ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Сапкина Наталья Владимировна

ВОССТАНОВЛЕНИЕ ЗАКОНОМЕРНОСТЕЙ НА ОСНОВЕ НЕЧЕТКИХ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Специальность 05.13.17 – Теоретические основы информатики

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор технических наук, профессор
Леденёва Татьяна Михайловна

Воронеж – 2014

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Актуальность темы. Использование компьютерных технологий привело к пониманию важности задач, связанных с обработкой накопленной информации для извлечения знаний. Все более востребованным становится создание систем интеллектуального анализа данных, позволяющих выявить полезную скрытую информацию на основе классификации, кластеризации, статистического анализа, поиска ассоциативных правил и других подходов. Если данные представлены в виде динамических рядов каких-то показателей или их можно сгенерировать из базы данных, то для восстановления закономерностей используется техника регрессионного моделирования, при этом подразумевается, что данные являются числовыми. Однако, если информация относится к другому типу, например, является частично или полностью приближенной, то классические методы регрессионного анализа не применимы, и этот факт обусловливает необходимость их модификации. Одним из способов формализации приближенной информации является использование понятия нечеткого множества и его частного случая – нечеткого числа. Задача разработки регрессионных моделей, ориентированных на нечеткие числа, решалась зарубежными (H.Tanaka, P. Diamond, D. Dubois, M.S. Yang, M. Sakawa, M. Albrecht) и отечественными (Р.А. Алиев, А.Э. Церковный, Г.А. Мамедова, Н.Г. Ярушкина и др.) учеными. В общем случае методы нечеткого регрессионного моделирования могут быть разделены на две группы: первая базируется на методе наименьших квадратов и его модификациях, а вторая – на линейном программировании. Анализ показал, что рассмотрены далеко не все возможные постановки задач, учитывающих нечеткость исходных данных и/или параметров модели, кроме того, во многих исследованиях отсутствует комплексность подхода к реализации всех этапов регрессионного моделирования. Построение нечетких регрессионных моделей опирается на математический аппарат, включающий определение арифметических операций над нечеткими числами и их сравнение. Только для некоторых типов нечетких чисел результат арифметической операции представляет собой нечеткое число того же типа. В других случаях требуется дополнительная аппроксимация. Необходимость совершенствования существующих методов нечеткого регрессионного моделирования за счет учета различных типов данных и параметров, представленных нечеткими числами *L*-*R*-типа, и их реализации в рамках информационной системы интеллектуального анализа данных обусловливает актуальность диссертационного исследования.

Диссертационная работа выполнена в соответствии с одним из основных научных направлений Воронежского государственного университета «Математическое моделирование, программное и информационное обеспечение, методы вычислительной и прикладной математики и их применение к фундаментальным исследованиям в естественных науках».

Объект исследования – информационная система интеллектуального анализа данных, в которой реализуются нечеткие линейные регрессионные модели с коэффициентами в виде нечетких чисел *L*-*R*-типа.

Предмет исследования – нечеткий линейный регрессионный анализ на множестве нечетких чисел *L*-*R*-типа.

Цель работы и задачи исследования. Цель диссертационной работы заключается в развитии подходов к решению задачи восстановления закономерностей в данных на основе нечеткого регрессионного моделирования.

Для достижения поставленной цели решаются следующие **задачи:**

1. Анализ существующих подходов к восстановлению закономерностей в данных на основе регрессионного моделирования и выявление путей их совершенствования на случай приближенной исходной информации.

2. Выявление алгебраических свойств операций над нечеткими числами *L*-*R*-типа и разработка теоретической основы нечеткого регрессионного моделирования.

3. Определение оценок параметров нечетких регрессионных моделей и модификация общей процедуры регрессионного моделирования для выявления закономерностей в приближенной информации.

4. Разработка программного комплекса с применением современных компьютерных технологий для анализа и интеллектуальной обработки данных на основе предложенных алгоритмов нейросетевого и нечеткого регрессионного моделирования.

Методы исследования. В диссертационной работе использовались методы нечеткого и нейросетевого моделирования, линейной алгебры, математической статистики, теории нечетких множеств и нечеткой арифметики, объектно-ориентированного и модульного программирования.

Основные результаты, выносимые на защиту, и их научная новизна:

1. Совокупность теоретических результатов, касающихся свойств арифметических операций над нечеткими числами L-R-типа и существования алгебраических структур, что позволяет осуществлять вычисления при построении нечетких регрессионных моделей.

2. Оценка параметров нечетких линейных (парной и множественной) регрессионных моделей для задач, в которых исходная информация является полностью или частично приближенной, а коэффициенты моделей представлены обычными и/или нечеткими числами, что позволяет модифицировать общую процедуру регрессионного моделирования для восстановления закономерностей в разнородных и приближенных данных на основе использования нечетких чисел L-R-типа.

3. Альтернативные подходы к выявлению множества существенных переменных в рамках нечеткого регрессионного моделирования, основанные на нечетком коэффициенте корреляции, стандартизированном уравнении нечеткой множественной линейной регрессии и применении автоассоциативных нейронных сетей, «работающих» с приближенной информацией, что позволяет обеспечить комплексность анализа данных на различных этапах процесса выявления закономерностей в данных.

4. Информационная система интеллектуального анализа данных и структура программного комплекса, включающего блок нечеткой арифметики, который

может использоваться как самостоятельное приложение, и средства для проведения нечеткого линейного регрессионного моделирования.

Область исследования. Диссертационная работа соответствует следующему пункту Паспорта специальности 05.13.17 «Теоретические основы информатики»: п. 5. «Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях...».

Практическая значимость работы. Разработанная информационная система, в которой реализован предложенный комплекс алгоритмов нейросетевого анализа и нечеткого регрессионного моделирования, предназначена для обработки приближенной информации, выявления в ней функциональных зависимостей и проведения исследований в ситуациях, когда традиционные методы неприменимы. Результаты диссертационной работы используются для оценки качества выпущенной продукции с целью обоснования управленческих решений по совершенствованию технологических процессов специалистами ЗАО ЛЦ «АВС Фарбен», а также в учебном процессе ФГБОУ ВПО «Воронежский государственный университет» при чтении спецкурсов и выполнении курсовых и выпускных квалификационных работ.

Апробация работы. Основные результаты, полученные в диссертационной работе, докладывались и обсуждались на следующих международных и всероссийских конференциях: Международная научно-практическая конференция «Глобальная научная интеграция» (Тамбов, 2011); Международная конференция «Актуальные проблемы прикладной математики, информатики и механики» (Воронеж, 2011-2012); Всероссийская молодежная научная школа «Инженерия знаний. Представление знаний: состояние и перспективы» (Воронеж, 2012); Международная конференция«ExploIT Dynamics PhD Workshop» (Германия, г. Ольденбург, 2012); Международная конференция «Современные методы прикладной математики, теории управления и компьютерных технологий» (Воронеж, 2013); Международный научный семинар «Emerging Trends in Informations Systems (IS)» (Нижний Новгород, 2013).

Публикации. Основные результаты диссертации опубликованы в 12 научных работах, в том числе 5 – в изданиях, рекомендованных ВАК РФ. В работах, опубликованных в соавторстве, лично соискателю принадлежат: [1] – метод оценки параметров нечеткой линейной множественной регрессионной модели, анализ данных; [10] – детальная разработка и наполнение шагов нечеткого парного линейного регрессионного анализа.

Объём и структура работы. Диссертация состоит из введения, четырёх глав, заключения, списка использованных источников из 110 наименований, двух приложений. Основная часть работы изложена на 151 странице и включает 42 рисунка и 17 таблиц.

Глава 1. Информационные системы анализа данных на основе регрессионного моделирования

В данной главе рассмотрены информационные системы интеллектуального анализа данных, технологии их построения и архитектура; представлена классификация задач интеллектуального анализа данных; рассмотрены теоретические основы и этапы регрессионного моделирования; приведен подход к отбору наиболее информативных признаков для проведения множественного регрессионного анализа данных; указаны принципы построения стандартизированного уравнения нечеткой линейной множественной регрессии; рассмотрены существующие подходы к восстановлению закономерностей на основе нечеткого регрессионного моделирования; приведены цели и задачи исследования.

1.1. Информационные системы интеллектуального анализа данных

1.1.1. Технологии построения информационных систем анализа данных

Для выполнения задач анализа данных и поиска решений необходимо накопление и хранение достаточно больших объемов данных. Этим целям служат программно-аппаратные комплексы, называемые автоматизированными информационными системами (АИС) [38]. Основой построения современных АИС являются базы данных – модели некоторой предметной области, состоящие из связанных между собой данных об объектах, их свойствах и характеристиках [8]. Предполагается, что создание базы данных, поддержание ее в актуальном состоянии и обеспечение эффективного доступа пользователей и их приложений к содержащейся в ней информации осуществляется с помощью специального программного инструметария – системы управления базами данных (СУБД) [34].

Чтобы сохранять данные согласно какой-либо модели предметной области, структура базы данных должна максимально соответствовать этой модели. Наиболее распространены в настоящее время реляционные СУБД, основанные на реляционной модели данных, имеющей солидный теоретический фундамент – теорию множеств и исчисление предикатов. СУБД реляционной модели должна обеспечивать выполнение операций над базой данных, предоставляя при этом возможность одновременной работы нескольким пользователям (с нескольких компьютеров) и гарантируя целостность данных. Для соблюдения этих правил в СУБД используется механизм управления транзакциями [8].

Информационная система анализа данных основана на интеллектуальном подходе. Она обрабатывает большие массивы данных, осуществляет автоматизированный поиск ранее неизвестных закономерностей, скрытых и неочевидных правил в базах данных. Полученные знания помогают оптимизировать процессы деятельности предприятия и могут быть использованы для принятия решений [2, 8].

В качестве основных причин, способствующих распространению систем интеллектуального анализа данных выступают следующие [2]:

– определение того, что в больших по объемам базах данных содержатся скрытые ценные знания, способствующие повышению эффективности управления;

– развитие технологии информационных хранилищ позволяет создать единое информационное пространство, собрав требуемые для анализа данные в центральной базе;

– благодаря внедрению информационных хранилищ увеличивается число сотрудников организаций, получающих доступ к информации и способных принимать решения в той или иной области.

Впервые информационные хранилища данных [38] были определены У. Инмоном как предметно-ориентированный, интегрированный, неизменчивый, поддерживающий хронологию набор данных, организованный для целей поддержки принятия решений [8, 12, 93]. Концептуальная модель хранилища данных представлена на рисунке 1.1.

Использование концепции информационного хранилища в системе интеллектуального анализа данных (ИАД) нацелено на своевременное обеспечение аналитиков всей информацией, необходимой для выработки решений, и создание единого справочника метаданных, т. е. справочника информации о данных, содержащихся в хранилище [8, 38].

С концепцией информационных хранилищ тесно связан оперативный анализ, который выполняется средствами OLAP-систем. OLAP (On-Line Analytical Processing) – технология оперативной аналитической обработки данных, использующая методы и средства для сбора, хранения и анализа многомерных данных в целях поддержки процессов принятия решений [8, 93].

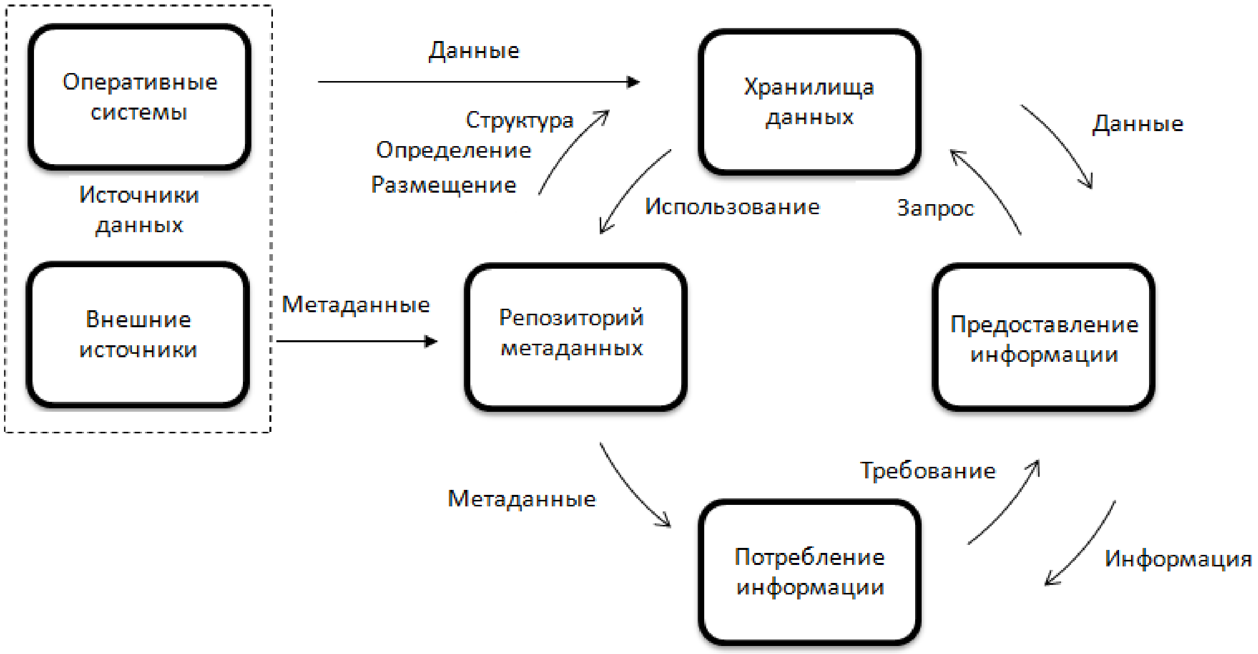


Рисунок 1.1 – Концептуальная модель информационного хранилища

Основное назначение OLAP-систем – поддержка аналитической деятельности, произвольных запросов пользователей-аналитиков. Целью

OLAP-анализа является проверка возникающих гипотез. Архитектура OLAP-системы включает в себя OLAP-сервер и OLAP-клиент. OLAP-сервер может быть реализован на основе многомерных баз данных (MOLAP), реляционных баз данных (ROLAP) или сочетания обеих моделей (HOLAP) [8].

1.1.2. Архитектура информационной системы интеллектуального анализа данных

Система интеллектуального анализа данных на основе технологии информационного хранилища состоит из четырех компонентов [2, 12]:

1) одного или нескольких серверов баз данных;
2) программного обеспечения для функционирования систем клиент/сервер;
3) программы загрузки данных в информационное хранилище из внешних источников, которая сопровождается предварительной обработкой данных;
4) клиентских приложений, предназначенных для поддержки принятия решений.

Процесс интеллектуального анализа осуществляется в три этапа [9] (рис. 1.2).

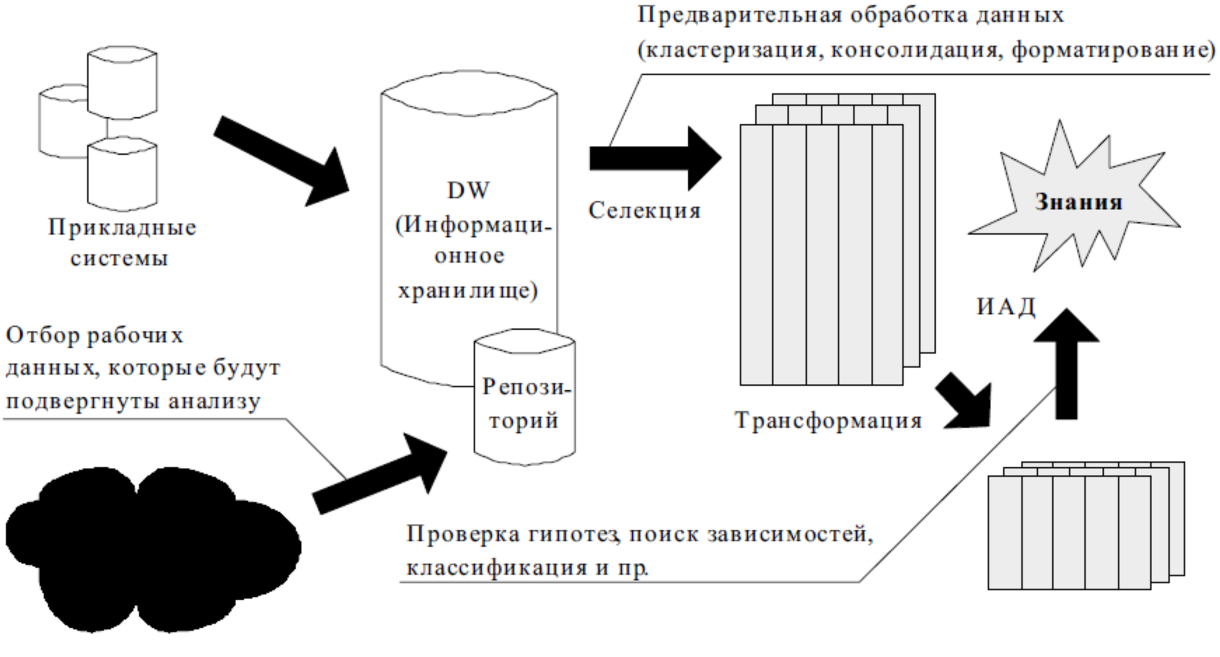


Рисунок 1.2 – Архитектура информационной системы ИАД

Выбор данных. Для решения конкретной задачи нужны не все данные из информационного хранилища, поэтому сначала выбирается подмножество, которое будет подвергнуто анализу. При этом может потребоваться объединение нескольких таблиц и фильтрация полученных записей.

Трансформация данных. После подготовки рабочих таблиц осуществляется предварительная обработка данных, характер которой определяется методами, применяемыми в ходе анализа. Трансформация может заключаться в удалении дублирующих записей или зашумленных данных, преобразовании типов, добавлении новых атрибутов и др.

Анализ. Трансформированные данные последовательно обрабатываются согласно определенной методике с целью извлечения требуемой информации или знаний.

В ходе ИАД выполняются операции, которые реализуются на основе различных алгоритмов. Методы ИАД могут быть разделены на два класса [2, 12]: операции проверки гипотез и операции поиска зависимостей, направленные на автоматическое выявление закономерностей или правил, которым подчиняются данные информационного хранилища. К недостаткам процедур первого класса можно отнести ограниченность анализа жесткими рамками заранее принятой гипотезы. Для второго класса система ИАД самостоятельно обрабатывает информацию с целью обнаружения внутренних закономерностей. Полученные результаты часто оказываются весьма неожиданными и ведут к нетривиальным выводам. Комбинируя операции этих двух классов, возможно реализовать самые различные стратегии анализа.

1.1.3. Классификация задач интеллектуального анализа данных

Методы интеллектуального анализа данных позволяют решить многие задачи, основными из которых являются: классификация, регрессия, поиск ассоциативных правил и кластеризация [8, 9].

Задача классификации сводится к определению класса объекта по его основным характеристикам. При этом заранее известно множество классов, к которым нужно отнести исследуемый объект [8].

Задача регрессии состоит в определении значения одного из параметров анализируемого объекта на основании значений других параметров. Определяемый параметр часто называют зависимой переменной, а параметры, участвующие в его определении, – независимыми переменными [7, 83].

Поиск ассоциативных правил нацелен на нахождение частых зависимостей (ассоциаций) между объектами и событиями. Полученные ассоциации представляются в виде правил и могут быть использованы как для лучшего понимания природы анализируемых данных, так и для предсказания появления событий [9].

Задача кластеризации заключается в поиске независимых групп (кластеров) и их характеристик во всем множестве анализируемых данных. Группировка однородных объектов позволяет сократить их число, а следовательно, облегчить анализ [8].

Перечисленные задачи делятся по назначению на описательные и предсказательные [8, 9].

Описательные задачи уделяют внимание улучшению понимания анализируемых данных. Ключевой момент в таких моделях – легкость и прозрачность результатов для восприятия человеком. К данному виду задач относятся кластеризация и поиск ассоциативных правил [9].

Решение *предсказательных задач* разбивается на два этапа. На первом строится модель на основании набора данных с известными результатами. На втором этапе она используется для предсказания результатов на основании новых наборов данных. При этом требуется, чтобы построенные модели работали максимально точно. К данному виду задач относятся классификация и регрессия [8].

1.2. Постановка задачи регрессионного анализа

1.2.1. Понятие регрессии и регрессионной модели

Регрессионный анализ – это раздел математической статистики, объединяющий практические методы исследования регрессионной зависимости между величинами на основе статистической информации. Исходные данные представляют собой пары значений зависимой (объясняемой) переменной y и независимой (объясняющей) переменной $X = \left(x^{(1)},\ldots,x^{(m)}\right)^T$ [55, 83]. Зависимая переменная выступает в роли функции, значения которой определяются (с некоторой случайной погрешностью) значениями объясняющих переменных, являющихся аргументами этой функции. Поэтому по своей природе переменная y всегда стохастична (случайна). Независимая переменная в существенной мере определяет процесс формирования значений зависимой переменной и может быть как случайной, так и неслучайной [17, 21, 59, 80].

Регрессией y по X называется зависимость $E(y\,|\,X) = f(X)$ математического ожидания некоторой случайной величины y от значения независимой переменной X [53, 83]. Задача регрессионного анализа заключается в поиске такой функции f, которая описывает эту зависимость. Регрессия может быть представлена в виде суммы неслучайной и случайной составляющих

$$y(X) = f(X) + \varepsilon, \tag{1.1}$$

где f – функция регрессионной зависимости, а ε – случайная составляющая [4, 42].

В качестве целей анализа регрессионных зависимостей вида (1.1) можно выделить [19, 29, 42]:

1. Определение степени детерминированности вариации зависимой переменной объясняющими переменными.

2. Предсказание значения зависимой переменной с помощью одного или нескольких значений независимых переменных.

3. Определение влияния каждой из объясняющих переменных на вариацию зависимой переменной.

Различают одномерную (парную) и многомерную (множественную) регрессию с одной и несколькими свободными переменными, а также линейную и нелинейную регрессию. Если регрессионная модель не является линейной комбинацией функций свободных переменных, то говорят о нелинейной регрессии. При этом модель может быть представлена в виде суперпозиции функций свободных переменных из некоторого набора. Нелинейные модели включают экспоненциальные, тригонометрические, и другие (например, радиальные базисные функции или персептрон Розенблатта) [21, 27, 59, 80].

Все выводы в регрессионном анализе строятся на основании имеющихся исходных статистических данных [4, 13].

Будем полагать, что задана выборка – множество $\{X_1,\ldots,X_n \mid X_i \in \Re^m\}$ значений анализируемых независимых переменных и множество $\{y_1,\ldots,y_n \mid y_i \in \Re\}$ соответствующих значений зависимой переменной на n статистически обследованных объектах. Эти множества обозначаются как D, множество исходных данных $\{(X,y)_i\}_{i=\overline{1,n}}$ [4, 83].

Обычно делается некоторое предположение о распределении y, однако чаще всего считается, что условные распределения y при каждом допустимом значении независимых переменных являются нормальными [42]. Объясненная часть в уравнении регрессии (1.1) представляет собой функцию вида [42, 83]:

$$y_e = f\left(x^{(1)},\ldots,x^{(m)}\right).$$

Наиболее естественным выбором объясненной части случайной величины y является ее среднее значение – условное математическое ожидание $M_{x^{(1)},\ldots,x^{(m)}}(y) = M_x(y)$, полученное при данном наборе значений независимых переменных. В самом деле, по своему смыслу объясненная часть – это ожидаемое значение зависимой переменной при заданных значениях независимых переменных. При таком выборе уравнение (1.1) приобретает вид [42]:

$$y = M_x(y) + \varepsilon, \tag{1.2}$$

где ε – случайная величина, называемая возмущением или ошибкой.

Уравнение (1.2) является уравнением регрессионной модели [42, 94].

Стоит отметить, что различают “математическую модель” и “регрессионную модель”. Первая предполагает участие аналитика в конструировании функции, которая описывает некоторую известную закономерность. При построении математической модели сначала создается параметрическое семейство функций, затем с помощью измеряемых данных выполняется “идентификация модели” – нахождение ее параметров. Известная функциональная зависимость объясняемой переменной от независимых переменных – основное отличие математического моделирования от регрессионного анализа [83].

Регрессионная модель объединяет широкий класс универсальных функций, которые описывают некоторую закономерность. При этом для построения модели в основном используются измеряемые данные, а не знание свойств исследуемой закономерности [82, 94].

Выделяют следующие основные этапы регрессионного моделирования [42]:

– определение цели исследования и формирование набора участвующих в модели переменных. В качестве цели моделирования обычно рассматривают анализ исследуемого объекта или процесса, прогноз его показателей, имитацию развития объекта при различных значениях переменных, выработку управленческих решений. При выборе переменных необходимо теоретическое обоснование каждой из них, они не должны быть связаны корреляционной зависимостью, так как это может привести к невозможности оценки параметров модели или к получению не имеющих реального смысла оценок;

– проведение анализа сущности изучаемого объекта, формализация известной до начала исследования информации;

– осуществление моделирования, т. е. выбор общего вида модели, выявление входящих в нее связей;

– осуществление статистического анализа модели и оценки ее параметров;

– проверка адекватности и качества модели;

– использование модели для анализа данных и прогнозирования.

1.2.2. Отбор наиболее существенных объясняющих переменных

При исследовании зависимости объясняемой переменной y от независимой (объясняющей) переменной $X = \left(x^{(1)},\ldots,x^{(m)}\right)^T$ необходимо решать ряд задач, характерных лишь для множественной регрессии и корреляции. К таким задачам относится отбор объясняющих переменных $x^{(1)},\ldots,x^{(m)}$, существенно влияющих на y, при этом необходимо учитывать следующие требования к переменным $x^{(1)},\ldots,x^{(m)}$ [41, 95]:

– объясняющие переменные не должны быть линейно зависимы, поскольку эта зависимость означает, что они характеризуют аналогичные свойства изучаемого явления (включение в модель линейно зависимых объясняющих переменных приводит к возникновению явления мультиколлинеарности);

– рекомендуется включать количественные объясняющие переменные (наряду с ними существуют и качественные);

– в одну модель нельзя включать обобщенную объясняющую переменную и образующие ее частные переменные – это приводит к неоправданно увеличенному их влиянию на зависимую переменную.

Особенностью множественной регрессии и корреляции является необходимость различать случаи корреляционной связи, когда переменные $x^{(1)},\ldots,x^{(m)}$ являются случайными величинами, регрессионной, если переменные $x^{(1)},\ldots,x^{(m)}$ – неслучайные величины, а также смешанный случай, когда некоторые из переменных – случайные величины, другие – неслучайные. В случае корреляционной зависимости следует вычислять и интерпретировать коэффициенты корреляции, при регрессионной зависимости это не имеет смысла, а при наличии как случайных, так и неслучайных переменных коэффициенты корреляции следует вычислять только между случайными переменными [41].

Рассмотрим отбор объясняющих переменных для построения множественной регрессионной зависимости, когда переменные $x^{(1)},...,x^{(m)}$ являются случайными величинами (обычно предполагается, что их совместное распределение нормальное). Необходимо определить, все ли переменные стоит включать в уравнение регрессии или есть переменные, которые существенно не влияют на величину y и их можно не учитывать [41, 42].

Для решения этого вопроса необходимо построить таблицу (табл. 1.1) коэффициентов парной корреляции для всех m объясняющих переменных [41, 95].

Таблица 1.1 – Таблица коэффициентов парной корреляции

	y	$x^{(1)}$	…	$x^{(m)}$
y	1	r_{yx^1}	…	r_{yx^m}
$x^{(1)}$	r_{x^1y}	1	…	$r_{x^1x^m}$
…	…	…	…	…
$x^{(m)}$	r_{x^my}	$r_{x^mx^1}$	…	1

В силу симметричности таблицы относительно главной диагонали можно свести её к треугольному виду.

Коэффициент корреляции случайных величин $x^{(1)},...,x^{(m)}$ и y определяется как отношение их корреляционного момента к произведению средних квадратических отклонений этих величин [1, 4, 42]:

$$r_{x^jy} = \frac{M(x^{(j)}y) - M(x^{(j)})M(y)}{\sqrt{D(x^{(j)})D(y)}}, \ j = \overline{1,m}.$$

Чем больше r_{x^jy}, тем теснее связь между y и $x^{(j)}$. Если $r_{x^jy} > 0{,}5$, то переменную $x^{(j)}$ лучше оставить, если же меньше – можно исключить. При каждом исключении переменной корреляционную таблицу следует пересчитывать [1, 4, 42].

Если $r_{x^s x^j}$, $s=\overline{1,m}$, мало, то считается, что связь между $x^{(s)}$ и $x^{(j)}$ мала, при этом обе переменные могут остаться; если значение велико, то нет смысла использовать обе переменные (по r_{yx^s} и r_{yx^j} как раз выбирается та, которая останется) [41, 42].

Таким образом, после отбора объясняющих переменных $x^{(1)},\ldots,x^{(m)}$, существенно влияющих на переменную y, будет сформировано множество значимых переменных $x^{(i_1)},\ldots,x^{(i_k)}$.

1.2.3. Парный линейный регрессионный анализ

Пусть установлен характер экспериментальных данных и выделен определенный набор объясняющих переменных.

Для нахождения объясненной части, т.е. величины $M_x(y)$ требуется знание условных распределений случайной величины y. При решении практических задач это почти никогда не имеет места, поэтому точное вычисление объясненной части невозможно.

В таких случаях применяется стандартная процедура сглаживания экспериментальных данных. Она включает в себя следующие два этапа [1, 42]:

1. Определяется параметрическое семейство, к которому принадлежит искомая функция $M_x(y)$. Это может быть множество линейных функций, показательных функций и т.д.;
2. Находятся оценки параметров этой функции на основе методов математической статистики.

Формально не существует никаких способов выбора параметрического семейства. Однако в большинстве случаев модель выбирается линейной. Кроме очевидного ее преимущества – относительной простоты, – для такого выбора имеются, по крайней мере, две существенные причины [4, 42]:

1. Если случайная величина (X, y) имеет совместное нормальное распределение, то, как известно, уравнения регрессии линейные. Предположение о нормальном распределении является вполне естественным и в ряде случаев может быть обосновано с помощью предельных теорем теории вероятностей. В других случаях сами величины y или X могут не иметь нормального распределения, но некоторые функции от них распределены нормально. Например, известно, что логарифм доходов населения – нормально распределенная случайная величина. Часто гипотеза о нормальном распределении принимается во многих случаях, когда нет явного ей противоречия, и, как показывает практика, подобная предпосылка оказывается вполне разумной.

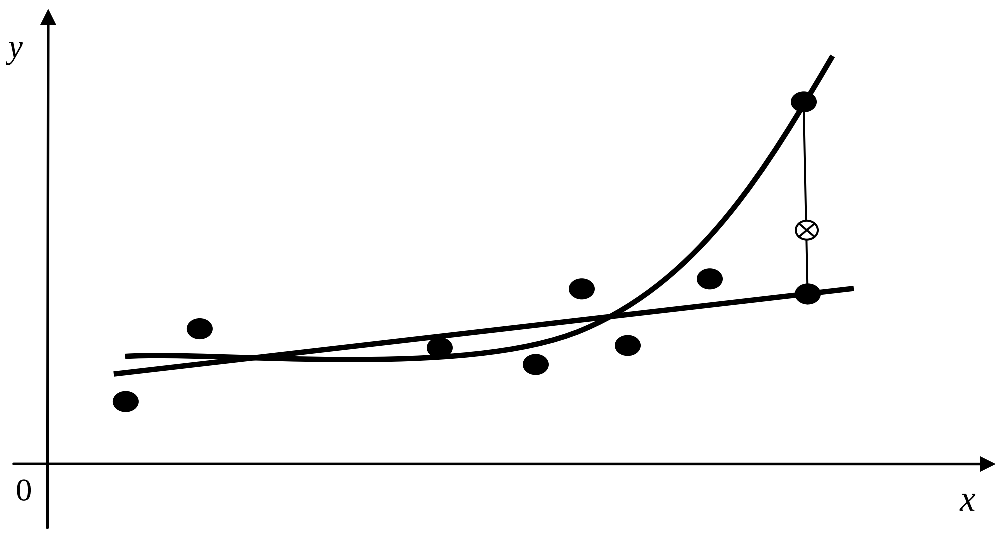


Рисунок 1.3 – Выбор функции регрессии

2. Меньший риск значительной ошибки прогноза. Рисунок 1.3 иллюстрирует два выбора функции регрессии – линейной и квадратичной. Как видно, имеющееся множество экспериментальных данных (точек) парабола сглаживает лучше, чем прямая. Однако парабола быстро удаляется от корреляционного поля и при добавлении наблюдения, обозначенного крестиком, теоретическое значение значительно отличается от эмпирического.

Можно придать точный математический смысл этому утверждению: ожидаемое значение ошибки прогноза оказывается меньше в том случае, если уравнение регрессии выбрано линейным [42].

Обратимся теперь к вопросам, связанным с априорными предположениями, оценкой коэффициентов и верификацией модели парной регрессии вида

$$y = a_0 + a_1 x + \varepsilon, \tag{1.3}$$

где y – зависимая переменная, состоящая из двух слагаемых: неслучайной составляющей $y_1 = a_0 + a_1 x$ (x – независимая переменная, a_0, a_1 – параметры уравнения) и случайного члена ε [4, 42].

Допустим, что данные представлены в двух видах: табличном (табл. 1.2) и графическом (рис. 1.4). Предположим, что истинная зависимость между x и y – линейная, т. е. существует некоторая прямая $y_1 = a_0 + a_1 x$, отражающая истинную зависимость. Задача регрессионного анализа состоит в получении оценок a_0, a_1, а следовательно, и положения прямой. На рисунке 1.4 такое уравнение построено с помощью пакета STATISTICA [6, 17, 26].

Таблица 1.2 – Исходные данные

x	2	4,3	6	9,1	12	13	14
y	2	5	3	6,5	5	8	12,5

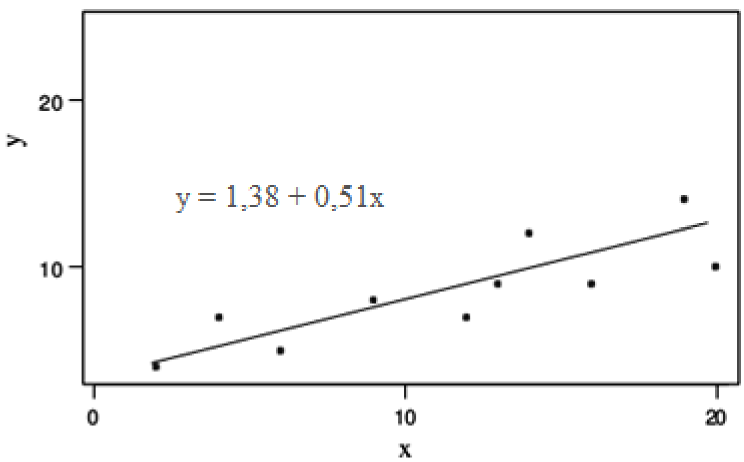


Рисунок 1.4 – Уравнение регрессии, построенное по исходным данным

Если имеется n наблюдений, то уравнение (1.3) можно представить в следующем виде [4, 42]:

$$y_i = a_0 + a_1 x_i + \varepsilon_i,\ i = \overline{1,n}. \tag{1.4}$$

Случайное слагаемое можно рассматривать как последовательность n случайных величин $\varepsilon_i,\ i = \overline{1,n}$.

Метод наименьших квадратов позволяет получить такие оценки $\tilde{a}_0, \tilde{a}_1$ параметров a_0, a_1, при которых сумма квадратов отклонений ε фактических значений признака y_i от расчетных (теоретических) $\tilde{y}_{1i}$ является минимальной [4, 95]:

$$Q(\tilde{a}_0, \tilde{a}_1) = \sum_{i=1}^{n} (y_i - \tilde{y}_{1i})^2 = \sum_{i=1}^{n} (y_i - \tilde{a}_0 - \tilde{a}_1 x_i)^2 = \sum_{i=1}^{n} \varepsilon_i^2 .$$

Минимум функции $Q(\tilde{a}_0, \tilde{a}_1)$ находится из условия равенства нулю производных по каждой переменной $\tilde{a}_0, \tilde{a}_1$ [4, 95]:

$$\frac{dQ}{d\tilde{a}_0} = -2\sum_{i=1}^{n} y_i + 2n\tilde{a}_0 + 2\tilde{a}_1 \sum_{i=1}^{n} x_i = 0,$$
$$\frac{dQ}{d\tilde{a}_1} = -2\sum_{i=1}^{n} y_i x_i + 2\tilde{a}_0 \sum_{i=1}^{n} x_i + 2\tilde{a}_1 \sum_{i=1}^{n} x_i^2 = 0.$$

После преобразований получаем систему уравнений

$$\begin{cases} n\tilde{a}_0 + \tilde{a}_1 \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} y_i, \\ \tilde{a}_0 \sum\limits_{i=1}^{n} x_i + \tilde{a}_1 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} y_i x_i. \end{cases} \tag{1.5}$$

Решением системы (1.5) являются оценки [42]

$$\begin{cases} \tilde{a}_0 = \bar{y} - \tilde{a}_1 \bar{x}, \\ \tilde{a}_1 = \dfrac{\operatorname{cov}(x, y)}{\sigma_x^2}, \end{cases} \tag{1.6}$$

где $\bar{x} = \sum\limits_{i=1}^{n} x_i / n$, $\bar{y} = \sum\limits_{i=1}^{n} y_i / n$, $\operatorname{cov}(x, y) = \overline{yx} - \bar{y}\bar{x}$;

$\overline{yx} = \sum\limits_{i=1}^{n} y_i x_i$, $\sigma_x^2 = \overline{x^2} - \bar{x}^2$, $\overline{x^2} = \sum\limits_{i=1}^{n} x_i^2 / n$.

Коэффициент $\tilde{a}_1$ при x называется выборочным коэффициентом регрессии. Он показывает среднее изменение зависимой переменной y при изменении независимой переменной x на единицу. Коэффициент $\tilde{a}_0$ указывает на значение результирующего признака при нулевом значении независимой переменной. Это важный индикатор для выбора вида уравнения регрессии. Например, если в результате вычислений коэффициент $\tilde{a}_0$ оказался отрицательным, а смысл задачи диктует положительность или равенство нулю показателя $\tilde{a}_0$, значит, выбор вида уравнения был неудачен [4, 42, 95].

Пример 1.1. Произведем расчеты для данных из таблицы 1.2.

Систему уравнений (1.5) запишем в виде

$$\begin{cases} 10\tilde{a}_0 + 115{,}3\tilde{a}_1 = 73, \\ 115{,}3\tilde{a}_0 + 1661{,}95\tilde{a}_1 = 1012{,}45. \end{cases}$$

Решив ее, получим следующие оценки параметров:

$$\begin{cases} \tilde{a}_0 = 1{,}38, \\ \tilde{a}_1 = 0{,}51. \end{cases}$$

Построим таблицу (табл. 1.3), содержащую исходные данные, расчетные значения $\tilde{y}_{1i} = 1{,}38 + 0{,}51x_i$ и остатки $\tilde{\varepsilon}_i = y_i - \tilde{y}_{1i} = y_i - 1{,}38 - 0{,}51x_i$.

Таблица 1.3 – Результаты регрессионного анализа

x	2	4,3	6	9,1	12	13	14
y	2	5	3	6,5	5	8	12,5
y_1	2,4	3,6	4,5	6,1	7,5	8,1	8,6
ε	-0,4	1,4	-1,5	0,4	-2,5	-0,1	3,9

Метод наименьших квадратов предполагает ряд ограничений на поведение случайного слагаемого ε – условия Гаусса-Маркова [4, 95]:

1) нулевое математическое ожидание, $M(\varepsilon_i) = 0, i = \overline{1,n}$;

2) равные дисперсии ошибок для всех наблюдений, $D(\varepsilon_i) = \sigma^2, i = \overline{1,n}$;

3) ошибки модели ε_i при разных наблюдениях независимы;

4) для всех $i = \overline{1,n}$ ошибки ε_i являются нормально распределенными случайными величинами.

Одним из показателей качества построенного уравнения регрессии является коэффициент детерминации R^2 [41, 95]. По определению

$$R^2 = \frac{D(y_1)}{D(y)}. \tag{1.7}$$

В свою очередь, $D(y) = D(y_1) + D(\varepsilon)$. Отсюда

$$R^2 = 1 - \frac{D(\varepsilon)}{D(y)}. \tag{1.8}$$

Таким образом, коэффициент детерминации можно интерпретировать как часть общей дисперсии y, которая объяснена с помощью расчетной переменной y_1 уравнения регрессии. Максимальное значение коэффициента детерминации R^2 равно 1. Это произойдет тогда, когда все остатки $\varepsilon_i = 0$, а уравнение прямой точно ляжет на экспериментальные точки y_i. Тем самым, при построении регрессии коэффициент детерминации желательно максимизировать. Именно это и делается при применении метода наименьших квадратов, так как из формулы (1.8) вытекает, что максимум R^2 достигается при минимуме $D(\varepsilon)$ [4, 42, 95].

Пример 1.2. Проведем дисперсионный анализ уравнения регрессии, построенного в примере 1.1. По данным таблицы 1.3, используя возможности пакета STATISTICA [17, 23, 92], получаем следующее (табл. 1.4).

Таблица 1.4 – Определение качества уравнения пакете STATISTICA

	df	*Sums of Squares*	*Mean Squares*	*F*	*p-level*
Regression	1	91,6853609	91,6853609	15,27798	0,004489228
Residual	8	45,9146391	5,739329887		
Total	9	131,6			

Итак, $D(\tilde{y}_1) = 91{,}6853609, D(\varepsilon) = 45{,}9146391, D(y) = 131{,}6$. Отсюда $\tilde{R}^2 = 0{,}7$.

Значение $\tilde{R}^2$ близко к единице. Это указывает на хорошее описание результирующей переменной y полученным уравнением регрессии.

1.2.4. Множественный линейный регрессионный анализ

Для построения модели с несколькими независимыми переменными, определяя при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель, используется множественная регрессия [83]. Она представляет собой регрессию результативного признака с двумя или большим числом независимых переменных вида

$$y = f\left(x^{(1)},\ldots,x^{(m)}\right) + \varepsilon, \tag{1.9}$$

где ε – случайная величина, характеризующая отклонения результативного признака от теоретического, найденного по уравнению регрессии [31].

В рамках множественного регрессионного анализа необходимо решать следующие задачи [4, 83, 31]:

– определение вида регрессии;

– отбор независимых переменных $x^{(1)},\ldots,x^{(m)}$ существенно влияющих на переменную y, при наличии возможностей внутренней взаимосвязи между $x^{(1)},\ldots,x^{(m)}$, то есть формирование множества значимых независимых переменных для y;

– оценка параметров регрессионной модели;

– оценка адекватности построенной регрессионной модели.

Пусть рассматривается множественная линейная регрессия, для которой произведен отбор значимых независимых переменных. Обозначим i-е наблюдение зависимой переменной через y_i, а i-е наблюдение независимых переменных через $x_i^{(1)},\ldots,x_i^{(m)}$. Тогда модель множественной линейной регрессии можно представить в виде [4, 22]:

$$y_i = a_0 + a_1 x_i^{(1)} + \ldots + a_m x_i^{(m)} + \varepsilon_i,\ i = \overline{1,n}, \tag{1.10}$$

где ε_i – случайные ошибки.

Для поиска неизвестных параметров модели $a_0, a_1, ..., a_m$ целесообразно перейти к матричным обозначениям для облегчения расчетных процедур [22].

Пусть $Y = (y_1, ..., y_n)'$ – вектор значений зависимой переменной размера n; $X = \begin{pmatrix} 1 & x_1^{(1)} & ... & x_1^{(m)} \\ 1 & x_2^{(1)} & ... & x_2^{(m)} \\ ... & ... & ... & ... \\ 1 & x_n^{(1)} & ... & x_n^{(m)} \end{pmatrix}$ – матрица значений независимых переменных размера $n \times (m+1)$; $a = (a_0, a_1, ..., a_m)'$ – вектор параметров размера $m+1$; $\varepsilon = (\varepsilon_1, ..., \varepsilon_n)'$ – вектор случайных ошибок модели.

Тогда в матричной форме модель (1.10) примет вид [4, 31]:

$$Y = Xa + \varepsilon. \tag{1.11}$$

Оценкой этой модели по выборке является уравнение

$$\tilde{Y} = X\tilde{a}, \tag{1.12}$$

где $\tilde{Y} = (\tilde{y}_1, ..., \tilde{y}_n)'$ – вектор оценок значений зависимой переменной, $\tilde{a} = (\tilde{a}_0, \tilde{a}_1, ..., \tilde{a}_m)'$ – вектор оценок параметров модели [4, 31].

Для оценки неизвестных параметров $a_0, a_1, ..., a_m$ используется метод наименьших квадратов. Тогда условие минимизации остаточной суммы квадратов запишется в виде [42]:

$$S = \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 = (Y - X\tilde{a})'(Y - X\tilde{a}) \to \min. \tag{1.13}$$

После раскрытия скобок получим [42]:

$$S = Y'Y - \tilde{a}'X'Y - Y'X\tilde{a} + \tilde{a}'X'X\tilde{a}.$$

На основании необходимого условия экстремума функции нескольких переменных $S = (\tilde{a}_0, \tilde{a}_1, ..., \tilde{a}_m)$, представляющей (1.13), необходимо приравнять к нулю вектор частных производных [22]:

$$\frac{\partial S}{\partial \tilde{a}} = \left(\frac{\partial S}{\partial \tilde{a}_0}, \frac{\partial S}{\partial \tilde{a}_1}, ..., \frac{\partial S}{\partial \tilde{a}_m} \right),$$

$$\frac{\partial S}{\partial \tilde{a}} = -2X'Y + 2X'X\tilde{a} = 0,$$

откуда получаем уравнение в матричной форме для определения вектора $\tilde{a}$ [22]:

$$X'X\tilde{a} = X'Y. \qquad (1.14)$$

Решением уравнения (1.14) является вектор оценки неизвестных параметров [4, 22, 42]:

$$\tilde{a} = \left(X'X\right)^{-1} X'Y. \qquad (1.15)$$

Для определения тесноты связи между переменной y и совокупностью объясняющих переменных $x_i^{(1)},...,x_i^{(m)}$ в случае линейной зависимости применяется коэффициент множественной корреляции R [4, 42].

Этот коэффициент изменяется в интервале от 0 до 1, причем, в отличие от коэффициентов парной корреляции, он берется всегда по абсолютной величине. Если $R = 0$, то линейной корреляционной связи между y и $x_i^{(1)},...,x_i^{(m)}$ нет. Если $R = 1$, то связь функциональная. Выражение, по которому вычисляют коэффициент множественной корреляции, имеет вид

$$R = \sqrt{\frac{\tilde{a}_1 r_{x^1 y}\sigma_{x^1} + ... + \tilde{a}_m r_{x^m y}\sigma_{x^m}}{\sigma_y}}. \qquad (1.16)$$

где $\tilde{a}_0, \tilde{a}_1, ..., \tilde{a}_m$ – коэффициенты регрессии уравнения (1.12), $r_{x^1 y},...,r_{x^m y}$ – парные коэффициенты корреляции, $\sigma_{x^1},...,\sigma_{x^m}$ – средние квадратические отклонения факторов, σ_y – среднее квадратическое отклонение y [4, 42].

При этом обычно интерпретируется не сам коэффициент корреляции R, а его квадрат R^2, который называется коэффициентом множественной детерминации. Он показывает, какая часть общей дисперсии объясняется за счет вариации линейной комбинации аргументов $x_i^{(1)},...,x_i^{(m)}$ при данных значениях коэффициентов регрессии [31, 41].

Существенность отличия от нуля выборочного коэффициента множественной корреляции проверяется на основе F-критерия (критерия

Фишера). Для каждого регрессионного уравнения вычисляется величина [41, 42]

$$F = R^2(n-m-1)/\left(1-R^2\right)m, \tag{1.17}$$

где R – множественный коэффициент корреляции, m – число независимых переменных, n – число наблюдений.

Найденное значение критерия F сравнивается с табличным $F^{табл}$ при числе степеней свободы m и $n-m-1$ и заданном уровне значимости α. Если расчетное значение превышает табличное, т. е. $F > F^{табл}$, то гипотеза о равенстве нулю коэффициента множественной корреляции отвергается, и связь считается существенной [41, 42].

1.2.5. Стандартизированное уравнение линейной регрессии

В линейной множественной регрессионной модели (1.10) численные значения коэффициентов регрессии зависят от выбранных единиц измерения каждой независимой переменной. Чтобы коэффициенты стали сравнимы, необходимо привести их к стандартизированному масштабу [41, 94].

Для этого все переменные выражают в безразмерных единицах измерения на основе следующих соотношений:

$$t_y = (y-\bar{y})/\sigma_y,$$

$$t_{x^j} = (x^{(j)}-\bar{x}^{(j)})/\sigma_{x^j}, \ j=\overline{1,m},$$

где $x^{(1)},\ldots,x^{(m)}$ и y – значения переменных в исходном масштабе, $\bar{x}^{(1)},\ldots,\bar{x}^{(m)}$ и $\bar{y}$ – их средние значения, $t_{x^1},\ldots,t_{x^m}$ и t_y – соответствующие значения переменных в стандартизированном масштабе [41].

Тогда уравнение (1.10) можно записать в виде:

$$t_y = \beta_1 t_{x^1} + \ldots + \beta_m t_{x^m}. \tag{1.18}$$

Все переменные уравнения выражены в сравнимых единицах измерения. Коэффициенты $\beta_1,\ldots,\beta_m$ называются коэффициентами регрессии в стандартизированном масштабе. Они определяются по формуле:

$$\beta_j = a_j \sigma_{x^j} / \sigma_y , \tag{1.19}$$

где σ_{x^j} и σ_y – средние квадратические отклонения [41].

Коэффициенты регрессии $\beta_1,...,\beta_m$ показывают влияние изменения каждой независимой переменной на изменение переменной y. Чем больше $|\beta_j|$, тем сильнее влияет соответствующая независимая переменная на зависимую величину [41].

