На правах рукописи

Лёзин Илья Александрович

# ПРОГРАММНЫЙ КОМПЛЕКС АППРОКСИМАЦИИ ДВУМЕРНЫХ ПЛОТНОСТЕЙ ВЕРОЯТНОСТИ

Специальность 05.13.18 – Математическое моделирование,
численные методы и комплексы программ

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Самара – 2009

Работа выполнена в Самарском государственном аэрокосмическом университете имени академика С.П. Королева на кафедре информационных систем и технологий

| | |
|---|---|
| Научный руководитель: | заслуженный работник высшей школы РФ,<br>доктор технических наук,<br>профессор Прохоров Сергей Антонович |
| Официальные оппоненты: | доктор физико-математических наук,<br>профессор Радченко Владимир Павлович |
| | доктор технических наук,<br>профессор Привалов Александр Юрьевич |
| Ведущая организация: | ОАО «Самарский научно-технический комплекс имени академика Н.Д. Кузнецова» |

Защита состоится "__"___________ 200__ г. в _____ часов на заседании диссертационного совета Д 212.215.05 при Самарском государственном аэрокосмическом университете имени академика С.П. Королева по адресу: 443086, г. Самара, Московское шоссе, 34

С диссертацией можно ознакомиться в библиотеке Самарского государственного аэрокосмического университета

Автореферат разослан "___"__________ 200__ г.

Ученый секретарь
диссертационного совета
д.т.н., профессор А.А. Калентьев

# ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы**

Исследование параметров различных объектов зачастую дает результат в виде больших массивов однородной информации, являющихся результатом многократных повторений измерений либо с целью повышения достоверности результата, либо с целью накопления большого количества опытных данных об исследуемом предмете. При этом объем выборки может достигать огромных размеров, и оперировать им становится не очень удобно. Кроме того, любые результаты несут в себе некоторую долю погрешности, вносимую по различным причинам. Если в условиях конкретной задачи можно исходить из предположения о том, что данная выборка распределена по какому-либо закону, пусть даже нам неизвестному, то в таком случае можно перейти от хранения информации в виде числовых массивов к хранению закона распределения числового ряда. Помимо удобства в хранении это позволит:

а) использовать аналитическое выражение закона распределения для дальнейшего анализа,

б) уменьшить влияние случайных погрешностей при получении данных на реальных объектах, сглаживая результаты,

в) упростить получение вероятностных характеристик числовых выборок и прочие характеристики законов распределения, например, моду, вероятность попадания отсчета в заданную область, минимальные границы области с заданной вероятностью попадания, условные вероятности и т.д., а также разбиение на отдельные составляющие в случае декомпозиции законов.

Аппроксимация плотностей вероятности методом моментов или параметрическими методами имеет некоторые недостатки, которые делают их неприменимыми в ряде случаев. Во-первых, мы не всегда можем предположить, какому именно закону подчинена имеющаяся у нас выборка, а во-вторых, исследуемые плотности вероятности могут существенно отличаться от имеющегося набора стандартных законов.

По этим причинам в качестве универсального метода восстановления аналитического выражения плотности вероятности можно предложить представление неизвестной плотности в виде суперпозиции базисных функций некоторого семейства с неизвестными коэффициентами. Такой подход нечувствителен к виду распределения входных данных, что расширяет область его применимости.

Предлагаемый метод основан на анализе исследуемой выборки, построении на ее основе двумерной гистограммы и вычислении некоторой непрерывной функции, проходящей через точки, вычисленные по столбцам гистограммы. В дальнейшем полученная гладкая функция может использоваться для нахождения числовых и функциональных характеристик выборки, а также для других видов анализа, оперирующих аналитическими выражениями плотности вероятности.

Данный метод может быть реализован с помощью аппроксимации плотности вероятности в некотором ортогональном базисе. Таким образом, неизвестная функция плотности вероятности будет представлена в виде конечной суммы ортогональных функций.

Кроме аппроксимации функций ортогональными рядами в последнее время все больше внимания уделяется приближению функций многих переменных с помощью линейных операций и суперпозиций функций одного переменного. Такое приближение осуществляется специальными формальными «устройствами» – нейронными сетями. Нейрон получает на входе вектор сигналов $\boldsymbol{x}$, вычисляет его скалярное произведение на вектор весов $\boldsymbol{w}$ и некоторую функцию одного переменного. Результат пересылается на входы других нейронов или передается на выход. Таким образом, нейронные сети вычисляют суперпозиции простых функций одного переменного и их линейных комбинаций. Для решения задачи аппроксимации с применением нейронной сети следует спроектировать структуру сети, адекватную поставленной задаче. В данной работе используется радиально-базисная сеть, являющаяся универсальным аппроксиматором.

Разработке методов аппроксимации неизвестной плотности вероятности ортогональными рядами и нейронными сетями, решению сопутствующих проблем и сравнению двух способов аппроксимации посвящена данная работа.

Вопросы разработки аппроксимативных методов и алгоритмов, а также вопросы, посвященные теории нейронных сетей в разное время исследовали Л.Деврой, Л.Дьерфи, В.Н.Вапник, Э.А.Надарая, С.А.Прохоров, Ф.П.Тарасенко, Н.Н.Ченцов, М.Розенблатт, А.Н.Горбань, С.Хайкин, С.Осовский и другие ученые.

Анализ существующих современных автоматизированных комплексов математических расчетов (Statistica, SPSS, MathLab, MathCad) показал, что они позволяют работать с одномерными ортогональными полиномами, однако в них отсутствуют алгоритмы представления функций, заданных в табличном виде, ортогональными рядами. При этом аппроксимация двумерных функций вообще не рассматривается. Также эти комплексы требуют дополнительной настройки или программирования для решения конкретных задач.

Существует довольно много универсальных программных пакетов для работы с нейронными сетями (Statistica Neural Networks, NeuroShell, Matlab Neural Network Toolbox, NeuroSolutions, BrainMaker). Однако для решения задач с помощью этих комплексов пользователь должен выполнить настройки нейронной сети (выбрать ее структуру, задать алгоритм обучения, подать на вход обучающие данные и т.д.), подходящие именно для конкретной решаемой задачи. Это, в свою очередь, требует от пользователя владения знаниями по теории нейронных сетей и умение программировать на языке используемого пакета.

В связи с этим актуальной представляется задача разработки алгоритмов аппроксимации двумерных плотностей вероятности ортогональными функциями и нейросетевыми моделями, а также построения комплекса программ, реализующего эти алгоритмы.

**Целью работы** является разработка алгоритмов и комплекса программ для аппроксимативного анализа двумерных плотностей вероятности в ортогональных базисах, представленных произведениями одномерных функций Лежандра, Чебышева, Лагерра и Эрмита, а также с помощью аппроксимации радиально-базисными сетями.

В соответствии с поставленной целью в диссертационной работе решаются следующие **задачи исследования**:

- анализ и сравнение имеющихся материалов в области аппроксимации двумерных функций и восстановления двумерных плотностей вероятности;

- разработка алгоритмов аппроксимации двумерных законов распределения с использованием ортогональных базисов и нейронных сетей;
- оценка результатов аппроксимации с помощью критерия Пирсона и величины погрешности;
- создание программного комплекса, реализующего разработанные алгоритмы;
- исследование и сравнительный анализ результатов аппроксимации ортогональными функциями и нейронными сетями;
- проведение экспериментальных исследований по обработке реальных данных с целью практического внедрения комплекса программ.

**Методы исследования**, используемые в диссертации, основаны на положениях теории вероятности и математической статистики, теории случайных процессов, теории оптимизации и аппроксимации, методах имитационного моделирования, численных методах, теории нейронных сетей.

**Научная новизна** работы заключается в следующих положениях:

- предложены модифицированные формулы расчета рекомендуемого числа коридоров в гистограмме по правилу Стёрджеса, методу Фридмана-Диакониса и методу Скотта для двумерного случая, а также выработаны рекомендации по выбору оптимального числа коридоров по осям;
- предложена методика аппроксимации двумерных плотностей вероятности семействами двумерных ортогональных функций, являющихся произведением одномерных ортогональных функций, с расчетом оптимальных значений коэффициентов масштаба для достижения минимума погрешности аппроксимации, а также предложена методика аппроксимации двумерных плотностей вероятности радиально-базисной сетью;
- предложена методика определения областей отрицательности полученных выражений и разработан ускоренный алгоритм нахождения базиса Грёбнера для решения систем двух полиномиальных уравнений с двумя неизвестными, а также разработана методика компенсации областей отрицательности функциями Гаусса через локализацию минимумов;
- выработаны рекомендации по применению предложенных моделей и выбору значений настраиваемых параметров для решения конкретных задач.

