

Федеральное государственное автономное образовательное учреждение
высшего образования «Сибирский федеральный университет»

На правах рукописи

Михов Евгений Дмитриевич

**ИДЕНТИФИКАЦИЯ И УПРАВЛЕНИЕ ПРОЦЕССАМИ СО
СТОХАСТИЧЕСКИ-ЗАВИСИМЫМИ ПЕРЕМЕННЫМИ
МЕТОДАМИ НЕПАРАМЕТРИЧЕСКОЙ СТАТИСТИКИ**

Специальность 05.13.01 – Системный анализ, управление и обработка
информации (космические и информационные технологии)

Диссертация на соискание ученой степени

кандидата технических наук

Научный руководитель:

д.т.н., профессор А.В.Медведев

Красноярск 2019

Содержание

Введение	5
1 Общие сведения об идентификации и управлении в условиях неполной информации	12
1.1 Характеристика объектов дискретно – непрерывного типа	13
1.2 Априорная информация, уровни априорной информации	15
1.3 Идентификация в «узком» и «широком» смыслах	16
1.4 Особые случаи моделируемых процессов	17
1.4.1 Н-процессы	18
1.4.2 Процессы протекающие в пространстве дробной размерности	22
1.4.3 Процессы протекающие в пространстве с изменяющейся размерностью	24
1.5 Выводы по главе	35
2 Модели дискретно непрерывных процессов	37
2.1 Постановка задачи	37
2.2 Выделение существенных переменных	39
2.3 Нахождение параметрической модели методом перебора	55
2.4 Нахождение параметрической модели при помощи генетического программирования	57
2.5 Модели основанные на локальной аппроксимации	62
2.5.1 Настройка вектора параметров размытости	65
2.5.2 Улучшение качества обучающей выборки	69
2.6 Н-модели	74
2.6.1 Общая информация об Н-моделях	74

2.6.2	Модель многомерного процесса с общей зоной определения входных переменных	82
2.6.3	Модель многомерного процесса без общей зоны определения входных переменных	88
2.7	Модели основанные на технологии нейросетей	90
2.7.1	Искусственный нейрон	91
2.7.2	Функции активации	92
2.7.3	Метод обратного распространения ошибки	93
2.8	Сравнение алгоритмов моделирования	99
2.9	Выводы по главе	106
3	Управление безынерционными стохастическими процессами при неполной информации	108
3.1	Постановка задачи	108
3.2	Стандартные алгоритмы управления	110
3.3	Дуальное управление	111
3.4	Непараметрическое дуальное управление	112
3.4.1	Общая информация о непараметрическом дуальном управлении	112
3.4.2	Управление процессом, имеющим одну выходную переменную	113
3.4.3	Управление многомерным N-процессом	115
3.5	Выводы по главе	120
4	Применение методов непараметрической статистики на практических задачах	122
4.1	Постановка задачи экологического мониторинга	122

4.2	Алгоритм непараметрического моделирования экологической обстановки	123
4.3	Результаты моделирования экологической обстановки	126
4.4	Результаты применения разработанных методов для управления котлоагрегатами ТЭЦ	136
4.5	Выводы по главе	136
	Заключение	138
	Приложение А	149

Введение

Актуальность работы. В диссертации рассматриваются многомерные безынерционные процессы с запаздыванием со стохастически зависимыми компонентами вектора входных переменных в условиях непараметрической неопределенности, то есть в условиях, когда представление о параметрической модели по различным каналам объекта с точностью до параметров отсутствует. Актуальность исследования подобных процессов обусловлена тем, что постоянная времени этих объектов (по-существу динамических) значительно меньше периодичности контроля. Это означает, что время протекания переходного процесса может быть значительно меньше, чем периодичность измерения некоторых выходных переменных. Например, переходный процесс завершается через 15-20 минут, а измерение выходной переменной осуществляется через 2 или более часов. Анализ подобных процессов показывает, что они протекают в пространствах дробной размерности. Все это приводит к необходимости решения задач идентификации не традиционными методами, а с учетом тех особенностей, о которых было сказано выше. В основе моделирования подобных процессов лежит непараметрическое оценивание функции регрессии по наблюдениям. Из соображений краткости процессы со стохастической зависимостью между компонентами вектора входных переменных в дальнейшем названы H -процессами, а их модели соответственно H -моделями. Подобные процессы типичны для различных отраслей промышленности: металлургической, стройиндустрии, нефтепромышленности, поскольку в них доминируют различные технологические аппараты, объекты, как-то: измельчения, помола, обжига, плавки и другие. Также эти процессы присутствуют в экологических и экономических областях.

Одним из актуальных направлений в теории идентификации являет-

ся задача выделения существенных переменных (показателей), влияющих на значение выходных переменных. Причем на различные компоненты вектора выходных переменных существенное влияние могут оказывать различные наборы входных переменных. Подобная задача рассматривается в диссертации при идентификации в условиях непараметрической неопределенности.

Задачи управления многомерными безынерционными N -процессами в условиях непараметрической неопределенности отличаются от традиционных постановок задач управления. Определяется это, прежде всего, тем, что задающие воздействия, определяющие желаемые значения компонент вектора выхода, не могут быть выбраны произвольно из области определения выходных переменных, как это принято в теории управления. В этом случае, особую важность приобретает задача определения задающих воздействий для компонент вектора выходных переменных. В диссертации рассматривается определение областей выбора задающих воздействий при управлении многомерным безынерционным процессом.

Степень разработанности работы. Идентификацией безынерционных процессов занимаются многие российские исследователи: С. Н. Паньчев, А. А. Мусаев, Н. Д. Поляхов. Из зарубежной литературы можно отметить работы J.A. Hertz, G.E. Hinton, K. Gurney, R. M. Evaristo.

Исследования по отбору существенных переменных изложены в работах А.М. Попова, П.М. Анохина, А. Лепского, Н. Г. Загоруйко. Использование методов непараметрической статистики для отбора существенных переменных рассмотрено в работах А.В. Медведева и Д. В. Безмен, но не был проведен анализ по эффективности работы этого метода и сравнение его с другими методами выделения существенных переменных.

Современная теория идентификации в основном базируется на параметрическом подходе – П. Эйкхофф, Н.С. Райбман, Я.З. Цыпкин, Л. Льюнг и др. При идентификации многомерных N -процессов необходимо провести выбор

такой параметрической структуры модели, в которой будут учтены зависимости между входными переменными. Однако во многих практических задачах выбор такой параметрической структуры представляет существенные трудности, связанные с недостатком априорной информации о процессе.

Н-процессам посвящены работы А.В. Медведева, А.А. Корнеевой, Н.В. Копляровой. В рассматриваемых работах не учитываются особенности, возникающие при моделировании и управлении Н-процессом с несколькими выходными переменными.

Цель данной работы состоит в повышении точности моделирования и эффективности управления многомерными безынерционными Н-процессами при помощи методов непараметрической статистики.

Для достижения поставленной цели необходимо решить следующие задачи:

- 1) разработать алгоритм выделения существенных входных переменных на основании методов непараметрической статистики;
- 2) разработать модификацию алгоритма для построения модели многомерного процесса, имеющего зависимость между входными переменными;
- 3) исследовать возможность уменьшения размерности вектора параметров размытости в моделировании при помощи методов локальной аппроксимации;
- 4) адаптировать алгоритм непараметрического дуального управления для случая управления Н-процессом с несколькими выходными переменными.

Научная новизна диссертационного исследования состоит в следующем:

- предложен новый алгоритм выделения существенных входных переменных, основанный на настройке параметров размытости в непараметрической оценке функции регрессии, позволяющий выбрать входные переменные,

имеющие наибольшее влияние на выходные переменные процесса;

– разработана модификация алгоритма моделирования, включающая в себя параметрическую модель и индикаторную функцию, основанную на методах локальной аппроксимации, и позволяющая строить модели многомерных N -процессов;

– адаптирован алгоритм непараметрического дуального управления с новым видом поискового шага для случая управления N -процессом с несколькими выходными переменными.

Теоретическая значимость Результаты диссертационной работы пополняют арсенал методов выделения существенных переменных. Кроме того, полученные результаты расширяют область применения методов непараметрической статистики для решения задачи моделирования и управления многомерными N -процессами.

Практическая значимость диссертационной работы состоит в следующем:

Разработанная модификация параметрических методов моделирования позволяет моделировать многомерные процессы с зависимыми компонентами вектора входных переменных. Подобные процессы встречаются в стройиндустрии, металлургии, нефтепромышленности, экологии.

Внедрение: Получен акт о применении на производстве результатов диссертационной работы, которым подтверждено, что результаты диссертационной работы будут использованы при разработке дополнений к технологическим инструкциям по эксплуатации котла БКЗ-320 в организации ОАО "Красноярская ТЭЦ-1".

Реализация результатов.

Диссертационное исследование выполнено при поддержке:

– РФФИ, №18-47-240006 "Методы и информационные технологии оценки рисков развития социально-природно-техногенных систем промышленно-

го региона";

– Министерства науки и высшего образования РФ (соглашение № 14.575.21.0142 от 26.09.2017 г., уникальный идентификатор проекта RFMEFI57517X0142);

– Российского научного фонда (проект №16-19-10089).

Методы исследования. При выполнении диссертационной работы использовались методы теории управления, адаптивных и обучающихся систем управления, параметрической и непараметрической идентификации, оптимизации, выделения информативных признаков.

Основные положения, выносимые на защиту:

1) алгоритм выделения существенных переменных, основанный на настройке параметров размытости в непараметрической оценке функции регрессии, позволяет выделить входные переменные, имеющие несущественное влияние на выходные переменные процесса, в условиях непараметрической неопределенности;

2) модификация алгоритма идентификации, включающая в себя параметрическую модель и индикаторную функцию, основанную на методах локальной аппроксимации, позволяет строить модели многомерных N -процессов;

3) алгоритм непараметрического дуального управления с новым видом поискового шага позволяет управлять безынерционным многомерным N -процессом;

4) модель экологической обстановки в городе Красноярске, основанная на методах непараметрической статистики, позволяет восстановить поля распределения различных примесей.

Личный вклад соискателя. Основные научные результаты, выносимые на защиту, получены автором лично. Алгоритм выделения существенных переменных и модель экологической обстановки в г. Красноярске были разработаны совместно с научным руководителем соискателя – д.т.н. А.В. Мед-

ведевым. Модификация алгоритма идентификации для построения моделей многомерных процессов со стохастически-зависимыми входными переменными и модификация алгоритма непараметрического дуального управления для управления многомерными Н-процессами разработаны автором лично.

Апробация работы

Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих конференциях:

"Решетневские чтения"(г. Красноярск, 2014 г.);

"Проблемы управления и моделирования в сложных системах"(г. Самара, 2014 г.);

X Международная конференция "Идентификация систем и задачи управления" SICPRO (г. Москва, 2015);

AMSA'15 (г. Новосибирск, 2015 г.);

"Решетневские чтения" (г. Красноярск, 2015 г.);

"Системы автоматизации в образовании, науке и производстве" (г. Новокузнецк, 2015 г.);

"Проблемы управления и моделирования в сложных системах" (г. Самара, 2015 г.);

"Решетневские чтения" (г. Красноярск, 2016 г.);

"Моделирование неравновесных систем" (г. Красноярск, 2016 г.);

AMSA'17 (г. Красноярск, 2017 г.);

"Системы автоматизации в образовании, науке и производстве" (г. Новокузнецк, 2017 г.).

Публикации. По теме диссертационной работы опубликованы 24 печатные работы, включая 9 работ в научных изданиях из перечня ВАК 5 из которых, в изданиях, индексируемых в международной базе Scopus.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения, списка литературы, включающего 105 наимено-

вания, и приложения. Общий объем работы – 149 страниц основного текста, включая 86 формул, 2 таблицы, 87 рисунков.

1 Общие сведения об идентификации и управлении в условиях неполной информации

Глава посвящена описанию объектов моделирования и управления. Существует множество способов моделирования или управления технологическими и другими видами процессов. Это регламентируется тем что перед исследователями возникают разные постановки задач, а также тем что природа объектов моделирования и управления может сильно различаться. Исходя из данных особенностей, перед тем как описывать те или иные методы моделирования или управления, необходимо дать постановку задачи. Назовем объекты моделирования и управления - исследуемыми процессами.

В первую очередь, необходимо дать общее описание исследуемым процессам. Рассматриваются процессы дискретно-непрерывного типа. В главе приведено схематичное описание рассматриваемым процессам. Приведенная схема является достаточно общей, но на ее примере будет рассмотрены некоторые особенности исследуемых процессов которые могут встретиться исследователю при решении реальных задач.

Отметим что идентификация процесса тесно связана со степенью знаний об исследуемом объекте. Очевидно что чем больше информации об объекте известно исследователю, тем проще решить поставленную задачу моделирования или управления. В главе описаны различные уровни априорной информации и дана краткая характеристика каждому уровню.

Исходя из того каким уровнем априорной информации об объекте обладает исследователь возникают различные задачи идентификации. В случае когда исследователю известно математическое описание (неравенство, система неравенств и т.д.), описывающее исследуемый процесс с точностью до параметров, то исследователю необходимо подобрать параметры при которых

это математическое описание наиболее точно отразит исследуемый процесс. С другой стороны, если исследователю неизвестно математическое описание, описывающее исследуемый процесс, то ему предстоит решить другую задачу идентификации, а именно - найти математическое описание исследуемого процесса. Исходя из этого различают идентификацию в "узком" и "широком" смысле. В главе подробно описан каждый из этих типов идентификации.

Стоит отметить что помимо уровня априорной информации разновидность исследуемого процесса также имеет большое влияние на процесс моделирования или управления. В главе описаны следующие разновидности процессов: Н-процессы; процессы, протекающие в пространстве дробной размерности; процессы, протекающие в пространстве изменяющейся размерности. Каждый из этих процессов был охарактеризован и раскрыты сложности, возникающие при моделировании или управлении этими процессами.

1.1 Характеристика объектов дискретно – непрерывного типа

Будем предполагать что возможно установить зависимость каждой из выходных переменных процесса от входных. Связь между входными и выходными переменными моделируемого процесса может характеризоваться графически, аналитически, или алгоритмически. Независимо от формы представления конструктора, описывающего эту связь, будем именовать его оператором «вход-выход».

Все объекты переменные которых (включая, при необходимости, время) могут принимать несчетное множество сколь угодно близких друг к другу значений, называются непрерывными. Подавляющее большинство реальных физических объектов, состояние которых характеризуется физическими величинами (температура, давление, скорость, ускорение, сила тока, напряженность электрического или магнитного полей и т.д.), обладают свойством

непрерывности.

Математические структуры, адекватно описывающие такие объекты, тоже должны быть непрерывными. Поэтому при модельном описании таких объектов используется, главным образом, аппарат дифференциальных и интегрально - дифференциальных уравнений. Объекты, переменные которых могут принимать некоторое, практически всегда конечное число наперед известных значений, называются дискретными. Примеры: релейно-контактные переключательные схемы, коммутационные системы АТС. Основой формализованного описания дискретных объектов является аппарат математической логики (логические функции, аппарат булевой алгебры, алгоритмические языки). В связи с развитием ЭВМ, дискретные методы анализа получили широкое распространение также для описания и исследования непрерывных объектов [1].

Рассмотрим теперь наиболее существенные, с точки зрения моделирования, внутренние свойства объектов разного класса. При этом будут использоваться понятие структура и параметры моделируемого объекта. Под структурой вид математического выражения, которое связывает его входные и выходные переменные (например: $y = au_1 + bu_2$). Параметры представляют собой коэффициенты, входящие в выражения, которыми описывается структура (а и b).

В реальных объектах, обладающих свойством непрерывности, считывание величин входов и выходов, часто, не может быть непрерывным, соответственно их считывают дискретно. Такие объекты называют дискретно-непрерывными. Они обладают всеми характеристиками непрерывных процессов, но при моделировании их описывают и дискретно.

Стоит сказать что способ моделирования (о них более подробно будет описано во 2 главе) выбирается исходя из нескольких факторов. Одним из самых важных факторов является уровень априорной информации.

1.2 Априорная информация, уровни априорной информации

Рассмотрим процессы с различным уровнем априорной информации [11]:

- процессы с параметрической неопределенностью. Параметрический уровень априорной информации предполагает наличие описание объекта с точностью до параметров и некоторых характеристик случайных помех обычными из которых являются нулевое математическое ожидание и ограниченная дисперсия. Для оценки параметров используются чаще всего разнообразные итеративные вероятностные процедуры. При этих условиях решается задача идентификации в "узком" смысле (более подробно идентификация в "узком" и "широком" смыслах будет описана в следующем разделе);

- процессы с непараметрической неопределенностью. Непараметрический уровень априорной информации не предполагает наличие модели но требует наличия некоторых сведений качественного характера о процессе, например, однозначность либо неоднозначность его характеристик, линейность для динамических процессов либо характер его нелинейности. Для решения задач идентификации на этом уровне априорной информации (идентификация в "широком" смысле) применяются методы непараметрической статистики;

- процессы с параметрической и непараметрической неопределенностью. Важными с точки зрения практики являются задачи идентификации многосвязных систем в условиях когда объем исходной информации не соответствует ни одному из вышеописанных типов. Например, для отдельных каналов многосвязного процесса, на основании физико-химических закономерностей, энергетических, закона сохранения массы, балансовых соотношений, могут быть выведены параметрические закономерности, а для других нет. Таким образом, мы находимся в ситуации когда задача идентификации фор-

мулируется в условиях и параметрической, и непараметрической априорной информации. Тогда и модели представляют собой взаимосвязанную систему параметрических и непараметрических соотношений.

1.3 Идентификация в «узком» и «широком» смыслах

В зависимости от априорной информации об объекте управления различают задачи идентификации в "узком" и "широком" смыслах [2].

Исследователь, на первом этапе, при идентификации в "узком" смысле, на основании априорной информации, определяет функционал, определенный с точностью до параметров $\vec{\alpha}$, описывающий зависимость выходных переменных объекта моделирования \vec{x} от входных \vec{u} (A^α):

$$\hat{\vec{x}}_a(t) = A^\alpha(\vec{u}(t), \vec{\alpha}) \quad (1.3.1)$$

где A^α — функционал, определенный с точностью до параметров $\vec{\alpha}$; $\vec{\alpha}$ — вектор параметров; $\vec{u}(t)$ — вектор входных переменных, размерностью n ; $\vec{x}(t)$ — вектор выходных переменных, размерностью k , $\hat{\vec{x}}_a(t)$ — оценка вектора выходных переменных $\vec{x}(t)$. Другие входные переменные опущены из соображений простоты.

Следующий этап состоит в оценке параметров $\vec{\alpha}$ и состояния системы по результатам наблюдений над входными и выходными переменными (выборки $(\vec{x}_i, \vec{u}_i, i = \overline{1, s})$, s — объем выборки), полученными в условиях функционирования объекта.

При идентификации в "узком" смысле требуется, чтобы априорная информация об объекте содержала в себе описание объекта с точностью до вектора параметров.

Априорной информации об объекте при идентификации в "широком" смысле меньше чем при идентификации в "узком" смысле. Априорной ин-

формации при идентификации в "широком" смысле недостаточно чтобы найти функционал, определенный с точностью до параметров $\vec{\alpha}$, описывающий зависимость выходных переменных от входных, поэтому необходимо решать большое число предварительных задач. К этим задачам относятся: задание класса модели, выбор структуры модели, оценка степени стационарности и линейности объекта, оценка степени и формы влияния входных переменных на выходные, выбор существенных переменных и т. д.

Стоит сказать что априорная информация достаточная для получения параметрической структуры объекта в реальных задачах встречается редко. В связи с этим можно признать что идентификация в "широком" смысле является более подходящей для решения реальных задач.

К настоящему времени накоплен большой опыт решения задач идентификации в "узком" смысле. Методы решения задач идентификации в "широком" смысле начали разрабатываться только в последние годы. Это, в первую очередь, можно объяснить трудностью задачи.

1.4 Особые случаи моделируемых процессов

Некоторые исследуемые процессы могут обладать рядом особенностей вследствие которых моделирование или управление этими процессами может быть затруднено. В диссертации рассмотрены следующие виды исследуемых процессов: Н-процессы; процессы, протекающие в пространстве дробной размерности; процессы, протекающие в пространстве с изменяющейся размерностью.

Рассмотрим каждый из этих процессов подробнее.

1.4.1 Н-процессы

При моделировании и управлении безынерционными объектами с запаздыванием часто возникает ситуация, когда компоненты вектора входных переменных \vec{u} стохастически — зависимы, причем характер этой зависимости, чаще всего, неизвестен.

Подобные процессы типичны для различных отраслей промышленности: металлургической, стройиндустрии, нефтепромышленности, поскольку в этих отраслях доминируют технологические аппараты, объекты, как-то: измельчения, помола, обжига, плавки и другие, а так же для экологических и экономических областей.

Моделирование дискретно-непрерывных процессов в условиях помех при неполной информации продолжает оставаться одной из актуальных проблем теории управления. Случай когда входные переменные (управляемые и неуправляемые) оказываются стохастически – зависимы представляет специальный интерес и может быть отнесен к категории новых задач идентификации. Последнее приводит к тому что исследуемый процесс протекает не только в многомерном кубе, определяемом набором входных и выходных переменных, но в некоторой его подобласти. В дальнейшем, из соображений краткости, подобные процессы будем называть Н – процессами.

В качестве примера можно привести любой процесс у которого входные переменные должны быть заданы в известных пропорциях, это такие процессы как: выплавка стали (в которой должны быть смешаны в нужной пропорции углерод и железо); замешивание теста (здесь нужно соблюдать пропорции добавления муки, воды, соли и других ингредиентов).

В трехмерном пространстве, Н-процесс может выглядеть в форме «трубки», линии или части плоскости.

В простом варианте, Н-процесс представляет собой прямую линию (Ри-

сунок 1.1).

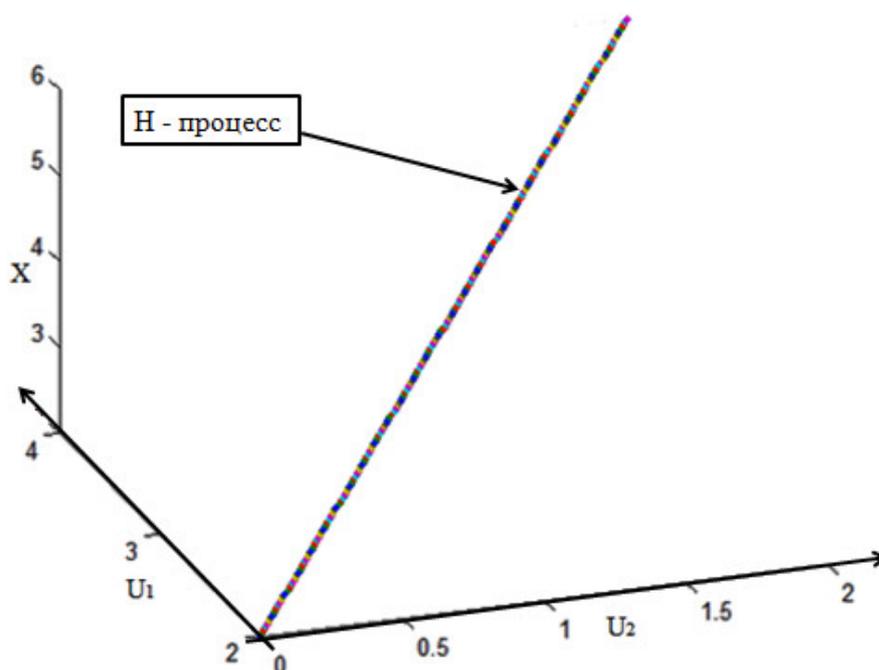


Рисунок 1.1 — Пример H-процесса. Линейная зависимость между входными переменными

Пример, изображенный на рисунке 1.1 является простым примером H-процесса. Здесь, u_2 имеет линейную зависимость от u_1 , нет помех при считывании информации, на процесс помехи также не оказывают влияние.

Конечно, ситуации когда на процесс не действуют помехи в реальных задачах встречаются крайне редко, но линейная зависимость может быть. Например, когда плавят сталь определенной марки имеется информация в каких пропорциях должно быть железо с углеродом. Также, если будет увеличен объем углерода но не увеличен объем железа, то будет получен чугун. В этом случае управление выйдет за рамки процесса получения стали.

Разумеется не все процессы существуют при линейно – зависимых переменных, в сложных технологических или химических процессах между компонентами вектора \vec{u} может быть и нелинейная зависимость (рисунок 1.2).

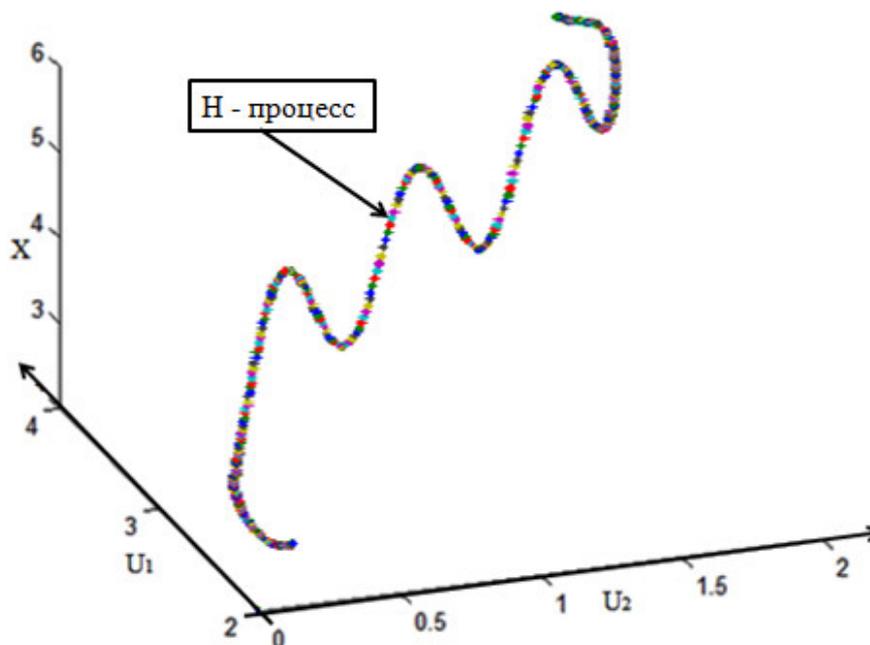


Рисунок 1.2 — Пример H-процесса. Нелинейная зависимость между входными переменными

На рисунке 1.2 изображен H-процесс в котором связь между элементами вектора входных переменных – нелинейная. Это — более сложные процессы чем процесс, отображенный на рисунке 1.1. В рассматриваемом процессе для успешного управления нужно знать не просто — в каких пропорциях должны быть входные переменные, а нужно точно знать - зависимость между ними. Представленные сложные процессы чаще встречаются на реальном производстве, чем рассмотренные ранее. Это и химические процессы, и добыча металлов, и процессы, происходящие в электронных схемах.

В рассмотренных процессах на рисунках 1.1 и 1.2, компоненты вектора входа находятся в функциональной зависимости. Иными словами, если входные параметры – заданы даже с незначительным отклонением от зависимостей то процесс сразу перестает существовать. Очевидно что для реально существующих процессов это – не так, ведь если, например, при плавке стали добавить незначительно меньше нужного количества углерода то все равно

– это будет сталь, или если незначительно налить больше воды при замешивании теста то это останется тестом, то есть процесс существует.

Учитывая вышесказанное приведем пример Н-процесса, более приближенного к реальным процессам (рисунок 1.3).

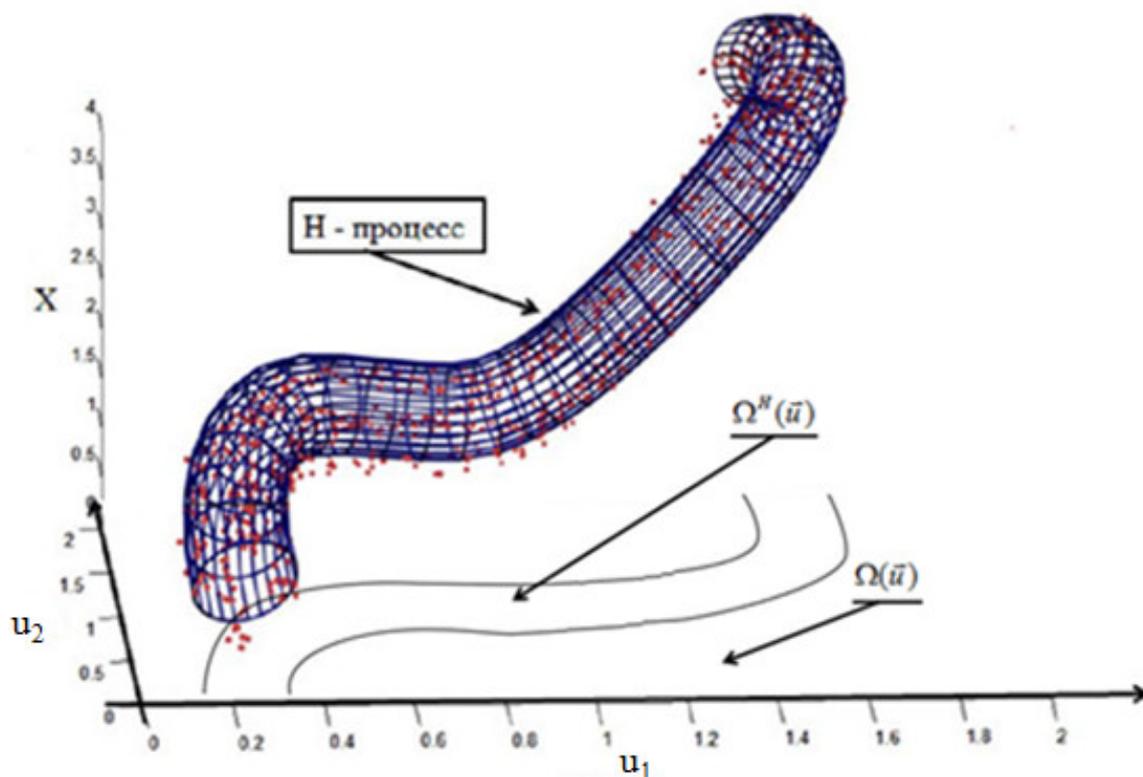


Рисунок 1.3 — Пример Н-процесса, наиболее близкого к реальным процессам

На рисунке 1.3 отображен наиболее близкий к действительности вид Н-процесса. Приняты следующие обозначения: $\Omega(\vec{u})$ – область входных переменных без учета взаимной связи между компонентами вектора входных воздействий \vec{u} . Например, в технологических процессах область $\Omega(\vec{u})$ определяется технологической картой. $\Omega^H(\vec{u})$ – область входных переменных Н-процесса, с учетом взаимной связи между компонентами вектора входных воздействий \vec{u} . Важной особенностью является то, что во многих сложных

технологических процессах факт наличия области $\Omega^H(\vec{u})$ неизвестен.

На рисунке 1.3, при помощи решетчатой структуры, указан объем и область существования Н-процесса. Как мы здесь видим, на рисунке 1.3 процесс проходит уже в какой – то области, а не существует только на линии, иными словами зависимость между компонентами вектора входных воздействий \vec{u} — стохастическая, а не функциональная.

Некоторые технологические процессы характеризуются множеством выходных переменных. В случае с Н - процессами это означает что для каждой выходной переменной x_i существует своя область определения $\Omega_i^H(\vec{u})i = \overline{1, k}$, где k — размерность вектора выходных переменных \vec{x} . Эта особенность вызывает значительные сложности при моделировании или управлении подобными процессами. В дальнейшем такие процессы будут называться многомерными Н-процессами. Подробнее моделирование многомерных Н-процессов раскрыто в главе 2.6.

Важно отметить, что часто в сложных технологических процессах параметрическая структура, описывающая исследуемый процесс — неизвестна. Также неизвестен вид параметрической зависимости между входными переменными. Все это вместе взятое приводит к необходимости разработки нетрадиционных методов для идентификации и управления подобными процессами.

1.4.2 Процессы протекающие в пространстве дробной размерности

Рассмотрим следующий частный случай идентификации безынерционной системы. Пусть объект описывается уравнением:

$$x(\vec{u}) = f(u_1, u_2, u_3), \quad (1.4.1)$$

где $\vec{u} = (u_1, u_2, u_3)$ — вектор входных переменных размерностью $n=3$, а x — выходная переменная. Традиционный путь построения модели процесса, описываемого (1.4.1) состоит в определении класса параметрических зависимостей $\hat{x}(\vec{u}) = \hat{f}(u_1, u_2, u_3, \vec{\alpha})$ и последующей оценки параметров $\vec{\alpha}$ тем или иным способом по выборке наблюдений $(\vec{u}_i, x_i, i = \overline{1, s})$, s — объем выборки.

Проанализируем этот пример с разных точек зрения. Пусть компоненты вектора входных переменных $\vec{u} = (u_1, u_2, u_3)$ стохастически никак не связаны т.е. независимы. В этом случае, естественно использовать обычный традиционный прием, описанный выше. Теперь предположим что объективно компоненты вектора входных переменных функционально связаны, например,

$$u_2 = \varphi_1(u_1), u_3 = \varphi_2(u_2) = \varphi_2(\varphi_1(u_1)). \quad (1.4.2)$$

Естественно исследователь не знает о существовании зависимостей (1.4.2), в противном случае можно было бы сделать подстановку (1.4.2) в (1.4.1), и получить следующую зависимость x уже от одной переменной u_1 вида:

$$x(u) = f(u_1, \varphi_1(u_1), \varphi_2(\varphi_1(u_1))). \quad (1.4.3)$$

Таким образом, зависимость (1.4.1) приведенных выше условиях может быть сведена к одномерной зависимости x от u_1 . В случае, если зависимость u_3 от u_2 объективно отсутствует, то (1.4.1) легко приводится к виду (1.4.4):

$$x(\vec{u}) = f(u_1, \varphi_1(u_1), u_3), \quad (1.4.4)$$

т.е. к двумерной зависимости x от u_1, u_3 . Отсюда можно заключить, что при наличии функциональной зависимости между компонентами вектора \vec{u} мы получаем зависимость x от \vec{u} в данном случае одно-двух-трехмерные. Подчеркнем еще раз что о наличии функциональных зависимостей между компонентами вектора входных переменных исследователю не известно. А теперь

проанализируем наиболее интересный случай, имеющий непосредственное отношение к Н-процессам.

Пусть u_3 и u_2 хотя и неизвестным образом но стохастически связаны. Подчеркнем – стохастически, а не функционально. Вернемся еще раз к анализу того, что произошло. Во-первых, если компоненты вектора \vec{u} независимы то исследуемый процесс описывается функцией трех переменных. Если две компоненты вектора входных переменных \vec{u} связаны функциональной зависимостью то процесс описывается функцией двух переменных. Наконец, если две переменные связаны стохастически то процесс описывается функцией более чем двух переменных но менее чем трех. Можно считать, что мы приходим к зависимости от дробного числа переменных и, следовательно, к пространству дробной размерности.

Естественно считать, что Н-процесс протекает именно в пространстве дробной размерности, так как входные воздействия в Н-процессе имеют стохастическую зависимость.