1.3. Существующие подходы к восстановлению закономерностей на основе нечеткого регрессионного моделирования

Регрессионный анализ может быть использован для моделирования взаимодействий между зависимыми и независимыми величинами в нечеткой среде. В традиционном регрессионном анализе остатки считаются следствием случайных ошибок, поэтому для получения оценки параметров модели в нем применяются статистические методы. Погрешность вычисления возникает иногда ввиду неопределенности структуры модели или по причине неточности данных наблюдения. Неопределенность в данном случае становится нечеткостью, а не случайностью. С тех пор, как Л. А. Заде (1965) предложил теорию нечетких множеств, понятие нечеткости получило больше внимания, а анализ нечеткой информации стал особенно важен [107].

Для того чтобы применить нечеткость к регрессионному анализу, H. Tanaka, H. Ishibuchi, S. Yoshikawa (1982) первыми предложили провести исследование нечеткой линейной регрессионной модели с коэффициентами в виде симметричных треугольных нечетких чисел [106]. В качестве параметров оценки таких моделей выступали точность модели и степень ее неопределенности. Проблемы оценки были сведены к задаче линейного программирования, в которой данные факторы выступали в виде критериев. В дальнейшем P. Diamond (1988) разработал нечеткий анализ методом наименьших квадратов как нечеткое расширение обычного метода наименьших квадратов путем введения нового понятия расстояния на множестве нечетких чисел. В общем случае все нечеткие регрессионные методы могут быть разделены на две группы: первая базируется на методе линейного программирования Танака [103-106], вторая – на методе наименьших квадратов Диамонда [99, 100, 110]. Преимущество первого метода состоит в простоте программной реализации и вычислений, в то время как второй минимизирует степень нечеткости между экспериментальными и прогнозируемыми значениями на основе метода наименьших квадратов.

Основы построения статистических моделей с нечеткими параметрами были предложены в работе Р.А. Алиева в 1991 году [5]. Он рассматривал математическую модель, представленную в виде нечеткого уравнения множественной регрессии:

$$Y = A_0 + A_1 x_1 + \ldots + A_k x_k, \tag{1.20}$$

в котором $x_i, i = \overline{1,k}$ – детерминированные значения, а Y – нечеткие.

Для оценки нечетких параметров построенной модели был использован критерий минимизации отклонения нечетких выходных значений $\tilde{Y}_j$, полученных по (1.20), от соответствующих выборочных значений Y_j по данным наблюдений $j = \overline{1,n}$:

$$J = \bigcup_{j=\overline{1,n}} \left(Y_j |-| \tilde{Y}_j\right)^2 \to \min .$$

Здесь $|-|$ – ограниченная разность нечетких чисел, определяемая по формуле

$$\mu_J(x) = \mu_{Y_j |-| \tilde{Y}_j}(x) = \max\left(0, \mu_{Y_j}(x) - \mu_{\tilde{Y}_j}(x)\right).$$

Иными словами, задача оценивания параметров уравнения регрессии (1.20) была сведена к минимизации многомерной функции с нечеткими переменными.

Предположив, что нечеткие коэффициенты $A_0, A_1, \ldots, A_k$ являются нормальными нечеткими множествами на $\Re$

$$A_i = \bigcup_{a_i \in \Re} \mu_{A_i}(a_i)/a_i,$$

Р.А. Алиев определил α-уровневые множества нечетких коэффициентов $\tilde{A}_i$:

$$a_i^{\alpha} = \{a_i : a_i \in \Re, \mu_{A_i}(a_i) \geq \alpha\}, i = \overline{0,k}, \alpha \in [0,1].$$

Для оценивания нечетких коэффициентов $A_0, A_1, \ldots, A_k$ он определял такие коэффициенты $a_0^{\alpha_s}, a_1^{\alpha_s}, \ldots, a_k^{\alpha_s}, s = \overline{1,p}$ на каждом уровне $\alpha_s, \left(\alpha : \{\alpha_0 = 0, \alpha_1, \ldots, \alpha_s, \ldots, \alpha_p = 1\}\right)$, которые удовлетворяют условию

$$J_s = \sum_{j=1}^{n}\left(y_i^{\alpha_s} - \tilde{y}_i^{\alpha_s}\right)^2 \to \min, s = \overline{1,p},$$

где

$$\tilde{y}_i^{\alpha_s} = a_0^{\alpha_s} + a_1^{\alpha_s} x_1 + \ldots + a_k^{\alpha_s} x_k.$$

Таким образом, исходная задача оценивания нечетких коэффициентов нечеткого уравнения регрессии (1.20) была сведена к классическим задачам оценивания параметров множественной регрессии [5].

Большинство современных нечетких регрессионных моделей содержат нечеткую зависимую переменную, нечеткие параметры и нечеткие независимые переменные, имеют линейную множественную структуру. В 1992 году M. Sakawa и H. Yano предложили к рассмотрению оценку нечетких параметров нечеткой линейной регрессионной модели:

$$Y_j = A_0 + A_1 X_{j1} + \ldots + A_k X_{jk}, j = 1, \ldots, n, \tag{1.21}$$

где входные данные $X_{j1}, \ldots X_{jk}$ и выходы модели Y_j являются нечеткими. Они сформулировали многокритериальные методы программирования для оценки модели наряду с подходом, основанным на линейном программировании [103].

M. Sakawa и H. Yano использовали следующие три равенства для определения взаимодействия между двумя нечеткими числами M и N:

$$Pos(M = N) = \sup_x \min\{\mu_M(x), \mu_N(x)\},$$

$$Nes(M \subset N) = \inf_x \max\{1 - \mu_M(x), \mu_N(x)\},$$

$$Nes(M \supset N) = \inf_x \max\{\mu_M(x), 1 - \mu_N(x)\},$$

где $\mu_M(x)$ и $\mu_N(x)$ – это функции принадлежности нечетких чисел M и N, Pos и Nes – сокращения от вероятности (Possibility) и необходимости (Necessity). Далее были сформулированы три типа многоцелевых задач программирования для оценки параметров нечетких линейных регрессионных моделей, а также подход, основанный на методе наименьших квадратов. Многокритериальный анализ нечетких линейных регрессионных моделей обеспечивает хорошую оценку параметров при использовании неопределенности модели (через отношения включения). Подход, основанный на методе наименьших квадратов, напрямую обращается к информации, заключенной во входных/выходных данных, и рассматривает меру наилучшего соответствия в условиях нечеткости.

В статье Miin-Shen Yang и Tzu-Shun Lin (2002) предлагается два типа нечеткого метода наименьших квадратов для оценки параметров нечеткой линейной регрессионной модели (1.21): интервальный метод и метод, основанный на аппроксимации по расстоянию. Они представляются неплохой альтернативой к многоцелевому программированию [109].

В ходе нечеткого метода наименьших квадратов, основанного на аппроксимации по расстоянию, рассматривается нечеткая линейная регрессионная модель (1.21), в которой $Y_j = \left(m_{y_j}, \alpha_{y_j}, \beta_{y_j}\right)_{LR}$ – выходы, $X_{ji} = \left(m_{x_{ji}}, \alpha_{x_{ji}}, \beta_{x_{ji}}\right)_{LR}$ – входы и $A_i = \left(m_{a_i}, \alpha_{a_i}, \beta_{a_i}\right)_{LR}$ – параметры модели, $i = 1,...,k$, $j = 1,...,n$, так что $M = (m, \alpha, \beta)_{LR}$ – нечеткое число L-R-типа [67] с функцией принадлежности

$$\mu_M(x) = \begin{cases} L\left(\dfrac{m-x}{\alpha}\right), x \le m \\ R\left(\dfrac{x-m}{\beta}\right), x \ge m \end{cases} (\alpha > 0, \beta > 0).$$

В $M = (m, \alpha, \beta)_{LR}$ m называется средним значением (модой), а α и β – левым и правым коэффициентами нечеткости соответственно. Если $L(x) = R(x) = 1 - x$, то $M=(m,\alpha,\beta)_{LR}$ называется треугольным нечетким числом и обозначается как $M=(m,\alpha,\beta)_{T}$ [67].

Особенностью рассматриваемой модели (1.21) является то, что результат операции A_iX_{ji} может не относиться ко множеству нечетких чисел *L*-*R*-типа.

В работах Дюбуа и Праде были предложены следующие формулы аппроксимации для $M=(m,\alpha,\beta)_{LR}$ и $N=(n,\gamma,\delta)_{LR}$ [101]:

(1) если $M>0$ и $N>0$, то

$$(m,\alpha,\beta)_{LR}(n,\gamma,\delta)_{LR}\approx(mn,m\gamma+n\alpha,m\delta+n\beta)_{LR};$$

(2) если $M<0$ и $N>0$, то

$$(m,\alpha,\beta)_{LR}(n,\gamma,\delta)_{LR}\approx(mn,n\alpha-m\delta,n\beta-m\gamma)_{LR};$$

(3) если $M<0$ и $N<0$, то

$$(m,\alpha,\beta)_{LR}(n,\gamma,\delta)_{LR}\approx(mn,-n\beta-m\delta,-n\alpha-m\gamma)_{RL}.$$

В результате была получена приближенная оценка

$$A_0+A_1X_{j1}+\ldots+A_kX_{jk}\approx\left(\tilde{m}_j,\tilde{\alpha}_j,\tilde{\beta}_j\right)_{LR},$$

где

$$\tilde{m}_j=m_{a_0}+\sum_{p=1}^{k}\left(m_{a_p}m_{x_{jp}}\right),$$

$$\tilde{\alpha}_j=\alpha_{a_0}+\sum_{A_p\in H_1}\left[s_{jp}\left(m_{a_p}\alpha_{x_{jp}}+m_{x_{jp}}\alpha_{a_p}\right)+\left(1-s_{jp}\right)\left(m_{a_p}\alpha_{x_{jp}}+m_{x_{jp}}\beta_{a_p}\right)\right]+$$
$$\sum_{A_p\in H_2}\left[s_{jp}\left(m_{x_{jp}}\alpha_{a_p}-m_{a_p}\beta_{x_{jp}}\right)+\left(1-s_{jp}\right)\left(-m_{a_p}\beta_{x_{jp}}-m_{x_{jp}}\beta_{a_p}\right)\right],$$

$$\tilde{\beta}_j=\beta_{a_0}+\sum_{A_p\in H_1}\left[s_{jp}\left(m_{a_p}\beta_{x_{jp}}+m_{x_{jp}}\beta_{a_p}\right)+\left(1-s_{jp}\right)\left(m_{a_p}\beta_{x_{jp}}+m_{x_{jp}}\alpha_{a_p}\right)\right]+$$
$$\sum_{A_p\in H_2}\left[s_{jp}\left(m_{x_{jp}}\beta_{a_p}-m_{a_p}\alpha_{x_{jp}}\right)+\left(1-s_{jp}\right)\left(-m_{a_p}\alpha_{x_{jp}}-m_{x_{jp}}\alpha_{a_p}\right)\right].$$

Так как $A_0+A_1X_{j1}+\ldots+A_kX_{jk}$ аппроксимируется нечетким числом *L*-*R*-типа, то M.S. Yang и C.H. Ko предложили вычисление расстояния d_{LR}^2 между двумя нечеткими числами *L*-*R*-типа по формуле [110]:

$$d_{LR}^2(X,Y)=\left(m_x-m_y\right)^2+\left[\left(m_x-l\alpha_x\right)-\left(m_y-l\alpha_y\right)\right]^2+\left[\left(m_x-r\beta_x\right)-\left(m_y+l\beta_y\right)\right]^2,$$

где $X=\left(m_x,\alpha_x,\beta_x\right)_{LR}$, $Y=\left(m_y,\alpha_y,\beta_y\right)_{LR}$ и $l=\int_0^1 L^{-1}(\omega)d\omega$, $r=\int_0^1 R^{-1}(\omega)d\omega$.

Таким образом, была получена следующая целевая функция:

$$J\left(A_0,A_1,\ldots,A_k\right)=\sum_{j=1}^{n} d_{LR}^2\left(Y_j,A_0+A_1X_{j1}+\ldots+A_kX_{jk}\right)=$$
$$=\sum_{j=1}^{n-1}\left[\left(m_{y_j}-\tilde{m}_j\right)^2+\left[\left(m_{y_j}-l\alpha_{y_j}\right)-\left(\tilde{m}_j-l\alpha_j\right)\right]^2+\left[\left(m_{y_j}-l\beta_{y_j}\right)-\left(\tilde{m}_j+l\beta_j\right)\right]^2\right].$$

Минимизация $J\left(A_0,A_1,\ldots,A_k\right)$ по A_i при $\alpha_{a_i}\geq 0$ и $\beta_{a_i}\geq 0$, $i=0,1,\ldots,k$ получила название нечеткого метода наименьших квадратов, основанного на аппроксимации по расстоянию [101, 109, 110].

Недостатком такого метода явилось то, что если коэффициенты нечеткости α и β достаточно велики по отношению к моде нечеткого числа, то снижается качество аппроксимации. В связи с этим был разработан еще один нечеткий метод наименьших квадратов.

Так как сумма $A_0+A_1X_{j1}+\ldots+A_kX_{jk}$ может не относиться к *L*-*R*-типу, M. Sakawa и H. Yano представили ее ω-уровневое множество с закрытым интервалом [103]:

$$\left(A_0+A_1X_{j1}+\ldots+A_kX_{jk}\right)_\omega=\left(\tilde{A}\otimes\tilde{X}_j\right)_\omega=\left[\left(\tilde{A}\otimes\tilde{X}_j\right)_\omega^L,\left(\tilde{A}\otimes\tilde{X}_j\right)_\omega^R\right],$$

$$\left(\tilde{A}\otimes\tilde{X}_j\right)_\omega^L=\sum_{i=0}^{k}\left\{\min\left\{\begin{matrix}\left(m_{a_i}-L_{a_i}^{-1}(\omega)\alpha_{a_i}\right)\left(m_{x_{ji}}-L_{x_{ji}}^{-1}(\omega)\alpha_{x_{ji}}\right),\\ \left(m_{a_i}-L_{a_i}^{-1}(\omega)\alpha_{a_i}\right)\left(m_{x_{ji}}+R_{x_{ji}}^{-1}(\omega)\beta_{x_{ji}}\right)\end{matrix}\right\}\right\},$$

$$\left(\tilde{A}\otimes\tilde{X}_j\right)_\omega^R=\sum_{i=0}^{k}\left\{\max\left\{\begin{matrix}\left(m_{a_i}+R_{a_i}^{-1}(\omega)\beta_{a_i}\right)\left(m_{x_{ji}}+R_{x_{ji}}^{-1}(\omega)\beta_{x_{ji}}\right),\\ \left(m_{a_i}+R_{a_i}^{-1}(\omega)\beta_{a_i}\right)\left(m_{x_{ji}}-L_{x_{ji}}^{-1}(\omega)\alpha_{x_{ji}}\right)\end{matrix}\right\}\right\}.$$

Здесь через $(M)_\omega$ обозначено ω-уровневое множество нечеткого числа M. Нечеткий выход Y_j имеет свое ω-уровневое множество $\left(Y_j\right)_\omega=\left[Y_{j\omega}^L,Y_{j\omega}^R\right]$, где

$$Y_{j\omega}^L=m_{y_j}-L_{y_j}^{-1}(\omega)\alpha_{y_j},\ Y_{j\omega}^R=m_{y_j}-R_{y_j}^{-1}(\omega)\beta_{y_j}.$$

Для определения функции цели нечеткого метода наименьших квадратов было введено расстояние D^2 между Y_j и $A_0+A_1X_{j1}+\ldots+A_kX_{jk}$:

$$D^2\left(Y_j, A_0 + A_1 X_{j1} + \ldots + A_k X_{jk}\right) = \int_0^1 \left(\left(Y_{j\omega}^L - \left(\tilde{A} \otimes \tilde{X}_j \right)_\omega^L \right)^2 + \left(Y_{j\omega}^R - \left(\tilde{A} \otimes \tilde{X}_j \right)_\omega^R \right)^2 \right) d\omega.$$

Таким образом, функция цели ρ была найдена по формуле:

$$\rho\left(A_0, A_1, \ldots, A_k\right) = \sum_{j=1}^{n} D^2\left(Y_j, A_0 + A_1 X_{j1} + \ldots + A_k X_{jk}\right) =$$
$$= \sum_{j=1}^{n} \int_0^1 \left(\left(Y_{j\omega}^L - \left(\tilde{A} \otimes \tilde{X}_j \right)_\omega^L \right)^2 + \left(Y_{j\omega}^R - \left(\tilde{A} \otimes \tilde{X}_j \right)_\omega^R \right)^2 \right) d\omega.$$

При рассмотрении операторов min и max в $\left(\tilde{A} \otimes \tilde{X}_j\right)_\omega$ параметры $A_0, A_1, \ldots, A_k$ разбиваются на следующие три группы:

$$P_1 = \left\{ A_i \mid m_{a_i} - L_{a_i}^{-1}(\omega)\alpha_{a_i} \geq 0 \right\},$$

$$P_2 = \left\{ A_i \mid -R_{a_i}^{-1}(\omega)\beta_{a_i} \leq m_{a_i} < -L_{a_i}^{-1}(\omega)\alpha_{a_i} \right\},$$

$$P_3 = \left\{ A_i \mid m_{a_i} < -R_{a_i}^{-1}(\omega)\beta_{a_i} \right\}.$$

Тогда $\left(\tilde{A} \otimes \tilde{X}_j\right)_\omega^L$ и $\left(\tilde{A} \otimes \tilde{X}_j\right)_\omega^R$ в функции цели $\rho\left(A_0, A_1, \ldots, A_k\right)$ находятся по формулам:

$$\left(\tilde{A} \otimes \tilde{X}_j\right)_\omega^L = \sum_{A_i \in P_1} \left(m_{a_i} - L_{a_i}^{-1}(\omega)\alpha_{a_i}\right)\left(m_{x_{ji}} - L_{x_{ji}}^{-1}(\omega)\alpha_{x_{ji}}\right) +$$
$$+ \sum_{A_i \in P_2 \cup P_3} \left(m_{a_i} - L_{a_i}^{-1}(\omega)\alpha_{a_i}\right)\left(m_{x_{ji}} + R_{x_{ji}}^{-1}(\omega)\beta_{x_{ji}}\right),$$

$$\left(\tilde{A} \otimes \tilde{X}_j\right)_\omega^R = \sum_{A_i \in P_1 \cup P_2} \left(m_{a_i} + R_{a_i}^{-1}(\omega)\beta_{a_i}\right)\left(m_{x_{ji}} + R_{x_{ji}}^{-1}(\omega)\beta_{x_{ji}}\right) +$$
$$+ \sum_{A_i \in P_3} \left(m_{a_i} + R_{a_i}^{-1}(\omega)\beta_{a_i}\right)\left(m_{x_{ji}} - L_{x_{ji}}^{-1}(\omega)\alpha_{x_{ji}}\right).$$

В работе [109] было показано, что функция цели $\rho\left(A_0, A_1, \ldots, A_k\right)$ может быть аналогично расширена до функции устойчивого типа

$$\rho^0\left(\mu, A_{0i}, A_{1i}, \ldots, A_{ki}\right) = \sum_{i=1}^{c+1} \sum_{j=1}^{n} \mu_{ij}^m d_{ij}^2, \tag{1.22}$$

где

$$d_{ij}^2 = \begin{cases} D^2\left(Y_j, A_{0i} + A_{1i} X_{j1} + \ldots + A_{ki} X_{jk}\right), i = 1, \ldots, c,\ j = 1, \ldots, n, \\ \delta^2, i = c+1,\ j = 1, \ldots, n. \end{cases}$$

$$\delta^2 = \lambda \left(\frac{\sum_{i=1}^{c}\sum_{j=1}^{n} d_{ij}^2}{nc} \right), \ \lambda > 0 \ \text{– константа.}$$

Условием минимизации ρ^0 по μ является

$$\mu_{ij} = \left(\sum_{p=1}^{c+1} \frac{\left(d_{ij}^2\right)^{1/(m-1)}}{\left(d_{pj}^2\right)^{1/(m-1)}} \right)^{-1}, i = 1,...,c+1, \ j = 1,...,n. \tag{1.23}$$

Данный алгоритм (1.22) и (1.23) получил название интервального нечеткого метода наименьших квадратов.

В статье [102] предложена процедура нечеткой гребневой регрессии, которая представляет собой незначительную модификацию метода наименьших квадратов и в качестве функции цели использует функцию

$$\min \gamma \left\| \tilde{A} \right\|^2 + \sum_{j=1}^{n} D^2\left(\tilde{A}_0 + \tilde{A}_1 x_{j1} + \ldots + \tilde{A}_k x_{jk}, \tilde{Y}_j\right), \tag{1.24}$$

где γ − фиксированная положительная константа, входные переменные являются действительными числами, а коэффициенты и результирующая переменная − нечеткими числами. Оценки параметров модели находятся при этом из условия равенства нулю частных производных функции.

Наряду с этим в [82, 102] рассматривается двойственная модель нечеткой гребневой регрессии:

$$\min \gamma \left\| \tilde{A} \right\|^2 + \sum_{i=1}^{m} \xi_{1i}^2 + \frac{1}{2} \sum_{i=1}^{m} \xi_{2i}^2 ,$$

где $\left\| \tilde{A} \right\|^2 = \left\| a \right\|^2 + \frac{1}{2} \left\| \sigma \right\|^2$ для вектора $\tilde{A} = \left(\tilde{A}_0, \tilde{A}_1, ..., \tilde{A}_n\right)$ оценочной модели

$$\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \ldots + \tilde{A}_n x_n, \ \tilde{A}_j = \left(a_j, \sigma_j\right), \ j = 0,...,n.$$

В качестве ограничений к задаче используются условия:

$$y_i - a \cdot x_i = \xi_{1i},$$

$$s_i - \sigma \cdot x_i = \xi_{2i}, \ i = 1,\ldots,m.$$

Затем вводятся множители Лагранжа α_{1i} и α_{2i}, $i=1,...,m$ и составляется функция Лагранжа в виде:

$$L=\gamma\left\|\tilde{A}\right\|^2+\sum_{i=1}^{m}\xi_{1i}^2+\frac{1}{2}\sum_{i=1}^{m}\xi_{2i}^2+\sum_{i=1}^{m}\alpha_{1i}\left(y_i-a\cdot x_i-\xi_{1i}\right)+\sum_{i=1}^{m}\alpha_{2i}\left(s_i-\sigma\cdot x_i-\xi_{2i}\right),$$

седловая точка которой находится из равенства нулю ее частных производных.

Прогноз $\tilde{Y}(y,s)$, полученный с помощью нечеткой гребневой регрессии, для неизвестного примера x имеет вид:

$$(a\cdot x,\sigma\cdot x)=\left(\frac{1}{2\gamma}\sum_{i=1}^{m}\alpha_{1i}(x_i\cdot x),\frac{1}{\gamma}\sum_{i=1}^{m}\alpha_{2i}(x_i\cdot x)\right)=\left(\frac{1}{2\gamma}(\alpha_1\cdot k),\frac{1}{\gamma}(\alpha_2\cdot k)\right)=$$
$$=\left(y^t(K+\gamma I)^{-1}k,s^t(K+\gamma I)^{-1}k\right),$$

где $k=\left(k_1,...,k_m\right)^t$ − вектор скалярного произведения:

$$k_i=x_i\cdot x,\ i=1,...,m.$$

Математико-статистические методы [57] используют линейную регрессионную модель для построения нечеткого группового метода обработки данных. Он заключается в следующем.

Для заданных наблюдаемых данных $\left(y_i,x_i\right)$ строится несколько частичных линейных моделей (они называются частичным представлением), которые упорядочиваются иерархически. В качестве частичного представления используется линейная интервальная модель:

$$Y=A_0+A_1x_p+A_2x_q+A_3x_p^2+A_4x_q^2+A_5x_px_q,$$

где $A_0,\ldots,A_5$ − интервальные коэффициенты, Y − интервальный выход, $p,q=1,...,N,\ p\neq q$. В качестве оценочного интервала Y берется треугольное нечеткое число с модой α, шириной c и функцией принадлежности $\mu_Y(y)$. Интервальные коэффициенты $A_0,\ldots,A_5$ получаются из решения проблемы линейного программирования [87].

Получив несколько частичных представлений, устанавливается критерий выбора: какое из частичных представлений дает наилучшую модель. В качестве такого критерия оценки определяется следующая функция:

$$J = \sum_{i=1}^{n} J_i^{(2)} / \sum_{i=1}^{n} J_i^{(1)}, \text{ где } J_i^{(1)} = \mu_{Y_i}(y_i),\ J_i^{(2)} = c|x_i|.$$

$J_i^{(1)}$ указывает меру близости наблюдаемых значений y_i к центру оценочного интервала Y_i. Это означает, что чем больше оценочное значение $J_i^{(1)}$, тем ближе наблюдаемое значение y_i к центру оценочного интервала Y_i. $J_i^{(2)}$ указывает ширину оценочного интервала Y_i. Следовательно, чем меньше оценочное значение $J_i^{(2)}$, тем меньше ширина и тем лучше оценка. Итак, оценочная модель с наименьшим оценочным значением J – это та модель, которая наиболее хорошо аппроксимирует заданные данные (y_i, x_i).

1.4. Цели и задачи исследования

В данной главе показано, что в настоящее время все более востребованным становится создание систем интеллектуального анализа данных, позволяющих выявить полезную скрытую информацию и восстановить закономерности, которые содержатся в исходных данных. При этом получение достоверных результатов и выводов на основе исследования осуществляется за счет применения различных подходов, основными из которых являются: классификация, поиск ассоциативных правил, кластеризация и др. Определение зависимости некоторого показателя от различных признаков и отображение их взаимосвязи в форме модели осуществляется в ходе проведения регрессионного анализа.

В современных вычислительных алгоритмах анализ данных зачастую характеризуется наличием слабоструктурированной, неполной, неточной и нечеткой информации, возникающей вследствие неопределенности, присущей моделям сложных процессов, а также необходимостью решения задач в тех областях, где существенная роль принадлежит суждениям и знаниям экспертов. Эта специфика значительно усложняет процесс построения моделей и обусловливает необходимость разработки специальных методов повышения их адекватности. С тех пор, как в работах L. Zadeh для формализации неопределенности стала использоваться теория нечетких множеств, интерес к анализу приближенной и, в частности, нечеткой информации значительно вырос.

Для определения зависимостей между признаками в нечеткой среде используется нечеткий регрессионный анализ, который позволяет решать различные задачи в ситуациях, когда традиционные методы неэффективны или совсем неприменимы из-за отсутствия достаточно точных сведений об исследуемых объектах. В общем случае все нечеткие регрессионные методы могут быть разделены на две группы: первая базируется на методе линейного программирования H. Tanaka, вторая – на методе наименьших квадратов P. Diamond.

Построение нечетких регрессионных моделей опирается на математический аппарат, включающий определение арифметических операций над нечеткими числами и их сравнение. Только для некоторых типов нечетких чисел результат арифметической операции представляет собой нечеткое число того же типа. В других случаях требуется дополнительная аппроксимация.

В связи с этим была поставлена цель диссертационной работы – усовершенствование существующих методов нечеткого регрессионного моделирования за счет применения нечетких чисел L-R-типа и использования моделей, в которых учитываются различные типы данных в выборке. Для достижения поставленной цели необходимо провести анализ свойств арифметических операций над нечеткими числами L-R-типа для выполнения вычислений в ходе нечеткого регрессионного моделирования. Также требуется разработать процедуры оценки параметров и проверки адекватности и точности нечетких регрессионных моделей. Одной из задач исследования является разработка и реализация алгоритма для решения задачи отбора приближенной информации, необходимой для анализа, на основе нейронных сетей. Необходимо создать программный комплекс с применением современных компьютерных технологий для анализа и интеллектуальной обработки данных на основе предложенных алгоритмов нейросетевого и нечеткого регрессионного моделирования.

Выводы по главе 1

1. В настоящее время все более востребованным становится создание систем интеллектуального анализа данных, позволяющих выявить полезную скрытую информацию и восстановить закономерности, которые содержатся в исходных данных.

2. Восстановление закономерностей в данных можно осуществлять различными методами интеллектуального анализа, одним из которых является регрессионное моделирование.

3. Регрессионное моделирование хорошо разработано теоретически, однако в нем отсутствует ориентация на приближенную информацию, поэтому необходимо модифицировать существующие подходы для анализа данных в нечеткой среде.

4. Построение нечетких регрессионных моделей опирается на математический аппарат, включающий определение арифметических операций над нечеткими числами и их сравнение. Только для некоторых типов нечетких чисел результат арифметической операции представляет собой нечеткое число того же типа. В других случаях требуется дополнительная аппроксимация.

5. Требуется усовершенствование существующих методов нечеткого регрессионного моделирования за счет применения нечетких чисел определенного типа и использования моделей, в которых учитываются различные типы данных в выборке.

6. Необходимо разработать процедуры проверки качества нечетких регрессионных моделей на основе оценки их адекватности и точности.

Глава 2. Алгебраические структуры на множествах нечетких чисел *L*-*R*-типа

В данной главе рассмотрены теоретические основы обработки нечеткой информации; приведено понятие нечеткого числа и его функции принадлежности, перечислен ряд требований, которые к ней предъявляются; рассмотрено понятие нечетких чисел *L*-*R*-типа и задание арифметических операций над ними на основе принципа обобщения; введено понятие группоида нечетких чисел *L*-*R*-типа, позволяющего в зависимости от выполняемых свойств (коммутативности, ассоциативности, дистрибутивности, наличия обратного и единичного элементов) определить тип алгебры с одной и двумя операциями; проведены доказательства арифметических операций над нечеткими числами *L*-*R*-типа, позволяющих осуществлять вычисления в ходе нечеткого регрессионного анализа.

2.1. Нечеткие множества и нечеткие числа

Пусть X есть множество, счетное или нет, и x – элемент из X. Нечетким подмножеством A множества X называется множество упорядоченных пар вида [41, 43, 67]:

$$A = \{(x / m_a(x))\}, \tag{2.1}$$

где для каждого элемента $x \in X$ степень его принадлежности множеству A задается с помощью функции принадлежности $m_a(x)$, при этом

$$m_a(x) \in [0,1].$$

Функция принадлежности ставит в соответствие каждому значению x заданной переменной некоторое число из интервала $[0,1]$

$$m_a(x) : X \to [0,1], \forall x \in X. \tag{2.2}$$

Определим ряд характеристик для нечетких множеств.

Высота нечеткого множества A определяется как максимальное из значений, принимаемых функцией принадлежности нечеткого множества на всей области определения X [51, 67]

$$h(A) = \sup_{x \in X} (m_a(x)). \tag{2.3}$$

Нечеткое множество A нормально, если его высота равна 1, иначе нечеткое множество называется субнормальным [67].

Нечеткое множество называется унимодальным, если его функция принадлежности является унимодальной, т. е.

$$\exists x_m \in X \; (m_a(x_m) = 1),$$

причем слева от x_m функция возрастает, а справа убывает, x_m называется модальным значением функции принадлежности $m_a(x)$ [51, 61].

Носитель нечеткого множества A представляет собой четкое подмножество множества X, содержащее все элементы, степени принадлежности которых множеству A отличны от нуля [71, 96]

$$S(A) = Supp(A) = \{x : m_a(x) > 0, x \in X\}. \tag{2.4}$$

Ядро нечеткого множества A – это четкое подмножество множества X, содержащее все элементы, принадлежащие множеству A, со степенью, равной 1 [3, 62]

$$C(A) = Core(A) = \{x : m_a(x) = 1, x \in X\}. \tag{2.5}$$

Пусть $\alpha \in (0,1]$. Для нечеткого множества A можно определить слабое уровневое множество

$$\sigma_\alpha(A) = \{x : m_a(x) \ge \alpha, x \in X\} \tag{2.6}$$

и сильное уровневое множество

$$\omega_\alpha(A) = \{x : m_a(x) > \alpha, x \in X\}. \tag{2.7}$$

Эти множества иначе называются подмножествами уровня α нечеткого множества A или α-срезами [36, 51].

Для описания объектов и явлений в условиях неопределенности широко используется понятие нечеткой переменной.

Нечеткая переменная задается тройкой $\langle \alpha, X, A \rangle$, где α – название переменной, X – область определения переменной α, A – нечеткое множество на X с функцией принадлежности $m_a(x)$, описывающее ограничения на значение нечеткой переменной α [48, 51, 58].

Нечеткая величина – это нечеткая переменная, определенная на множестве действительных чисел $\Re$. Нечеткую величину и соответствующее ей нечеткое множество обозначают одной и той же переменной. Множество нечетких величин на $\Re$ обозначают через $F(\Re)$ [39, 48, 51].

Все понятия и определения, введенные для нечетких множеств (формулы (2.1)-(2.7)), справедливы и для нечетких величин. Некоторые из них представлены на рисунке 2.1 [39, 48, 51].

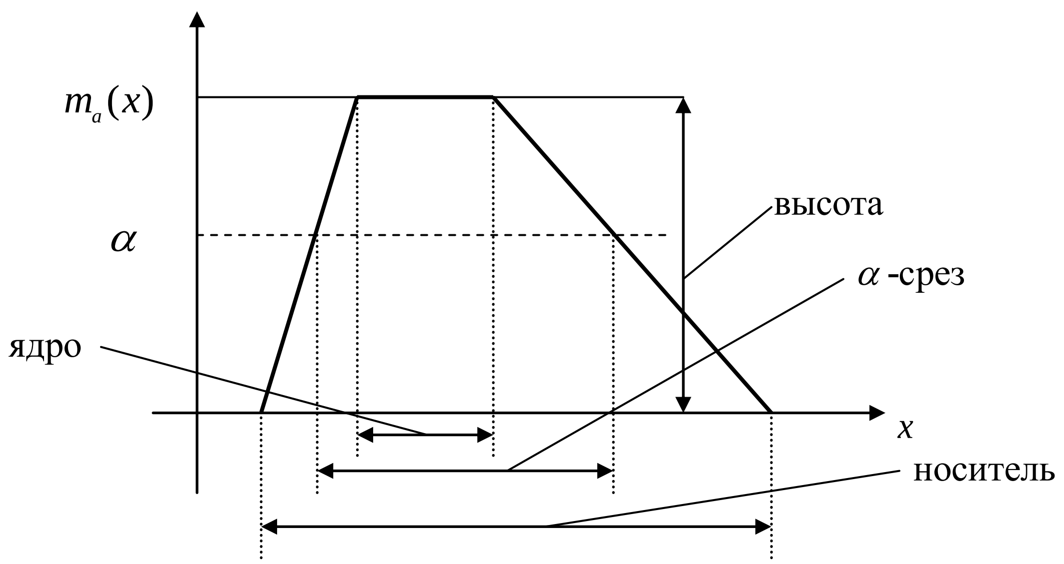


Рисунок 2.1 – Носитель, высота, ядро и α -срез

Как правило, к функции принадлежности нечеткой величины A предъявляется ряд требований. Среди наиболее существенных выделяют следующие [39, 43, 51]:

1) $m_a(x)$ непрерывна (это означает, что в результате небольшого изменения значения аргумента значение функции также изменится мало);

2) функция принадлежности нечеткой величины является нормальной, т. е. $\sup\limits_{x \in \Re}(m_a(x)) = 1$;

3) $m_a(x)$ выпукла (выпукла вверх), т. е.

$$\forall x_1, x_2 (x_1 < x_2), \forall \gamma \in [0,1] \; m_a(\gamma x_1 + (1-\gamma)x_2) \geq \min\{m_a(x_1), m_a(x_2)\}.$$

Нечеткая величина $A \in F(\Re)$ называется выпуклой, если ее функция принадлежности $m_a(x)$ выпукла. Если неравенство выполняется только для $\gamma = 0$ и $\gamma = 1$, то нечеткая величина A называется строго выпуклой [51, 67].

Нетрудно заметить, что если A – строго выпуклая нечеткая величина, то функция принадлежности $m_a(x)$ на носителе $Supp(A)$ является строго возрастающей или строго убывающей или строго возрастает до некоторого x, а затем строго убывает, т. е. является унимодальной [51].

Предположение о выпуклости нечетких величин является существенным, поскольку большинство функций принадлежности, используемых в приложениях, выпуклы. Если нечеткая величина не является выпуклой, то ее можно представить в виде объединения выпуклых нечетких величин. Поэтому без ограничения общности в дальнейшем будем рассматривать только выпуклые нечеткие величины [39, 51].

Среди нечетких величин принято выделять два типа – нечеткое число и нечеткий интервал. Следует заметить, что нечеткое число в общем случае является частным случаем нечеткого интервала, что полностью согласуется с обычными числами и интервалами на множестве действительных чисел. Нечеткий интервал – форма представления неточных величин, более богатая информацией, чем обычный интервал [51].

Нечетким числом A называется нечеткая величина A, функция принадлежности $m_a(x)$ которой является выпуклой и унимодальной на $\Re$. Нечеткое число с модальным значением μ можно рассматривать как нечеткое значение высказывания “x приблизительно равно μ” [16, 63].

Нечеткое число A положительно, если его модальное значение положительно, т. е. $\mu > 0$, и отрицательно, если $\mu < 0$ [67].

Нечеткий интервал с ядром $[a, b]$ обобщает понятие нечеткого числа и рассматривается как нечеткое значение высказывания “x приблизительно находится в интервале $[a, b]$” [51].

Решение задач математического моделирования сложных систем с применением аппарата нечетких множеств требует выполнения большого объема операций над нечеткими переменными. Для удобства использования операций, а также для ввода-вывода и хранения данных желательно работать с функциями принадлежности стандартного вида. Рассмотрим некоторые типы нечетких чисел. Наибольшее применение нашли простейшие частные случаи нечетких чисел и интервалов, получившие свое название по виду их функции принадлежности [51].

Треугольным нечетким числом A (рис. 2.2) с центром в точке μ, левой шириной $\alpha > 0$ и правой шириной $\beta > 0$ называется нечеткое множество A с функцией принадлежности [51, 67, 108]:

$$m_a(x) = \begin{cases} 1-(\mu - x)/\alpha, \textit{если } \mu - \alpha \le x < \mu, \\ 1, \textit{если } x = \mu, \\ 1-(x-\mu)/\beta, \textit{если } \mu < x \le \mu + \beta, \\ 0, \textit{иначе.} \end{cases}$$

Условно треугольное нечеткое число обозначается в виде тройки параметров $A = (\mu, \alpha, \beta)$.

Трапециевидным нечетким числом (рис. 2.3) с отрезком толерантности $[a, b]$, левой шириной α и правой шириной β называется нечеткое множество A с функцией принадлежности вида [51, 67]:

$$m_a(x) = \begin{cases} 1-(a-x)/\alpha, \textit{если } a-\alpha \le x < a, \\ 1, \textit{если } x \in [a, b], \\ 1-(x-b)/\beta, \textit{если } b < x \le b + \beta, \\ 0, \textit{иначе.} \end{cases}$$

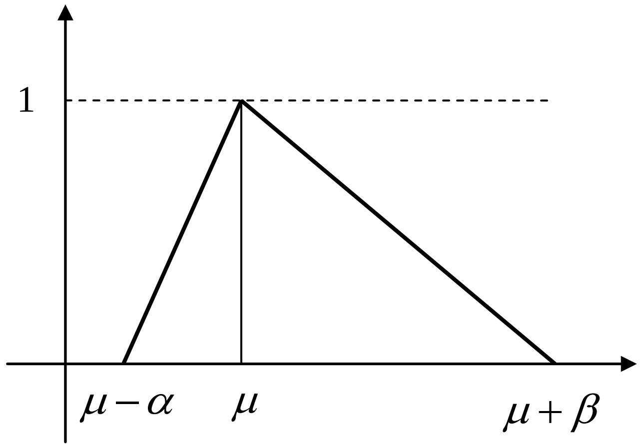


Рисунок 2.2 – Треугольное нечеткое число

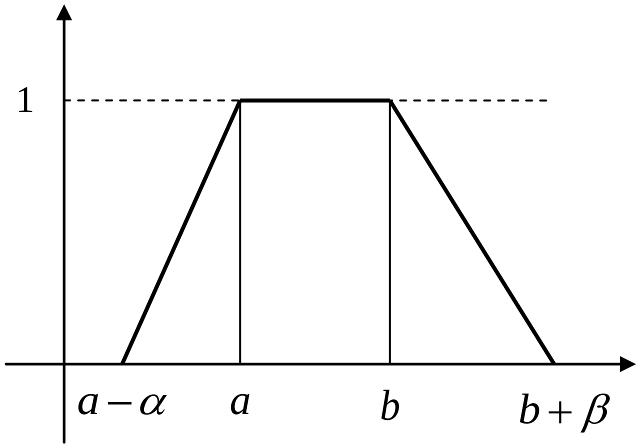


Рисунок 2.3 – Трапециевидное нечеткое число

2.2. Нечеткие числа *L-R* типа и операции над ними

В общем виде нечеткие числа и интервалы задаются с помощью так называемых L-R-функций, обладающих одними и теми же свойствами.

L-функция есть отображение $L:\Re^+ \to [0,1]$, удовлетворяющее следующим условиям [51, 67, 108]:

1) $L(x)$ – непрерывная, невозрастающая функция;

2) $L(x)$ – четная функция, т. е. $L(-x) = L(x)$;

3) $L(0)=1$ и $\forall x > 0 : L(x) < 1$,

а также выполняется одно из двух условий:

4) $\forall x < 1 : L(x) > 0$ и $L(1) = 0$;

5) $\forall x > 0 : L(x) > 0$ и $\lim\limits_{x\to\infty} L(x) = 0$.

Пусть $L(x)$ и $R(x)$ – функции, удовлетворяющие перечисленным выше условиям. A есть унимодальное нечеткое число L-R-типа, если существуют константы $\alpha, \beta > 0$, такие, что функция принадлежности нечеткого числа A имеет вид [67, 78, 108]:

$$m_a(x) = \begin{cases} L\left(\dfrac{\mu - x}{\alpha}\right), если\ x \le \mu, \\ R\left(\dfrac{x - \mu}{\beta}\right), если\ x \ge \mu. \end{cases} \tag{2.8}$$

где α и β – соответственно левый и правый коэффициенты нечеткости, μ – мода нечеткого числа. Условно нечеткое число A обозначается в виде тройки параметров (μ, α, β). Если $\alpha = \beta$, то нечеткое число A называется *симметричным нечетким числом L-R-типа* и обозначается как $A = (\mu, \alpha)$. Заметим, что если $\alpha = \beta = 0$, то получается обычное число из $\Re$ [78].

Пример. Зададим нечеткое число A с помощью функций $L(x) = 1/(1 + x^2)$, $R(x) = 1/(1 + 2|x|)$. Пусть $\alpha = 2$, $\beta = 3$ и $\mu = 5$, тогда функция принадлежности $m_a(x)$ примет вид:

$$m_a(x) = \begin{cases} L\left(\dfrac{5 - x}{2}\right) = \dfrac{1}{1 + \left(\dfrac{5 - x}{2}\right)^2}, если\ x \le 5, \\ R\left(\dfrac{x - 5}{3}\right) = \dfrac{1}{1 + 2\left|\dfrac{x - 5}{3}\right|}, если\ x \ge 5. \end{cases}$$

Приведем некоторые свойства нечетких чисел L-R-типа, основываясь на работах [37, 67, 78, 96].

Если число A задано с помощью L-R-представления $A = (\mu_a, \alpha_a, \beta_a)$, то L-R-представление противоположного ему числа $-A$ имеет вид:

$$-A = (-\mu_a, \beta_a, \alpha_a). \tag{2.9}$$

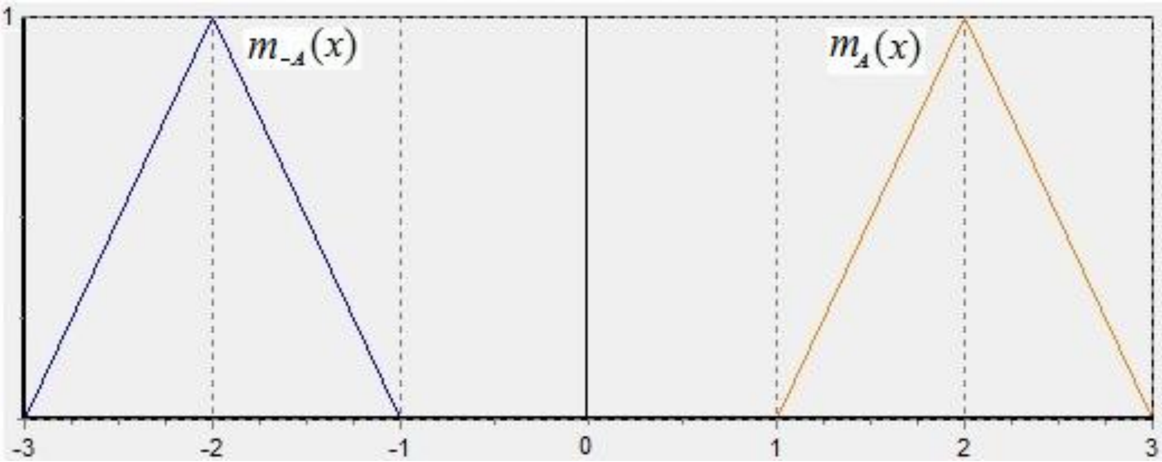


Рисунок 2.4 – Число A = «примерно 2» и число $-A$ = «примерно -2», построенные с помощью *L*-*R*-представления в виде треугольных нечетких чисел

Как видно из рисунка 2.4, нечеткое число A и противоположное ему число $-A$ симметричны относительно оси ординат.

Пример. Пусть $A=(1{,}0.5{,}0.8)$, тогда $-A=(-1{,}0.8{,}0.5)$.

Если число A задано с помощью *L*-*R*-представления $A=(\mu_a,\alpha_a,\beta_a)$, то *L*-*R*-представление обратного нечеткого числа A^{-1} можно выразить в виде формулы:

$$A^{-1}=\left(\frac{1}{\mu_a},\frac{\beta_a}{\mu_a(\mu_a+\beta_a)},\frac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right). \tag{2.10}$$

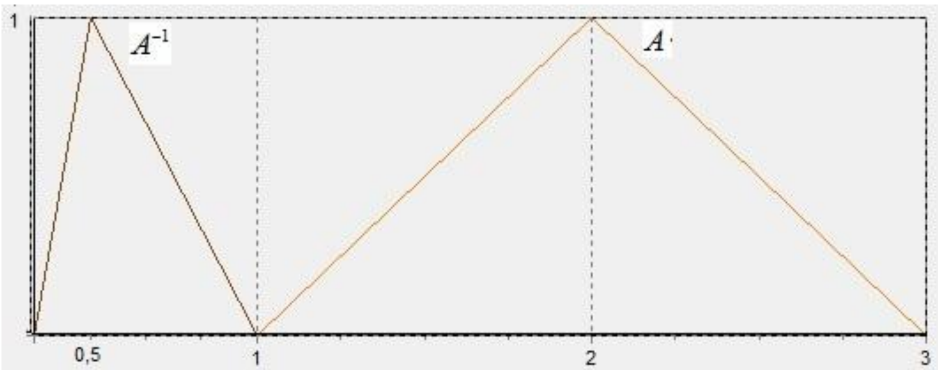


Рисунок 2.5 – Нечеткое число A = «примерно 2» и обратное к нему число A^{-1}

Пример. Пусть $A=(2{,}1{,}0.5)$, тогда

$$A^{-1}=\left(\frac{1}{2},\frac{0.5}{2(2+0.5)},\frac{1}{2(2-1)}\right)=(0.5{,}0.1{,}0.5).$$

Если число A задано с помощью *L*-*R*-представления $A=(\mu_a,\alpha_a,\beta_a)$, то произведение нечеткого числа A на некоторую константу $\lambda\in\Re$ имеет вид:

$$\lambda A=(\lambda\mu_a,\lambda\alpha_a,\lambda\beta_a),\ \lambda>0, \qquad (2.11)$$

$$\lambda A=(\lambda\mu_a,-\lambda\beta_a,-\lambda\alpha_a),\ \lambda<0. \qquad (2.12)$$

Пример. Пусть $A=(2,1,0.5)$, тогда $2A=(4,2,1)$ и $-2A=(-4,1,2)$.

Если $A=(\mu_a,\alpha_a,\beta_a)$ и $B=(\mu_b,\alpha_b,\beta_b)$ являются нечеткими числами *L*-*R*-типа, то расстояние между ними определяется по формуле Евклида:

$$D=\sqrt{(\mu_a-\mu_b)^2+(\alpha_a-\alpha_b)^2+(\beta_a-\beta_b)^2}. \qquad (2.13)$$

Пример. Пусть $A=(2,1,0.5)$ и $B=(1,0.5,0.8)$. Тогда $D=\sqrt{1^2+0{,}5^2+0{,}3^2}=\sqrt{1{,}34}\approx 1{,}16$.

Рассмотрим теперь арифметические операции над нечеткими числами *L*-*R*-типа $A=(\mu_a,\alpha_a,\beta_a)$ и $B=(\mu_b,\alpha_b,\beta_b)$, используя [67, 78]. Для этого применим принцип обобщения, при котором операции над нечеткими числами сводятся к операциям над соответствующими функциями принадлежности.

Пусть $m_a,m_b:\Re\to[0,1]$ являются непрерывными и сюръективными функциями принадлежности нечетких чисел A и B, а $\tilde{\circ}$ – нечеткая бинарная операция, соответствующая непрерывной и монотонной алгебраической операции $\circ$ над обычными числами. Тогда $A\tilde{\circ}B$ – нечеткое число, функция принадлежности которого непрерывна, сюръективна и определяется по формуле:

$$m_{A\tilde{\circ}B}(z)=\sup_{z=x\circ y}\min(m_a(x),m_b(y)).$$

Таким образом, результат сложения, вычитания, умножения и деления нечетких чисел *L*-*R*-типа должен быть точно или приближенно равен некоторому нечеткому числу с теми же функциями *L*-типа и *R*-типа, а коэффициенты нечеткости α и β должны однозначно зависеть от аналогичных параметров исходных нечетких чисел *L*-*R*-типа.

В дальнейшем при рассмотрении арифметических операций над нечеткими числами L-R-типа будем всюду предполагать, что функции L и R заданы.