**Практическая ценность** заключается в разработке алгоритмического и программного обеспечения автоматизированного программного комплекса аппроксимативного анализа, позволяющего решать следующие задачи:

- моделирование двумерных случайных последовательностей с заданным законом распределения, являющимся произведением стандартных одномерных законов;
- аппроксимация двумерных плотностей вероятности ортогональными функциями, являющимися произведениями одномерных ортогональных базисов Лежандра, Чебышева, Лагерра и Эрмита;
- аппроксимации двумерных плотностей вероятности радиально-базисными сетями;
- обеспечение условий неотрицательности и нормировки полученных выражений;
- обработка данных натурного эксперимента.

**Положения, выносимые на защиту:**

- двумерная модификация правил Стёрджеса, Скотта и Фридмана-Диакониса для определения оптимального числа коридоров по осям гистограммы;
- методика и алгоритмы аппроксимации двумерных плотностей вероятности ортогональными функциями, являющимися произведениями одномерных ортогональных функций Лежандра, Чебышева, Лагерра и Эрмита, с использованием первичного приближения сплайнами;
- методика определения оптимальных значений коэффициентов масштаба при аппроксимации функциями Лагерра и Эрмита;
- ускоренный алгоритм нахождения базиса Грёбнера для решения системы двух полиномиальных уравнений с двумя неизвестными;
- методика нахождения и компенсации отрицательно определенных областей полученного аппроксимирующего выражения;
- методика и алгоритмы аппроксимации двумерных плотностей вероятности радиально-базисной сетью;
- программный комплекс аппроксимативного анализа двумерных плотностей вероятности ортогональными базисами и радиально-базисными сетями.

**Внедрение результатов работы**

Результаты работы внедрены в учебном процессе кафедры «Информационных систем и технологий» СГАУ при подготовке студентов по специальности 230102, а также в ООО «НТФ Протон» для исследования оптимальных пропорций компонентов при компаундировании бензинов.

**Апробация работы**

Основные положения и результаты работы докладывались и обсуждались на XXX Юбилейной Самарской областной студенческой научной конференции (Самара, 2004), международном симпозиуме «Надежность и качество» (Пенза, 2004), международной научно-технической конференции «Информационные, измерительные и управляющие системы (ИИУС-2005)» (Самара, 2005), международной научно-технической конференции, посвященной 110-летию изобретения радио и 75-летию Саратовского государственного технического университета «Радиотехника и связь» (Саратов, 2005), Всероссийской молодежной научной конференции с международным участием «VIII Королевские чтения» (Самара, 2005), Всероссийской межвузовской научно-практической конференции «Компьютерные технологии в науке, практике и образовании» (Самара, 2005), третьей международной научно-технической конференции «Радиотехника и связь» (Саратов, 2006), научно-технической конференции с международным участием «Перспективные информационные технологии в научных исследованиях, проектировании и обучении» (ПИТ-2006) (Самара, 2006), Всероссийской научной конференции «Инновационные технологии в управлении, образовании, промышленности» («АСТИНТЕХ-2007») (Астрахань, 2007), международной научно-технической конференции «Проблемы автоматизации и управления в технических системах» (Пенза, 2007), II межрегиональной научно-практической конференции «Информационные технологии в высшем профессиональном образовании» (Тольятти-Самара, 2007), четвертой международной научно-технической конференции «Радиотехника и связь» (Саратов, 2007), международном симпозиуме «Надежность и качество» (Пенза, 2009), международной научно-технической конференции «Проблемы и перспективы развития двигателестроения» (Самара, 2009).

**Публикации**

По результатам исследований опубликовано 18 печатных работ, в том числе 1 монография, 15 статей, из них 3 – в изданиях, определенных Высшей аттестационной комиссией Российской Федерации.

**Объем и структура работы**

Диссертация состоит из введения, трех глав и заключения. Основное содержание работы изложено на 112 страницах, включая 20 рисунков и 18 таблиц. Список использованных источников включает 67 наименований, 1 приложение размещено на 4 страницах.

## СОДЕРЖАНИЕ РАБОТЫ

**Во введении** показана актуальность темы диссертации, определена цель работы, изложена научная новизна и практическая значимость полученных результатов, сформулированы основные положения, выносимые на защиту.

**В первой главе** проанализированы основные принципы и особенности аппроксимативного подхода к решению задач обработки экспериментальной информации, рассмотрены эволюция и современное состояние существующих методов аппроксимации двумерных функций ортогональными рядами и моделями, построенными на использовании нейронных сетей, а также прочими методами параметрической аппроксимации.

Одной из самых сложных и плохо формализуемых задач, от правильного решения которой во многом зависит точность, достоверность полученных результатов, простота технической реализации, является выбор модели функциональной характеристики. Модель функциональной характеристики выбирается на основе априорной информации о свойствах процесса и может представлять собой:

- линейную комбинацию функций (аппроксимация функциями заданного вида);
- разложение в ряд по функциям определенного класса (аппроксимация степенными рядами, экспоненциально-степенная, рядами по дисперсиям производных, тригонометрическим многочленом (гармонический анализ), ортогональными полиномами и функциями, асимптотическими рядами).

Существующие современные автоматизированные системы математических расчетов не позволяют аппроксимировать двумерные плотности вероятности в различных ортогональных базисах. В связи с этим разработка и реализация алгоритмов аппроксимативного анализа в различных ортогональных базисах и нейросетевых моделях (в нашем случае анализируются алгоритмы аппроксимации ортогональными функциями Лежандра, Чебышева, Лагерра и Эрмита) представляется обоснованной.

**Во второй главе** рассмотрены методы аппроксимации двумерных плотностей вероятности на основе выборочных данных с помощью ортогональных функций и радиально-базисных сетей.

Для аппроксимации плотностей вероятности, определенных на прямоугольной области

$$G = \{(x, y): a < x < b, c < y < d\}, \tag{1}$$

рассмотрим ортонормированные функции по двум переменным

$$F_{nm}(x,y)=P_n(x)Q_m(y), \tag{2}$$

являющиеся произведениями одномерных ортогональных функций $P_n(x)$ и $Q_m(y)$, с весовой функцией с разделяющимися переменными

$$\mu(x,y)=\mu_x(x)\mu_y(y). \tag{3}$$

Аппроксимирующее выражение в таком случае выглядит следующим образом:

$$\tilde{f}(x,y)=\sum_{n=0}^{N}\sum_{m=0}^{M}\beta_{nm}F_{nm}(x,y). \tag{4}$$

Виды исследованных одномерных функций представлены в таблице 1.

Таблица 1 – Виды одномерных функций

| Одномерный базис | Аналитическое выражение |
|---|---|
| Полиномы Лежандра, $[-1;1]$ | $P_k(x)=\sum_{s=0}^{[k/2]}a_s\cdot x^{k-2s},\ a_s=(-1)^s\cdot\frac{\sqrt{2\cdot k+1}}{\sqrt{2}}\cdot\frac{(2k-2s)!}{2k\cdot s!\cdot(k-s)!\cdot(k-2s)!}.$ |
| Полиномы Чебышева I рода, $[-1;1]$ | $T_0(x)=\frac{1}{\sqrt{\pi}};\ T_k(x)=\sum_{s=0}^{[k/2]}a_s\cdot x^{k-2s},\ a_s=(-1)^s\cdot\sqrt{\frac{2}{\pi}}\cdot 2^{k-2s-1}\cdot\frac{k(k-s-1)!}{s!\cdot(k-2s)!}.$ |
| Полиномы Чебышева II рода, $[-1;1]$ | $U_0(x)=\sqrt{\frac{2}{\pi}};\ U_k(x)=\sum_{s=0}^{[k/2]}a_s\cdot x^{k-2s},\ a_s=(-1)^s\cdot\sqrt{\frac{2}{\pi}}\cdot 2^{k-2s}\cdot\frac{(k-s)(k-s-1)!}{s!\cdot(k-2s)!}.$ |
| Функции Лагерра, $[0;\infty)$ | $L_0(x)=e^{-x/2};\ L_k(x)=\sum_{s=0}^{k}a_s\cdot x^s\cdot e^{-x/2},\ a_s=(-1)^s\cdot\frac{k!}{(s!)^2\cdot(k-s)!}.$ |
| Функции Эрмита, $(-\infty;+\infty)$ | $H_k(x)=\sum_{s=0}^{[k/2]}a_s\cdot x^{k-2s}\cdot e^{-x^2/2},\ a_s=(-1)^s\sqrt{\frac{k!}{2^k\cdot\sqrt{\pi}}}\cdot\frac{2^{k-2s}}{s!\cdot(k-2s)!}.$ |

Коэффициенты перед функциями $F_{nm}(x,y)$ рассчитываются по формуле:

$$\beta_{nm}=\iint_G f(x,y)F_{nm}(x,y)\mu(x,y)dxdy, \tag{5}$$

где $f(x,y)$ – аппроксимируемая функция плотности вероятности.