Помимо этого у исследуемых процессов зависимость между входными переменными как и сам объект может изменяться с течением времени. Это обусловлено тем, что большинство исследуемых процессов — нестационарные.

1.4.3 Процессы протекающие в пространстве с изменяющейся размерностью

При решении задач идентификации многомерных систем исследователь может столкнуться с ситуацией, когда некоторые входные-выходные переменные, характеризующие состояние объекта, утрачивают свое значение при его описании. В некоторых случаях, по разным причинам, в процессе исследования или работы с объектом, возникают некоторые новые переменные. Это

может быть вызвано множеством причин. Одна из самых часто встречаемых заключается в том, что исследуемый объект не является стационарным, то есть исследуемый объект изменяется со временем. В качестве примера, можно привести "старение" оборудования.

Например, пусть объект моделирования описывается уравнением:

$$x(u) = f(u_1, u_2, u_3) \quad (1.4.5)$$

В случае стохастической зависимости между переменными $u_2(u_1), u_3(u_1)$ по имеющимся в наличии обучающим выборкам можно вычислить квадратичную ошибку прогноза δ оценки $\hat{u}_{2s}(u_1), \hat{u}_{3s}(u_1)$. $\hat{u}_{is}(u_j)$ - модель компоненты вектора входных переменных u_i , построенная по значению компоненты вектора входных переменных u_j .

$$\delta_{21} = \sum_{i=1}^s (u_2 - \hat{u}_{2s}(u_1))^2 / \sigma_{u_2}^2, \delta_{31} = \sum_{i=1}^s (u_3 - \hat{u}_{3s}(u_1))^2 / \sigma_{u_3}^2, \quad (1.4.6)$$

где δ_{ij} это среднеквадратичная ошибка прогноза u_i от u_j .

«Силу» стохастической связи λ между двумя произвольными переменными можно, например, вычислить по формуле:

$$\lambda_{ij} = 1 - \delta_{ij}, \quad (1.4.7)$$

$$\delta_{ij} = \sum_{i=1}^n (u_i - \hat{u}_i(u_j))^2 / \sigma_{u_i}^2. \quad (1.4.8)$$

В формуле (1.4.8) $\sigma_{u_i}^2$ - дисперсия компоненты вектора входных переменных u_i , $\hat{u}_i(u_j)$ - оценка компонента вектора входных переменных u_i , построенная при помощи значений компоненты вектора входных переменных u_j .

Из формулы видно что в случае функциональной связи между компонентами вектора входных переменных $\lambda = 1$, в случае отсутствия зависимости между компонентами вектора входных переменных $\lambda = 0$, при стохастической зависимости между входными переменными $0 < \lambda < 1$.

Если в более общем случае такого рода процессы интерпретировать как функции многих переменных, то изменчивость этой функции во времени может быть, например, показана на нижеследующей цепочке соотношений, действующих во времени.

$$x(\vec{u}, t) = \begin{cases} f(t, u_1, \mathbf{u}_2, \mathbf{u}_3, u_4, u_5) - T_1; \\ f(t, u_1, u_2, \mathbf{u}_3, u_4, u_5) - T_2; \\ f(t, u_1, u_2, u_3, u_4, u_5) - T_3; \\ f(t, u_1, u_2, u_3, u_4, u_5, u_6) - T_4; \\ f(t, u_1, u_2, u_3, u_4, u_5, u_6) - T_5; \\ f(t, u_1, u_2, u_3, u_4, u_5, u_6) - T_6; \\ f(t, u_1, \mathbf{u}_2, u_3, u_4, \mathbf{u}_5, u_6) - T_7; \\ f(t, \mathbf{u}_1, \mathbf{u}_2, u_3, u_4, \mathbf{u}_5, \mathbf{u}_6, u_7) - T_8. \end{cases} \quad (1.4.9)$$

Наиболее темным цветом (\mathbf{u}_1) обозначены компоненты вектора входных воздействий \vec{u} которые оказывают самое сильное влияние на выходную переменную x (функциональная зависимость). Менее темное обозначение (u_1) говорит о более слабом влиянии компоненты \vec{u} на выходную переменную x (возможно, достаточно сильная, стохастическая зависимость), ещё более слабое влияние на выходную переменную x оказывают компоненты вектора входных переменных u_1 и u_1 . $T_i, i = \overline{1, 8}$ — интервалы времени существования соответствующих зависимостей.

Таким образом в реально действующих процессах подобного рода сте-

пени влияния входных переменных на выходную — изменчивы. Из приведенных выше зависимостей видно что некоторые компоненты вектора входных воздействий \vec{u} утрачивают свое значение, некоторые — утрачивают, а потом восстанавливают, а некоторые новые компоненты вектора входных воздействий \vec{u} , появляются впервые, как например u_6, u_7 .

Если сохранить математический «облик» интерпретации функции многих переменных как точку многомерного пространства, то мы приходим к наличию пространства дробной размерности F^λ . Оценку размерности F^λ можно осуществить, например, так:

$$\dim F^\lambda = (n + 1) - \sum_{i=1}^{n-1} \lambda_{i,i+1} \quad (1.4.10)$$

где n — размерность вектора \vec{u} , а $\lambda_{i,i+1}$ означает «силу» стохастической связи между u_i и u_{i+1} .

Могут быть предложены и другие способы оценки размерности пространства в котором протекает процесс. Например,

$$\dim F^\lambda = (n + 1) - \sum_{i=1}^{n-1} \lambda_{1,i} \quad (1.4.11)$$

$\lambda_{1,i}$ зависимость всех компонент вектора u_i от одной компоненты u_1 .

Проведем некоторые вычислительные эксперименты, по вычислению размерности пространства в котором протекает процесс.

Пусть процесс описывается функцией $x = f(u_1, u_2)$, и находится под воздействием помехи $\xi(t)$. $f(u_1, u_2) = u_1^2 + 2u_2 + \xi(t)$.

Для начала построим график зависимости между оценкой размерности пространства в котором протекает процесс $\dim F^\lambda$ и уровнем помехи ξ в процентах (Рисунок 1.4). Объем выборки наблюдений примем равной 500, входные переменные — независимы.

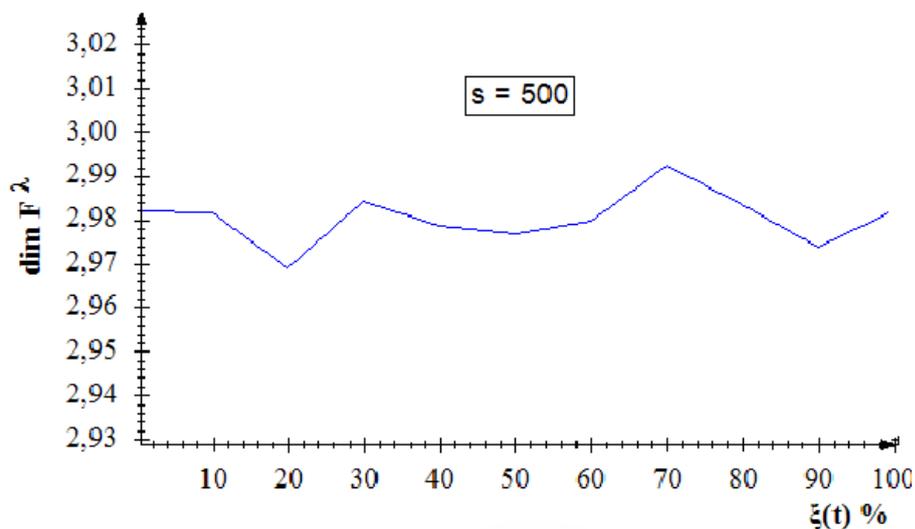


Рисунок 1.4 — Зависимость размерности пространства F^λ от уровня помех

Как можно увидеть на рисунке 1.4, уровень помехи ξ не влияет на оценку размерности пространства в котором протекает процесс $\dim F^\lambda$ в случае если компоненты вектора входных переменных — независимы. Это можно объяснить так: размерность пространства в котором протекает процесс равна размерности вектора входных воздействий плюс один (выходная переменная) и уменьшается только в случае если между компонентами вектора входных переменных имеется стохастическая или функциональная зависимость. Так как все компоненты вектора u_i — независимы, то увеличение уровня помех, хоть и приведет к ухудшению качества модели но никак не повлияет на оценку размерности пространства в котором протекает процесс $\dim F^\lambda$.

В следующем эксперименте построим график зависимости между оценкой размерности пространства в котором протекает процесс $\dim F^\lambda$ и объемом выборки наблюдений s (Рисунок 1.5). Компоненты вектора \vec{u} — независимы. Проведем оценку размерности пространства в котором протекает процесс $\dim F^\lambda$ в двух случаях: когда на процесс не действует помеха и когда на процесс действует помеха $\xi = 10\%$.

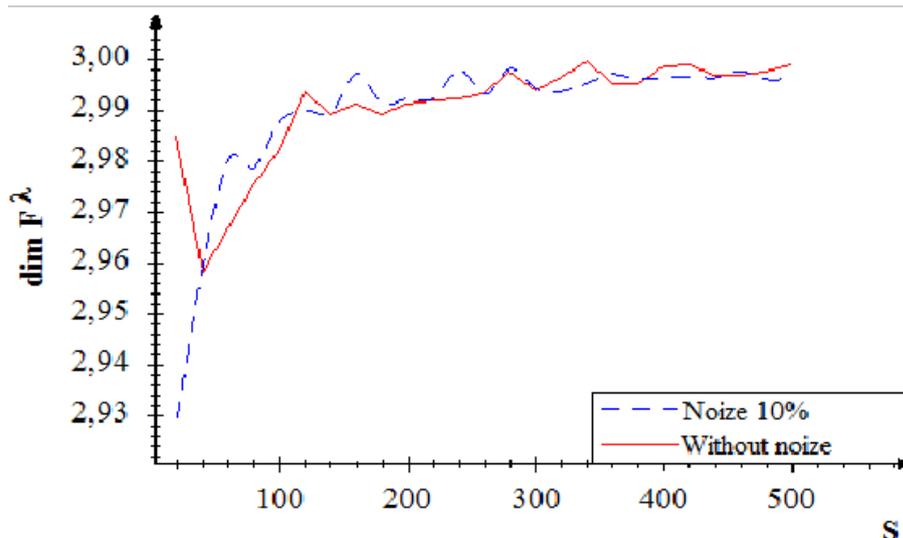


Рисунок 1.5 — Зависимость размерности пространства F^λ от объема выборки

Рисунок 1.5 иллюстрирует что при небольшой выборке размерность F^λ мала. Это можно объяснить тем что при малом объеме выборки наблюдений между компонентами вектора входных переменных может быть "иллюзорная" зависимость. Например, при объеме выборки наблюдений $s=2$ можно предположить что обе точки наблюдений находятся на одной линии (так как через 2 точки пространства всегда можно провести одну линию), и предположить что между ними есть зависимость. Увеличение объема выборки позволяет точнее провести оценку размерности пространства в котором протекает процесс $dim F^\lambda$, приближая ее к значению 3. Помеха (ξ), как и в прошлом случае, не влияет на оценку размерности пространства в котором протекает процесс $dim F^\lambda$.

Для следующих экспериментов рассмотрим процесс у которого между компонентами вектора входных воздействий имеется зависимость то есть N-процесс. Зависимость между входными переменными, в данном примере, пусть будет следующая: $u_2 = 3u_1$.

Построим график зависимости между оценкой размерности простран-

ства в котором протекает Н-процесс $dimF^\lambda$ и уровнем помехи ξ , в процентах (Рисунок 1.6). Объем выборки наблюдений примем равной 500.

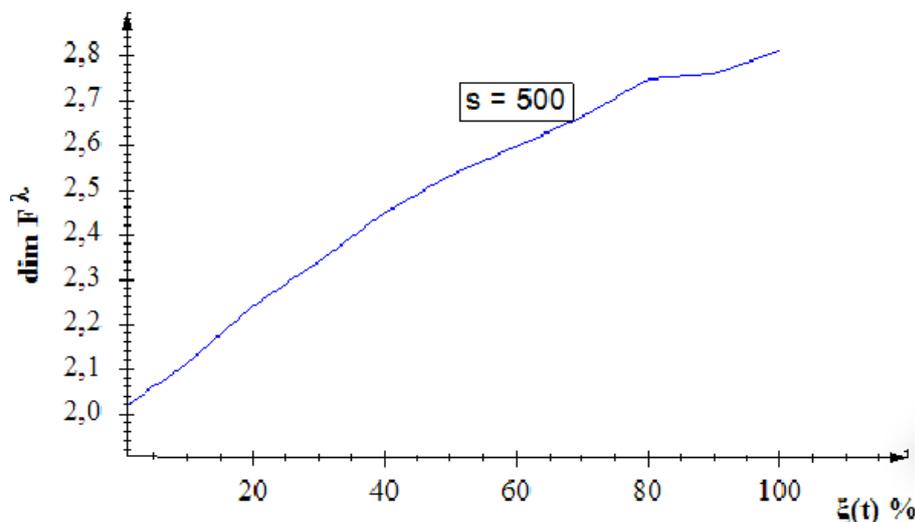


Рисунок 1.6 — Зависимость размерности пространства F^λ от уровня помех

Как можно видеть на рисунке 1.6 при низком уровне помехи $\xi \approx 0$, оценка размерности пространства в котором протекает процесс $dimF^\lambda$ близка к 2. Учитывая что между компонентами вектора входных воздействий имеется функциональная зависимость, процесс протекает в пространстве двух переменных (одно входное воздействие и одна выходная переменная). Поэтому при низком уровне помехи оценка размерности пространства в котором протекает процесс $dimF^\lambda$ производится точно. При увеличении помехи зависимость между входными воздействиями ослабевает и, как следствие, пространство в котором протекает процесс - увеличивается, как и отображено на рисунке 1.6.

Теперь построим график зависимости между оценкой размерности пространства в котором протекает процесс $dimF^\lambda$ и объемом выборки наблюдений s (Рисунок 1.7). Как и в прошлый раз, проведем оценку размерности пространства в котором протекает процесс $dimF^\lambda$ в двух случаях: когда на процесс не действует помеха и когда на процесс действует помеха $\xi = 10\%$.

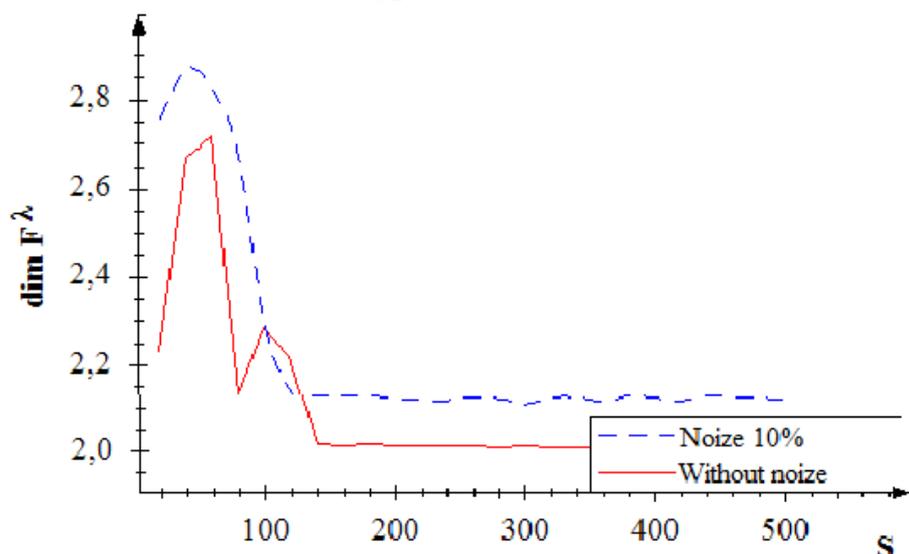


Рисунок 1.7 — Зависимость размерности пространства F^λ от объема выборки

Как и предполагалось на рисунке 1.7 видно что при небольшом объеме выборки наблюдений оценка размерности пространства в котором протекает процесс $\dim F^\lambda$ производится не точно. При увеличении объема выборки наблюдений оценка уточняется и $\dim F^\lambda$ приближается к значению 2. Важно отметить что при воздействии $\xi = 10\%$ помехи, $\dim F^\lambda \approx 2,1$. Учитывая что помеха $\xi = 10\%$, разумно предположить что размерность пространства в котором протекает процесс, увеличится на 0,1, что видно на рисунке 1.7. Это демонстрирует что предложенный алгоритм оценки размерности пространства в котором протекает процесс, работает в условиях помех.

Проведем вычислительный эксперимент для случая когда на входе процесса действуют 10 независимых переменных, а выход, без нарушения общности, скалярный.

Построим график зависимости между оценкой размерности пространства в котором протекает процесс $\dim F^\lambda$ и уровнем помехи ξ в процентах (Рисунок 1.8). Объем выборки наблюдений примем равной 500, входные переменные – независимы.

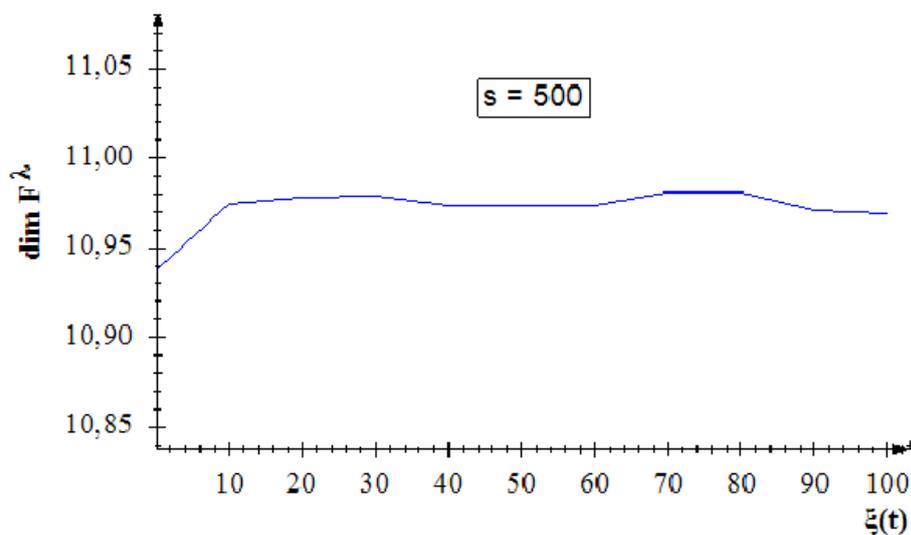


Рисунок 1.8 — Зависимость размерности пространства F^λ от уровня помех

Как можно увидеть на рисунке 1.8 в случае если компоненты вектора входных переменных \vec{u} - независимы то уровень помех не влияет на оценку размерности пространства в котором протекает процесс $\dim F^\lambda$. Это подтверждает что предложенный способ оценки размерности пространства в котором протекает процесс $\dim F^\lambda$ можно использовать для многомерных случаев.

Теперь построим график зависимости между оценкой размерности пространства в котором протекает процесс $\dim F^\lambda$, и объемом выборки наблюдений s (Рисунок 1.9). Как и в прошлые разы, проведем оценку размерности пространства в котором протекает процесс $\dim F^\lambda$ в двух случаях: когда на процесс не действует помеха и когда на процесс действует помеха $\xi = 10\%$.

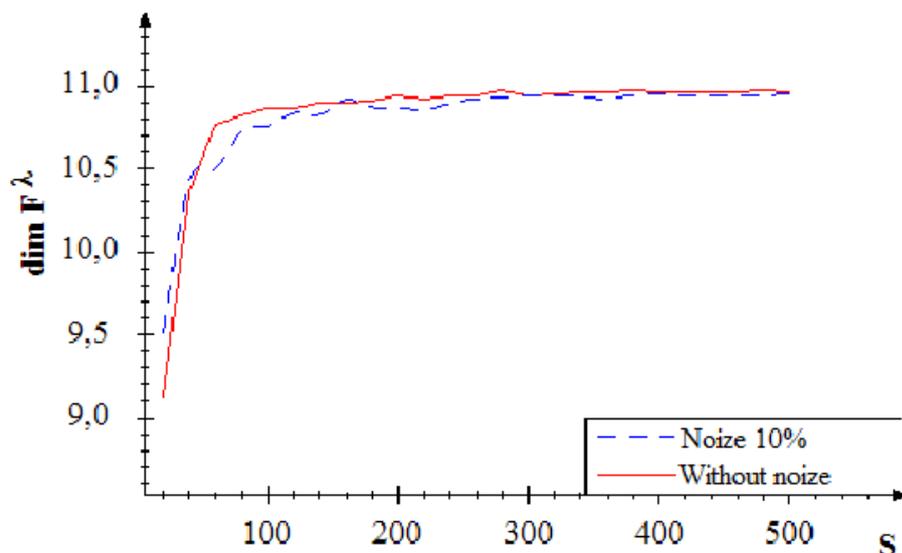


Рисунок 1.9 — Зависимость размерности пространства F^λ от объема выборки

Рисунок 1.9 демонстрирует что в случае если компоненты вектора входных переменных - независимы (рисунок 1.5), то при небольшой выборке размерность F^λ мала. Увеличение объема выборки позволяет точнее провести оценку размерности пространства в котором протекает процесс $dim F^\lambda$, приближая ее к значению 11. Помеха (ξ) не влияет на оценку размерности пространства в котором протекает процесс $dim F^\lambda$.

Для следующих экспериментов рассмотрим процесс у которого компоненты вектора входных воздействий \vec{u} находятся в функциональной зависимости.

Построим график зависимости оценки размерности пространства в котором протекает процесс $dim F^\lambda$ от уровня помех (Рисунок 1.10).

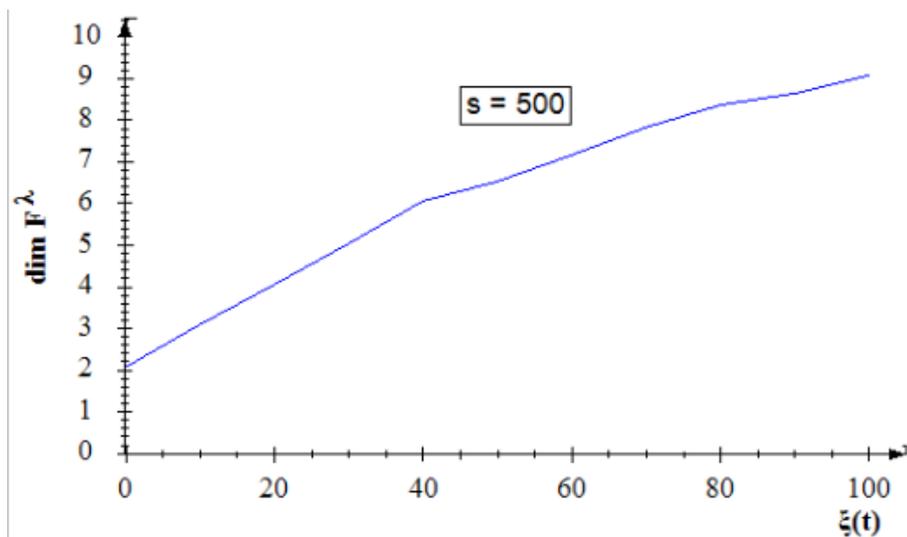


Рисунок 1.10 — Зависимость размерности пространства F^λ от уровня помех

Рисунок 1.10 демонстрирует что, как и предполагалось, оценка размерности пространства в котором протекает процесс $\dim F^\lambda$ при функциональной зависимости компонент вектора входа равна двум. Увеличение значения уровня помехи $\xi(t)$ на входные переменные и, как следствие, уменьшение этой зависимости приводит к увеличению оценки $\dim F^\lambda$.

Теперь построим график зависимости между оценкой размерности пространства в котором протекает процесс $\dim F^\lambda$, и объемом выборки наблюдений s (Рисунок 1.9). Как всегда, проведем оценку размерности пространства в котором протекает процесс $\dim F^\lambda$ в двух случаях: когда на процесс не действует помеха и когда на процесс действует помеха $\xi = 10\%$.

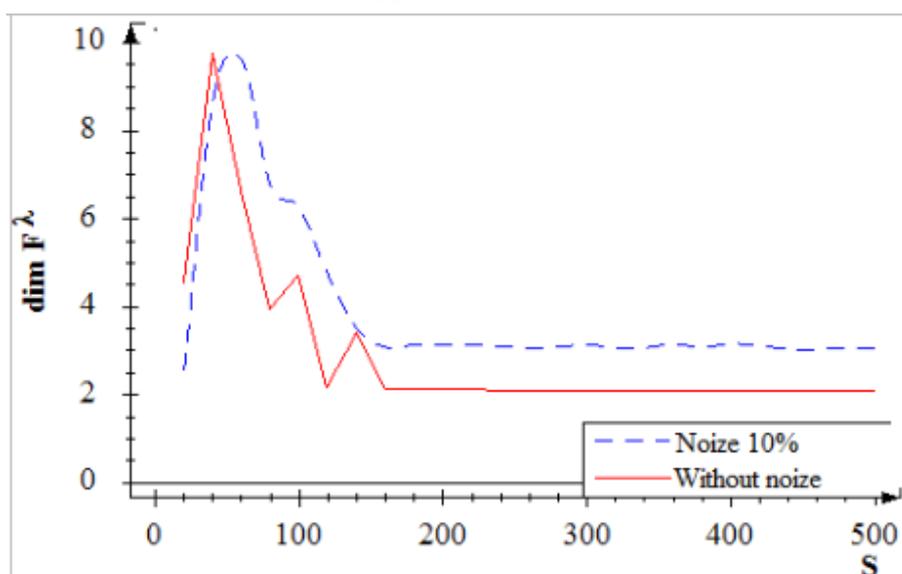


Рисунок 1.11 — Зависимость размерности пространства F^λ от объема выборки

На рисунке 1.11 видно что при небольшом объеме выборки наблюдений оценка размерности пространства в котором протекает процесс $\dim F^\lambda$ производится не точно. При увеличении объема выборки наблюдений оценка уточняется и $\dim F^\lambda$ приближается к значению 11. Важно отметить, что при воздействии помехи $\xi = 10\%$, $\dim F^\lambda \approx 3$. Это объясняется тем что помеха $\xi = 10\%$, воздействует на каждую компоненту вектора входных воздействий \vec{u} , соответственно размерность пространства в котором протекает процесс увеличится на следующее значение: $0,1 * n$ ($n = 10$), где n — размерность вектора входных переменных.

1.5 Выводы по главе

Глава посвящена общим вопросам, связанным с моделированием процессов дискретно-непрерывного вида.

В первую очередь дана характеристика исследуемым процессам. Описана природа процессов дискретно-непрерывного вида.

В дальнейшем были рассмотрены различные уровни априорной информации. Подробно описаны особенности характерные для каждого уровня априорной информации. Уделено внимание вопросам моделирования процессов для различных уровней априорной информации.

Также в главе затронут вопрос идентификации в "узком" и "широком" смыслах. Обращено внимание что большая часть исследований посвящена идентификации в "узком" смысле.

Последняя часть главы посвящена исследуемым процессам которые обладают рядом особенностей. Рассмотренные следующие виды процессов: Н-процессы; процессы, протекающие в пространстве дробной размерности; процессы, протекающие в пространстве с изменяющейся размерностью.

Н-процессы - процессы у которых компоненты вектора входных воздействий \vec{u} имеют стохастическую или функциональную зависимость. В главе раскрыта особенность, возникающая при моделировании подобных процессов, которая заключается в том, что процесс протекает не во всей области, определяемой вектором \vec{u} и выходных \vec{x} переменных — $\Omega(\vec{u}, \vec{x})$, а только в некоторой подобласти $\Omega^H(\vec{u}, \vec{x})$.

В главе продемонстрировано что процессы у которых компоненты вектора входных воздействий \vec{u} имеют стохастическую или функциональную зависимость протекают в дробной размерности. Был рассмотрен способ оценки размерности в котором протекает процесс.

В конце главы была рассмотрена природа процессов которые протекают в пространстве с изменяющейся размерностью.

Особенности рассмотренных процессов приводит к тому что необходимо разработать качественно новые методы идентификации и управления такими процессами.

2 Модели дискретно непрерывных процессов

Моделирование технологических, социальных и других процессов является одной из актуальных задач уже на протяжении многих лет. Не является удивительным тот факт, что человек пробует описать тот или иной процесс, чтобы предсказать его поведение, а в дальнейшем, возможно, и управлять им.

Перед тем как описывать способы моделирования, безусловно необходимо дать постановку задачи, ее специфику, и возможный результат. С данного описания и начинается глава.

В главе описаны некоторые из методов идентификации. Приведены такие методы как: подбор параметрической структуры; генетическое программирование; МНК; локальная аппроксимация; нейросети.

В качестве отдельной задачи рассмотрена задача выделения существенных переменных. Вне зависимости от того какой метод моделирования выберет исследователь, ему необходимо будет выбрать переменные на основании которых будет строиться модель. Данная необходимость встречается как в задачах моделирования, так и в задачах классификации (необходимо выделить признаки на основании которых будет классифицирован объект). Исходя из вышесказанного можно утверждать, что выделение существенных переменных - актуальная задача.

2.1 Постановка задачи

Рассмотрим классическую схему идентификации [3, 4]:

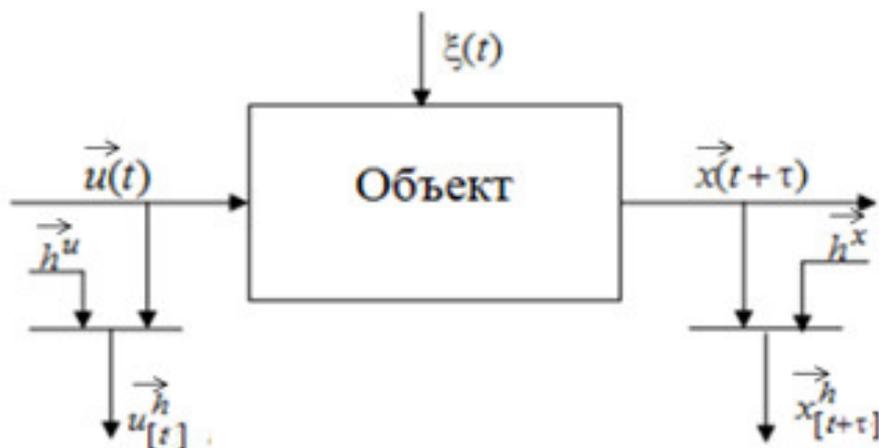


Рисунок 2.1 — Классическая схема идентификации

На рисунке 2.1 приняты следующие обозначения: $\vec{u}(t)$ – это вектор входных переменных, размерности n , в момент времени t ; $\vec{x}(t+\tau)$ – вектор выходных переменных, размерности k , в момент времени $(t+\tau)$, τ – запаздывание. $\xi(t)$ – помеха, действующая на объект. \vec{h}^u, \vec{h}^x – помехи в каналах связи, $\vec{u} = (u_1, u_2, \dots, u_n) \in \Omega(\vec{u}) \subset R_n$, где $\Omega(\vec{u})$ – область, которой принадлежат значения (u_1, u_2, \dots, u_n) . Стоит сказать, что каждая компонента вектора \vec{u} , $u_i \in [a_i, b_i], i = \overline{1, n}$, где n – количество входных переменных, а $x_j \in [c_j, d_j], j = \overline{1, k}$, где k – количество выходных переменных. При исследовании реальных процессов значения коэффициентов a_i, b_i, c_j, d_j чаще всего известны. В технологических процессах, например, значения этих коэффициентов задаются технологическим регламентом (картой).

Задача стоит в описании поведения данного процесса при помощи математического языка (равенства, системы уравнений, дифференциальных уравнений и т. д.).

Описание поведения изучаемого процесса называется математической моделью. Математические модели бывают нескольких типов:

- модели с полной информацией о процессе. В этом случае оператор процесса известен точно, а случайные помехи, действующие на объект и в

каналах связи отсутствуют, кроме того

$$\vec{u}(t) \in \Omega(\vec{u}(t)) \quad (2.1.1)$$

где $\Omega(\vec{u}(t))$ – некоторая известная область допустимых значений $\vec{u}(t)$;

- модели с неполной информацией о процессе. Это модели с независимым (пассивным) накоплением информации. Класс операторов не известен точно, но имеются предположения о плотности распределения всех случайных факторов. Обычно плотности вероятности случайных факторов, действующих на объект и в каналах измерения переменных, предполагаются нормальными и аддитивными. Ясно что в этом случае необходимо наличие выборки $(\vec{u}_i, \vec{x}_i), i = \overline{1, s}$, а сами наблюдения предполагаются статистически независимыми, т.е., например, наблюдения $(\vec{u}_1, \vec{x}_1), (\vec{u}_2, \vec{x}_2)$ – независимы;

- модели с активным накоплением информации. Задачи идентификации и задачи управления здесь могут быть объединены. Элементы выборки измерений поступают последовательно в обучающую модель и систему управления, т.е. $(\vec{u}_1, \vec{x}_1), (\vec{u}_2, \vec{x}_2), (\vec{u}_3, \vec{x}_3)$. Таким образом, в случае объединения этих задач, выработка входных воздействий носит двойственный характер, они должны быть одновременно и изучающими, и управляющими, т.е. носить дуальный характер (более подробно о дуальном управлении будет сказано в главе 3.4). Однако, если помехи, действующие на процесс, аддитивны и в каналах измерения, то в целом система дуального управления может быть приведена к разомкнутой, т.е. темп накопления информации оказывается независимым от значений входных переменных. Такие процессы называют приводимыми к разомкнутому или нейтральными.

2.2 Выделение существенных переменных

В настоящее время задача моделирования процесса решается множеством способов. Известны методы моделирования при помощи ядерной ап-

проксимации, нейросетей, генетического программирования и другие.

Моделируемые процессы могут иметь множество входных воздействий, но важно помнить, что не все входные воздействия имеют одинаковое влияние на выходные. Те входные воздействия, которые имеют слабое влияние на выходные переменные являются несущественными, а те значения которых имеют сильное влияние на выходные переменные - существенными. Стоит задача выделить среди компонентов вектора входных воздействий \vec{u} - существенные.

Приведем пример, демонстрирующий актуальность задачи выделения существенных переменных.

Для моделирования объекта необходимо наличие большой выборки, и чем больше переменных имеется у объекта, тем больше должна быть выборка. В качестве примера можно привести цитату председателя Экспертного совета Российского фонда фундаментальных исследований по математике, механике и информатике, академика Евгения Моисеева: "Существует очень важная проблема, как сделать заключение на основании малого - в смысле математической статистики - количества данных. У нас на рассмотрении была задача из области медицины примерно с 20 параметрами и всего 600 данными. По всем правилам нельзя делать выводы по 600 данным, если меняются 20 параметров".

Стоит сказать что в подобных задачах выделение существенных переменных имеет особенное значение. Создание модели при большом количестве входных переменных затруднительно и требует большого объема выборки наблюдений. Если выделить из компонентов вектора входных воздействий - только существенные и только их учитывать в модели, то результаты моделирования будут точнее из-за уменьшения размерности задачи моделирования.

Учитывая вышесказанное, можно утверждать, что выделение существенных переменных - является актуальной задачей.

Безусловно, существенность переменных понятие относительное и его определение - непростая задача.