Суммой нечетких чисел A и B является нечеткое число вида

$$A+B=(\mu_a+\mu_b,\alpha_a+\alpha_b,\beta_a+\beta_b). \quad (2.14)$$

Пример. Пусть $A=(5,1,1)$ и $B=(7,2,2)$. Тогда $A+B=(12,3,3)$.

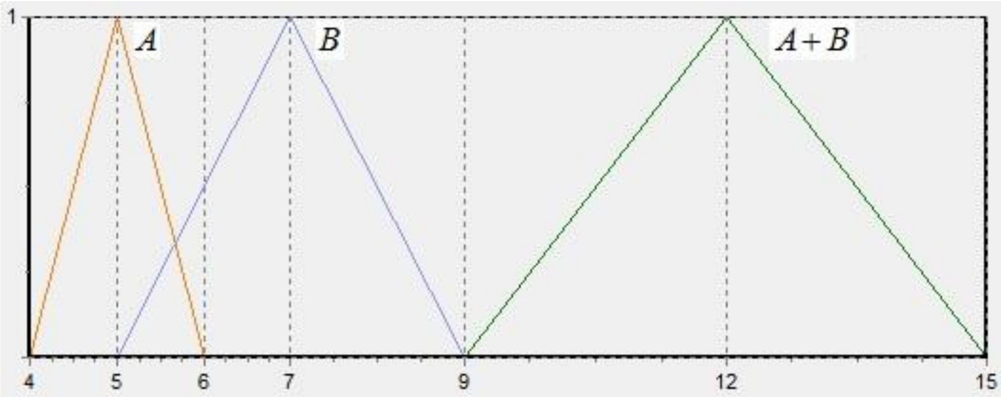


Рисунок 2.6 – Сложение нечетких чисел A = «примерно 5» и B = «примерно 7»

Разностью нечетких чисел A и B является нечеткое число

$$A-B=(\mu_a-\mu_b,\alpha_a+\beta_b,\alpha_b+\beta_a). \quad (2.15)$$

Пример. Пусть $A=(5,1,1)$ и $B=(7,2,2)$. Тогда $A-B=(-2,3,3)$.

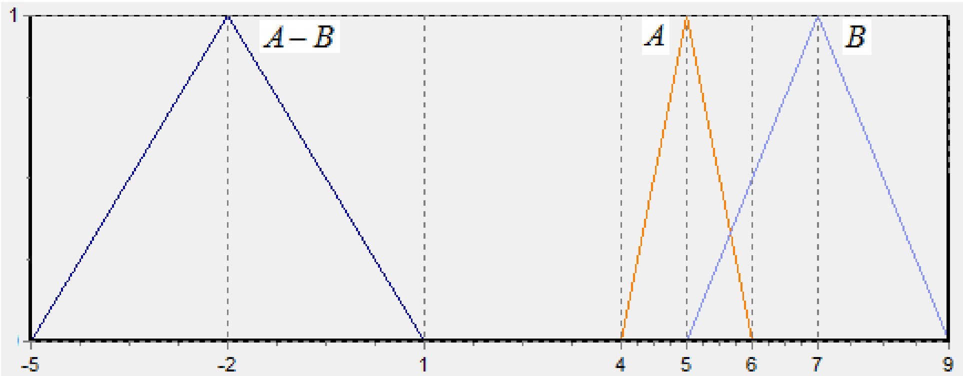


Рисунок 2.7 – Вычитание двух нечетких чисел A = «примерно 5» и B = «примерно 7»

Произведение положительных нечетких чисел A и B определяется выражением:

$$A\cdot B=(\mu_a\cdot\mu_b,\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b,\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b),\ A>0,\ B>0. \quad (2.16)$$

Пример. Пусть $A=(5{,}3{,}2)$ и $B=(7{,}2{,}3)$. Тогда $A\cdot B=(35{,}25{,}35)$.

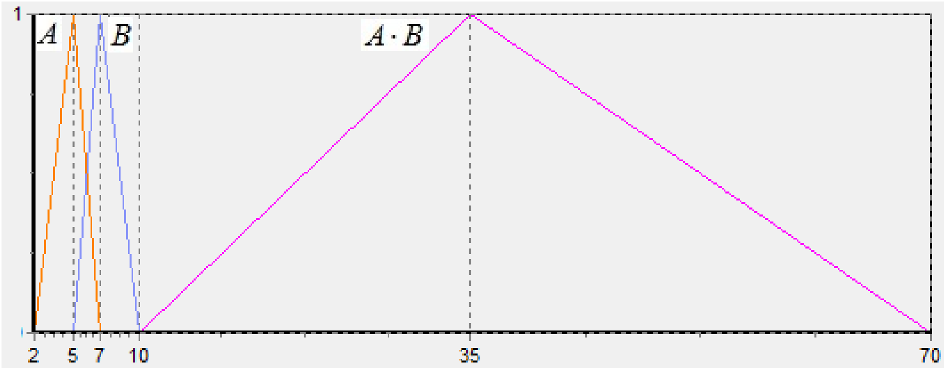


Рисунок 2.8 – Умножение двух нечетких чисел A = «примерно 5» и B = «примерно 7»

Под положительным нечетким числом здесь подразумевается нечеткое число, мода которого положительна. У отрицательного нечеткого числа мода отрицательна.

L-*R*-представление произведения нечетких чисел $A<0$ и $B>0$ имеет вид:

$$A\cdot B=\left(\mu_a\cdot\mu_b,-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b,-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a\right),\ A<0,\ B>0. \quad (2.17)$$

Пример. Пусть $A=(-5{,}3{,}2)$ и $B=(7{,}2{,}3)$. Тогда $A\cdot B=(-35{,}45{,}27)$.

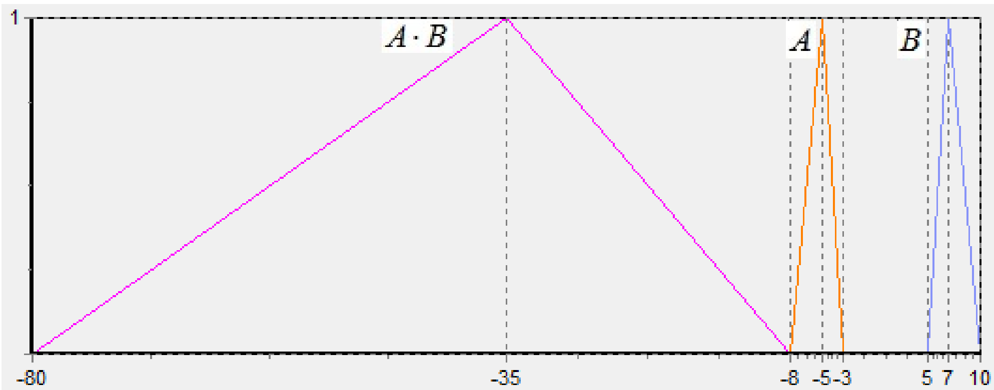


Рисунок 2.9 – Умножение двух нечетких чисел A = «примерно -5» и B = «примерно 7»

Произведение нечетких чисел $A>0$ и $B<0$ выражается формулой:

$$A\cdot B=\left(\mu_a\cdot\mu_b,\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a,\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b\right),\ A>0,\ B<0. \quad (2.18)$$

Пример. Пусть $A=(5{,}3{,}2)$ и $B=(-7{,}2{,}3)$. Тогда $A\cdot B=(-35{,}28{,}27)$.

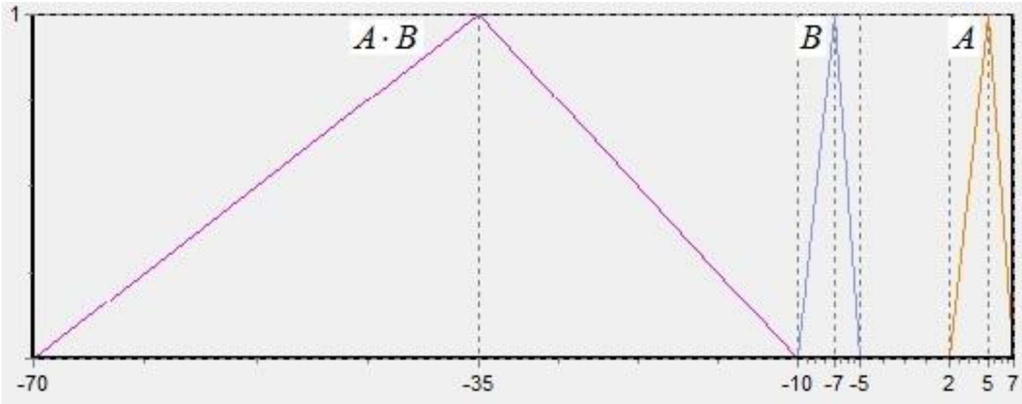


Рисунок 2.10 – Умножение двух нечетких чисел A = «примерно 5» и B = «примерно -7»

Произведение отрицательных нечетких чисел A и B определяется выражением:

$$A \cdot B = \left(\mu_a \cdot \mu_b, -\mu_a\beta_b - \mu_b\beta_a - \beta_a\beta_b, -\mu_a\alpha_b - \mu_b\alpha_a + \alpha_a\alpha_b\right),\ A < 0,\ B < 0. \quad (2.19)$$

Пример. Пусть $A = (-5,3,2)$ и $B = (-7,2,3)$. Тогда $A \cdot B = (35,23,37)$.

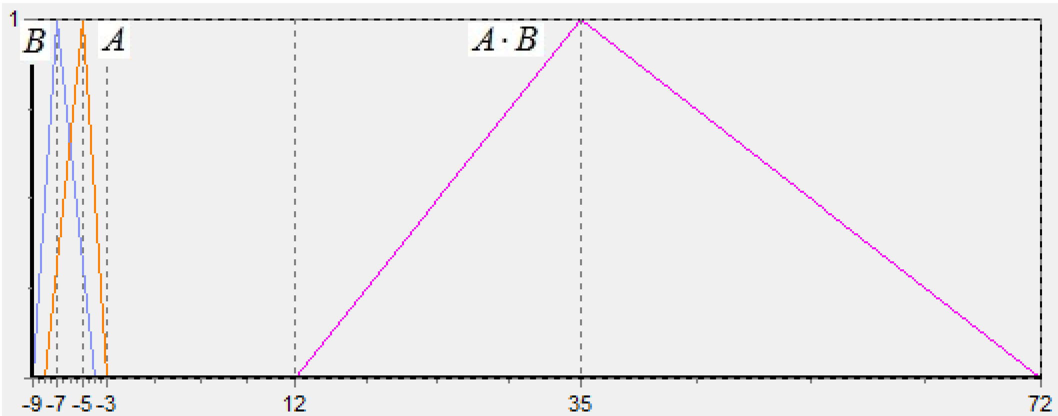


Рисунок 2.11 – Умножение двух нечетких чисел A = «примерно -5» и B = «примерно -7» с помощью L-R-представления

Произведение нечетких чисел A и B произвольного знака имеет вид:

$$A \cdot B = \left(\begin{array}{l} \mu_a\mu_b, \mu_a\mu_b - \min\begin{bmatrix} (\mu_a - \alpha_a)(\mu_b - \alpha_b), (\mu_a - \alpha_a)(\mu_b + \beta_b), \\ (\mu_a + \beta_a)(\mu_b - \alpha_b), (\mu_a + \beta_a)(\mu_b + \beta_b) \end{bmatrix}, \\ \max\begin{bmatrix} (\mu_a - \alpha_a)(\mu_b - \alpha_b), (\mu_a - \alpha_a)(\mu_b + \beta_b), \\ (\mu_a + \beta_a)(\mu_b - \alpha_b), (\mu_a + \beta_a)(\mu_b + \beta_b) \end{bmatrix} - \mu_a\mu_b \end{array} \right). \quad (2.20)$$

2.3. Закон нечеткой внутренней композиции

2.3.1. Понятие закона композиции. Нечеткий группоид

Пусть E – универсальное множество и A и B – подмножества множества E, т. е. $A \subset E$ и $B \subset E$. Обозначим множество нечетких подмножеств множества E через $Fuz(E)$. Тогда $A \subset Fuz(E)$ и $B \subset Fuz(E)$. Введем понятие закона нечеткой внутренней композиции [39, 77].

Законом нечеткой внутренней композиции на множестве E называется отображение из $Fuz(E) \times Fuz(E)$ в $Fuz(E)$, которое каждой упорядоченной паре (A, B), где $A \subset Fuz(E)$ и $B \subset Fuz(E)$, ставит в соответствие единственное нечеткое подмножество $C \subset Fuz(E)$ [39, 77].

Этот закон будем изображать символом *, который, располагаясь между A и B, служит для обозначения нечеткого подмножества, соответствующего упорядоченной паре (A, B). Таким образом,

$$A * B = C.$$

Упорядоченная пара, состоящая из множества $Fuz(E)$ и закона нечеткой внутренней композиции *, определенного на этом множестве всюду, называется нечетким группоидом и обозначается $\langle Fuz(E), * \rangle$. Тем самым, если $A \subset Fuz(E)$ и $B \subset Fuz(E)$, то $A * B \subset Fuz(E)$ [39, 77]. Если * – операция типа сложения, то группоид называется аддитивным, если * – операция типа умножения, то группоид называется мультипликативным [18].

Пусть $Fuz^{LR}(\Re)$ – это семейство нечетких чисел *L*-*R*-типа, определенных на множестве $\Re$, а * – арифметическая операция над нечеткими числами. Тогда, в соответствии с принципом обобщения Заде, справедливо следующее утверждение [77].

Утверждение. Результат операции * над нечеткими числами *L*-*R*-типа представляет собой нечеткое число *L*-*R*-типа.

Исходя из этого утверждения, на основе определения нечеткого группоида заключаем следующее.

Утверждение. Упорядоченная пара, состоящая из множества $Fuz^{LR}(\Re)$ нечетких чисел *L*-*R*-типа и арифметической операции * над нечеткими числами, образует нечеткий группоид. Обозначим его через $\left\langle Fuz^{LR}(\Re),*\right\rangle$ [39, 77].

2.3.2. Основные свойства группоида нечетких чисел *L*-*R*-типа

Пусть * есть закон внутренней композиции нечеткого группоида $\left\langle Fuz^{LR}(\Re),*\right\rangle$. Определим основные свойства группоида через арифметические операции над нечеткими числами *L*-*R*-типа [77].

Коммутативность. Пусть $A=(m_a,\alpha_a,\beta_a)$ и $B=(m_b,\alpha_b,\beta_b)$ являются нечеткими числами *L*-*R*-типа, т. е. $A\in Fuz^{LR}(\Re)$, $B\in Fuz^{LR}(\Re)$. Тогда если для всех упорядоченных пар $(A,B)\in Fuz^{LR}(\Re)\times Fuz^{LR}(\Re)$ выполняется условие:

$$A*B=B*A, \tag{2.21}$$

то говорят, что закон внутренней композиции * коммутативен. Также говорят, что группоид коммутативен [39, 77].

Доказательство.

а) Пусть * – операция типа сложения. Покажем, что

$$A+B=B+A.$$

По правилам сложения нечетких чисел *L*-*R*-типа (формула (2.14)) имеем:

$$A+B=(\mu_a,\alpha_a,\beta_a)+(\mu_b,\alpha_b,\beta_b)=(\mu_a+\mu_b,\alpha_a+\alpha_b,\beta_a+\beta_b).$$

Так как $\mu_a,\alpha_a,\beta_a,\mu_b,\alpha_b,\beta_b\in\Re$, то, воспользовавшись коммутативностью действительных чисел, далее получаем:

$$A+B=(\mu_b+\mu_a,\alpha_b+\alpha_a,\beta_b+\beta_a)=(\mu_b,\alpha_b,\beta_b)+(\mu_a,\alpha_a,\beta_a)=B+A.$$

б) Пусть теперь * – операция типа умножения. Покажем, что выполняется условие:

$$A\cdot B=B\cdot A.$$

Ввиду того, что произведение нечетких чисел *L*-*R*-типа зависит от знака множителей, рассмотрим следующие случаи:

1. Нечеткие числа A и B имеют одинаковый знак.

Пусть $A>0, B>0$. Тогда, используя свойство коммутативности действительных чисел и формулу (2.16), имеем:

$$A\cdot B=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b,\alpha_b,\beta_b)=(\mu_a\mu_b,\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b,\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b)=$$
$$=(\mu_b\mu_a,\mu_b\alpha_a+\mu_a\alpha_b-\alpha_b\alpha_a,\mu_b\beta_a+\mu_a\beta_b+\beta_b\beta_a)=(\mu_b,\alpha_b,\beta_b)\cdot(\mu_a,\alpha_a,\beta_a)=B\cdot A.$$

Пусть теперь $A<0, B<0$. Применяя формулу произведения отрицательных нечетких чисел *L*-*R*-типа (2.19), аналогично показываем, что

$$A\cdot B=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b,\alpha_b,\beta_b)=(\mu_a\mu_b,-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b,-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b)=$$
$$=(\mu_b\mu_a,-\mu_b\beta_a-\mu_a\beta_b-\beta_b\beta_a,-\mu_b\alpha_a-\mu_a\alpha_b+\alpha_b\alpha_a)=(\mu_b,\alpha_b,\beta_b)\cdot(\mu_a,\alpha_a,\beta_a)=B\cdot A.$$

2. Нечеткие числа A и B имеют разный знак.

Пусть $A>0, B<0$. Воспользовавшись формулой (2.9), приходим к произведению положительных нечетких чисел:

$$A\cdot B=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b,\alpha_b,\beta_b)=-(\mu_a,\alpha_a,\beta_a)\cdot(-\mu_b,\beta_b,\alpha_b),\text{ где } A>0, -B>0.$$

Далее по формулам (2.9) и (2.16) имеем:

$$A\cdot B=-(\mu_a,\alpha_a,\beta_a)\cdot(-\mu_b,\beta_b,\alpha_b)=$$
$$=-(-\mu_b\mu_a,\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b,\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a)=$$
$$=(\mu_b\mu_a,\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a,\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)=B\cdot A.$$

Доказательство аналогично для случая $A<0, B>0$.

Таким образом было показано, что нечеткий группоид $\langle Fuz^{LR}(\Re),*\rangle$, состоящий из множества положительных и отрицательных нечетких чисел *L*-*R*-типа, обладает свойством коммутативности.

Ассоциативность. Пусть $A=(m_a,\alpha_a,\beta_a)$, $B=(m_b,\alpha_b,\beta_b)$ и $C=(m_c,\alpha_c,\beta_c)$ являются нечеткими числами *L*-*R*-типа. Если $\forall A,B,C\in Fuz^{LR}(\Re)$ выполняется условие:

$$(A*B)*C=A*(B*C), \tag{2.22}$$

то говорят, что закон внутренней композиции * ассоциативен. Также говорят, что группоид ассоциативен [39, 77].

Доказательство.

а) Пусть * – операция типа сложения. Покажем, что

$$\left(A+B\right)+C=A+\left(B+C\right).$$

По правилам сложения нечетких чисел *L*-*R*-типа получаем:

$$\left(A+B\right)+C=\left(\mu_a+\mu_b,\alpha_a+\alpha_b,\beta_a+\beta_b\right)+\left(\mu_c,\alpha_c,\beta_c\right)=$$
$$=\left(\mu_a+\mu_b+\mu_c,\alpha_a+\alpha_b+\alpha_c,\beta_a+\beta_b+\beta_c\right).$$

Так как $\mu_a,\alpha_a,\beta_a,\mu_b,\alpha_b,\beta_b,\mu_c,\alpha_c,\beta_c\in\Re$, то, воспользовавшись ассоциативностью действительных чисел, имеем:

$$\left(A+B\right)+C=\left(\mu_a+\left(\mu_b+\mu_c\right),\alpha_a+\left(\alpha_b+\alpha_c\right),\beta_a+\left(\beta_b+\beta_c\right)\right)=$$
$$=\left(\mu_a,\alpha_a,\beta_a\right)+\left(\mu_b+\mu_c,\alpha_b+\alpha_c,\beta_b+\beta_c\right)=A+(B+C).$$

б) Пусть теперь * – операция типа умножения. Покажем выполнение условия

$$\left(A\cdot B\right)\cdot C=A\cdot\left(B\cdot C\right).$$

Ввиду того, что произведение нечетких чисел *L*-*R*-типа зависит от знака множителей, рассмотрим следующие случаи:

1. Пусть $A\cdot B>0, C>0$, где $A>0, B>0, C>0$.

Вследствие коммутативности и ассоциативности действительных чисел, используя формулу (2.16), получаем:

$$(A\cdot B)\cdot C=\left(\mu_a\mu_b,\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b,\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b\right)\cdot\left(\mu_c,\alpha_c,\beta_c\right)=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,\mu_a\mu_b\alpha_c+\mu_c\left(\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b\right)-\alpha_c\left(\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b\right),\\ \mu_a\mu_b\beta_c+\mu_c\left(\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b\right)+\beta_c\left(\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b\right)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,\left(\mu_a\mu_b\alpha_c-\mu_a\mu_c\alpha_b-\mu_a\alpha_b\alpha_c\right)+\mu_b\mu_c\alpha_a+\left(-\alpha_a\mu_b\alpha_c-\alpha_a\mu_c\alpha_b+\alpha_a\alpha_b\alpha_c\right),\\ \left(\mu_a\mu_b\beta_c+\mu_a\mu_c\beta_b+\mu_a\beta_b\beta_c\right)+\mu_b\mu_c\beta_a+\left(\beta_a\mu_b\beta_c+\beta_a\mu_c\beta_b+\beta_a\beta_b\beta_c\right)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\left(\mu_b\mu_c\right),\mu_a\left(\mu_b\alpha_c-\mu_c\alpha_b-\alpha_b\alpha_c\right)+\mu_b\mu_c\alpha_a-\alpha_a\left(\mu_b\alpha_c+\mu_c\alpha_b-\alpha_b\alpha_c\right),\\ \mu_a\left(\mu_b\beta_c+\mu_c\beta_b+\beta_b\beta_c\right)+\mu_b\mu_c\beta_a+\beta_a\left(\mu_b\beta_c+\mu_c\beta_b+\beta_b\beta_c\right)\end{pmatrix}=A\cdot(B\cdot C).$$

2. Пусть $A\cdot B>0,\ C<0$, где $A<0,\ B<0,\ C<0$.

Применяя формулы произведения нечетких чисел (2.18) и (2.19), приходим к следующему результату:

$$\left(A\cdot B\right)\cdot C=\left(\mu_a\cdot\mu_b,-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b,-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b\right)\cdot\left(\mu_c,\alpha_c,\beta_c\right)=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,\mu_a\mu_b\alpha_c-\mu_c\left(-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b\right)+\alpha_c\left(-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b\right),\\ \mu_a\mu_b\beta_c-\mu_c\left(-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b\right)-\left(-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b\right)\beta_c\end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a\mu_b\mu_c, (\mu_a\mu_b\alpha_c+\mu_a\mu_c\alpha_b-\mu_a\alpha_b\alpha_c)+\mu_b\mu_c\alpha_a+\begin{pmatrix}-\alpha_a\mu_b\alpha_c-\alpha_a\mu_c\alpha_b+\\ +\alpha_a\alpha_b\alpha_c\end{pmatrix}, \\ (\mu_a\mu_b\beta_c+\mu_a\mu_c\beta_b+\mu_a\beta_b\beta_c)+\mu_b\mu_c\beta_a+(\mu_b\beta_a\beta_c+\mu_c\beta_a\beta_b+\beta_a\beta_b\beta_c) \end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a\mu_b\mu_c, -\mu_a(-\mu_b\alpha_c-\mu_c\alpha_b+\alpha_b\alpha_c)+\mu_b\mu_c\alpha_a+\alpha_a(-\mu_b\alpha_c-\mu_c\alpha_b+\alpha_b\alpha_c), \\ -\mu_a(-\mu_b\beta_c-\mu_c\beta_b-\beta_b\beta_c)+\mu_b\mu_c\beta_a-(-\mu_b\beta_c-\mu_c\beta_b-\beta_b\beta_c)\beta_a \end{pmatrix}=$$

$$=A\cdot(B\cdot C).$$

3. Пусть $A\cdot B>0,\ C>0$, где $A<0,\ B<0,\ C>0$.

Основываясь на формулах (2.16), (2.17) и (2.19), учитывая коммутативность и ассоциативность действительных чисел, имеем

$$(A\cdot B)\cdot C=(\mu_a\cdot\mu_b, -\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b, -\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b)\cdot(\mu_c,\alpha_c,\beta_c)=$$

$$=\begin{pmatrix} \mu_a\mu_b\mu_c, \mu_a\mu_b\alpha_c+\mu_c(-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b)-(-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b)\alpha_c, \\ \mu_a\mu_b\beta_c+\mu_c(-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b)+\beta_c(-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b) \end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a\mu_b\mu_c, (\mu_a\mu_b\alpha_c-\mu_a\mu_c\beta_b+\mu_a\alpha_c\beta_b)-\mu_b\mu_c\beta_a+(\mu_b\alpha_c\beta_a-\mu_c\beta_b\beta_a+\alpha_c\beta_b\beta_a), \\ (\mu_a\mu_b\beta_c-\mu_a\mu_c\alpha_b-\mu_a\alpha_b\beta_c)-\mu_b\mu_c\alpha_a+(-\alpha_a\mu_b\beta_c+\alpha_a\mu_c\alpha_b+\alpha_a\alpha_b\beta_c) \end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a(\mu_b\mu_c), -\mu_a(\mu_b\alpha_c+\mu_c\beta_b-\alpha_c\beta_b)-\mu_b\mu_c\beta_a-(-\mu_b\alpha_c+\mu_c\beta_b-\alpha_c\beta_b)\beta_a, \\ -\mu_a(-\mu_b\beta_c+\mu_c\alpha_b+\alpha_b\beta_c)-\mu_b\mu_c\alpha_a+\alpha_a(-\mu_b\beta_c+\mu_c\alpha_b+\alpha_b\beta_c) \end{pmatrix}=A\cdot(B\cdot C).$$

4. Пусть $A\cdot B>0, C<0$, где $A>0, B>0, C<0$.

Ввиду коммутативности и ассоциативности действительных чисел, а также по формулам (2.16) и (2.18) получаем

$$(A\cdot B)\cdot C=(\mu_a\mu_b, \mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b, \mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b)\cdot(\mu_c,\alpha_c,\beta_c)=$$

$$=\begin{pmatrix} \mu_a\mu_b\mu_c, \mu_a\mu_b\alpha_c-\mu_c(\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b)+\alpha_c(\mu_a\beta_b+\mu_b\beta_a+\beta_a\beta_b), \\ \mu_a\mu_b\beta_c-\mu_c(\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b)-(\mu_a\alpha_b+\mu_b\alpha_a-\alpha_a\alpha_b)\beta_c \end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a\mu_b\mu_c, (\mu_a\mu_b\alpha_c-\mu_a\mu_c\beta_b+\mu_a\alpha_c\beta_b)-\mu_b\mu_c\beta_a+(\mu_b\alpha_c\beta_a-\mu_c\beta_b\beta_a+\alpha_c\beta_b\beta_a), \\ (\mu_a\mu_b\beta_c-\mu_a\mu_c\alpha_b-\mu_a\alpha_b\beta_c)-\mu_b\mu_c\alpha_a+(-\alpha_a\mu_b\beta_c+\alpha_a\mu_c\alpha_b+\alpha_a\alpha_b\beta_c) \end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a(\mu_b\mu_c), \mu_a(\mu_b\alpha_c-\mu_c\beta_b+\alpha_c\beta_b)-\mu_b\mu_c\beta_a+(\mu_b\alpha_c-\mu_c\beta_b+\alpha_c\beta_b)\beta_a, \\ \mu_a(\mu_b\beta_c-\mu_c\alpha_b-\alpha_b\beta_c)-\mu_b\mu_c\alpha_a-\alpha_a(\mu_b\beta_c-\mu_c\alpha_b-\alpha_b\beta_c) \end{pmatrix}=A\cdot(B\cdot C).$$

5. Пусть $A\cdot B<0,\ C>0$, где $A>0,\ B<0,\ C>0$.

Используя свойство коммутативности и ассоциативности действительных чисел и формулы (2.18) и (2.19), имеем

$$(A\cdot B)\cdot C=(\mu_a\mu_b,\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a,\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,-\mu_a\mu_b\beta_c+\mu_c(\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a)+(\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a)\beta_c,\\ -\mu_a\mu_b\alpha_c+\mu_c(\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)-\alpha_c(\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,(-\mu_a\mu_b\beta_c+\mu_a\mu_c\alpha_b+\mu_a\alpha_b\beta_c)-\mu_b\mu_c\beta_a+(-\beta_a\mu_b\beta_c+\beta_a\mu_c\alpha_b+\alpha_b\beta_a\beta_c),\\ (-\mu_a\mu_b\alpha_c+\mu_a\mu_c\beta_b-\mu_a\alpha_c\beta_b)-\mu_b\mu_c\alpha_a+(\alpha_a\mu_b\alpha_c-\alpha_a\mu_c\beta_b+\alpha_a\alpha_c\beta_b)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a(\mu_b\mu_c),\mu_a(-\mu_b\beta_c+\mu_c\alpha_b+\alpha_b\beta_c)-\mu_b\mu_c\beta_a+(-\mu_b\beta_c+\mu_c\alpha_b+\alpha_b\beta_c)\beta_a,\\ \mu_a(-\mu_b\alpha_c+\mu_c\beta_b-\alpha_c\beta_b)-\mu_b\mu_c\alpha_a-\alpha_a(-\mu_b\alpha_c+\mu_c\beta_b-\alpha_c\beta_b)\end{pmatrix}=A\cdot(B\cdot C).$$

6. Пусть $A\cdot B<0,\ C>0$, где $A<0,\ B>0,\ C>0$.

Воспользуемся, как и прежде, свойствами действительных чисел, а также формулами (2.16) и (2.17). Тогда получим

$$(A\cdot B)\cdot C=(\mu_a\mu_b,-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b,-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)\cdot(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,-\mu_a\mu_b\beta_c+\mu_c(-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b)+(-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b)\beta_c,\\ -\mu_a\mu_b\alpha_c+\mu_c(-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)-\alpha_c(-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,(-\mu_a\mu_b\beta_c-\mu_a\mu_c\beta_b-\mu_a\beta_b\beta_c)+\mu_b\mu_c\alpha_a+(\alpha_a\mu_b\beta_c+\alpha_a\mu_c\beta_b+\alpha_a\beta_b\beta_c),\\ (-\mu_a\mu_b\alpha_c-\mu_a\mu_c\alpha_b+\mu_a\alpha_b\alpha_c)+\mu_b\mu_c\beta_a+(-\mu_b\alpha_c\beta_a-\mu_c\alpha_b\beta_a+\alpha_b\alpha_c\beta_a)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a(\mu_b\mu_c),-\mu_a(\mu_b\beta_c+\mu_c\beta_b+\beta_b\beta_c)+\mu_b\mu_c\alpha_a+\alpha_a(\mu_b\beta_c+\mu_c\beta_b+\beta_b\beta_c),\\ -\mu_a(\mu_b\alpha_c+\mu_c\alpha_b-\alpha_b\alpha_c)+\mu_b\mu_c\beta_a-(\mu_b\alpha_c+\mu_c\alpha_b-\alpha_b\alpha_c)\beta_a\end{pmatrix}=A\cdot(B\cdot C).$$

7. Пусть $A\cdot B<0, C<0$, где $A>0, B<0, C<0$.

Применяя формулы произведения нечетких чисел (2.18) и (2.19), а также свойства действительных чисел, имеем

$$(A\cdot B)\cdot C=(\mu_a\mu_b,\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a,\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,-\mu_a\mu_b\beta_c-\mu_c(\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)-(\mu_a\beta_b-\mu_b\alpha_a-\alpha_a\beta_b)\beta_c,\\ -\mu_a\mu_b\alpha_c-\mu_c(\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a)+(\mu_a\alpha_b-\mu_b\beta_a+\alpha_b\beta_a)\alpha_c\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\mu_b\mu_c,(-\mu_a\mu_b\beta_c-\mu_a\mu_c\beta_b-\mu_a\beta_b\beta_c)+\mu_b\mu_c\alpha_a+(\alpha_a\mu_b\beta_c+\alpha_a\mu_c\beta_b+\alpha_a\beta_b\beta_c),\\ (-\mu_a\mu_b\alpha_c-\mu_a\mu_c\alpha_b+\mu_a\alpha_b\alpha_c)+\mu_b\mu_c\beta_a+(-\beta_a\mu_b\alpha_c-\beta_a\mu_c\alpha_b+\beta_a\alpha_b\alpha_c)\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a(\mu_b\mu_c),\mu_a(-\mu_b\beta_c-\mu_c\beta_b-\beta_b\beta_c)+\mu_b\mu_c\alpha_a-\alpha_a(-\mu_b\beta_c-\mu_c\beta_b-\beta_b\beta_c),\\ \mu_a(-\mu_b\alpha_c-\mu_c\alpha_b+\alpha_b\alpha_c)+\mu_b\mu_c\beta_a+\beta_a(-\mu_b\alpha_c-\mu_c\alpha_b+\alpha_b\alpha_c)\end{pmatrix}=$$
$$=A\cdot(B\cdot C).$$

8. Пусть $A\cdot B<0, C<0$, где $A<0, B>0, C<0$.

Учитывая формулы (2.17), (2.18) и (2.19) и свойства действительных чисел, приходим к следующему результату

$$(A\cdot B)\cdot C=(\mu_a\mu_b,-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b,-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)(\mu_c,\alpha_c,\beta_c)=$$

$$=\begin{pmatrix}\mu_a\mu_b\mu_c,-\mu_a\mu_b\beta_c-\mu_c(-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)-(-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)\beta_c,\\ -\mu_a\mu_b\alpha_c-\mu_c(-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b)+(-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b)\alpha_c\end{pmatrix}=$$

$$=\begin{pmatrix}\mu_a\mu_b\mu_c,(-\mu_a\mu_b\beta_c+\mu_a\mu_c\alpha_b+\mu_a\alpha_b\beta_c)-\mu_b\mu_c\beta_a+(-\beta_a\mu_b\beta_c+\beta_a\mu_c\alpha_b+\beta_a\alpha_b\beta_c),\\ (-\mu_a\mu_b\alpha_c+\mu_a\mu_c\beta_b-\mu_a\alpha_c\beta_b)-\mu_b\mu_c\alpha_a+(\alpha_a\mu_b\alpha_c-\alpha_a\mu_c\beta_b+\alpha_a\alpha_c\beta_b)\end{pmatrix}=$$

$$=\begin{pmatrix}\mu_a(\mu_b\mu_c),-\mu_a(\mu_b\beta_c-\mu_c\alpha_b-\alpha_b\beta_c)-\mu_b\mu_c\beta_a-\beta_a(\mu_b\beta_c-\mu_c\alpha_b-\alpha_b\beta_c),\\ -\mu_a(\mu_b\alpha_c-\mu_c\beta_b+\alpha_c\beta_b)-\mu_b\mu_c\alpha_a+\alpha_a(\mu_b\alpha_c-\mu_c\beta_b+\alpha_c\beta_b)\end{pmatrix}=A\cdot(B\cdot C).$$

Таким образом было показано, что нечеткий группоид $\langle Fuz^{LR}(\Re),*\rangle$, состоящий из множества положительных и отрицательных нечетких чисел *L*-*R*-типа, обладает свойством ассоциативности.

Дистрибутивность. Пусть $A=(m_a,\alpha_a,\beta_a)$ – это положительное или отрицательное нечеткое число *L*-*R*-типа, а нечеткие числа $B=(m_b,\alpha_b,\beta_b)$ и $C=(m_c,\alpha_c,\beta_c)$ оба положительны или оба отрицательны и пусть * и *' представляют собой два закона внутренней композиции, определенных на одном и том же множестве $Fuz^{LR}(S)$, состоящем только из положительных или только из отрицательных нечетких чисел. Если для $A,B,C\in Fuz^{LR}(S)$ выполняется условие:

$$A*(B*'C)=(A*B)*'(A*C), \tag{2.23}$$

то говорят, что закон * дистрибутивен слева относительно закона *' [39, 77].

Аналогично, если для $A,B,C\in Fuz^{LR}(S)$ справедливо равенство:

$$(A*'B)*C=(A*C)*'(B*C), \tag{2.24}$$

то говорят, что закон * дистрибутивен справа относительно закона *' [51, 89].

Если закон * дистрибутивен относительно другого закона *' и слева, и справа, то говорят, что он дистрибутивен относительно *' [39, 77]. Тогда можно записать

$$(A*'B)*(C*'D)=(A*C)*'(A*D)*'(B*C)*'(B*D). \tag{2.25}$$

Доказательство.

Пусть * – операция типа умножения, а *' – операция типа сложения.

а) Покажем, что закон * дистрибутивен слева относительно закона *'. Рассмотрим следующие варианты:

1. $A > 0, B + C > 0$.

Пользуясь коммутативностью и дистрибутивностью действительных чисел, а также формулами произведения (2.16) и сложения (2.14) положительных нечетких чисел, получаем:

$$\begin{aligned}
&A \cdot (B + C) = (\mu_a, \alpha_a, \beta_a) \cdot (\mu_b + \mu_c, \alpha_b + \alpha_c, \beta_b + \beta_c) = \\
&= \begin{pmatrix} \mu_a(\mu_b + \mu_c), \mu_a(\alpha_b + \alpha_c) + (\mu_b + \mu_c)\alpha_a - \alpha_a(\alpha_b + \alpha_c), \\ \mu_a(\beta_b + \beta_c) + (\mu_b + \mu_c)\beta_a + \beta_a(\beta_b + \beta_c) \end{pmatrix} = \\
&= \begin{pmatrix} \mu_a\mu_b + \mu_a\mu_c, (\mu_a\alpha_b + \mu_b\alpha_a - \alpha_a\alpha_b) + (\mu_a\alpha_c + \mu_c\alpha_a - \alpha_a\alpha_c), \\ (\mu_a\beta_b + \mu_b\beta_a + \beta_a\beta_b) + (\mu_a\beta_c + \mu_c\beta_a + \beta_a\beta_c) \end{pmatrix} = \\
&= (\mu_a, \alpha_a, \beta_a) \cdot (\mu_b, \alpha_b, \beta_b) + (\mu_a, \alpha_a, \beta_a)(\mu_c, \alpha_c, \beta_c) = (A \cdot B) + (A \cdot C), \\
&A > 0, B > 0, C > 0.
\end{aligned}$$

Из проведенного доказательства следует, что в случае $A > 0, B + C > 0$ закон дистрибутивности выполняется, если $B > 0, C > 0$. Иначе равенство $A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ не верно.

2. $A > 0, B + C < 0$.

Ввиду коммутативности и дистрибутивности действительных чисел, а также по формулам произведения (2.18) и сложения (2.14) нечетких чисел имеем:

$$\begin{aligned}
&A \cdot (B + C) = (\mu_a, \alpha_a, \beta_a) \cdot (\mu_b + \mu_c, \alpha_b + \alpha_c, \beta_b + \beta_c) = \\
&= \begin{pmatrix} \mu_a(\mu_b + \mu_c), \mu_a(\alpha_b + \alpha_c) - (\mu_b + \mu_c)\beta_a + (\alpha_b + \alpha_c)\beta_a, \\ \mu_a(\beta_b + \beta_c) - (\mu_b + \mu_c)\alpha_a - \alpha_a(\beta_b + \beta_c) \end{pmatrix} = \\
&= \begin{pmatrix} \mu_a\mu_b + \mu_a\mu_c, (\mu_a\alpha_b - \mu_b\beta_a + \alpha_b\beta_a) + (\mu_a\alpha_c - \mu_c\beta_a + \alpha_c\beta_a), \\ (\mu_a\beta_b - \mu_b\alpha_a - \alpha_a\beta_b) + (\mu_a\beta_c - \mu_c\alpha_a - \alpha_a\beta_c) \end{pmatrix} = \\
&= (\mu_a, \alpha_a, \beta_a) \cdot (\mu_b, \alpha_b, \beta_b) + (\mu_a, \alpha_a, \beta_a)(\mu_c, \alpha_c, \beta_c) = (A \cdot B) + (A \cdot C), \\
&A > 0, B < 0, C < 0.
\end{aligned}$$

На основе данного доказательства выполнения закона дистрибутивности при $A > 0, B + C < 0$ можно сделать вывод о том, что нечеткие числа B и C

должны быть оба отрицательны. В противном случае равенство $A\cdot(B+C)=(A\cdot B)+(A\cdot C)$ не является верным.

3. $A<0, B+C>0$.

Воспользуемся, как и прежде, свойствами действительных чисел, а также формулами (2.14) и (2.17). Тогда получим:

$$
\begin{aligned}
&A\cdot(B+C)=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b+\mu_c,\alpha_b+\alpha_c,\beta_b+\beta_c)=\\
&=\begin{pmatrix}\mu_a(\mu_b+\mu_c),-\mu_a(\beta_b+\beta_c)+(\mu_b+\mu_c)\alpha_a+\alpha_a(\beta_b+\beta_c),\\ -\mu_a(\alpha_b+\alpha_c)+(\mu_b+\mu_c)\beta_a-(\alpha_b+\alpha_c)\beta_a\end{pmatrix}=\\
&=\begin{pmatrix}\mu_a\mu_b+\mu_a\mu_c,(-\mu_a\beta_b+\mu_b\alpha_a+\alpha_a\beta_b)+(-\mu_a\beta_c+\mu_c\alpha_a+\alpha_a\beta_c),\\ (-\mu_a\alpha_b+\mu_b\beta_a-\alpha_b\beta_a)+(-\mu_a\alpha_c+\mu_c\beta_a-\alpha_c\beta_a)\end{pmatrix}=\\
&=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b,\alpha_b,\beta_b)+(\mu_a,\alpha_a,\beta_a)(\mu_c,\alpha_c,\beta_c)=(A\cdot B)+(A\cdot C),\\
&A<0, B>0, C>0.
\end{aligned}
$$

Проведенное доказательство показывает, что при условии, когда $A<0, B+C>0$, закон дистрибутивности выполняется, если $B>0, C>0$. Иначе равенство $A\cdot(B+C)=(A\cdot B)+(A\cdot C)$ не верно.

4. $A<0, B+C<0$.

Применяя формулы (2.14) и (2.19), имеем:

$$
\begin{aligned}
&A\cdot(B+C)=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b+\mu_c,\alpha_b+\alpha_c,\beta_b+\beta_c)=\\
&=\begin{pmatrix}\mu_a(\mu_b+\mu_c),-\mu_a(\beta_b+\beta_c)-(\mu_b+\mu_c)\beta_a-\beta_a(\beta_b+\beta_c),\\ -\mu_a(\alpha_b+\alpha_c)-(\mu_b+\mu_c)\alpha_a+\alpha_a(\alpha_b+\alpha_c)\end{pmatrix}=\\
&=\begin{pmatrix}\mu_a\mu_b+\mu_a\mu_c,(-\mu_a\beta_b-\mu_b\beta_a-\beta_a\beta_b)+(-\mu_a\beta_c-\mu_c\beta_a-\beta_a\beta_c),\\ (-\mu_a\alpha_b-\mu_b\alpha_a+\alpha_a\alpha_b)+(-\mu_a\alpha_c-\mu_c\alpha_a+\alpha_a\alpha_c)\end{pmatrix}=\\
&=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_b,\alpha_b,\beta_b)+(\mu_a,\alpha_a,\beta_a)(\mu_c,\alpha_c,\beta_c)=(A\cdot B)+(A\cdot C),\\
&A<0, B<0, C<0.
\end{aligned}
$$

Данное доказательство свидетельствует о том, что оба нечетких числа B и C должны быть отрицательны, в противном случае равенство $A\cdot(B+C)=(A\cdot B)+(A\cdot C)$, где $A<0, B+C<0$, не будет являться верным.

б) Покажем теперь, что закон * дистрибутивен справа относительно закона *', т. е. справедливо равенство (2.24). Рассмотрим следующие случаи:

1. $A+B>0, C>0$.

Пользуясь свойствами коммутативности и дистрибутивности действительных чисел, а также формулами (2.14) и (2.16), получаем:

$$(A+B)\cdot C=(\mu_a+\mu_b,\alpha_a+\alpha_b,\beta_a+\beta_b)\cdot(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix}(\mu_a+\mu_b)\mu_c,(\mu_a+\mu_b)\alpha_c+\mu_c(\alpha_a+\alpha_b)-(\alpha_a+\alpha_b)\alpha_c,\\ (\mu_a+\mu_b)\beta_c+\mu_c(\beta_a+\beta_b)+(\beta_a+\beta_b)\beta_c\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\mu_c+\mu_b\mu_c,(\mu_a\alpha_c+\mu_c\alpha_a-\alpha_a\alpha_c)+(\mu_b\alpha_c+\mu_c\alpha_b-\alpha_b\alpha_c),\\ (\mu_a\beta_c+\mu_c\beta_a+\beta_a\beta_c)+(\mu_b\beta_c+\mu_c\beta_b+\beta_b\beta_c)\end{pmatrix}=$$
$$=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_c,\alpha_c,\beta_c)+(\mu_b,\alpha_b,\beta_b)(\mu_c,\alpha_c,\beta_c)=(A\cdot C)+(B\cdot C),$$
$$A>0,B>0,C>0.$$

На основе проведенного доказательства можно сделать вывод о том, что в случае, когда $A+B>0,C>0$, закон дистрибутивности выполняется, если оба слагаемых в сумме $A+B$ положительны, т. е. при $A>0,B>0$.

2. $A+B>0,C<0$.

Учитывая формулы (2.14) и (2.18) и свойства действительных чисел, приходим к следующему результату:

$$(A+B)\cdot C=(\mu_a+\mu_b,\alpha_a+\alpha_b,\beta_a+\beta_b)\cdot(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix}(\mu_a+\mu_b)\mu_c,(\mu_a+\mu_b)\alpha_c-\mu_c(\beta_a+\beta_b)+\alpha_c(\beta_a+\beta_b),\\ (\mu_a+\mu_b)\beta_c-\mu_c(\alpha_a+\alpha_b)-(\alpha_a+\alpha_b)\beta_c\end{pmatrix}=$$
$$=\begin{pmatrix}\mu_a\mu_c+\mu_b\mu_c,(\mu_a\alpha_c-\mu_c\beta_a+\alpha_c\beta_a)+(\mu_b\alpha_c-\mu_c\beta_b+\alpha_c\beta_b),\\ (\mu_a\beta_c-\mu_c\alpha_a-\alpha_a\beta_c)+(\mu_b\beta_c-\mu_c\alpha_b-\alpha_b\beta_c)\end{pmatrix}=$$
$$=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_c,\alpha_c,\beta_c)+(\mu_b,\alpha_b,\beta_b)(\mu_c,\alpha_c,\beta_c)=(A\cdot C)+(B\cdot C),$$
$$A>0,B>0,C<0.$$

Из доказательства следует, что слагаемые в сумме $A+B$ должны быть положительны, иначе закон дистрибутивности не будет выполнен.

3. $A+B<0,C>0$.

Ввиду формул (2.14) и (2.17), получаем следующее:

$$(A+B)\cdot C=(\mu_a+\mu_b,\alpha_a+\alpha_b,\beta_a+\beta_b)\cdot(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix}(\mu_a+\mu_b)\mu_c,-(\mu_a+\mu_b)\beta_c+\mu_c(\alpha_a+\alpha_b)+(\alpha_a+\alpha_b)\beta_c,\\ -(\mu_a+\mu_b)\alpha_c+\mu_c(\beta_a+\beta_b)-\alpha_c(\beta_a+\beta_b)\end{pmatrix}=$$

$$=\begin{pmatrix} \mu_a\mu_c+\mu_b\mu_c, (-\mu_a\beta_c+\mu_c\alpha_a+\alpha_a\beta_c)+(-\mu_b\beta_c+\mu_c\alpha_b+\alpha_b\beta_c), \\ (-\mu_a\alpha_c+\mu_c\beta_a-\alpha_c\beta_a)+(-\mu_b\alpha_c+\mu_c\beta_b-\alpha_c\beta_b) \end{pmatrix}=$$
$$=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_c,\alpha_c,\beta_c)+(\mu_b,\alpha_b,\beta_b)(\mu_c,\alpha_c,\beta_c)=(A\cdot C)+(B\cdot C),$$
$$A<0, B<0, C>0.$$

Заметим, что закон дистрибутивности выполняется только в том случае, когда оба слагаемых в сумме $A+B$ будут меньше нуля, т. е. $A<0, B<0$, и при этом $C>0$.

4. $A+B<0, C<0$.

Воспользуемся формулами (2.14) и (2.19), а также коммутативностью и дистрибутивностью действительных чисел. Тогда имеем

$$(A+B)\cdot C=(\mu_a+\mu_b, \alpha_a+\alpha_b, \beta_a+\beta_b)\cdot(\mu_c,\alpha_c,\beta_c)=$$
$$=\begin{pmatrix} (\mu_a+\mu_b)\mu_c, -(\mu_a+\mu_b)\beta_c-\mu_c(\beta_a+\beta_b)-(\beta_a+\beta_b)\beta_c, \\ -(\mu_a+\mu_b)\alpha_c-\mu_c(\alpha_a+\alpha_b)+(\alpha_a+\alpha_b)\alpha_c \end{pmatrix}=$$
$$=\begin{pmatrix} \mu_a\mu_c+\mu_b\mu_c, (-\mu_a\beta_c-\mu_c\beta_a-\beta_a\beta_c)+(-\mu_b\beta_c-\mu_c\beta_b-\beta_b\beta_c), \\ (-\mu_a\alpha_c-\mu_c\alpha_a+\alpha_a\alpha_c)+(-\mu_b\alpha_c-\mu_c\alpha_b+\alpha_b\alpha_c) \end{pmatrix}=$$
$$=(\mu_a,\alpha_a,\beta_a)\cdot(\mu_c,\alpha_c,\beta_c)+(\mu_b,\alpha_b,\beta_b)(\mu_c,\alpha_c,\beta_c)=(A\cdot C)+(B\cdot C),$$
$$A<0, B<0, C<0.$$

Из приведенного доказательства следует, что слагаемые в сумме $A+B$ должны быть оба отрицательны, чтобы выполнялся закон дистрибутивности в случае, когда $A+B<0, C<0$.

