Заявка № **2014660963**

Дата поступления **28 октября 2014 г.**

Дата государственной регистрации в Реестре программ для ЭВМ ***12 декабря 2014 г.***

Врио руководителя Федеральной службы по интеллектуальной собственности

Л.Л. Кирий

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2018617469

Библиотека для восстановления параметров многомерных нечётких распределений

Правообладатели: ***Сорокин Сергей Владимирович (RU), Сорокина Ирина Владимировна (RU)***

Авторы: ***Сорокин Сергей Владимирович (RU), Сорокина Ирина Владимировна (RU)***

Заявка № **2018612900**

Дата поступления **19 марта 2018 г.**

Дата государственной регистрации в Реестре программ для ЭВМ ***25 июня 2018 г.***

Руководитель Федеральной службы по интеллектуальной собственности

Г.П. Ивлиев