Существует два способа оценки существенности переменных: прямой и косвенный [9].

Прямой состоит в следующем: необходимо выделить набор переменных при котором средняя ошибка моделирования достаточно мала.

Исходя из косвенного пути степень существенности переменных оценивается на основании свойств распределения: $(M(u_1), \dots, M(u_n), D(u_1), \dots, D(u_n))$.

Рассмотрим косвенные способы оценки существенности.

Один из способов выделения существенных переменных - это построение корреляционной матрицы. Данный способ несмотря на свою известность имеет неточность в определении существенных переменных, как было доказано А.Г. Французом [12].

В настоящее время наибольшей популярностью пользуется оценка степени существенности при помощи критерия Фишера. Согласно этому критерию, чем меньше дисперсия выборки и больше максимальная разница между двумя элементами выборки наблюдений - тем более существенная переменная.

Данный способ, несмотря на свои преимущества, не учитывает зависимости между входными переменными. Это значит что при выборе переменных на основании которых будет строиться модель, могут быть выбраны взаимозаменяемые переменные.

Также большой популярностью пользуется метод дисперсионного анализа, но в данном методе тот же недостаток, что и в критерии Фишера.

Также стоит упомянуть про такой метод как метод главных компонент. Данный метод позволяет уловить линейную зависимость между входными переменными, но не улавливает нелинейную зависимость, что также может

повлиять на точность определения существенных переменных.

Рассмотрим прямые способы оценки существенности.

Предположим стоит задача из n переменных выбрать m ($m < n$) существенных. Существует несколько алгоритмов для выполнения данной задачи.

Алгоритм Del

Согласно этому алгоритму первоначально из вектора входных переменных \vec{u} , необходимо исключить компоненту u_1 и рассчитать $\sigma(u_2, \dots, u_m)$ без учета исключенной компоненты вектора входных переменных. $\sigma(u_2, \dots, u_n)$ — средняя ошибка моделирования. Затем те же действия, поочередно выполняем над переменными u_2, \dots, u_n и получаем $\sigma(u_1, u_3, \dots, u_n), \dots, \sigma(u_1, \dots, u_{n-1})$, соответственно. Наименее существенная переменная выделяется по правилу: $\max \sigma(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n) \rightarrow \min I(u_i)$ где u_i - наименее существенная переменная (она исключается из вектора \vec{u}), $I(u_i)$ — степень существенности. Проще говоря, переменная u_i считается наименее существенной если модель, построенная без её учета, была самой грубой среди других. После этого остается $n - 1$ переменная и алгоритм необходимо повторять, пока в итоге не останется m существенных переменных.

При данном подходе количество проделанных итераций для выделения существенных переменных вычисляется по формуле (2.2.1).

$$L = n + (n - 1) + (n - 2) + \dots + (m + 1) = \sum_{i=1}^{n-m} (n - i) \quad (2.2.1)$$

Алгоритм Ad

Согласно этому алгоритму необходимо поочередно построить модели исследуемых процессов основываясь на единственной переменной u_i , $i = \overline{(1, n)}$. Та переменная у которой ошибка моделирования наименьшая - будет наиболее существенной. К найденной переменной u_i мы добавляем поочередно переменные $u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n$, которых уже $n - 1$ и получаем

соответствующие наборы переменных $(u_i, u_1), \dots, (u_i, u_n)$. Далее необходимо построить модели, основываясь на полученных наборах. Та переменная, которую добавили в набор, на основании которого получили наиболее точную модель, является существенной. Данную операцию необходимо повторять пока не будет набран набор из m переменных.

При данном подходе количество проделанных итераций для выделения существенных переменных равна количеству итераций в алгоритме Del.

Алгоритм AdDel

Продемонстрированные алгоритмы выделения существенных переменных входят в класс так называемых «жадных» алгоритмов.

Проблема таких алгоритмов в том, что при получении оптимального решения на каждом шаге они не обеспечивают нахождения глобального минимума. Под глобальным минимумом здесь имеется ввиду набор из самых существенных компонентов вектора \vec{u} .

Для увеличения вероятности выделения самых существенных компонентов вектора \vec{u} используют «релаксационный» алгоритм (алгоритм AdDel). Согласно этому алгоритму исследователь сначала при помощи алгоритма Ad находит a_1 существенных переменных, затем при помощи алгоритма Del исключает a_2 ($a_2 < a_1$) переменных из выбранных. Данный алгоритм продолжается пока не будет найден набор состоящий из m переменных.

У алгоритма AdDel есть одна особенность. Критерий качества (в данном случае, средняя точность моделирования) при использовании данного алгоритма изменяется, согласно рисунку 2.2.

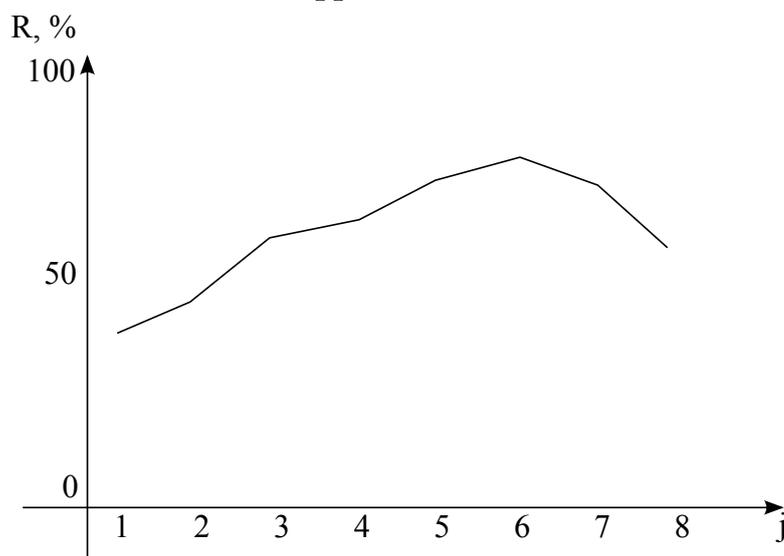


Рисунок 2.2 — Зависимость точности моделирования и количества используемых переменных

Как видно на рисунке 2.2, увеличение количества переменных сначала приводит и к увеличению точности, но затем в набор переменных попадают мешающие (шумовые) переменные, что приводит к ухудшению качества классифицирования. Этот «перегиб» функции $\sigma(\vec{u})$ позволяет автоматически указать оптимальное (под критерием оптимальности понимается средняя точность моделирования) количество переменных, что очень важно для задач моделирования.

Алгоритм на основе настройки вектора параметров размытости

Перед использованием алгоритма выделения существенных переменных на основе настройки вектора параметров размытости необходимо провести центрирование и нормирование элементов вектора \vec{u} из выборки наблюдений.

Предложенный алгоритм основан на оптимизации вектора параметров размытости, в непараметрической оценке функции регрессии по наблюдениям.

Непараметрическая оценка функции регрессии по наблюдениям имеет вид:

$$\hat{x}_l(\vec{u}, \vec{c}_s) = \frac{\sum_{i=1}^s x_i \prod_{j=1}^n \Phi\left(\frac{u_j - u_{ji}}{c_{sj}}\right)}{\sum_{i=1}^s \prod_{j=1}^n \Phi\left(\frac{u_j - u_{ji}}{c_{sj}}\right)}, l = (\overline{1, k}), \quad (2.2.2)$$

где $\hat{x}_l(\vec{u}, \vec{c}_s)$ это непараметрическая оценка функции $x_l(\vec{u})$ в точке \vec{u} ; \vec{u} — вектор входных переменных, k — размерность вектора выходных переменных \vec{x} ; $\Phi\left(\frac{u_j - u_{ji}}{c_{sj}}\right)$ — колокообразная функция; $c_{sj}, j = (\overline{1, n})$ — параметры размытости.

$\Phi\left(\frac{u - u_i}{c_s}\right)$ — функция, удовлетворяющая следующим условиям (2.2.3) и (2.2.4):

$$\frac{1}{s} \int_{-\infty}^{\infty} \Phi\left(\frac{u - u_i}{c_s}\right) du = 1 \quad (2.2.3)$$

$$\lim_{c_s \rightarrow 0} \frac{1}{s} \int_{-\infty}^{\infty} \varphi(u) \Phi\left(\frac{u - u_i}{c_s}\right) du = \varphi(u_i) \quad (2.2.4)$$

В качестве $\Phi\left(\frac{u - u_i}{c_{sj}}\right)$ может быть использована треугольная функция (2.2.5).

$$\Phi\left(\frac{u - u_i}{c_{sj}}\right) = \begin{cases} 1 - \left|\frac{u - u_i}{c_{sj}}\right|, & \left|\frac{u - u_i}{c_{sj}}\right| < 1 \\ 0, & \left|\frac{u - u_i}{c_{sj}}\right| > 1 \end{cases} \quad (2.2.5)$$

Как видно из (2.2.2) с каждой компонентой вектора \vec{u} мы связали соответствующую компоненту вектора \vec{c}_s .

Предложенный алгоритм выделения существенных входных переменных для выходной переменной x_i выглядит следующим образом:

1. Введем вектор $\vec{u}^{<i>}$, в который входят компоненты вектора \vec{u} , влияющие на x_i . Компоненты вектора \vec{u} , влияющие на x_i выбираются на основании априорной информации об исследуемом процессе. В случае если априорной информации недостаточно для выделения компонентов вектора $\vec{u}^{<i>}$, то вводятся все компоненты вектора \vec{u} ;

2. Решить задачу (2.2.6):

$$\sigma_i(\vec{u}^{<i>}, \vec{c}_s^{<i>}) = \sqrt{\frac{1}{s} \sum_{j=1}^s (\hat{x}_i(\vec{u}^{<i>}, \vec{c}_s^{<i>}) - x_{ji})^2} \rightarrow \min_{\vec{c}_s^{<i>}} \quad (2.2.6)$$

3. Считать кандидатом на исключение, как наименее существенную, компоненту вектора $\vec{u}^{<i>}$, у которой, после решения задачи (2.2.6), соответствующий элемент вектора $\vec{c}_s^{<i>}$ – наибольший. Обозначим указанную компоненту как $u_m^{<i>}$;

4. Рассчитать значение $\sigma_i(\vec{u}^{<i>}, \vec{c}_s^{<i>})$ без учета $u_m^{<i>}$;

5. Если полученное значение $\sigma_i(\vec{u}^{<i>}, \vec{c}_s^{<i>})$ увеличилось, перейти к шагу

7. В случае если полученное значение $\sigma_i(\vec{u}^{<i>}, \vec{c}_s^{<i>})$ не изменилось, исключить $u_m^{<i>}$ из вектора $\vec{u}^{<i>}$, как оказывающую наименьшее влияние на результат аппроксимации;

6. Если после исключения вектор $\vec{u}^{<i>}$ состоит только из одной компоненты, перейти к шагу 7, иначе перейти к шагу 2;

7. Оставшиеся компоненты вектора $\vec{u}^{<i>}$ считать существенными переменными.

Таким образом, в результате оптимизации (2.2.6) по $\vec{c}_s^{<i>}$, получена цепочка неравенств, например $c_{s2}^* < c_{s9}^* < c_{s1}^* < \dots < c_{s4}^* < c_{s23}^*$, откуда следует, что претендентом на отбрасывание является компонента вектора u_{23} , что подтверждается вычислением при его отсутствии (отбрасывании).

Доказательство работы метода

Стоит задача построения модели, описывающей процесс $x = f(\vec{u})$. В качестве метода моделирования был использован метод локальной аппроксимации:

$$\hat{x}_l(\vec{u}^{<l>}) = \frac{\sum_{i=1}^s x_i \prod_{j=1}^n \Phi((u_{jl} - u_{ji})/c_{sj})}{\sum_{i=1}^s \prod_{j=1}^n \Phi((u_{jl} - u_{ji})/c_{sj})}, \quad (2.2.7)$$

где s — объем выборки наблюдений, n — количество входных воздействий. Φ — колокообразная функция.

Качество построенной модели оценивается на основании значения среднеквадратичной ошибки (2.2.8).

$$\sigma(\vec{c}_s) = \sqrt{\frac{1}{s} \sum_{i=1}^s (x_i - \hat{x}_i(\vec{u}_i, \vec{c}_s))^2}, \quad (2.2.8)$$

где $\hat{x}_i(\vec{u}_i, \vec{c}_s)$ — это непараметрическая оценка функции, $(\vec{u}_i, x_i), i = (\overline{1, s})$ — выборка наблюдений.

Очевидно, что значения вектора \vec{c}_s имеют влияние на $\sigma(\vec{c}_s)$. Исходя из этого возникает задача оптимизации вектора \vec{c}_s .

Предположим, что задача оптимизации вектора \vec{c}_s была решена и был получен наилучший вектор \vec{c}_s^* .

Логично предположить, что элементы вектора \vec{c}_s^* можно составить в цепочку неравенств, например $c_{s6}^* < c_{s3}^* < c_{s12}^* < \dots < c_{s1}^* < c_{s9}^*$.

Предположим, что один из элементов $c_{sm} \rightarrow \infty$. Это значит, что $\forall u_i, ((u_{mj} - u_{ij})/c_{sm}) \rightarrow 0$, значит, $\Phi((u_{mj} - u_{ij})/c_{sm}) \rightarrow 1$. Учитывая, что модель описывается уравнением (2.2.9), можно утверждать, что элемент u_m не влияет на результат моделирования процесса.

$$\hat{x}_l = \frac{\sum_{i=1}^s x_i \prod_{j=1}^n \Phi((u_{jl} - u_{ji})/c_{sj})}{\sum_{i=1}^s \prod_{j=1}^n \Phi((u_{jl} - u_{ji})/c_{sj})} = \frac{\int_1^s x_i \prod_{j=1}^n \Phi((u_{jl} - u_{ji})/c_{sj})}{\int_1^s \prod_{j=1}^n \Phi((u_{jl} - u_{ji})/c_{sj})}, \quad (2.2.9)$$

Так как модель описывает поведение объекта, то можно утверждать, что элемент u_m не влияет и на сам процесс (при условии, что модель достаточно точно описывает процесс).

Исходя из вышесказанного, можно утверждать что компонент вектора входных переменных \vec{u} у которого соответствующий параметр размытости является наибольшим - имеет наименьшее влияние на процесс.

Численные эксперименты

Проведем вычислительные эксперименты с целью продемонстрировать выделение существенных переменных при помощи различных алгоритмов, и провести сравнение этих алгоритмов.

В моделируемом объекте количество входных переменных будет меняться от 2 до 40. В качестве выходных данных эксперимента будут представлены графики, которые отображают скорость с которой происходит выделение существенных переменных. Ось абсцисс представляет количество входных переменных (n), а ось ординат представляет время, затраченное на решение задачи выделения существенных переменных, в секундах.

Выделение существенных переменных будет происходить при помощи следующих алгоритмов: алгоритм Ad; алгоритм Del; алгоритм AdDel; алгоритм, основанный на настройке вектора \vec{c}_s .

Продemonстрируем выделение существенных переменных при помощи алгоритма Ad для разного количества переменных.

Зависимость количества входных переменных и времени, затраченного на решение задачи выделения существенных переменных для данного алгоритма представлена на рисунке (2.3)

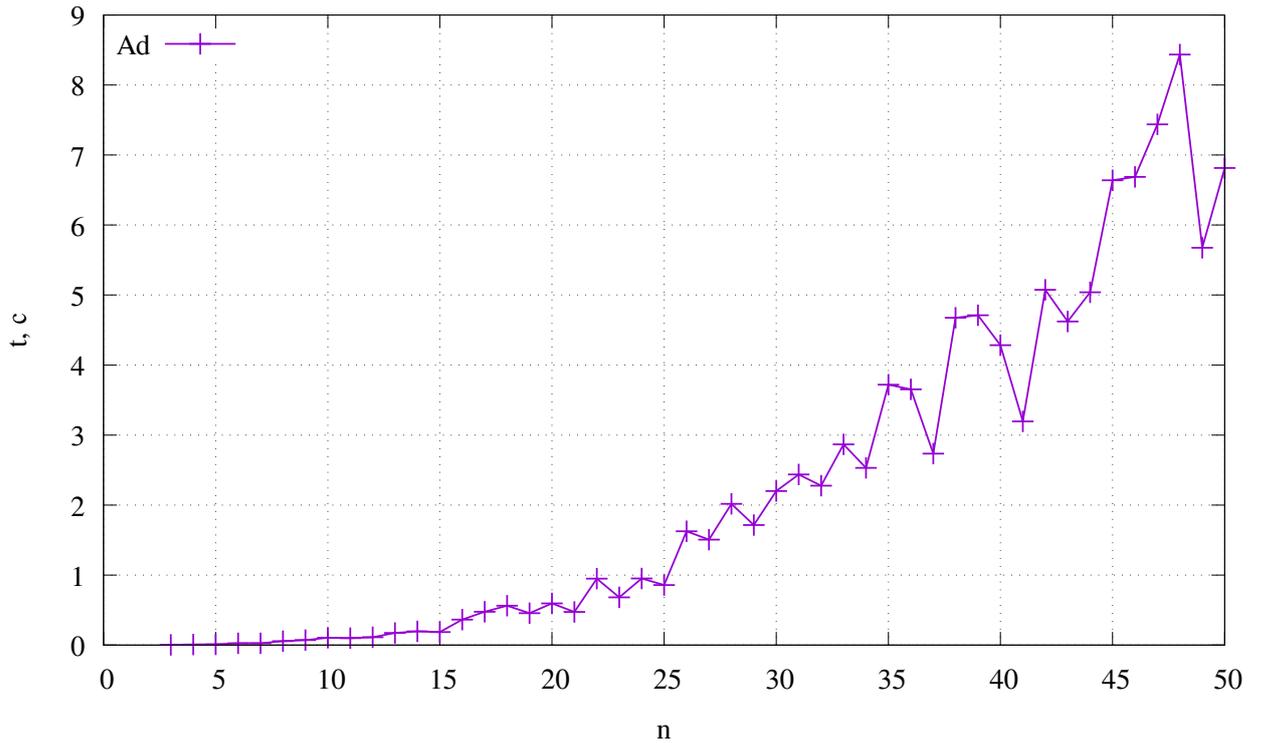


Рисунок 2.3 — Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных.

Работа алгоритма Ad

Как видно на рисунке (2.3), зависимость имеет нелинейный характер. Это означает что при большом количестве входных переменных, выделение существенных переменных при помощи алгоритма Ad потребует значительное количество вычислительных ресурсов.

Продemonстрируем выделение существенных переменных, при помощи алгоритма Del.

Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных для данного алгоритма представлена на рисунке (2.4)

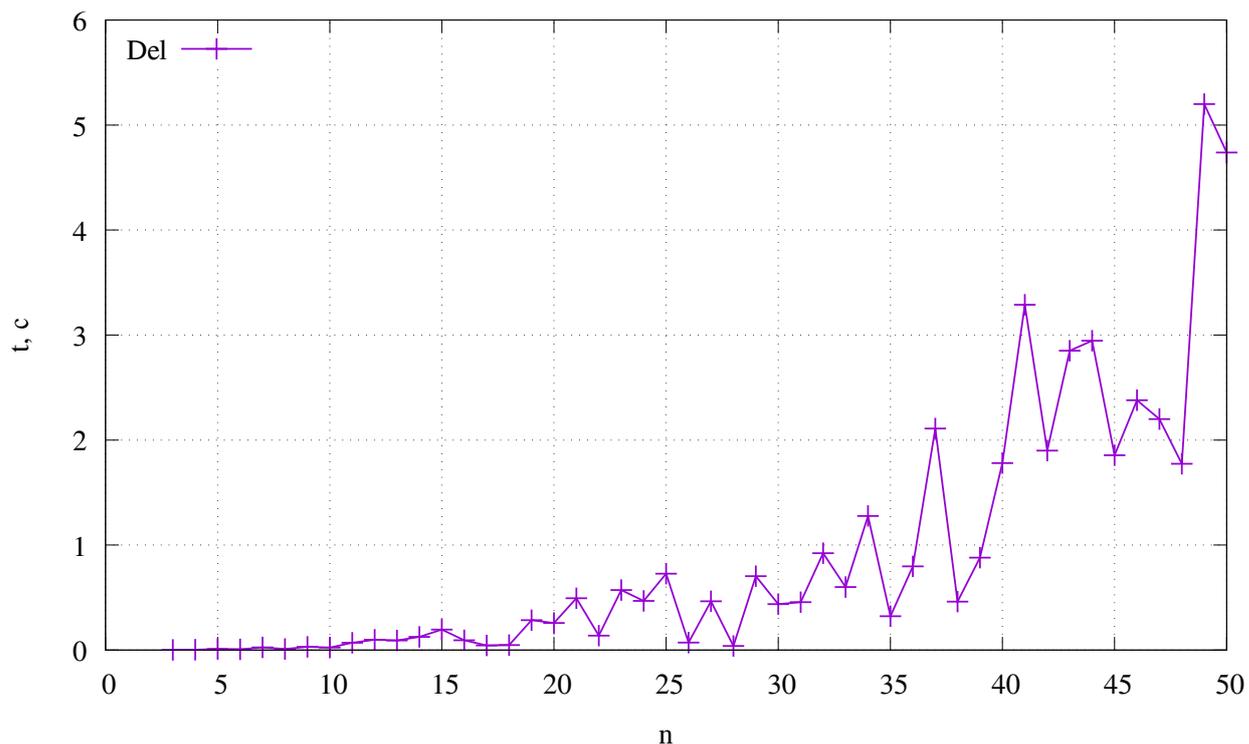


Рисунок 2.4 — Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных.

Работа алгоритма Del

Как видно на рисунке (2.4), зависимость также имеет нелинейный характер.

Стоит отметить, что алгоритмы Ad и Del относятся к категории, так называемых, "жадных" алгоритмов, это обозначает что они при нахождении локального минимума не всегда находят вектор наиболее существенных переменных. Для исправления данного недостатка был разработан алгоритм AdDel.

Продemonстрируем выделение существенных переменных при помощи алгоритма AdDel. Как было указано выше, алгоритм AdDel основан на алгоритмах Ad и Del, исследователь сначала при помощи алгоритма Ad находит a_1 самых существенных переменных, затем при помощи алгоритма Del

исключает a_2 ($a_2 < a_1$) переменных из выбранных. В вычислительном эксперименте $a_1 = 3$, $a_2 = 1$.

Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных для данного алгоритма представлена на рисунке (2.5)

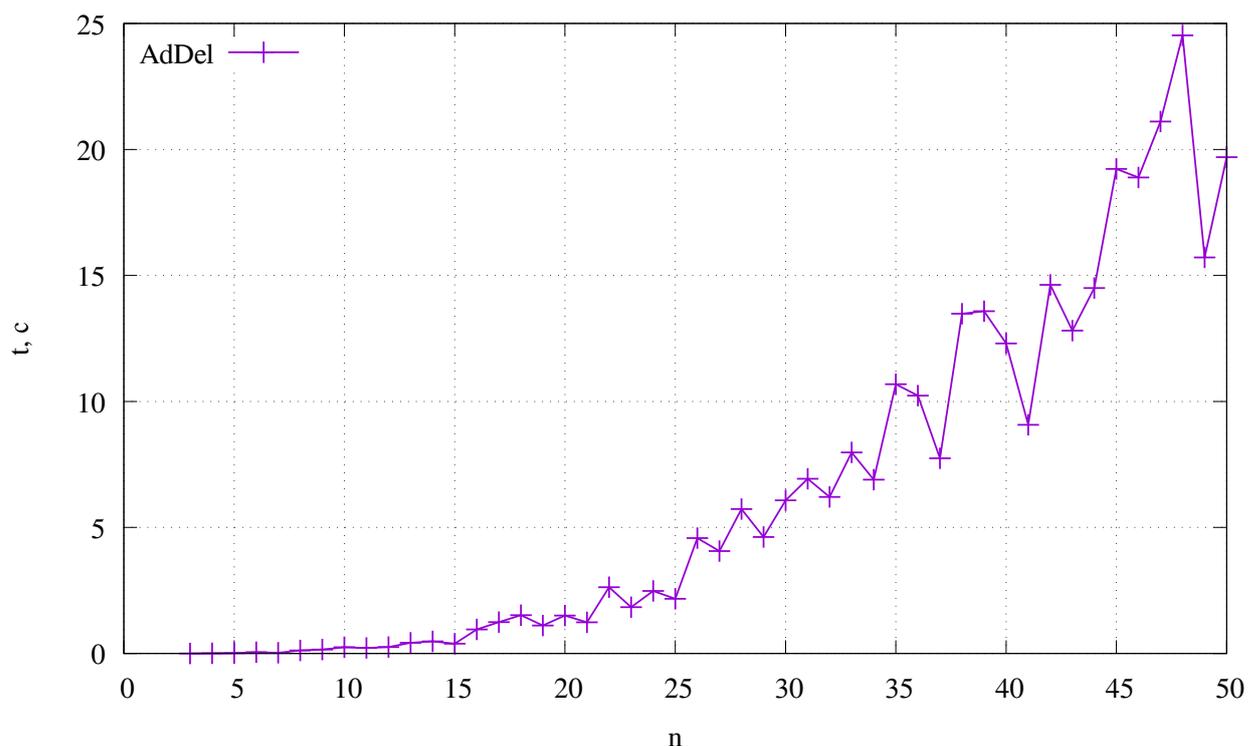


Рисунок 2.5 — Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных.

Работа алгоритма AdDel

Как видно на рисунке (2.5), зависимость имеет не только нелинейный но и экспоненциальный характер. Это обозначает что данный алгоритм имеет еще большие требования к вычислительным мощностям чем алгоритмы, представленные на рисунках (2.3 - 2.4).

Теперь представим алгоритм, основанный на настройке вектора \vec{c}_s .

Стоит отдельно сказать о работе данного алгоритма. Так как мы будем

находить такой вектор $\vec{c}_s, \sigma(\vec{c}_s) \rightarrow 0$, то необходимо решить задачу оптимизации. Оптимизировать мы будем, используя метод деформируемого многогранника. Вследствие особенности данного метода, скорость сходимости будет сильно зависеть от изначально выбранных точек (c_{s1}, \dots, c_{sn}) .

Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных для данного алгоритма представлена на рисунке (2.6)

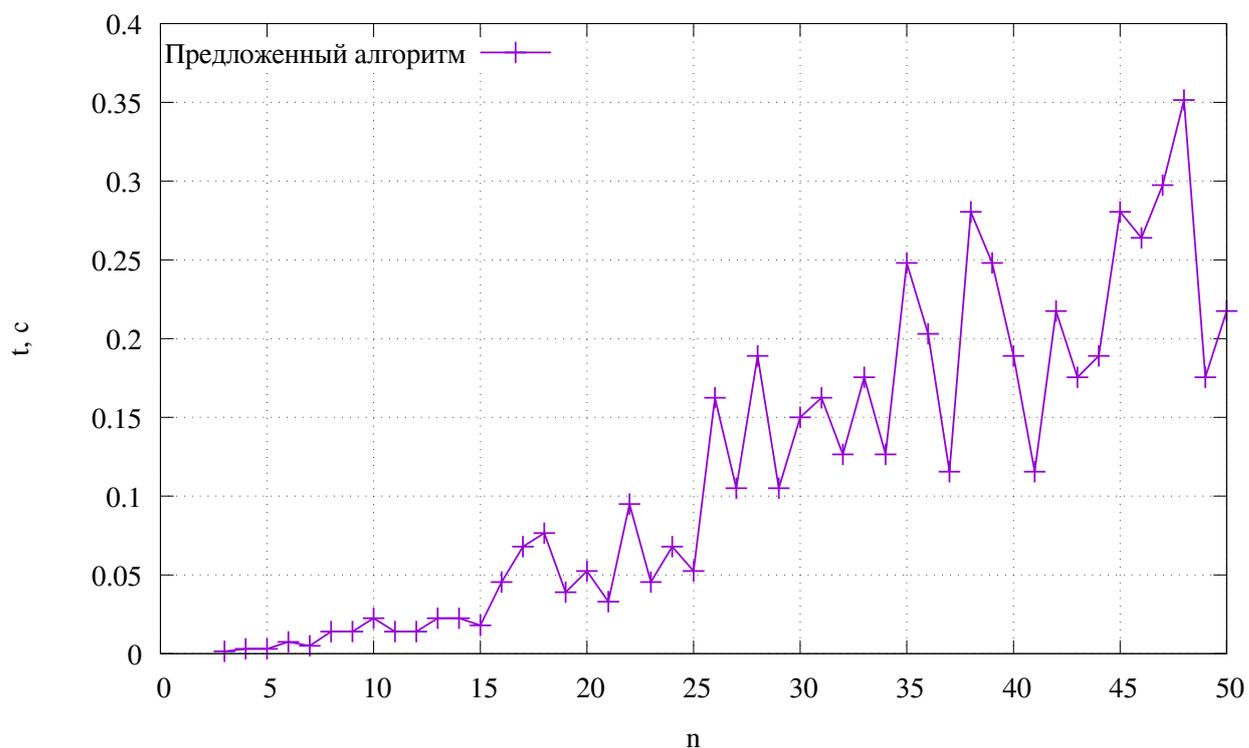


Рисунок 2.6 — Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных.

Работа алгоритма, основанном на настройке \vec{c}_s

Как видно на рисунке 2.6, зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных, для алгоритма, основанном на настройке \vec{c}_s имеет линейный характер, в отличие от зависимостей ранее продемонстрированных методов. И при большом

количестве входных переменных, этот метод требует меньшее количество вычислений, чем представленные на рисунках (2.3 - 2.5).

В современных задачах часто возникает ситуация когда при большом количестве переменных существенность которых необходимо оценить, имеется малый объем данных. Задача моделирования или выделения существенных переменных, в таких случаях, может оказаться крайне затруднительной. Вследствие этого был предложен модифицированный метод выделения существенных переменных на основе настройки вектора \vec{c}_s .

Суть модификации заключается в следующем: все входные переменные процесса разбиваются на группы, критерием разбивки является оценочная степень существенности по мнению эксперта. Для каждой группы входных переменных назначается один компонент вектора \vec{c}_s . Например, предположим, что исследуется процесс с 60 входными переменными. Эксперты разбили эти переменные на 3 группы по степени существенности. На каждую группу назначен свой параметр размытости. И теперь, вместо того чтобы решать задачу оптимизации по 60 параметрам, необходимо будет решить задачу оптимизации только по 3 параметрам. Важной особенностью данной модификации является то, что после решения задачи 2.2.6 исключается не один компонент вектора входных переменных \vec{u} , а вся группа, у которой соответствующий компонент вектора \vec{c}_s - наибольший.

Таким образом решается две задачи. Во первых - задача оптимизации упрощается т.к. уменьшается количество параметров в задаче оптимизации. Во вторых - увеличивается точность прогноза, вследствие того, что создание решающего правила для меньшего числа переменных проще, и, как следствие, более точное.

Недостатком данной модификации является то, что по ошибке эксперта в группу к несущественным переменным может попасть существенная переменная и наоборот. Также в качестве недостатка можно отметить что для

работы данного алгоритма необходима помощь эксперта.

Скорость расчета для данного модифицированного алгоритма представлена на рисунке (2.7)

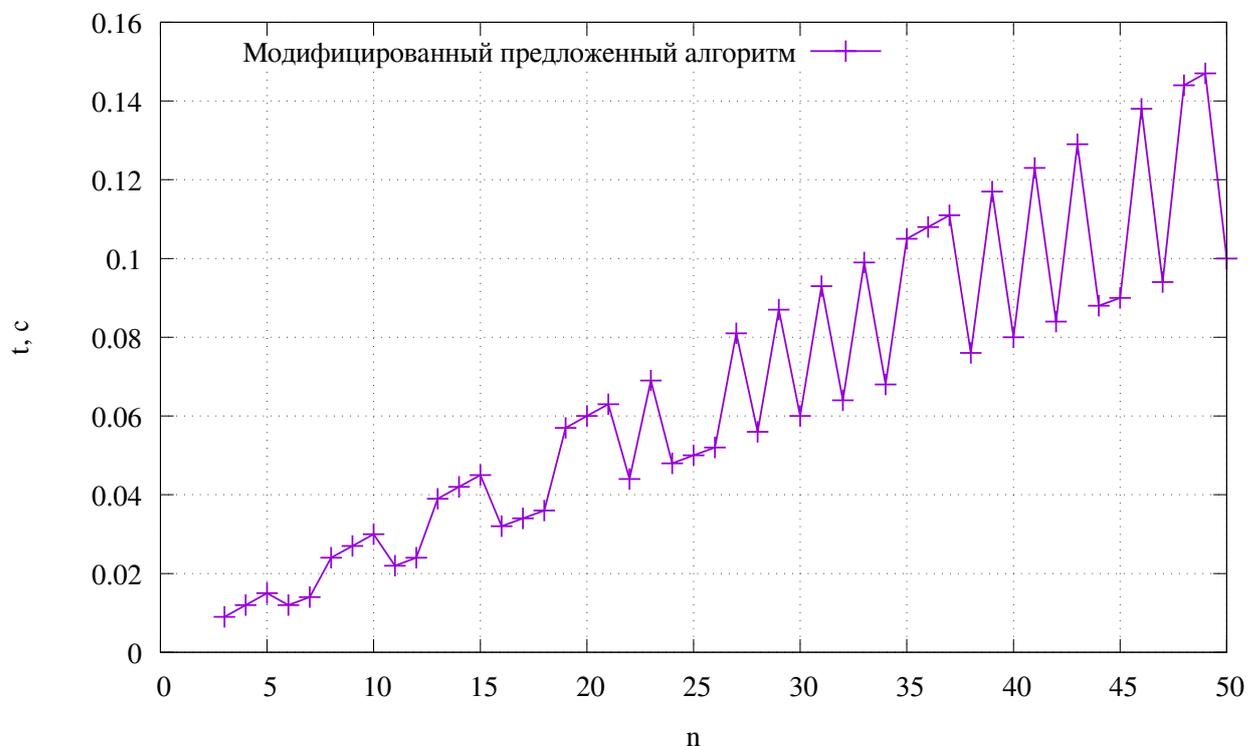


Рисунок 2.7 — Зависимость количества переменных и времени, затраченного на решение задачи выделения существенных переменных при помощи модифицируемого алгоритма, основанном на настройке \vec{c}_s

Как видно на рисунке (2.7), время, затраченное на решение задачи выделения существенных переменных сильно уменьшилось, что очень важно для задач с большим количеством входных переменных.

Отметим, что в задачах идентификации может возникнуть ситуация, когда $\sigma(\vec{u}_1) \approx 0, \sigma(\vec{u}_2) \approx 0, \vec{u}_1 \neq \vec{u}_2$, и даже $|\vec{u}_1| \neq |\vec{u}_2|$. Иными задачами задача моделирования может одинаково успешно решаться при помощи различных компонент вектора входных переменных \vec{u} . Например, зависимость между входными переменными подразумевает, что при моделировании мо-

жет быть использована любая из зависимых переменных, и не использоваться остальные зависимые переменные.

2.3 Нахождение параметрической модели методом перебора

Во многих учебниках, диссертациях и статьях, посвященных вопросам математического моделирования стохастических процессов, можно часто встретить фразу: «Пусть объект описывается следующим образом...». Далее дана формула, описывающая объект с точностью до параметров. В дальнейшем эти параметры определяются при помощи различных алгоритмов (один из самых известных - Метод наименьших квадратов).