Так как рассматриваемая задача подразумевает отсутствие исходного аналитического выражения для плотности вероятности, то данная функция представлена сплайн-моделью, построенной по двумерной гистограмме. То есть на практике мы имеем дело с билинейным или бикубическим сплайном либо с их производными, если строилась не частотная, а кумулятивная гистограмма:

$$s^{(p)}(x,y)=s_{ij}^{(p)}(x,y),\ x_i\le x<x_{i+1},\ y_j\le y<y_{j+1},\ s_{ij}^{(p)}(x,y)=\sum_{l=0}^{p}\sum_{m=0}^{p}a_{ij}^{(lm)}x^l y^m. \tag{6}$$

Рекомендациями по выбору количества столбцов в гистограмме являются модифицированные правила для одномерных гистограмм, расширенные для двумерных случаев. Самыми широко распространенными в литературе для одномерных распределений являются правила Стёрджеса, Скотта и Фридмана-Диакониса. Их двумерные модификации представлены уравнениями, решения которых определяются итеративно в целых числах и зависят от количества столбцов гистограммы $M_x$ и $M_y$ по осям.

Формула Стёрджеса:

$$M_x = \left[\log_2 \frac{N}{M_y} + 1\right] = \left[\log_2 N - \log_2\left[\log_2 \frac{N}{M_x} + 1\right] + 1\right],$$
$$M_y = \left[\log_2 \frac{N}{M_x} + 1\right] = \left[\log_2 N - \log_2\left[\log_2 \frac{N}{M_y} + 1\right] + 1\right]. \tag{7}$$

Правило Скотта:

$$M_x = \left[\frac{\sqrt[3]{N/M_y}\,(x_{\max} - x_{\min})}{3.5\sqrt{DX}}\right] = \left[\sqrt[3]{N\Big/\left[\frac{\sqrt[3]{N/M_x}\,(y_{\max} - y_{\min})}{3.5\sqrt{DY}}\right]}\,\frac{(x_{\max} - x_{\min})}{3.5\sqrt{DX}}\right],$$
$$M_y = \left[\frac{\sqrt[3]{N/M_x}\,(y_{\max} - y_{\min})}{3.5\sqrt{DY}}\right] = \left[\sqrt[3]{N\Big/\left[\frac{\sqrt[3]{N/M_y}\,(x_{\max} - x_{\min})}{3.5\sqrt{DX}}\right]}\,\frac{(y_{\max} - y_{\min})}{3.5\sqrt{DY}}\right]. \tag{8}$$

Правило Фридмана-Диакониса ( $IQR(X)$ – интерквартильный размах):

$$M_x = \left[\frac{\sqrt[3]{N/M_y}\,(x_{\max} - x_{\min})}{2 \cdot IQR(X)}\right] = \left[\sqrt[3]{N\Big/\left[\frac{\sqrt[3]{N/M_x}\,(y_{\max} - y_{\min})}{2 \cdot IQR(Y)}\right]}\,\frac{(x_{\max} - x_{\min})}{2 \cdot IQR(X)}\right],$$
$$M_y = \left[\frac{\sqrt[3]{N/M_x}\,(y_{\max} - y_{\min})}{2 \cdot IQR(Y)}\right] = \left[\sqrt[3]{N\Big/\left[\frac{\sqrt[3]{N/M_y}\,(x_{\max} - x_{\min})}{2 \cdot IQR(X)}\right]}\,\frac{(y_{\max} - y_{\min})}{2 \cdot IQR(Y)}\right]. \tag{9}$$

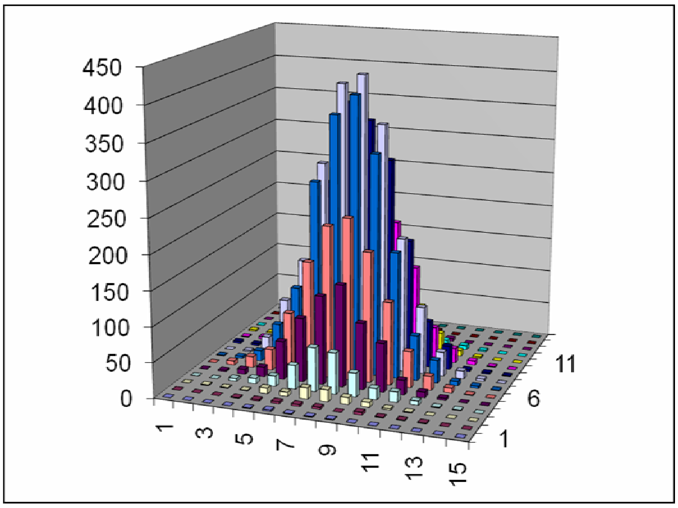


Рисунок 1

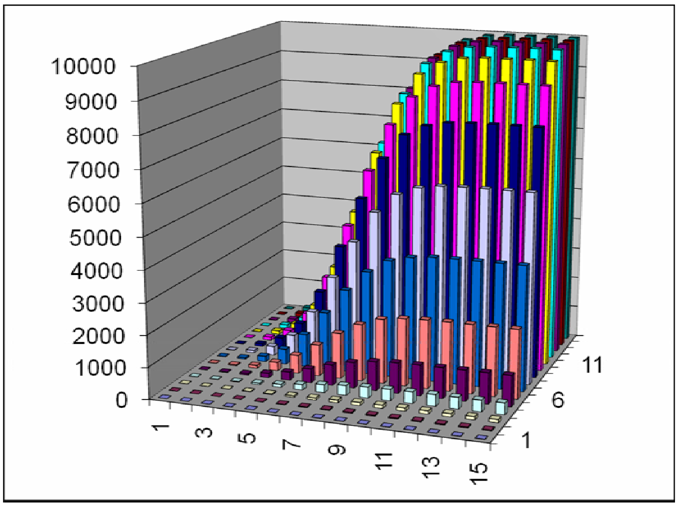


Рисунок 2

Для расчета билинейного и бикубического сплайнов нужен набор точек, через которые они проходят. Эти точки получаются построением двумерных гистограмм на основе выборочных данных (частотная на рис. 1 и кумулятивная на рис. 2).

В общем случае область $G$ определения семейства ортогональных функций и область $D$ определения плотности вероятности не совпадают. Поэтому необходимо ввести понятие семейств функций, ортонормированных на произвольной прямоугольной области. Это достигается путем использования функций вида (2) в сочетании с коэффициентами переноса $\beta_x$ и $\beta_y$ и масштаба $\alpha_x$ и $\alpha_y$:

$$\tilde{F}_{nm}(x,y) = \frac{1}{\sqrt{\alpha_x \alpha_y}} F_{nm}\left(\frac{x-\beta_x}{\alpha_x}, \frac{y-\beta_y}{\alpha_y}\right),\ \tilde{\mu}(x,y) = \mu\left(\frac{x-\beta_x}{\alpha_x}, \frac{y-\beta_y}{\alpha_y}\right). \tag{10}$$

Подставляя выражение (6) в формулу (5), мы получаем формулу расчета коэффициентов разложения неизвестной плотности через аппроксимацию сплайн-модели $s^{(p)}(x,y)$. Величина взвешенной погрешности определяется по формуле:

$$\Delta_{мет.}^{(N,M)} = \iint_D \left(s^{(p)}(x,y) - \hat{f}(x,y)\right)^2 \hat{\mu}(x,y)dxdy = \iint_D \left(s^{(p)}(x,y)\right)^2 \hat{\mu}(x,y)dxdy - \sum_{n=0}^{N}\sum_{m=0}^{M}\hat{\beta}_{nm}^2 . \tag{11}$$

Поскольку для функций Лагерра и Эрмита невозможно без дополнительных исследований определить оптимальное значение коэффициента масштаба, представим коэффициенты разложения в общем виде как функции от неизвестных коэффициентов масштаба $\alpha_x$ и $\alpha_y$:

$$\hat{\beta}_{nm} = \hat{\beta}_{nm}(\alpha_x, \alpha_y) = \frac{1}{\sqrt{\alpha_x \alpha_y}} \iint_D s^{(p)}(x,y) F_{nm}\left(\frac{x-\beta_x}{\alpha_x}, \frac{y-\beta_y}{\alpha_y}\right)\mu\left(\frac{x-\beta_x}{\alpha_x}, \frac{y-\beta_y}{\alpha_y}\right)dxdy . \tag{12}$$

Погрешность (11) восстановления исходной функции в большой степени зависит от выбора коэффициентов масштаба. Если хотя бы на одной из осей для аппроксимации выбираются ортогональные функции Лагерра или Эрмита, то возникает задача определения оптимального значения коэффициента масштаба (или сразу двух).