Таким образом, в ходе доказательств были выявлены условия, при которых мультипликативный закон * является дистрибутивным слева и дистрибутивным справа относительно аддитивного закона *', т. е. свойство дистрибутивности справедливо для нечетких чисел *L*-*R*-типа определенного знака. Стоит отметить, что закон * дистрибутивен слева и справа относительно закона *', если A – это положительное или отрицательное нечеткое число *L*-*R*-типа, а нечеткие числа B и C оба положительны или оба отрицательны, иными словами, свойство дистрибутивности выполняется на множестве только положительных или только отрицательных нечетких чисел.

Единичный элемент. Нечеткий группоид $\left\langle Fuz^{LR}(\Re),*\right\rangle$ с внутренним законом композиции * обладает левой единицей E, если $\forall A \in Fuz^{LR}(\Re)$

$$E * A = A, \tag{2.26}$$

и правой единицей E', если $\forall A \in Fuz^{LR}(\Re)$

$$A * E' = A. \tag{2.27}$$

Элемент, который является одновременно и левой, и правой единицей, называется единицей [39, 77]. Тогда нечеткий группоид имеет единицу E, если $\forall A \in Fuz^{LR}(\Re)$

$$E * A = A * E = A. \tag{2.28}$$

а) Определим единичный элемент в нечетком группоиде $\left\langle Fuz^{LR}(\Re),+\right\rangle$ с операцией сложения.

1. Покажем существование левой единицы. Рассмотрим некоторое нечеткое число $A \in Fuz^{LR}(\Re)$ и будем искать левую единицу $E = (\mu_e, \alpha_e, \beta_e)$ из условия (2.26). Тогда по формуле (2.14) получаем

$$E + A = (\mu_e + \mu_a, \alpha_e + \alpha_a, \beta_e + \beta_a) = A = (\mu_a, \alpha_a, \beta_a) \Rightarrow \begin{cases} \mu_e = 0, \\ \alpha_e = 0, \\ \beta_e = 0. \end{cases}$$

Таким образом, левой единицей для нечеткого группоида $\left\langle Fuz^{LR}(\Re),+\right\rangle$ с операцией сложения является элемент $E = (0,0,0)$.

2. Найдем теперь правую единицу $E' = (\mu'_e, \alpha'_e, \beta'_e)$ нечеткого группоида $\left\langle Fuz^{LR}(\Re),+\right\rangle$. Применяя формулу (2.14), из условия (2.27) имеем

$$A + E' = (\mu_a + \mu'_e, \alpha_a + \alpha'_e, \beta_a + \beta'_e) = A = (\mu_a, \alpha_a, \beta_a) \Rightarrow \begin{cases} \mu'_e = 0, \\ \alpha'_e = 0, \\ \beta'_e = 0. \end{cases}$$

Тем самым получаем, что правой единицей для нечеткого группоида $\left\langle Fuz^{LR}(\Re),+\right\rangle$ с операцией сложения является элемент $E' = (0,0,0)$.

Так как левая и правая единица совпадают, то элемент $E=(0,0,0)$ будет являться единицей для нечеткого группоида $\langle Fuz^{LR}(\Re),+\rangle$ с операцией сложения на множестве нечетких чисел L-R-типа.

б) Определим единичный элемент в нечетком группоиде $\langle Fuz^{LR}(\Re),\times\rangle$ с операцией умножения.

1. Покажем сначала, что существует левая единица. Рассмотрим некоторое нечеткое число $A\in Fuz^{LR}(\Re)$ такое, что $A>0$. Будем искать левую единицу $E=(\mu_e,\alpha_e,\beta_e)$ из условия (2.26). Применяя формулу (2.16), имеем

$$E\cdot A=(\mu_e\mu_a,\mu_e\alpha_a+\mu_a\alpha_e-\alpha_e\alpha_a,\mu_e\beta_a+\mu_a\beta_e+\beta_e\beta_a)=A=(\mu_a,\alpha_a,\beta_a)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_e\mu_a=\mu_a,\\ \mu_e\alpha_a+\mu_a\alpha_e-\alpha_e\alpha_a=\alpha_a,\\ \mu_e\beta_a+\mu_a\beta_e+\beta_e\beta_a=\beta_a.\end{cases}\Rightarrow\begin{cases}\mu_e=1,\\ \alpha_e(\mu_a-\alpha_a)=0,\\ \beta_e(\mu_a+\beta_a)=0.\end{cases}\Rightarrow\begin{cases}\mu_e=1,\\ \left[\begin{array}{l}\alpha_e=0,\\ \mu_a=\alpha_a.\end{array}\right.\\ \left[\begin{array}{l}\beta_e=0,\\ \mu_a=-\beta_a.\end{array}\right.\end{cases}$$

В этом случае получаем следующие варианты для представления левой единицы:

1) если $\alpha_e=0,\beta_e=0$, то $E=(1,0,0)$;

2) если $\mu_a=\alpha_a$ и $\mu_a\neq-\beta_a$, то $E=(1,\alpha_e,0)$, α_e – любое положительное число.

Остальные ситуации не являются возможными, так как $A>0\Rightarrow\mu_a>0$, а из определения нечеткого числа L-R-типа следует, что и $\beta_a\geq 0$.

Пусть теперь $A<0$. Воспользовавшись формулой (2.18), получаем:

$$E\cdot A=(\mu_e\mu_a,\mu_e\alpha_a-\mu_a\beta_e+\alpha_a\beta_e,\mu_e\beta_a-\mu_a\alpha_e-\alpha_e\beta_a)=A=(\mu_a,\alpha_a,\beta_a)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_e\mu_a=\mu_a,\\ \mu_e\alpha_a-\mu_a\beta_e+\alpha_a\beta_e=\alpha_a,\\ \mu_e\beta_a-\mu_a\alpha_e-\alpha_e\beta_a=\beta_a.\end{cases}\Rightarrow\begin{cases}\mu_e=1,\\ \beta_e(\alpha_a-\mu_a)=0,\\ \alpha_e(-\mu_a-\beta_a)=0.\end{cases}\Rightarrow\begin{cases}\mu_e=1,\\ \left[\begin{array}{l}\beta_e=0,\\ \mu_a=\alpha_a.\end{array}\right.\\ \left[\begin{array}{l}\alpha_e=0,\\ \mu_a=-\beta_a.\end{array}\right.\end{cases}$$

Отсюда получаем следующие варианты для представления левой единицы:

1) если $\alpha_e = 0, \beta_e = 0$, то $E = (1{,}0{,}0)$;

2) если $\mu_a \neq \alpha_a$ и $\mu_a = -\beta_a$, то $E = (1, \alpha_e, 0)$, где α_e – любое положительное число.

Остальные ситуации не представляются возможными, так как $A < 0 \Rightarrow \mu_a < 0$, а из определения нечеткого числа *L*-*R*-типа следует, что $\alpha_a \geq 0$.

Таким образом, левой единицей для нечеткого группоида $\langle Fuz^{LR}(\Re), \times \rangle$ с операцией умножения является элемент $E = (1{,}0{,}0)$.

2. Найдем правую единицу $E' = (\mu'_e, \alpha'_e, \beta'_e)$ нечеткого группоида $\langle Fuz^{LR}(\Re), \times \rangle$. Рассмотрим сначала некоторое нечеткое число $A \in Fuz^{LR}(\Re)$ такое, что $A > 0$. Тогда, применяя формулу (2.16), на основе условия (2.27) имеем

$$A \cdot E' = (\mu_a \mu'_e, \mu_a \alpha'_e + \mu'_e \alpha_a - \alpha_a \alpha'_e, \mu_a \beta'_e + \mu'_e \beta_a + \beta_a \beta'_e) = A = (\mu_a, \alpha_a, \beta_a) \Rightarrow$$

$$\Rightarrow \begin{cases} \mu_a \mu'_e = \mu_a, \\ \mu_a \alpha'_e + \mu'_e \alpha_a - \alpha_a \alpha'_e = \alpha_a, \\ \mu_a \beta'_e + \mu'_e \beta_a + \beta_a \beta'_e = \beta_a. \end{cases} \Rightarrow \begin{cases} \mu'_e = 1, \\ \alpha'_e (\mu_a - \alpha_a) = 0, \\ \beta'_e (\mu_a + \beta_a) = 0. \end{cases} \Rightarrow \begin{cases} \mu'_e = 1, \\ \left[\begin{array}{l} \alpha'_e = 0, \\ \mu_a = \alpha_a. \end{array} \right. \\ \left[\begin{array}{l} \beta'_e = 0, \\ \mu_a = -\beta_a. \end{array} \right. \end{cases}$$

В этом случае получаем следующие варианты для представления правой единицы:

1) если $\alpha'_e = 0, \beta'_e = 0$, то $E' = (1{,}0{,}0)$;

2) если $\mu_a = \alpha_a$ и $\mu_a \neq -\beta_a$, то $E' = (1, \alpha'_e, 0)$, где α'_e – любое положительное число.

Другие ситуации не являются возможными, так как $A > 0 \Rightarrow \mu_a > 0$, а из определения нечеткого числа *L*-*R*-типа следует, что и $\beta_a \geq 0$.

Рассмотрим теперь случай, когда $A<0$. Воспользовавшись формулой (2.17), получаем

$$A\cdot E'=\left(\mu_a\mu_e',-\mu_a\beta_e'+\mu_e'\alpha_a+\alpha_a\beta_e',-\mu_a\alpha_e'+\mu_e'\beta_a-\alpha_e'\beta_a\right)=A=\left(\mu_a,\alpha_a,\beta_a\right)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_e'\mu_a=\mu_a,\\ -\mu_a\beta_e'+\mu_e'\alpha_a+\alpha_a\beta_e'=\alpha_a,\\ -\mu_a\alpha_e'+\mu_e'\beta_a-\alpha_e'\beta_a=\beta_a.\end{cases}\Rightarrow\begin{cases}\mu_e'=1,\\ \beta_e'\left(\alpha_a-\mu_a\right)=0,\\ \alpha_e'\left(-\mu_a-\beta_a\right)=0.\end{cases}\Rightarrow\begin{cases}\mu_e'=1,\\ \left[\begin{array}{l}\beta_e'=0,\\ \mu_a=\alpha_a.\end{array}\right.\\ \left[\begin{array}{l}\alpha_e'=0,\\ \mu_a=-\beta_a.\end{array}\right.\end{cases}$$

Отсюда получаем следующие варианты для представления правой единицы:

1) если $\alpha_e'=0,\ \beta_e'=0$, то $E'=(1,0,0)$;

2) если $\mu_a\neq\alpha_a$ и $\mu_a=-\beta_a$, то $E'=(1,\alpha_e',0)$, где α_e' – любое положительное число.

Остальные ситуации не представляются возможными, так как $A<0\Rightarrow\mu_a<0$, а из определения нечеткого числа *L*-*R*-типа следует, что $\alpha_a\geq 0$.

Тем самым заключаем, что правой единицей для нечеткого группоида $\left\langle Fuz^{LR}(\Re),\times\right\rangle$ с операцией умножения является элемент $E'=(1,0,0)$.

Так как левая и правая единица совпадают, то элемент $E=(1,0,0)$ будет являться единицей для нечеткого группоида $\left\langle Fuz^{LR}(\Re),\times\right\rangle$ с операцией умножения на множестве нечетких чисел *L*-*R*-типа.

Обратный элемент. Рассмотрим закон композиции *, для которого существует единичный элемент $E=(0,0,0)$, если * – операция сложения, и единичный элемент $E=(1,0,0)$, если * – операция умножения. Нечеткий группоид $\left\langle Fuz^{LR}(\Re),*\right\rangle$ на множестве нечетких чисел *L*-*R*-типа с внутренним законом композиции * имеет левый обратный элемент $\overline{A}\in Fuz^{LR}(\Re)$, если $\forall A\in Fuz^{LR}(\Re)$

$$\overline{A}*A=E, \tag{2.29}$$

и правый элемент $\overline{A}'$, если $\forall A \in Fuz^{LR}(\Re)$

$$A * \overline{A}' = E. \tag{2.30}$$

Элемент, который является одновременно и левым, и правым обратным, называется обратным для $A \in Fuz^{LR}(\Re)$ [39, 77]. В этом случае говорят, что нечеткий группоид имеет обратный элемент $\overline{A}$, если $\forall A \in Fuz^{LR}(\Re)$

$$\overline{A} * A = A * \overline{A} = E. \tag{2.31}$$

а) Определим обратный элемент в нечетком группоиде $\langle Fuz^{LR}(\Re), + \rangle$ с операцией сложения.

1. Проверим существование левого обратного элемента. Рассмотрим некоторое нечеткое число $A \in Fuz^{LR}(\Re)$. Будем искать левый обратный элемент $\overline{A} = (\mu_{\overline{a}}, \alpha_{\overline{a}}, \beta_{\overline{a}})$ из условия (2.29) и формулы (2.14). Тогда имеем

$$\overline{A} + A = (\mu_{\overline{a}} + \mu_a, \alpha_{\overline{a}} + \alpha_a, \beta_{\overline{a}} + \beta_a) = E = (0,0,0) \Rightarrow \begin{cases} \mu_{\overline{a}} = -\mu_a, \\ \alpha_{\overline{a}} = -\alpha_a, \\ \beta_{\overline{a}} = -\beta_a. \end{cases}$$

В рамках предположения, что $\alpha_a \ge 0$ и $\beta_a \ge 0$, приходим к выводу, что левый обратный элемент не может быть определен для нечеткого группоида $\langle Fuz^{LR}(\Re), + \rangle$ с операцией сложения.

2. Покажем, что правый обратный элемент также не существует. Для этого воспользуемся условием (2.30) и формулой (2.14). Получим следующее:

$$A + \overline{A}' = (\mu_a + \mu'_{\overline{a}}, \alpha_a + \alpha'_{\overline{a}}, \beta_a + \beta'_{\overline{a}}) = E = (0,0,0) \Rightarrow \begin{cases} \mu'_{\overline{a}} = -\mu_a, \\ \alpha'_{\overline{a}} = -\alpha_a, \\ \beta'_{\overline{a}} = -\beta_a. \end{cases}$$

Так как $\alpha_a \ge 0$ и $\beta_a \ge 0$, то имеем, что правый обратный элемент не
3.
может быть определен для нечеткого группоида $\langle Fuz^{LR}(\Re), + \rangle$ с операцией сложения.

Таким образом, приходим к выводу, что нечеткий группоид $\langle Fuz^{LR}(\Re), + \rangle$ с операцией сложения не имеет обратного элемента.

б) Определим обратный элемент в нечетком группоиде $\left\langle Fuz^{LR}(\Re),\times\right\rangle$ с операцией умножения.

1. Покажем сначала существование левого обратного элемента нечеткого группоида $\left\langle Fuz^{LR}(\Re),\times\right\rangle$. Рассмотрим некоторое нечеткое число $A\in Fuz^{LR}(\Re)$ такое, что $A>0$. Будем искать левый обратный элемент $\overline{A}=\left(\mu_{\overline{a}},\alpha_{\overline{a}},\beta_{\overline{a}}\right)$ из условия (2.29). Применяя формулу (2.16), имеем

$$\overline{A}\cdot A=\left(\mu_{\overline{a}}\mu_a,\mu_{\overline{a}}\alpha_a+\mu_a\alpha_{\overline{a}}-\alpha_{\overline{a}}\alpha_a,\mu_{\overline{a}}\beta_a+\mu_a\beta_{\overline{a}}+\beta_{\overline{a}}\beta_a\right)=E=(1,0,0)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_{\overline{a}}\mu_a=1,\\ \mu_{\overline{a}}\alpha_a+\mu_a\alpha_{\overline{a}}-\alpha_{\overline{a}}\alpha_a=0,\\ \mu_{\overline{a}}\beta_a+\mu_a\beta_{\overline{a}}+\beta_{\overline{a}}\beta_a=0.\end{cases}\Rightarrow\begin{cases}\mu_{\overline{a}}=\dfrac{1}{\mu_a},\\ \alpha_a+\mu_a^2\alpha_{\overline{a}}-\mu_a\alpha_{\overline{a}}\alpha_a=0,\\ \beta_a+\mu_a^2\beta_{\overline{a}}+\mu_a\beta_{\overline{a}}\beta_a=0.\end{cases}\Rightarrow\begin{cases}\mu_{\overline{a}}=\dfrac{1}{\mu_a},\\ \alpha_{\overline{a}}=\dfrac{\alpha_a}{\mu_a(\alpha_a-\mu_a)},\\ \beta_{\overline{a}}=-\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)}.\end{cases}$$

Пусть теперь $A<0$. Воспользовавшись формулой (2.19), получаем

$$\overline{A}\cdot A=\left(\mu_{\overline{a}}\mu_a,-\mu_{\overline{a}}\beta_a-\mu_a\beta_{\overline{a}}-\beta_{\overline{a}}\beta_a,-\mu_{\overline{a}}\alpha_a-\mu_a\alpha_{\overline{a}}+\alpha_{\overline{a}}\alpha_a\right)=E=(1,0,0)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_{\overline{a}}\mu_a=1,\\ -\mu_{\overline{a}}\beta_a-\mu_a\beta_{\overline{a}}-\beta_{\overline{a}}\beta_a=0,\\ -\mu_{\overline{a}}\alpha_a-\mu_a\alpha_{\overline{a}}+\alpha_{\overline{a}}\alpha_a=0.\end{cases}\Rightarrow\begin{cases}\mu_{\overline{a}}=\dfrac{1}{\mu_a},\\ -\beta_a-\mu_a^2\beta_{\overline{a}}-\mu_a\beta_{\overline{a}}\beta_a=0,\\ -\alpha_a-\mu_a^2\alpha_{\overline{a}}+\mu_a\alpha_{\overline{a}}\alpha_a=0.\end{cases}\Rightarrow\begin{cases}\mu_{\overline{a}}=\dfrac{1}{\mu_a},\\ \beta_{\overline{a}}=-\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)},\\ \alpha_{\overline{a}}=\dfrac{\alpha_a}{\mu_a(\alpha_a-\mu_a)}.\end{cases}$$

Тем самым, в обоих случаях получаем одно и то же нечеткое число

$$\overline{A}=\left(\frac{1}{\mu_a},\frac{\alpha_a}{\mu_a(\alpha_a-\mu_a)},-\frac{\beta_a}{\mu_a(\mu_a+\beta_a)}\right)=\left(\frac{-1}{\mu_a},\frac{\beta_a}{\mu_a(\mu_a+\beta_a)},\frac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right),$$

которое будет являться левым обратным элементом нечеткого группоида $\left\langle Fuz^{LR}(\Re),\times\right\rangle$ с операцией умножения на множестве положительных и отрицательных нечетких чисел *L*-*R*-типа.

2. Найдем правый обратный элемент $\overline{A}'=\left(\mu'_{\overline{a}},\alpha'_{\overline{a}},\beta'_{\overline{a}}\right)$ нечеткого группоида $\left\langle Fuz^{LR}(\Re),\times\right\rangle$. Рассмотрим некоторое нечеткое число $A\in Fuz^{LR}(\Re)$ такое, что $A>0$. Тогда, применяя формулу (2.16), на основе условия (2.30) имеем

$$A\cdot\overline{A}'=\left(\mu_a\mu'_{\overline{a}},\mu_a\alpha'_{\overline{a}}+\mu'_{\overline{a}}\alpha_a-\alpha_a\alpha'_{\overline{a}},\mu_a\beta'_{\overline{a}}+\mu'_{\overline{a}}\beta_a+\beta_a\beta'_{\overline{a}}\right)=E=(1,0,0)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_a\mu'_{\overline{a}}=1,\\ \mu_a\alpha'_{\overline{a}}+\mu'_{\overline{a}}\alpha_a-\alpha_a\alpha'_{\overline{a}}=0,\\ \mu_a\beta'_{\overline{a}}+\mu'_{\overline{a}}\beta_a+\beta_a\beta'_{\overline{a}}=0.\end{cases}\Rightarrow\begin{cases}\mu'_{\overline{a}}=\dfrac{1}{\mu_a},\\ \mu_a^2\alpha'_{\overline{a}}+\alpha_a-\mu_a\alpha_a\alpha'_{\overline{a}}=0,\\ \mu_a^2\beta'_{\overline{a}}+\beta_a+\mu_a\beta_a\beta'_{\overline{a}}=0.\end{cases}\Rightarrow\begin{cases}\mu'_{\overline{a}}=\dfrac{1}{\mu_a},\\ \alpha'_{\overline{a}}=\dfrac{\alpha_a}{\mu_a(\alpha_a-\mu_a)},\\ \beta'_{\overline{a}}=-\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)}.\end{cases}$$

Рассмотрим теперь случай, когда $A<0$. Воспользовавшись формулой (2.19), получаем:

$$A\cdot\overline{A}'=\left(\mu_a\mu'_{\overline{a}},-\mu_a\beta'_{\overline{a}}-\mu'_{\overline{a}}\beta_a-\beta_a\beta'_{\overline{a}},-\mu_a\alpha'_{\overline{a}}-\mu'_{\overline{a}}\alpha_a+\alpha_a\alpha'_{\overline{a}}\right)=E=(1,0,0)\Rightarrow$$

$$\Rightarrow\begin{cases}\mu_a\mu'_{\overline{a}}=1,\\ -\mu_a\beta'_{\overline{a}}-\mu'_{\overline{a}}\beta_a-\beta_a\beta'_{\overline{a}}=0,\\ -\mu_a\alpha'_{\overline{a}}-\mu'_{\overline{a}}\alpha_a+\alpha_a\alpha'_{\overline{a}}=0.\end{cases}\Rightarrow\begin{cases}\mu'_{\overline{a}}=\dfrac{1}{\mu_a},\\ -\mu_a^2\beta'_{\overline{a}}-\beta_a-\mu_a\beta_a\beta'_{\overline{a}}=0,\\ -\mu_a^2\alpha'_{\overline{a}}-\alpha_a+\mu_a\alpha_a\alpha'_{\overline{a}}=0.\end{cases}\Rightarrow\begin{cases}\mu'_{\overline{a}}=\dfrac{1}{\mu_a},\\ \beta'_{\overline{a}}=-\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)},\\ \alpha'_{\overline{a}}=\dfrac{\alpha_a}{\mu_a(\alpha_a-\mu_a)}.\end{cases}$$

Таким образом, правым обратным элементом нечеткого группоида $\langle Fuz^{LR}(\Re),\times\rangle$ с операцией умножения является элемент

$$\overline{A}'=\left(\frac{1}{\mu_a},\frac{\alpha_a}{\mu_a(\alpha_a-\mu_a)},-\frac{\beta_a}{\mu_a(\mu_a+\beta_a)}\right)=\left(\frac{-1}{\mu_a},\frac{\beta_a}{\mu_a(\mu_a+\beta_a)},\frac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right).$$

Так как левый и правый обратные элементы совпадают, то элемент $\overline{A}=\left(\dfrac{-1}{\mu_a},\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)},\dfrac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right)$ будет являться обратным элементом нечеткого группоида $\langle Fuz^{LR}(\Re),\times\rangle$ с операцией умножения на множестве нечетких чисел *L*-*R*-типа. Заметим, что обратный элемент существует при выполнении следующих условий:

$$\begin{cases}\mu_a\neq 0,\\ \mu_a\neq\beta_a,\\ \mu_a\neq\alpha_a,\end{cases}$$

поскольку в этом случае параметры нечеткого числа не определены. Обозначим

$$S = \{(\mu_a, \alpha_a, \beta_a) : \mu_a = 0 \vee \mu_a = \beta_a \vee \mu_a = \alpha_a\},$$

тогда нечеткий группоид $\langle Fuz^{LR}(\Re) \setminus S, \times \rangle$ имеет обратный элемент.

2.3.3. Типы алгебр с одной и двумя арифметическими операциями

Рассмотрим нечеткий группоид $\langle Fuz^{LR}(\Re), * \rangle$ с внутренним законом композиции *, где $Fuz^{LR}(\Re)$ образует семейство нечетких чисел *L*-*R*-типа. Перенесем некоторые понятия алгебры с одной и двумя арифметическими операциями из теории обычных множеств на нечеткие подмножества.

Ранее было показано, что нечеткий группоид $\langle Fuz^{LR}(\Re), * \rangle$ обладает свойствами коммутативности и ассоциативности, имеет единичный элемент $E = (0,0,0)$, если * – операция сложения, и единичный элемент $E = (1,0,0)$, если * – операция умножения. При этом нечеткий группоид $\langle Fuz^{LR}(\Re) \setminus S, \times \rangle$ имеет обратный элемент $\overline{A} = \left(\frac{-1}{\mu_a}, \frac{\beta_a}{\mu_a(\mu_a + \beta_a)}, \frac{\alpha_a}{\mu_a(\mu_a - \alpha_a)} \right)$ на множестве нечетких чисел *L*-*R*-типа.

Введем определения полугруппы и абелевой полугруппы [18, 35].

Полугруппа – это группоид $\langle Fuz^{LR}(\Re), * \rangle$, в котором операция * является ассоциативной.

Абелева полугруппа – это полугруппа $\langle Fuz^{LR}(\Re), * \rangle$ с коммутативной операцией *.

Тогда справедливы следующие утверждения.

Утверждение 1. Полугруппа $\langle Fuz^{LR}(\Re), * \rangle$, содержащая единичный элемент, образует моноид.

Утверждение 2. Моноид $\langle Fuz^{LR}(\Re) \setminus S, \times \rangle$ с обратным элементом является группой.

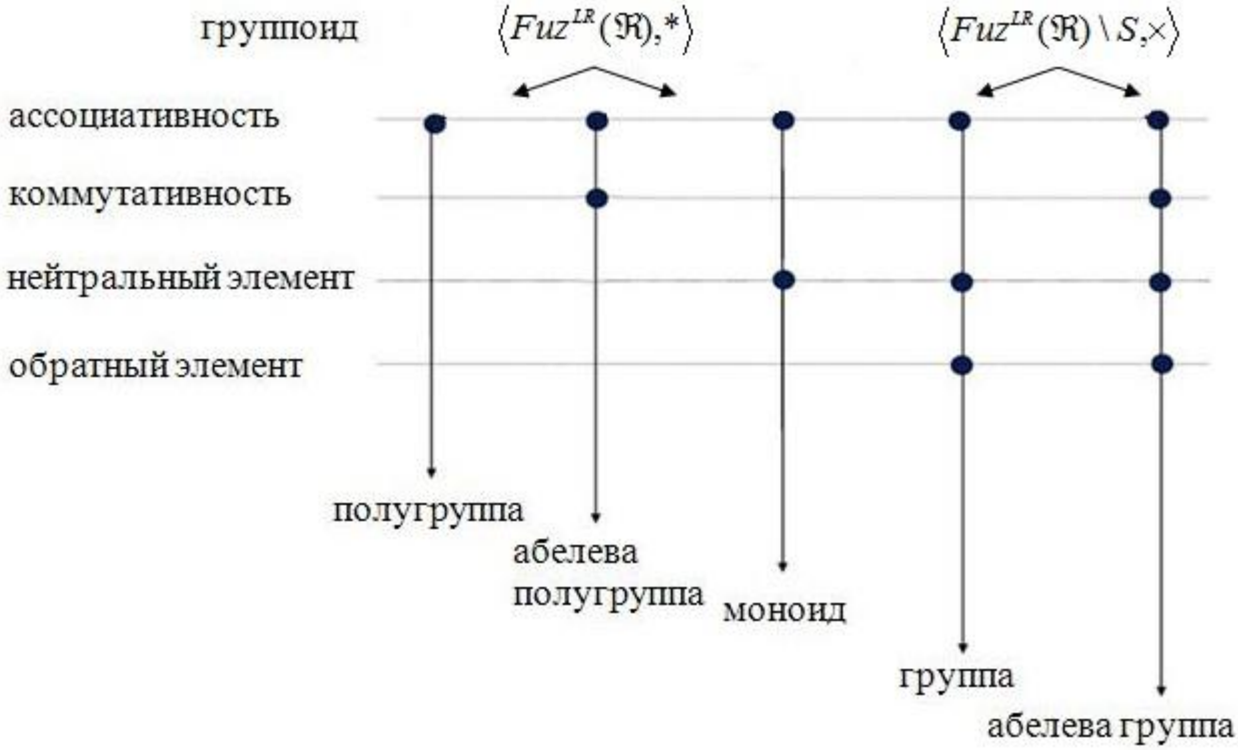


Рисунок 2.12 – Схема образования понятий для алгебры с одной операцией

Утверждение 3. Абелева полугруппа $\langle Fuz^{LR}(\Re)\setminus S,\times\rangle$, содержащая единичный и обратный элементы, образует абелеву группу.

Введенные понятия схематически представлены на рисунке 2.12.

Пусть заданы операция типа сложения (+) и операция типа умножения (×) на множестве нечетких чисел *L*-*R*-типа.

Ввиду того что группоид $\langle Fuz^{LR}(\Re),+\rangle$ с операцией сложения не имеет обратного элемента, для группоида $\langle Fuz^{LR}(\Re),+,\times\rangle$ на множестве нечетких чисел *L*-*R*-типа не могут быть определены алгебраические структуры с двумя операциями.

2.4. Некоторые дополнительные свойства операций над нечеткими числами *L*-*R*-типа

Используя основные понятия нечеткой арифметики, докажем следующие свойства операций над нечеткими числами [78].

1. Пусть нечеткое число A задано с помощью L-R-представления $A=(\mu_a,\alpha_a,\beta_a)$ и $-A$ – это противоположное ему число. Тогда

$$-(-A)=(\mu_a,\alpha_a,\beta_a)=A. \qquad (2.32)$$

Доказательство.

В самом деле, используя формулу (2.9), получим

$$-(-A)=-(-(\mu_a,\alpha_a,\beta_a))=-(-\mu_a,\beta_a,\alpha_a)=(\mu_a,\alpha_a,\beta_a)=A.$$

2. Пусть $A=(\mu_a,\alpha_a,\beta_a)$ является нечетким числом L-R-типа. Тогда инволюция числа A имеет вид:

$$\left(A^{-1}\right)^{-1}=(\mu_a,\alpha_a,\beta_a)=A. \qquad (2.33)$$

Доказательство.

Пусть $A^{-1}=(\mu,\alpha,\beta)$, где $\mu=\dfrac{1}{\mu_a},\alpha=\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)},\beta=\dfrac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}$ по формуле (2.10). Тогда получим

$$\left(A^{-1}\right)^{-1}=\left(\frac{1}{\mu},\frac{\beta}{\mu(\mu+\beta)},\frac{\alpha}{\mu(\mu-\alpha)}\right),$$

где $\dfrac{1}{\mu}=\mu_a$,

$$\frac{\beta}{\mu(\mu+\beta)}=\frac{\alpha_a\mu_a}{\mu_a(\mu_a-\alpha_a)\left(\dfrac{1}{\mu_a}+\dfrac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right)}=\frac{\alpha_a\mu_a}{\mu_a(\mu_a-\alpha_a)\left(\dfrac{\mu_a-\alpha_a+\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right)}=\alpha_a,$$

$$\frac{\alpha}{\mu(\mu-\alpha)}=\frac{\beta_a\mu_a}{\mu_a(\mu_a+\beta_a)\left(\dfrac{1}{\mu_a}-\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)}\right)}=\frac{\beta_a\mu_a}{\mu_a(\mu_a+\beta_a)\left(\dfrac{\mu_a+\beta_a-\beta_a}{\mu_a(\mu_a+\beta_a)}\right)}=\beta_a.$$

Таким образом находим, что

$$\left(A^{-1}\right)^{-1}=\left(\frac{1}{\mu},\frac{\beta}{\mu(\mu+\beta)},\frac{\alpha}{\mu(\mu-\alpha)}\right)=(\mu_a,\alpha_a,\beta_a)=A.$$

3. Пусть нечеткое число A задано с помощью L-R-представления $A=(\mu_a,\alpha_a,\beta_a)$. Тогда произведение нечеткого числа A на обратное ему число A^{-1} вычисляется по формулам

$$A\cdot A^{-1}=\left(1,\frac{\alpha_a+\beta_a}{\mu_a+\beta_a},\frac{\alpha_a+\beta_a}{\mu_a-\alpha_a}\right),\ \text{если}\ A>0, \tag{2.34}$$

$$A\cdot A^{-1}=\left(1,\frac{\alpha_a+\beta_a}{\alpha_a-\mu_a},-\frac{\alpha_a+\beta_a}{\mu_a+\beta_a}\right),\ \text{если}\ A<0. \tag{2.35}$$

Доказательство.

а) Рассмотрим сначала случай, когда нечеткое число $A=(\mu_a,\alpha_a,\beta_a)>0$, а следовательно, и $A^{-1}=\left(\frac{1}{\mu_a},\frac{\beta_a}{\mu_a(\mu_a+\beta_a)},\frac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right)>0$ (формула (2.10)). Тогда ввиду (2.16) найдем формулу произведения нечеткого числа $A>0$ на обратное ему число $A^{-1}>0$. Получим

$$A\cdot A^{-1}=(\mu,\alpha,\beta),$$

где $\mu=\mu_a\frac{1}{\mu_a}=1$,

$$\alpha=\mu_a\frac{\beta_a}{\mu_a(\mu_a+\beta_a)}+\frac{1}{\mu_a}\alpha_a-\alpha_a\frac{\beta_a}{\mu_a(\mu_a+\beta_a)}=\frac{\mu_a\beta_a-\alpha_a\beta_a+\alpha_a\mu_a+\alpha_a\beta_a}{\mu_a(\mu_a+\beta_a)}=\frac{\alpha_a+\beta_a}{\mu_a+\beta_a},$$

$$\beta=\mu_a\frac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}+\frac{1}{\mu_a}\beta_a+\frac{\beta_a\alpha_a}{\mu_a(\mu_a-\alpha_a)}=\frac{\mu_\alpha\alpha_a+\alpha_a\beta_a+\mu_\alpha\beta_a-\alpha_a\beta_a}{\mu_a(\mu_a-\alpha_a)}=\frac{\alpha_a+\beta_a}{\mu_a-\alpha_a}.$$

Таким образом,

$$A\cdot A^{-1}=(\mu,\alpha,\beta)=\left(1,\frac{\alpha_a+\beta_a}{\mu_a+\beta_a},\frac{\alpha_a+\beta_a}{\mu_a-\alpha_a}\right),\ \text{если}\ A>0.$$

Пример. Пусть $A=(2{,}1{,}0.5)$ и $A^{-1}=(0.5{,}0.1{,}0.5)$. Тогда $A\cdot A^{-1}=(1{,}3/5{,}1.5)$.

б) Рассмотрим теперь нечеткое число $A=(\mu_a,\alpha_a,\beta_a)<0$, а, следовательно, и $A^{-1}=\left(\frac{1}{\mu_a},\frac{\beta_a}{\mu_a(\mu_a+\beta_a)},\frac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right)<0$ (формула (2.10)). Тогда, используя (2.19), найдем формулу для произведения нечеткого числа $A<0$ на обратное ему число $A^{-1}<0$, предполагая, что результат также является нечетким числом L-R-типа:

$$A\cdot A^{-1}=(\mu,\alpha,\beta),$$

где $\mu = \mu_a \dfrac{1}{\mu_a} = 1$,

$$\alpha = -\mu_a \frac{\alpha_a}{\mu_a(\mu_a - \alpha_a)} - \frac{1}{\mu_a}\beta_a - \beta_a \frac{\alpha_a}{\mu_a(\mu_a - \alpha_a)} = \frac{-\mu_a\alpha_a + \alpha_a\beta_a - \mu_a\beta_a - \alpha_a\beta_a}{\mu_a(\mu_a - \alpha_a)} = \frac{\alpha_a + \beta_a}{\alpha_a - \mu_a}$$

,

$$\beta = -\frac{\mu_a\beta_a}{\mu_a(\mu_a + \beta_a)} - \frac{1}{\mu_a}\alpha_a + \frac{\alpha_a\beta_a}{\mu_a(\mu_a + \beta_a)} = \frac{-\mu_a\beta_a - \alpha_a\beta_a - \alpha_a\mu_a + \alpha_a\beta_a}{\mu_a(\mu_a + \beta_a)} = -\frac{\alpha_a + \beta_a}{\mu_a + \beta_a}.$$

Таким образом имеем, что при $A < 0$

$$A \cdot A^{-1} = (\mu, \alpha, \beta) = \left(1, \frac{\alpha_a + \beta_a}{\alpha_a - \mu_a}, -\frac{\alpha_a + \beta_a}{\mu_a + \beta_a}\right).$$

4. Пусть $A = (\mu_a, \alpha_a, \beta_a)$ и $B = (\mu_b, \alpha_b, \beta_b)$ являются нечеткими числами *L*-*R*-типа и при этом мода нечеткого числа B равна нулю. Тогда их произведение является нечетким нулем вида

$$A \cdot B = \begin{pmatrix} 0, -\min[-(\mu_a - \alpha_a)\alpha_b, (\mu_a - \alpha_a)\beta_b, -(\mu_a + \beta_a)\alpha_b, (\mu_a + \beta_a)\beta_b], \\ \max[-(\mu_a - \alpha_a)\alpha_b, (\mu_a - \alpha_a)\beta_b, -(\mu_a + \beta_a)\alpha_b, (\mu_a + \beta_a)\beta_b] \end{pmatrix}. \quad (2.36)$$

Доказательство следует из правила умножения нечетких чисел произвольного знака (формула (2.20)).

5. Пусть нечеткие числа $A = (\mu_a, \alpha_a, \beta_a)$ и $B = (\mu_b, \alpha_b, \beta_b)$ имеют *L*-*R*-представление и мода нечеткого числа B равна единице. Тогда их произведение является нечетким числом с модой μ_a вида

$$A \cdot B = (\mu_a, \mu_a\alpha_b + \alpha_a - \alpha_a\alpha_b, \mu_a\beta_b + \beta_a + \beta_a\beta_b), \text{ если } A > 0, \quad (2.37)$$

Пример. Пусть $A = (5,3,2)$ и $B = (1,1,2)$. Тогда $A \cdot B = (5,5,16)$.

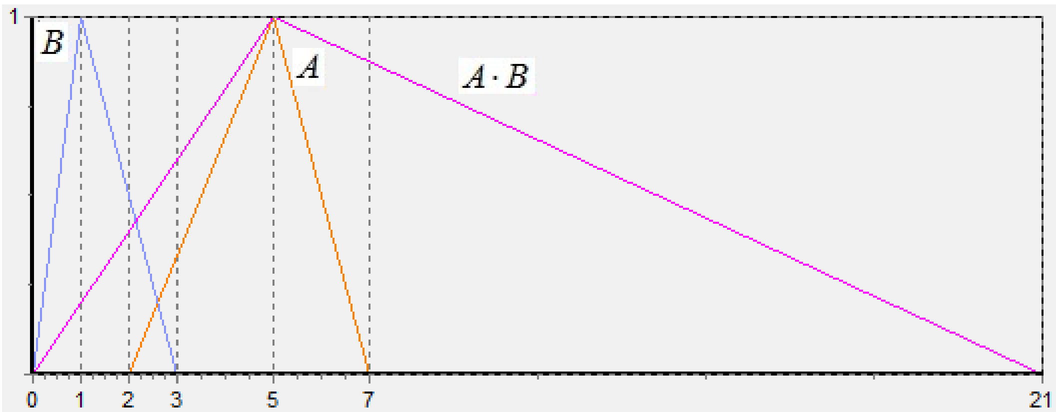


Рисунок 2.13 – Умножение чисел A = «примерно 5» и B = «примерно 1»

$$A \cdot B = \left(\mu_a, -\mu_a\beta_b + \alpha_a + \alpha_a\beta_b, -\mu_a\alpha_b + \beta_a - \alpha_b\beta_a\right), \text{ если } A < 0. \quad (2.38)$$

Доказательство следует из (2.16) и (2.17) при $\mu_b = 1$.

Пример. Пусть $A = (-5,3,2)$ и $B = (1,1,2)$. Тогда $A \cdot B = (-5,19,6)$.

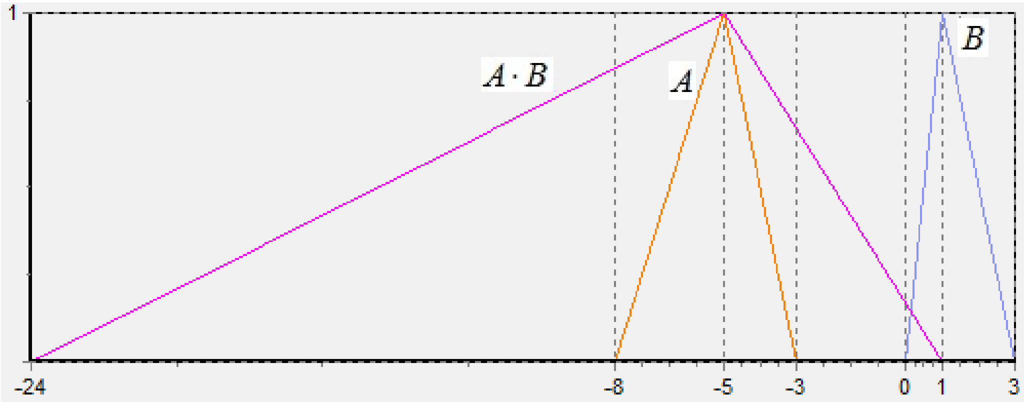


Рисунок 2.14 – Умножение чисел A = «примерно -5», B = «примерно 1»

6. Пусть $A = (\mu_a, \alpha_a, \beta_a)$ является нечетким числом *L*-*R*-типа, тогда

$$A^2 = \left(\mu_a^2, 2\mu_a\alpha_a - \alpha_a^2, 2\mu_a\beta_a + \beta_a^2\right), \text{ при } A > 0, \quad (2.39)$$

$$A^2 = \left(\mu_a^2, -2\mu_a\beta_a - \beta_a^2, -2\mu_a\alpha_a + \alpha_a^2\right), \text{ при } A < 0. \quad (2.40)$$

Доказательство.

а) Пусть $A > 0$. По формуле (2.16) находим, что

$$A^2 = \left(\mu_a^2, \mu_a\alpha_a + \mu_a\alpha_a - \alpha_a^2, \mu_a\beta_a + \mu_a\beta_a + \beta_a^2\right) = \left(\mu_a^2, 2\mu_a\alpha_a - \alpha_a^2, 2\mu_a\beta_a + \beta_a^2\right).$$

Пример. Пусть $A = (2,1,3)$. Тогда $A^2 = (4,3,21)$.

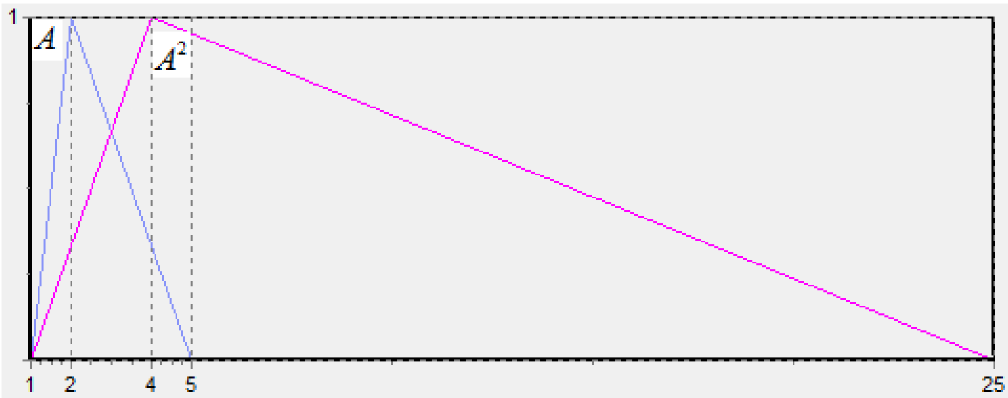


Рисунок 2.15 – Квадрат нечеткого числа A = «примерно 2»,
найденный с помощью *L*-*R*-представления

б) При $A < 0$ на основе формулы (2.19) получим

$$A^2 = \left(\mu_a^2, -\mu_a\beta_a - \mu_a\beta_a - \beta_a^2, -\mu_a\alpha_a - \mu_a\alpha_a + \alpha_a^2\right) = \left(\mu_a^2, -2\mu_a\beta_a - \beta_a^2, -2\mu_a\alpha_a + \alpha_a^2\right).$$

Пример. Пусть $A = (-2{,}1{,}3)$. Тогда $A^2 = (4{,}7{,}5)$.

Рисунок 2.16 – Квадрат нечеткого числа A = «примерно -2»

7. Пусть нечеткие числа $A = (\mu_a, \alpha_a, \beta_a)$ и $B = (\mu_b, \alpha_b, \beta_b)$ имеют L-R-представление, тогда:

$$(A + B)^2 = A^2 + 2AB + B^2, \text{ если } A > 0,\ B > 0 \text{ или } A < 0,\ B < 0. \quad (2.41)$$

Доказательство.

а) Положим, что $A + B > 0$. Тогда на основе (2.14) и (2.16) и из формулы (2.39) получим

$$(A+B)^2 = \begin{pmatrix} (\mu_a + \mu_b)^2, 2(\mu_a + \mu_b)(\alpha_a + \alpha_b) - (\alpha_a + \alpha_b)^2, \\ 2(\mu_a + \mu_b)(\beta_a + \beta_b) + (\beta_a + \beta_b)^2 \end{pmatrix} =$$

$$= \left(\mu_a^2, 2\mu_a\alpha_a - \alpha_a^2, 2\mu_a\beta_a + \beta_a^2\right) + 2\left(\mu_a\mu_b, \mu_a\alpha_b + \mu_b\alpha_a - \alpha_a\alpha_b, \mu_a\beta_b + \mu_b\beta_a + \beta_a\beta_b\right) +$$

$$+ \left(\mu_b^2, 2\mu_b\alpha_b - \alpha_b^2, 2\mu_b\beta_b + \beta_b^2\right) = A^2 + 2AB + B^2,$$

при условии, что $A > 0,\ B > 0$.

б) При $A + B < 0$, основываясь на (2.14), (2.19) и (2.40), получим

$$(A+B)^2 = \begin{pmatrix} (\mu_a + \mu_b)^2, -2(\mu_a + \mu_b)(\beta_a + \beta_b) - (\beta_a + \beta_b)^2, \\ -2(\mu_a + \mu_b)(\alpha_a + \alpha_b) + (\alpha_a + \alpha_b)^2 \end{pmatrix} =$$

$$= \left(\mu_a^2, 2\mu_a\alpha_a - \alpha_a^2, 2\mu_a\beta_a + \beta_a^2\right) + 2\begin{pmatrix} \mu_a\mu_b, -\mu_a\beta_b - \mu_b\beta_a - \beta_a\beta_b, \\ -\mu_a\alpha_b - \mu_b\alpha_a + \alpha_a\alpha_b \end{pmatrix} +$$

$$+ \left(\mu_b^2, 2\mu_b\alpha_b - \alpha_b^2, 2\mu_b\beta_b + \beta_b^2\right) = A^2 + 2AB + B^2,$$

при условии, что $A < 0$ и $B < 0$.

Выводы по главе 2

1. Нечеткие числа *L*-*R*-типа могут быть использованы для формализации неопределенности.

2. Результат операции над нечеткими числами *L*-*R*-типа также принадлежит множеству нечетких чисел *L*-*R*-типа.

3. Нечеткий группоид $\langle Fuz^{LR}(\Re),*\rangle$ обладает свойствами коммутативности и ассоциативности, имеет единичный элемент $E=(0,0,0)$, если * – операция сложения, и единичный элемент $E=(1,0,0)$, если * – операция умножения. При этом нечеткий группоид $\langle Fuz^{LR}(\Re)\setminus S,\times\rangle$ имеет обратный элемент $\overline{A}=\left(\dfrac{-1}{\mu_a},\dfrac{\beta_a}{\mu_a(\mu_a+\beta_a)},\dfrac{\alpha_a}{\mu_a(\mu_a-\alpha_a)}\right)$ и обладает свойством дистрибутивности на множестве нечетких чисел *L*-*R*-типа.

4. Для группоида нечетких чисел *L*-*R*-типа могут быть определены некоторые алгебраические структуры с одной операцией, но, ввиду отсутствия обратного элемента у группоида с операцией сложения, алгебраические структуры с двумя операциями не могут быть выявлены.

5. Получены формулы для нахождения инволюции, произведения нечеткого числа на нечеткий ноль и единицу, формула квадрата нечеткого числа, а также произведения суммы двух нечетких чисел и произведения квадрата нечеткого числа на обратное ему число, которые позволяют осуществлять вычисления при проведении нечеткого регрессионного анализа.

Глава 3. Разработка нечетких регрессионных моделей для восстановления закономерностей в данных, содержащих приближенную информацию

В данной главе рассмотрена задача оценки параметров нечеткой линейной парной и множественной регрессионных моделей, в которых учитываются различные типы данных в выборке: а) четкие независимые переменные, нечеткие коэффициенты и зависимая величина, б) нечеткие зависимая и независимые переменные модели с четкими коэффициентами; проведена оценка адекватности построенных моделей на основе вычисления нечеткого коэффициента корреляции и его интерпретации, анализа остаточной последовательности ошибок, а также точности моделей путем вычисления средней относительной ошибки; предложены методы отбора существенных независимых переменных на основе построения стандартизированного уравнения регрессии и использования автоассоциативной нейронной сети.

3.1. Нечеткая парная линейная регрессионная модель

3.1.1. Оценка параметров нечеткой парной линейной регрессионной модели

Рассмотрим метод оценки неизвестных параметров нечеткой линейной парной регрессионной модели. Положим, что исходные данные состоят из выборки $\{(x_i, Y_i)\}_{i=\overline{1,n}}$, где $x_i \in \Re$, а $Y_i = (\mu_{yi}, \alpha_{yi}, \beta_{yi})$ – нечеткие числа *L*-*R*-типа [67].

$$Y^{мод} = B_0 + B_1 x, \qquad (3.1)$$

где x – независимая (объясняющая) переменная, $x \in \Re$; Y – зависимая (объясняемая) переменная, $Y = (\mu_y, \alpha_y, \beta_y)$ – нечеткое число *L*-*R*-типа; B_0, B_1 –

неизвестные коэффициенты регрессии (параметры уравнения), $B_0=(\mu_{b0},\alpha_{b0},\beta_{b0})$ и $B_1=(\mu_{b1},\alpha_{b1},\beta_{b1})$ – нечеткие числа *L*-*R*-типа.