За этой фразой скрывается огромная задача по нахождению параметрической структуры исследуемого процесса, которую исследователю необходимо решить. От того насколько точно подобрана параметрическая структура исследуемого процесса зависит точность построенной модели. Поэтому актуальность задачи нахождения параметрической структуры не вызывает сомнений.

В некоторых случаях параметрическую структуру можно восстановить используя фундаментальные законы физики, химии, астрономии или других наук. Данный способ хорош тем, что полученная структура будет точно отображать исследуемый процесс и, таким образом, исследователи получат наилучшую математическую модель исследуемого процесса.

К сожалению подобный способ, несмотря на его явные преимущества, редко используется. Обусловлено это тем что большинство исследуемых процессов весьма сложны и восстановить параметрическую структуру подобных процессов, используя фундаментальные законы, оказывается невыполнимой или слишком затрудненной задачей.

Вышеописанные сложности заставили исследователей подбирать пара-

метрическую структуру иными способами. Было замечено, что необходимость в нахождении именно той параметрической структуры исследуемого процесса которая «заложена природой» часто отсутствует. Достаточно, чтобы найденная параметрическая структура в достаточной мере (мера определяется исследователем) отражала исследуемый процесс в области определения входных воздействий $u_i \in (a_i \dots b_i), i = (\overline{1, n})$, где a_i и b_i - начальная и конечная граница области определения входного воздействия u_i , n — размерность вектора входных переменных \vec{u} . В технологических процессах области определения входных воздействий задаются технологическим регламентом.

В качестве примера возьмем процесс который задается формулой $f(u) = 5 * \sin(u/3) + u^2 + \xi$, с областью определения входных переменных $u \in [0; 3]$. Исследователь не обязан находить именно ту формулу, которой описывается процесс, он может использовать и другие, которые приближенно отражают данный процесс, на заданном промежутке, например линейная зависимость: $f_2(u) = 3,65 * u + 0.1$, отражает объект на промежутке $u \in (\overline{0, 3})$ (Рисунок 2.8).

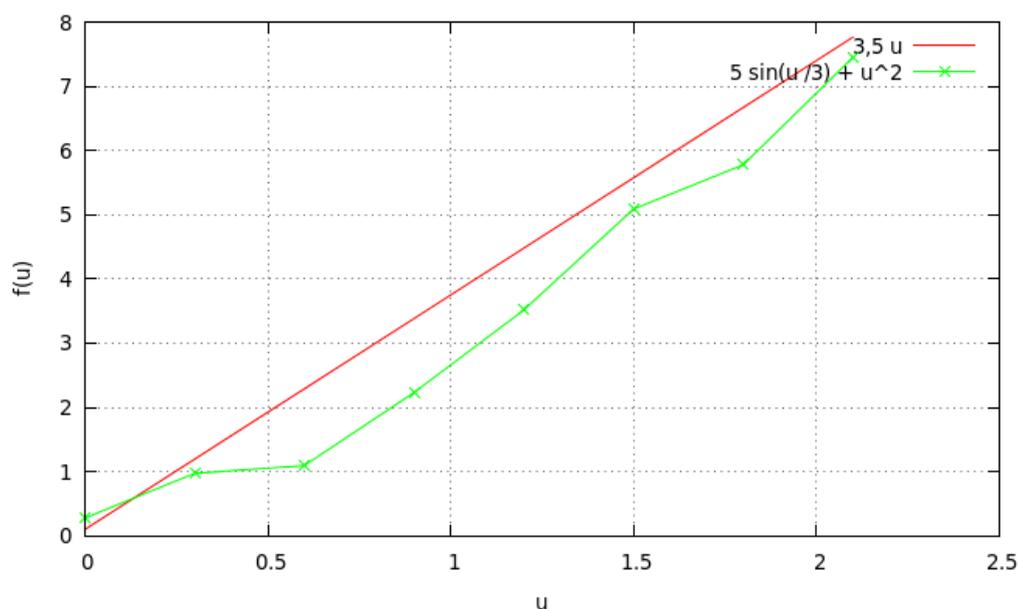


Рисунок 2.8 — Упрощенная параметрическая структура сложного объекта

Как видно на рисунке 2.8, предложенная модель достаточно точно описывает исследуемый процесс, в области определения входных переменных, хотя "истинный" параметрический вид процесса сильно отличается от подобранной структуры.

Одним из первых способов нахождения параметрической структуры является обыкновенный ручной подбор, который заключается в следующем. Исследователь или группа исследователей выбирает критерий $P(f(\vec{u}), \hat{f}(\vec{u}))$ который минимизируется при нахождении параметрической структуры. Таким критерием может быть, например, среднеквадратичная ошибка или средняя ошибка по модулю.

После того как выбран оптимизируемый критерий, решается задача оптимизации $P(f(\vec{u}), \hat{f}(\vec{u})) \rightarrow \min_{\hat{f}(\vec{u})}$. Исследователи находят модель $\hat{f}(\vec{u})$ у которой $P(f(\vec{u}), \hat{f}(\vec{u}))$ - наименьший, а потом добавляют в эту модель новые переменные, подбирают параметры, в конечном итоге минимизируя критерий $P(f(\vec{u}), \hat{f}(\vec{u}))$. В случае если модель была подобрана неточно, что исследователи узнают после того как структура будет апробирована на новых реальных данных, исследователи начинают подбор заново.

Подобный подход требует много времени у исследователей, более того при моделировании сложного технологического процесса подобрать структуру может быть — невыполнимой задачей.

2.4 Нахождение параметрической модели при помощи генетического программирования

Алгоритм генетического программирования основан на одноименном алгоритме оптимизации.

Данный алгоритм применяется, когда знания об исследуемом объекте соответствуют 2 или 3 уровню априорной информации, иными словами, когда

не известно каким уравнением описывается исследуемый процесс, но имеется выборка наблюдений $(\vec{u}_i, \vec{x}_i), i = (\overline{1, s})$, где \vec{u}_i — вектор входных переменных, \vec{x}_i — вектор выходных переменных, s — количество элементов в выборке.

Генетическое программирование позволяет на основании имеющейся выборки наблюдений восстановить параметрическое выражение, описывающее исследуемый процесс.

Процесс нахождения параметрической структуры процесса (параметрического выражения) при помощи генетического программирования является итерационным. В качестве критерия оптимальности могут быть выбраны такие характеристики как: средняя ошибка, среднеквадратичная ошибка, средняя относительная ошибка и иные характеристики, отражающие — насколько точно найденная модель описывает исследуемый процесс.

Процесс генетического программирования делится на следующие этапы:

- 1) случайным образом создаются первые генетические деревья;
- 2) получение потомков от деревьев. Получение потомков, может производиться при помощи следующих действий:
 - 2.1) репликация;
 - 2.2) размножение;
 - 2.3) мутация.
- 3) уничтожаются деревья, наименее удовлетворяющие условиям оптимума;
- 4) проверка удовлетворению заданных условий.

Заметно, что процесс генетического программирования сильно схож с генетическим алгоритмом.

Опишем некоторые центральные понятия, связанные с генетическим программированием:

В качестве описания структуры объекта в генетическом программи-

ровании используется понятие «Генетическое дерево». Генетическое дерево представляет собой связанный граф где узлы олицетворяют действия, а конечные элементы (листья) — переменные.

В качестве примера приведем генетическое дерево следующего выражения: $f(x_1, x_2) = x_1(x_2 - 7)(2/\sin(x_1))$ (Рисунок 2.9)

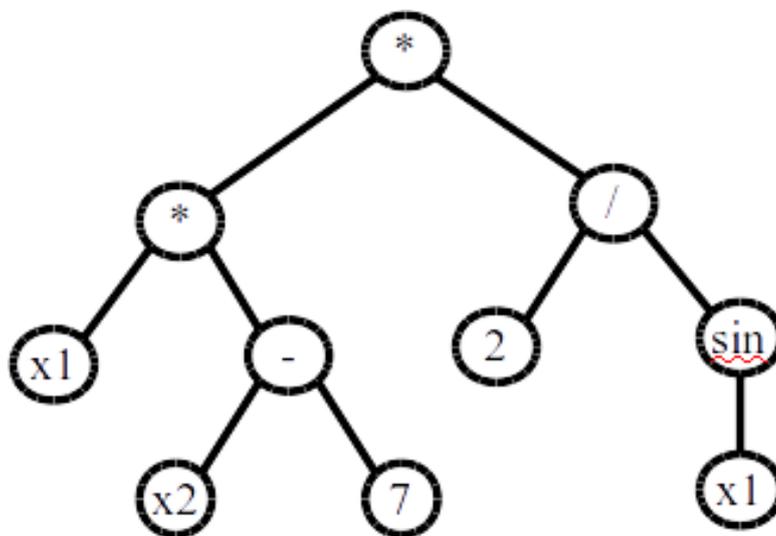


Рисунок 2.9 — Пример генетического дерева

Восстановление искомого выражения из представленного дерева начинается с листьев.

Также важно описать такие термины как: репликация, размножение и мутация.

Репликация — действие над деревом в результате которого получается точная копия дерева-родителя (Рисунок 2.10).

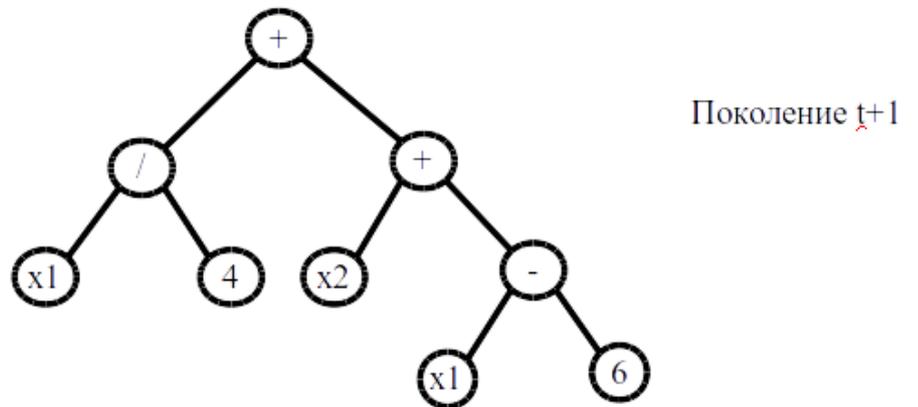
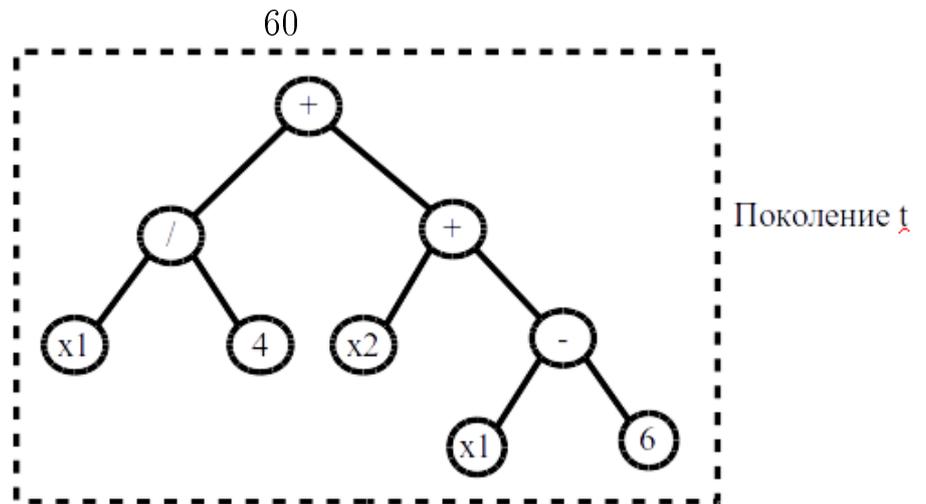


Рисунок 2.10 — Пример репликации

Размножение - получение потомков путем замены местами случайно выбранных ветвей родителей. (Рисунок 2.11).

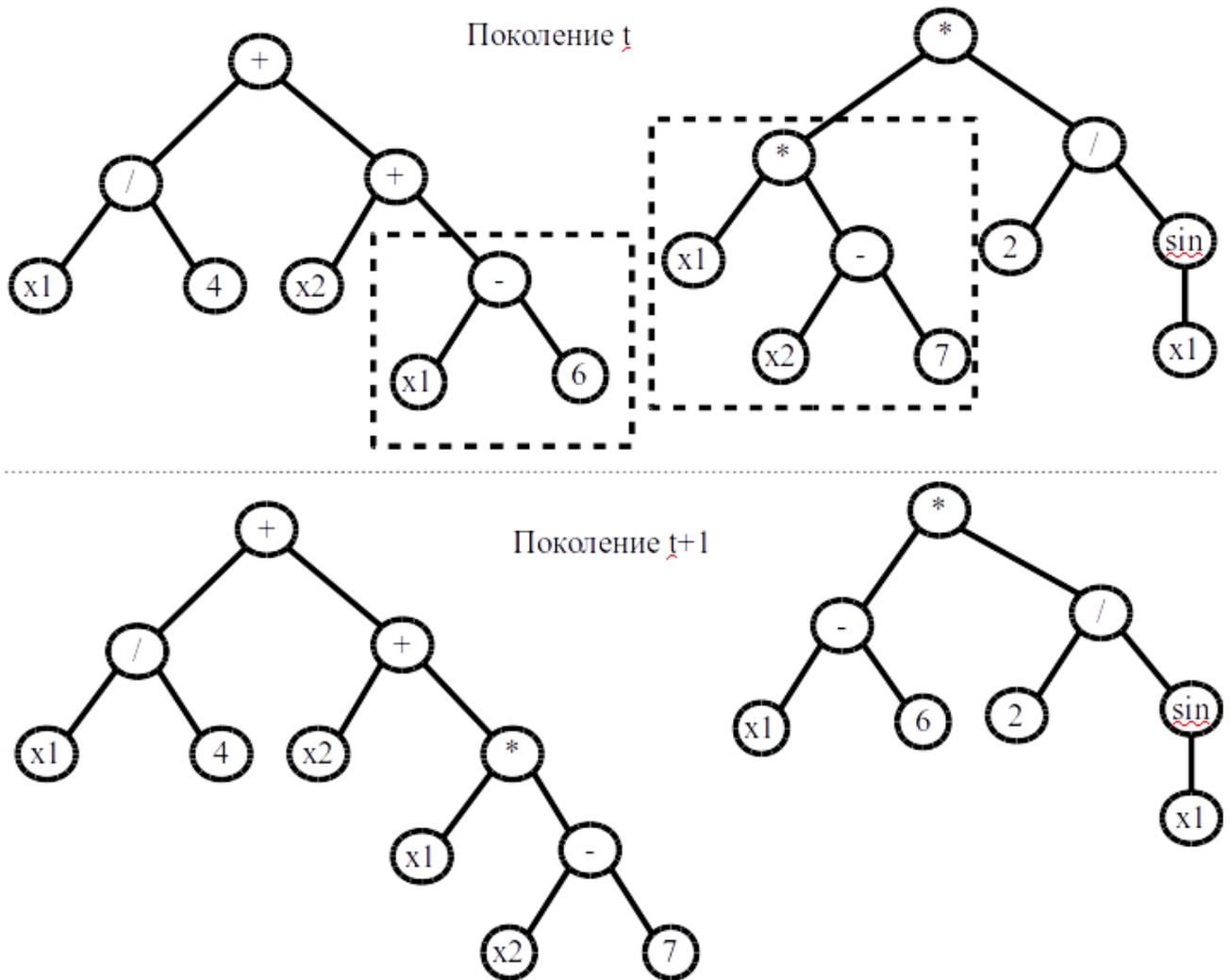


Рисунок 2.11 — Пример размножения

Мутация — получение потомка, путем изменение ветви дерева-родителя на случайно-созданную (Рисунок 2.12).

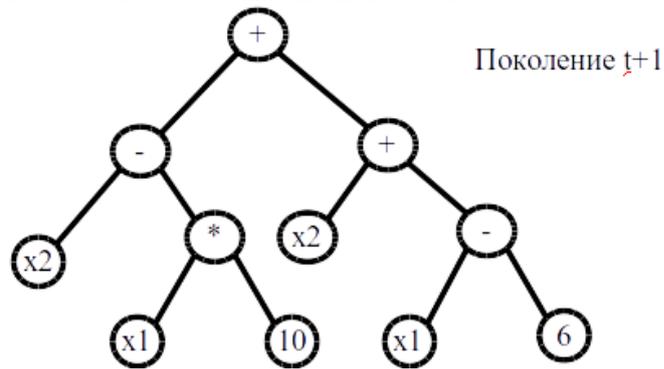
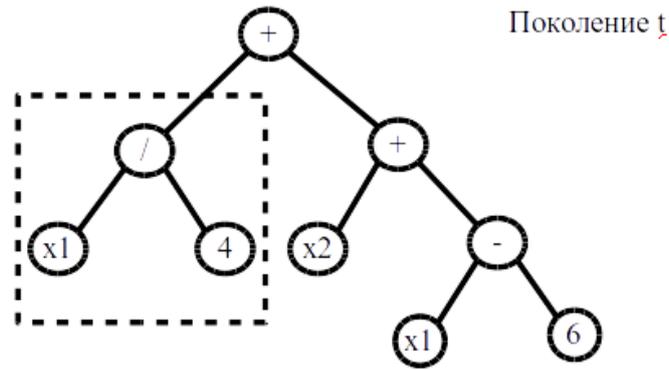


Рисунок 2.12 — Пример мутации

Как писалось ранее, последний этап в генетическом программировании — это проверка удовлетворению заданных условий.

Если проверка подтверждает что заданные условия были достигнуты, то исследователь использует для расчетов полученное дерево, и при помощи его моделирует процесс. Если заданные условия не были достигнуты, то процесс генетического программирования повторяется, начиная с этапа 2.

Стоит сказать что алгоритм генетического программирования во многом повторяет описанный в предыдущей главе метод подбора, отличием является то, что он легко автоматизируется.

2.5 Модели основанные на локальной аппроксимации

В случае когда исследователь не располагает информацией о параметрической структуре моделируемого процесса, имеет смысл использовать ме-

тоды моделирования, которые свободны от этапа выбора параметрической структуры.

Одним из таких методов является локальная аппроксимация.

Основная идея локальной аппроксимации проста. Работу этого метода можно представить себе следующим образом: вокруг аппроксимирующей точки строится прямоугольник, каждому элементу выборки внутри данного прямоугольника назначается вес (чем больше евклидово расстояние между элементом выборки и точки построения прогноза, тем меньше вес), а затем происходит усреднение взвешенных значений.

В случае если при локальной аппроксимации используется регрессия нулевого порядка, то взвешенные значения последовательности, которые попали в прямоугольник, просто усредняются.

Обратимся к выражению, определяющему процедуру оценки функции в точке \vec{u} [6].

$$\hat{f}(\vec{u}) = \frac{\sum_{i=1}^s f_i \prod_{j=1}^n \Phi((u_j - u_{ji})/c_{sj})}{\sum_{i=1}^s \prod_{j=1}^n \Phi((u_j - u_{ji})/c_{sj})}, \quad (2.5.1)$$

где: f — последовательность длиной s ; \vec{c}_s — вектор параметров размытости; Φ — колокообразная функция.

Колокообразная функция $\Phi(*)$ — функция, удовлетворяющая следующим условиям (2.5.2) и (2.5.3):

$$\frac{1}{c_s} \int_{-\infty}^{+\infty} \Phi\left(\frac{u - u_i}{c_s}\right) du = 1, \quad (2.5.2)$$

$$\lim_{c_s \rightarrow 0} \frac{1}{c_s} \int_{-\infty}^{+\infty} \varphi(u) \Phi\left(\frac{u - u_i}{c_s}\right) du = \varphi(u_i), \quad (2.5.3)$$

где u любая произвольная компонента вектора \vec{u} , u_i — элемент выборки, $i = (\overline{1, s})$, $\varphi(u)$ — произвольная функция.

Как следует из выражения (2.5.1), плотность в точке \vec{u} вычисляется как сумма значений колокообразных функций для величин, определяемых разностями между значением \vec{u} и значениями последовательности. При этом точки \vec{u} , в которых вычисляется плотность, могут и не совпадать со значениями элементов выборки.

Существует несколько наиболее часто используемых колокообразных функций. Приведем несколько из них:

Квадратичная функция:

$$\Phi\left(\frac{u - u_i}{c_s}\right) = \begin{cases} \left(\frac{15}{16}\right)\left(1 - \left(\frac{u-u_i}{c_s}\right)^2\right)^2 & \left|\frac{u-u_i}{c_s}\right| \leq 1 \\ 0 & \left|\frac{u-u_i}{c_s}\right| > 1 \end{cases} \quad (2.5.4)$$

Функция Епанечникова:

$$\Phi\left(\frac{u - u_i}{c_s}\right) = \begin{cases} 0.75\left(1 - \left(\frac{u-u_i}{c_s}\right)^2\right) & \left|\frac{u-u_i}{c_s}\right| \leq 1 \\ 0 & \left|\frac{u-u_i}{c_s}\right| > 1 \end{cases} \quad (2.5.5)$$

Функция Гаусса:

$$\Phi\left(\frac{u - u_i}{c_s}\right) = (2\pi)^{(-1/2)} \exp\left(-\left(\frac{u - u_i}{c_s}\right)^2 / 2\right) \quad (2.5.6)$$

Треугольная функция:

$$\Phi\left(\frac{u - u_i}{c_s}\right) = \begin{cases} \left(1 - \left|\frac{u-u_i}{c_s}\right|\right) & \left|\frac{u-u_i}{c_s}\right| \leq 1 \\ 0 & \left|\frac{u-u_i}{c_s}\right| > 1 \end{cases} \quad (2.5.7)$$

Прямоугольная функция:

$$\Phi\left(\frac{u - u_i}{c_s}\right) = \begin{cases} 1/2 & \left|\frac{u-u_i}{c_s}\right| \leq 1 \\ 0 & \left|\frac{u-u_i}{c_s}\right| > 1 \end{cases} \quad (2.5.8)$$

Для большинства задач вид колокообразной функции не имеет большого значения. Локальная аппроксимация при различных колокообразных функциях, дает близкие по значению прогнозы. Гораздо более важным является настройка вектора параметров размытости \vec{c}_s .

2.5.1 Настройка вектора параметров размытости

От выбранного вектора параметров размытости напрямую зависит качество построенной модели. Вектор параметров размытости определяет степень участия элементов выборки наблюдений в построении прогноза (Рисунок 2.13)

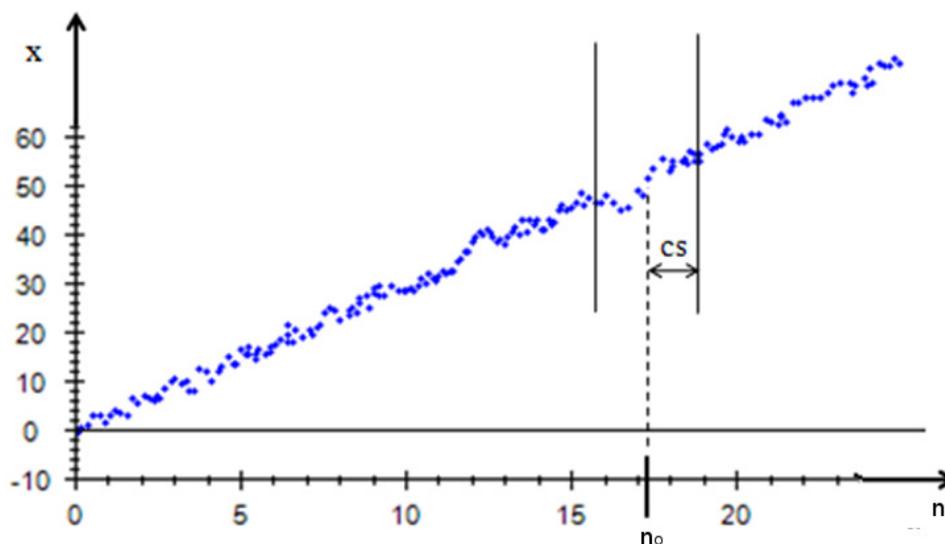


Рисунок 2.13 — Определение параметра размытости

Как видно на рисунке 2.13, участвуют в оценке в точке n_o только те элементы выборки, у которых $|n - n_o| < c_s$.

Исходя из большого влияния вектора \vec{c}_s на качество построенной модели, у исследователей часто возникает задача настройки данного вектора, то есть необходимо решить задачу оптимизации (в качестве критерия оптими-

зации имеется ввиду точность построенной модели).

В настоящее время существуют множество методов оптимизации как для нахождения локальных экстремумов, так и для нахождения глобальных (хотя, стоит сказать, что методы для нахождения глобальных экстремумов - вероятностные).

Для того чтобы понять, какой из методов оптимизации предпочтительнее, необходимо изучить оптимизируемую величину. Для этого построим график зависимости между \vec{c}_s и ошибкой моделирования. Ошибка моделирования будет рассчитываться как среднеквадратичное отклонение.

$$\sigma(\vec{c}_s) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}(\vec{c}_s))^2} \quad (2.5.9)$$

Были проведены несколько вычислительных экспериментов, в которых строились графики зависимости между \vec{c}_s и ошибкой моделирования но они все были близкого вида. Приведем пример одного из таких графиков (Рисунок 2.14).

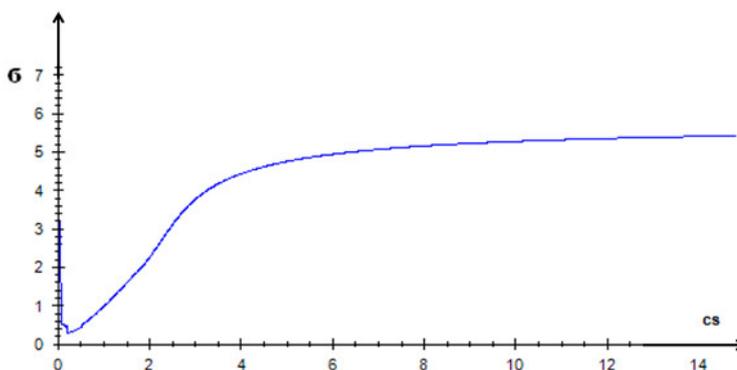


Рисунок 2.14 — Зависимость между c_s и ошибкой моделирования

Исходя из графика 2.14, можно предположить, что при среднеквадратичном отклонении в качестве критерия оптимизации, существует единственный глобальный минимум функции $\sigma(\vec{c}_s)$.

Учитывая, что количество компонент вектора \vec{c}_s равно количеству входных переменных n , при больших значениях n , решение задачи оптимизации будет представлять существенные сложности. В этой связи при моделировании процессов с большим количеством входных переменных логично уменьшить размер вектора \vec{c}_s . Для этого необходимо связать каждый компонент вектора \vec{c}_s с несколькими входными переменными. Например, предположим что у моделируемого процесса 50 входных переменных $n = 50$. Вместо того чтобы вводить вектор \vec{c}_s размером 50, предлагается ввести вектор \vec{c}_s размером, например, 5, чтобы каждой компоненте вектора \vec{c}_s соответствовало 10 компонент вектора входных переменных \vec{u} . Это позволит уменьшить размерность задачи оптимизации с 50 до 5.

Для того чтобы убедиться, что уменьшение размера вектора \vec{c}_s не приведет к ухудшению качества модели, были проведены соответствующие вычислительные эксперименты. В экспериментах решалась задача оптимизации в двух случаях: когда размер вектора \vec{c}_s равен n и когда размер вектора \vec{c}_s равен 1. Сравнивалась скорость решения задачи оптимизации и точность построенной модели. Вычислительные эксперименты проводились над различными процессами.

Опишем пример вычислительного эксперимента.

Моделируемый процесс имеет два входных воздействия и одну выходную переменную. Объем выборки наблюдений - 300. Помеха, воздействующая на объект, равна 3%. Критерием оптимизации выбрана среднеквадратичная ошибка.

Для решения задачи оптимизации использовались следующие методы оптимизации: метод деформируемого многогранника, генетический алгоритм, перебор возможных значений (осуществляется по сетке в диапазоне $(0,3]$ с поисковым шагом 0.01).

Результаты решения задачи оптимизации представлены в таблице 1

Таблица 1 — Результаты настройки вектора параметров размытости

Метод оптимизации	Размерность вектора \vec{c}_s	Время, необходимое для нахождения оптимального вектора \vec{c}_s , мс	Средне-квадратичная ошибка, σ
МДМ	2	1 118	0,75
МДМ	1	100	0,76
ПВЗ	2	26 808	0,78
ПВЗ	1	934	0,78
ГА	2	3 906	0,76
ГА	1	494	0,76

В таблице 1 введены следующие обозначения: МДМ - метод деформируемого многогранника; ПВЗ - перебор возможных значений; ГА - Генетический алгоритм.

Как можно видеть из таблицы 1 уменьшение размерности вектора \vec{c}_s не привело к ухудшению качества модели.

Большинство других вычислительных экспериментов подтвердили этот вывод. Это обозначает что во многих задачах можно уменьшить размерность вектора \vec{c}_s без ущерба для качества модели.

Также стоит отметить что метод деформируемых многогранников гораздо быстрее решает задачу оптимизации, чем перебор значений или генетический алгоритм.

2.5.2 Улучшение качества обучающей выборки

При всех преимуществах моделирования при помощи локальной аппроксимации не стоит забывать о том что модель, построенная при помощи методов непараметрической статистики, имеет важный недостаток. Дело в том, что при построении подобных моделей точность этих моделей будет сильно коррелировать от качества обучающей выборки. Безусловно, каким бы методом не была построена модель объекта зависимость качества построенной модели от качества обучающей выборки будет иметь место, но при создании модели при помощи локальной аппроксимации эта зависимость имеет гораздо большее значение.

Исходя из описанной особенности, появилась задача улучшения качества обучающей выборки. Улучшить качество обучающей выборки можно разными способами, в диссертации рассмотрено избавление от выбросов.

Для начала стоит объяснить понятие выброса. Выброс – это такой элемент выборки наблюдений (\vec{u}_i, x_i) , в котором $x_i \notin \Omega(\vec{u}, x)$, где $\Omega(\vec{u}, x)$ - область протекания процесса. Иными словами, данный элемент не характеризует изучаемый процесс и, вследствие этого, мешает при моделировании процесса.

Чем больше выброс отдален от области протекания процесса, тем сильнее искажение которое выброс вносит в создаваемую модель.

Природа появления выбросов в обучающей выборке может быть различной: выброс может быть допущен во время измерения из-за технического сбоя, или человеческого фактора.

Предложенный способ выделения выбросов заключается в том чтобы для каждого элемента обучающей выборки исследователь строил прогноз при тех же входных переменных, что и у элемента (\vec{u}_i, x_i) (сам элемент (\vec{u}_i, x_i) исключается из процесса построения прогноза). Если полученный прогноз

отличается от выходной характеристики элемента обучающей выборки более чем на заданное значение, то данный элемент (\vec{u}_i, x_i) считается выбросом и его следует исключить из выборки наблюдений.

Представим данный алгоритм в математическом виде.

$\forall(\vec{u}_i, x_i), |\hat{x}(\vec{u}_i) - x_i| > q \rightarrow (\vec{u}_i, x_i)$ - выброс, где q - значение заданное исследователем; $\hat{x}(\vec{u}_i)$ - прогноз, построенный в точке \vec{u}_i .

$$\hat{x}(\vec{u}_i) = \frac{\sum_{j=0}^s x_j \prod_{l=1}^n \Phi\left(\frac{u_{il}-u_{jl}}{c_s}\right)}{\sum_{j=0}^s \prod_{l=1}^n \Phi\left(\frac{u_{il}-u_{jl}}{c_s}\right)}, \quad (2.5.10)$$

где $j = (\overline{1, s}), j \neq i$.

При помощи предложенного метода можно улучшить качество выборки наблюдений.

Представим пример работы данного метода на примере моделирования процесса с двумя входами и одним выходом.

Пусть процесс описывается следующей формулой:

$$f(u_1, u_2) = 6u_1 - 3u_2 + 4\sin\left(\frac{u_1 + u_2}{30}\right) \quad (2.5.11)$$

Стоит отметить, что исследователь не знает о структуре объекта, она используется только для генерации выборки.

Объем обучающей выборки равен 300, $u \in (\overline{0, 30})$, $\xi = 5\%$, выбросов 5% от размера обучающей выборки.

Для начала построим модель на основании выборки которая не прошла предварительной обработки и, как следствие, не избавленной от выбросов. Построенная модель представлена на рисунке 2.15. Моделирование проведено с использованием методов локальной аппроксимации.

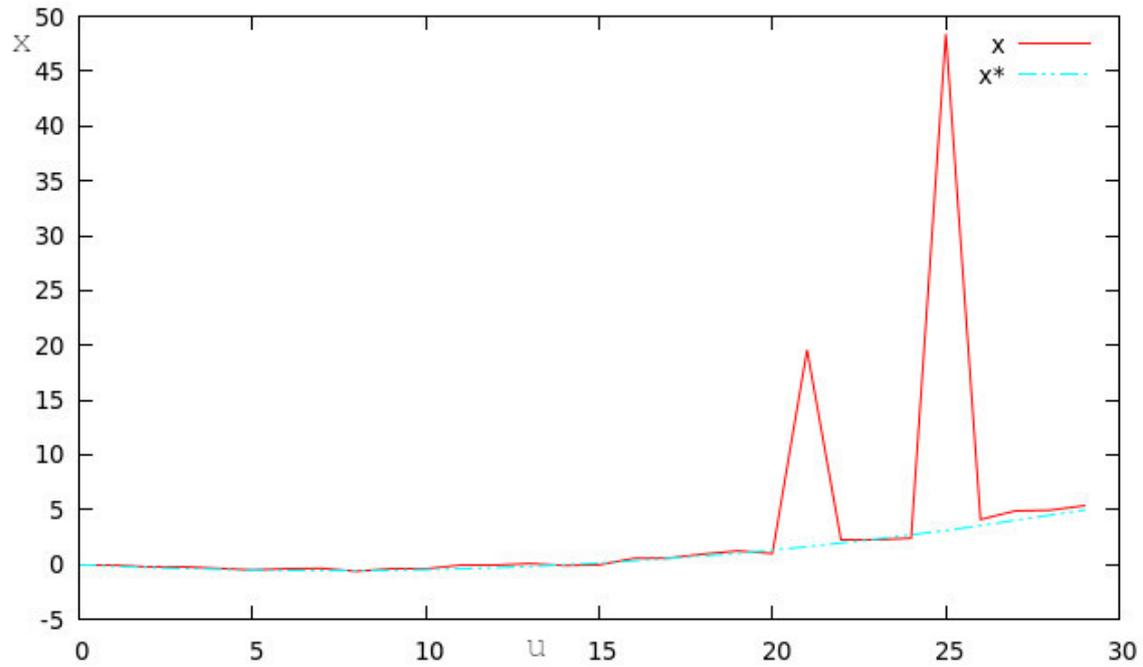


Рисунок 2.15 — Модель объекта, построенная на обучающей выборке с выбросами без предварительной обработки

На рисунке 2.15 x — выходная переменная процесса, x^* — выходная переменная модели процесса.