Остановимся на рассмотрении случая, когда по обеим осям выбраны функции Лагерра или Эрмита. Задача оптимизации коэффициентов масштаба решается через минимизацию функции (11). Так как функции Лагерра и Эрмита имеют единичную весовую функцию и левая часть погрешности (11) становится константой, требуется минимизировать выражение:

$$-\sum_{n=0}^{N}\sum_{m=0}^{M}\hat{\beta}_{nm}^2(\alpha_x, \alpha_y) \to \min . \tag{13}$$

Введем обозначение интеграла слагаемого сплайна для формулы (12):

$$I_x^{(kn)}(a,b) = \int_a^b x^k P_n(x)\mu_x(x)dx . \tag{14}$$

Подставив в (13) формулу для расчета коэффициентов, получим:

$$-\sum_{n=0}^{N}\sum_{m=0}^{M}\left(\sum_{i=0}^{m_x-1}\sum_{j=0}^{m_y-1}\sum_{k=0}^{p}\sum_{l=0}^{p} a_{D,ij}^{(kl)} I_x^{(kn)}\left(\frac{x_i-\beta_x}{\alpha_x}, \frac{x_{i+1}-\beta_x}{\alpha_x}\right) I_y^{(lm)}\left(\frac{y_j-\beta_y}{\alpha_y}, \frac{y_{j+1}-\beta_y}{\alpha_y}\right)\right)^2 \to \min . \tag{15}$$

Поскольку для функций Эрмита выражение (15) нельзя представить в аналитическом виде, в общем случае задача минимизации решается численно. Например, можно воспользоваться методом Ньютона как алгоритмом наискорейшего спуска, обеспечивающим более быструю сходимость на широком классе функций, чем остальные градиентные методы.

Полученное выражение плотности вероятности должно удовлетворять условию неотрицательной определенности и условию нормировки. Чтобы упростить решение этой задачи, нужно представить аппроксимирующее выражение в виде суммы мономов, а не суммы полиномов. В случае функций Лагерра и Эрмита это будет сумма произведений мономов на экспоненциальную функцию.

Представим функции из таблицы 1 для упрощения вычислений в виде:

$$P_k(x) = \sum_{s=0}^{k} b_s^{(k)} x^s \varphi_x(x). \tag{16}$$

где $\varphi_x(x)$ – некоторая функция-сомножитель. Значения новых коэффициентов $b_s^{(k)}$ и вид сомножителей приведены в таблице 2.

Таблица 2 – Значения новых коэффициентов $b_s^{(k)}$ и вид сомножителей

| Вид | Коэффициенты мономов | Сомножитель |
|---|---|---|
| Полиномы Лежандра, Чебышева | $b_s^{(k)} = \begin{cases} a_{(k-s)/2}, (k-s) - четное, \\ 0, иначе. \end{cases}$ | 1 |
| Функции Лагерра | $b_s^{(k)} = a_s$ | $e^{-x/2}$ |
| Функции Эрмита | $b_s^{(k)} = \begin{cases} a_{(k-s)/2}, (k-s) - четное, \\ 0, иначе. \end{cases}$ | $e^{-x^2/2}$ |

Подставляя выражение (16), получим:

$$\hat{f}(x,y) = \sum_{n=0}^{N}\sum_{m=0}^{M}\frac{\hat{\beta}_{nm}}{\sqrt{\alpha_x\alpha_y}}\left(\sum_{r=0}^{n} b_{r,x}^{(n)}\left(\frac{x-\beta_x}{\alpha_x}\right)^r \varphi_x\left(\frac{x-\beta_x}{\alpha_x}\right)\right)\left(\sum_{s=0}^{m} b_{s,y}^{(m)}\left(\frac{y-\beta_y}{\alpha_y}\right)^s \varphi_y\left(\frac{y-\beta_y}{\alpha_y}\right)\right). \qquad (17)$$

Раскроем внутренние скобки:

$$\hat{f}(x,y) = \sum_{n=0}^{N}\sum_{m=0}^{M}\frac{\hat{\beta}_{nm}}{\sqrt{\alpha_x\alpha_y}}\left(\sum_{r=0}^{n} c_{r,x}^{(n)} x^r \varphi_x\left(\frac{x-\beta_x}{\alpha_x}\right)\right)\left(\sum_{s=0}^{m} c_{s,y}^{(m)} y^s \varphi_y\left(\frac{y-\beta_y}{\alpha_y}\right)\right). \qquad (18)$$

где пересчет коэффициентов осуществляется по формуле:

$$c_{r,x}^{(n)} = \sum_{i=r}^{n} C_i^r \frac{b_{i,x}^{(n)}}{\alpha_x^i}(-\beta_x)^{i-r}\ . \qquad (19)$$

Перейдем к более краткому представлению вида

$$\hat{f}(x,y) = \sum_{n=0}^{N}\sum_{m=0}^{M}\gamma_{nm} x^n y^m \varphi_x\left(\frac{x-\beta_x}{\alpha_x}\right)\varphi_y\left(\frac{y-\beta_y}{\alpha_y}\right), \qquad (20)$$

где коэффициенты $\gamma_{nm}$ вычисляются по формуле:

$$\gamma_{nm} = \sum_{i=n}^{N}\sum_{j=m}^{M}\frac{\hat{\beta}_{ij}}{\sqrt{\alpha_x\alpha_y}} c_{n,x}^{(i)} c_{m,y}^{(j)}\ . \qquad (21)$$

Чтобы компенсировать области в $D$, на которых выражение (21) становится отрицательным, необходимо определить все экстремумы функции (21), в которых она принимает отрицательные значения. Определение точек, подозрительных на экстремум, сводится к решению следующей системы уравнений:

$$\begin{cases} g(x,y) = \sum_{n=0}^{N'}\sum_{m=0}^{M'}\gamma_{nm}'^{(x)} x^n y^m = 0, \\ h(x,y) = \sum_{n=0}^{N'}\sum_{m=0}^{M'}\gamma_{nm}'^{(y)} x^n y^m = 0. \end{cases} \qquad (22)$$

Системы полиномиальных уравнений в общем случае решаются построением базиса Грёбнера из базиса, составленного многочленами рассматриваемой системы, и решением системы уравнений, левые части которых являются многочленами, составляющими базис Грёбнера.

Рассмотрим новый ускоренный алгоритм нахождения базиса Грёбнера, усовершенствованный по отношению к алгоритму Бухбергера с целью минимизации вычислительной сложности. Блок-схема алгоритма представлена на рисунке 3. Итак, пусть

$K$ – поле вещественных чисел, $K[x,y]$ – кольцо многочленов от двух переменных с коэффициентами из $K$. Рассмотрим некоторый идеал $I$ над кольцом многочленов, в который входят полиномы из системы (22). Множество $H=\{g,h\}$ является базисом идеала $I$. Зададим лексикографическое упорядочение мономов в многочленах:

$$x^{p_x}y^{p_y} > x^{q_x}y^{q_y} \leftrightarrow (p_x > q_x)\cup(p_x = q_x)\cap(p_y > q_y). \tag{23}$$