Пусть n раз измерены значения переменной x и соответствующие значения Y. Тогда получим следующую линейную парную регрессионную модель с нечеткими коэффициентами (теоретическое линейное уравнение регрессии) [74, 76]:

$$Y_i = B_0 + B_1 x_i + E_i,\, i=\overline{1,n}, \tag{3.2}$$

где $x_i \in \Re$, $Y_i=(\mu_{yi},\alpha_{yi},\beta_{yi})$ – нечеткие числа *L*-*R*-типа, $B_0=(\mu_{b0},\alpha_{b0},\beta_{b0})$ и $B_1=(\mu_{b1},\alpha_{b1},\beta_{b1})$ – теоретические коэффициенты (параметры) регрессии, $E_i=(\mu_{ei},\alpha_{ei},\beta_{ei})$ – случайные ошибки (отклонения), нечеткие числа *L*-*R*-типа, $i=\overline{1,n}$ – номер наблюдения.

Для получения наиболее подходящих оценок $\tilde{B}_0,\tilde{B}_1$ регрессионных параметров B_0,B_1 применим метод наименьших квадратов с использованием формулы расстояния (2.13) между двумя нечеткими числами [76, 79]. Следует учитывать при этом, что $Y_i=(\mu_{yi},\alpha_{yi},\beta_{yi})$, $i=\overline{1,n}$, и $B_0=(\mu_{b0},\alpha_{b0},\beta_{b0})$ и $B_1=(\mu_{b1},\alpha_{b1},\beta_{b1})$ имеют одинаковую функцию принадлежности, и после надлежащего преобразования данных можно сделать все $x_i>0$, $i=\overline{1,n}$.

После подстановки оценок параметров $\tilde{B}_0,\tilde{B}_1$ в уравнение регресии (3.2) приходим к следующей эмпирической (оценочной) нечеткой линейной парной регрессионной модели [74]:

$$\tilde{Y}_i = \tilde{B}_0 + \tilde{B}_1 x_i,\, i=\overline{1,n}, \tag{3.3}$$

где $x_i \in \Re$ – наблюдаемые значения независимой (объясняющей) переменной, $\tilde{Y}_i=(\tilde{\mu}_{yi},\tilde{\alpha}_{yi},\tilde{\beta}_{yi})$ – оценки значений зависимой (объясняемой) переменной, $\tilde{B}_0=(\tilde{\mu}_{b0},\tilde{\alpha}_{b0},\tilde{\beta}_{b0})$ и $\tilde{B}_1=(\tilde{\mu}_{b1},\tilde{\alpha}_{b1},\tilde{\beta}_{b1})$ – оценки неизвестных параметров B_0,B_1, называемые эмпирическими (выборочными) коэффициентами регрессии.

В соответствии с методом наименьших квадратов оценки параметров регресии находятся из решения задачи минимизации функции [74, 76]:

$$F\left(\widetilde{B}_0,\widetilde{B}_1\right)=\sum_{i=1}^{n}D^2\left(Y_i,\widetilde{Y}_i\right)\to \min , \qquad (3.4)$$

где $Y_i=\left(\mu_{yi},\alpha_{yi},\beta_{yi}\right)$ и $\widetilde{Y}_i=\left(\widetilde{\mu}_{yi},\widetilde{\alpha}_{yi},\widetilde{\beta}_{yi}\right)$ вычисляются по формулам (3.2), (3.3), $D\left(Y_i,\widetilde{Y}_i\right)$ – функция расстояния между нечеткими переменными Y_i и $\widetilde{Y}_i$ (формула (2.13)).

Согласно формулам (2.11), (2.13) и (2.14) и с учетом коммутативности (2.21) целевую функцию можно записать в виде:

$$F\left(\widetilde{B}_0,\widetilde{B}_1\right)=\sum_{i=1}^{n}\left(\widetilde{B}_0+\widetilde{B}_1x_i-Y_i\right)^2=$$

$$=\sum_{i=1}^{n}\left(\left(\widetilde{\mu}_{b0}+\widetilde{\mu}_{b1}x_i-\mu_{yi}\right)^2+\left(\widetilde{\alpha}_{b0}+\widetilde{\alpha}_{b1}x_i-\alpha_{yi}\right)^2+\left(\widetilde{\beta}_{b0}+\widetilde{\beta}_{b1}x_i-\beta_{yi}\right)^2\right)\to \min .$$

Нетрудно заметить, что функция F является квадратичной функцией двух параметров $\widetilde{B}_0=\left(\widetilde{\mu}_{b0},\widetilde{\alpha}_{b0},\widetilde{\beta}_{b0}\right)$ и $\widetilde{B}_1=\left(\widetilde{\mu}_{b1},\widetilde{\alpha}_{b1},\widetilde{\beta}_{b1}\right)$ ($F=F\left(\widetilde{B}_0,\widetilde{B}_1\right)$), поскольку $\left\{\left(x_i,Y_i\right)\right\}_{i=\overline{1,n}}$ – известные данные наблюдений. Так как функция F непрерывна, выпукла и ограничена снизу ($F>0$), то она имеет минимум [74].

Необходимым условием существования минимума функции (3.4) является равенство нулю ее частных производных по неизвестным величинам $\widetilde{\mu}_{b0},\widetilde{\alpha}_{b0},\widetilde{\beta}_{b0}$ и $\widetilde{\mu}_{b1},\widetilde{\alpha}_{b1},\widetilde{\beta}_{b1}$ [79]:

$$\begin{cases}\dfrac{\partial F}{\partial \widetilde{\mu}_{b0}}=2\sum\limits_{i=1}^{n}\left(\widetilde{\mu}_{b0}+\widetilde{\mu}_{b1}x_i-\mu_{yi}\right)=0,\\ \dfrac{\partial F}{\partial \widetilde{\alpha}_{b0}}=2\sum\limits_{i=1}^{n}\left(\widetilde{\alpha}_{b0}+\widetilde{\alpha}_{b1}x_i-\alpha_{yi}\right)=0,\\ \dfrac{\partial F}{\partial \widetilde{\beta}_{b0}}=2\sum\limits_{i=1}^{n}\left(\widetilde{\beta}_{b0}+\widetilde{\beta}_{b1}x_i-\beta_{yi}\right)=0.\end{cases} \qquad (3.5)$$

$$\begin{cases}\dfrac{\partial F}{\partial \widetilde{\mu}_{b1}}=2\sum\limits_{i=1}^{n}\left(\widetilde{\mu}_{b0}+\widetilde{\mu}_{b1}x_i-\mu_{yi}\right)x_i=0,\\ \dfrac{\partial F}{\partial \widetilde{\alpha}_{b1}}=2\sum\limits_{i=1}^{n}\left(\widetilde{\alpha}_{b0}+\widetilde{\alpha}_{b1}x_i-\alpha_{yi}\right)x_i=0,\\ \dfrac{\partial F}{\partial \widetilde{\beta}_{b1}}=2\sum\limits_{i=1}^{n}\left(\widetilde{\beta}_{b0}+\widetilde{\beta}_{b1}x_i-\beta_{yi}\right)x_i=0.\end{cases} \qquad (3.6)$$

После преобразований получим системы нормальных уравнений для определения параметров нечеткой линейной парной регрессии [76, 79]:

$$\begin{cases} n\tilde{\mu}_{b0} + \tilde{\mu}_{b1}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu_{yi}, \\ n\tilde{\alpha}_{b0} + \tilde{\alpha}_{b1}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \alpha_{yi}, \\ n\tilde{\beta}_{b0} + \tilde{\beta}_{b1}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \beta_{yi}. \end{cases} \qquad (3.7)$$

$$\begin{cases} \tilde{\mu}_{b0}\sum_{i=1}^{n} x_i + \tilde{\mu}_{b1}\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i\mu_{yi}, \\ \tilde{\alpha}_{b0}\sum_{i=1}^{n} x_i + \tilde{\alpha}_{b1}\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i\alpha_{yi}, \\ \tilde{\beta}_{b0}\sum_{i=1}^{n} x_i + \tilde{\beta}_{b1}\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i\beta_{yi}. \end{cases} \qquad (3.8)$$

Теперь, разделив обе части уравнений систем (3.7) и (3.8) на n, получим следующие системы нормальных уравнений [74, 76]:

$$\begin{cases} \tilde{\mu}_{b0} + \tilde{\mu}_{b1}\bar{x} = \bar{\mu}_y, \\ \tilde{\alpha}_{b0} + \tilde{\alpha}_{b1}\bar{x} = \bar{\alpha}_y, \\ \tilde{\beta}_{b0} + \tilde{\beta}_{b1}\bar{x} = \bar{\beta}_y, \end{cases} \qquad (3.9)$$

$$\begin{cases} \tilde{\mu}_{b0}\bar{x} + \tilde{\mu}_{b1}\overline{x^2} = \overline{x\mu_y}, \\ \tilde{\alpha}_{b0}\bar{x} + \tilde{\alpha}_{b1}\overline{x^2} = \overline{x\alpha_y}, \\ \tilde{\beta}_{b0}\bar{x} + \tilde{\beta}_{b1}\overline{x^2} = \overline{x\beta_y}, \end{cases} \qquad (3.10)$$

где соответствующие средние определяются по формулам [74, 76]:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \qquad (3.11)$$

$$\overline{x^2} = \frac{\sum_{i=1}^{n} x_i^2}{n}, \qquad (3.12)$$

$$\overline{\mu_y} = \frac{\sum_{i=1}^{n} \mu_{yi}}{n}, \; \overline{\alpha_y} = \frac{\sum_{i=1}^{n} \alpha_{yi}}{n}, \; \overline{\beta_y} = \frac{\sum_{i=1}^{n} \beta_{yi}}{n}, \qquad (3.13)$$

$$\overline{x\mu_y} = \frac{\sum_{i=1}^{n} x_i \mu_{yi}}{n}, \ \overline{x\alpha_y} = \frac{\sum_{i=1}^{n} x_i \alpha_{yi}}{n}, \ \overline{x\beta_y} = \frac{\sum_{i=1}^{n} x_i \beta_{yi}}{n}. \tag{3.14}$$

Выразим $\tilde{\mu}_{b0}, \tilde{\alpha}_{b0}, \tilde{\beta}_{b0}$ в системе (3.9) через $\tilde{\mu}_{b1}, \tilde{\alpha}_{b1}, \tilde{\beta}_{b1}$.

$$\begin{cases} \tilde{\mu}_{b0} = \bar{\mu}_y - \tilde{\mu}_{b1}\bar{x}, \\ \tilde{\alpha}_{b0} = \bar{\alpha}_y - \tilde{\alpha}_{b1}\bar{x}, \\ \tilde{\beta}_{b0} = \bar{\beta}_y - \tilde{\beta}_{b1}\bar{x}. \end{cases} \tag{3.15}$$

Подставив выражения из системы (3.15) в уравнения (3.10), получим [74]:

$$\begin{cases} \left(\bar{\mu}_y - \tilde{\mu}_{b1}\bar{x}\right)\bar{x} + \tilde{\mu}_{b1}\overline{x^2} = \overline{x\mu_y}, \\ \left(\bar{\alpha}_y - \tilde{\alpha}_{b1}\bar{x}\right)\bar{x} + \tilde{\alpha}_{b1}\overline{x^2} = \overline{x\alpha_y}, \\ \left(\bar{\beta}_y - \tilde{\beta}_{b1}\bar{x}\right)\bar{x} + \tilde{\beta}_{b1}\overline{x^2} = \overline{x\beta_y}. \end{cases} \tag{3.16}$$

Решая систему (3.16), находим

$$\begin{cases} \tilde{\mu}_{b1} = \dfrac{\overline{x\mu_y} - \bar{\mu}_y\bar{x}}{\overline{x^2} - \bar{x}^2} = \dfrac{\operatorname{cov}(x, \mu_y)}{\sigma_x^2}, \\ \tilde{\alpha}_{b1} = \dfrac{\overline{x\alpha_y} - \bar{\alpha}_y\bar{x}}{\overline{x^2} - \bar{x}^2} = \dfrac{\operatorname{cov}(x, \alpha_y)}{\sigma_x^2}, \\ \tilde{\beta}_{b1} = \dfrac{\overline{x\beta_y} - \bar{\beta}_y\bar{x}}{\overline{x^2} - \bar{x}^2} = \dfrac{\operatorname{cov}(x, \beta_y)}{\sigma_x^2}, \end{cases} \tag{3.17}$$

где σ_x^2 – выборочная дисперсия переменной x:

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = \left(\sum_{i=1}^{n} x_i^2\right)/n - \left(\left(\sum_{i=1}^{n} x_i\right)/n\right)^2, \tag{3.18}$$

$\operatorname{cov}(x, \mu_y)$, $\operatorname{cov}(x, \alpha_y)$, $\operatorname{cov}(x, \beta_y)$ – выборочные ковариации [74]:

$$\begin{cases} \operatorname{cov}(x, \mu_y) = \overline{x\mu_y} - \bar{\mu}_y\bar{x} = \dfrac{\sum_{i=1}^{n} x_i\mu_{yi}}{n} - \dfrac{\sum_{i=1}^{n} \mu_{yi}}{n} \cdot \dfrac{\sum_{i=1}^{n} x_i}{n}, \\ \operatorname{cov}(x, \alpha_y) = \overline{x\alpha_y} - \bar{\alpha}_y\bar{x} = \dfrac{\sum_{i=1}^{n} x_i\alpha_{yi}}{n} - \dfrac{\sum_{i=1}^{n} \alpha_{yi}}{n} \cdot \dfrac{\sum_{i=1}^{n} x_i}{n}, \\ \operatorname{cov}(x, \beta_y) = \overline{x\beta_y} - \bar{\beta}_y\bar{x} = \dfrac{\sum_{i=1}^{n} x_i\beta_{yi}}{n} - \dfrac{\sum_{i=1}^{n} \beta_{yi}}{n} \cdot \dfrac{\sum_{i=1}^{n} x_i}{n}. \end{cases} \tag{3.19}$$

Пример 3.1. По данным таблицы 3.1 найти уравнение регрессии Y по x.

Таблица 3.1 – Исходные данные

x	11	12	13	14	15	16	17	18	19	20	21	22	23	24
μ_y	11,2	12,5	12,9	14,1	14,8	16,1	17,5	18,9	18,9	20	21,1	22,2	23,3	24,4
α_y	0,9	0,1	1	0,5	1,1	1,2	0,1	0,5	0,3	1,3	1,1	0,2	1,44	0,7
β_y	0,2	0,4	0,5	1,1	0,7	0,1	0,08	1,2	0,7	0,48	1,9	0,4	0,8	0,71

Решение. Вычислим все необходимые суммы (табл. 3.2):

Таблица 3.2 – Результаты вычислений

$\sum_{i=1}^{14} x_i$	$\sum_{i=1}^{14} x_i^2$	$\sum_{i=1}^{14} \mu_{yi}$	$\sum_{i=1}^{14} \alpha_{yi}$	$\sum_{i=1}^{14} \beta_{yi}$	$\sum_{i=1}^{14} x_i \mu_{yi}$	$\sum_{i=1}^{14} x_i \alpha_{yi}$	$\sum_{i=1}^{14} x_i \beta_{yi}$
245	4515	247,9	10,44	9,27	4567,7	186,62	171

Затем по формулам (3.11)-(3.19) найдем выборочные характеристики уравнения регрессии:

$$\bar{x} = \frac{245}{14} = 17{,}5; \quad \overline{x^2} = \frac{1415}{14} = 322{,}5;$$

$$\overline{\mu_y} = \frac{247{,}9}{14} = 17{,}7; \quad \overline{\alpha_y} = \frac{10{,}44}{14} = 0{,}75, \quad \overline{\beta_y} = \frac{9{,}27}{14} = 0{,}662;$$

$$\overline{x\mu_y} = \frac{4567{,}7}{14} = 326{,}3; \quad \overline{x\alpha_y} = \frac{186{,}62}{14} = 13{,}3; \quad \overline{x\beta_y} = \frac{171}{14} = 12{,}2;$$

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 322{,}5 - 306{,}25 = 16{,}25;$$

$$\begin{cases} \operatorname{cov}(x, \mu_y) = \overline{x\mu_y} - \bar{\mu}_y \bar{x} = 326{,}3 - 17{,}7 \cdot 17{,}5 = 16{,}55, \\ \operatorname{cov}(x, \alpha_y) = \overline{x\alpha_y} - \bar{\alpha}_y \bar{x} = 13{,}3 - 0{,}75 \cdot 17{,}5 = 0{,}175, \\ \operatorname{cov}(x, \beta_y) = \overline{x\beta_y} - \bar{\beta}_y \bar{x} = 12{,}2 - 0{,}662 \cdot 17{,}5 = 0{,}615. \end{cases}$$

Теперь вычислим параметры уравнения регрессии:

$$\begin{cases} \tilde{\mu}_{b1} = \dfrac{\operatorname{cov}(x, \mu_y)}{\sigma_x^2} = \dfrac{16{,}55}{16{,}25} = 1{,}02, \\ \tilde{\alpha}_{b1} = \dfrac{\overline{x\alpha_y} - \overline{\alpha}_y \overline{x}}{\overline{x^2} - \overline{x}^2} = \dfrac{0{,}175}{16{,}25} = 0{,}01, \\ \tilde{\beta}_{b1} = \dfrac{\overline{x\beta_y} - \overline{\beta}_y \overline{x}}{\overline{x^2} - \overline{x}^2} = \dfrac{0{,}615}{16{,}25} = 0{,}037. \end{cases}$$

$$\begin{cases} \tilde{\mu}_{b0} = \overline{\mu}_y - \tilde{\mu}_{b1}\overline{x} = 17{,}7 - 1{,}02 \cdot 17{,}5 = -0{,}15, \\ \tilde{\alpha}_{b0} = \overline{\alpha}_y - \tilde{\alpha}_{b1}\overline{x} = 0{,}75 - 0{,}01 \cdot 17{,}5 = 0{,}575, \\ \tilde{\beta}_{b0} = \overline{\beta}_y - \tilde{\beta}_{b1}\overline{x} = 0{,}662 - 0{,}037 \cdot 17{,}5 = 0{,}0145. \end{cases}$$

Итак, уравнение регрессии Y по x:

$$\tilde{Y} = (-0.15, 0.575, 0.0145) + (1.02, 0.01, 0.037)x.$$

3.1.2. Оценка качества нечеткой парной линейной регрессионной модели

Качество построенной нечеткой парной линейной регрессионной модели оценивается по точности и адекватности.

Адекватность модели устанавливается на основе анализа остаточной последовательности $\{E_i = Y_i - \tilde{Y}_i, i = \overline{1,n}\}$, которая проверяется на выполнение свойств случайной компоненты: близость нулю математического ожидания, случайный характер отклонений, отсутствие автокорреляции и нормальность закона распределения (условия Гаусса-Маркова [3, 87]). Эту проверку будем проводить теми же методами, что и в случае четкого парного регрессионного анализа. При этом для перехода от нечетких величин $E_i = (\mu_{ei}, \alpha_{ei}, \beta_{ei})$ к четким e_i, $i = \overline{1,n}$, применим следующие подходы:

1) дефаззификацию на основе результирующей функции принадлежности $m_{ei}(x)$ нечеткого числа $E_i = (\mu_{ei}, \alpha_{ei}, \beta_{ei})$ при заданных *L*-*R*-функциях (формула 2.8), которая вычисляет четкое числовое значение погрешности e_i. Данная операция выполняется посредством механизма дефаззификации, который

определяет метод вычисления [67]. Примером может служить метод центра тяжести [51]:

$$e_i = Defuz(E_i) = D(m_{ei}) = \frac{\int\limits_{\min}^{\max} x \cdot m_{ei}(x)dx}{\int\limits_{\min}^{\max} m_{ei}(x)dx}, \text{ где } \min = \mu_{ei} - \alpha_{ei}, \max = \mu_{ei} + \beta_{ei}.$$

2) дефаззификацию на основе результирующих функций принадлежности $m_{yi}(x)$ и $m_{\tilde{y}i}(x)$ нечетких чисел $Y_i = (\mu_{yi}, \alpha_{yi}, \beta_{yi})$ и $\tilde{Y}_i = (\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}, \tilde{\beta}_{yi})$ соответственно при заданных *L*-*R*-функциях (формула 2.8). Данная операция позволяет вычислить точные значения зависимой переменной и ее оценок посредством одного из методов дефаззификации [51, 67] и, тем самым, найти погрешность e_i через разность этих двух значений:

$$e_i = Defuz(Y_i) - Defuz(\tilde{Y}_i) = D(m_{yi}) - D(m_{\tilde{y}i}).$$

3) построение γ-срезов (формулы (2.6), (2.7)) по значениям зависимой переменной $Y_i = (\mu_{yi}, \alpha_{yi}, \beta_{yi})$ и ее оценок $\tilde{Y}_i = (\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}, \tilde{\beta}_{yi})$:

$$\left[\underline{Y_i}(\gamma), \overline{Y_i}(\gamma)\right] = [\mu_y - \alpha_y(1-\gamma), \mu_y + \beta_y(1-\gamma)],\ \gamma \in [0,1],$$

$$\left[\underline{\tilde{Y}_i}(\gamma), \overline{\tilde{Y}_i}(\gamma)\right] = [\tilde{\mu}_y - \tilde{\alpha}_y(1-\gamma), \tilde{\mu}_y + \tilde{\beta}_y(1-\gamma)],\ \gamma \in [0,1],$$

и составление остаточной последовательности для каждого γ-среза через определение ε-точки:

$$\varepsilon \in \left[\underline{Y_i}(\gamma), \overline{Y_i}(\gamma)\right] \Rightarrow \lambda_i^\gamma = \underline{Y_i}(\gamma) \cdot \varepsilon + \overline{Y_i}(\gamma) \cdot (1-\varepsilon),\ \gamma \in [0,1],$$

$$\varepsilon \in \left[\underline{\tilde{Y}_i}(\gamma), \overline{\tilde{Y}_i}(\gamma)\right] \Rightarrow \tilde{\lambda}_i^\gamma = \underline{\tilde{Y}_i}(\gamma) \cdot \varepsilon + \overline{\tilde{Y}_i}(\gamma) \cdot (1-\varepsilon),\ \gamma \in [0,1].$$

В результате четкие значения погрешности e_i, $i = \overline{1,n}$ будем вычислять по формуле:

$$e_i = \lambda_i^\gamma - \tilde{\lambda}_i^\gamma,\ i = \overline{1,n},\ \gamma \in [0,1].$$

О качестве модели нечеткой парной регрессии можно судить также по значениям коэффициента корреляции, который показывает тесноту линейной связи между зависимой и независимой переменными.

Перейдем к оценке тесноты корреляционной зависимости. Подставим полученные в (3.15) значения модальной величины и левого и правого коэффициентов нечеткости параметра $\tilde{B}_0$ в уравнение регресии (3.3). Тогда получим [74]:

$$\left(\tilde{\mu}_y, \tilde{\alpha}_y, \tilde{\beta}_y\right) = \left(\tilde{\mu}_{b0}, \tilde{\alpha}_{b0}, \tilde{\beta}_{b0}\right) + \left(\tilde{\mu}_{b1}, \tilde{\alpha}_{b1}, \tilde{\beta}_{b1}\right)x$$

или более подробно

$$\left(\tilde{\mu}_y, \tilde{\alpha}_y, \tilde{\beta}_y\right) = \left(\overline{\mu}_y - \tilde{\mu}_{b1}\overline{x}, \overline{\alpha}_y - \tilde{\alpha}_{b1}\overline{x}, \overline{\beta}_y - \tilde{\beta}_{b1}\overline{x}\right) + \left(\tilde{\mu}_{b1}, \tilde{\alpha}_{b1}, \tilde{\beta}_{b1}\right)x.$$

Воспользовавшись формулой сложения нечетких чисел (2.14), имеем:

$$\left(\tilde{\mu}_y, \tilde{\alpha}_y, \tilde{\beta}_y\right) = \left(\overline{\mu}_y - \tilde{\mu}_{b1}\overline{x} + \tilde{\mu}_{b1}, \overline{\alpha}_y - \tilde{\alpha}_{b1}\overline{x} + \tilde{\alpha}_{b1}, \overline{\beta}_y - \tilde{\beta}_{b1}\overline{x} + \tilde{\beta}_{b1}\right) =$$
$$= \left(\overline{\mu}_y + \tilde{\mu}_{b1}(x - \overline{x}), \overline{\alpha}_y + \tilde{\alpha}_{b1}(x - \overline{x}), \overline{\beta}_y + \tilde{\beta}_{b1}(x - \overline{x})\right).$$

Таким образом приходим к следующей системе уравнений:

$$\begin{cases} \tilde{\mu}_y = \overline{\mu}_y + \tilde{\mu}_{b1}(x - \overline{x}), \\ \tilde{\alpha}_y = \overline{\alpha}_y + \tilde{\alpha}_{b1}(x - \overline{x}), \\ \tilde{\beta}_y = \overline{\beta}_y + \tilde{\beta}_{b1}(x - \overline{x}). \end{cases} \tag{3.20}$$

Представим уравнения системы (3.20) в эквивалентном виде [74]:

$$\begin{cases} \dfrac{\tilde{\mu}_y - \overline{\mu}_y}{\sigma_\mu} = \tilde{\mu}_{b1} \dfrac{\sigma_x}{\sigma_\mu} \dfrac{(x - \overline{x})}{\sigma_x}, \\ \dfrac{\tilde{\alpha}_y - \overline{\alpha}_y}{\sigma_\alpha} = \tilde{\alpha}_{b1} \dfrac{\sigma_x}{\sigma_\alpha} \dfrac{(x - \overline{x})}{\sigma_x}, \\ \dfrac{\tilde{\beta}_y - \overline{\beta}_y}{\sigma_\beta} = \tilde{\beta}_{b1} \dfrac{\sigma_x}{\sigma_\beta} \dfrac{(x - \overline{x})}{\sigma_x}, \end{cases} \tag{3.21}$$

где σ_x^2 – выборочная дисперсия переменной x, определяемая по формуле (3.18), а $\sigma_\mu^2, \sigma_\alpha^2, \sigma_\beta^2$ – выборочные дисперсии величин μ_y, α_y, β_y [74]:

$$\sigma_\mu^2 = \overline{\mu_y^2} - \overline{\mu_y}^2 = \frac{\sum_{i=1}^{n} \mu_{yi}^2}{n} - \left(\frac{\sum_{i=1}^{n} \mu_{yi}}{n}\right)^2, \tag{3.22}$$

$$\sigma_\alpha^2 = \overline{\alpha_y^2} - \overline{\alpha_y}^2 = \frac{\sum_{i=1}^{n} \alpha_{yi}^2}{n} - \left(\frac{\sum_{i=1}^{n} \alpha_{yi}}{n} \right)^2, \tag{3.23}$$

$$\sigma_\beta^2 = \overline{\beta_y^2} - \overline{\beta_y}^2 = \frac{\sum_{i=1}^{n} \beta_{yi}^2}{n} - \left(\frac{\sum_{i=1}^{n} \beta_{yi}}{n} \right)^2. \tag{3.24}$$

Из системы (3.21) следует, что величина

$$R = \left(\tilde{\mu}_{b1} \frac{\sigma_x}{\sigma_\mu}, \tilde{\alpha}_{b1} \frac{\sigma_x}{\sigma_\alpha}, \tilde{\beta}_{b1} \frac{\sigma_x}{\sigma_\beta} \right) \tag{3.25}$$

показывает , на сколько величин $\sigma_\mu, \sigma_\alpha, \sigma_\beta$ изменится в среднем Y, когда x увеличится на величину σ_x [41, 42].

Величина R показывает тесноту связи между переменными Y и x и является выборочным коэффициентом корреляции (или просто коэффициентом корреляции) [64, 89].

Для нахождения четкого значения r коэффициента корреляции R могут быть использованы следующие методы:

1) дефаззификация на основе результирующей функции принадлежности $m_r(x)$ нечеткого коэффициента корреляции $R = (\mu_r, \alpha_r, \beta_r)$ при заданных *L*-*R*-функциях (формула 2.8). Данная операция может быть выполнена с помощью метода центра тяжести [51]:

$$r = Defuz(R) = D(m_r) = \frac{\int_{\min}^{\max} x \cdot m_r(x) dx}{\int_{\min}^{\max} m_r(x) dx}, \text{ где } \min = \mu_r - \alpha_r, \max = \mu_r + \beta_r.$$

2) построение γ-среза по значениям нечеткого коэффициента корреляции $R = (\mu_r, \alpha_r, \beta_r)$:

$$\left[\underline{R}(\gamma), \overline{R}(\gamma) \right] = [\mu_r - \alpha_r(1-\gamma), \mu_r + \beta_r(1-\gamma)], \ \gamma \in [0,1],$$

и вычисление четкого коэффициента корреляции для каждого γ-среза через определение ε-точки:

$$\varepsilon \in \left[\underline{R}(\gamma), \overline{R}(\gamma)\right] \Rightarrow r = \underline{R}(\gamma) \cdot \varepsilon + \overline{R}(\gamma) \cdot (1-\varepsilon),\ \gamma \in [0,1].$$

Чем ближе абсолютная величина полученного четкого коэффициента r к 1, тем теснее связь между переменными Y и x и, следовательно, с тем большей уверенностью можно судить об адекватности построенной модели.

Учитывая (3.17), формулу для R представим в виде:

$$R = \left(\frac{\operatorname{cov}(x, \mu_y)}{\sigma_x \sigma_\mu}, \frac{\operatorname{cov}(x, \alpha_y)}{\sigma_x \sigma_\alpha}, \frac{\operatorname{cov}(x, \beta_y)}{\sigma_x \sigma_\beta} \right). \qquad (3.26)$$

Отметим другие модификации формулы для R, полученные из формулы (3.26) с помощью формул (3.17)-(3.19) [74]:

$$R = \left(\frac{\sum_{i=1}^{n} (x_i - \overline{x})(\mu_{yi} - \overline{\mu_y})}{n\sigma_x \sigma_\mu}, \frac{\sum_{i=1}^{n} (x_i - \overline{x})(\alpha_{yi} - \overline{\alpha_y})}{n\sigma_x \sigma_\alpha}, \frac{\sum_{i=1}^{n} (x_i - \overline{x})(\beta_{yi} - \overline{\beta_y})}{n\sigma_x \sigma_\beta} \right), \qquad (3.27)$$

$$R = (\mu_r, \alpha_r, \beta_r), \qquad (3.28)$$

где μ_r, α_r, β_r вычисляются по формулам:

$$\mu_r = \frac{n\sum_{i=1}^{n} x_i \mu_{yi} - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} \mu_{yi}\right)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} \mu_{yi}^2 - \left(\sum_{i=1}^{n} \mu_{yi}\right)^2}},$$

$$\alpha_r = \frac{n\sum_{i=1}^{n} x_i \alpha_{yi} - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} \alpha_{yi}\right)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} \alpha_{yi}^2 - \left(\sum_{i=1}^{n} \alpha_{yi}\right)^2}},$$

$$\beta_r = \frac{n\sum_{i=1}^{n} x_i \beta_{yi} - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} \beta_{yi}\right)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\sum_{i=1}^{n} \beta_{yi}^2 - \left(\sum_{i=1}^{n} \beta_{yi}\right)^2}}.$$

Для практических расчетов наиболее удобна формула (3.28), так как по ней R находится непосредственно из данных наблюдений, и на значении R не скажутся округления данных, связанные с расчетом средних и отклонений от них.

Точность модели характеризует близость модельных и фактических значений по каждому наблюдению. Для характеристики степени близости будем использовать среднюю относительную ошибку

$$E_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \tilde{Y}_i}{Y_i}\right| \cdot 100\%,$$

вычисление которой может быть проведено с помощью следующих подходов:

1) дефаззификации на основе результирующих функций принадлежности $m_{yi}(x)$ и $m_{\tilde{y}i}(x)$ нечетких чисел $Y_i = (\mu_{yi}, \alpha_{yi}, \beta_{yi})$ и $\tilde{Y}_i = (\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}, \tilde{\beta}_{yi})$ соответственно при заданных *L*-*R*-функциях (формула 2.8):

$$e_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Defuz(Y_i) - Defuz(\tilde{Y}_i)}{Defuz(Y_i)}\right| \cdot 100\%.$$

2) построения γ-срезов по значениям зависимой переменной $Y_i = (\mu_{yi}, \alpha_{yi}, \beta_{yi})$ и ее оценок $\tilde{Y}_i = (\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}, \tilde{\beta}_{yi})$:

$$\left[\underline{Y}_i(\gamma), \overline{Y}_i(\gamma)\right] = [\mu_y - \alpha_y(1-\gamma), \mu_y + \beta_y(1-\gamma)],\ \gamma \in [0,1],$$

$$\left[\underline{\tilde{Y}}_i(\gamma), \overline{\tilde{Y}}_i(\gamma)\right] = [\tilde{\mu}_y - \tilde{\alpha}_y(1-\gamma), \tilde{\mu}_y + \tilde{\beta}_y(1-\gamma)],\ \gamma \in [0,1],$$

и нахождения средней относительной ошибки для каждого γ-среза через определение ε-точки:

$$\varepsilon \in \left[\underline{Y}_i(\gamma), \overline{Y}_i(\gamma)\right] \Rightarrow \lambda_i^{\gamma} = \underline{Y}_i(\gamma)\cdot\varepsilon + \overline{Y}_i(\gamma)\cdot(1-\varepsilon),\ \gamma \in [0,1],$$

$$\varepsilon \in \left[\underline{\tilde{Y}}_i(\gamma), \overline{\tilde{Y}}_i(\gamma)\right] \Rightarrow \tilde{\lambda}_i^{\gamma} = \underline{\tilde{Y}}_i(\gamma)\cdot\varepsilon + \overline{\tilde{Y}}_i(\gamma)\cdot(1-\varepsilon),\ \gamma \in [0,1].$$

В результате четкое значение ошибки $e_{отн}$ будем вычислять по формуле:

$$e_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\lambda_i^{\gamma} - \tilde{\lambda}_i^{\gamma}}{\lambda_i^{\gamma}}\right| \cdot 100\%,\ \gamma \in [0,1].$$

Чем ближе найденная средняя относительная ошибка к нулю, тем точнее построенная нечеткая линейная регрессионная модель. Если $e_{отн}$ не превосходит 15%, то точность модели считается приемлемой. Величина менее 5% свидетельствует о хорошем уровне точности.

Пример 3.2. По данным таблицы 3.1 вычислить коэффициент корреляции между переменными Y и x, провести оценку адекватности полученной модели.

Решение. В примере 3.1 были вычислены:

$$\sum_{i=1}^{14} x_i = 245; \ \sum_{i=1}^{14} x_i^2 = 4515; \ \sum_{i=1}^{14} \mu_{yi} = 247,9; \ \sum_{i=1}^{14} \alpha_{yi} = 10,44; \ \sum_{i=1}^{14} \beta_{yi} = 9,27;$$

$$\sum_{i=1}^{14} x_i \mu_{yi} = 4567,7; \ \sum_{i=1}^{14} x_i \alpha_{yi} = 186,62; \ \sum_{i=1}^{14} x_i \beta_{yi} = 171.$$

Найдем теперь суммы, необходимые для определения коэффициента корреляции (табл. 3.3):

Таблица 3.3 – Результаты вычислений

$\sum_{i=1}^{14} \mu_{yi}^2$	$\sum_{i=1}^{14} \alpha_{yi}^2$	$\sum_{i=1}^{14} \beta_{yi}^2$
4622,13	10,5736	9,2409

По формуле (3.28) вычислим коэффициент корреляции между переменными Y и x.

$$\mu_r = \frac{14 \cdot 4567,7 - 245 \cdot 247,9}{\sqrt{14 \cdot 4515 - 245^2} \cdot \sqrt{14 \cdot 4622,13 - 247,9^2}} = \frac{3212,3}{3220,01} = 0,998,$$

$$\alpha_r = \frac{14 \cdot 186,62 - 245 \cdot 10,44}{\sqrt{14 \cdot 4515 - 245^2} \cdot \sqrt{14 \cdot 10,5736 - 10,44^2}} = \frac{54,88}{352,605} = 0,156,$$

$$\beta_r = \frac{14 \cdot 171 - 245 \cdot 9,27}{\sqrt{14 \cdot 4515 - 245^2} \cdot \sqrt{14 \cdot 9,2409 - 9,27^2}} = \frac{122,85}{371,91} = 0,33,$$

т. е. $R = (0.998, 0.156, 0.33)$.

Для оценки адекватности полученной в примере 3.1 модели построим γ-срез по значениям найденного нечеткого коэффициента корреляции $R = (0.998, 0.156, 0.33)$:

$$\left[\underline{R}(\gamma), \overline{R}(\gamma)\right] = [0.998 - 0.156 \cdot (1-\gamma), 0.998 + 0.33 \cdot (1-\gamma)], \ \gamma \in [0,1],$$

и вычислим четкий коэффициент корреляции для каждого γ-среза, взяв за ε-точку середину отрезка $\left[\underline{R}(\gamma), \overline{R}(\gamma)\right]$:

$$r = \underline{R}(\gamma)\cdot\varepsilon + \overline{R}(\gamma)\cdot(1-\varepsilon), \varepsilon = \frac{\underline{R}(\gamma)+\overline{R}(\gamma)}{2},\ \gamma\in[0,1].$$

При $\gamma\in[0,1]$ $r\in[0.795,0.988]$, что свидетельствует о тесной линейной связи между переменными Y и x и адекватности построенной модели.

3.2. Нечеткая линейная множественная регрессионная модель

3.2.1. Оценка параметров нечеткой линейной множественной регрессионной модели. Адекватность и точность модели

Исследуем зависимость переменной Y, представляющей собой нечеткое число *L*-*R*-типа, от объясняющих переменных $x_1, x_2, \ldots, x_n$, являющихся вещественными числами. Пусть i-е наблюдение зависимой переменной равно Y_i, а независимых – $x_{i1}, x_{i2}, \ldots, x_{in}$. Рассмотрим следующую обобщенную нечеткую линейную регрессионную модель [52]:

$$Y_i = A_0 + A_1 x_{i1} + A_2 x_{i2} + \ldots + A_n x_{in} + E_i, i = \overline{1,m}, \tag{3.29}$$

где $x_{ik}\in R, k=\overline{1,n}$, $Y_i=(\mu_{yi},\alpha_{yi},\beta_{yi})$, $i=\overline{1,m}$ – нечеткие числа с модой μ_{yi} и левым и правым коэффициентами нечеткости α_{yi} и β_{yi} соответственно; $A_j=(\mu_{aj},\alpha_{aj},\beta_{aj})$, $j=\overline{0,n}$ – нечеткие регрессионные параметры, имеющие функцию принадлежности, подобную Y_i; $E_i=(\mu_{ei},\alpha_{ei},\beta_{ei}), i=\overline{1,m}$ – случайные ошибки, нечеткие числа *L*-*R*-типа; $i=\overline{1,m}$ – номер наблюдения, $k=\overline{1,n}$ – номер независимой переменной.

Будем использовать метод наименьших квадратов для получения наиболее подходящих оценок нечетких регрессионных параметров A_j, $j=\overline{0,n}$ [72, 75].

Следует учитывать при этом, что $Y_i=(\mu_{yi},\alpha_{yi},\beta_{yi})$, $i=\overline{1,m}$, и $A_j=(\mu_{aj},\alpha_{aj},\beta_{aj})$, $j=\overline{0,n}$, имеют одинаковую функцию принадлежности, и после надлежащего преобразования данных можно сделать все $x_{ik}>0$, $i=\overline{1,m}$, $k=\overline{1,n}$.

Оценкой модели (3.29) является уравнение:

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \ldots + \tilde{A}_n x_{in}, i = \overline{1, m}, \tag{3.30}$$

где $\tilde{Y}_i = \left(\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}, \tilde{\beta}_{yi}\right)$, $i = \overline{1, m}$ – оценки зависимой переменной, $\tilde{A}_j = \left(\tilde{\mu}_{aj}, \tilde{\alpha}_{aj}, \tilde{\beta}_{aj}\right)$, $j = \overline{0, n}$ – оценки нечетких коэффициентов регрессии [52].

Так как $x_{ik} > 0$, $i = \overline{1, m}$, $k = \overline{1, n}$, за счет преобразования данных, то уравнение (3.30) может быть записано в виде:

$$\left(\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}, \tilde{\beta}_{yi}\right) = \left(\tilde{\mu}_{a0}, \tilde{\alpha}_{a0}, \tilde{\beta}_{a0}\right) + \left(\tilde{\mu}_{a1}, \tilde{\alpha}_{a1}, \tilde{\beta}_{a1}\right) x_{i1} + \ldots + \left(\tilde{\mu}_{an}, \tilde{\alpha}_{an}, \tilde{\beta}_{an}\right) x_{in},\ i = \overline{1, m}. \tag{3.31}$$

Согласно формуле расстояния Евклида оценкой параметров μ_{aj}, α_{aj} и β_{aj}, основанной на методе наименьших квадратов, являются величины $\tilde{\mu}_{aj}, \tilde{\alpha}_{aj}$ и $\tilde{\beta}_{aj}$, минимизирующие величину D^2 [72, 75], где

$$D^2 = \sum_{i=1}^{m} \left[\begin{array}{l} \left(\mu_{yi} - \left(\tilde{\mu}_{a0} + \tilde{\mu}_{a1} x_{i1} + \ldots + \tilde{\mu}_{an} x_{in}\right)\right)^2 + \left(\alpha_{yi} - \left(\tilde{\alpha}_{a0} + \tilde{\alpha}_{a1} x_{i1} + \ldots + \tilde{\alpha}_{an} x_{in}\right)\right)^2 + \\ + \left(\beta_{yi} - \left(\tilde{\beta}_{a0} + \tilde{\beta}_{a1} x_{i1} + \ldots + \tilde{\beta}_{an} x_{in}\right)\right)^2 \end{array} \right]. \tag{3.32}$$

Пусть $\|\vec{v}\|$ – длина вектора $\vec{v}$, тогда, используя векторные и матричные обозначения, величину D^2 можно переписать в виде $D^2 = \left\|X\tilde{\mu}_a - \mu_y\right\|^2 + \left\|X\tilde{\alpha}_a - \alpha_y\right\|^2 + \left\|X\tilde{\beta}_a - \beta_y\right\|^2$, где X – матрица объясняющих переменных размером $m \times (n+1)$, $\tilde{\mu}_a = (\tilde{\mu}_{a0}, \tilde{\mu}_{a1}, \ldots, \tilde{\mu}_{an})^T$, $\tilde{\alpha}_a = (\tilde{\alpha}_{a0}, \tilde{\alpha}_{a1}, \ldots, \tilde{\alpha}_{an})^T$, $\tilde{\beta}_a = (\tilde{\beta}_{a0}, \tilde{\beta}_{a1}, \ldots, \tilde{\beta}_{an})^T$, $\tilde{\mu}_y = (\tilde{\mu}_{y1}, \ldots, \tilde{\mu}_{ym})$, $\tilde{\alpha}_y = (\tilde{\alpha}_{y1}, \ldots, \tilde{\alpha}_{ym})$ и $\tilde{\beta}_y = (\tilde{\beta}_{y1}, \ldots, \tilde{\beta}_{ym})$. Тогда значения параметров $\tilde{\mu}_a, \tilde{\alpha}_a$ и $\tilde{\beta}_a$, минимизирующих величину D^2, находятся из условий [72, 75]:

$$\begin{cases} \dfrac{\partial D^2}{\partial \tilde{\mu}_a} = 0, \\ \dfrac{\partial D^2}{\partial \tilde{\alpha}_a} = 0, \\ \dfrac{\partial D^2}{\partial \tilde{\beta}_a} = 0, \end{cases} \tag{3.33}$$

где частные производные вычисляются по формулам:

$$\frac{\partial D^2}{\partial \tilde{\mu}_a} = \left(\frac{\partial D^2}{\partial \tilde{\mu}_{a0}}, \frac{\partial D^2}{\partial \tilde{\mu}_{a1}}, \dots, \frac{\partial D^2}{\partial \tilde{\mu}_{an}} \right), \ \frac{\partial D^2}{\partial \tilde{\alpha}_a} = \left(\frac{\partial D^2}{\partial \tilde{\alpha}_{a0}}, \frac{\partial D^2}{\partial \tilde{\alpha}_{a1}}, \dots, \frac{\partial D^2}{\partial \tilde{\alpha}_{an}} \right),$$

$$\frac{\partial D^2}{\partial \tilde{\beta}_a} = \left(\frac{\partial D^2}{\partial \tilde{\beta}_{a0}}, \frac{\partial D^2}{\partial \tilde{\beta}_{a1}}, \dots, \frac{\partial D^2}{\partial \tilde{\beta}_{an}} \right).$$

$$D^2 = \left\| X\tilde{\mu}_a - \mu_y \right\|^2 + \left\| X\tilde{\alpha}_a - \tilde{\alpha}_y \right\|^2 + \left\| X\tilde{\beta}_a - \tilde{\beta}_y \right\|^2 =$$
$$= \left(X\tilde{\mu}_a - \mu_y \right)^T \left(X\tilde{\mu}_a - \mu_y \right) + \left(X\tilde{\alpha}_a - \tilde{\alpha}_y \right)^T \left(X\tilde{\alpha}_a - \tilde{\alpha}_y \right) + \left(X\tilde{\beta}_a - \tilde{\beta}_y \right)^T \left(X\tilde{\beta}_a - \tilde{\beta}_y \right),$$

$$D^2 = \tilde{\mu}_a{}^T X^T X\tilde{\mu}_a - \mu_y{}^T X\tilde{\mu}_a - \tilde{\mu}_a{}^T X^T \mu_y + \mu_y{}^T \mu_y + \tilde{\alpha}_a{}^T X^T X\tilde{\alpha}_a - \alpha_y{}^T X\tilde{\alpha}_a -$$
$$- \tilde{\alpha}_a{}^T X^T \alpha_y + \alpha_y{}^T \alpha_y + \tilde{\beta}_a{}^T X^T X\tilde{\beta}_a - \beta_y{}^T X\tilde{\beta}_a - \tilde{\beta}_a{}^T X^T \beta_y + \beta_y{}^T \beta_y =$$
$$= \tilde{\mu}_a{}^T X^T X\tilde{\mu}_a - 2\tilde{\mu}_a{}^T X^T \mu_y + \mu_y{}^T \mu_y + \tilde{\alpha}_a{}^T X^T X\tilde{\alpha}_a - 2\tilde{\alpha}_a{}^T X^T \alpha_y + \alpha_y{}^T \alpha_y +$$
$$+ \tilde{\beta}_a{}^T X^T X\tilde{\beta}_a - 2\tilde{\beta}_a{}^T X^T \beta_y + \beta_y{}^T \beta_y,$$

$$\begin{cases} \dfrac{\partial D^2}{\partial \tilde{\mu}_a} = 2X^T X\tilde{\mu}_a - 2X^T \mu_y = 0, \\ \dfrac{\partial D^2}{\partial \tilde{\alpha}_a} = X^T X\tilde{\alpha}_a - X^T \alpha_y = 0, \\ \dfrac{\partial D^2}{\partial \tilde{\beta}_a} = X^T X\tilde{\beta}_a - X^T \beta_y = 0. \end{cases} \tag{3.34}$$

В результате получим следующие оценки параметров регрессии:

$$\begin{cases} \tilde{\mu}_a = (X^T X)^{-1} X^T \mu_y, \\ \tilde{\alpha}_a = (X^T X)^{-1} X^T \alpha_y, \\ \tilde{\beta}_a = (X^T X)^{-1} X^T \beta_y. \end{cases} \tag{3.35}$$

Из (3.35) следует, что оценки зависят только от наблюдаемых значений переменных. По свойствам действия с нечеткими числами уравнение (3.29) может быть представлено в виде системы [52]:

$$\begin{cases} \mu_{yi} = \mu_{a0} + \mu_{a1} x_{i1} + \dots + \mu_{an} x_{in} + \mu_{ei}, \\ \alpha_{yi} = \alpha_{a0} + \alpha_{a1} x_{i1} + \dots + \alpha_{an} x_{in} + \alpha_{ei}, \\ \beta_{yi} = \beta_{a0} + \beta_{a1} x_{i1} + \dots + \beta_{an} x_{in} + \beta_{ei}. \end{cases} \tag{3.36}$$

Соотношения (3.36) удобно записать в матричной форме:

$$\begin{cases} \mu_y = X\mu_a + \mu_e, \\ \alpha_y = X\alpha_a + \alpha_e, \\ \beta_y = X\beta_a + \beta_e. \end{cases} \quad (3.37)$$

где X – матрица размерностью $m \times (n+1)$, $\mu_a = (\mu_{a0}, \mu_{a1}, ..., \mu_{an})^T$, $\alpha_a = (\alpha_{a0}, \alpha_{a1}, ..., \alpha_{an})^T$, $\beta_a = (\beta_{a0}, \beta_{a1}, ..., \beta_{an})^T$, $\mu_y = (\mu_{y1}, ..., \mu_{ym})$, $\alpha_y = (\alpha_{y1}, ..., \alpha_{ym})$, $\beta_y = (\beta_{y1}, ..., \beta_{ym})$, $\mu_e = (\mu_{e1}, ..., \mu_{em})$, $\alpha_e = (\alpha_{e1}, ..., \alpha_{em})$, $\beta_e = (\beta_{e1}, ..., \beta_{em})$.