Среднеквадратичная ошибка данной модели $\sigma = 8.90$.

Теперь построим модель на основании обработанной выборки в которой выбросы были исключены. Построенная модель представлена на рисунке 2.16

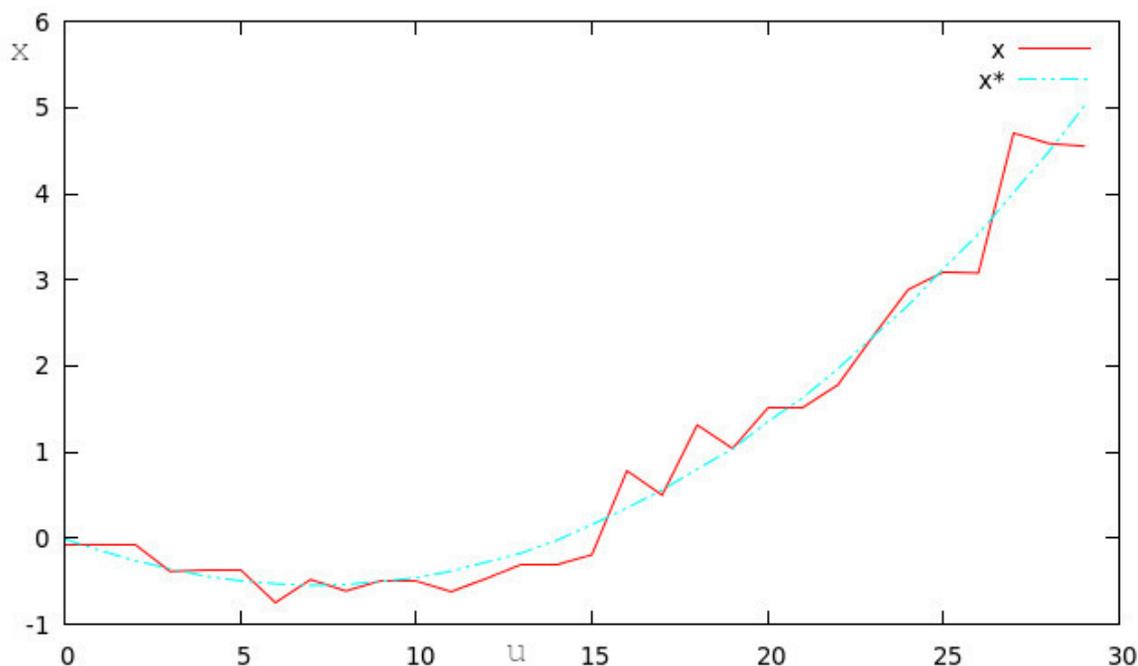


Рисунок 2.16 — Модель объекта, построенная на обучающей выборке с выбросами с предварительной обработки

Как видно на рисунке 2.16, модель исследуемого процесса получилась более точной в сравнении с моделью на рисунке 2.15.

Среднеквадратичная ошибка данной модели $\sigma = 0.25$, что на 8.65 лучше чем у модели, построенной на основании необработанной обучающей выборки.

На основании вышесказанного возможно заключить, что избавление от выбросов представленным способом, позволяет улучшить качество моделирования.

Логично предположить что существует зависимость между эффективностью представленного алгоритма и тем сколько выбросов было в выборке наблюдений. Построим график зависимости между количеством выбросов в выборке наблюдений и эффективностью представленного алгоритма, в случае моделирования вышеописанного процесса.

Зависимость между количеством выбросов в обучающей выборке и эффективностью предложенного алгоритма представлена на рисунке 2.17.

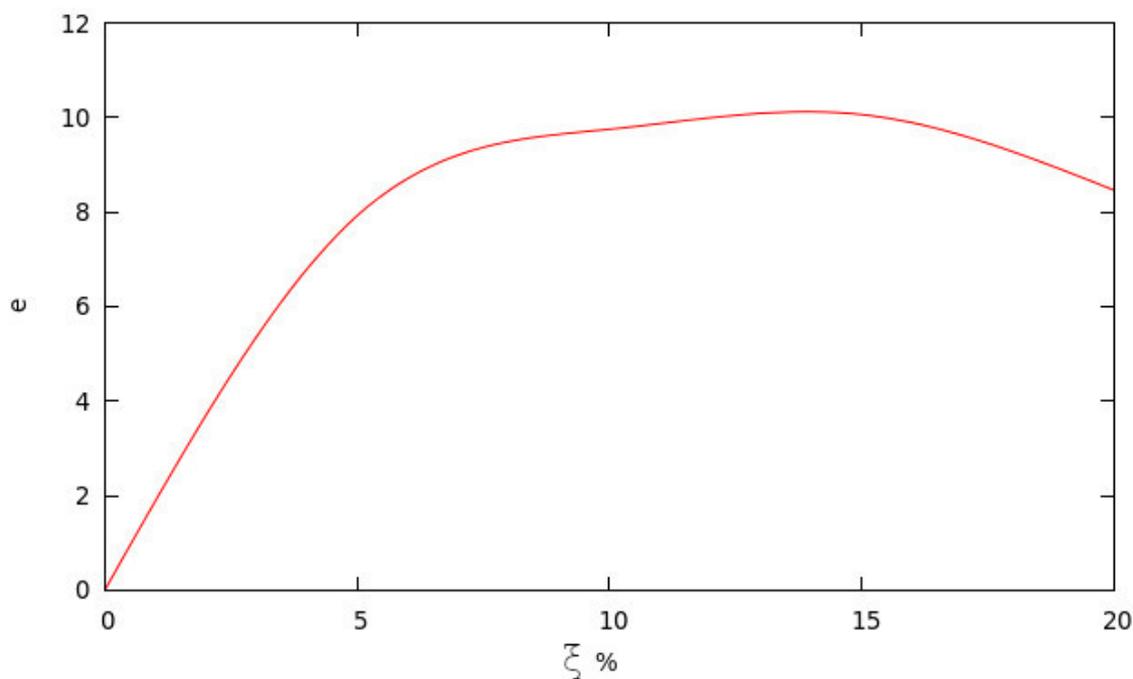


Рисунок 2.17 — Зависимость между количеством выбросов в выборке наблюдений и эффективностью предложенного алгоритма

На рисунке 2.17 приняты следующие обозначения: ξ — количество выбросов в выборке наблюдений (в %); e - разница между среднеквадратичной ошибки модели, построенной по выборке наблюдений без предварительной обработки, и среднеквадратичной ошибки модели, построенной по выборке наблюдений с предварительной обработкой.

Если посмотреть на рисунок 2.17 можно заметить что при количестве выбросов более 15% эффективность алгоритма начинает падать. Это объясняется тем, что количество выбросов становится настолько большим, что алгоритму сложнее определить - является ли величина выбросом или нет.

2.6 Н-модели

2.6.1 Общая информация об Н-моделях

Рассмотрим классическую схему идентификации (Рисунок 2.18)

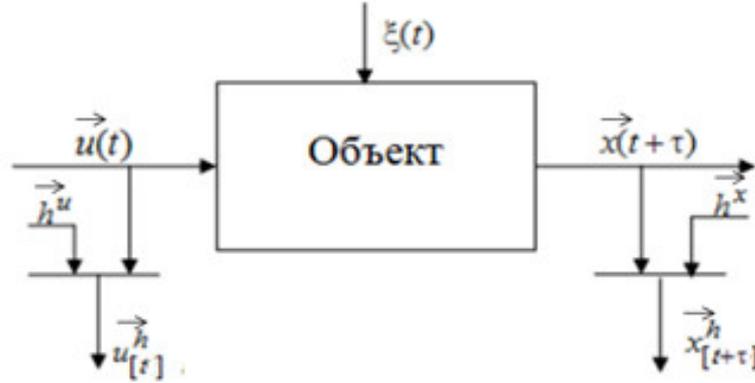


Рисунок 2.18 — Классическая схема идентификации

На рисунке 2.18 приняты следующие обозначения: $\vec{u}(t)$ – это вектор входных переменных в момент времени t ; $\vec{x}(t+\tau)$ – вектор выходных переменных в момент времени $(t+\tau)$; τ – запаздывание; $\xi(t)$ – помеха, действующая на объект; h^u, h^x – помехи в каналах связи $\vec{u} = (u_1, u_2, \dots, u_n) \in \Omega(\vec{u}) \subset R$, где $\Omega(\vec{u})$ – область определения (u_1, u_2, \dots, u_n) . Стоит сказать, что каждая компонента вектора $u_i \in [a_i, b_i]$, а $x_j \in [c_j, d_j], i = (1, \dots, n), j = (1, \dots, k)$, где n – количество входных переменных, k – количество выходных переменных. При исследовании реальных процессов значения коэффициентов a_i, b_i, c_j, d_j часто известны.

Если входные воздействия на объект стохастически – зависимы друг от друга, то процесс протекает не во всей области, определяемой векторами входных-выходных переменных $\Omega(\vec{x}, \vec{u})$, а в некой подобласти $\Omega^H(\vec{x}, \vec{u})$ [10]. Рассмотрим особенности, возникающие при моделировании подобных процессов.

Из соображений краткости, в дальнейшем, будем называть процессы с зависимыми компонентами вектора входных переменных — Н-процессами.

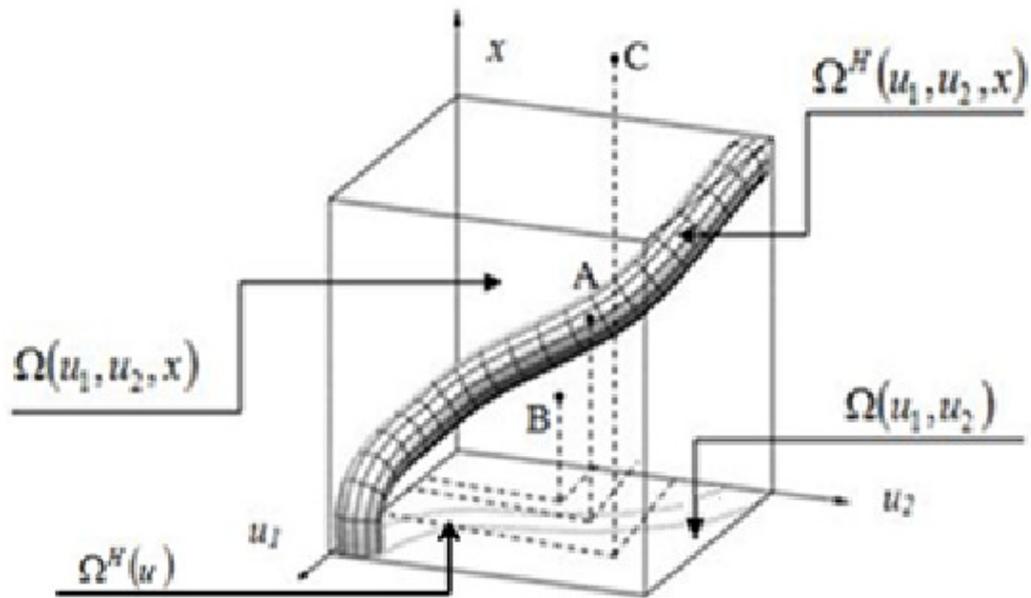


Рисунок 2.19 — Н - процесс

На рисунке 2.19 представлено схематичное изображение Н-процесса. На рисунке 2.19 точка A при входных значениях $\vec{u} = (u_1, u_2, \dots, u_n) \in \Omega(\vec{u})$ соответствует величине выходной переменной, из области $\Omega^H(u_1, u_2, x)$. Таким образом при $\vec{u} = (u_1, u_2, \dots, u_n) \in \Omega^H(\vec{u})$ объект имеет значения выхода из области $\Omega^H(u_1, u_2, x)$. Точка B – точка которая при входных значениях $\vec{u} = (u_1, u_2, \dots, u_n) \in \Omega(\vec{u})$, но не принадлежит $\Omega^H(u_1, u_2, x)$, хотя при этом – все-еще находится в пределах $\Omega(u_1, u_2, x)$. Это происходит потому что $\vec{u} = (u_1, u_2, \dots, u_n) \notin \Omega^H(\vec{u})$. При описанных случаях процесс не выходит за рамки $\Omega(u_1, u_2, x)$. А если посмотреть на точку C , то здесь видно, что хотя $\vec{u} = (u_1, u_2, \dots, u_k) \in \Omega(\vec{u})$, но $\vec{x} \notin \Omega(u_1, u_2, x)$.

Особенность при работе с Н-процессами заключается в следующем: когда будет предпринята попытка построить параметрическую модель для этих процессов, не учитывая зависимости между компонентами вектора входных

переменных \vec{u} , то будут использоваться только те точки, которые принадлежат $\Omega^H(\vec{u}, x)$. В частном случае, если представить себе прямую в пространстве, то моделей такого процесса в виде плоскости может быть сколько угодно много, поскольку через прямую в пространстве можно провести бесконечное число плоскостей, что иллюстрирует рисунок 2.20. В случае если зависимость носит нелинейный характер, то по различным исходным данным, восстанавливаемых поверхностей может быть также достаточно много.

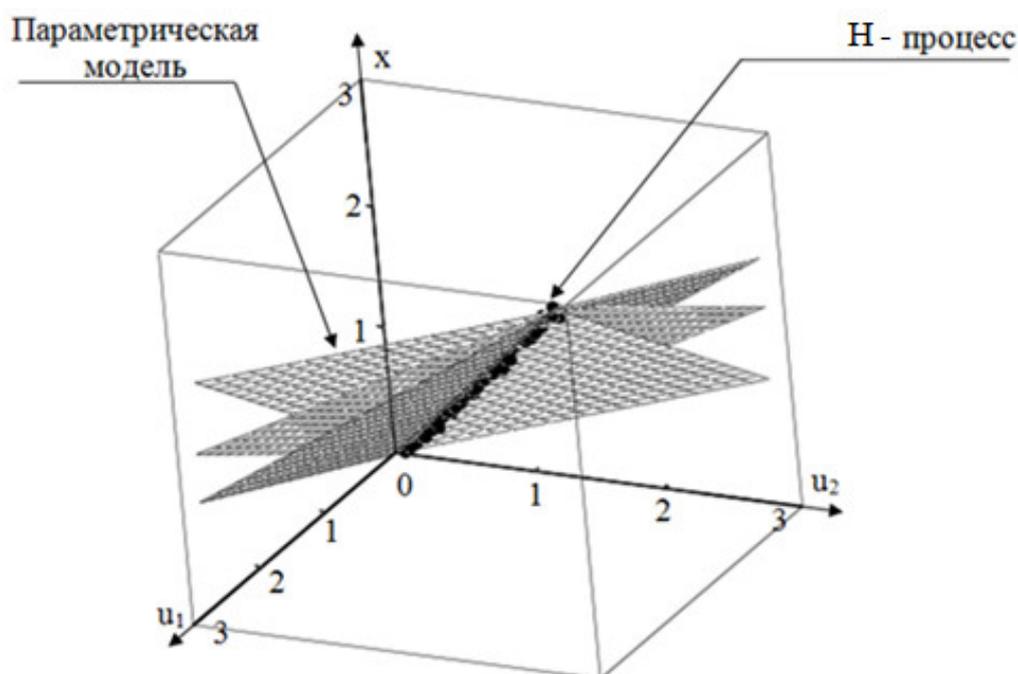


Рисунок 2.20 — Параметрические модели трубчатых процессов (H – процесс)

Также при попытке прогнозирования H – процесса могут возникнуть сложности, представленные на рисунке 2.21.

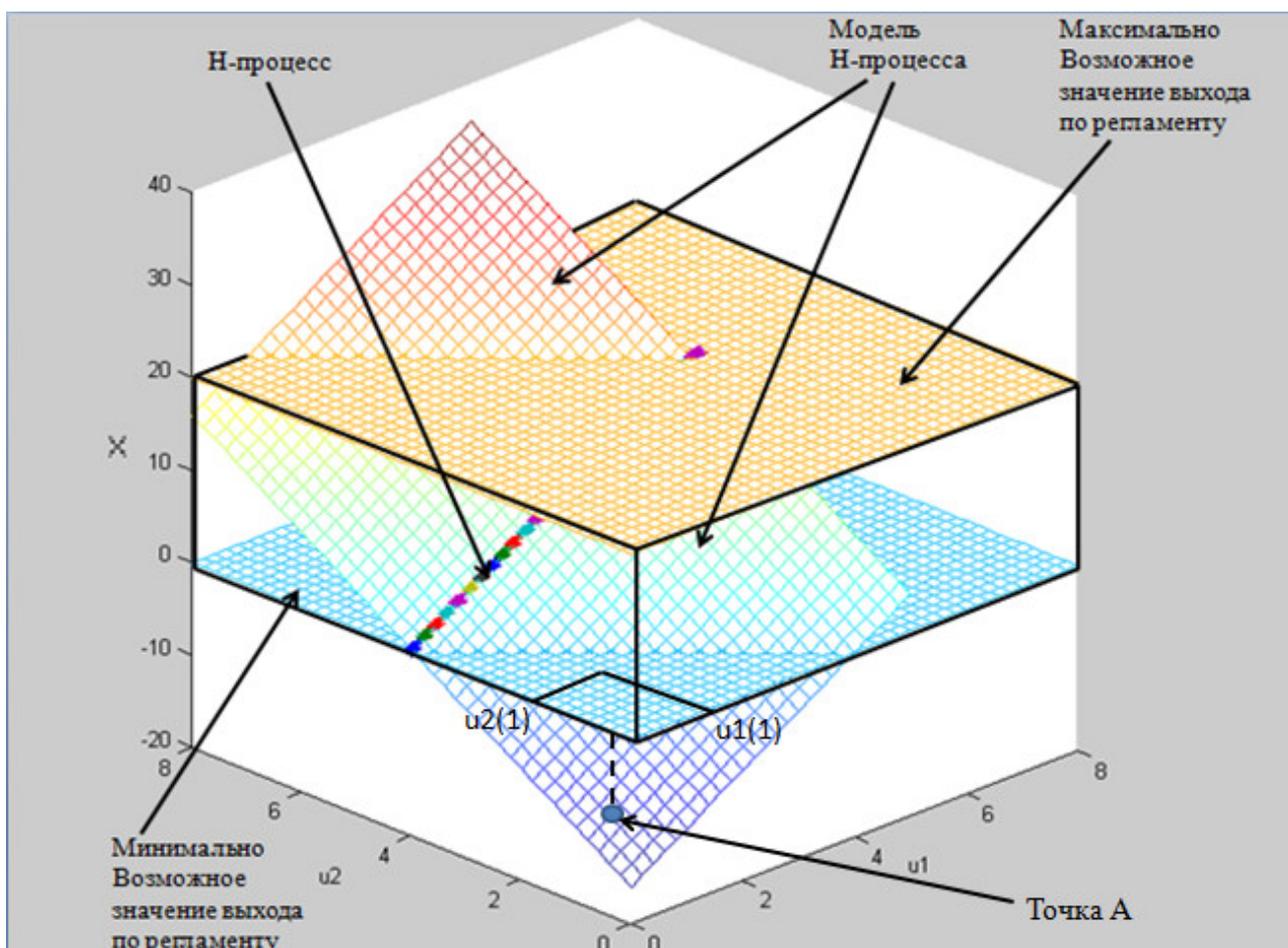


Рисунок 2.21 — Построение прогноза H-процессов

В качестве примера можно привести точку А. В этой точке u_1 и $u_2 \in \Omega(\vec{u})$, но, не принадлежат $\Omega^H(\vec{u})$, значит если u_1 или $u_2 \notin \Omega^H(\vec{u})$, то используя параметрическую модель, мы можем получить прогнозы - физически нереализуемые.

Для того чтобы построенная модель могла реагировать на вышеописанную проблему её дополняют индикаторной функцией $I(\vec{u})$. Индикаторная функция $I(\vec{u})$ - это функция, которая позволяет определить принадлежит ли вектор входных воздействий \vec{u} области определения $\Omega^H(\vec{u})$.

Индикаторная функция $I(\vec{u})$ имеет следующий вид:

$$I(\vec{u}) = \begin{cases} 1 & , \vec{u} \in \Omega^H(\vec{u}) \\ 0 & , \vec{u} \notin \Omega^H(\vec{u}) \end{cases}. \quad (2.6.1)$$

В качестве оценки индикаторной функции $I_s(\vec{u})$ можно принять следующее приближение:

$$I_s(\vec{u}) = \text{sgn} \sum_{i=1}^s \prod_{j=1}^n \Phi(c_s^{-1}(u^j - u_i^j)), \quad (2.6.2)$$

где

$$\Phi(c_s^{-1}(u^j - u_i^j)) = \begin{cases} (1 - |c_s^{-1}(u^j - u_i^j)|) & , |c_s^{-1}(u^j - u_i^j)| \leq 1 \\ 0 & , |c_s^{-1}(u^j - u_i^j)| > 1 \end{cases} \quad (2.6.3)$$

При использовании индикаторной функции, стандартную модель (2.6.4) можно изменить следующим образом (2.6.5)

$$\hat{x} = \hat{f}(\vec{u}, \alpha_s) \quad (2.6.4)$$

$$\hat{x} = \hat{f}(\vec{u}, \alpha_s) I(\vec{u}) \quad (2.6.5)$$

Подчеркнем еще раз, что область $\Omega^H(\vec{u})$ не известна, а известна только выборка наблюдений $(x_i, \vec{u}_i)_{i = \overline{1, s}}$. Если индикаторная функция равна нулю, то оценка \hat{x} не может быть вычислена, т.е. при таких значениях компонент вектора $\vec{u} \in \Omega(\vec{u})$ процесс протекать не может. Если индикаторная функция $I(\vec{u})$ при любом значении $\vec{u} \in \Omega(\vec{u})$ равна единице (моделируемый процесс не является Н-процессом), то модель 2.6.5 совпадает со стандартной 2.6.4.

Продемонстрируем работу индикаторной функции, сравнив ее с построенной модели при помощи алгоритма МНК.

Для начала отобразим обучающую выборку (Рисунок 2.22):

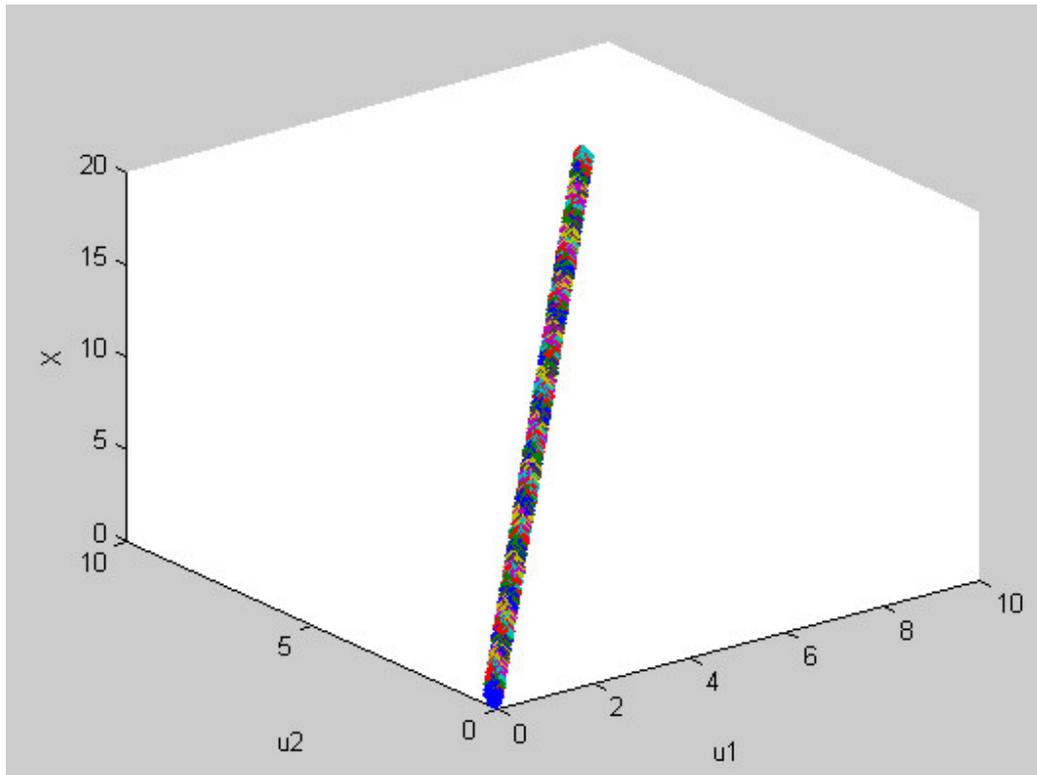


Рисунок 2.22 — Обучающая выборка

Как видно на рисунке 2.22, данная выборка действительно является выборкой из N -процесса. Были добавлены помехи для расширения области выборки и еще большей визуализации сложностей, связанных с моделированием. Теперь, посмотрим модель, построенную при помощи подбора структуры и расчета параметров при помощи метода МНК (Рисунок 2.23).

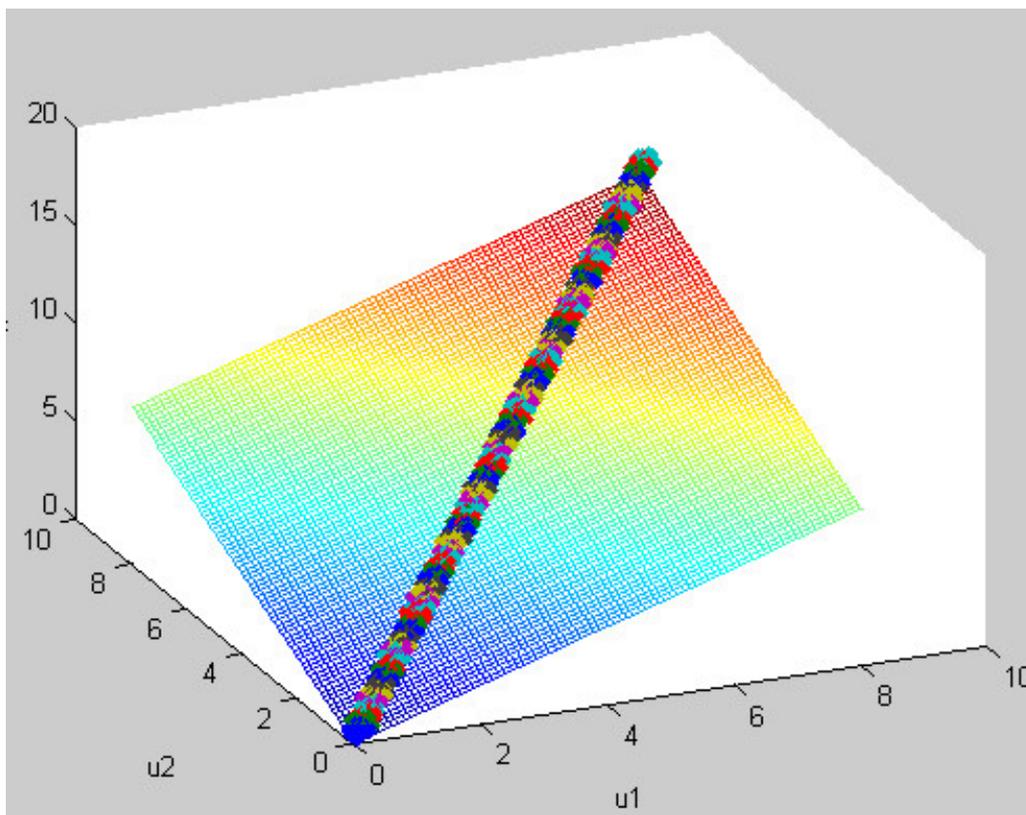


Рисунок 2.23 — Модель, построенная при помощи подбора структуры и расчета параметров при помощи метода МНК

Как видно на рисунке 2.23, была получена модель Н-процесса. Данная модель не выходит за рамки регламента хоть и не отражает истинную природу процесса. Заметить что данная модель может работать некорректно крайне сложно, так как все прогнозы, которые она будет давать – физически реализуемы. Хотя на практике, прогнозы, почти наверняка, будут неверны (в точках, которые находятся за областью протекания процесса).

Теперь построим модель Н – процесса, с использованием индикаторной функции (Рисунок 2.24).

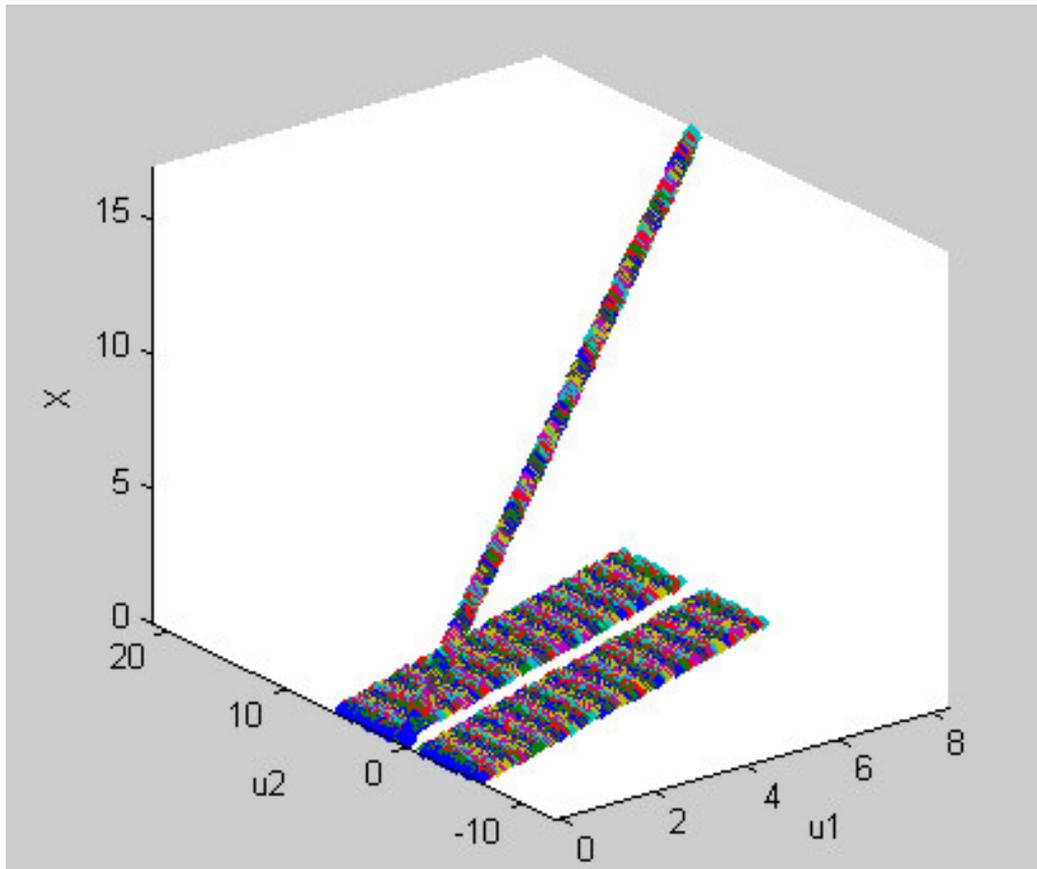


Рисунок 2.24 — Модель, построенная с использованием индикаторной функции

Как видно на рисунке 2.24, модель Н-процесса, построенная использованием индикаторной функции, ближе к действительной природе моделируемого процесса в отличие от модели, представленной ранее. Здесь все точки, не входящие область $\Omega^H(\vec{u})$ "зануляются отражая что моделируемый процесс не протекает вне области $\Omega^H(\vec{u})$.

Таким образом было продемонстрировано что для Н-процесса можно построить достаточно точную модель при помощи методов непараметрической статистики.

2.6.2 Модель многомерного процесса с общей зоной определения входных переменных

H -процессы, описанные выше, имеют одну выходную переменную. Безусловно, в реальных технологических процессах подобная ситуация нереалистична. Чаще всего технологический процесс характеризуется множеством выходных переменных $x_j, j = (\overline{1, k})$, где k - количество выходных переменных.

Логично предположить что некоторые из этих выходных переменных могут протекать не во всей области входных переменных $\Omega_j(\vec{u})$, а в некоторой подобласти $\Omega_j^H(\vec{u})$.

В качестве примера можно представить следующий процесс с двумя входными и двумя выходными переменными (Рисунок 2.25):

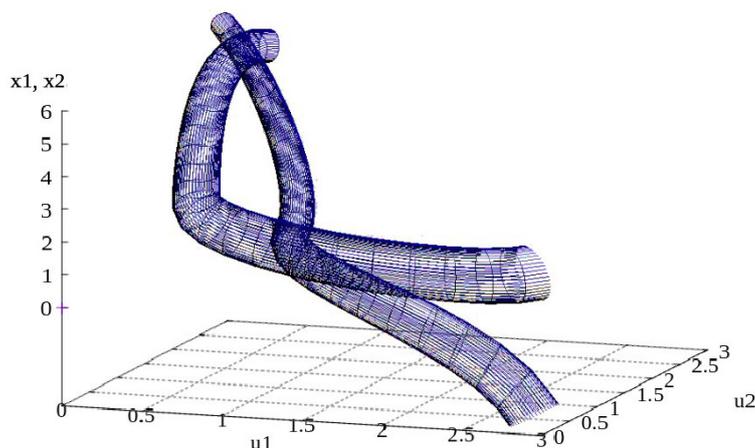


Рисунок 2.25 — H -процесс с двумя входными и выходными переменными

Наличие подобных процессов предполагает появление качественно новых задач в сфере моделирования и управления.

Безусловно возникает вопрос о моделировании подобных процессов.

Рассмотрим частный случай H -процессов с несколькими выходными переменными.

Предположим, что изучаемый процесс имеет двое входных и двое выходных переменных. Обучающая выборка размера s $(x_{1i}, x_{2i}, \vec{u}_i), i = \overline{1, s}$.

На рисунке 2.26 отображена структура изучаемого процесса

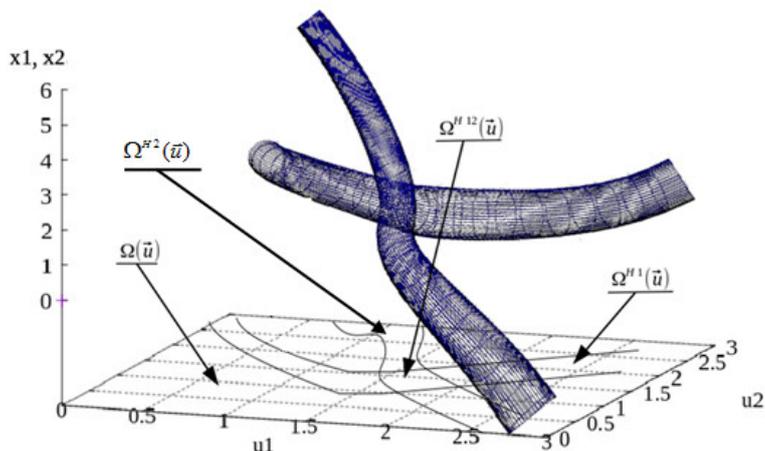


Рисунок 2.26 — Процесс с двумя выходными переменными, имеющими общую область определения

На рисунке приняты следующие обозначения: $\Omega(\vec{u})$ - область определения входных-выходных переменных; $\Omega^{H1}(\vec{u})$ - область определения входных переменных для x_1 ; $\Omega^{H2}(\vec{u})$ - область определения входных переменных для x_2 ; $\Omega^{H12}(\vec{u})$ - пересечение областей $\Omega^{H1}(\vec{u})$ и $\Omega^{H2}(\vec{u})$; u_1, u_2 , - входные переменные; x_1, x_2 - выходные переменные.