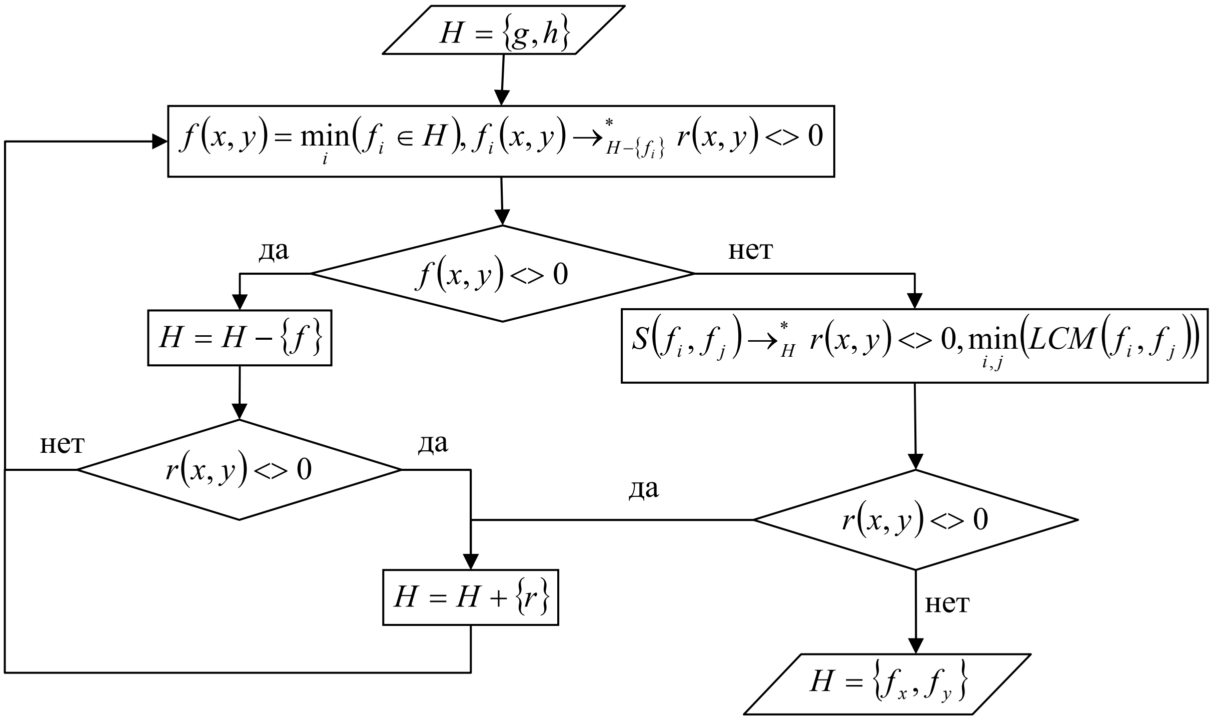


Рисунок 3

Используя операции сложения и умножения над многочленами, входящими в базис $H$, требуется редуцировать его так, чтобы построить базис Грёбнера идеала $I$.

Введем понятие S-полинома:

$$S(f_i,f_j)=\frac{LCM(f_i,f_j)}{LT(f_i)}f_i-\frac{LCM(f_i,f_j)}{LT(f_j)}f_j. \tag{24}$$

$LT(f_i)$ представляет старший терм многочлена $f_i(x,y)$, а $LCM(f_i,f_j)$ равен наименьшему моному, который делится на старшие мономы многочленов $f_i(x,y)$ и $f_j(x,y)$. Операцией редуцирования произвольного полинома $f(x,y)$ из $K[x,y]$ относительно $H$ является нахождение такого минимального многочлена $r(x,y)$, который выражается в виде:

$$f(x,y)\rightarrow^*_H r(x,y)=f(x,y)-\sum_i a_i(x,y)f_i(x,y), \tag{25}$$

где $a_i(x,y)\in K[x,y]$ и $f_i(x,y)\in H$.

Базис $H$ можно назвать базисом Грёбнера, если для любых двух полиномов $f_i(x,y)$ и $f_j(x,y)$ из $H$ выполняется условие:

$$S(f_i,f_j)\rightarrow^*_H 0. \tag{26}$$

При заданном лексикографическом упорядочении (23) базис Грёбнера будет составлен парой многочленов, удовлетворяющих условию (26), вида:

$$\begin{cases} f_x(x,y) = \sum\limits_{n=1}^{M_{xx}} k_n^{(x)} x^n + \sum\limits_{m=0}^{M_{xy}} k_m^{(xy)} y^m, \\ f_y(x,y) = \sum\limits_{m=0}^{M_{yy}} k_m^{(y)} y^m. \end{cases} \tag{27}$$

Множество решений системы, составленной из полиномов (27), будет соответствовать множеству решений системы (22). Полиномиальные уравнения одной переменной решаются методом Дюрана-Кернера.

Для каждого рассчитанного центра локальных минимумов $\{\bar{x}_i, \bar{y}_i\}$ требуется определить функцию, которая соответствовала бы следующим условиям:

$$\begin{aligned} &\breve{f}_i(x,y) \le \tilde{f}(x,y), \forall x, y \in D_i, \\ &\breve{f}_i(x,y) \to -0, \forall x, y \notin D_i, \end{aligned} \tag{28}$$

где $D_i$ задает некоторую область отрицательности в окрестностях локального минимума.

Скомпенсированное аппроксимирующее выражение примет вид:

$$\tilde{f}_C(x,y) = \sum_{n=0}^{N} \sum_{m=0}^{M} \gamma_{nm} x^n y^m \varphi_x\left(\frac{x-\beta_x}{\alpha_x}\right) \varphi_y\left(\frac{y-\beta_y}{\alpha_y}\right) - \sum_{i=1}^{\breve{N}} \breve{f}_i(x,y). \tag{29}$$

Отрицательную область $D_i$ в окрестности точки $\{\bar{x}_i, \bar{y}_i\}$ локализуем с помощью уравнения поверхности, сечениями которой являются эллипсы:

$$\left(a_{i,\max} x + b_{i,\max} y + c_{i,\max}\right)^2 + \left(a_{i,\min} x + b_{i,\min} y + c_{i,\min}\right)^2 = 1. \tag{30}$$

Ни рисунке 4 показана итерация определения неизвестных параметров большой и малой полуосей $r_{i,\max}$ и $r_{i,\min}$ для расчета коэффициентов эллипса (30).

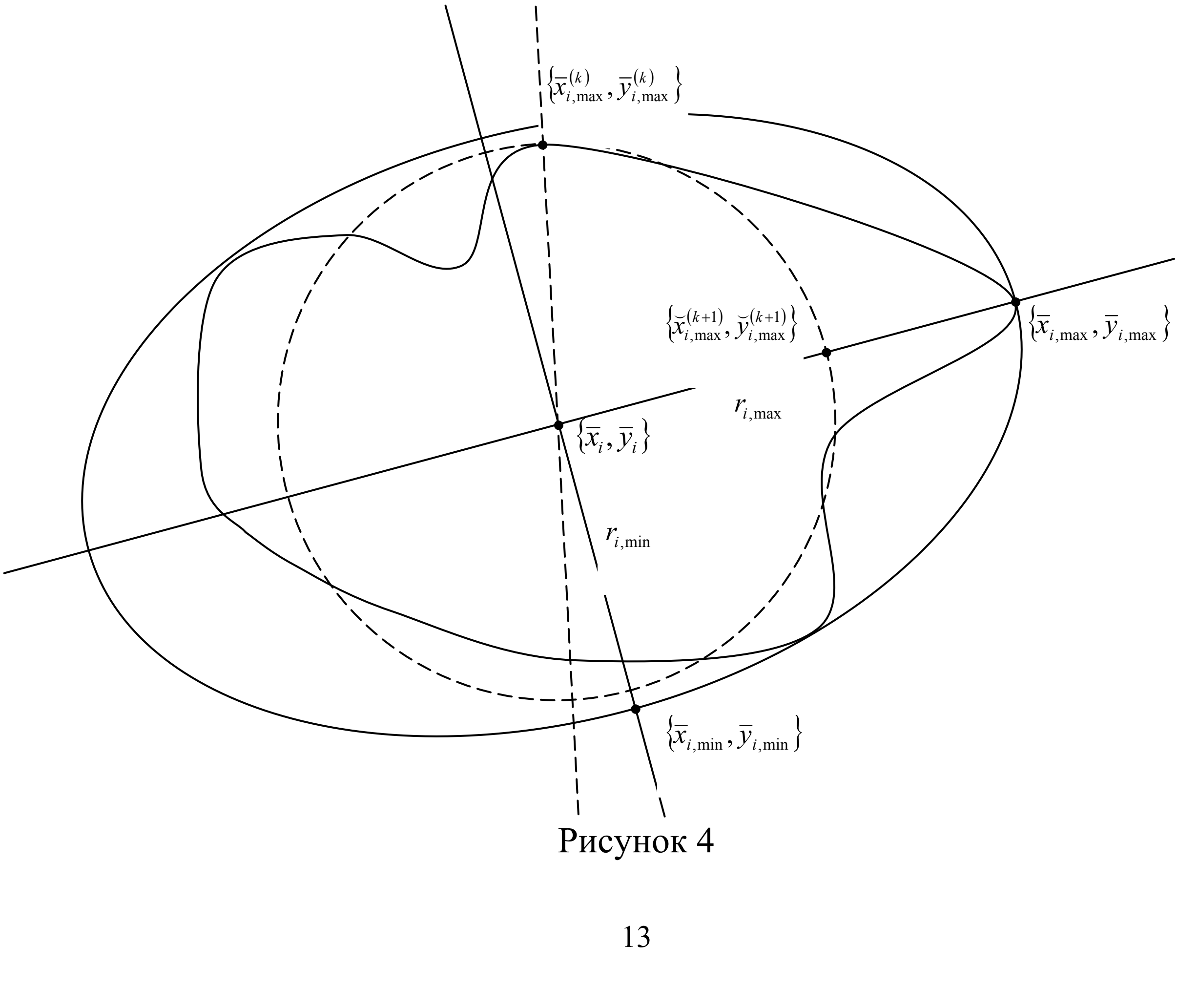


Рисунок 4

В качестве компенсирующей на области $D_i$ используется функция Гаусса:

$$\breve{f}_i(x,y)=\breve{k}_i e^{-\left(\left(a_{i,\max}x+b_{i,\max}y+c_{i,\max}\right)^2+\left(a_{i,\min}x+b_{i,\min}y+c_{i,\min}\right)^2\right)/\sigma_i^2}. \qquad (31)$$

Ее параметры рассчитываются из соображений минимизации объема:

$$\breve{V}_i=-\breve{k}_i\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty}e^{-\left(u^2/r_{i,\max}^2+v^2/r_{i,\min}^2\right)/\sigma_i^2}dudv=\breve{k}_i\int\limits_{-\infty}^{+\infty}e^{-u^2/r_{i,\max}^2\sigma_i^2}du\int\limits_{-\infty}^{+\infty}e^{-v^2/r_{i,\min}^2\sigma_i^2}dv=\pi\cdot r_{i,\max}r_{i,\min}\breve{k}_i\sigma_i^2. \qquad (32)$$

Альтернативой аппроксимации ортогональными функциями является использование нейронных сетей. Для восстановления аналитического выражения функции плотности вероятности из набора «узловых» точек используется алгоритм нейросетевой аппроксимации. В качестве базовой модели берется *RBF*-сеть, нейроны скрытого слоя которой являются двумерными функциями Гаусса вида:

$$G(x,y)=e^{-(v_x x-c_x)^2/2-(v_y y-c_y)^2/2}. \qquad (33)$$

Эти функции используются для построения аппроксимирующей модели, которая выглядит следующим образом:

$$\tilde{f}_{Ne}(x,y)=w_0+\sum_{k=1}^{K}w_k\cdot G_k(x,y)=w_0+\sum_{k=1}^{K}w_k\cdot e^{-(v_{x,k}x-c_{x,k})^2/2-(v_{y,k}y-c_{y,k})^2/2}. \qquad (34)$$

Неизвестные коэффициенты $w_k$, $v_{x,k}$, $c_{x,k}$, $v_{y,k}$ и $c_{y,k}$ в выражении (34) являются настраиваемыми, а их значения определяются в процессе обучения нейронной сети.

Для определения коэффициентов сети используется градиентный алгоритм наискорейшего спуска, в котором для расчета вектора градиента применяется метод обратного распространения ошибки. На вход сети подаются координаты «узловых» точек $\{x_i, y_j\}$, а на выходе полученный результат сравнивается с ожидаемым значением в «узловой» точке. Корректировка весов производится по алгоритму обратного распространения ошибки. Основу алгоритма обучения составляет целевая функция ошибки, определенная для каждой обучающей пары (под парой понимается вектор входных значений координатной сетки и значение функции в точке):

$$E_{ij}=\frac{1}{2}\left(\sum_{k=0}^{K}w_k\cdot G_k(x_i,y_j)-f(x_i,y_j)\right)^2. \qquad (35)$$

где $G_0(x,y)=1$.

Целью алгоритма обучения является минимизация значений целевой функции для каждой пары из обучающей выборки.

Если пороговые функции $G_k(x,y)$ упрощенно представить в виде

$$\varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right)=e^{-u_{ij}^{(k)}/2},u_{ij}^{(k)}=\left(v_{x,k}x_i-c_{x,k}\right)^2+\left(v_{y,k}y_j-c_{y,k}\right)^2, \qquad (36)$$

то величина ошибки будет рассчитываться по следующей формуле:

$$\Delta_{ij}=\sum_{k=0}^{K}w_k\cdot\varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right)-f(x_i,y_j). \qquad (37)$$

Так как метод является градиентным, для корректировки коэффициентов сети необходимо рассчитать вектор градиента. Конкретные компоненты градиента рассчитываются дифференцированием зависимости (35).

В первую очередь подбираются веса нейронов выходного слоя:

$$w_k=w_k-\eta\cdot\frac{\partial E_{ij}}{\partial w_k}=w_k-\eta\cdot\Delta_{ij}\cdot\frac{\partial\Delta_{ij}}{\partial w_k}=w_k-\eta\cdot\Delta_{ij}\cdot\varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right). \qquad (38)$$

Здесь $\eta$ – обучающий коэффициент, значение которого изменяется в пределах $(0{,}1)$ в зависимости от задачи и определяется эмпирически. Компоненты градиента относительно неизвестных коэффициентов нейронов скрытого слоя определяются по тому же принципу:

$$c_k = c_k - \eta \cdot \frac{\partial E_{ij}}{\partial c_k} = c_k - \eta \cdot \Delta_{ij} \cdot \frac{\partial \Delta_{ij}}{\partial c_k} = w_k + \eta \cdot \Delta_{ij} \cdot w_k \cdot \varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right) \cdot \frac{\partial u_{ij}^{(k)}}{\partial c_k}. \tag{39}$$

Таким образом, оставшиеся коэффициенты рассчитываются по следующим формулам:

$$\begin{cases} c_{x,k} = c_{x,k} - \eta \cdot \Delta_{ij} \cdot \varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right) \cdot \left(v_{x,k} x_i - c_{x,k}\right), \\ c_{y,k} = c_{y,k} - \eta \cdot \Delta_{ij} \cdot \varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right) \cdot \left(v_{y,k} y_j - c_{y,k}\right), \\ v_{x,k} = v_{x,k} + \eta \cdot \Delta_{ij} \cdot \varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right) \cdot \left(v_{x,k} x_i - c_{x,k}\right) \cdot x_i, \\ v_{y,k} = v_{y,k} + \eta \cdot \Delta_{ij} \cdot \varphi_{ij}^{(k)}\left(u_{ij}^{(k)}\right) \cdot \left(v_{y,k} y_j - c_{y,k}\right) \cdot y_j. \end{cases} \tag{40}$$

Все обучающие пары применяются по очереди в цикле, пока не закончилось заданное количество итераций либо величина ошибки не стала меньше определенной величины.

**В третьей главе** приведено описание разработанного комплекса программных средств, предназначенного для аппроксимативного анализа двумерных плотностей вероятности.

Система осуществляет аппроксимативный анализ, который позволяет производить аппроксимацию двумерных плотностей вероятности ортогональными функциями, являющимися произведениями функций Лежандра, Чебышева, Лагерра и Эрмита, а также с помощью радиально-базисных сетей. После аппроксимации система проверяет качество аппроксимации с помощью критерия Пирсона:

$$\chi_{xy}^2 = N \sum_{i=1}^{M_x} \frac{\left( \sum_{j=1}^{M_y} \hat{p}_{ij} - \int\limits_{x_i - \Delta x/2}^{x_i + \Delta x/2} \int\limits_{y_{\min}}^{y_{\max}} \hat{f}(x,y)\,dxdy \right)^2}{\int\limits_{x_i - \Delta x/2}^{x_i + \Delta x/2} \int\limits_{y_{\min}}^{y_{\max}} \hat{f}(x,y)\,dxdy}, \quad \chi_{yx}^2 = N \sum_{j=1}^{M_y} \frac{\left( \sum_{i=1}^{M_x} \hat{p}_{ij} - \int\limits_{x_{\min}}^{x_{\max}} \int\limits_{y_j - \Delta y/2}^{y_j + \Delta y/2} \hat{f}(x,y)\,dxdy \right)^2}{\int\limits_{x_{\min}}^{x_{\max}} \int\limits_{y_j - \Delta y/2}^{y_j + \Delta y/2} \hat{f}(x,y)\,dxdy}. \tag{41}$$

В данной главе рассмотрены аспекты аппроксимации в различных базисах, проведены исследования методом имитационного моделирования и выданы рекомендации по выбору параметров аппроксимативной модели и методике аппроксимации тех или иных функций.