Подставив (3.37) в (3.35), получим оценки параметров

$$\begin{cases} \tilde{\mu}_a = (X^T X)^{-1} X^T (X\mu_a + \mu_e) = \mu_a + (X^T X)^{-1} X^T \mu_e, \\ \tilde{\alpha}_a = (X^T X)^{-1} X^T (X\alpha_a + \alpha_e) = \alpha_a + (X^T X)^{-1} X^T \alpha_e, \\ \tilde{\beta}_a = (X^T X)^{-1} X^T (X\beta_a + \beta_e) = \beta_a + (X^T X)^{-1} X^T \beta_e. \end{cases} \quad (3.38)$$

математическое ожидание которых равно:

$$\begin{cases} M(\tilde{\mu}_a) = \mu_a + (X^T X)^{-1} X^T M(\mu_e), \\ M(\tilde{\alpha}_a) = \alpha_a + (X^T X)^{-1} X^T M(\alpha_e), \\ M(\tilde{\beta}_a) = \beta_a + (X^T X)^{-1} X^T M(\beta_e). \end{cases} \quad (3.39)$$

Оценки, получаемые по методу наименьших квадратов, представляют собой несмещенные оценки параметров уравнения (3.29),

$$\begin{cases} M(\tilde{\mu}_a) = \mu_a, \\ M(\tilde{\alpha}_a) = \alpha_a, \\ M(\tilde{\beta}_a) = \beta_a, \end{cases} \quad (3.40)$$

тогда и только тогда, когда независимые переменные являются точными и математическое ожидание неучтенных переменных равно нулю [52]:

$$\begin{cases} M(\mu_e) = 0, \\ M(\alpha_e) = 0, \\ M(\beta_e) = 0. \end{cases} \quad (3.41)$$

Найдем зависимость оцениваемых параметров модели (3.29) от погрешности вычисления. Из (3.35) и (3.38) имеем:

$$\begin{cases}\mu_a + (X^T X)^{-1} X^T \mu_e = (X^T X)^{-1} X^T \mu_y,\\ \alpha_a + (X^T X)^{-1} X^T \alpha_e = (X^T X)^{-1} X^T \alpha_y,\\ \beta_a + (X^T X)^{-1} X^T \beta_e = (X^T X)^{-1} X^T \beta_y.\end{cases} \Rightarrow \begin{cases}\mu_a = (X^T X)^{-1} X^T \left(\mu_y - \mu_e\right),\\ \alpha_a = (X^T X)^{-1} X^T \left(\alpha_y - \alpha_e\right),\\ \beta_a = (X^T X)^{-1} X^T \left(\beta_y - \beta_e\right).\end{cases} \quad (3.42)$$

Как и в случае нечеткой парной регрессионной модели оценка качества нечеткой множественной модели может быть проведена по адекватности модели на основе анализа остаточной последовательности $\left\{E_i = Y_i - \tilde{Y}_i,\ i = \overline{1,n}\right\}$ и по точности модели путем вычисления средней относительной ошибки $E_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \tilde{Y}_i}{Y_i}\right| \cdot 100\%$.

Рассмотрим пример, который иллюстрируют полученные результаты. Исходные данные представляют собой набор из десяти трехмерных векторов-входов с четкими компонентами и десяти соответствующих им симметричных нечетких выходов (табл. 3.4) [52].

Найдем оценки параметров $\tilde{A}_j = \left(\tilde{\mu}_{aj}, \tilde{\alpha}_{aj}\right)$ модели (3.30) из (3.35).

Таблица 3.4 – Исходные данные

№ вектора	Переменные x_{i1}, x_{i2}, x_{i3}	Нечеткие выходы $Y_i = \left(\mu_{yi}, \alpha_{yi}\right)$
1	(3, 5, 9)	(96, 42)
2	(14, 8, 3)	(120, 47)
3	(7, 1, 4)	(52, 33)
4	(11, 7, 3)	(106, 45)
5	(7, 12, 15)	(189, 79)
6	(8, 15, 10)	(194, 65)
7	(3, 9, 6)	(107, 42)
8	(12, 15, 11)	(216, 78)
9	(10, 5, 8)	(108, 52)
10	(9, 7, 4)	(103, 44)

$$X^T X = \begin{pmatrix} 10 & 84 & 84 & 73 \\ 84 & 822 & 735 & 581 \\ 84 & 735 & 888 & 711 \\ 73 & 581 & 711 & 677 \end{pmatrix},$$

$$\left(X^T X\right)^{-1} = \frac{1}{14472471}\begin{pmatrix} 20385687 & -1422645 & 198675 & -1185903 \\ -1422645 & 167190 & -73857 & 87486 \\ 198675 & -73857 & 157364 & -123306 \\ -1185903 & 87486 & -123306 & 203670 \end{pmatrix},$$

$$\tilde{\mu}_a = (X^T X)^{-1} X^T \mu_y \approx \begin{pmatrix} -1{,}39 \\ 3{,}25 \\ 7{,}92 \\ 5{,}03 \end{pmatrix}, \ \tilde{\alpha}_a = (X^T X)^{-1} X^T \alpha_y \approx \begin{pmatrix} 8{,}01 \\ 1{,}64 \\ 1{,}2 \\ 2{,}85 \end{pmatrix}. \qquad (3.43)$$

Для каждого номера i вычислим оценку $\tilde{Y}_i = \left(\tilde{\mu}_{yi}, \tilde{\alpha}_{yi}\right)$:

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \tilde{A}_3 x_{i3},\ i = \overline{1{,}10},$$

а также ошибку $E_i = \left(\mu_{ei}, \alpha_{ei}\right), i = \overline{1,m}$, где $\mu_{ei} = \mu_{yi} - \tilde{\mu}_{yi}, \alpha_{yi} = \alpha_{yi} - \tilde{\alpha}_{yi}, i = \overline{1,m}$.

Таблица 3.5 – Полученные данные

№	$\tilde{\mu}_{yi}$	$\tilde{\alpha}_{yi}$	μ_{ei}	α_{ei}	$\hat{\mu}_{ei}$	$\hat{\alpha}_{ei}$
1	93,23	44,58	2,77	-2,58	0,971	0,939
2	122,56	49,12	-2,56	-2,12	0,979	0,955
3	49,4	32,09	2,6	0,91	0,950	0,972
4	104,89	43	1,11	2	0,990	0,956
5	191,85	76,64	-2,85	2,36	0,985	0,970
6	193,71	67,63	0,29	-2,63	0,999	0,960
7	109,82	40,83	-2,82	1,17	0,974	0,972
8	211,74	77,04	4,26	0,96	0,980	0,988
9	110,95	53,21	-2,95	-1,21	0,973	0,977
10	103,42	42,57	-0,42	1,43	0,996	0,968

Точность полученных результатов (табл. 3.5) определим по формулам:

$$\hat{\mu}_{ei} = 1 - \left|\frac{\mu_{ei}}{\mu_{yi}}\right|,\ \hat{\alpha}_{ei} = 1 - \left|\frac{\alpha_{ei}}{\alpha_{yi}}\right|.$$

Данные таблицы свидетельствуют от том, что оценки (3.43) были вычислены с точностью 97% (в среднем) [52].

Рассмотрим теперь пример нахождения параметров модели (3.29) с заранее заданной точностью 95%. Для этого воспользуемся формулой (3.42). Получим следующие результаты (табл. 3.6) [52]:

$$\left(X^T X\right)^{-1} = \frac{1}{14472471}\begin{pmatrix} 20385687 & -1422645 & 198675 & -1185903 \\ -1422645 & 167190 & -73857 & 87486 \\ 198675 & -73857 & 157364 & -123306 \\ -1185903 & 87486 & -123306 & 203670 \end{pmatrix},$$

$$\mu_a = (X^T X)^{-1} X^T (\mu_y - \mu_e) \approx \begin{pmatrix} -1{,}32 \\ 3{,}08 \\ 7{,}52 \\ 4{,}78 \end{pmatrix},\ \alpha_a = (X^T X)^{-1} X^T (\alpha_y - \alpha_e) \approx \begin{pmatrix} 7{,}61 \\ 1{,}56 \\ 1{,}14 \\ 2{,}71 \end{pmatrix}.$$

Таблица 3.6 – Результаты вычислений с точностью 95%

№	μ_{ei}	α_{ei}	$\tilde{\mu}_{yi}$	$\tilde{\alpha}_{yi}$	$\mu_{yi} - \tilde{\mu}_{yi}$	$\alpha_{yi} - \tilde{\alpha}_{yi}$	$\hat{\mu}_{ei}$	$\hat{\alpha}_{ei}$
1	4,8	2,1	91,2	39,9	7,46	-0,38	0,922	0,991
2	6	2,35	114	44,65	3,7	0,3	0,969	0,994
3	2,6	1,65	49,4	31,35	5,12	2,49	0,902	0,925
4	5,3	2,25	100,7	42,75	6,46	4,12	0,939	0,908
5	9,45	3,95	179,55	75,05	6,82	6,14	0,964	0,922
6	9,7	3,25	184,3	61,75	10,08	0,71	0,948	0,989
7	5,35	2,1	101,65	39,9	2,72	3,19	0,975	0,924
8	10,8	3,9	205,2	74,1	14,98	4,76	0,931	0,939
9	5,4	2,6	102,6	49,4	2,68	1,41	0,975	0,973
10	5,15	2,2	97,85	41,8	4,84	3,53	0,953	0,920

Проведя расчет точности найденных параметров по формулам:

$$\hat{\mu}_{ei} = 1 - \left| \frac{\mu_{yi} - \tilde{\mu}_{yi}}{\mu_{yi}} \right|,\ \hat{\alpha}_{ei} = 1 - \left| \frac{\alpha_{yi} - \tilde{\alpha}_{yi}}{\alpha_{yi}} \right|,$$

можно убедиться в том, что средняя погрешность вычислений действительно составляет 0,05% (табл. 3.6) [52].

3.2.2. Стандартизированное уравнение нечеткой линейной множественной регрессионной модели

Если рассматривать параметры нечеткой линейной множественной регрессионной модели в качестве показателей влияния объясняющих переменных, то стоит учесть, что коэффициенты регрессии $A_j = (\mu_{aj}, \alpha_{aj}, \beta_{aj})$, $j = \overline{0,n}$, представляющие собой нечеткие числа *L*-*R*-типа, в уравнении

$$Y = A_0 + A_1 x_1 + \ldots + A_n x_n,\ Y = (\mu_y, \alpha_y, \beta_y),\ x_p \in \Re,\ p = \overline{1,n}, \tag{3.44}$$

между собой прямо не сравнимы. Их численные значения зависят от выбранных единиц измерения каждой объясняющей переменной. Чтобы коэффициенты регрессии стали сравнимы, приведем их к стандартизированному масштабу [42, 98].

Для этого выразим все переменные в безразмерных, стандартизированных единицах измерения при помощи следующих соотношений:

$$T_y = (\mu_{ty}, \alpha_{ty}, \beta_{ty}) = \left(\frac{\mu_y - \bar{\mu}_y}{\sigma_\mu}, \frac{\alpha_y - \bar{\alpha}_y}{\sigma_\alpha}, \frac{\beta_y - \bar{\beta}_y}{\sigma_\beta} \right), \tag{3.45}$$

$$t_{xp} = \frac{(x_p - \bar{x}_p)}{\sigma_{xp}},\ p = \overline{1,n}, \tag{3.46}$$

где $Y = (\mu_y, \alpha_y, \beta_y)$ и x_p – значения зависимой и независимой переменных регрессионной модели в исходном масштабе; $\bar{Y} = (\bar{\mu}_y, \bar{\alpha}_y, \bar{\beta}_y)$ и $\bar{x}_p$ – средние значения зависимой и независимой переменных; $\sigma_\mu, \sigma_\alpha, \sigma_\beta$ – средние квадратические отклонения величин μ_y, α_y, β_y; σ_{xp} – средние квадратические отклонения переменных x_p; T_y и t_{xp} – соответствующие значения зависимой и независимой переменных регрессионной модели в стандартизированном масштабе.

Средние значения в (3.45) и (3.46) вычисляются по формулам:

$$\bar{x}_p = \frac{\sum_{i=1}^{m} x_{ip}}{m}, \tag{3.47}$$

$$\overline{\mu_y} = \frac{\sum_{i=1}^{m} \mu_{yi}}{m}, \ \overline{\alpha_y} = \frac{\sum_{i=1}^{m} \alpha_{yi}}{m}, \ \overline{\beta_y} = \frac{\sum_{i=1}^{m} \beta_{yi}}{m}, \tag{3.48}$$

а средние квадратические отклонения имеют вид:

$$\sigma_{xp} = \sqrt{\frac{\sum_{i=1}^{m} x_{ip}^2}{m} - \left(\frac{\sum_{i=1}^{m} x_{ip}}{m} \right)^2}, \tag{3.49}$$

$$\sigma_{\mu} = \sqrt{\frac{\sum_{i=1}^{m} \mu_{yi}^2}{m} - \left(\frac{\sum_{i=1}^{m} \mu_{yi}}{m} \right)^2}, \ \sigma_{\alpha} = \sqrt{\frac{\sum_{i=1}^{m} \alpha_{yi}^2}{m} - \left(\frac{\sum_{i=1}^{m} \alpha_{yi}}{m} \right)^2}, \ \sigma_{\beta} = \sqrt{\frac{\sum_{i=1}^{m} \beta_{yi}^2}{m} - \left(\frac{\sum_{i=1}^{m} \beta_{yi}}{m} \right)^2}. \tag{3.50}$$

Свободный член $A_0 = (\mu_{a0}, \alpha_{a0}, \beta_{a0})$ в стандартизированном уравнении нечеткой линейной множественной регрессии отсутствует, т. е. уравнение (3.44) можно записать как

$$T_y = B_1 t_{x1} + B_2 t_{x2} + \ldots + B_p t_{xp}, \ p = \overline{1,n}, \tag{3.51}$$

где нечеткие коэффициенты $B_p = (\mu_{bp}, \alpha_{bp}, \beta_{bp})$, $p = \overline{1,n}$, называются коэффициентами регрессии в стандартизированном масштабе, и так же, как и $A_p = (\mu_{ap}, \alpha_{ap}, \beta_{ap})$, $p = \overline{1,n}$, они являются нечеткими числами *L*-*R*-типа. Для их определения необязательно снова решать систему уравнений. Переход от коэффициентов $A_p = (\mu_{ap}, \alpha_{ap}, \beta_{ap})$ к $B_p = (\mu_{bp}, \alpha_{bp}, \beta_{bp})$, $p = \overline{1,n}$, и обратно осуществляется по формулам:

$$\begin{cases} \mu_{bp} = \dfrac{\mu_{ap} \cdot \sigma_{xp}}{\sigma_{\mu}}, \\ \alpha_{bp} = \dfrac{\alpha_{ap} \cdot \sigma_{xp}}{\sigma_{\alpha}}, \\ \beta_{bp} = \dfrac{\beta_{ap} \cdot \sigma_{xp}}{\sigma_{\beta}}, \end{cases} \tag{3.52}$$

где σ_{xp}, $\sigma_{\mu}, \sigma_{\alpha}, \sigma_{\beta}$ – соответственно, средние квадратические отклонения.

Коэффициенты $B_p = (\mu_{bp}, \alpha_{bp}, \beta_{bp})$, $p = \overline{1,n}$, позволяют выявить наиболее существенные независимые переменные, оказывающие влияние на зависимую величину. Это приводит к сокращению количества независимых переменных, участвующих в модели.

Пример. Пусть исходные данные заданы таблицей 3.4, на основе которых составлено уравнение регрессии с параметрами (3.43) в виде симметричных нечетких чисел *L*-*R*-типа:

$$\tilde{Y} = (-1.39{,}8.01) + (3.25{,}1.64)x_1 + (7.92{,}1.2)x_2 + (5.03{,}2.85)x_3 .$$

Необходимо записать его в стандартизированном масштабе. Сравнить влияние рассматриваемых независимых переменных x_1, x_2, x_3 на переменную $\tilde{Y}$.

Решение. Вычислим коэффициенты регрессии уравнения в стандартизированном масштабе по формулам (3.52). Для этого найдем сначала средние квадратические отклонения.

Вычислим все необходимые суммы (табл. 3.7).

Таблица 3.7 – Результаты вычислений

$\sum_{i=1}^{10} x_{i1}$	$\sum_{i=1}^{10} x_{i1}^2$	$\sum_{i=1}^{10} x_{i2}$	$\sum_{i=1}^{10} x_{i2}^2$	$\sum_{i=1}^{10} x_{i3}$	$\sum_{i=1}^{10} x_{i3}^2$	$\sum_{i=1}^{10} \mu_{yi}$	$\sum_{i=1}^{10} \mu_{yi}^2$	$\sum_{i=1}^{10} \alpha_{yi}$	$\sum_{i=1}^{10} \alpha_{yi}^2$
84	822	84	888	73	677	1291	191291	527	30041

Теперь по формулам (3.49)-(3.50) получим:

$\sigma_{x1}=\sqrt{\frac{822}{10}-\left(\frac{84}{10}\right)^2}=3{,}41$; $\sigma_{x2}=\sqrt{\frac{888}{10}-\left(\frac{84}{10}\right)^2}=4{,}27$;

$\sigma_{x3}=\sqrt{\frac{677}{10}-\left(\frac{73}{10}\right)^2}=3{,}796$, $\sigma_{\mu}=\sqrt{\frac{191291}{10}-\left(\frac{1291}{10}\right)^2}=49{,}62$,

$\sigma_{\alpha}=\sqrt{\frac{30041}{10}-\left(\frac{527}{10}\right)^2}=15{,}06.$

Найдем коэффициенты регрессии:

$$B_1=\left(\frac{\mu_{a1}\cdot\sigma_{x1}}{\sigma_{\mu}},\frac{\alpha_{a1}\cdot\sigma_{x1}}{\sigma_{\alpha}}\right)=\left(\frac{3{,}25\cdot 3{,}41}{49{,}62},\frac{1{,}64\cdot 3{,}41}{15{,}06}\right)=(0.24{,}0.37),$$

$$B_2=\left(\frac{\mu_{a2}\cdot\sigma_{x2}}{\sigma_{\mu}},\frac{\alpha_{a2}\cdot\sigma_{x2}}{\sigma_{\alpha}}\right)=\left(\frac{7{,}92\cdot 4{,}27}{49{,}62},\frac{1{,}2\cdot 4{,}27}{15{,}06}\right)=(0.68{,}0.34),$$

$$B_3=\left(\frac{\mu_{a3}\cdot\sigma_{x3}}{\sigma_{\mu}},\frac{\alpha_{a3}\cdot\sigma_{x3}}{\sigma_{\alpha}}\right)=\left(\frac{5{,}03\cdot 3{,}796}{49{,}62},\frac{2{,}85\cdot 3{,}796}{15{,}06}\right)=(0.38{,}0.72).$$

Таким образом, уравнение в стандартизированном масштабе имеет вид:

$$T_y=(0.24{,}0.37)t_{x1}+(0.68{,}0.34)t_{x2}+(0.38{,}0.72)t_{x3}.$$

Коэффициенты регрессии $B_p=(\mu_{bp},\alpha_{bp})$, $p=\overline{1{,}3}$, показывают влияние изменения каждой объясняющей переменной x_1,x_2,x_3 на изменение переменной $\tilde{Y}$. Все коэффициенты выражены в сравнимых единицах измерения. Чем больше B_p, $p=\overline{1{,}3}$, тем сильнее влияет соответствующая независимая переменная на зависимую величину. Исходя из полученных результатов, наибольшее влияние на $\tilde{Y}$ оказывают вторая и третья независимые переменные, а первая переменная имеет меньшее влияние.

3.2.3. Метод наименьших квадратов для модели с четкими коэффициентами и нечеткими данными

Предложим теперь метод оценки неизвестных параметров нечеткой множественной линейной регрессионной модели с четкими коэффициентами. Положим, что исходные данные состоят из пары (X,Y), где $X=(\mu_x,\alpha_x,\beta_x)$

$Y=(\mu_y,\alpha_y,\beta_y)$ – нечеткие числа L-R-типа. Исследуем следующую множественную модель [73]:

$$Y^{мод}=a_0+a_1X_1+...+a_nX_n, \tag{3.53}$$

где X_k, $k=\overline{1,n}$ – независимые (объясняющие) переменные, $Y^{мод}$ – зависимая (объясняемая) переменная, $a_j\in R^n$, $j=\overline{0,n}$ – неизвестные коэффициенты регрессии (параметры уравнения). Пусть m раз измерены значения независимых переменных $X_1,X_2,...,X_n$ и соответствующие значения $Y^{мод}$. Тогда:

$$Y_i=a_0+a_1X_{i1}+...+a_nX_{in}+E_i,\ i=\overline{1,m}, \tag{3.54}$$

где $Y_i=(\mu_{yi},\alpha_{yi},\beta_{yi}), i=\overline{1,m}$ – значения зависимой переменной, нечеткие числа L-R-типа (μ_{yi} – мода, α_{yi} и β_{yi} – левый и правый коэффициенты нечеткости, соответственно); $E_i=(\mu_{ei},\alpha_{ei},\beta_{ei}), i=\overline{1,m}$ – случайные ошибки, нечеткие числа L-R-типа, $i=\overline{1,m}$ – номер наблюдения; $X_{ik}=(\mu_{xik},\alpha_{xik},\beta_{xik})$ – значения независимой переменной, $k=\overline{1,n}$ – номер независимой переменной [73].

Для получения наиболее подходящих оценок регрессионных параметров a_j, $j=\overline{0,n}$, применим метод наименьших квадратов с использованием формулы (2.13) расстояния между двумя нечеткими числами [72, 75].

Следует учитывать при этом, что $Y_i=(\mu_{yi},\alpha_{yi},\beta_{yi})$, $i=\overline{1,m}$, и $X_{ik}=(\mu_{xik},\alpha_{xik},\beta_{xik})$, $k=\overline{1,n}$, имеют одинаковую функцию принадлежности, и после надлежащего преобразования данных можно считать все $a_j>0,\ j=\overline{0,n}$.

После подстановки оценок параметров $\tilde{a}_j$, $j=\overline{0,n}$, уравнение регрессии (3.54) приводится к виду:

$$\tilde{Y}_i=\tilde{a}_0+\tilde{a}_1X_{i1}+...+\tilde{a}_nX_{in},\ i=\overline{1,m}, \tag{3.55}$$

где $\tilde{Y}_i=(\tilde{\mu}_{yi},\tilde{\alpha}_{yi},\tilde{\beta}_{yi}), i=\overline{1,m}$ – оценки зависимой переменной, $\tilde{a}_j\in R^n$, $j=\overline{0,n}$ – оценки коэффициентов регрессии.

В соответствии с методом наименьших квадратов параметры коэффициентов $\widetilde{a}_j$ находятся из решения следующей задачи [73]:

$$F(\widetilde{a}_0, \widetilde{a}_1, \ldots, \widetilde{a}_n) = \sum_{i=1}^{m} D^2(Y_i, \widetilde{Y}_i) \to \min, \quad (3.56)$$

где Y_i и $\widetilde{Y}_i$ вычисляются по формулам (3.54) и (3.55) соответственно.

Согласно формулам (2.11), (2.13) и (2.14) целевую функцию можно переписать в виде:

$$F(\widetilde{a}_0, \widetilde{a}_1, \ldots, \widetilde{a}_n) = \sum_{i=1}^{m} \begin{pmatrix} (\widetilde{a}_0 + \widetilde{a}_1 \mu_{xi1} + \ldots + \widetilde{a}_n \mu_{xin} - \mu_{yi})^2 + \\ + (\widetilde{a}_1 \alpha_{xi1} + \ldots + \widetilde{a}_n \alpha_{xin} - \alpha_{yi})^2 + \\ + (\widetilde{a}_1 \beta_{xi1} + \ldots + \widetilde{a}_n \beta_{xin} - \beta_{yi})^2 \end{pmatrix} \to \min.$$

$$F(\widetilde{a}_0, \widetilde{a}_1, \ldots, \widetilde{a}_n) = \sum_{i=1}^{m} \left(([\widetilde{a}, \mu_{xi}] - \mu_{yi})^2 + ([\widetilde{a}, \alpha_{xi}] - \alpha_{yi})^2 + ([\widetilde{a}, \beta_{xi}] - \beta_{yi})^2 \right) \to \min, \quad (3.57)$$

где $\mu_{xi} = (1, \mu_{xi1}, \ldots, \mu_{xin})$, $\widetilde{a} = (\widetilde{a}_0, \widetilde{a}_1, \ldots, \widetilde{a}_n)$, $\alpha_{xi} = (0, \alpha_{xi1}, \ldots, \alpha_{xin})$, $\beta_{xi} = (0, \beta_{xi1}, \ldots, \beta_{xin})$.

Пусть $\|\vec{v}\|$ – длина вектора $\vec{v}$, тогда, используя векторные и матричные обозначения, величину F можно переписать в виде

$$F = \|\mu_x \widetilde{a} - \mu_y\|^2 + \|\alpha_x \widetilde{a} - \alpha_y\|^2 + \|\beta_x \widetilde{a} - \beta_y\|^2,$$

где μ_x, α_x, β_x – матрица размером $m \times (n+1)$, $\widetilde{a} = (\widetilde{a}_0, \widetilde{a}_1, \ldots, \widetilde{a}_n)^T$, $\alpha_y = (\alpha_{y1}, \ldots, \alpha_{ym})^T$, $\mu_y = (\mu_{y1}, \mu_{y2}, \ldots, \mu_{ym})^T$ и $\beta_y = (\beta_{y1}, \ldots, \beta_{ym})^T$. Значения вектора параметров $\widetilde{a}$, минимизирующих величину F, находятся из условий:

$$\frac{\partial F}{\partial \widetilde{a}} = 0, \quad (3.58)$$

где $\dfrac{\partial F}{\partial \widetilde{a}} = \left(\dfrac{\partial F}{\partial \widetilde{a}_0}, \dfrac{\partial F}{\partial \widetilde{a}_1}, \ldots, \dfrac{\partial F}{\partial \widetilde{a}_n} \right)$.

$$F = \|\mu_x \widetilde{a} - \mu_y\|^2 + \|\alpha_x \widetilde{a} - \alpha_y\|^2 + \|\beta_x \widetilde{a} - \beta_y\|^2 == (\mu_x \widetilde{a} - \mu_y)^T (\mu_x \widetilde{a} - \mu_y) +$$
$$+ (\alpha_x \widetilde{a} - \alpha_y)^T (\alpha_x \widetilde{a} - \alpha_y) + (\beta_x \widetilde{a} - \beta_y)^T (\beta_x \widetilde{a} - \beta_y),$$

$$F = \widetilde{a}^T \mu_x{}^T \mu_x \widetilde{a} - \mu_y{}^T \mu_x \widetilde{a} - \widetilde{a}^T \mu_x{}^T \mu_y + \mu_y{}^T \mu_y + \widetilde{a}^T \alpha_x{}^T \alpha_x \widetilde{a} - \alpha_y{}^T \alpha_x \widetilde{a} - \widetilde{a}^T \alpha_x{}^T \alpha_y +$$
$$+ \alpha_y{}^T \alpha_y + \widetilde{a}^T \beta_x{}^T \beta_x \widetilde{a} - \beta_y{}^T \beta_x \widetilde{a} - \widetilde{a}^T \beta_x{}^T \beta_y + \beta_y{}^T \beta_y,$$

$$\frac{\partial F}{\partial \widetilde{a}} = 2\mu_x^{\ T}\mu_x\widetilde{a} - 2\mu_x^{\ T}\mu_y + 2\alpha_x^{\ T}\alpha_x\widetilde{a} - 2\alpha_x^{\ T}\alpha_y + 2\beta_x^{\ T}\beta_x\widetilde{a} - 2\beta_x^{\ T}\beta_y = 0. \quad (3.59)$$

В результате получим следующие оценки параметров регрессии:

$$\widetilde{a} = (\mu_x^{\ T}\mu_x + \alpha_x^T\alpha_x + \beta_x^T\beta_x)^{-1}\left(\mu_x^{\ T}\mu_y + \alpha_x^T\alpha_y + \beta_x^T\beta_y\right). \quad (3.60)$$

3.3. Отбор независимых переменных в нечетком регрессионном анализе на основе нейронных сетей

На этапе построения нечеткой линейной регрессионной модели возникает необходимость отбора объясняющих переменных $x_1, x_2, ..., x_n$, которые оказывают влияние на переменную Y, при наличии внутренней взаимосвязи между переменными $x_1, x_2, ..., x_n$. Предварительно можно указать множество различных признаков, однако не все из них будут являться информативными и наилучшими для достижения высокой точности регрессионного моделирования. В связи с этим возникает задача уменьшения до минимума количества объясняющих переменных $x_1, x_2, ..., x_n$ без существенного увеличения вероятности ошибки проведения анализа. Иными словами, необходим поиск компромисса между избыточностью данных и точностью построения нечеткой линейной регрессионной модели [25, 54].

Для решения задачи оценки информативности и отбора входных признаков будем использовать автоассоциативную нейронную сеть, способную извлекать знания из данных [15, 47, 91, 98]. В общем случае она содержит три слоя нейронов: входной и выходной слой, а также средний слой – так называемое "узкое горло", который в результате обучения выдает сжатое представление данных (вектор z). Число выходов n совпадает с числом входов, а внутренний слой содержит меньшее количество нейронов $m < n$ [25, 30, 54].

Сеть самообучается на воспроизведение входов – то есть ответ нейросети считается правильным, когда значения сигналов на каждом выходе совпадает со значением на соответствующем ему входе ($x_i = \widetilde{x}_i$) (рис. 3.1) [25, 54].

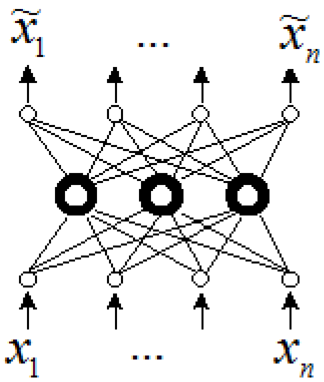


Рисунок 3.1 – Архитектура автоассоциативной нейронной сети с узким горлом

Если удается обучить такую нейросеть, то она способна решать задачу сокращения размерности – и тогда сигнал необходимо снимать с нейронов "горла" сети. Трехслойная автоассоциативная сеть сначала линейно преобразует входные данные в меньшую размерность промежуточного слоя, а затем снова линейно разворачивает их в выходном слое [25, 45, 54].

Рассмотрим алгоритм обучения автоассоциативной сети. Он заключается в следующем (рис. 3.2).

Пусть на наборе n-мерных данных обучается m линейных нейронов. Выходное значение каждого из этих нейронов в случае четких данных вычисляется по формуле:

$$z_i = \sum_{j=1}^{n} w_{ij} x_j ;\, z_i, w_{ij}, x_j \in \Re, i = \overline{1,m},\ j = \overline{1,n}. \tag{3.61}$$

Если входные признаки представляют собой нечеткие числа *L*-*R*-типа, то формула (3.61) имеет вид:

$$Z_i = \left(\mu_{zi}, \alpha_{zi}, \beta_{zi}\right) = \sum_{j=1}^{n} W_{ij} X_j = \sum_{j=1}^{n} \left(\mu_{wij}, \alpha_{wij}, \beta_{wij}\right) \cdot \left(\mu_{xj}, \alpha_{xj}, \beta_{xj}\right), i = \overline{1,m}, \tag{3.62}$$

где произведение $\left(\mu_{wij}, \alpha_{wij}, \beta_{wij}\right) \cdot \left(\mu_{xj}, \alpha_{xj}, \beta_{xj}\right)$ вычисляется по правилам (2.16)-(2.20) умножения нечетких чисел в зависимости от их знака. В случае (3.62) весовые коэффициенты так же, как и выходные значения промежуточного слоя будут являться нечеткими числами *L*-*R*-типа.

Мы хотим, чтобы амплитуды выходных нейронов были набором независимых индикаторов, максимально полно отражающих информацию о многомерном входе сети. Для этого логично было бы предложить правило, по которому значения входов восстанавливаются по всей выходной информации. Такое правило обучения называется правилом Ойя [44, 65, 87, 97], согласно которому весовые коэффициенты в случае четких данных изменяются следующим образом [25, 54]:

$$w_{ij}^{t+1} = w_{ij}^{t} + \Delta w_{ij}^{t}, \tag{3.63}$$

где $\Delta w_{ij}^{t} = \eta z_i^t \left(x_j^t - \tilde{x}_j^t \right) = \eta z_i^t \left(x_j^t - \sum_{k=1}^{m} z_k^t w_{kj}^t \right), i = \overline{1, m},\ j = \overline{1, n}$.

В нечетком случае формула (3.63) приобретает вид [73]:

$$W_{ij}^{t+1} = W_{ij}^{t} + \Delta W_{ij}^{t} = \left(\mu_{wij}^t + \Delta\mu_{wij}^t, \alpha_{wij}^t + \Delta\alpha_{wij}^t, \beta_{wij}^t + \Delta\beta_{wij}^t \right), \text{ где} \tag{3.64}$$

$$\Delta W_{ij}^{t} = \eta Z_i^t \left(X_j^t - \tilde{X}_j^t \right) = \eta \left(\mu_{zi}^t, \alpha_{zi}^t, \beta_{zi}^t \right) \left(\left(\mu_{xj}^t, \alpha_{xj}^t, \beta_{xj}^t \right) - \sum_{k=1}^{m} \left(\mu_{zk}^t, \alpha_{zk}^t, \beta_{zk}^t \right) \cdot \left(\mu_{wkj}^t, \alpha_{wkj}^t, \beta_{wkj}^t \right) \right),$$

$i = \overline{1, m},\ j = \overline{1, n}.$

В равенстве (3.64) все алгебраические операции над нечеткими числами проводятся на основе правил (2.14)-(2.20) вычисления суммы, разности и произведения.

Нейроны выходного слоя являются линейными с тождественной функцией активации [54]

$$\tilde{x}_j = \sum_{k=1}^{m} z_k w_{kj};\ z_k, w_{kj}, \tilde{x}_j \in \Re,\ j = \overline{1, n} \tag{3.65}$$

или для нечетких данных *L*-*R*-типа [73]

$$\tilde{X}_j = \left(\tilde{\mu}_{xj}, \tilde{\alpha}_{xj}, \tilde{\beta}_{xj} \right) = \sum_{k=1}^{m} Z_k W_{kj} = \sum_{k=1}^{m} \left(\mu_{zk}, \alpha_{zk}, \beta_{zk} \right) \cdot \left(\mu_{wkj}, \alpha_{wkj}, \beta_{wkj} \right),\ j = \overline{1, n}. \tag{3.66}$$

Таким образом, сеть с узким горлом из скрытых линейных нейронов обучается воспроизводить на выходе значения своих входов. Скрытый слой такой сети при этом осуществляет оптимальное кодирование входных данных и содержит максимально возможное при данных ограничениях количество информации [73].

Для подобных целей могут быть использованы алгоритмы стандартного статистического анализа, однако изложенный выше итеративный нейросетевой алгоритм обладает следующими преимуществами [25, 46, 70]:

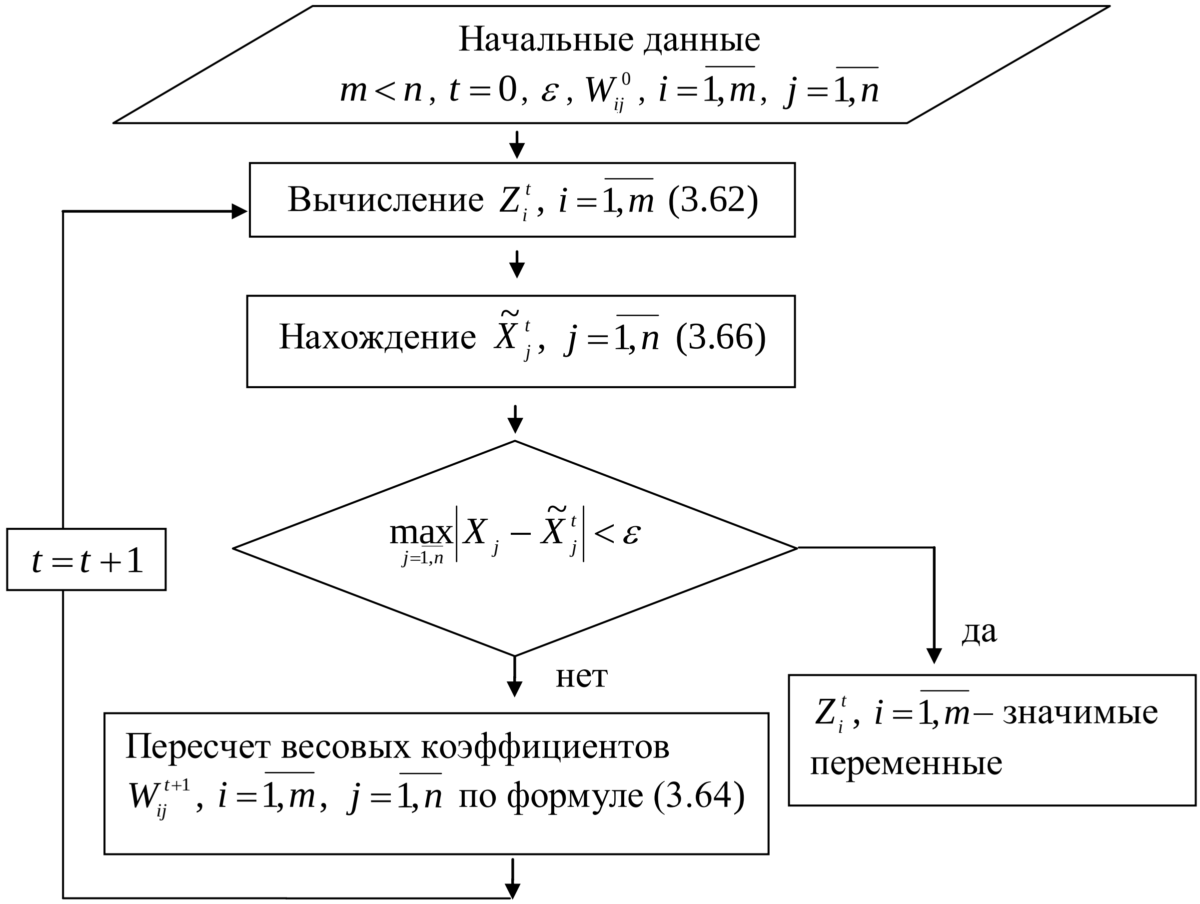


Рисунок 3.2– Схема обучения автоассоциативной нейронной сети

1. Во-первых, иногда обучение необходимо проводить в режиме on-line, т. е. на ходу адаптироваться к меняющемуся потоку данных. В этой ситуации идеально подходят итерационные методы, когда нет возможности собрать воедино весь набор примеров и произвести необходимые матричные операции над ним.

2. Во-вторых, нейроалгоритм легко обобщается на случай нелинейного сжатия информации, когда никаких явных решений уже не существует. Для этого необходимо заменить линейные нейроны в описанной автоассоциативной нейронной сети нелинейными. С минимальными видоизменениями нейроалгорим будет работать и в этом случае, всегда находя оптимальное сжатие при наложенных ограничениях.

Выводы по главе 3

1. Получена оценка параметров нечеткой парной и множественной линейной регрессионной модели с помощью метода наименьших квадратов.

2. Осуществлена оценка адекватности модели на основе анализа нечеткого коэффициента корреляции и остаточной последовательности ошибок вычисления.

3. В качестве одного из подходов определения точности модели было предложено вычисление средней относительной ошибки.

4. Разработан нечеткий тип классического регрессионного анализа, предназначенный для нахождения функционального взаимодействия между зависимыми и независимыми переменными в нечеткой среде и позволяющий решать различные задачи в ситуациях, когда традиционные методы неэффективны или совсем неприменимы из-за отсутствия достаточно точных сведений об исследуемых объектах.

5. Для выявления наиболее существенных независимых переменных, оказывающих влияние на зависимую величину, а также с целью сокращения количества независимых переменных, участвующих в модели, было построено стандартизированное уравнение нечеткой множественной регрессии.

6. Для решения задачи отбора независимых переменных в регрессионной модели с нечеткими данными разработан подход, основанный на использовании автоассоциативной нейронной сети, принцип работы которой был адаптирован для анализа информации в нечеткой среде.

Глава 4. Программный комплекс для проведения интеллектуального анализа данных на основе нечеткого регрессионного моделирования

В данной главе разработана система интеллектуального анализа данных на основе нечеткого регрессионного моделирования, позволяющая выявить полезную скрытую информацию для определения зависимостей, которые содержатся в данных; приведена структура хранилища данных информационной системы; рассмотрены функции системы администрирования; определены этапы процесса интеллектуального анализа данных; разработано программное обеспечение для проведения нечеткого регрессионного анализа данных и моделирования, включающее реализацию калькулятора нечетких чисел для осуществления различных арифметических операций над нечеткими числами L-R-типа и построения графиков их функций принадлежности, выполнение отбора существенных независимых переменных на основе автоассоциативных нейронных сетей, проведение нечеткого линейного парного и множественного регрессионного анализа, построение стандартизированного уравнения нечеткой множественной линейной регрессии.

4.1. Разработка информационной системы интеллектуального анализа данных

4.1.1. Структура информационной системы интеллектуального анализа данных на основе нечеткого регрессионного моделирования

Разработанная информационная система интеллектуального анализа данных (ИС ИАД) основана на технологии информационного хранилища [28, 85]. Она обрабатывает большие массивы данных, осуществляет автоматизированный поиск ранее неизвестных закономерностей и скрытых и неочевидных правил в базе данных. Полученные знания помогают оптимизировать процессы деятельности предприятия и могут быть использованы для принятия решений [2, 10, 86].

Структура системы ИАД (рис. 4.1) предусматривает наличие двух приложений – аналитического (основного) и системы администрирования (вспомогательного). Последнее предназначено для составления правильных SQL-запросов к базам данных информационной системы при участии пользователя системы – аналитика [14]. Данные о формах и переходах содержатся в специальной базе ИС, которая может быть локальной или удаленной.

Процесс проведения интеллектуального анализа данных на основе нечеткого регрессионного моделирования предназначен для восстановления закономерностей с помощью предложенных в главе 3 методов и алгоритмов выявления знаний в нечеткой среде.

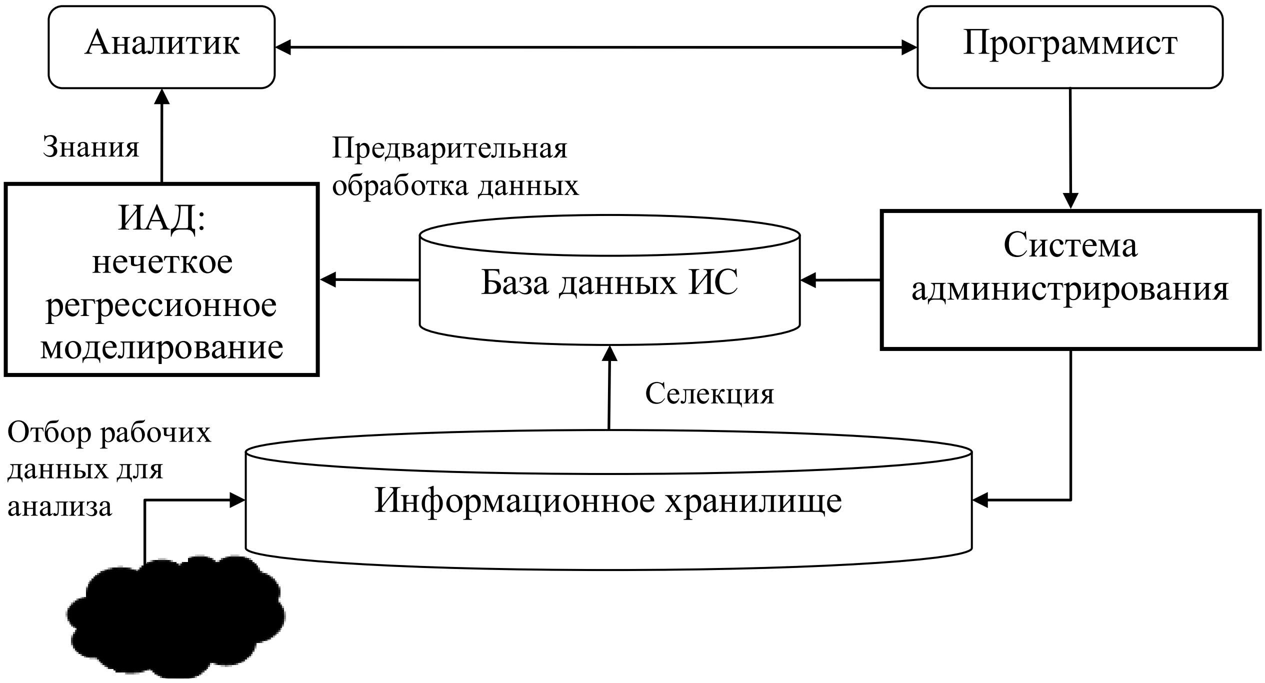


Рисунок 4.1 – Структура информационной системы ИАД

На рисунке 4.2 приведена физическая модель данных блока идентификации пользователя, который получает доступ к системе.

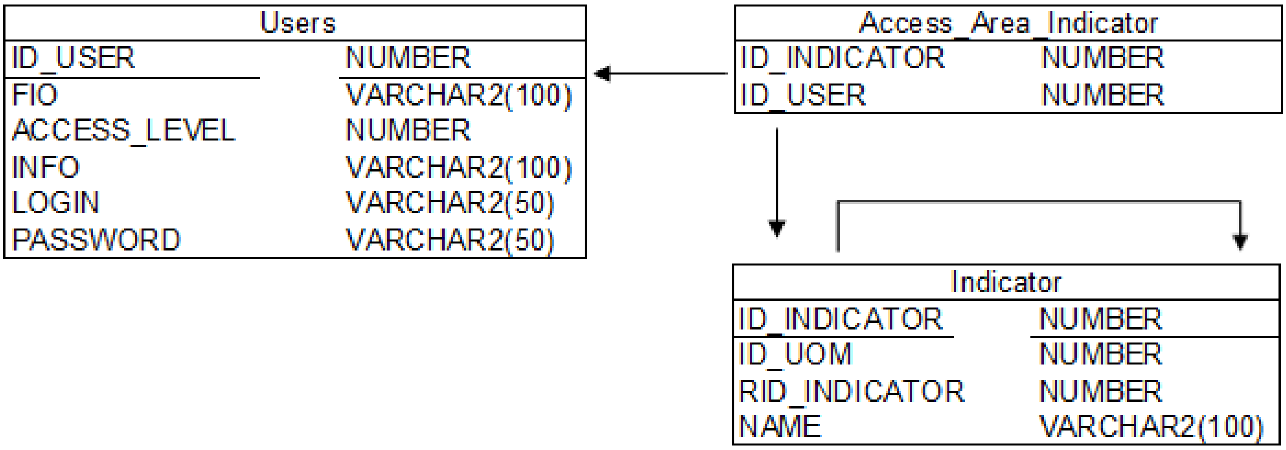


Рисунок 4.2 – Физическая модель данных блока идентификации пользователя

Сущность “Пользователь” является основной в блоке идентификации. Для разделения прав доступа служит атрибут “Уровень доступа”. Хранение служебной информации о пользователе осуществляется с помощью атрибута “Информация”. В сущности содержатся также имя пользователя и пароль для входа в систему.

Блок идентификации пользователя связан с сущностью “Показатель” ядра информационного хранилища через “Область доступа”, которая служит для определения показателей, доступных пользователю.

4.1.2. Информационное хранилище системы ИАД

Комплексное хранилище данных информационной системы ИАД содержит всю информацию, необходимую для аналитической обработки данных. При этом применяется следующий подход [14, 56, 60]:

– выделяется сущность (показатель, данные, единица измерения, период действия);

– составляется ее полное описание – перечисляются все характеристики сущности;

– данные загружаются строго по составленному описанию.

Концептуальная схема всей системы приведена на рисунке 4.3. Хранилище данных состоит из уровней иерархии данных и уровней иерархии процедур. Иерархия данных включает в себя [68, 81]:

1. данные с точки зрения СУБД – записи, хранящиеся в таблицах пользователя;
2. метаданные с точки зрения СУБД – имена и структура таблиц, имена и типы столбцов, связи между таблицами и т.д.;
3. данные с точки зрения предметной области – значения отдельных показателей и сроки их применения;
4. метаданные с точки зрения предметной области – имена, типы, свойства показателей, их связь между собой, ограничения, накладываемые на показатели.

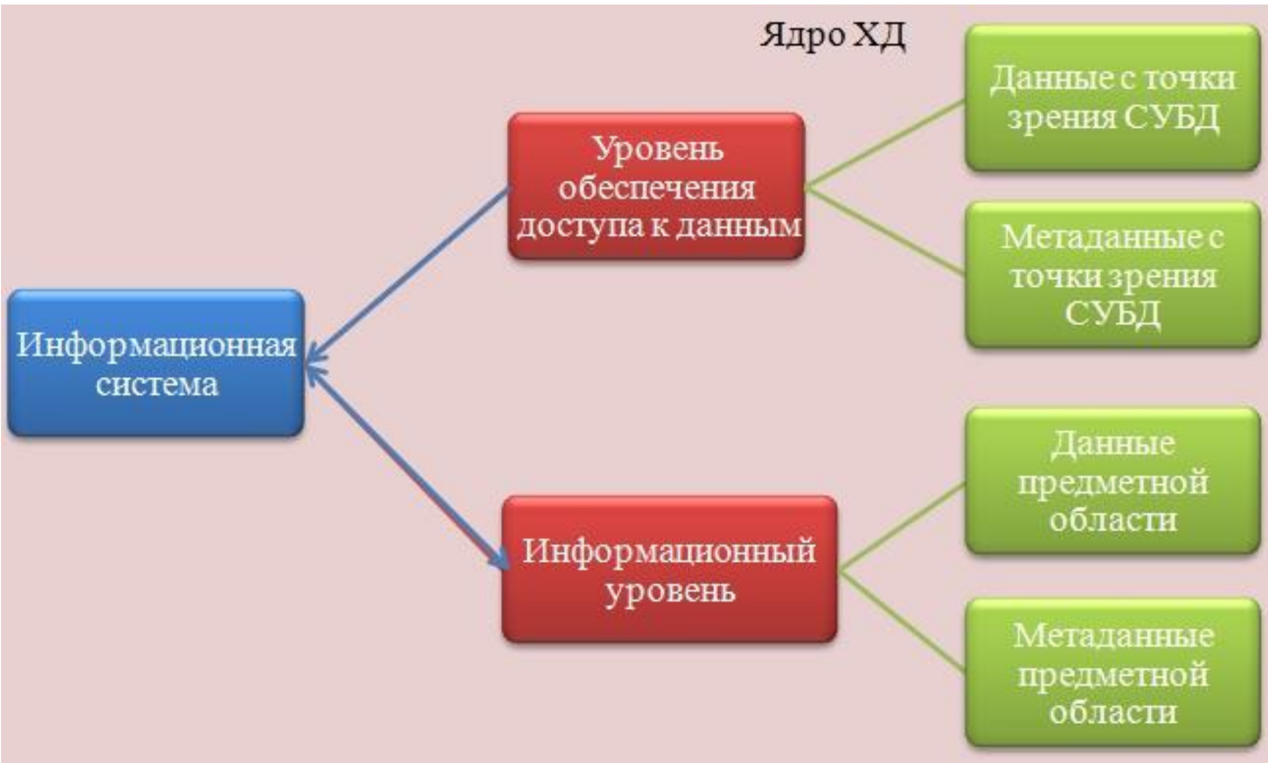


Рисунок 4.3 – Концептуальная схема системы

Иерархия процедур представлена двумя уровнями [32, 33, 49]:

1. уровнем обеспечения доступа к данным для уровней иерархии данных 1 и 2 реляционной СУБД. Он содержит набор процедур для извлечения, преобразования и загрузки данных;

2. информационным уровнем для уровней иерархии данных 3 и 4 предметной области, который включает данные для построения отчетов и анализа.