Еще раз напомним, выходная переменная x_1 определена только в области $\Omega^{H1}(\vec{u})$, вне области $\Omega^{H1}(\vec{u})$, часть процесса с выходной переменной x_1 не существует, то же касается и x_2 с областью $\Omega^{H2}(\vec{u})$.

Исходя из данного рисунка видно, что процесс с двумя выходными переменными x_1 и x_2 , определен только в области $\Omega^{H12}(\vec{u})$.

Исследователь не знает о существовании $\Omega^{H1}(\vec{u})$ и $\Omega^{H2}(\vec{u})$, следовательно он не знает и о области $\Omega^{H12}(\vec{u})$, что может отрицательно сказаться на построенной модели (она может дать прогнозы при входных воздействиях,

не входящих в области $\Omega^{H1}(\vec{u})$ и $\Omega^{H2}(\vec{u})$, то есть где процесс не определен, а значит прогнозы будут заведомо неверными).

Логичным решением задачи моделирования с несколькими выходными переменными будет использование двух индикаторных функций.

При использовании индикаторных функций модель будет выглядеть следующим образом:

$$\begin{aligned} x_1 &= I_1(\vec{u})I_2(\vec{u})f_1(\vec{u}) \\ x_2 &= I_1(\vec{u})I_2(\vec{u})f_2(\vec{u}) \end{aligned} \quad (2.6.6)$$

где $I_1(\vec{u})$ - индикаторная функция, определяющая принадлежность вектора входных воздействий \vec{u} к области $\Omega^{H1}(\vec{u})$; $I_2(\vec{u})$ - индикаторная функция, определяющая принадлежность вектора входных воздействий \vec{u} к области $\Omega^{H2}(\vec{u})$; x_1 - первая выходная переменная процесса; x_2 - вторая выходная переменная процесса.

$$I_1(\vec{u}) = \begin{cases} 1, & \vec{u} \in \Omega^{H1}(\vec{u}) \\ 0, & \vec{u} \notin \Omega^{H1}(\vec{u}) \end{cases} . \quad (2.6.7)$$

$$I_2(\vec{u}) = \begin{cases} 1, & \vec{u} \in \Omega^{H2}(\vec{u}) \\ 0, & \vec{u} \notin \Omega^{H2}(\vec{u}) \end{cases} . \quad (2.6.8)$$

В общем виде предложенная модель N-процесса с k выходными переменными выглядит следующим образом:

$$\begin{aligned}
\hat{x}_1 &= \hat{f}_1(\vec{u}^{<1>}, \vec{\alpha}_{s1}) \prod_{i=0}^k I_i(\vec{u}^{<1>}, \vec{u}_s) \\
\hat{x}_2 &= \hat{f}_2(\vec{u}^{<2>}, \vec{\alpha}_{s2}) \prod_{i=0}^k I_i(\vec{u}^{<2>}, \vec{u}_s) \\
&\dots\dots\dots \\
\hat{x}_k &= \hat{f}_k(\vec{u}^{<k>}, \vec{\alpha}_{sk}) \prod_{i=0}^k I_i(\vec{u}^{<k>}, \vec{u}_s)
\end{aligned} \tag{2.6.9}$$

где k – размерность вектора выходных переменных \vec{x} ; $\vec{u}^{<j>}$ – вектор входных воздействий, который содержит компоненты вектора \vec{u} , участвующие в формировании \hat{x}_j ; $I_j(\vec{u}^{<i>}, \vec{u}_s)$ – индикаторная функция, определяющая принадлежность $\vec{u}^{<j>}$ к области $\Omega_j^H(\vec{u})$, $j = (\overline{1, k})$.

В случае если одна из компонент вектора выходных переменных \vec{x} определена во всей области $\Omega(\vec{u})$, то соответствующая индикаторная функция будет всегда иметь значение 1, и, соответственно, не будет влиять на выражение $\prod_{l=0}^k I_l(\vec{u})$.

Помимо этого, в случае если процесс не является процессом типа Н, то все индикаторы примут значение 1 в любой точке из области $\Omega(\vec{u})$, а значит $\prod_{l=0}^k I_l(\vec{u}) = 1$. В этом случае система (2.6.9) принимает вид (2.6.10):

$$\begin{aligned}
x_1 &= \hat{f}_1(\vec{u}^{<1>}, \vec{\alpha}_{s1}) \\
x_2 &= \hat{f}_2(\vec{u}^{<2>}, \vec{\alpha}_{s2}) \\
&\dots\dots\dots \\
x_k &= \hat{f}_k(\vec{u}^{<k>}, \vec{\alpha}_{sk})
\end{aligned} \tag{2.6.10}$$

Это подтверждает, что система (2.6.9) является более общим видом модели процесса с несколькими выходными переменными.

Продемонстрируем результат моделирования, с использованием индикаторных функций.

В качестве исследуемого процесса, возьмем следующую систему уравнений:

$$\begin{aligned}x_1 &= u_1 + u_2 \\x_2 &= u_1 + 0.6u_2,\end{aligned}\tag{2.6.11}$$

$\vec{u} \in (\overline{0, 3})$. на процесс воздействует помеха 5%. Были созданы две параметрические модели, отражающие данные системы.

Данный процесс является Н-процессом. Для первого выходного воздействия выполняется следующее равенство:

$$u_2 = \sin(u_1/1.3) * 2 + \xi_1, \xi_1 \in (-0.2; 0.2).\tag{2.6.12}$$

Для второго выходного воздействия выполняется следующее равенство.

$$u_2 = 3 - u_1 + \xi_2, \xi_2 \in (-0.2; 0.2).\tag{2.6.13}$$

Были созданы две параметрических модели, описывающие поведение каждой выходной переменной.

Представим результаты моделирования при помощи параметрических моделей без использования индикаторных функций (Рисунок 2.27)

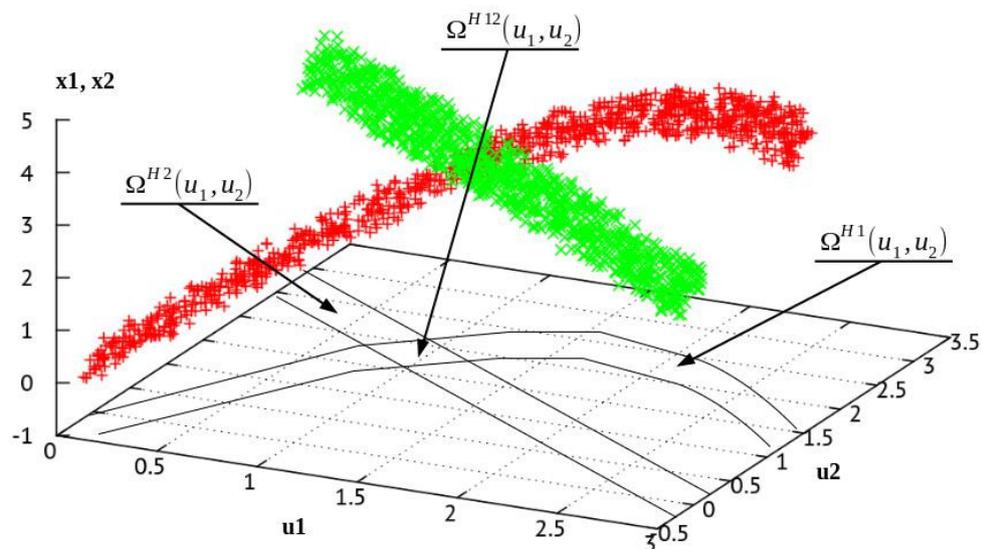


Рисунок 2.27 — Результаты моделирования при помощи параметрических моделей, без использования индикаторных функций

На рисунке 2.27 видно что представленная модель работает и вне области $\Omega^{H12}(\vec{u})$, иными словами эта модель дает оценку там, где процесс не протекает. Представим результаты моделирования при помощи параметрических моделей с использованием индикаторных функций 2.28

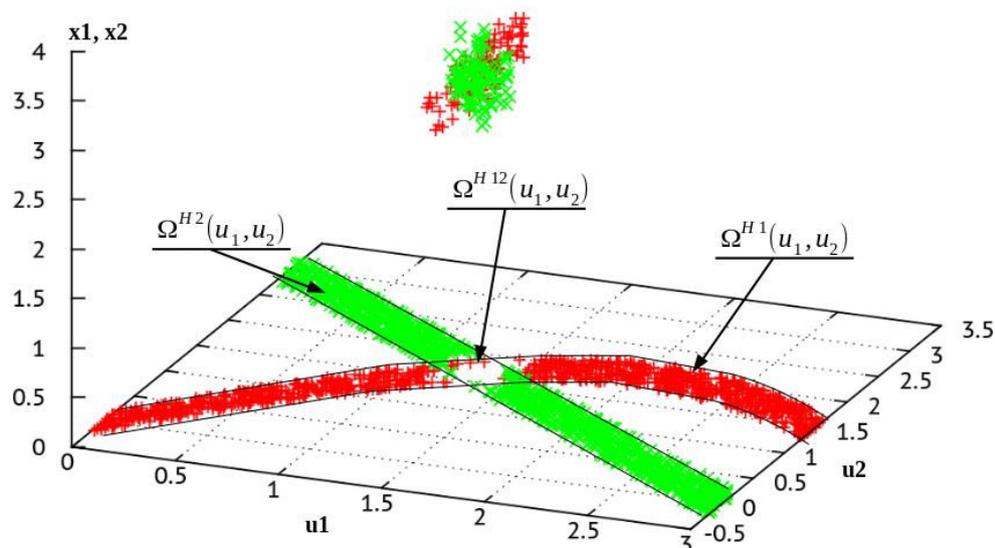


Рисунок 2.28 — Результаты моделирования при помощи параметрических моделей, с использованием индикаторных функций

Из второго рисунка видно, что модель строит прогнозы, только для $\vec{u} \in \Omega^{H12}(\vec{u})$, что подтверждает возможность использования предложенного алгоритма, для моделирования многомерных Н-процессов.

2.6.3 Модель многомерного процесса без общей зоны определения входных переменных

Представим случай в котором входные воздействия исследуемого процесса определены в области $\Omega^H(\vec{u}_i), i = \overline{1, k} \in \Omega(\vec{u})$, но при этом данные области не имеют общей области пересечения, $\prod_{i=1}^k \Omega^H(\vec{u}_i) = \emptyset$.

Представим данный случай в виде схематичного рисунка (Рисунок 2.29)

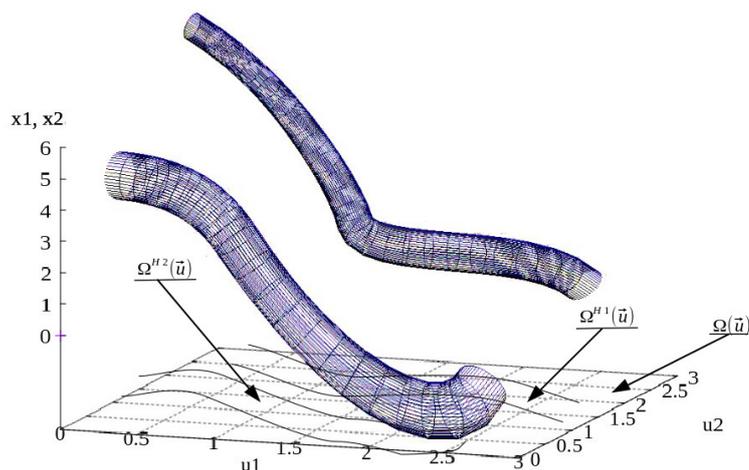


Рисунок 2.29 — H-Процесс без общей области определения входных переменных

Отсутствие пересечения является свидетельством того, что исследуемый процесс не может протекать с данным набором выходных переменных. Это может происходить в случае если были выбраны взаимоисключающие выходные переменные.

Данная ситуация может быть характерна для активных процессов (процессов, которые проходят с участием людей), организационных процессов, реже - технологических процессов. При обнаружении отсутствия пересечения между выбранными выходными переменными, эксперту необходимо изменить набор выходных переменных на такой набор, при котором $\prod_{i=1}^k \Omega^H(\vec{u}_i) \neq \emptyset$. Стоит отметить, что этот набор должен быть произведен только при помощи людей-экспертов.

Описанный случай характеризует переход от теории управления к теории принятия решения и является новым для задач идентификации. Помимо этого необходимо сказать, что множество активных процессов будут подвержены особенности, продемонстрированной выше, поэтому её изучение является актуальным.

2.7 Модели основанные на технологии нейросетей

Одной из первых искусственных нейронных сетей способных обучению является перцептрон Розенблатта. Перцептрон рассматривался Розенблатом не как конкретное техническое вычислительное устройство, а как модель работы мозга. Элементарный перцептрон Розенблатта представлен на рисунке 2.30

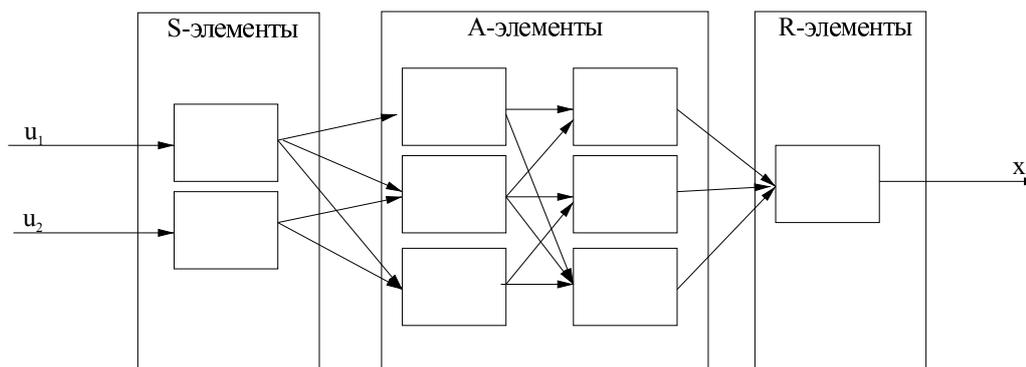


Рисунок 2.30 — Элементарный перцептрон Розенблатта

Перцептрон содержит элементы трех типов: вектор входных переменных \vec{u} поступает на S-элементы, которые его в дальнейшем передают в нейросеть.

Далее \vec{u} поступают в слой ассоциативных или А-элементов. Ассоциативные элементы, представляющие собой слой нейронов, выполняют нелинейную обработку информации и имеют изменяемые веса связей. R-элементы с фиксированными весами формируют сигнал реакции перцептрона на входной стимул.

Однослойный перцептрон характеризуется матрицей весов связей (W) от S- к А-элементам. Элемент матрицы W_{ij} отвечает весу связи, ведущей от i -го S-элемента к j -му А-элементу.

В работах Розенблатта был сделано заключение о том, что нейронная сеть рассмотренной архитектуры будет способна к воспроизведению любой функции. Стоит отметить, что позднее М.Минским и С.Пейпертом[13], было

показано что этот вывод оказался неточным. Были выявлены принципиальные неустранимые ограничения однослойных персептронов, и в последствии стал, в основном, рассматриваться многослойный вариант персептрона.

2.7.1 Искусственный нейрон

Нейрон состоит из элементов трех типов: умножителей (синапсов), сумматора и нелинейного преобразователя. Синапсы осуществляют связь между нейронами, умножают входной сигнал на число, характеризующее силу связи, (вес синапса). Сумматор выполняет сложение сигналов, поступающих по синаптическим связям от других нейронов, и внешних входных сигналов. Нелинейный преобразователь реализует нелинейную функцию одного аргумента - выхода сумматора. На рисунке (2.31) показана его структура.

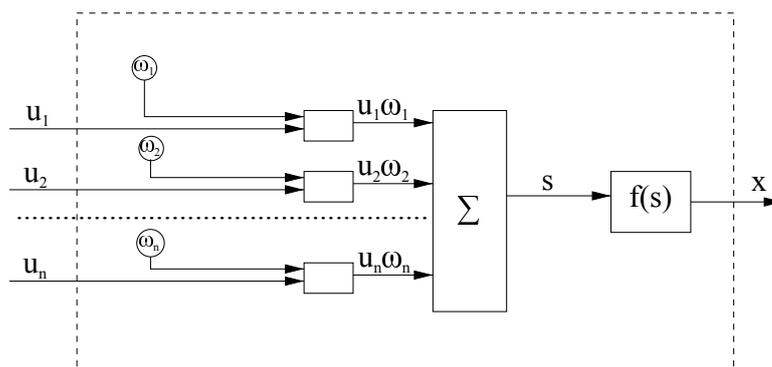


Рисунок 2.31 — Структура нейрона

Функция $f(s)$ называется функцией активации нейрона.

Нейрон реализует скалярную функцию векторного аргумента. Математическая модель нейрона:

$$x = f(s), \quad (2.7.1)$$

$$s = \sum_{i=1}^n (w_i u_i) \quad (2.7.2)$$

где w_i — вес синапса; s — результат суммирования; u_i - компонент входного вектора (входной сигнал); x — выходной сигнал нейрона; n — число входов нейрона; $i = \overline{1, n}$.

Выход (x) определяется видом функции активации и может быть как действительным, так и целым.

Синаптические связи с положительными весами называют возбуждающими, с отрицательными весами - тормозящими.

2.7.2 Функции активации

Одной из наиболее распространенных функций активации является нелинейная функция активации с насыщением, так называемая логистическая функция или сигмоид (функция S-образного вида):

$$f(s) = \frac{1}{1 + e^{-as}} \quad (2.7.3)$$

При уменьшении a сигмоид становится более пологим, в пределе при $a = 0$ вырождаясь в горизонтальную линию на уровне 0.5, при увеличении a сигмоид приближается к виду функции Хевисайда.

Одно из ценных свойств сигмоидальной функции - простое выражение для ее производной (2.7.4):

$$f'(s) = a * f(s)(1 - f(x)) \quad (2.7.4)$$

Следует отметить, что сигмоидальная функция дифференцируема на всей оси абсцисс, что используется в некоторых алгоритмах обучения. Кроме того, она обладает свойством усиливать слабые сигналы лучше, чем большие,

и предотвращает насыщение от больших сигналов, так как они соответствуют областям аргументов, где сигмоид имеет пологий наклон.

2.7.3 Метод обратного распространения ошибки

Рассмотрим процесс обучения нейронной сети, с использованием алгоритма обратного распространения ошибки.

Метод обратного распространения ошибки заключается в итерационной поправке весовых коэффициентов синапсов (связях между нейронами). Продемонстрируем каким образом вычисляются корректировки весов на каждой итерации.

Воспользуемся дельта-правилом:

$$\delta w_{jk} = h \delta_k x_j, \quad (2.7.5)$$

где h — норма обучения. Это число задается исследователем перед началом обучения. x_j — выход, нейрона j поступающий ко входу k нейрону k , δ_k — ошибка нейрона k .

Таким образом, в процессе обучения, на вход сети подаются элементы обучающей выборки и обновляются значения весовых коэффициентов. Обычно обучение заканчивается когда для всех вводимых образцов величина ошибки станет меньше определенной величины. После этого сеть подвергается тестированию при помощи новых данных, которые не участвовали в обучении. И по результатам этого тестирования уже можно сделать выводы, хорошо или нет справляется сеть со своими задачами.

Определим ошибку δ :

$$\delta = y - \hat{y} \quad (2.7.6)$$

Алгоритм определяет два «потока» в сети. Входные сигналы двигают-

ся в прямом направлении, в результате чего мы получаем выходной сигнал, из которого мы получаем значение ошибки. Величина ошибки двигается в обратном направлении, в результате происходит корректировка весовых коэффициентов связей сети.

Итак, для корректировки весовых значений будет использоваться дельта-правило. Необходимо определить универсальное правило для вычисления ошибки каждого элемента сети, после прохождения через элемент (при обратном распространении ошибок). Данное правило будет выглядеть следующим образом (2.7.7):

$$\delta_j = f'(y_j) \sum_{k=1}^p \delta_k * w_{jk} \quad (2.7.7)$$

Продемонстрируем работу нейросетей на примере. Создадим нейросеть со следующей архитектурой (Рисунок 2.32)

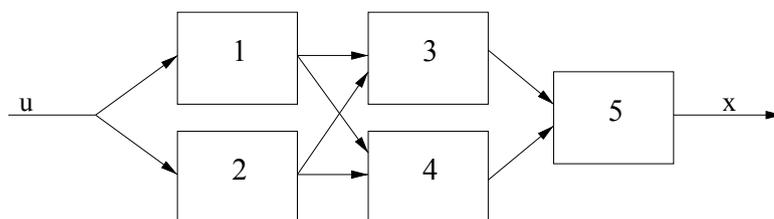


Рисунок 2.32 — Структура созданного перцептрона

В качестве функции активации возьмем функцию (2.7.3):

Математическая модель данного перцептрона выглядит следующим образом (2.7.8):

$$x = \frac{1}{1 + \exp^{\omega_{35} * u_3 + \omega_{45} * u_4}}, \quad (2.7.8)$$

где ω_{ij} это вес на синапсе между нейроном под номером i и нейроном под номером j , а u_i - это сигнал (выходное значение) из нейрона под номером i ,

$$u_3 = \frac{1}{1 + \exp^{\omega_{13} * u_1 + \omega_{23} * u_2}}, \quad (2.7.9)$$

$$u_4 = \frac{1}{1 + \exp^{\omega_{14} * u_1 + \omega_{24} * u_2}} \quad (2.7.10)$$

$$u_1 = \frac{1}{1 + \exp^{\omega_{01} * u}}, \quad (2.7.11)$$

$$u_2 = \frac{1}{1 + \exp^{\omega_{02} * u}}, \quad (2.7.12)$$

Здесь ω_{0i} это вес, между входом в нейросеть и нейроном под номером i . Объединяя формулы (2.7.8 - 2.7.12) получаем итоговую математическую модель процесса (2.7.13):

$$x = 1 / (1 + \exp(-(\omega_{35} * (1) / (1 + \exp^{\omega_{13} * (1) / (1 + \exp^{\omega_{01} * u}) + \omega_{23} * (1) / (1 + \exp^{\omega_{02} * u})}) + \omega_{45} * (1) / (1 + \exp^{\omega_{14} * (1) / (1 + \exp^{\omega_{01} * u}) + \omega_{24} * (1) / (1 + \exp^{\omega_{02} * u})})))))) \quad (2.7.13)$$

По формуле (2.7.13) минимальное значение модели равно 0, а максимальное 1. Для того чтобы использовать данную нейросеть в исследуемых процессах, нужно чтобы минимальный возможный выход объекта соответствовал выходу нейросети 0, а максимальный - выходу нейросети 1. Для этого проведем преобразование (2.7.14):

$$x_r = x * (max - min) + min. \quad (2.7.14)$$

где x_r - преобразованный выход, max - максимальный выход объекта, min - минимальный выход объекта.

Как видно из формулы (2.7.13) единственные неизвестные параметры - это веса между нейронами. В зависимости от вида моделируемого объекта эти веса меняются. Рассчитаем веса для различных процессов.

Рассчитаем веса для процесса, описываемого формулой (2.7.15)

$$x = u * 3 \quad (2.7.15)$$

После обучения нейросети веса на связях будут определены следующим образом:

$$w_{01} = 0.50; w_{02} = 0.50; w_{13} = -0.94; w_{14} = 1.025; w_{23} = -0.99; w_{24} = 0.84; w_{35} = -35.79; w_{45} = 9.86.$$

Подставив веса в формулу (2.7.13) можно построить график зависимости x от u (Рисунок 2.33)

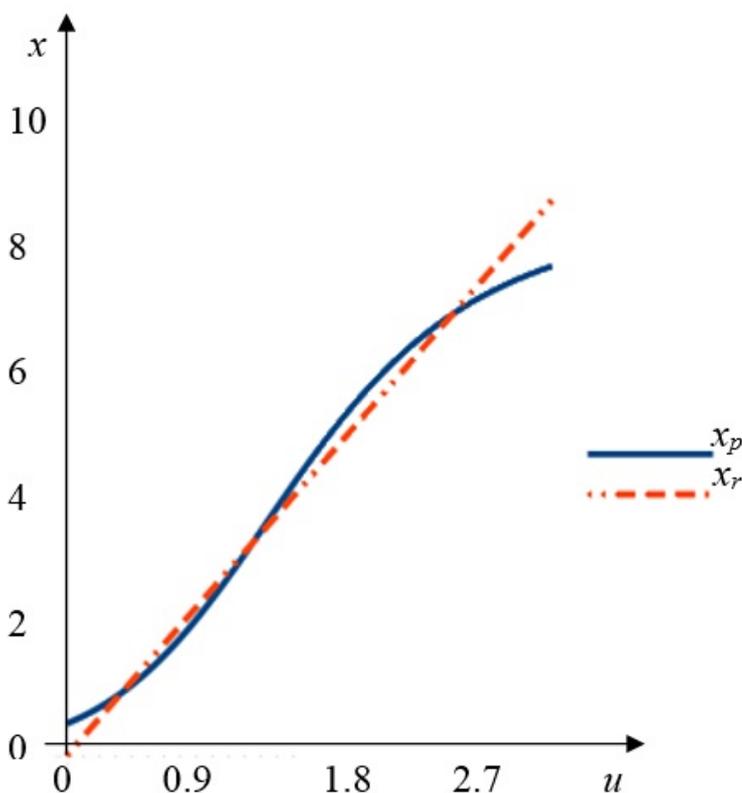


Рисунок 2.33 — Модель процесса, обладающего линейной структурой

На рисунке 2.33 приняты следующие обозначения: x_r - истинный выход процесса; x_p - выход модели на основе нейросети.

Рисунок 2.33 отражает что данный вид персептрона при определенной

настройке весов может моделировать поведение линейного процесса.

Изменим вид моделируемого процесса. Пусть процесс описывается формулой (2.7.16).

$$x = u^2 \quad (2.7.16)$$

После обучения нейросети веса на связях будут определены следующим образом:

$$w_{01} = -0.47, w_{02} = 0.44, w_{13} = 32.32, w_{14} = -3.34, w_{23} = -5.871, w_{24} = 4.18, w_{35} = -15.46, w_{45} = 17.12.$$

Подставив полученные веса в формулу (2.7.13) построим график зависимости x от u (Рисунок 2.34)

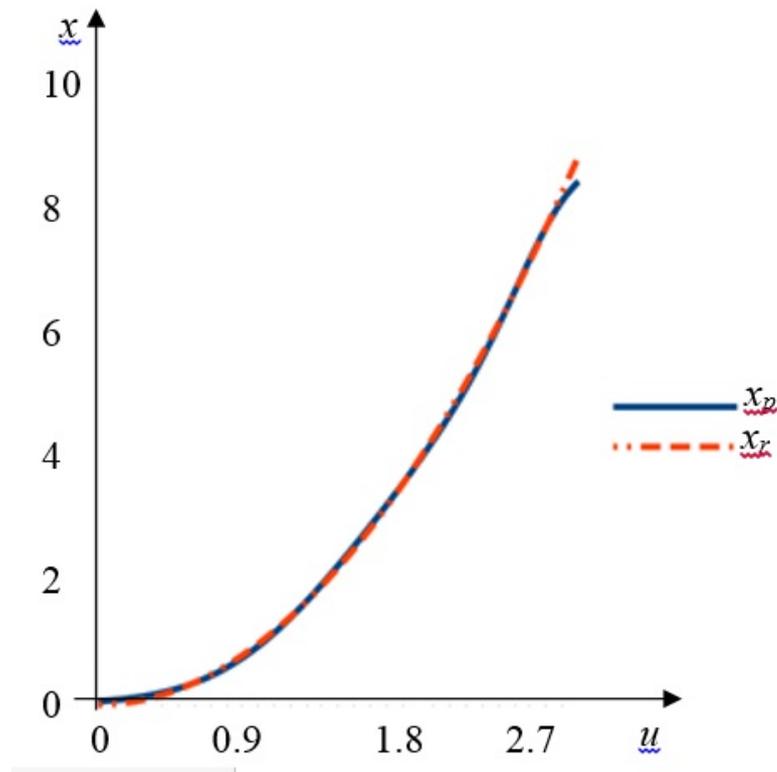


Рисунок 2.34 — Модель процесса, обладающего квадратичной структурой

Рисунок 2.34 отражает что данный вид персептрона при определенной

настройке весов может также моделировать поведение и квадратичного процесса.

Последнее, что будет продемонстрировано - моделирование процесса с синусоидальной структурой.

Еще раз изменим вид моделируемого процесса. Пусть процесс описывается формулой (2.7.17)

$$x = \sin(u) \quad (2.7.17)$$

После обучения нейросети веса на связях будут определены следующим образом:

$$w_{01} = 1.49, w_{02} = 3.10, w_{13} = -3.60, w_{14} = -61.42, w_{23} = 1.42, w_{24} = 60.11, w_{35} = -40.73, w_{45} = 8.78.$$

Подставив веса в формулу (2.7.13), построим график зависимости x от u (Рисунок 2.35)

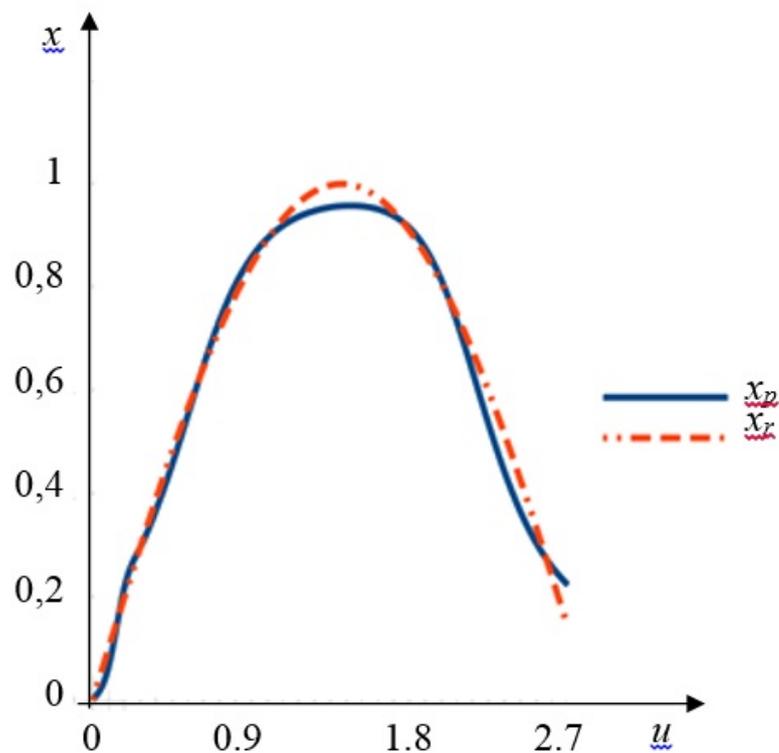


Рисунок 2.35 — Модель процесса, обладающего синусоидальной структурой

Рисунок 2.35 отражает что данный вид персептрона при определенной настройке весов может также моделировать поведение и синусоидального процесса.

Вышеприведенные примеры отражают то, что модель на основе нейросети может принимать различные виды. Все что требуется это подстройка весов таким образом, чтобы полученная модель максимально отражала процесс.

2.8 Сравнение алгоритмов моделирования

В рамках диссертационного исследования было проведено сравнение между двумя известными методами идентификации: многослойным персептроном Розенблата и алгоритмом локальной аппроксимации.

Продемонстрируем результаты моделирования при помощи персептрона Розенблата и локальной аппроксимации. Строились модели различных объектов и в различных условиях. Изменялись следующие условия: количество входных переменных (3, 5, 10), уровень помехи (0%, 1% и 10%), размер обучающей выборки (100, 500, 1000), качество выборки (2 - наличие выбросов и разряженностей, 1 - отсутствие выбросов и разряженностей). Часть результатов моделирования представлена в таблице 2.

Таблица 2 — Сравнение алгоритмов моделирования

1	2	3	4	5	6	7	8	9	10
5	500	3	1	1 030	0,63	4	47	0,59	595
10	500	3	1	1 090	0,62	35	29	0,66	123
0	1 000	3	1	2 093	0,02	21	28	0,003	97
5	1 000	3	1	2 062	0,04	19	28	0,004	100
10	1 000	3	1	2 063	0,08	21	29	0,013	250
5	100	6	1	403	2,43	92	36	3,64	830
0	1 000	10	1	6 554	0,02	125	45	0,0052	672
5	1 000	10	1	6 657	0,28	126	45	0,17	183
5	100	3	2	207	1,82	75	29	1,40	790
10	100	3	2	207	2,33	32	29	1,18	66
0	500	3	2	1 039	0,23	12	30	0,29	1138
5	500	6	2	2 007	0,54	12	38	0,864	256
5	1 000	6	2	3 950	3,66	56	37	3,86	144
10	1 000	6	2	4 026	3,34	57	37	3,80	121

В таблице 2 приняты следующие обозначения столбцов:

- 1 - Уровень помехи (в %);
- 2 - Объем выборки наблюдений;
- 3 - Размерность вектора входных переменных \vec{u} ;
- 4 - Качество выборки;
- 5 - Время, в миллисекундах, потраченное на построение прогноза алгоритмом локальной аппроксимации;

- 6 - Средняя ошибка прогноза у модели, построенной на основе алгоритма локальной аппроксимации;
- 7 - Время оптимизации вектора \vec{c}_s , в секундах;
- 8 - время, в миллисекундах, потраченное на построение прогноза при помощи нейросети;
- 9 - Средняя ошибка прогноза у модели, построенной на основе нейросети;
- 10 - Время обучения нейросети, в секундах.

Как видно из таблицы 2 нет такого алгоритма среди исследуемых, который имеет абсолютное преимущество в точности перед другим, имеются примеры, когда точность у нейросетей выше, чем у модели, построенной на основе алгоритма локальной аппроксимации (например в таблице 2 при помехе 5% обучающей выборке 500, количестве входных переменных 3, хорошем качестве выборки), но также имеются и примеры когда нейросеть уступает модели, построенной на основе локальной аппроксимации, в точности (при помехе 10%, обучающей выборке 500, количестве входных переменных 3, хорошем качестве выборки).

У нейросети имеется преимущество перед ядерной аппроксимацией в скорости расчета, в то время как настройка вектора \vec{c}_s у ядерной аппроксимации происходит значительно быстрее, чем обучение нейросетей.

Одной из особенностей работы с нейросетью является большая затрата времени на обучение. Усложнение объекта и увеличение количества входных переменных влечет за собой увеличение времени обучения.

Учитывая вышесказанное, имеет смысл построить график зависимости между количеством входных переменных и временем обучения нейросети. Указанная зависимость отображена на рисунке 2.36.