Данный программный комплекс используется при обработке результатов натурного эксперимента по оценке погрешности базирования соединения типа «ласточкин хвост» лопаток компрессора ГТД при установке в наладке ПОМКЛ-БЛИК с использованием координатно-измерительной машины.

КИМ DEA Global Performance 07.10.07 используется на кафедре производства двигателей летательных аппаратов СГАУ для измерения параметров различных деталей авиационных двигателей, в частности лопаток компрессора. С ее помощью для каждой лопатки проводится серия измерений, которая дает набор результатов, являющихся координатами измерения определенных точек для нескольких сечений лопатки (рисунок 5).

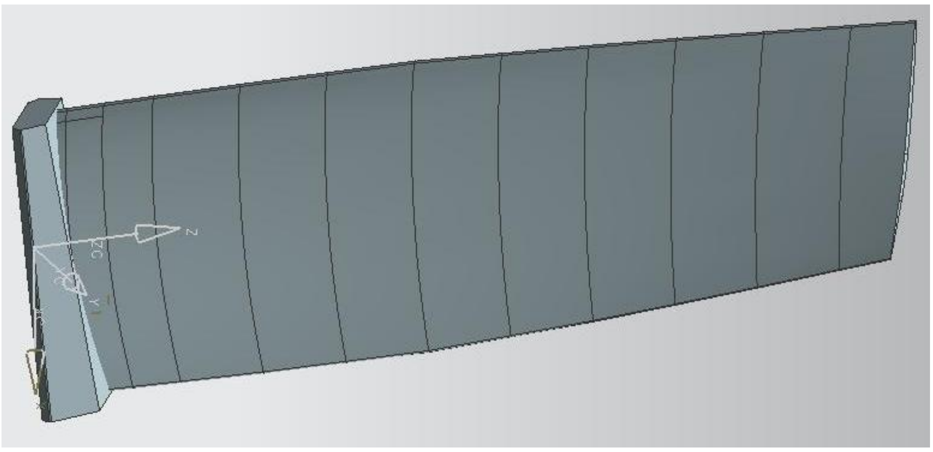

Рисунок 5

С целью оценки качества измерений была поставлена задача определения характеристик набора полученных значений. С помощью программного комплекса аппроксимативного анализа строится аналитическая модель для плотности вероятности, характеризующей распределение отклонений измеренных координат точек лопаток от их эталонных значений.

Измерения с помощью КИМ проводятся с базированием лопатки в приспособлении, аналогичном используемому на приборе ПОМКЛ-БЛИК, что означает наследование части погрешностей базирования из прибора ПОМКЛ, который используется на ОАО «Моторостроитель».

Серии измерений, проведенных с помощью КИМ, позволяют нам оперировать двумя наборами данных – набором отклонений измеренных машиной координат точек лопаток и набором отклонений измеренных координат точек контрольного эталона.

Каждая серия измерений представляет собой многократный повтор снятия координат в заранее определенных точках. Отклонения полученных результатов от требуемых и есть тот набор данных, которым мы оперируем при оценивании погрешности измерений.

Погрешность, полученная при измерении эталона, является суммой двух составляющих – погрешности базирования наладки ПОМКЛ в системе координат КИМ и погрешностью самой измерительной машины, которая допускает ошибки при снятии координат точек деталей. Погрешность же измерения точек на спинке и корытце лопаток ГТД является более сложной, поскольку к уже имеющимся составляющим погрешности добавляется погрешность изготовления поверхности самой лопатки. Если обозначить указанные погрешности как случайные величины $A$ и $B$, то через них можно оценить погрешность $Z = B - A$ изготовления лопаток ГТД.

Оперируя данными, полученными в ходе экспериментов, с помощью системы аппроксимации двумерных плотностей вероятности находим выражения плотностей $f_A(x,y)$ и $f_B(x,y)$ для двумерных случайных величин $A$ и $B$. Теперь величина погрешности $Z$ является суммой погрешностей $-A$ и $B$.

Для упрощения аналитического выражения плотности вероятности компенсирующие функции, если они присутствуют, нужно разложить в ряд аналогичной длины в том же базисе, что был выбран для восстановления неизвестной плотности. Учитывая, что в случае аппроксимации плотностей вероятности распределения погрешностей используются функции Эрмита, искомая плотность вероятности в таком случае определяется через известные выражения плотностей $f_{-A}(x,y)$ и $f_B(x,y)$. Процесс вычисления выражения сводится к вычислению интегралов в следующей формуле:

$$f_Z(x,y)=\sum_{n=0}^{N_A}\sum_{m=0}^{M_A}\sum_{k=0}^{N_B}\sum_{l=0}^{M_B}\sum_{i=0}^{k}\sum_{j=0}^{l}(-1)^{n+m+k-i+l-j}C_k^iC_l^j\gamma_{nm,A}\gamma_{kl,B}x^iy^j\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x_A^{n+k-i}y_A^{m+l-j}\cdot$$
$$e^{-\left(x_A-\beta_{x,A}/\alpha_{x,A}\right)^2/2-\left(y_A-\beta_{y,A}/\alpha_{y,A}\right)^2/2-\left(x-x_A-\beta_{x,B}/\alpha_{x,B}\right)^2/2-\left(y-y_A-\beta_{y,B}/\alpha_{y,B}\right)^2/2}dx_Ady_A. \qquad (42)$$

После расчета интегралов, раскрытия скобок и приведения слагаемых выражение (42) приводится к виду (20). Используя данное выражение, можно определить различные характеристики: точку с максимальной вероятностью, математическое ожидание, границы области, в которой измеренная характеристика окажется с заданной вероятностью, вероятность выхода за пределы зоны допустимой погрешности и т.д.

Подобным способом решается задача компаундирования бензинов. По составу автомобильные бензины представляют собой смесь компонентов, получаемых в результате различных технологических процессов: прямой перегонки нефти, каталитического риформинга, каталитического крекинга и гидрокрекинга вакуумного газойля, изомеризации прямогонных фракций, алкилирования, ароматизации термического крекинга, висбрекинга, замедленного коксования. Важнейшими показателями, характеризующими качество марочного бензина, являются октановое число $\Omega$ и плотность бензина $\rho$. Для того чтобы бензин имел соответствующий уровень качества, его октановое число и плотность должны находиться в заданных пределах. Поскольку же бензин является смесью нескольких компонентов, то его характеристики напрямую зависят от характеристик компонентов.

Используя статистику по характеристикам продуктов переработки нефти, можно построить оценки двумерных плотностей распределения для октанового числа и плотности, как связанных характеристик, любого из компонентов. Оперируя данными выражениями, можно рассчитать совместную плотность распределения двух и более компонент моторных топлив, составленных в заданных пропорциях. Если два компонента рассматривать как некоторые случайные величины $A$ и $B$, то их смесь в пропорциях $\eta:1-\eta$ также является случайной величиной $Z=\eta A+(1-\eta)B$. Рассчитав плотности вероятностей для величин $\eta A$ и $(1-\eta)B$, можно перейти к плотности вероятности случайной величины $Z$ аналогично вышеизложенной задаче с лопатками ГТД. Так характеристики смеси можно оценить заранее без физического смешения компонентов топлив, а анализ значения коэффициента $\eta$ позволит рассчитать оптимальное соотношение компонентов в смеси. Данная методика внедрена в ООО «НТФ «Протон».

В рамках разработанного комплекса также есть возможность решать другие задачи. Например, использование радиально-базисных сетей для получения аппроксимативного выражения дает результат в виде суммы двумерных функций Гаусса. Если объектом обработки является многомодальное распределение, то вид полученного выражения позволяет выделить отдельные составляющие из общего выражения (разбить выражение на несколько сумм), решая задачу декомпозиции.

Таким образом, в третьей главе рассмотрены задачи, возможность для решения которых предоставляет предложенная методика аппроксимации и разработанный на ее основе программный комплекс.

## ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

В диссертации решена актуальная задача по разработке алгоритмов и комплекса программ для восстановления аналитических выражений двумерных плотностей вероятности по выборочным данным методом аппроксимативного анализа с использованием ортогональных базисов и радиально-базисной сети. Комплекс является системой, открытой для добавления новых алгоритмов и для последующего развития.