На рисунке 4.4 приведена физическая модель ядра информационного хранилища в виде таблиц сущностей, которые взаимосвязаны между собой [63].

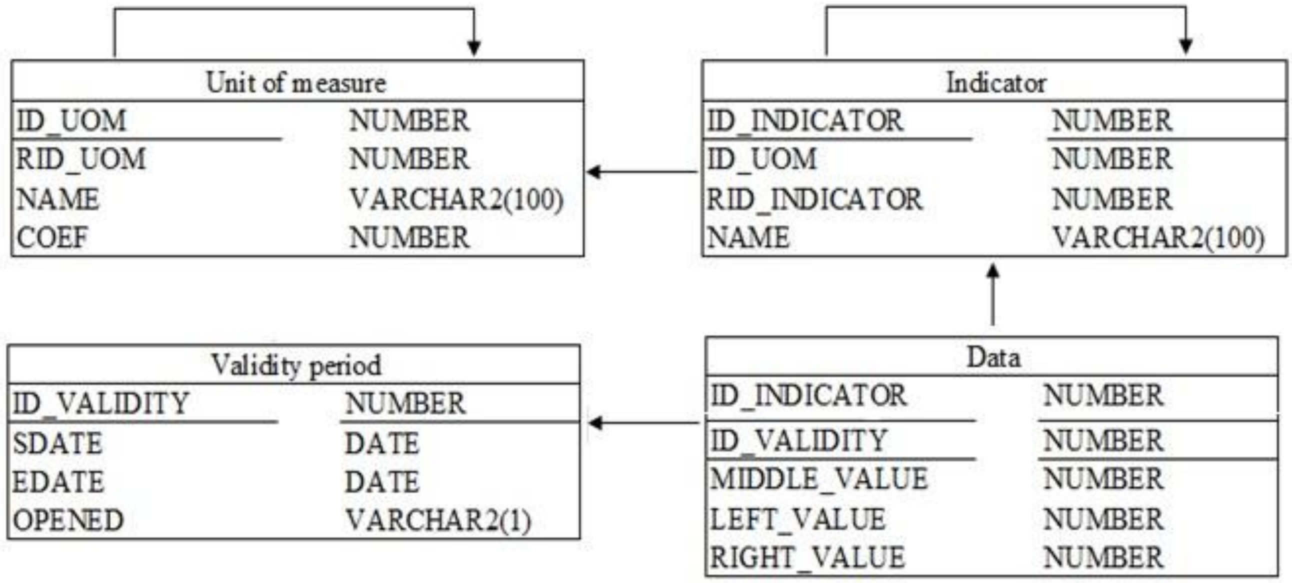


Рисунок 4.4 – Физическая модель ядра информационного хранилища

Сущность “Показатель” включает описание экономических, технических и иных показателей, которые необходимы для проведения аналитической работы. В ядре организована некоторая иерархическая структура показателей: в качестве родителя выступает группа технических, финансовых и других видов показателей, которые в дальнейшем детализируются. Отношение иерархичности определено связью сущности с самой собой и предназначено для классификации имеющихся показателей, используемых в ходе анализа данных.

Сущность “Единица измерения” создана для осуществления хранения данных в единой форме и содержит в иерархическом виде информацию об используемых измерениях. На верхнем уровне находятся базовые единицы (метр, килограмм и т. д.), дочерними к которым выступают производные единицы измерения. Атрибут “Множитель” определяет коэффициент пересчета между базовыми и дочерними единицами измерения.

Сущность “Данные” содержит значения показателей по каждому измерению в виде трех атрибутов для внесения центра и левой и правой границ нечеткого числа. Кроме того, она ссылается на идентификатор показателя, период действия и дату актуальности, что позволяет при необходимости иметь информацию о значениях определенных показателей в более детальном виде.

Период действия включает дату начала и окончания проведения измерений. При этом возможность добавления данных за определенный период контролируется с помощью атрибута “Открыт для заполнения”. Невозможно внести информацию в закрытый период.

4.1.3. Система администрирования

Данная система предназначена для осуществления администрирования информационного хранилища данных. Она включает в себя следующие задачи [34]:

- авторизация пользователей;
- управление пользователями;
- управление метаданными;
- управление периодами действия;
- загрузку данных в информационное хранилище.

Функция авторизации пользователя предполагает наличие идентификации и аутентификации пользователей, которые зарегистрированы в системе и имеют доступ к работе.

Функция управления пользователями предназначена для создания, изменения и удаления пользователя, просмотра списка всех пользователей системы, основные данные которых должны соответствовать сущности “Пользователь”.

Функция управления метаданными служит для управления сущностями “Единица измерения” и “Показатель” и позволяет работать с иерархической структурой данных, разделяя единицы измерения и показатели на базовые и производные от них. Также она предоставляет возможность добавления, редактирования, удаления данных и просмотра их списка.

Функция управления периодами действия предназначена для создания новых отчетных периодов и закрытия старых. Данные должны соответствовать сущности “Период действия”.

Функция загрузки данных в информационное хранилище служит для внесения данных по имеющемуся списку показателей и текущему периоду действия.

Входными данными программного обеспечения администрирования информационного хранилища служат данные, входящие в объекты пользователь, единица измерения, показатель и период действия. Выходными данными при этом являются количество загруженных записей в информационное хранилище, а также совокупность записей за определенный отчетный период.

4.1.4. Процесс интеллектуального анализа данных

Процесс выполнения интеллектуального анализа данных, основанного на нечетком регрессионном моделировании, состоит из следующих этапов (рис. 4.5) [12, 50, 84]:

– на первом выполняется осмысление поставленной задачи и уточнение целей, которые должны быть достигнуты методами ИАД. От этого зависит дальнейшая эффективность всего процесса;

– второй этап состоит в приведении данных к форме, пригодной для применения алгоритмов и методов нечеткого регрессионного моделирования. На этом шаге также осуществляется отбор значимых переменных с помощью автоассоциативной нейронной сети;

– третий этап – применение методов ИАД;

– следующий этап – проверка адекватности и точности построенной модели;

– последний этап – интерпретация проведенного моделирования человеком в целях использования полученных результатов для принятия решений.

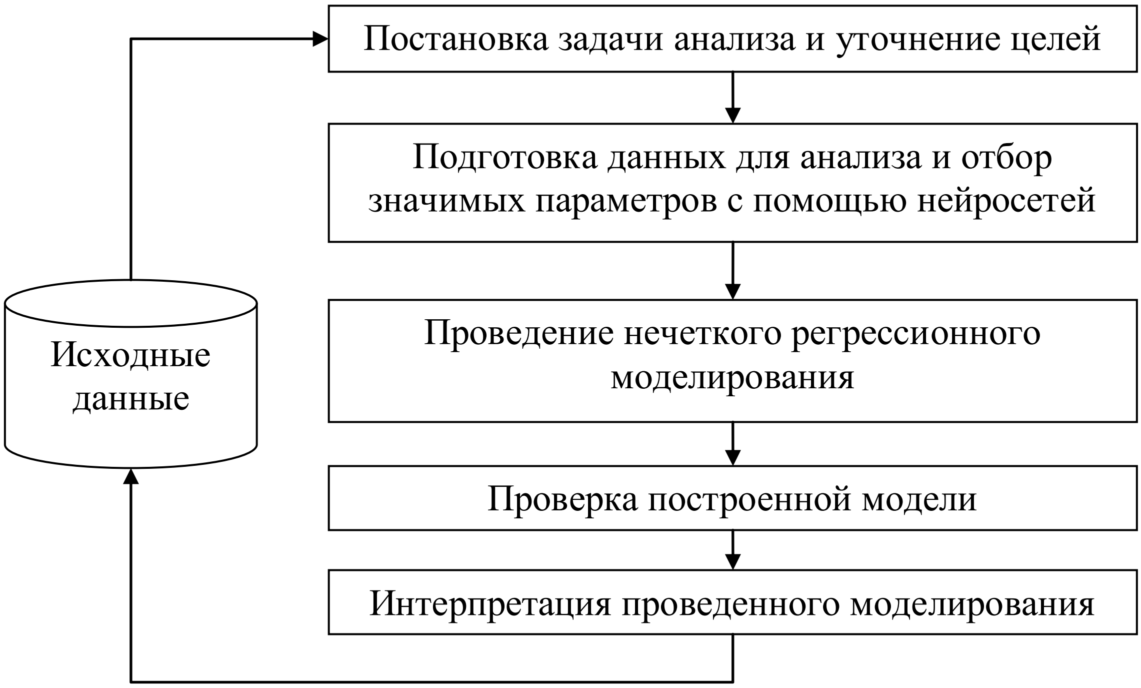


Рисунок 4.5 – Этапы интеллектуального анализа данных

Выполнение интеллектуального анализа данных осуществляется на трех уровнях обработки и представления информации (рис. 4.6). На уровне базы данных происходит селекция данных из информационного хранилища и их предварительная обработка (выбор существенных переменных) для проведения дальнейшего анализа. Программный комплекс нечеткого регрессионного моделирования запускается на уровне приложений с целью извлечения требуемой информации или знаний. Третий уровень – представление данных. Он содержит пользовательский интерфейс для определения параметров по выбору информации из базы данных и ее предварительной обработки, а также для отображения полученных результатов [10, 38].

В ходе интеллектуального анализа пользователю доступен только уровень представления данных. При этом техническая реализация информационного хранилища, программное обеспечение алгоритмов и методов нечеткого регрессионного моделирования выступают для него в качестве черного ящика (рис. 4.7).

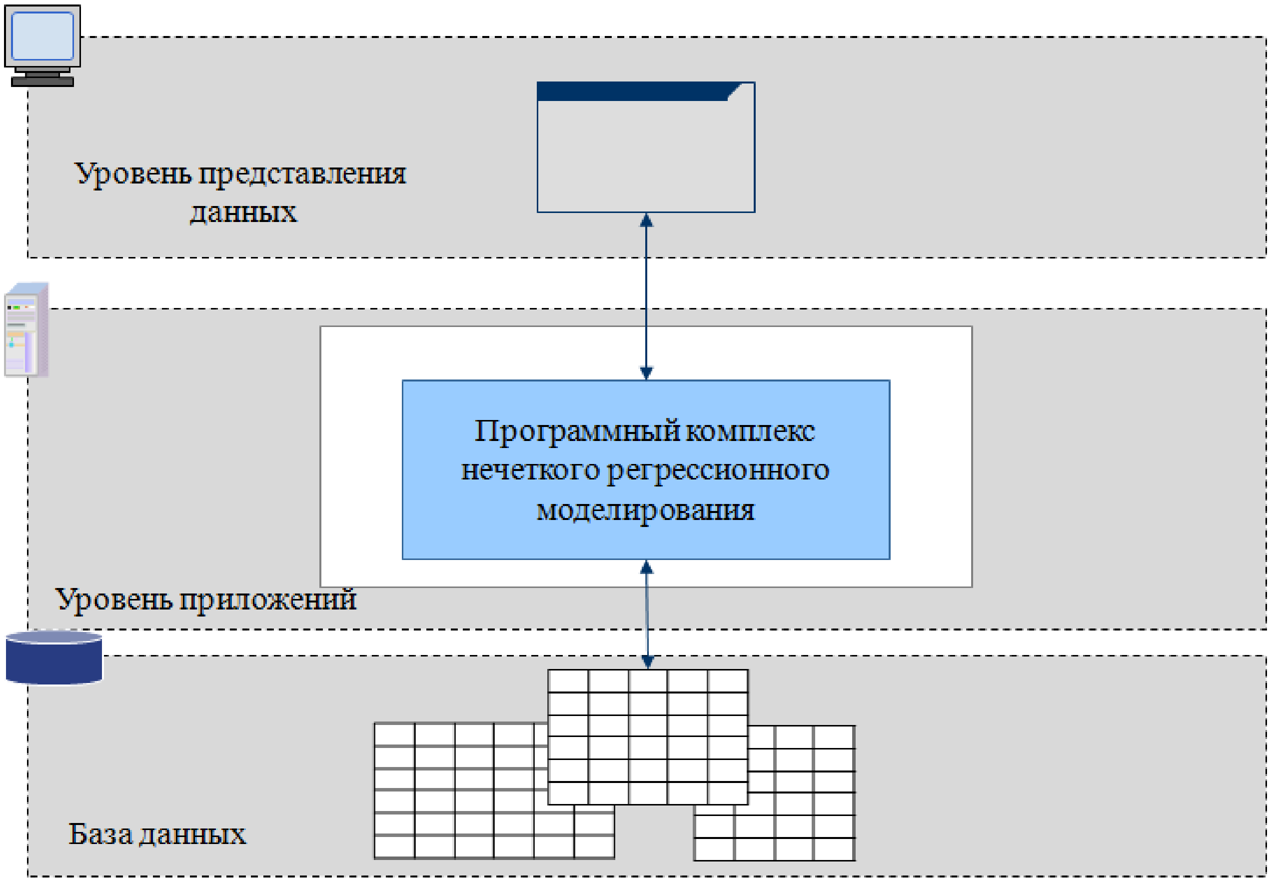


Рисунок 4.6 – Архитектура клиент/сервер для проведения ИАД

В разработанной информационной системе ИАД предусмотрены возможности для извлечения данных из хранилища, их предварительной обработки и анализа. Для этого реализованы следующие функции [2, 12, 69]:

– выбор данных из информационного хранилища;
– обновление данных в информационном хранилище;
– трансформация данных;
– анализ данных.

Функция запроса данных из информационного хранилища используется для выборки данных по заданному периоду действия и типу показателя для последующего их использования в ходе анализа. При этом может потребоваться объединение нескольких таблиц и фильтрация полученных записей.

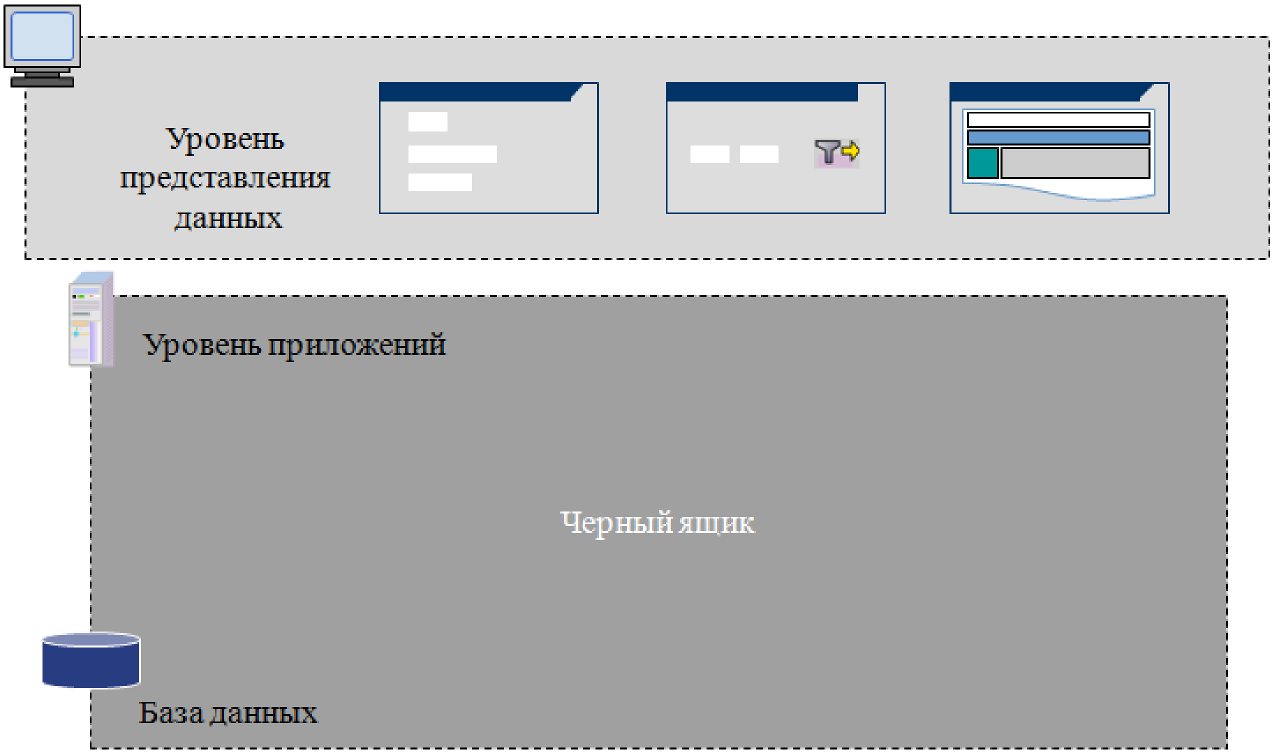


Рисунок 4.7 – Уровень доступности для пользователя

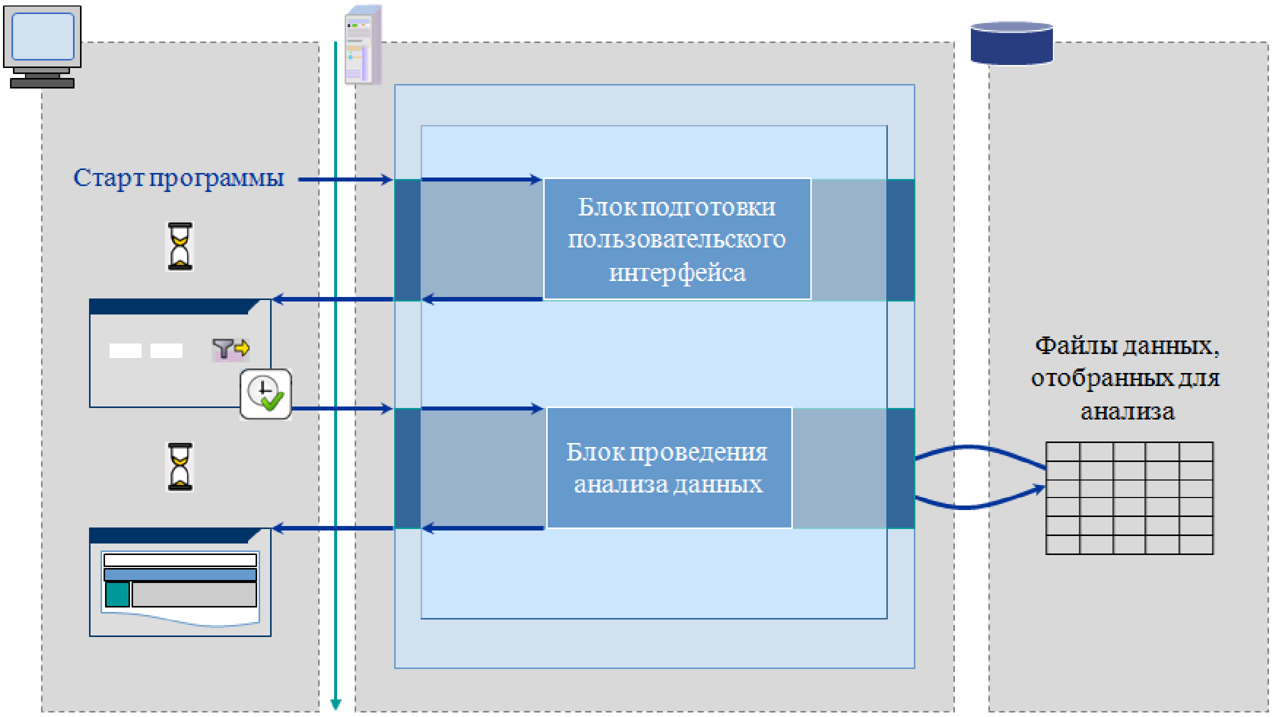


Рисунок 4.8 – Взаимодействие уровней анализа данных

Функция обновления данных предназначена для вставки новых, удаления старых или же изменения уже существующих записей информационного хранилища. Примером является внесение более подробных сведений по определенному показателю.

С помощью функции трансформации данных осуществляется предварительная обработка данных, характер которой определяется методами, применяемыми в ходе анализа. Трансформация заключается также в удалении дублирующих записей или зашумленных данных, преобразовании типов, добавлении новых атрибутов.

Функция анализа данных позволяет с помощью интеллектуального аппарата провести последовательную обработку информации согласно разработанным алгоритмам и методам нечеткого регрессионного моделирования.

4.2. Программное обеспечение нечеткого регрессионного моделирования

Принципы нечеткого регрессионного моделирования, описанные в предыдущих главах, были использованы для разработки специального программного обеспечения в среде CodeGear 2009 Borland C++ Builder [21, 66, 88, 90], которое было зарегистрировано в Роспатенте (приложение Б, рис. 1-2). Программа реализует возможности калькулятора нечетких чисел для осуществления различных арифметических операций над нечеткими числами и построения графиков их функций принадлежности, позволяет провести нечеткий линейный парный и множественный регрессионный анализ с нахождением коэффициентов модели и средней ошибки вычислений. С помощью разработанного программного обеспечения можно также построить стандартизированное уравнение нечеткой множественной линейной регрессии и провести отбор существенных переменных на основе автоассоциативных нейронных сетей.

Структура программного обеспечения нечеткого регрессионного моделирования приведена на рисунке 4.9.

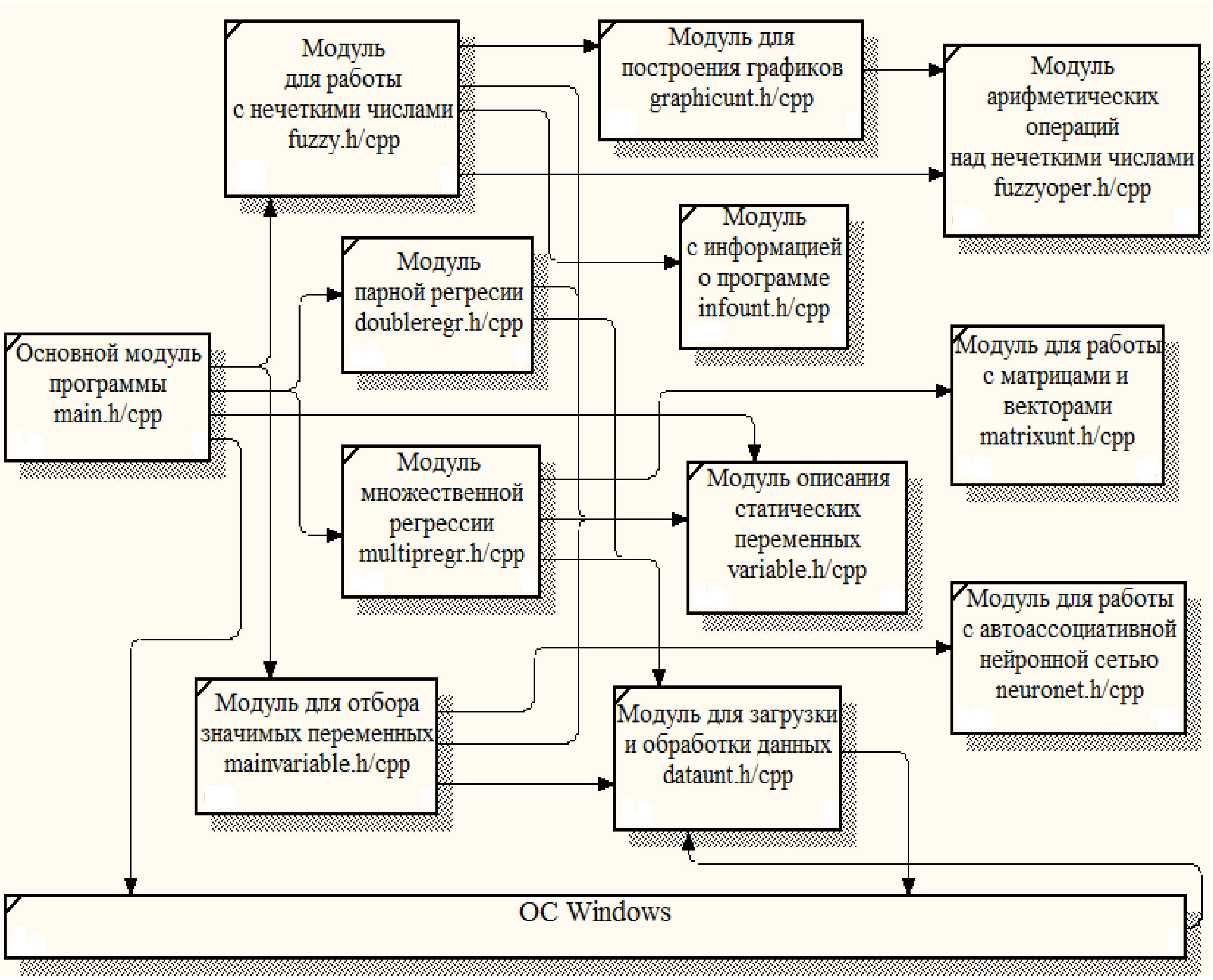


Рисунок 4.9 – Структура ПО нечеткого регрессионного моделирования

Процесс выполнения анализа нечеткой информации начинается с определения необходимого модуля (рис. 4.10): калькулятора нечетких чисел, нечеткой парной линейной регрессии, нечеткой множественной линейной регрессии и отбора существенных переменных с помощью автоассоциативной нейронной сети.

При выборе “Калькулятора” (рис. 4.11) автоматически открывается новая форма для работы с нечеткими числами *L-R*-типа, которая включает выполнение арифметических операций над ними и построение графиков.

В верхней части окна расположена строка состояния, в которой отображаются результаты проведенных операций. Ниже находятся поля для ввода с клавиатуры центра и левой и правой границ нечеткого числа.

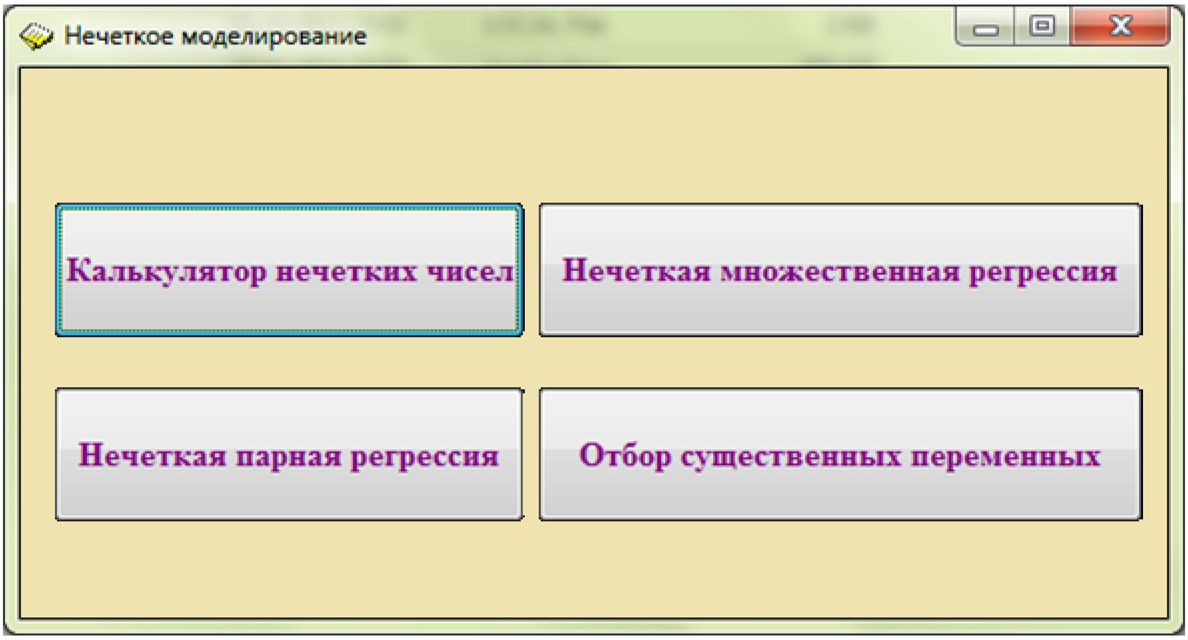


Рисунок 4.10 – Стартовое окно программы

Клавиатура калькулятора содержит кнопки ввода чисел и выполнения операций и функций. Она включает следующие клавиши [24]:

- Цифровые – десять клавиш с цифрами от 0 до 9.
- Десятичная запитая – ввод десятичного разделителя.
- Арифметические операции – ввод операций сложения (2.14), вычитания (2.15), умножения (2.16)-(2.20), возведения в квадрат (2.39)-(2.40), а также нахождение обратного (2.10), противоположного нечеткого числа (2.9) и расстояния между двумя нечеткими числами (2.13).
- Знак равенства – выполнение последней операции в цепочных вычислениях.
- Очистка – обнуление значения в поле ввода и отмены операции, если таковая была введена (клавиша “C”), или удаление последнего введенного символа (клавиша “←”).
- Табуляция – переключение фокуса ввода в поля центра и левой и правой границ нечеткого числа (клавиша “Tab”).
- Регистр памяти – очистка регистра памяти (клавиша “MC”), копирование значения из регистра памяти в поля ввода (клавиша “MR”), сохранение в

регистре текущего значения (клавиша “MS”), выполнение операций сложения и вычитания между текущим значением в регистре памяти и значением в полях ввода нечеткого числа с помещением результата снова в регистр памяти (клавиши “M+” и “M-”).

Из пункта меню “Файл” доступна команда построения графиков функций принадлежности нечетких чисел, над которыми была проведена текущая операция, а также нечеткого числа как результата этой операции (рис. 4.12). Пункт меню “Справка” содержит сведения по функционалу “Калькулятора”.

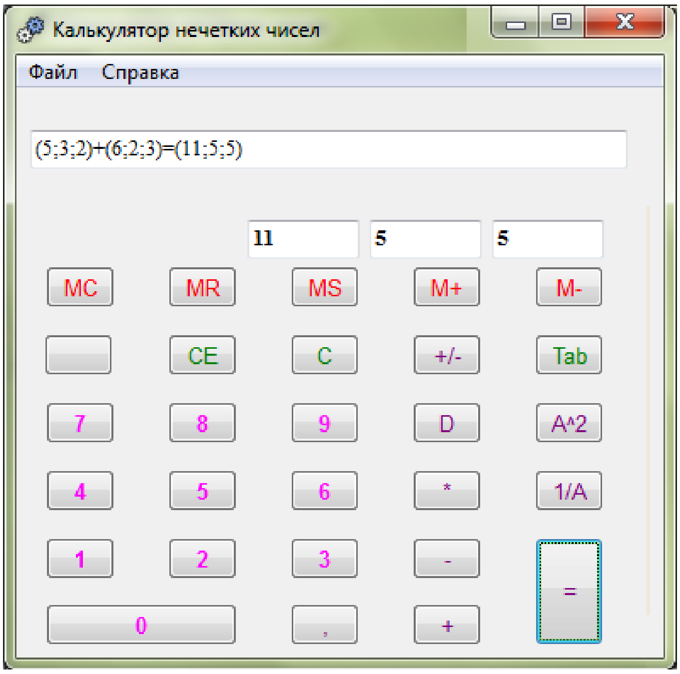


Рисунок 4.11 – Калькулятор нечетких чисел

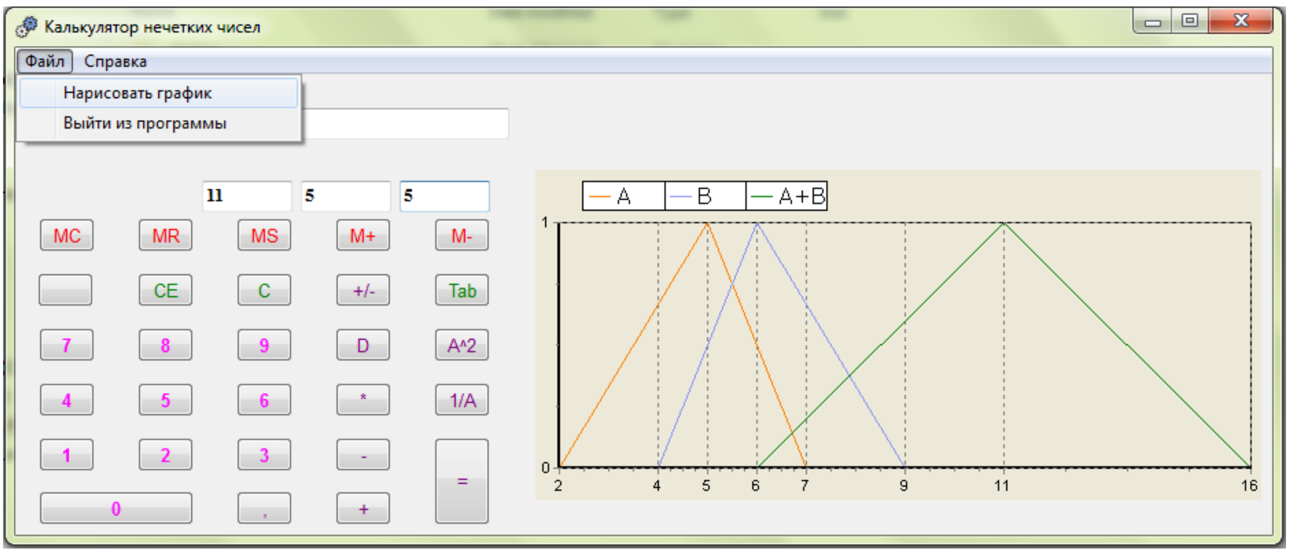


Рисунок 4.12 – Работа с калькулятором нечетких чисел

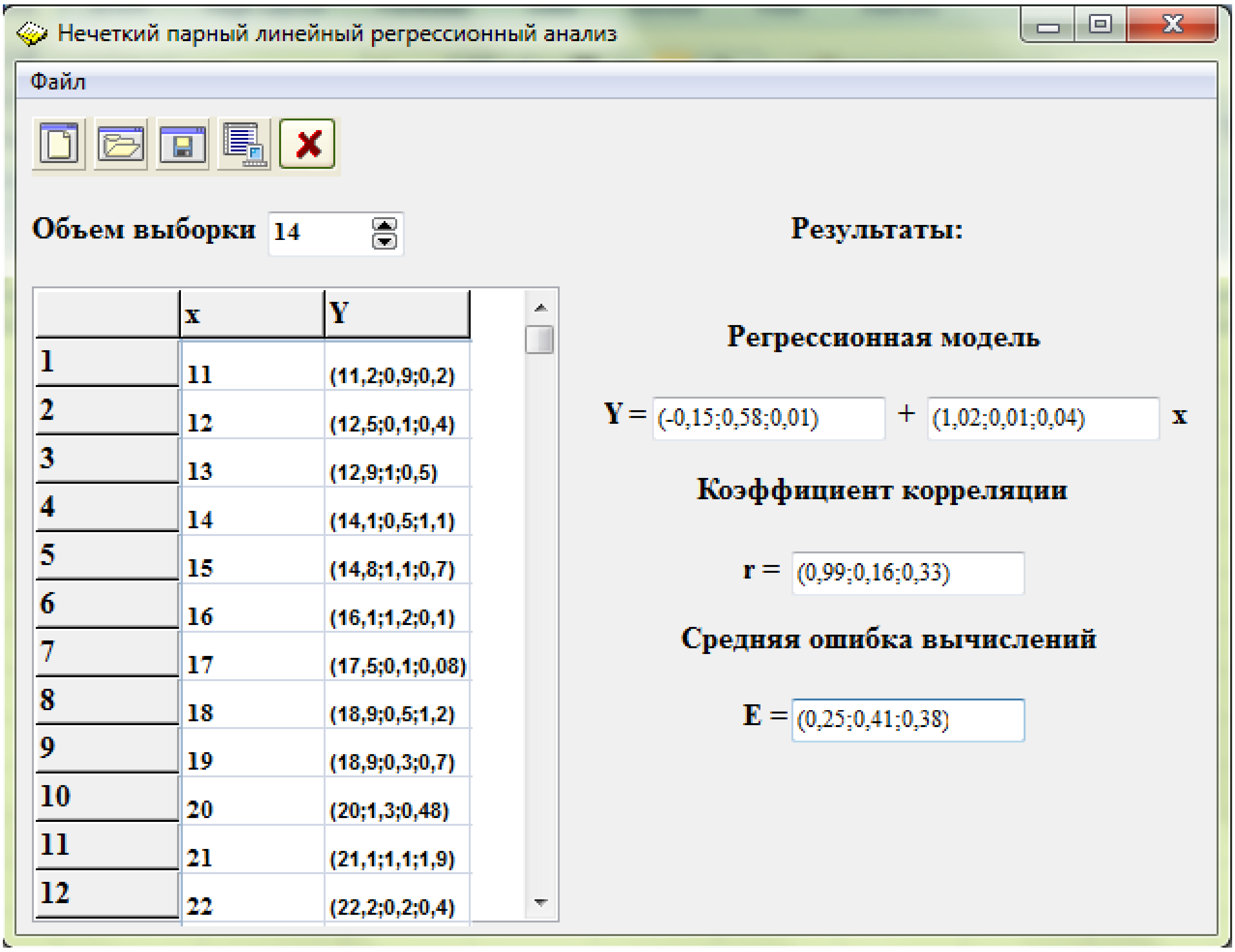


Рисунок 4.13 – Проведение нечеткого парного линейного регрессионного анализа

Проведение нечеткого регрессионного анализа (парного и множественного) (рис. 4.13-4.14) включает в себя следующие этапы [38]:

- инициализация;
- ввод данных;
- предобработка;
- выполнение нечеткого регрессионного моделирования;
- постобработка;
- вывод данных.

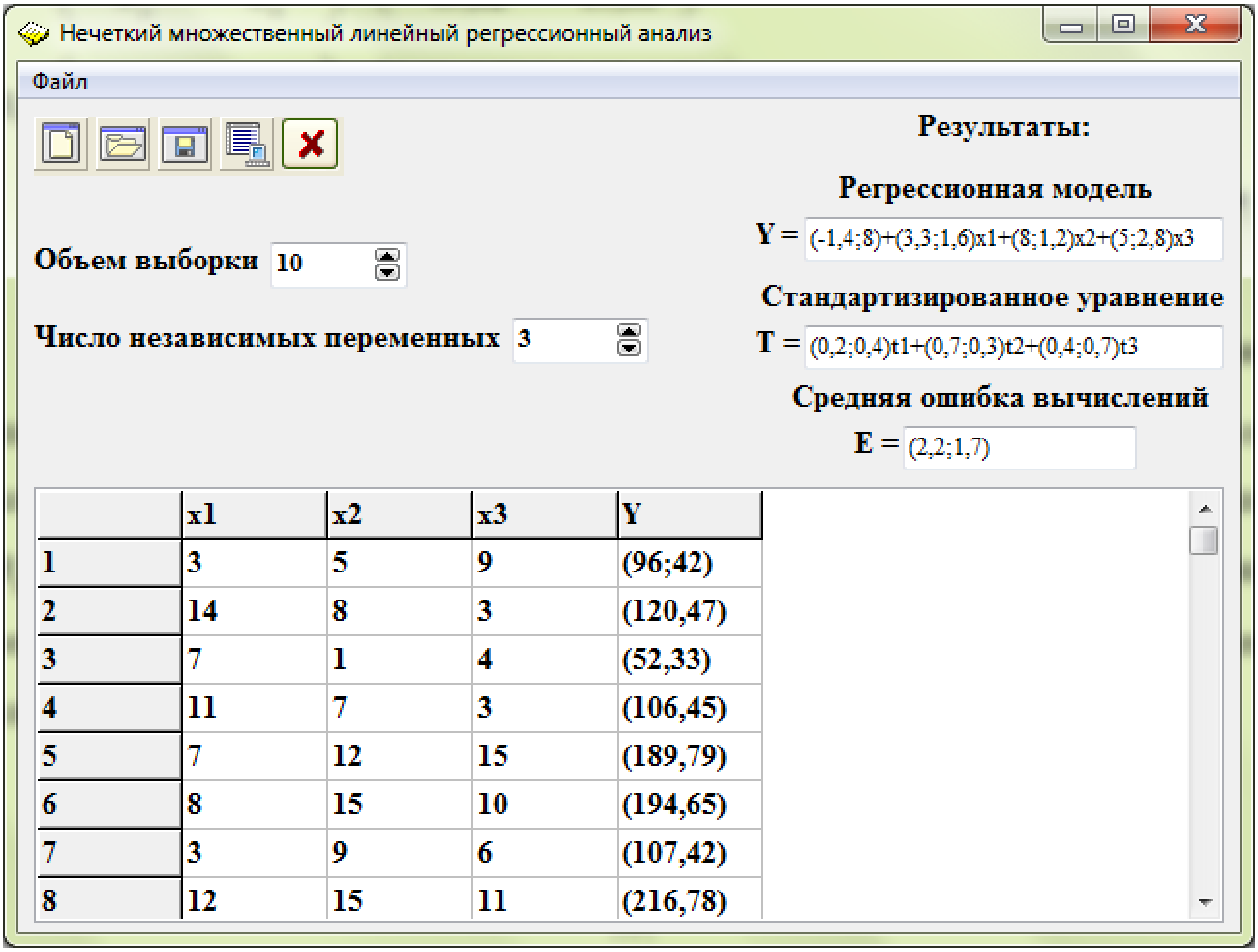


Рисунок 4.14 – Форма для проведения нечеткого множественного линейного регрессионного анализа

Алгоритм функционирования нечеткого регрессионного моделирования представлен на рисунке 4.15. На шаге инициализации производится присвоение начальных значений переменным, а также задание объема выборки (числа наблюдений) и количества независимых переменных (в случае множественной модели). Последующие этапы повторяются циклически до тех пор, пока не будет выполняться условие завершения обработки.

Исходные данные программа получает из входного файла или берет значения, введенные пользователем. В системе реализован доступ к файлам, состоящим из записей. Предобработка используется для преобразования введенной информации, изменения значений переменных [38].

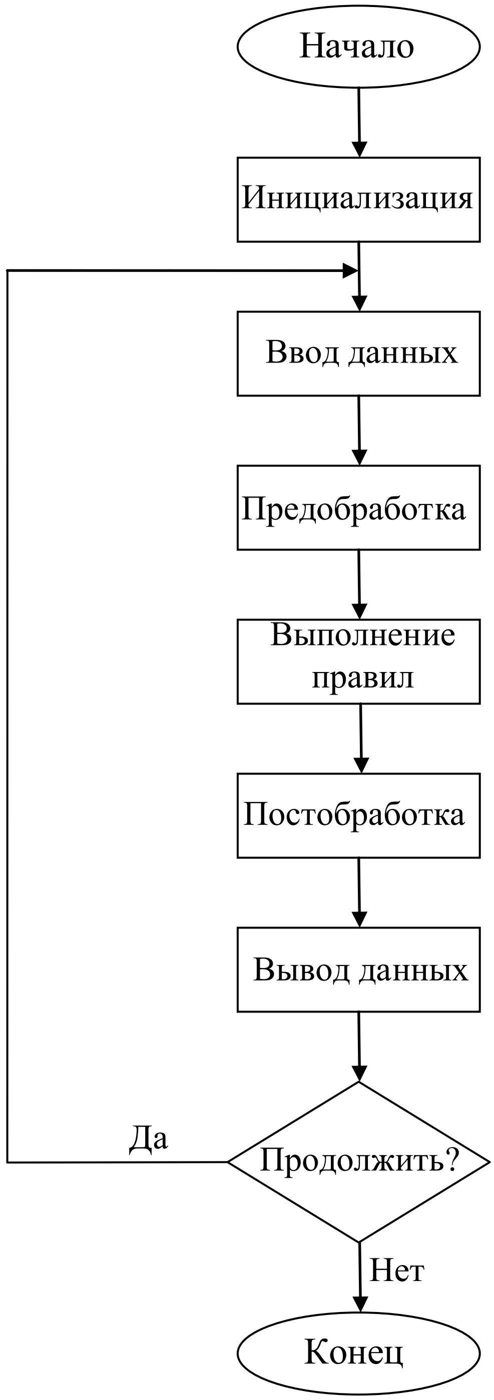
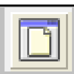
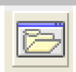
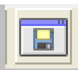
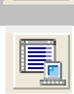


Рисунок 4.15 – Алгоритм нечеткого регрессионного моделирования

Этап проведения нечеткого регрессионного моделирования предназначен для нахождения параметров нечеткой линейной парной или множественной регрессии. Последовательность выполняемых действий была описана в главе 3. На стадии постобработки реализуются операции преобразования результатов вывода, присвоения значений внутренним переменным, используемым для записи результатов в файл. Кроме того на этом шаге находится коэффициент корреляции нечеткой парной регресии и коэффициенты стандартизированного уравнения нечеткой множественной регресионной модели, определяется средняя ошибка вычислений. На последнем этапе проверяется условие завершения работы.

В левой верхней части окна расположена инструментальная панель с кнопками, используемыми для выполнения основных операций. Ниже приведена таблица 4.1 назначений каждой из этих клавиш.

Таблица 4.1 – Назначение кнопок программы регрессионного анализа

Кнопка (пункт меню)	**Назначение**
	создать новый расчет
	загрузить данные из файла
	сохранить результаты анализа в файл
	провести регрессионный анализ данных
	очистить экран
Файл→Выход	выход из программы с сохранением результатов

При отборе существенных переменных с помощью автоассоциативной нейронной сети (рис. 4.16) осуществляется следующая последовательность шагов:

1. создание нового расчета;
2. определение объема выборки и исходного числа независимых переменных;
3. ввод значений с клавиатуры или загрузка данных из файла;
4. проведение нейросетевого анализа, настройка весов или загрузка ранее сохраненных весов из файла;
5. определение значимых переменных;

6. сохранение результатов.

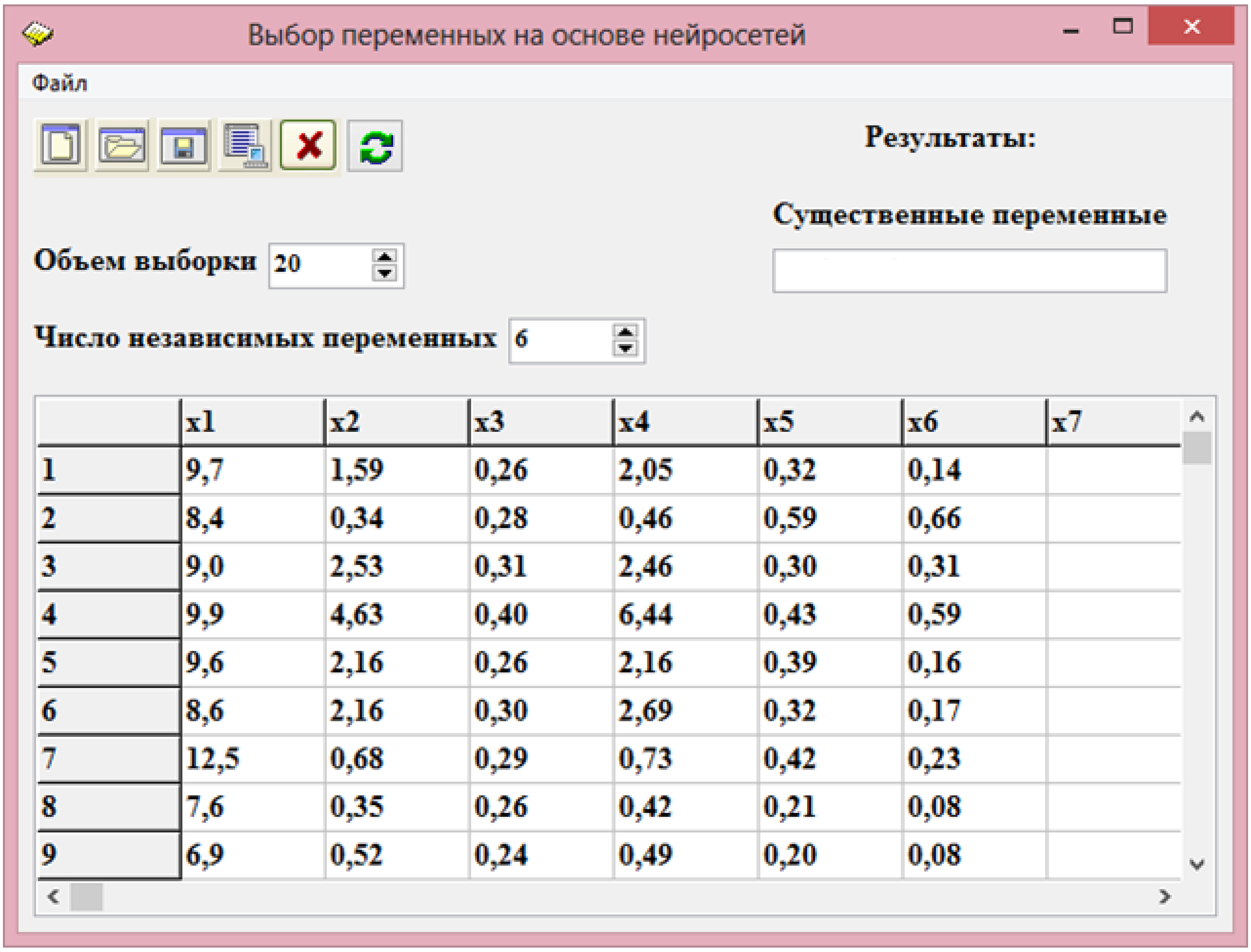
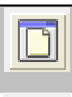
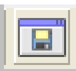
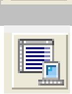


Рисунок 4.16 – Отбор существенных переменных

Выбор команд из меню можно осуществить как мышью, так и с помощью клавиатуры. В таблице 4.2 указано назначение пункта меню “Файл” и каждой из кнопок, которые отвечают за создание нового расчета, загрузку данных из файла и сохранение результатов в файл, проведение отбора значимых объясняющих переменных, сохранение и загрузку весовых коэффициентов, очистку экрана и выполнение выхода из программы с возможностью сохранения текущих результатов.