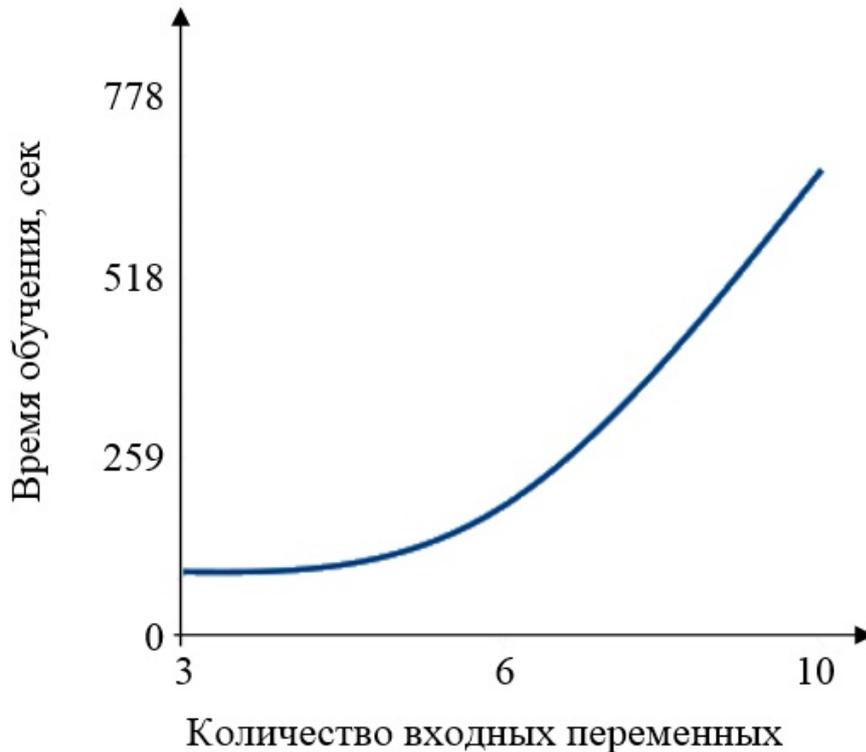


Рисунок 2.36 — Зависимость между количеством входных переменных и временем обучения нейросети

Зависимость, отображенная на рисунке 2.36 имеет квадратичный или экспоненциальный характер.

Исходя из рисунка 2.36, при построении модели сложных объектов со множеством входов исследователь может столкнуться с проблемой слишком долгого обучения.

Также стоит посмотреть данную зависимость у модели, построенной на основе локальной аппроксимации (вместо обучения будет взято время настройки вектора \vec{c}_s). Зависимость отображена на рисунке 2.37

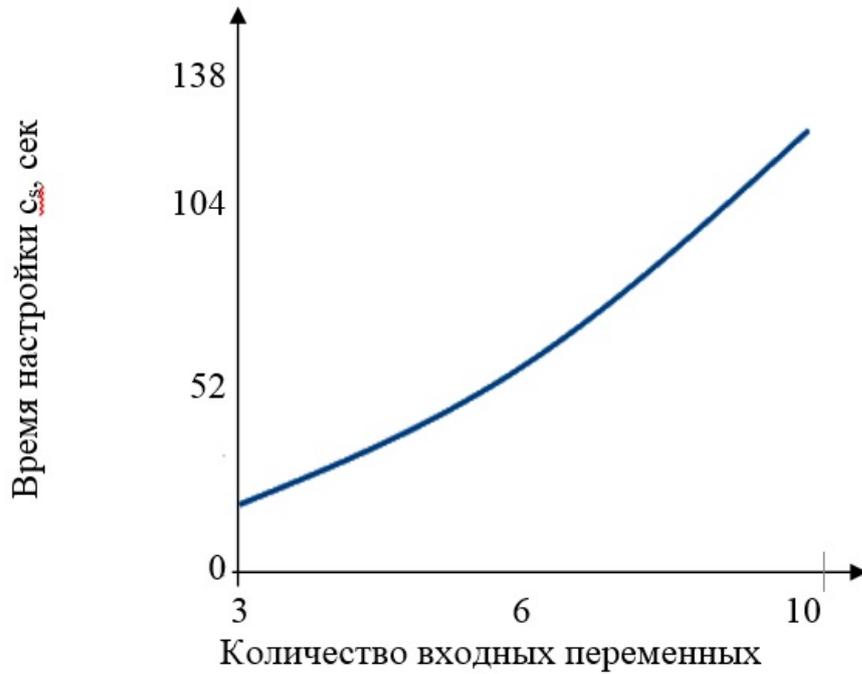


Рисунок 2.37 — Зависимость между количеством входных переменных и временем настройки \vec{c}_s

Как можно увидеть на рисунке 2.37 зависимость имеет линейный характер.

На основании рисунков 2.36 и 2.37, можно предположить, что при моделировании сложных процессов с большим количеством входных переменных, подстройка вектора \vec{c}_s будет происходить значительно быстрее, чем обучение нейросети.

Также для полноценного анализа необходимо построить график зависимости между скоростью вычисления и размером обучающей выборки. Связано это с тем, что существуют процессы при моделировании которых скорость расчета модели имеет существенную роль.

Построим график зависимости между скоростью расчета прогноза и размером обучающей выборки, у модели, построенной на основе локальной аппроксимации (Рисунок 2.38)

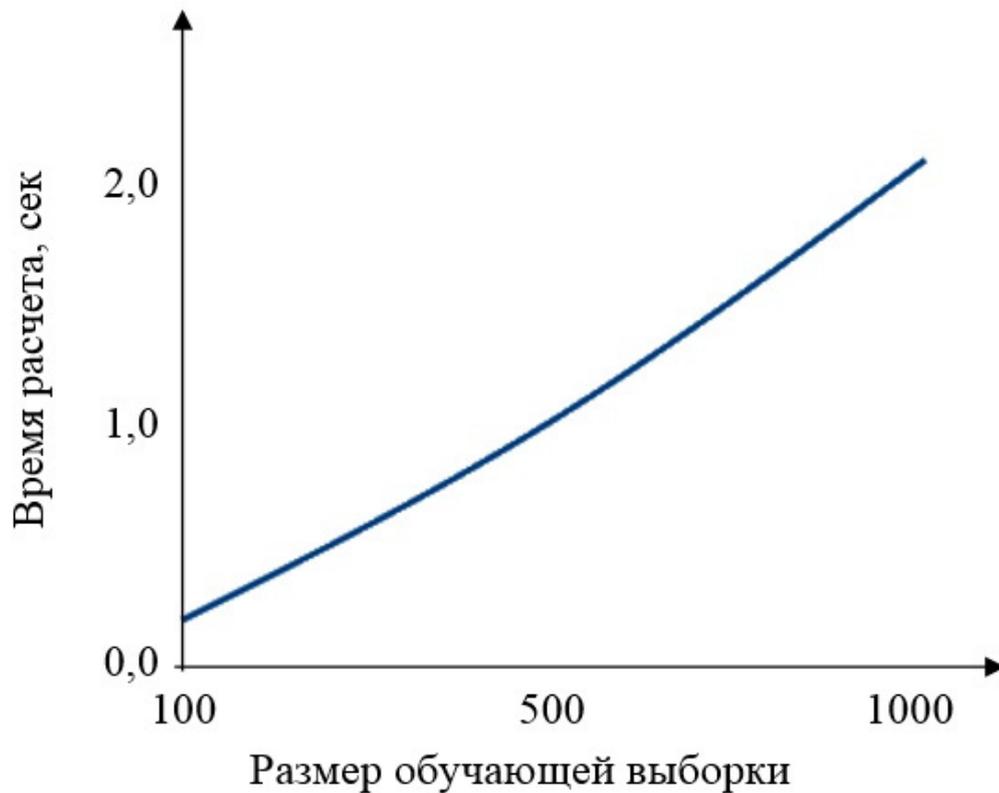


Рисунок 2.38 — Зависимость между скоростью расчета прогноза и размером обучающей выборки, у модели, построенной на основе локальной аппроксимации

Как видно из рисунка 2.38 зависимость имеет линейных характер.

Теперь построим график зависимости между скоростью расчета прогноза и размером обучающей выборки в модели построенной на основе нейросети (Рисунок 2.39)

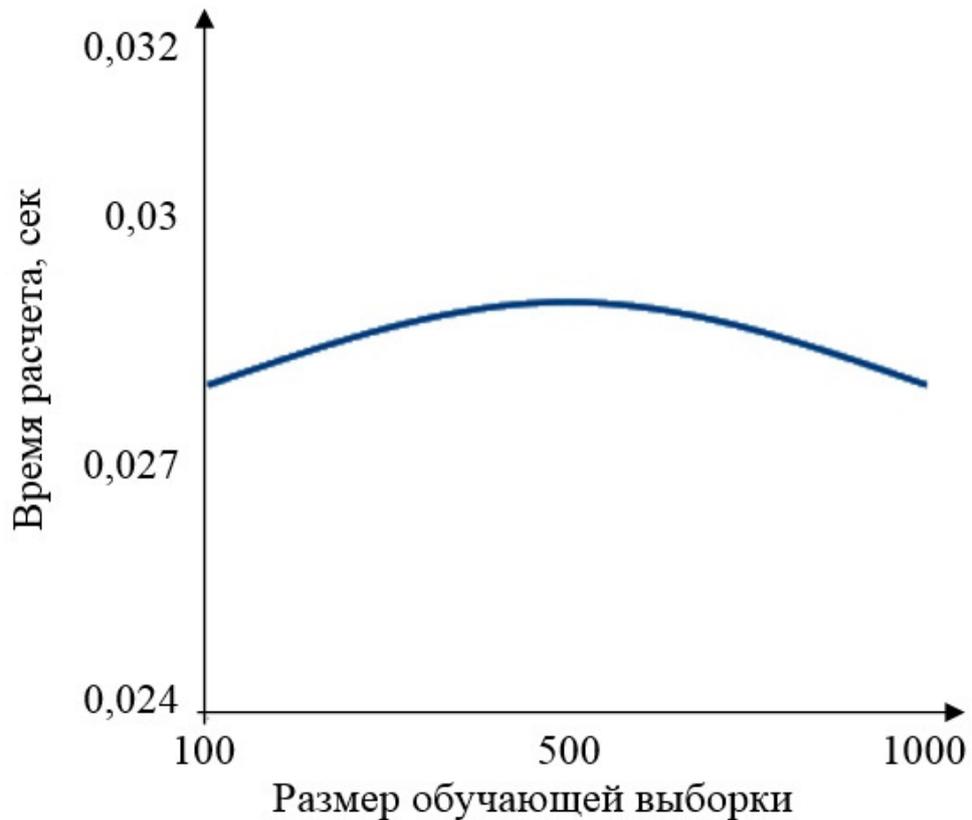


Рисунок 2.39 — Зависимость между скоростью расчета прогноза и размером обучающей выборки в модели построенной на основе нейросети

Рисунок 2.39 подтверждает что скорость расчета прогноза у моделей построенной на основе нейросетей не зависит от объема выборки наблюдений.

Анализируя графики 2.38 и 2.39, можно прийти к выводу, что скорость вычисления у нейросети не только выше чем у модели, построенной на основе ядерной аппроксимации, но и не зависит от количества обучающей выборки, что может быть преимуществом при моделировании процессов, где требуется высокая скорость построения прогноза и имеется большое количество обучающей информации.

2.9 Выводы по главе

Глава посвящена вопросам моделирования стохастических безынерционных процессов.

В первую очередь была представлена постановка задачи моделирования дискретно непрерывных процессов.

В дальнейшем рассмотрен вопрос выделения существенных переменных. Рассмотрены существующие алгоритмы выделения существенных переменных, их преимущества и недостатки. Также представлен новый алгоритм, основанный на настройке вектора параметров размытости в модели построенной на основе локальной аппроксимации. В ходе сравнения выяснилось, что предложенный алгоритм быстрее решает задачу выделения существенных переменных чем такие алгоритмы как Ad, Del, AdDel. Особенностью предложенного алгоритма является то, что он основан на методах непараметрической статистики, и, соответственно работает в условиях непараметрической неопределенности.

Были рассмотрены несколько методов решения задачи идентификации.

В первую очередь были рассмотрены параметрические методы решения задачи идентификации. Несмотря на распространенность данных методов, моделировать при помощи подобных методов сложные технологические процессы затруднительно. Основная сложность при использовании параметрических методов заключается в подборе "удачной" параметрической структуры.

Были рассмотрены непараметрические методы решения задачи идентификации. В качестве отдельной задачи, был исследован вопрос настройки вектора параметров размытости в модели, построенной на основе методов локальной аппроксимации.

Описан процесс улучшения качества выборки наблюдений, при помощи

алгоритма избавления от выбросов.

Следующий тип рассматриваемых моделей - Н-модели, которые являются объединением параметрической модели и индикаторной функции, основанной на непараметрическом методе локальной аппроксимации. Было продемонстрировано превосходство данных методов моделирования перед параметрическим, при построении моделей Н-процессов.

В главе были рассмотрены методы моделирования, основанные на технологии нейросетей, описаны общие принципы работы, их преимущества и недостатки.

В конце главы проведен сравнительный анализ между алгоритмом локальной аппроксимации и многослойным перцептроном Розенблатта-Парзена. В ходе сравнения выяснилось, что многослойный перцептрон Розенблатта-Парзена больше времени тратит на обучение, чем модель, построенная на основе локальной аппроксимации. В то же время модель, построенная на основе локальной аппроксимации, тратит большее время на расчет прогноза. Точность прогноза в моделях, основанных на локальной аппроксимации близка к точности прогноза в моделях, созданных при помощи многослойного перцептрона.

3 Управление безынерционными стохастическими процессами при неполной информации

В главе рассмотрена задача управления многомерными безынерционными процессами с запаздыванием в условиях непараметрической неопределенности. При управлении многомерными H -процессами есть существенное отличие от традиционных постановок задач управления. И определяется это, прежде всего тем, что задающие воздействия, определяющие желаемые значения компонент вектора выхода не могут быть выбраны произвольно из области определения задающих воздействий, как это принято в теории управления. В этом случае, возникает задача нахождения множества допустимых значений, для определения задающих воздействий. В главе представлен способ определения задающих воздействий, для случая управления многомерным H -процессом.

Помимо этого, рассмотрена другая особенность задачи управления H -процессами которая заключается в том, что устройство управления должно задать значения компонент вектора входных воздействий с учетом зависимости между ними. В главе предложен модифицированный метод непараметрического дуального управления, учитывающий зависимость между входными переменными.

3.1 Постановка задачи

Стоит задача управления многомерными безынерционными процессами с запаздыванием в условиях непараметрической неопределенности. На рисунке 3.1 представлена классическая схема управления объектом.

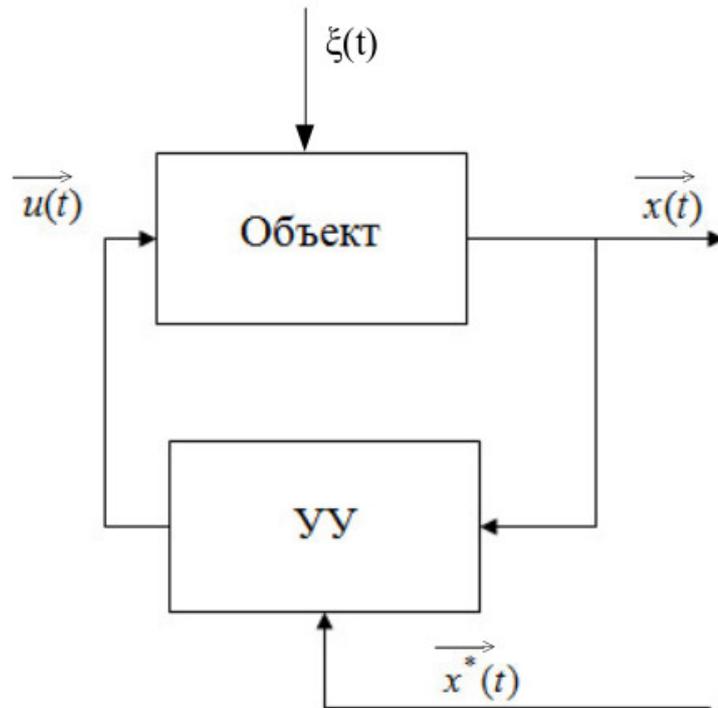


Рисунок 3.1 — Классическая схема управления объектом

где $\vec{u}(t)$ — вектор входных управляемых воздействий, размерности n ; $\vec{x}(t)$ — вектор выходных переменных процесса, размерности k ; \vec{x}^* — вектор задающих воздействий (задание); $\xi(t)$ — помеховое воздействие на процесс.

В зависимости от уровня априорной информации исследователь будет использовать различные алгоритмы управления. Например, предположим что необходимо разработать алгоритм управления для процесса у которого известна зависимость компонентов вектора выходных переменных от компонентов вектора входных переменных на основании физических законов. В этом случае, логично использовать известную зависимость для управления исследуемым процессом.

К сожалению в большинстве реальных задач параметрическая зависимость между вектором входных переменных \vec{u} и вектором выходных переменных \vec{x} — неизвестна. В этом случае, исследователь работает с объектом

управления как с "черным ящиком" в котором известны только значения компонентов вектора выходных переменных и значения компонентов вектора входных переменных. С таким видом процессов чаще всего сталкивается исследователь. В диссертации рассмотрена задача управления именно такими процессами.

3.2 Стандартные алгоритмы управления

Для управления различными процессами, как правило, используют типовые регуляторы, названия которых соответствуют названиям типовых звеньев:

П-регулятор (пропорциональный регулятор). Расчет выходного сигнала П-регулятора: $U_P(s) = K_p * (e(s) - e(s - 1))$. Принцип действия заключается в том, что регулятор вырабатывает управляющее воздействие на объект пропорционально величине ошибки (чем больше ошибка E , тем больше управляющее воздействие Y).

И - регулятор (интегрирующий регулятор). Выходной сигнал И - регулятора рассчитывается следующим образом: $U_I(s + 1) = K_i^{discr} e(s)$.

$$K_i^{discr} = K_p K_{ip} T \quad (3.2.1)$$

T — время дискретизации. Управляющее воздействие пропорционально интегралу от ошибки.

Д-регулятор, дифференцирующий регулятор. Расчет выходного сигнала Д-регулятора: $U_D(s + 1) = K_d^{discr} (e(s) - 2e(s - 1) + e(s - 2))$. Д-регулятор генерирует управляющее воздействие только при изменении регулируемой величины.

$$K_d^{discr} = \frac{K_p K_{ip}}{T} \quad (3.2.2)$$

На практике данные простейшие П, И, Д регуляторы комбинируются в регуляторы вида ПИ, ПД, ПИД:

ПИД-регулятор представляет собой сочетание П-, И- и Д-регуляторов. Расчет выходного сигнала ПИД- регулятора: $U_{PID}(s+1) = K_p(e(s) - e(s-1)) + K_i^{discr} e(s) + K_d^{discr} *(e(s) - 2e(s-1) + e(s-2))$. На практике наиболее часто используется ПИД-регулятор, поскольку он сочетает в себе достоинства всех трех типовых регуляторов.

Из-за простоты реализации ПИД-регуляторы широко распространены в различных схемах управления. Недостатком этих регуляторов является то, что они не являются адаптивными, иными словами они не обучаются.

3.3 Дуальное управление

Недостаточность априорной информации приводит к необходимости совмещать в определенном смысле изучение объекта и управление им. Мы не можем успешно управлять объектом, не зная его характеристик, но мы можем изучать объект, управляя им, и тем самым иметь возможность улучшать управление. В этой ситуации управляющие воздействия носят двойственный характер. Они служат как средством изучения, познания объекта, так и средством направления его к желаемому состоянию. Управление при котором управляющие воздействия носят двойственный характер называют дуальным управлением.

Успешное управление возможно если свойства объекта хорошо известны и управляющее устройство быстро реагирует на изменение состояния объекта. Но выяснение этих свойств, т. е. идентификация объекта, требует определенного времени. Вряд ли можно ожидать, что слишком поспешное управление, без достаточной информации о свойствах объекта, с одной стороны, и слишком осторожное управление, хотя и основанное на накопленной инфор-

мации, но действующее, когда надобность в нем миновала, с другой стороны, могут привести к успешному результату.

Дуальное управление было открыто и существенно развито А. А. Фельдбаумом [5] на основе теории статистических решений. Для работы алгоритма дуального управления необходимо, чтобы была известна плотность распределения внешних воздействий и параметров управляемого объекта. Важно отметить, что часто при решении реальных задач такой информации нет.

3.4 Непараметрическое дуальное управление

3.4.1 Общая информация о непараметрическом дуальном управлении

Алгоритм непараметрического дуального управления – это усовершенствование алгоритма, предложенного А.А. Фельдбаумом [5], которое позволяет уйти от определения параметрической модели.

Он был разработан и исследован А.В. Медведевым [7].

Аналитическое выражение этого алгоритма управления имеет следующий вид:

$$\vec{u}_{s+1} = \vec{u}_s(\vec{x}^*) + \delta \vec{u}_{s+1}(\vec{x}^*, \vec{x}_s). \quad (3.4.1)$$

$\vec{u}_s(\vec{x}^*)$ — составляющая, отвечающая за накопление информации об управляемом процессе (изучение объекта управления). На начальных этапах регулирования эта переменная близка к нулю и не влияет на управление процессом, но при накоплении обучающей выборки, её роль возрастает и становится доминирующей; \vec{x}^* — задающее воздействие; \vec{x}_s — выборка наблюдений, состоящая из значений компонентов вектора выходных переменных \vec{x} .

Составляющая $\vec{u}_s(\vec{x}^*)$ является математическим ожиданием \vec{u} при заданном значении \vec{x}^* :

$$\vec{u}_s(\vec{x}^*) = M(\vec{u}(\vec{x}^*)) \quad (3.4.2)$$

В алгоритме непараметрического дуального управления, в качестве оценки $M(\vec{u}(\vec{x}^*))$, используется непараметрическая оценка функции регрессии по наблюдениям [8]:

$$M(u_l(\vec{x}^*)) = \frac{\sum_{i=1}^s u_i \prod_{j=1}^k \Phi\left(\frac{x_j^* - x_{i,j}}{c_{sj}}\right)}{\sum_{i=1}^s \prod_{j=1}^k \Phi\left(\frac{x_j^* - x_{i,j}}{c_{sj}}\right)}, l = (\overline{1, n}), \quad (3.4.3)$$

где $\Phi(*)$ – колокообразная функция, а $c_{sj}, j = \overline{1, k}$ – параметры размытости.

$\delta\vec{u}_{s+1}(\vec{x}^*, \vec{x}_s)$ – это “поисковый шаг” алгоритма. На начальных этапах регулирования эта переменная вносит основной вклад в управление процессом, но при накоплении обучающей выборки, её роль в управлении становится несущественной.

Составляющая $\delta\vec{u}_{s+1}(\vec{x}^*, \vec{x}_s)$ рассчитывается по формуле 3.4.4:

$$\delta u_{s+1,j}(\vec{x}^*, \vec{x}_s) = m \sum_{i=1}^k (x_i^* - x_{si}), j = (\overline{1, n}), \quad (3.4.4)$$

где $(x_i^* - x_{si})$ – разница между заданием и прошлым выходом объекта, а m – параметр, отвечающий за величину «шага».

У воздействий, вырабатываемых непараметрическим дуальным алгоритмом имеется сразу два назначения: изучение объекта и управление им. Этим и обусловлена его дуальность.

В дальнейшем будет продемонстрировано использование данного алгоритма для различных объектов.

3.4.2 Управление процессом, имеющим одну выходную переменную

Для примера сравним работу алгоритма непараметрического алгоритма дуального управления и ПИ-регулятора.

Пусть процесс описывается следующим уравнением:

$$x = 2u + \sin(u) + 3 \quad (3.4.5)$$

Для ПИ-регулятора коэффициенты равны 0.02 и 0.1, у адаптивного регулятора параметр $k = 0.02$, $c_s = 1.2$, обучающей выборки в начале процесса управления нет, иными словами идет процесс активного накопления обучающей выборки.

Процесс управления представлен на рисунке 3.2

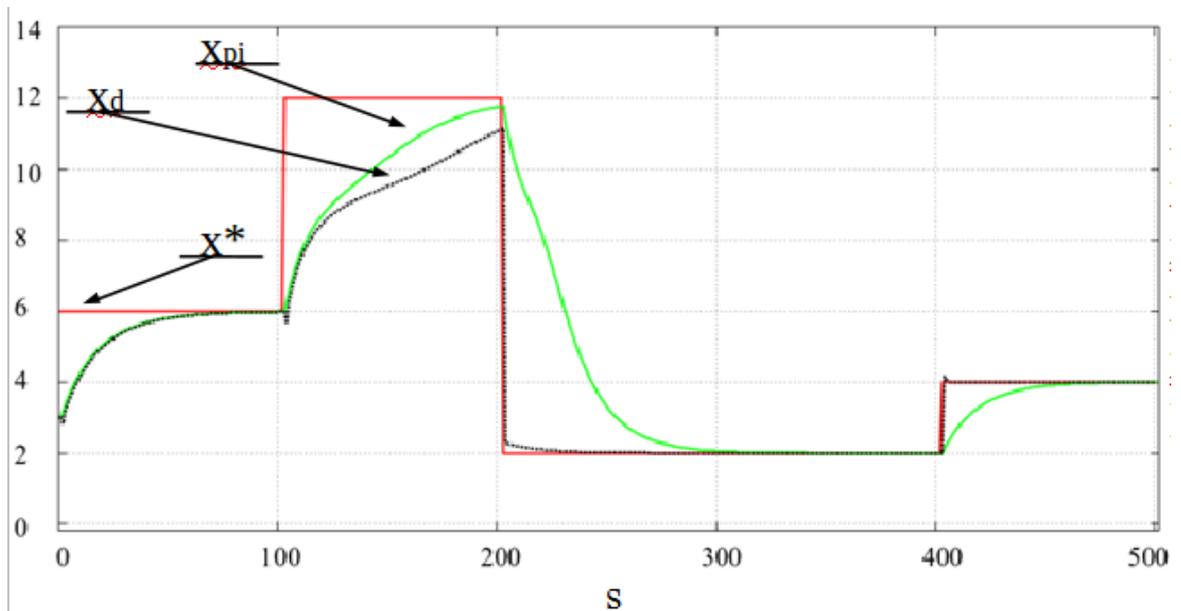


Рисунок 3.2 — Сравнение алгоритмов управления

На рисунке приняты следующие обозначения: x^* — задающее воздействие; x_d — процесс управления при помощи алгоритма непараметрического дуального управления; x_{pi} — процесс управления при помощи алгоритма ПИ-регулятора.

Как видно на рисунке, сначала адаптивный регулятор собирает информацию об объекте и скорость регулирования уступает либо равна скорости регулирования ПИ-регулятора. Разница в том что после обучения адаптивный

регулятор гораздо быстрее приводит объект к заданию чем ПИ регулятор.

Можно сделать вывод, что процесс обучения данного алгоритма сходни процессу обучения человека в котором человек изначально ничего не знает объекте управления, а в дальнейшем после обучения уверенно им управляет.

3.4.3 Управление многомерным Н-процессом

Отметим что при управлении многомерным Н-процессом задающие воздействия, определяющие желаемые значения компонент вектора выхода, не могут быть выбраны произвольно, как это принято в теории управления. Это обусловлено тем, что, возможно задать такой вектор задающих воздействий, что $\prod_{i=1}^k \Omega_i^H(\vec{x}^*) = \emptyset$, иными словами это задающее воздействие не достижимо по всем компонентам вектора \vec{x}^* . В связи с этим необходимо сначала выбрать достижимые (согласованные) задающие воздействия $\vec{x}^* \in \prod_{i=1}^k \Omega_i^H(\vec{x}^*)$, то есть определить $x_1^*, x_2^*, \dots, x_k^*$.

Предлагается следующий способ решения поставленной задачи:

1. Рассчитать значение $\sum_{i=1}^s \prod_{j=1}^k \Phi\left(\frac{x_j^* - x_{ij}}{c_{sj}}\right)$, где \vec{x}^* – задающее воздействие, s – объем выборки наблюдений;
2. Если рассчитанное значение $\sum_{i=1}^s \prod_{j=1}^k \Phi\left(\frac{x_j^* - x_{ij}}{c_{sj}}\right)$ не равно 0, то это значит, что задающее воздействие достижимо, иначе считаем, что задающее воздействие не может быть достигнуто.

Напомним, что в непараметрическом дуальном управлении расчет поискового шага $\delta\vec{u}_{s+1}$ производится по формуле (3.4.4). В случае управления Н-процессом с несколькими выходными переменными описанный метод расчета поискового шага $\delta\vec{u}_{s+1}$ не подходит, так как помимо приведения объекта к заданию, необходимо учитывать, что входное воздействие должно принадлежать области $\prod_{j=1}^k \Omega^H(\vec{u})$.

Исходя из описанной особенности, для расчета $\delta\vec{u}_{s+1}$ предлагается ис-

пользовать алгоритм с наказанием случайностью.

Модифицированный алгоритм непараметрического дуального управления имеет следующий вид:

- 1) рассчитывается $\vec{u}_s(\vec{x}^*)$;
- 2) выбирается случайный вектор $\delta\vec{u}_{s+1}$;
- 3) рассчитывается величина $\vec{u}_{s+1}(x^*)$;
- 4) если $\vec{u}_{s+1} \in \Omega^H(\vec{u})$, то используем \vec{u}_{s+1} в качестве управляющего воздействия, иначе возвращаемся к шагу 2;
- 5) если $\sum_{i=1}^k |x_{i,s+1} - x_i^*| < \sum_{i=1}^k |x_{i,s} - x_i^*|$, то в качестве следующей величины поискового шага $\delta\vec{u}_{s+2}$ необходимо взять $\delta\vec{u}_{s+1}$, иначе, опять в качестве $\delta\vec{u}_{s+2}$ выбирается случайный вектор;
- 6) Возвращаемся к шагу 1.

Длина вектора $\delta\vec{u}$ равна $m|x^* - x_s|$, где m – настраиваемый коэффициент.

Предложенный алгоритм использовался для управления Н-процессом, имеющего следующую структуру:

$$\begin{cases} x_1 = u_1 + u_2 + 2 \\ x_2 = -2u_1 + 3u_2 + 1 \end{cases} \quad (3.4.6)$$

$$u_2 = \sin\left(\frac{u_1}{2}\right) + \xi \quad (3.4.7)$$

Вычислительные эксперименты были проведены для 3 случаев:

1. когда алгоритм не имеет обучающей выборки;
2. когда алгоритм имеет малую обучающую выборку;
3. когда алгоритм имеет большую обучающую выборку.

Для начала рассмотрим первый случай, в котором объем обучающей выборки $s=0$ (Рисунок 3.3):

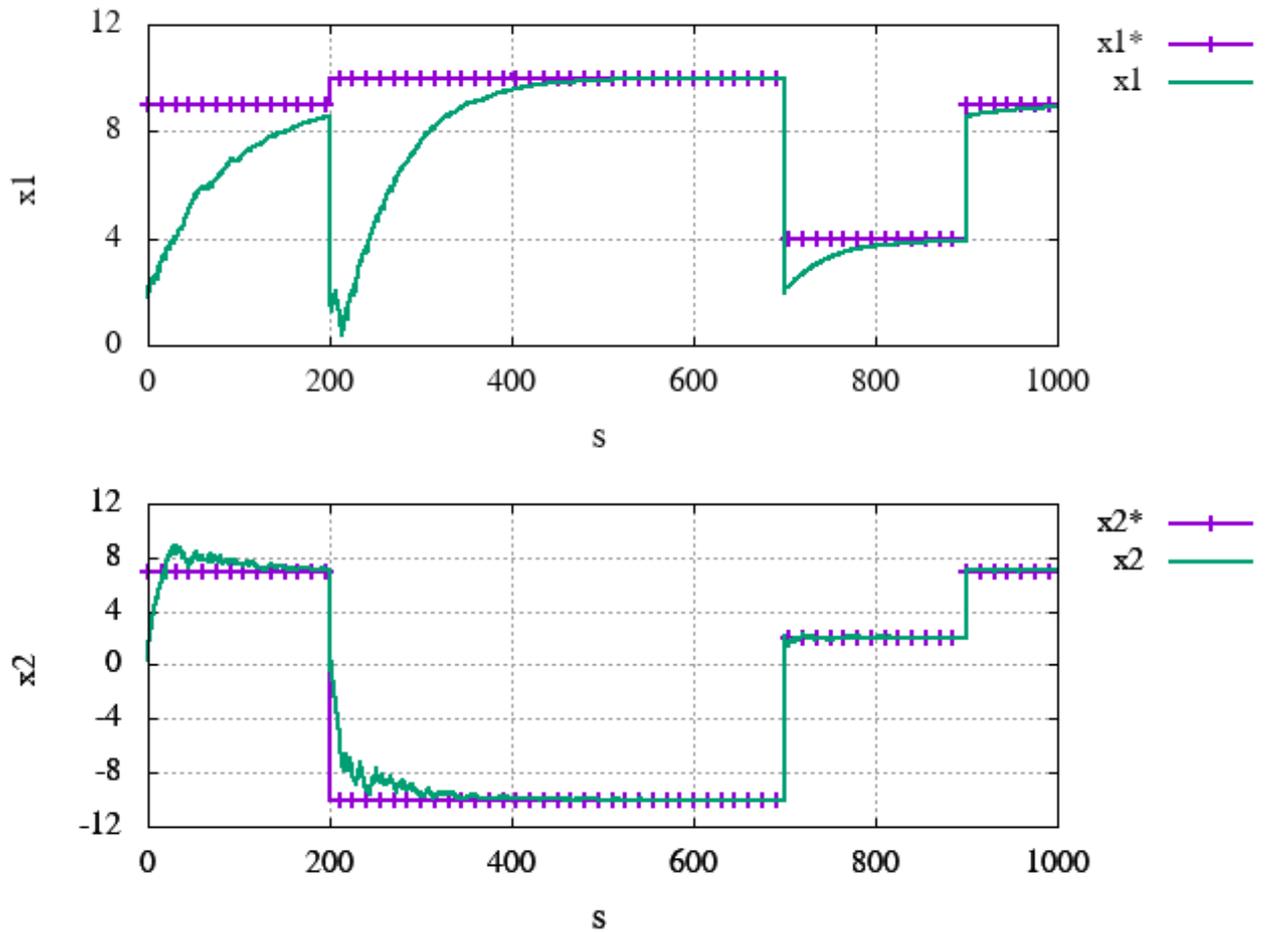


Рисунок 3.3 — Процесс управления в случае, когда алгоритм не имеет обучающей выборки

Как можно увидеть по рисунку, в начале алгоритм довольно грубо управляет объектом, но затем улучшает свои характеристики за счет накопившегося опыта.