Основные результаты работы состоят в следующем:

1. Исследованы имеющиеся материалы в области аппроксимации двумерных функций и восстановления двумерных плотностей вероятности.
2. Разработаны двумерные модификации правил Стёрджеса, Скотта и Фридмана-Диакониса и предложен численный метод расчета оптимального числа коридоров гистограммы в соответствии с данными правилами.
3. Разработана методика и алгоритмы аппроксимации двумерных плотностей вероятности на прямоугольных областях ортогональными функциями, являющимися произведениями функций Лежандра, Чебышева, Лагерра и Эрмита, с использованием первичного приближения сплайн-моделью.
4. Разработан ускоренный алгоритм нахождения базисов Грёбнера при решении систем из двух полиномиальных уравнений с двумя неизвестными.
5. Разработана методика компенсации отрицательно определенных областей двумерными функциями Гаусса.
6. Предложена методика аппроксимации двумерных плотностей вероятности радиально-базисными сетями.
7. Разработана структура, программное и методическое обеспечение программного комплекса, реализованного на языке Object Pascal в среде визуального проектирования Delphi.
8. Произведен анализ результатов аппроксимации различными ортогональными функциями и нейросетевыми моделями.
9. Получены рекомендации по использованию и настройке аппроксимативных моделей в зависимости от вида аппроксимируемой характеристики.
10. Разработанные методы, алгоритмы и комплекс программ внедрены в учебном процессе СГАУ, что подтверждается соответствующим актом о внедрении, а также использованы при обработке данных натурного эксперимента.

## ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

**Монография**

1. Прикладной анализ случайных процессов [Текст] / Под ред. Прохорова С.А / Лёзин И.А., Лёзина И.В. / СНЦ РАН, Самара, 2007. – 5.10 Автоматизированная система аппроксимативного анализа законов распределения ортогональными полиномами и нейросетевыми функциями. – С.391-406. – ISBN 978-5-93424-283-2.

**Статьи в изданиях, определенных ВАК России**

2. Прохоров, С. А. Аппроксимация законов распределения ортогональными полиномами [Текст] / Прохоров С. А., Лёзин И.А., Солдатова И.В. // Вестник Самарского государственного технического университета. Серия Физико-математические науки №34 – 2005, с.128-136. – Библиогр.: с.136.

3. Прохоров, С. А. Определение функциональных характеристик случайных процессов методами аппроксимации и нейросетевого анализа и их сравнение [Текст] / Прохоров С. А., Лёзин И.А., Лёзина И.В. // Вестник Самарского государственного технического университета. Серия Технические науки №33 – 2005, с.340-346. – Библиогр.: с. 346.

4. Лёзин, И.А. Разложение двумерных плотностей вероятности в ортогональных базисах [Текст] / Лёзин И.А. // Вестник Самарского государственного технического университета. Серия физико-математические науки №1 (18) – 2009, с. 169-174. – Библиогр.: с. 174.

**Статьи в других изданиях**

5. Прохоров, С. А. Автоматизированная система аппроксимативного анализа законов распределения случайных процессов [Текст] / Прохоров С. А., Солдатова И.В., Лёзин И.А. // Надежность и качество. Труды международного симпозиума. Под ред. Н.К. Юркова. – Пенза: Изд-во Пенз. гос. ун-та, 2004., с.57-63. – Библиогр.: с.63.

6. Прохоров, С. А. Определение функциональных характеристик случайных процессов методами аппроксимации и нейросетевого анализа и их сравнение [Текст] / Прохоров С. А., Лёзин И.А., Лёзина И.В. // Информационные, измерительные и управляющие системы (ИИУС-2005). Материалы Международной научно-технической конференции, Самара, изд-во Самарского государственного технического университета, 2005, с.270-272. – Библиогр.: с.272.

7. Солдатова, О.П. Применение нейросетевых моделей для аппроксимации временных рядов полиномами [Текст] / Солдатова О.П., Лёзин И.А., Васильева Ю.В. // Радиотехника и связь. Материалы международной научно-технической конференции, посвященной 110-летию изобретения радио и 75-летию Саратовского государственного технического университета. Саратов, 2005, с. 12-16. – Библиогр.: с.16.

8. Прохоров, С. А. Сравнение качества восстановления функций различными нейросетевыми моделями [Текст] / Прохоров С. А., Лёзин И.А., Лёзина И.В. // Компьютерные технологии в науке, практике и образовании. Труды Всероссийской межвузовской научно-практической конференции. Самара, изд-во Самарского государственного технического университета, 2005, с.59-62. – Библиогр.: с.62.

9. Лёзин, И.А. Декомпозиция сигналов с использованием нейросетевых моделей [Текст] / Лёзин И.А. // Радиотехника и связь. Материалы третьей международной научно-технической конференции. Саратов, 2006, с. 30-34. – Библиогр.: с.34.

10. Лёзин, И.А. Автоматизированный комплекс аппроксимации функций ортогональными полиномами и нейронными сетями [Текст] / Лёзин И.А., Лёзина И.В., Прохоров С. А. // Перспективные информационные технологии в научных исследованиях, проектировании и обучении (ПИТ-2006). Труды научно-технической конференции с международным участием. Том 1. Самара, 2006, с. 106-112. – Библиогр.: с.112.

11. Прохоров, С. А. Аппроксимация плотности вероятности случайных процессов ядерными функциями и ортогональными полиномами [Текст] / Прохоров С. А., Лёзин И.А., Лёзина И.В., Соболева А.Е. // Инновационные технологии в управлении, образовании, промышленности «АСТИНТЕХ-2007»: материалы Всероссийской научной конференции 18-20 апреля 2007 г.: в 2 ч./ сост. И.Ю.Петрова. – Астрахань: Издательский дом «Астраханский университет», 2007. Ч.2, с. 136-139. – Библиогр.: с.139.

12. Прохоров, С. А. Автоматизированная информационная система аппроксимации двумерных плотностей вероятностей нейронными сетями [Текст] / Прохоров С. А., Лёзин И.А., Лёзина И.В. // Проблемы автоматизации и управления в технических системах: труды Международной научно-технической конференции под ред. д.т.н., проф. М.А.Щербакова. – Пенза, Информационно-издательский центр ПГУ, 2007, с.143-146. – Библиогр.: с.146.

13. Лёзин И.А. Автоматизированный комплекс аппроксимативного анализа двумерных законов распределения ортогональными полиномами и нейронными сетями [Текст] / Лёзин И.А. // Информационные технологии в высшем профессиональном образовании: Сборник докладов II межрегиональной научно-практической конференции / Под.ред. О.А. Тарабрина, А.В. Очеповского – Тольятти-Самара: Самарский государственный аэрокосмический университет, 2007, с.84-87.

14. Прохоров, С. А. Аппроксимация двумерной плотности вероятности ортогональными полиномами [Текст] / Прохоров С. А., Лёзин И.А., Лёзина И.В. // Радиотехника и связь. Материалы четвертой международной научно-технической конференции. Саратовский государственный технический университет – Саратов, 2007, с. 17-22. – Библиогр.: с.22.

15. Лёзин И.А. Подбор оптимальных значений коэффициентов масштабирования при аппроксимации двухмерных функций [Текст] / Радиотехника и связь: сборник научных трудов. – Саратов: СГТУ, 2008. С. 8-13.

16. Лёзин И.А. Нахождение минимумов двумерных функций при компенсации областей отрицательности плотности вероятностей [Текст] / Лёзин И.А., Лёзина И.В. // Надежность и качество: труды Международного симпозиума: в 2-х т. / под ред. Н.К. Юркова. – Пенза: Информационно-издательский центр ПензГУ, 2009. –1 т., с. 336-338. – Библиогр.: с.338. – ISBN 978-5-94170-218-3.

**Тезисы докладов**

17. Лёзин, И.А. Декомпозиция сигналов с использованием нейросетевых моделей [Текст] / Лёзин, И.А., под рук. Прохорова С. А. // Всероссийская молодежная научная конференция с международным участием «VIII Королевские чтения». Тезисы докладов. Изд-во Самарского государственного аэрокосмического университета имени академика С.П. Королева, Самара, 2005, с. 316.

18. Болотов, М.А. Исследование погрешностей базирования в наладке ПОМКЛ-БЛИК в системе аппроксимативного анализа двумерных плотностей вероятности [Текст] / Болотов М.А., Жидяев А.Н., Лёзин И.А., Сурков О.С., Шитарев И.Л. // Международная научно-техническая конференция «Проблемы и перспективы развития двигателестроения». Материалы докладов. Изд-во Самарского государственного аэрокосмического университета имени академика С.П. Королева, Самара, 2009. – В 2 ч. Ч.1, с. 173-174.