В разделе представлено описание разработанного комплекса программ для нечеткого регрессионного моделирования на основе предложенных алгоритмов проведения парного и множественного нечеткого линейного регрессионного анализа, а также отбора существенных переменных с помощью автоассоциативной нейронной сети.

Таблица 4.2 – Назначение кнопок программы нейросетевого анализа

Кнопка (пункт меню)	**Назначение**
	создать новый расчет
	загрузить данные из файла
	сохранить результаты в файл
	провести отбор значимых переменных
	очистить экран
	сохранить/загрузить веса
Файл→Выход	осуществить выход из программы с возможностью сохранения результатов

4.3. Анализ данных на основе приближенной информации по выпускаемой лакокрасочной продукции

Разработанная информационная система использовалась для анализа данных по выпускаемой лакокрасочной продукции и проведения оценки качества товаров с целью принятия управленческих решений по совершенствованию технологических процессов. Проведение интеллектуального анализа данных включало следующие этапы:

1. Подготовка данных, которая заключалась в предварительной обработке информации для проведения нечеткого регрессионного моделирования. В таблице 4.3 представлены исходные данные, в которых в качестве показателей (независимых переменных) выступают проценты наполнителей, растворителей и связующих веществ эмалированной краски, а результирующая величина (зависимая переменная) представляет собой примерное качество выпускаемой продукции и принадлежит множеству нечетких чисел *L*-*R*-типа.

Таблица 4.3 – Исходные данные

x_1	11	10	11	12	13,5	14	15	16	17	17,5	19	20
x_2	10	10,5	12,5	12	13	13,5	14	16	14,5	15	17	16,5
x_3	12	12,5	13	14,5	16	16,5	17	18	19,5	20,5	21	22
μ_y	11,2	12,5	12,9	14,1	14,8	16,1	17,5	18,9	18,9	20	21,1	22,2
α_y	0,9	0,1	1	0,5	1,1	1,2	0,1	0,5	0,3	1,3	1,1	0,2
β_y	0,2	0,4	0,5	1,1	0,7	0,1	0,08	1,2	0,7	0,48	1,9	0,4

В результате обучения автоассоциативной нейронной сети с одним нейроном среднего слоя было получено сжатое представление независимых переменных: вектор $x = (11, 11.5, 13, 14, 15, 15.5, 16, 17, 18, 19, 21, 22)$.

2. Проведение нечеткого линейного регрессионного моделирования с целью нахождения оценок неизвестных параметров парной регрессии для полученного набора данных (табл. 4.4).

Таблица 4.4 – Отобранные данные для анализа

x_1	11	10	11	12	13,5	14	15	16	17	17,5	19	20
μ_y	11,2	12,5	12,9	14,1	14,8	16,1	17,5	18,9	18,9	20	21,1	22,2
α_y	0,9	0,1	1	0,5	1,1	1,2	0,1	0,5	0,3	1,3	1,1	0,2
β_y	0,2	0,4	0,5	1,1	0,7	0,1	0,08	1,2	0,7	0,48	1,9	0,4

Предварительно были вычислены все необходимые суммы (табл. 4.5):

Таблица 4.5 – Результаты вычислений

$\sum_{i=1}^{12} x_i$	$\sum_{i=1}^{12} x_i^2$	$\sum_{i=1}^{12} \mu_{yi}$	$\sum_{i=1}^{12} \alpha_{yi}$	$\sum_{i=1}^{12} \beta_{yi}$	$\sum_{i=1}^{12} x_i \mu_{yi}$	$\sum_{i=1}^{12} x_i \alpha_{yi}$	$\sum_{i=1}^{12} x_i \beta_{yi}$
193	3238,5	200,2	8,3	7,76	3356,6	133,85	132,85

Затем по формулам (3.11)-(3.19) были найдены выборочные характеристики уравнения регрессии:

$$\bar{x} = \frac{193}{12} = 16.08\,;\ \overline{x^2} = \frac{3238.5}{12} = 269.875\,;$$

$$\overline{\mu_y} = \frac{200.2}{12} = 16.7\,;\ \overline{\alpha_y} = \frac{8.3}{12} = 0.69,\ \overline{\beta_y} = \frac{7.76}{12} = 0.65\,;$$

$$\overline{x\mu_y} = \frac{3356.6}{12} = 279.72\,;\ \overline{x\alpha_y} = \frac{133.85}{12} = 11.15\,;\ \overline{x\beta_y} = \frac{132.85}{12} = 11.07\,;$$

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 269.875 - 258.57 = 11.3\,;$$

$$\begin{cases} \operatorname{cov}(x, \mu_y) = \overline{x\mu_y} - \bar{\mu}_y \bar{x} = 279.72 - 16.7 \cdot 16.08 = 10.85, \\ \operatorname{cov}(x, \alpha_y) = \overline{x\alpha_y} - \bar{\alpha}_y \bar{x} = 11.15 - 0.69 \cdot 16.08 = 0.05, \\ \operatorname{cov}(x, \beta_y) = \overline{x\beta_y} - \bar{\beta}_y \bar{x} = 11.07 - 0.65 \cdot 16.08 = 0.6. \end{cases}$$

Далее были определены параметры уравнения регрессии:

$$\begin{cases} \tilde{\mu}_{b1} = \dfrac{\operatorname{cov}(x, \mu_y)}{\sigma_x^2} = \dfrac{10.85}{11.3} = 1, \\ \tilde{\alpha}_{b1} = \dfrac{\overline{x\alpha_y} - \bar{\alpha}_y \bar{x}}{\overline{x^2} - \bar{x}^2} = \dfrac{0.05}{11.3} = 0.004, \\ \tilde{\beta}_{b1} = \dfrac{\overline{x\beta_y} - \bar{\beta}_y \bar{x}}{\overline{x^2} - \bar{x}^2} = \dfrac{0.6}{11.3} = 0.04. \end{cases}$$

$$\begin{cases} \tilde{\mu}_{b0} = \bar{\mu}_y - \tilde{\mu}_{b1}\bar{x} = 16.7 - 0{,}96 \cdot 16.08 = 0.6, \\ \tilde{\alpha}_{b0} = \bar{\alpha}_y - \tilde{\alpha}_{b1}\bar{x} = 0.69 - 0.004 \cdot 16.08 = 0.6, \\ \tilde{\beta}_{b0} = \bar{\beta}_y - \tilde{\beta}_{b1}\bar{x} = 0.65 - 0.04 \cdot 16.08 = 0{,}01. \end{cases}$$

В результате было получено уравнение регрессии Y по x (табл. 4.6):

$$\tilde{Y} = (0.6{,}0.6{,}0.01) + (1{,}0.004{,}0.04)x\,.$$

Таблица 4.6 – Полученные значения результирующей переменной

$\tilde{\mu}_y$	11,5	12,1	13,6	14,6	15,6	16,1	16,6	17,6	18,6	19,6	21,6	22,6
$\tilde{\alpha}_y$	0,65	0,65	0,65	0,66	0,66	0,66	0,67	0,67	0,68	0,68	0,68	0,69
$\tilde{\beta}_y$	0,45	0,47	0,53	0,57	0,61	0,63	0,65	0,69	0,73	0,77	0,85	0,89

Для определения тесноты связи между переменными Y и x был найден коэффициент корреляции на основе данных таблицы 4.4.

$$\sum_{i=1}^{12}\mu_{yi}^2 = 3483.88;\ \sum_{i=1}^{12}\alpha_{yi}^2 = 8.01;\ \sum_{i=1}^{12}\beta_{yi}^2 = 8.0968.$$

По формуле (3.28) в итоге был получен нечеткий коэффициент корреляции $R = (\mu_r, \alpha_r, \beta_r) = (0.68, 0.02, 0.4)$:

$$\mu_r = \frac{12\cdot 3356.6 - 193\cdot 200.2}{\sqrt{12\cdot 3238.5 - 193^2}\cdot\sqrt{12\cdot 3483.88 - 200.2^2}} = 0{,}68,$$

$$\alpha_r = \frac{12\cdot 133.85 - 193\cdot 8.3}{\sqrt{12\cdot 3238.5 - 193^2}\cdot\sqrt{12\cdot 8.01 - 8.3^2}} = \frac{4.3}{208.8} = 0{,}02,$$

$$\beta_r = \frac{12\cdot 132.85 - 193\cdot 7.76}{\sqrt{12\cdot 3238.5 - 193^2}\cdot\sqrt{12\cdot 8.0968 - 7.76^2}} = \frac{96.52}{240.96} = 0{,}4.$$

Для оценки адекватности полученной модели построим γ-срез по значениям найденного нечеткого коэффициента корреляции $R = (0.998, 0.156, 0.33)$ $R = (0.68, 0.02, 0.4)$:

$$\left[\underline{R}(\gamma), \overline{R}(\gamma)\right] = [0.68 - 0.02\cdot(1-\gamma), 0.68 + 0.4\cdot(1-\gamma)],\ \gamma\in[0,1],$$

и вычислим четкий коэффициент корреляции для каждого γ-среза, взяв за ε-точку середину отрезка $\left[\underline{R}(\gamma), \overline{R}(\gamma)\right]$:

$$r = \underline{R}(\gamma)\cdot\varepsilon + \overline{R}(\gamma)\cdot(1-\varepsilon),\ \varepsilon = \frac{\underline{R}(\gamma) + \overline{R}(\gamma)}{2},\ \gamma\in[0,1].$$

При $\gamma\in[0,1]$ $r\in[0.68, 0.87]$, что свидетельствует о тесной линейной связи между переменными Y и x и адекватности построенной модели.

3. Проверка построенной модели, включающая оценку ее качества путем нахождения средней абсолютной погрешности вычислений по данным таблиц 4.4 и 4.6.

$$\varepsilon = (\mu_\varepsilon, \alpha_\varepsilon, \beta_\varepsilon) = \left(\frac{\left|\sum_{i=1}^{12}\left(\tilde{\mu}_{yi} - \mu_{yi}\right)\right|}{12}, \frac{\left|\sum_{i=1}^{12}\left(\tilde{\alpha}_{yi} - \alpha_{yi}\right)\right|}{12}, \frac{\left|\sum_{i=1}^{12}\left(\tilde{\beta}_{yi} - \beta_{yi}\right)\right|}{12}\right) = (0.4, 0.4, 0.3).$$

По каждому наблюдению $i=\overline{1,12}$ был построен γ-срез для точного и регрессионного результирующего значения: $[\mu_y-\alpha_y(1-\gamma),\mu_y+\beta_y(1-\gamma)]$ и $[\tilde{\mu}_y-\tilde{\alpha}_y(1-\gamma),\tilde{\mu}_y+\tilde{\beta}_y(1-\gamma)]$ соответственно (рис. 4.17, 4.18).

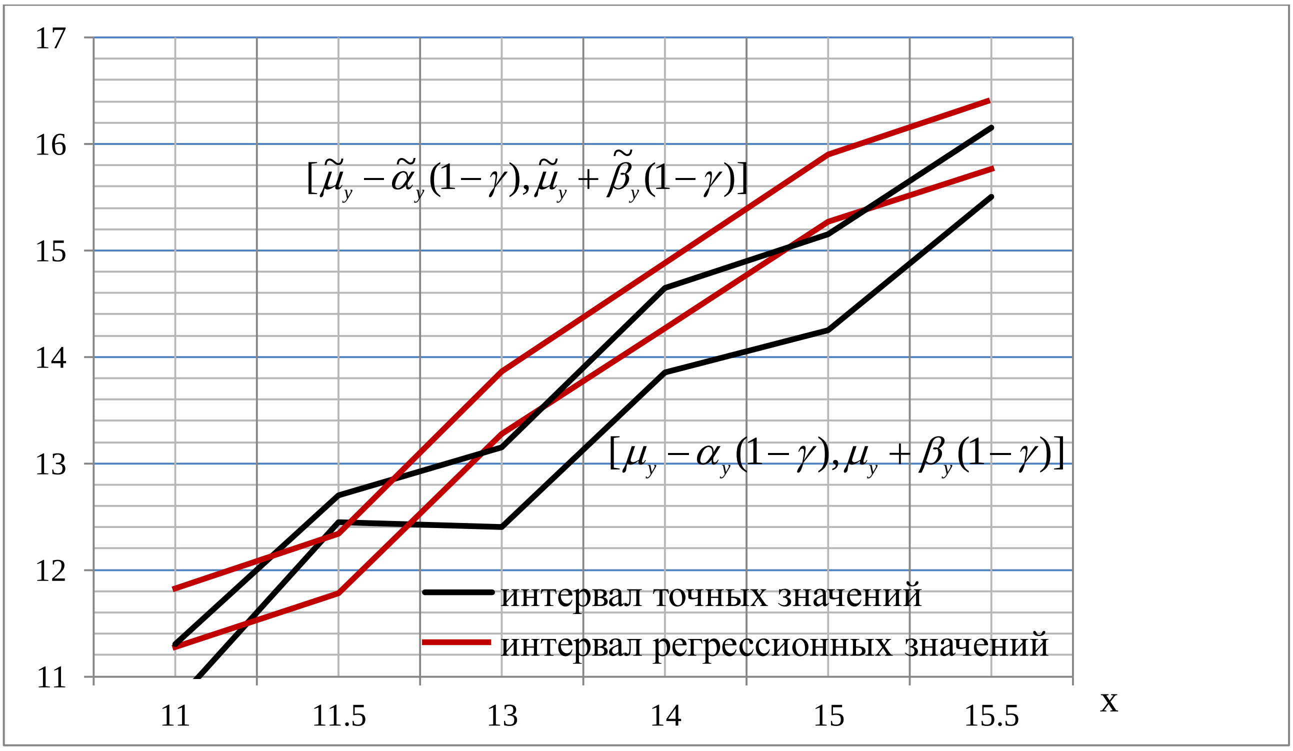


Рисунок 4.17 – Интервалы точных и регрессионных значений на 0.5-срезе

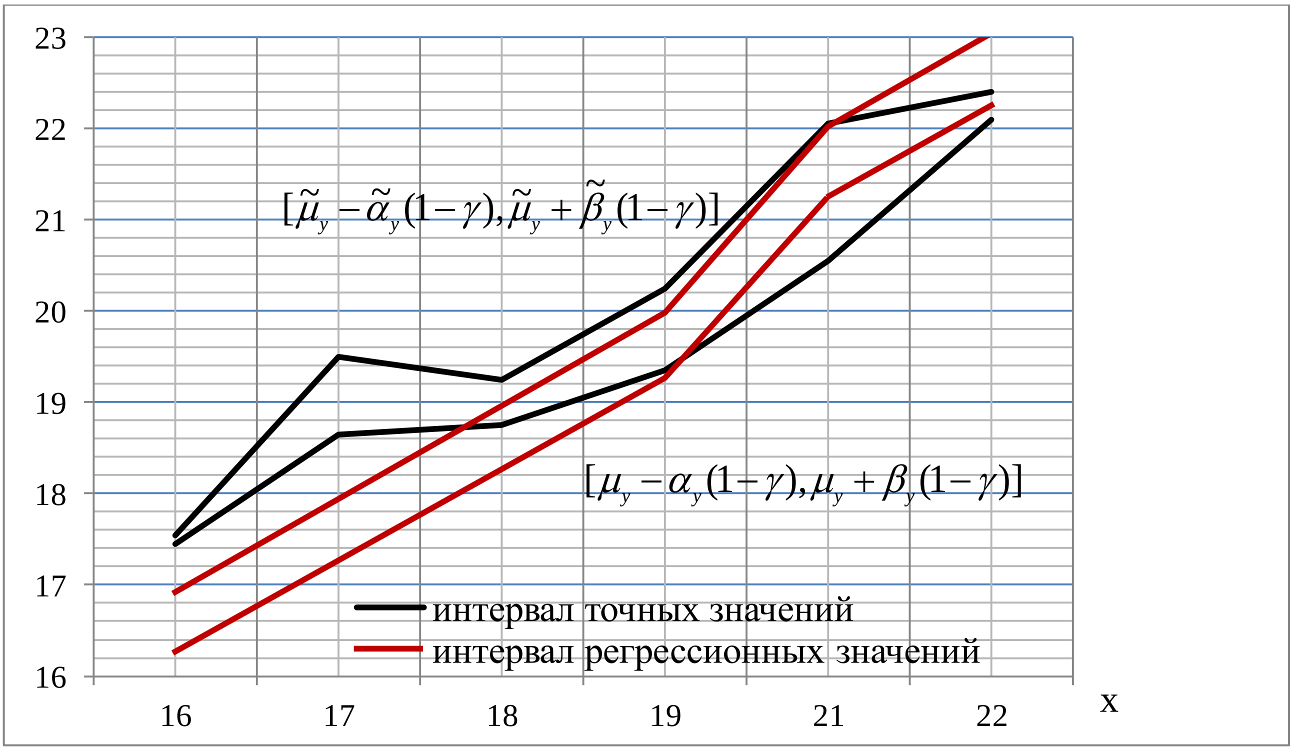


Рисунок 4.18 – Интервалы точных и регрессионных значений на 0.5-срезе

Точность вычисления в зависимости от выбранного γ-среза приведена на рисунке 4.20, где $\tilde{\varepsilon} = \left(\sum_{i=1}^{12} \left| Y - \tilde{Y} \right| \right) / 12$ при $Y = \mu_y + 0{,}5(1-\gamma)(\beta_y - \alpha_y)$, $\tilde{Y} = \tilde{\mu}_y + 0{,}5(1-\gamma)(\tilde{\beta}_y - \tilde{\alpha}_y)$. В ходе анализа было выявлено, что погрешность уменьшается при увеличении значения γ.

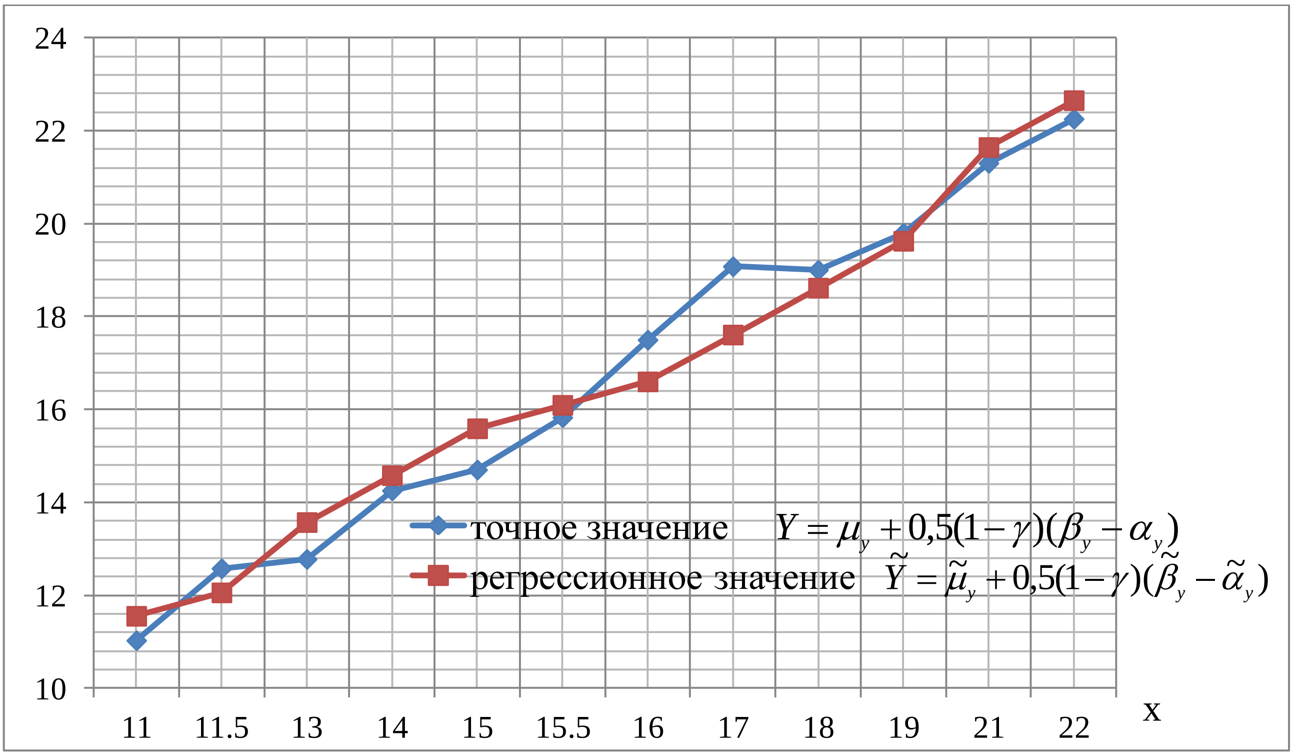


Рисунок 4.19 – График регрессии на 0.5-срезе

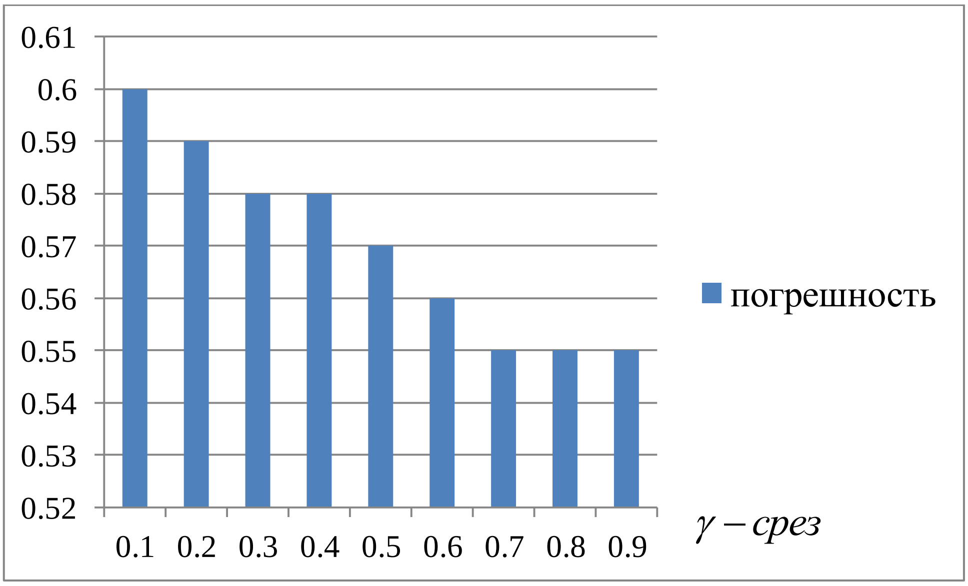


Рисунок 4.20 – Погрешность вычислений $\tilde{\varepsilon}$ в зависимости от γ-среза

4. Интерпретация проведенного моделирования, результаты которого были проанализированы и использованы при формировании бюджета закупки сырья.

Выводы по главе 4

1. Разработанная информационная система интеллектуального анализа данных основана на технологии информационного хранилища и способна обрабатывать большие массивы данных, осуществлять автоматизированный поиск ранее неизвестных закономерностей и скрытых и неочевидных правил в базе данных. Полученные знания помогают оптимизировать процессы деятельности предприятия и могут быть использованы для принятия решений.

2. Предложенный подход к проведению интеллектуального анализа данных предусматривает возможность для извлечения данных из информационного хранилища, их предварительную обработку и анализ.

3. Разработанное программное обеспечение для проведения нечеткого регрессионного анализа данных и моделирования осуществляет реализацию калькулятора нечетких чисел для вполнения различных арифметических операций над нечеткими числами L-R-типа и построения графиков их функций принадлежности, проведение отбора существенных независимых переменных на основе автоассоциативных нейронных сетей, реализацию нечеткого линейного парного и множественного регрессионного анализа, построение стандартизированного уравнения нечеткой множественной линейной регрессии.

4. Проведенный вычислительный эксперимент показывает хорошее качество предложенных регрессионных моделей.

ЗАКЛЮЧЕНИЕ

В диссертационной работе были получены следующие основные результаты:

1. Совокупность теоретических результатов, касающихся свойств арифметических операций над нечеткими числами *L*-*R*-типа и существования алгебраических структур, что позволяет осуществлять вычисления при построении нечетких регрессионных моделей.

2. Формулы для нахождения инволюции, произведения нечеткого числа на нечеткий ноль и единицу, формула квадрата нечеткого числа, а также произведения суммы двух нечетких чисел и произведения квадрата нечеткого числа на обратное ему число, которые позволяют проводить вычисления в ходе нечеткого регрессионного моделирования.

3. Оценка параметров нечеткой парной и множественной линейной регрессионной модели с помощью метода наименьших квадратов, оценка адекватности построенной модели на основе анализа нечеткого коэффициента корреляции и остаточной последовательности ошибок вычисления, определение точности модели путем вычисления средней относительной ошибки.

4. Альтернативные подходы к выявлению множества существенных переменных в рамках нечеткого регрессионного моделирования, основанные на нечетком коэффициенте корреляции, стандартизированном уравнении нечеткой множественной линейной регрессии и применении автоассоциативных нейронных сетей, «работающих» с приближенной информацией, что позволяет обеспечить комплексность анализа данных на различных этапах процесса выявления закономерностей в данных.

5. Информационная система интеллектуального анализа данных и структура программного комплекса, включающего блок нечеткой арифметики, который может использоваться как самостоятельное приложение, и средства для проведения нечеткого линейного регрессионного моделирования.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Аббакумов В.Л. Бизнес-анализ информации. Статистические методы / В.Л. Аббакумов. – М.: Экономика, 2009. – 194 с.

2. Абдрашитов Р.Т. Структура задачи интеллектуального анализа данных / Р.Т. Абдрашитов, Ю.В. Полищук // Вестник ОГУ. Серия Естественные и технические науки. – Оренбург: ИПЦ ОГУ, 2005. – №10. – С. 116-122.

3. Аверкин А.Н. Нечеткие множества в моделях управления и искусственного интеллекта / А.Н. Аверкин. – М.: Книга по требованию, 2012. – 312 с.

4. Айвазян С.А. Прикладная статистика. Основы эконометрики / С.А. Айвазян. – М.: ЮНИТИ, 2001. – 432 с.

5. Алиев Р.А. Управление производством при нечеткой исходной информации / Р.А. Алиев, А.Э. Церковный, Г.А. Мамедова. – М.: Энергоатомиздат, 1991. – 240 с.

6. Афанасьев В.Н. Эконометрика в пакете Statistica 6.0 / В.Н. Афанасьев, А.П. Цыпин. – Оренбург: ГОУ ОГУ, 2010. – 198 с.

7. Бабешко Л.О. Основы эконометрического моделирования / Л.О. Бабешко. – М.: КомКнига, 2006. – 432 с.

8. Барсегян А.А. Анализ данных и процессов / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И. Холод. – Спб.: БХВ-Петербург, 2009. – 512 с.

9. Барсегян А.А. Методы и модели анализа данных: OLAP и Data Mining / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И. Холод. – Спб.: БХВ-Петербург, 2004. – 336 с.

10. Барсегян А.А. Технология анализа данных: Data Mining, Visual Mining, Text Mining, OLAP / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И. Холод. – Спб.: БХВ-Петербург, 2007. – 384 с.

11. Башмаков А.И. Интеллектуальные информационные технологии: учеб. пособие / А.И. Башмаков, И.А. Башмаков. – М.: Издат-во МГТУ им. Н.Э. Баумана, 2005. – 384 с.

12. Белов В.С. Информационно-аналитические системы. Основы проектирования и применения: учеб. Пособие / В.С. Белов. – М.: МЭСИ, 2005. – 111 с.

13. Бережной Е.В. Математические методы моделирования экономических систем / Е.В. Бережной, В.И. Бережная. – М.: Финансы и статистика, 2001. – 368 с.

14. Блюмин С.Л. Базовые алгебраические структуры универсальных интеллектуальных информационно-аналитических систем / С.Л. Блюмин, А.К. Погодаев, П.В. Сараев // Международная научно-практическая конференция "Теория активных систем": Труды. – Т. 1. – М.: ИПУ РАН, 2009. – С. 42-45.

15. Борисов В.В. Нечеткие модели и сети / В.В. Борисов, В.В. Круглов, А.С. Федулов. – М.: Горячая линия – Телеком, 2007. – 284 с.

16. Борисов В.В. Основы нечеткой арифметики: учеб. пособие / В.В. Борисов, А.С. Федулов, М.М. Зернов. – М.: Горячая линия – Телеком, 2014. – 98 с.

17. Боровиков В.П. Прогнозирование в системе STATISTICA в среде Windows. Основы теории и интенсивная практика на компьютере: учебное пособие / В.П. Боровиков, Г.И. Ивченко. – М.: Финансы и статистика, 1999. – 368 с.

18. Бурбаки Н. Алгебра. Алгебраические структуры. Линейная и полилинейная алгебра / Н. Бурбаки. – М.: Оникс, 2012. – 516 с.

19. Бутакова М.М. Экономическое прогнозирование: методы и приемы практических расчетов / М.М. Бутакова. – М.: КНОРУС, 2010. – 168 с.

20. Виллемер А. Программирование на С++ / Пер. с англ. М. Райтман. – М.: Эксмо, 2013. – 528 с.

21. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М.: Высшая школа, 2003. – 479 с.

22. Доугерти К. Введение в эконометрику / Пер. с англ. Е.Н. Лукаш, О.Ю. Шибалкин; науч.ред. О.О. Замков. – М.: ИНФРА-М, 2009. – 465 с.

23. Дуброва Т.А. Статистические методы прогнозирования в экономике / Т.А. Дуброва. – М.: МЭСИ, 2004. – 136 с.

24. Дьяконов В.П. Современные зарубежные микрокалькуляторы / В.П. Дьяконов. – М.: СОЛОН-Р, 2002. – 400 с.

25. Ежов А.А. Нейрокомпьютинг и его приложения в экономике и бизнесе. /А.А. Ежов, С.А. Шумский. – М.:МИФИ, 1998. – 224с.

26. Елисеева И.И. Практикум по эконометрике: учебное пособие / И.И. Елисеева [и др.]. – М.: Финансы и статистика, 2002. – 192 с.

27. Елисеева И.И. Эконометрика / И.И. Елисеева. – М.: Проспект, 2011. – 288 с.

28. Исаев Г.Н. Проектирование информационных систем / Г.Н. Исаев. – М.: Омега-Л, 2012. – 342 с.

29. Карасев А.И. Курс высшей математики для экономических вузов. Теория вероятностей и математическая статистика. Линейное программирование / А.И. Карасев, З.М. Аксютина, Т.И. Савельева. – М.: Высш. школа, 1982. – Ч. 2. – 320 с.

30. Каширина И.Л. Нейросетевые технологии / И.Л. Каширина. – Воронеж: ИПЦ ВГУ, 2008. – 70 с.

31. Кендалл М. Многомерный статистический анализ и временные ряды / М. Кендал, А. Стьюарт. – М.: Наука, 1976. – 736 с.

32. Кириллов В.В. Введение в реляционные базы данных / В.В. Кириллов, Г.Ю. Громов. – Спб.: БХВ-Петербург, 2009. – 464 с.

33. Коваленко В.В. Проектирование информационных систем / В.В. Коваленко. – М.: Форум, 2012. – 320 с.

34. Коннолли Т. Базы данных. Проектирование, реализация и сопровождение. Теория и практика / Т. Коннолли, К. Бегг. – М.: Издательский дом “Вильямс”, 2003. – 1436 с.

35. Кононюк А.Е. Дискретная математика. Алгебры (четкие и нечеткие) / А.Е. Кононюк. – К.: Освита Украины, 2011. – Кн. 2, ч. 1. – 452 с.

36. Кононюк А.Е. Дискретная математика. Множества, отношения, пространства (четкие и нечеткие) / А.Е. Кононюк. – К.: Освита Украины, 2011. – Кн. 1, ч. 2. – 536 с.

37. Конышева Л. К. Основы теории нечетких множеств. – Спб.: Питер, 2011. – 192 с.

38. Корнеев В.В. Базы данных. Интеллектуальная обработка информации / В.В. Корнеев, А.Ф. Гарев, С.В. Васютин, В.В. Райх. – М.: Нолидж, 2000. – 352 с.

39. Кофман А. Введение в теорию нечетких множеств / А. Кофман. – М.: Радио и связь, 1982. – 432 с.

40. Кофман А. Введение теории нечетких множеств в управлении предприятиями / А. Кофман, Х. Хил Алуха. – Минск: Высш. школа, 1992 – 224 с.

41. Красс М.С. Математика для экономистов / М.С. Красс, Б.П. Чупрынов. – СПб.: Питер, 2010. – 464 с.

42. Кремер Н.Ш. Эконометрика: учебник для студ. вузов / Н.Ш. Кремер, Б.А. Путко. – М.: ЮНИТИ, 2002. – 311 с.

43. Кричевский М.Л. Интеллектуальный анализ данных в менеджменте / М.Л. Кричевский. – Спб.: Питер, 2005. – 208 с.

44. Кричевский М.Л. Интеллектуальные методы в менеджменте / М.Л. Кричевский. – Спб.: Питер, 2005. – 304 с.

45. Круг П.Г. Нейронные сети и нейрокомпьютеры / П.Г. Круг. – М.: Московский энергетический институт, 2002. – 176 с.

46. Круглов В.В. Искусственные нейронные сети. Теория и практика / В.В. Круглов, В.В. Борисов. – М.: Горячая линия – Телеком, 2001. – 382 с.

47. Круглов В.В. Нечеткая логика и искусственные нейронные сети / В.В. Круглов, М.И. Дли, Р.Ю. Голунов. – М.: Физматлит, 2001. – 224 с.

48. Крянев А.В. Математические методы обработки неопределенных данных / А.В. Крянев, Г.В. Лукин. – М.: Физматлит, 2006. – 216 с.

49. Кузин А.В. Базы данных: учеб. пособие / А.В. Кузин, С. Левонисова. – М.: Издательский центр «Академия», 2008. – 320 с.

50. Кулаичев А.П. Методы и средства комплексного анализа данных / А.П. Кулаичев. – М.: Форум, 2011. – 512 с.

51. Леденева Т. М. Обработка нечеткой информации / Т. М. Леденева. – Воронеж: Воронежский государственный университет, 2006. – 233 с.

52. Леденева Т.М. Нечеткая множественная линейная регрессионная модель для симметричных нечетких чисел L-R-типа / Т.М. Леденева, Н.В. Сапкина // Современная экономика: проблемы и решения: науч.-практ. журнал.– Воронеж: ИПЦ ВГУ, 2011. – № 10. – С. 174-181.

53. Лугинин О.Е. Экономико-математические методы и модели: теория и практика с решением задач / О.Е. Лугинин, В.Н. Фомишина. – Ростов н/Д: Феникс, 2009. – 440с.

54. Львович Я.Е. Нейросетевой подход к отбору наиболее информативных признаков для функционального диагностирования жидкостных ракетных двигателей. / Я.Е. Львович, И.Л. Каширина, А.А. Шостак // Вестник ВГТУ. – 2012. – Том 8. № 8. – С. 21-23.

55. Магнус Я.Р. Эконометрика: начальный курс / Я.Р. Магнус. – М.: Дело, 2000. – 399 с.

56. Марков А.В. Создание универсального хранилища данных промыслово-геофизических исследований / А.В. Марков, Д.Е. Черкунов, В.С. Шерстнев // Вестник науки Сибири. Серия Информационные технологии и системы управления. – Томск: 2011. – №1 (1). – С. 308-314.

57. Математико-статистические методы исследования взаимосвязей в экономике: Из теории и практики ГДР / Пер. с нем. А.Г. Закурдаева, Х.Н. Цаллагова; науч.ред. К. Отто, В.В. Швыркова. – М.: Статистика, 1977. – 181 с.

58. Могиленко А.В. Элементарные понятия теории нечетких множеств / А.В. Могиленко, А.В. Балуев. – Новосибирск, 2003. – 40 с.

59. Мхитарян В.С. Эконометрика: учеб. пособие / В.С. Мхитарян. – М.: Проспект, 2010. – 384 с.

60. Нестеров С.А. Базы данных. Интеллектуальный анализ данных / С.А. Нестеров. – Спб.: Изд-во Политехн. ун-та, 2011. – 272 с.

61. Нечеткие множества в моделях управления и искусственного интеллекта / под ред. Д.А. Поспелова. – М.: Наука, 1986. – 288 с.

62. Нечеткие множества и теория возможностей. Последние достижения: перевод с англ. / под ред. Р.Р. Ягера. – М.: Радио и связь, 1986. – 408 с.

63. Новак В. Математические принципы нечеткой логики / В. Новак, И. Перфильева, И. Мочкорж. – М.: Физматлит, 2006. – 352 с.

64. Носко В.П. Эконометрика. Элементарные методы и введение в регрессионный анализ временных рядов / В.П. Носко. – М.: ИЭПП, 2004. – 501 с.

65. Осовский С. Нейронные сети для обработки информации / С. Осовский. – М.: Финансы и статистика, 2004. – 344 с.

66. Павловская Т.А. C/C++: Программирование на языке высокого уровня / Т.А. Павловская. – Спб.: Питер, 2003. – 460 с.

67. Пегат А. Нечеткое моделирование и управление / А. Пегат. – М.: Бином, 2009. – 798 с.

68. Пирогов В.Ю. Информационные системы и базы данных. Организация и проектирование / В.Ю. Пирогов. – Спб.: БХВ-Петербург, 2009. – 528 с.

69. Погодаев А.К. Универсальное информационное и программное обеспечение для аналитической обработки данных / А.К. Погодаев, П.В. Сараев, Е.П. Татаринов // Информационные технологии моделирования и управления. – Воронеж: 2010. – №4 (63). – С. 543-550.

70. Рутковская Р. Нейронные сети, генетические алгоритмы и нечеткие системы / Р. Рутковская, М. Пилиньский, Л. Рутковский. – М.: Горячая линия – Телеком, 2013. – 384 с.

71. Рыжов А.П. Элементы теории нечетких множеств и измерения нечеткости / А.П. Рыжов. – М.: Диалог-МГУ, 2003. – 81 с.

72. Сапкина Н.В. Метод наименьших квадратов для нечеткой линейной регрессионной модели / Н.В. Сапкина // Актуальные проблемы прикладной

математики, информатики и механики: сб. тр. междунар. конф., Воронеж, 26-28 сентября 2011 г. – Воронеж: ИПЦ ВГУ, 2011.– С. 344-345.

73. Сапкина Н.В. Нечеткая линейная множественная регрессионная модель с четкими коэффициентами. Отбор значимых переменных модели с помощью нейросетей / Н.В. Сапкина // Системы управления и информационные технологии. – Москва-Воронеж: ИПЦ «Научная книга», 2013. – №4 (54). – С. 27-30.

74. Сапкина Н.В. Нечеткая парная линейная регрессия и корреляция / Н.В. Сапкина // Современная экономика: проблемы и решения: науч.-практ. журнал.– Воронеж: ИПЦ ВГУ, 2013.– № 10 (46). – С. 178-189.

75. Сапкина Н.В. Нечеткие линейные регрессионные модели. Метод наименьших квадратов для модели с четкими входами и гауссовым нечетким выходом / Н.В. Сапкина // Глобальная научная интеграция: сб. материалов междунар. науч.-практ. конф., Тамбов, 30 июня 2011 г.– Тамбов: ТМБпринт, 2011.– С. 68-71.

76. Сапкина Н.В. Нечеткий парный линейный регрессионный анализ / Н.В. Сапкина // Актуальные проблемы прикладной математики, информатики и механики: сб. тр. междунар. конф., Воронеж, 26-28 ноября 2012 г. – Воронеж: ИПЦ ВГУ, 2012.– Ч. 1. – С. 331-334.

77. Сапкина Н.В. Свойства группоида нечетких чисел LR-типа / Н.В. Сапкина // Современные методы прикладной математики, теории управления и компьютерных технологий: Сборник трудов VI Международной конференции, Воронеж, 10-16 сентября 2013г. – Воронеж: ИПЦ ВГУ, 2013. – С. 216-218.

78. Сапкина Н.В. Свойства операций над нечеткими числами / Н.В. Сапкина // Вестник ВГУ. Серия Системный анализ и информационные технологии. – Воронеж: ИПЦ ВГУ, 2013. – №1. – С. 23-28.

79. Сапкина Н.В. Нечеткий парный линейный регрессионный анализ / Н.В. Сапкина, А.А. Татаринцев // Инженерия знаний. Представление знаний: состояние и перспективы: материалыВсероссийской молодежной научной

школы, Воронеж, 29-30 июня 2012 г. – Воронеж: ИПЦ «Научная книга», 2012. – С. 260-261.

80. Себер Дж. Линейный регрессионный анализ / Пер. с англ. В.П. Носко; науч.ред. М.Б. Малютова. – М.: Мир, 1980. – 456 с.

81. Семакин И.Г. Информационные системы и модели / И.Г. Семакин, Е.К. Хеннер. – М.: Бином, Лаб. знаний, 2007. – 304 с.

82. Стрижов В.В. Методы выбора регрессионных моделей / В.В. Стрижов, Е.А. Крымова. – М.: ВЦ РАН, 2010. – 60 с.

83. Стрижов В.В. Методы индуктивного порождения регрессионных моделей / В.В. Стрижов. – М.: ВЦ РАН, 2008. – 61 с.

84. Тарасик В.П. Математическое моделирование технических систем / В.П. Тарасик. – Минск: ДизайнПРО, 1997. – 640 с.

85. Трофимов В.В. Информационные системы и технологии в экономике и управлении / В.В. Трофимов. – М.: Юрайт, 2013. – 544 с.

86. Трояновский В.М. Математическое моделирование в менеджменте / В.М. Трояновский. – М.: Русская деловая литература, 1999. – 240 с.

87. Уоссерман Ф. Нейрокомпьютерная техника: теория и практика / Ф. Уоссерман. – М.: Мир, 1992. – 240 с.

88. Ускова О.Ф. Программирование алгоритмов обработки данных / О.Ф. Ускова, И.Е. Воронина, М.В. Бакланов, В.М. Мельников. – Спб.: БХВ-Петербург, 2004. – 192 с.

89. Ферстер Э. Методы корреляционного и регрессионного анализа / Э. Ферстер, Б. Ренц; перевод с нем. В.М. Ивановой. – М.: Финансы и статистика, 1983. – 304 с.

90. Хабибуллин И.Ш. Программирование на языке высокого уровня. C/C++ / И.Ш. Хабибуллин. – Спб.: БХВ-Петербург, 2006. – 512 с.

91. Хайкин С. Нейронные сети: полный курс / Пер. с англ. Н.Н. Куссуль, А.Ю. Шелестова; науч.ред. Н.Н. Куссуль. – М.: Издательский дом «Вильямс», 2006. – 1104 с.

92. Халафян А. Statistica 6. Статистический анализ данных / А. Халафян. – Спб.: Бином-Пресс, 2010. – 528 с.

93. Хоббс Л. Oracle9i: разработка и эксплуатация хранилищ баз данных / Л. Хоббс, С. Хилсон, Ш. Лоуенд. – М.: Кудиц-Образ, 2004. – 592 с.

94. Шашков В.Б. Прикладной регрессионный анализ (многофакторная модель) / В.Б. Шашков. – Оренбург: ГОУ ОГУ, 2003. – 363 с.

95. Яновский Л.П. Введение в эконометрику: учебное пособие / Л.П. Яновский, А.Г. Буховец; под ред. Л.П. Яновского. – М.: КНОРУС, 2007. – 256 с.

96. Ярушкина Н.Г. Основы теории нечетких и гибридных систем / Н.Г. Ярушкина. – М.: Финансы и статистика, 2009. – 320 с.

97. Ясницкий Л.Н. Введение в искусственный интеллект / Л.Н. Ясницкий. – М.: Издательский центр «Академия», 2005. – 176с.

98. Яхъяева Г.Э. Нечеткие множества и нейронные сети / Г.Э. Яхъяева. – М.: Бином, Лаб. знаний, 2006. – 320 с.

99. Albrecht M. Approximation of functional relationships to fuzzy observations / M. Albrecht // Fuzzy Sets and Systems. – 1992. – V. 49(3). – P. 301-305.

100. Diamond P. Fuzzy least squares / P. Diamond // Information Science. – 1988. – V. 46. – P. 141-157.

101. Dubois D. Fuzzy Sets and Systems: Theory and Applications / D. Dubois, H. Prade. – Academic Press, 1980. – 393 p.

102. Hong D.H. Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output / D.H. Hong, C. Hwang, C. Ahn // Fuzzy Sets and Systems. – 2004. – V. 142. – P. 307-319.

103. Sakawa M. Multiobjective fuzzy linear regression analysis for fuzzy input–output data / M. Sakawa, H. Yano // Fuzzy Sets and Systems. – 1992. – V. 47. – P. 173-181.

104. Tanaka H. Exponential possibility regression analysis / H. Tanaka, H. Ishibuchi, S. Yoshikawa // Fuzzy Sets and Systems. – 1995. – V. 69(3). – P. 305-318.

105. Tanaka H. Identication of possibilistic linear systems by quadratic membership functions of fuzzy parameters / H. Tanaka, H. Ishibuchi // Fuzzy Sets and Systems. – 1991. – V. 41. – P. 145-160.

106. Tanaka H. Linear regression analysis with fuzzy model / H. Tanaka, S. Vegima, K. Asai // Transactions on Systems, Man and Cybernetics. – 1982. – V. 12(6). – P. 903-907.

107. Zadeh L. Fuzzy sets / L. Zadeh // Information and Control. – 1965. – V. 8. – P. 338-353.

108. Zimmermann H.J. Fuzzy Set Theory and its Applications / H.J. Zimmermann. – Kluwer Academic Publishers, 1997. – 429 p.

109. Yang M.S. Fuzzy least-squares linear regression analysis for fuzzy input–output data / M.S. Yang, T.S. Lin // Fuzzy Sets and Systems. – 2002. – V. 126(3). – P. 389-399.

110. Yang M.S. On cluster-wise fuzzy regression analysis / M.S. Yang, C.H. Ko // Transactions on Systems, Man and Cybernetics. – 1997. – V. 27(1). – P. 1-13.

Приложение А

Копия акта о внедрении результатов диссертационного исследования

ЗАО ЛЦ «АВС Фарбен»
Россия, 394000, г. Воронеж, пр-т Революции 1/1
тел./факс: +7 (473)255-30-01,255-30-34, 255-30-15, 255-30-25
info@abcfarben.ru

АКТ

о внедрении результатов диссертационного исследования
Н.В. Сапкиной «Восстановление закономерностей на основе нечетких регрессионных моделей» в процесс планирования качества продукции
ЗАО ЛЦ «АВС Фарбен»

Программный комплекс для нечеткого регрессионного моделирования, разработанный Сапкиной Натальей Владимировной в ходе диссертационного исследования «Восстановление закономерностей на основе нечетких регрессионных моделей» был внедрен в практическую деятельность ЗАО ЛЦ «АВС Фарбен» в июле 2013 года. Данный комплекс используется в качестве аналитической системы оценки качества выпущенной лакокрасочной продукции по нечетким показателям и для обоснования управленческих решений по совершенствованию технологических процессов.

Проведение многофакторного регрессионного анализа оценки качества лакокрасочной продукции в условиях нечеткой информации по качественным признакам позволяет выявить показатели, влияющие на качество производимой продукции и отражающие эффективность производственных процессов. Результаты моделирования были проанализированы и использованы при формировании бюджета закупки сырья и инвестиционного бюджета на 2014 год.

Руководитель отдела автоматизации бизнес процессов — Белобродский А.А.

Главный Бухгалтер — Епрынцева М.Ф.

«15» июля 2013г.

ИНН/КПП 3661047564 /366601001 ОГРН 1093668031016 от 24.07.2009 г. Р/С 40702810100250004103
в филиале ОАО Банка ВТБ в г. Воронеже К/С 30101810100000000835
БИК 042007835 ОКАТО 20401000000 ОКПО 61978211

Приложение Б

Копии свидетельств о государственной регистрации программ

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2013660211

«Реализация нечеткого множественного линейного регрессионного анализа»

Правообладатель: ***федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Воронежский государственный университет» (RU)***

Автор: ***Сапкина Наталья Владимировна (RU)***

Заявка № **2013618279**

Дата поступления **13 сентября 2013 г.**

Дата государственной регистрации

в Реестре программ для ЭВМ ***28 октября 2013 г.***

Руководитель Федеральной службы по интеллектуальной собственности

Б.П. Симонов

Рисунок 1 – Свидетельство о государственной регистрации программы для проведения нечеткого регрессионного анализа

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2013660206

«Отбор наиболее значимых факторов с помощью нейронной сети»

Правообладатель: *федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Воронежский государственный университет» (RU)*

Автор: *Сапкина Наталья Владимировна (RU)*

Заявка № **2013618061**

Дата поступления **10 сентября 2013 г.**

Дата государственной регистрации

в Реестре программ для ЭВМ ***28 октября 2013 г.***

Руководитель Федеральной службы по интеллектуальной собственности

Б.П. Симонов

Рисунок 2 – Свидетельство о государственной регистрации программы для отбора значимых переменных с помощью автоассоциативной нейронной сети