Рассмотрим второй случай, $s=20$ (Рисунок 3.4):

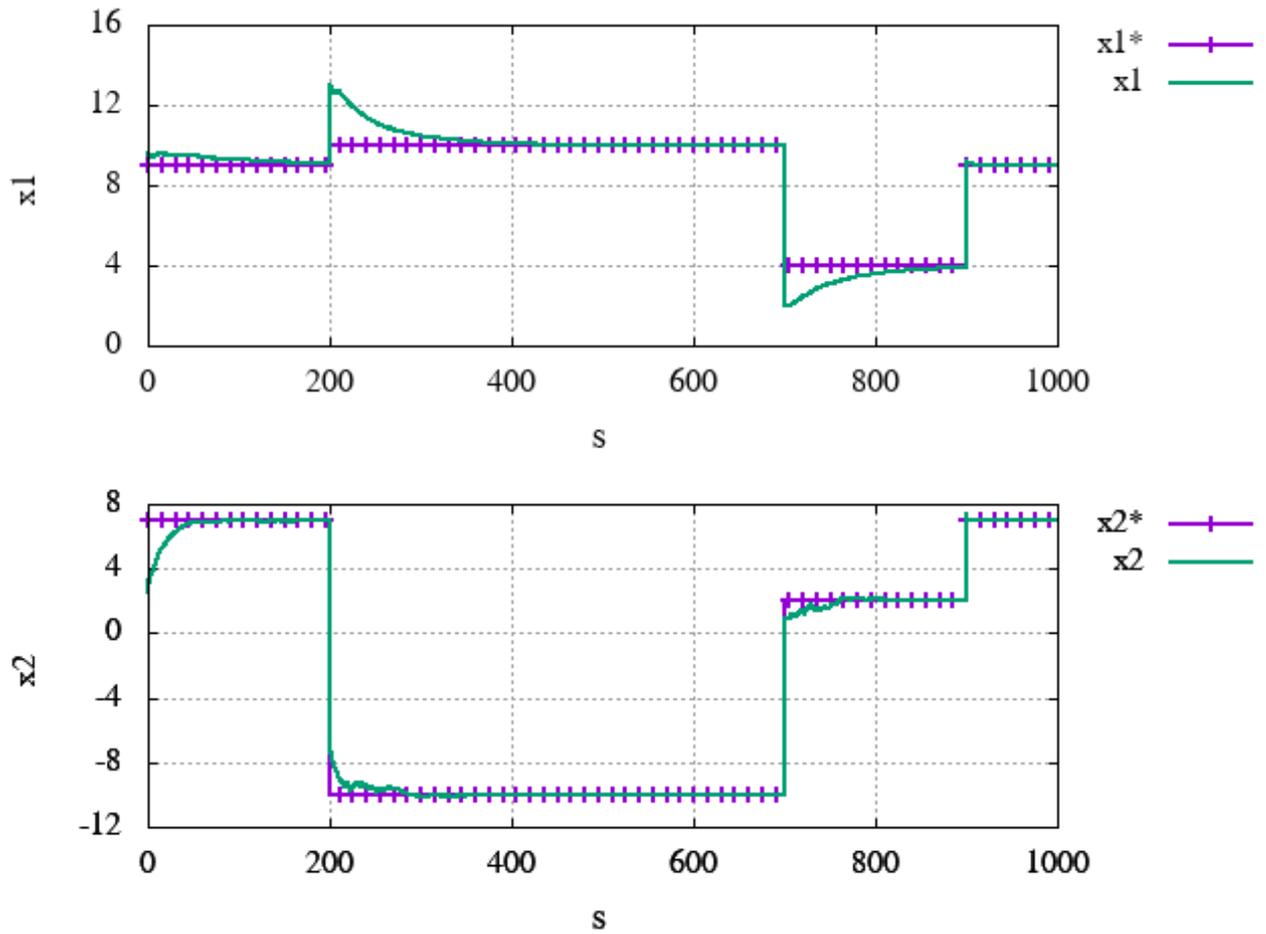


Рисунок 3.4 — Процесс управления в случае когда алгоритм имеет небольшой объем обучающей выборки

Как видно на рисунке 3.4, качество управления резко улучшилось, при наличии даже небольшой обучающей выборки.

Рассмотрим последний случай, $s=100$ (Рисунок 3.5)

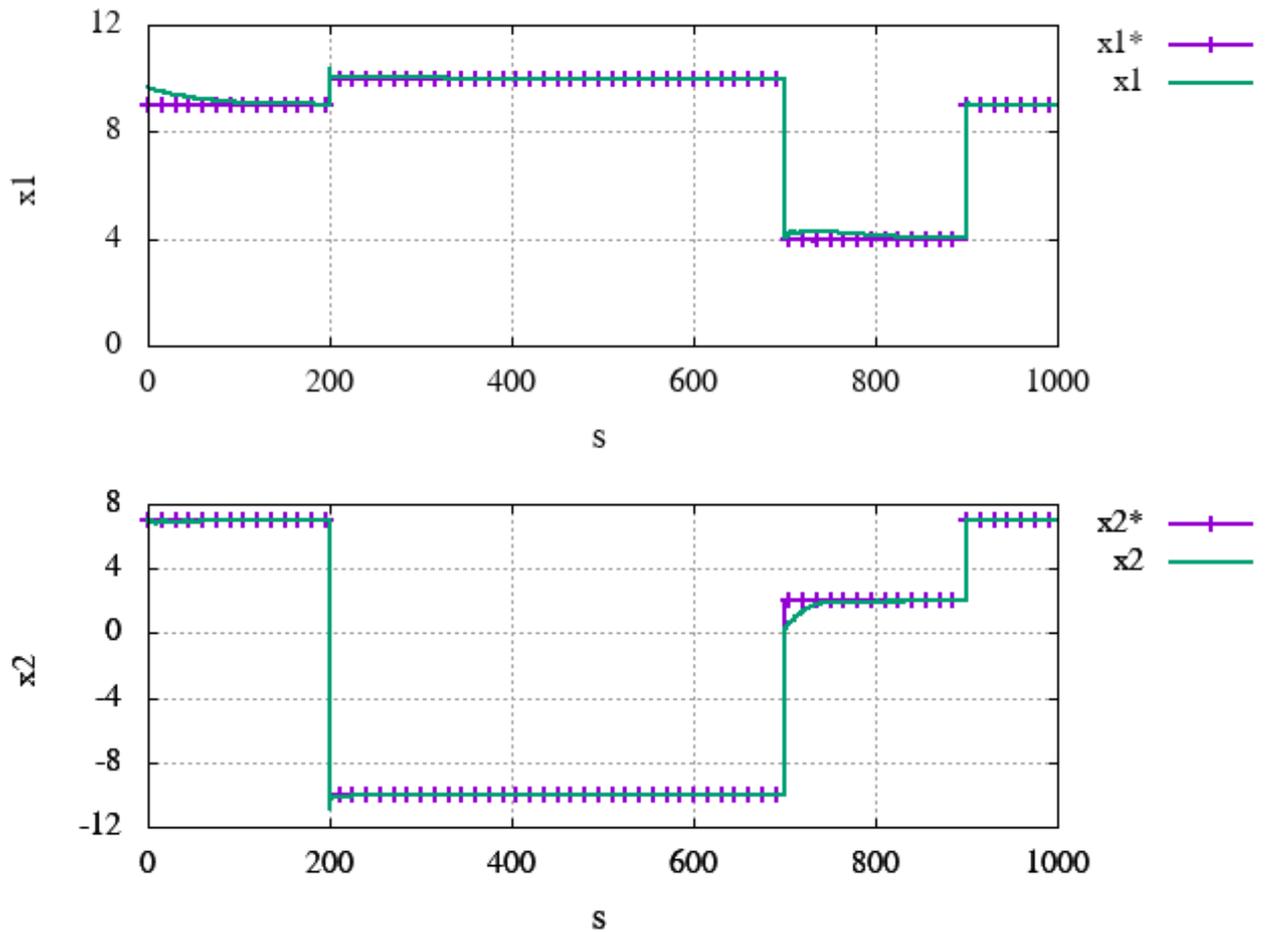


Рисунок 3.5 — Процесс управления в случае когда алгоритм имеет большой объем обучающей выборки

Как и предполагалось, большой объем обучающей выборки, привел к точному управлению N -процессом.

Были проведены многочисленные вычислительные эксперименты, по управлению различными многомерными N -процессами, при помощи предложенного алгоритма. Проведенные эксперименты подтвердили, что предложенная модификация непараметрического дуального управления позволяет успешно управлять многомерным N -процессом.

3.5 Выводы по главе

В данной главе рассмотрен вопрос управления безынерционными стохастическими процессами.

Важно отметить что при управлении N -процессом задающие воздействия, определяющие желаемые значения компонент вектора выхода не могут быть выбраны произвольно из области определения задающих воздействий, как это принято в теории управления. В этом случае возникает задача нахождения множества допустимых значений для определения задающих воздействий. В главе предложен способ определения областей выбора (областей достижимости) задающих воздействий, при управлении многомерным N -процессом.

Рассмотрены стандартные алгоритмы управления (П, ПИ и ПИД регуляторы), их достоинства и недостатки.

Следующий рассмотренный алгоритм управления, алгоритм дуального управления, который объединяет управление объектом и его изучение. К сожалению, из-за того что данный метод требует большого объема априорной информации, его использование для управления реальными процессами сопряжено с существенными трудностями.

Вследствие вышеописанной проблемы была разработана модификация алгоритма дуального управления — алгоритм непараметрического дуального управления. У данного алгоритма гораздо меньшие требования к объему априорной информации об исследуемом процессе вследствие чего использование данного алгоритма для решения реальных задач существенно проще.

Проведен сравнительный анализ между алгоритмом непараметрического дуального управления и ПИ регулятором. Было продемонстрировано, что алгоритм непараметрического дуального управления после обучения быстрее приводит объект к заданию чем ПИ регулятор.

Также был изучен вопрос управления Н-процессом с несколькими выходными переменными. Был предложен способ нахождения поискового шага, учитывающий границу протекания процесса. Продемонстрировано, что предложенный алгоритм успешно управляет Н-процессом с несколькими выходными переменными.

4 Применение методов непараметрической статистики на практических задачах

В главе рассматривается применение методов непараметрической статистики на практике. Решалась задача восстановления полей распределения различных примесей в воздухе города Красноярск. Было признано что Красноярск занимает одно из первых мест в стране по уровню загрязненности воздуха. Вследствие этого разработка систем моделирования, позволяющих восстановить поля распределения различных примесей в воздухе является весьма актуальным направлением.

В качестве источников данных использовались стационарные и передвижные пункты наблюдения за экологической обстановкой. В главе предложена модель объединяющая информацию от всех источников данных, в единую систему моделирования.

Внедрение подобной системы позволит решить задачу позиционирования новых стационарных пунктов наблюдения за экологической обстановкой, а также позволит проложить маршруты для передвижных пунктов наблюдения за экологической обстановкой. Решение этих задач, позволит улучшить качество восстанавливаемых полей распределения различных примесей в воздухе города Красноярск.

4.1 Постановка задачи экологического мониторинга

Увеличение антропогенной нагрузки на окружающую среду требует использования разносторонней и детальной информации о её состоянии. Такая информация нужна для оценки качества атмосферного воздуха, прогноза его изменения, организации рациональной системы природопользования. Репрезентативность наблюдений за состоянием загрязнения атмосферы в городе

зависит от правильности расположения пунктов отбора проб на обследуемой территории.

Программы наблюдений формируются по принципу выбора приоритетных (подлежащих первоочередному определению) загрязняющих веществ и интегральных (отражающих группу явлений, процессов, веществ) характеристик.

Всего постов наблюдения за загрязнением атмосферы в г. Красноярске 8. Для построения качественной модели 8 источников данных недостаточно. Несмотря на это было решено провести исследование с целью выяснить, возможно ли восстановить поля распределения примесей в таких условиях, используя методы непараметрической статистики.

Увеличение числа пунктов наблюдения также не всегда приводит к статистически значимому улучшению оценки качества воздуха.

Сложный рельеф местности, а также число и размещение основных промышленных и транспортных источников, расстояние от них, метеорологические условия и т.д. могут существенным образом повлиять на распределение концентраций поллютантов в пределах рассматриваемой зоны и сделать его существенно неравномерным. Для решения этих задач проводят математическое моделирование распространения примесей в атмосферном воздухе, которое позволяет оценить степень загрязнения атмосферы в заданной точке местности, не проводя в ней натурных измерений.

4.2 Алгоритм непараметрического моделирования экологической обстановки

Данные об экологической обстановке в г. Красноярске получены с постов наблюдения за загрязнением атмосферы Среднесибирского УГМС. Их распределение по территории города отображено на рисунке 4.1.

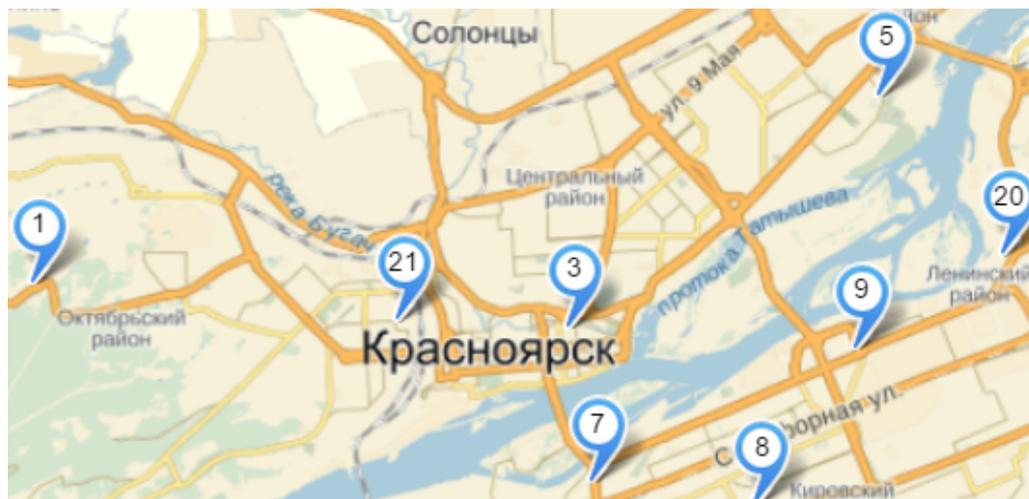


Рисунок 4.1 — Распределение постов наблюдения по территории города
Красноярск

На постах выполняются измерения концентраций следующих веществ: SO₂, сульфат, CO, NO₂, NO, H₂S, CS₂, Фенол, HF, Cl₂, HCl, NH₃, Бензол, Ксилол, Толуол, Этилбензол. Также определяются температура и влажность воздуха, направление и скорость ветра. Отбор проб осуществляется 4 раза в сутки: 01.00, 07.00, 13.00 и 19.00.

Существующие модели экологической обстановки используют данные, которые сложно (а иногда и невозможно) получить при решении реальных задач. Для примера можно привести алгоритм расчета концентрации загрязняющего вещества при выбросе из одиночного точечного источника из Приказа Министерства природных ресурсов и экологии Российской Федерации от 06.06.2017 № 273 «Об утверждении методов расчетов рассеивания выбросов вредных (загрязняющих) веществ в атмосферном воздухе»:

$$c_m = \frac{A \cdot M \cdot F \cdot t \cdot n \cdot \eta}{H^2 \cdot \sqrt[3]{V_t \cdot \Delta T}}, \quad (4.2.1)$$

где c_m – концентрация загрязняющего вещества при выбросе из одиночного точечного источника; A – коэффициент, зависящий от температурной страти-

фикации атмосферы, определяющий условия горизонтального и вертикального рассеивания загрязняющих веществ в атмосферном воздухе; M – масса загрязняющего вещества, выбрасываемого в атмосферный воздух в единицу времени (мощность выброса), г/с; F – безразмерный коэффициент, учитывающий скорость оседания загрязняющего вещества в атмосферном воздухе; m и n – безразмерные коэффициенты, учитывающие условия выброса из устья источника выброса; η – безразмерный коэффициент, учитывающий влияние рельефа местности; H – высота источника выброса, м; V_t – расход газовой смеси, м³/с; ΔT – разность между температурой выбрасываемой газовой смеси и температурой атмосферного воздуха °С. Значения некоторых переменных, указанных в формуле (4.2.1), сложно получить при решении реальных задач, например значения переменной M – массы загрязняющего вещества, выбрасываемого в атмосферный воздух в единицу времени источником. В связи с этим предлагается разработать новые модели восстановления полей распределения различных примесей в воздухе.

В качестве обучающей выборки была принята выборка последних двух месяцев, чтобы отследить разницу между временами года.

В диссертации предложена модель восстановления полей распределения различных примесей в атмосферном воздухе города Красноярск, основанная на методах непараметрической статистики. В связи с тем, что источников данных всего 8, есть вероятность, что при восстановлении поля распределения примесей в воздухе возникнет неопределенность типа «0/0». Для того чтобы избежать данной неопределенности, было решено в качестве колообразной функции для широты и долготы измерения использовать гауссову функцию (4.2.3) которая не обращается в 0.

Таким образом, модель будет представлять собой следующий вид:

$$x(\vec{u}, \vec{c}_s, t) = p(\vec{u}) \frac{\sum_{i=1}^s x_i \Phi_1\left(\frac{u_1 - u_{1,i}}{c_{s1}}\right) \cdot \Phi_1\left(\frac{u_2 - u_{2,i}}{c_{s2}}\right) \cdot \Phi_2\left(\frac{t - t_i}{c_{s3}}\right)}{\sum_{i=1}^s \Phi_1\left(\frac{u_1 - u_{1,i}}{c_{s1}}\right) \cdot \Phi_1\left(\frac{u_2 - u_{2,i}}{c_{s2}}\right) \cdot \Phi_2\left(\frac{t - t_i}{c_{s3}}\right)}, \quad (4.2.2)$$

здесь $u_{1,i}$ — долгота точки измерения; $u_{2,i}$ — широта точки измерения; t_i — время измерения; $i = \overline{1, s}$ — номер измерения; s — размер выборки наблюдений; Φ_1 — гауссова функция (4.2.3); Φ_2 — треугольная функция (4.2.4):

$$\Phi_1\left(\frac{u - u_i}{c_s}\right) = \frac{1}{\sqrt{2 * \pi}} * \exp\left(-\left(\frac{u - u_i}{c_s}\right)^2 / 2\right) \quad (4.2.3)$$

$$\Phi_2\left(\frac{u - u_i}{c_s}\right) = \begin{cases} 1 - \left|\frac{u - u_i}{c_s}\right|, & \left|\frac{u - u_i}{c_s}\right| < 1 \\ 0, & \left|\frac{u - u_i}{c_s}\right| > 1 \end{cases} \quad (4.2.4)$$

Также были учтены направление и скорость ветра. Изменяя коэффициенты размытости ядер c_{s1} и c_{s2} , возможно менять форму колокообразной функции, под который будут попадать точки для построения модели, например, его можно вытянуть вдоль осей u_1 или u_2 , в зависимости от скорости ветра. Направление можно учитывать поворотом координатной сетки на нужный угол, тогда координаты пунктов наблюдения пересчитываются по следующим формулам:

$$\begin{aligned} u'_1 &= u_1 * \cos(\theta) + u_2 * \sin(\theta); \\ u'_2 &= u_2 * \cos(\theta) - u_1 * \sin(\theta), \end{aligned} \quad (4.2.5)$$

где u'_1 и u'_2 - координаты точек, в системе координат, повернутой на угол θ .

4.3 Результаты моделирования экологической обстановки

С помощью предложенной модели выполнены сценарные расчеты содержания СО в атмосферном воздухе г. Красноярска. Оксид углерода часто используют в качестве прогнозируемого параметра в моделях качества атмосферного воздуха, в силу его химической инертности. Время нахождения

в воздухе этого вещества без вступления в химические реакции достигает 5 дней.

Распределение концентраций СО по территории города Красноярска в разные периоды по результатам математического моделирования представлено на рисунках 4.2-4.3.

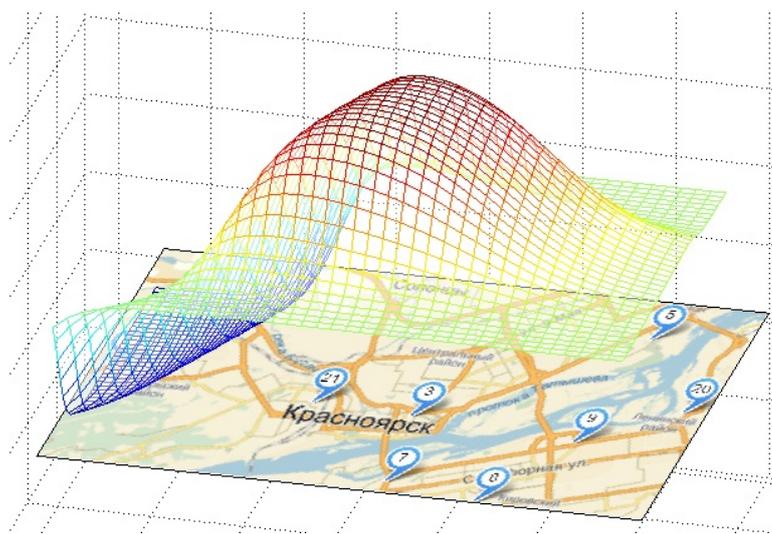


Рисунок 4.2 — Пространственное распределение СО в г. Красноярск 14.01.2004 в 13.00 (расчетные данные)

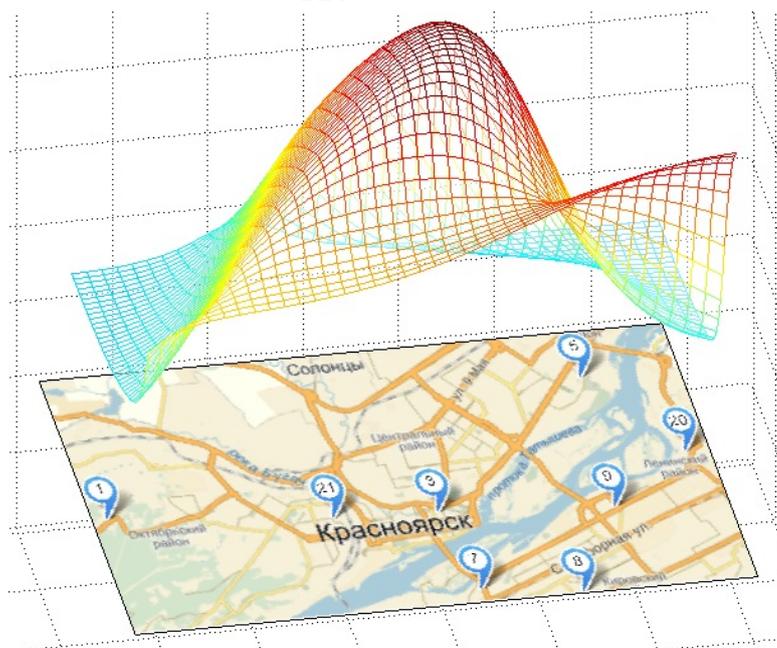


Рисунок 4.3 — Пространственное распределение CO в г. Красноярск 15.01.2004 в 19.00 (расчетные данные)

Согласно выполненным расчетам, наибольшие уровни загрязнения наблюдаются в центральном районе города, что объясняется наибольшей плотностью в центре города автотранспортных потоков, являющихся основным источником выбросов оксида углерода.

Также были сделаны анализы распространения других веществ в воздухе для подтверждения универсальности алгоритма.

Распределение концентраций оксида азота по территории города Красноярска, в разные периоды, по результатам математического моделирования, представлено на рисунках 4.4-4.5.

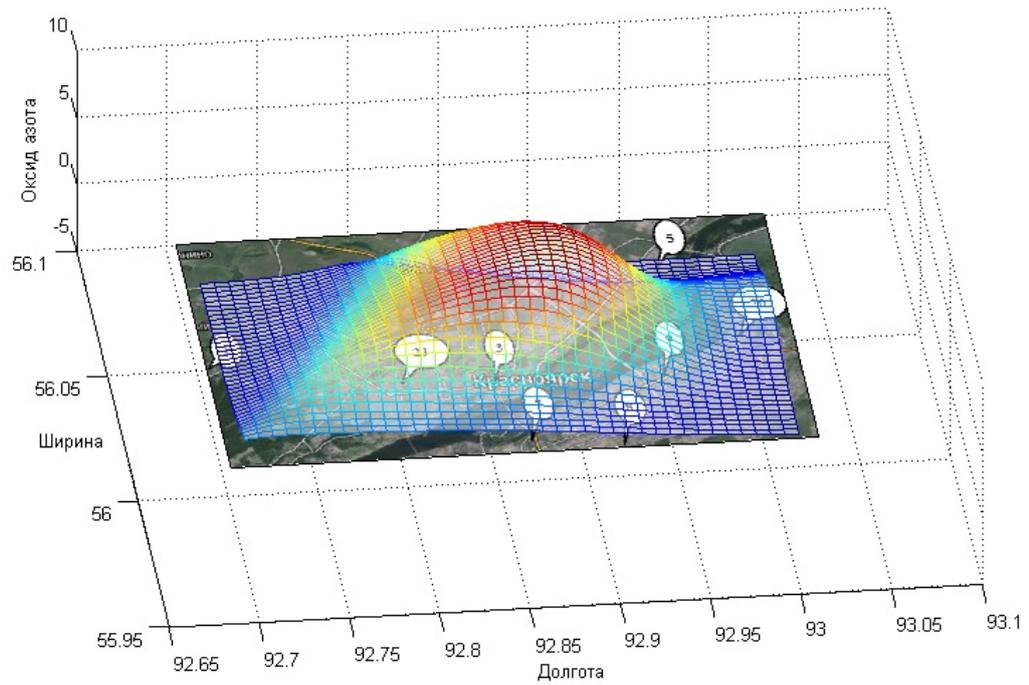


Рисунок 4.4 — Пространственное распределение оксида азота в г. Красноярск 04.06.2004 в 19.00 (расчетные данные)

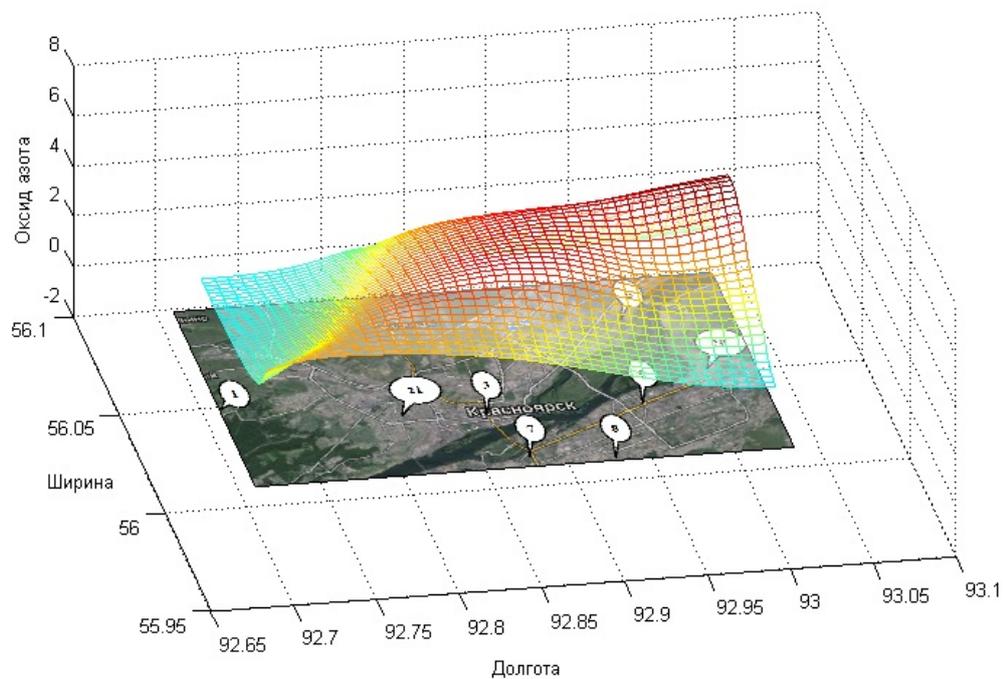


Рисунок 4.5 — Пространственное распределение оксида азота в г. Красноярск 15.01.2004 в 20.00 (расчетные данные)

Распределение концентраций формальдегида по территории города Красноярска в разные периоды по результатам математического моделирования, представлено на рисунках 4.6-4.7.

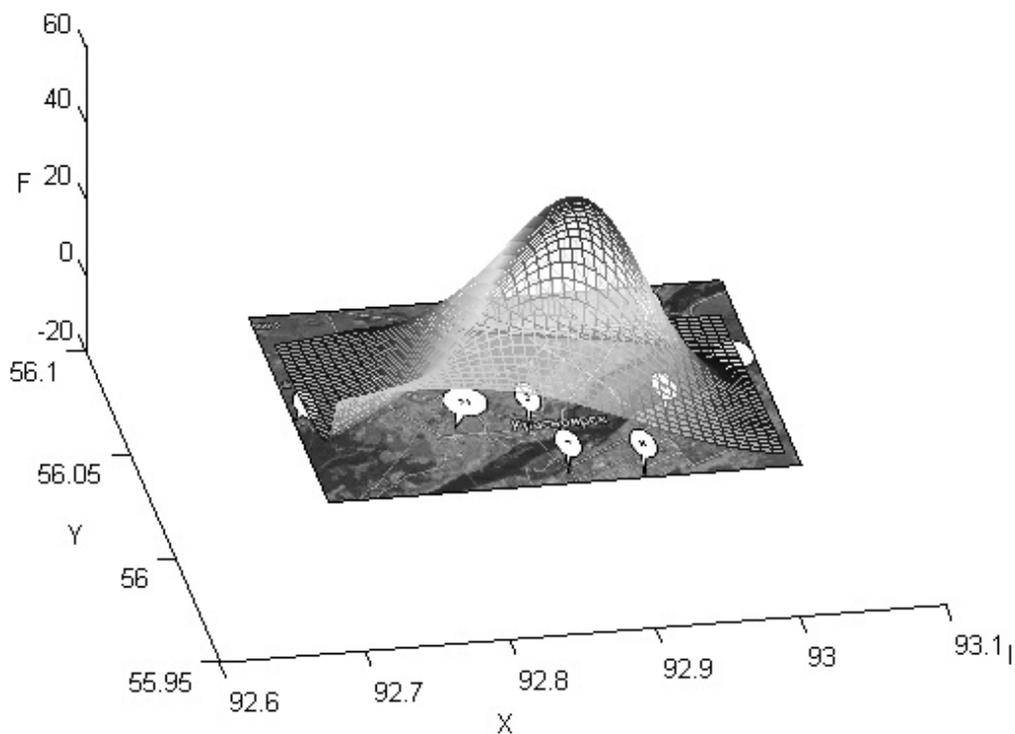


Рисунок 4.6 — Пространственное распределение формальдегида в г. Красноярск 23.09.2004 в 7.00 (расчетные данные)

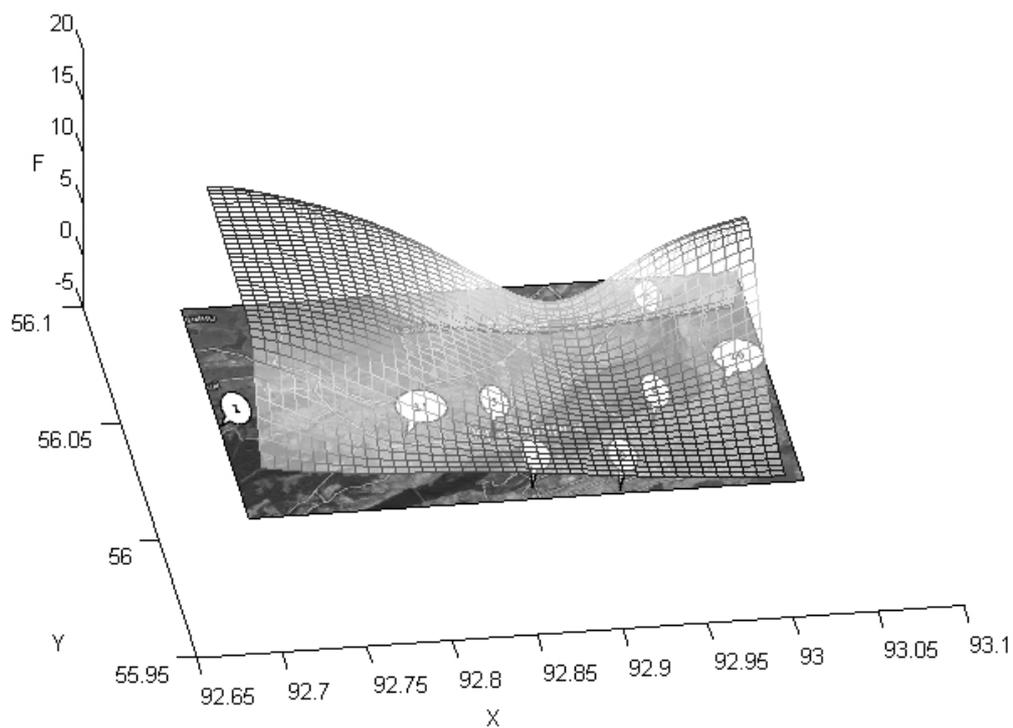


Рисунок 4.7 — Пространственное распределение формальдегида в г. Красноярск 25.09.2004 в 10.00 (расчетные данные)

Распределение концентраций пыли по территории города Красноярска в разные периоды по результатам математического моделирования, представлено на рисунках 4.8-4.9.

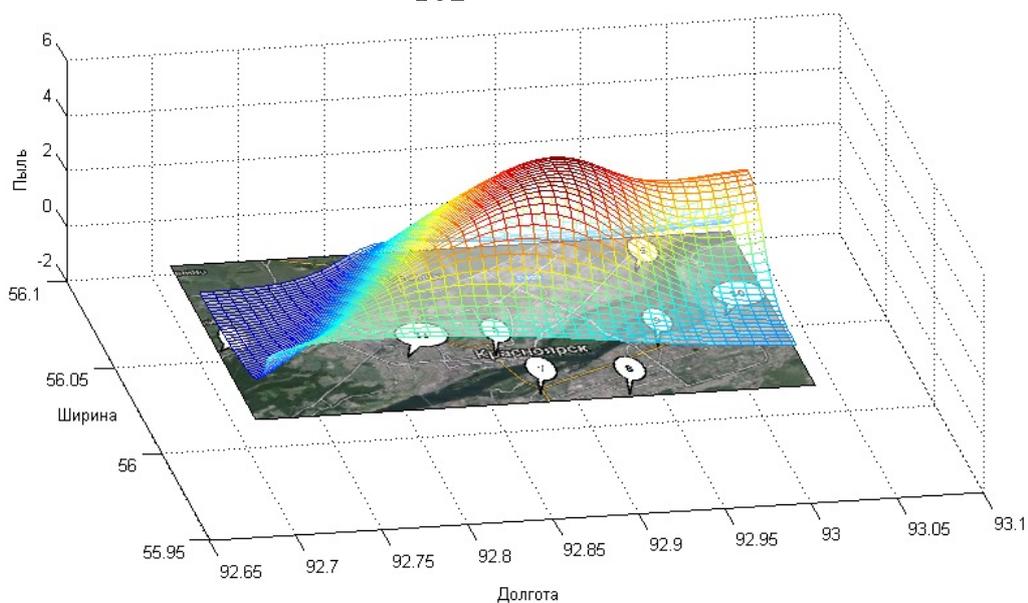


Рисунок 4.8 — Пространственное распределение пыли в г. Красноярск
13.08.2004 в в 9.00 (расчетные данные)

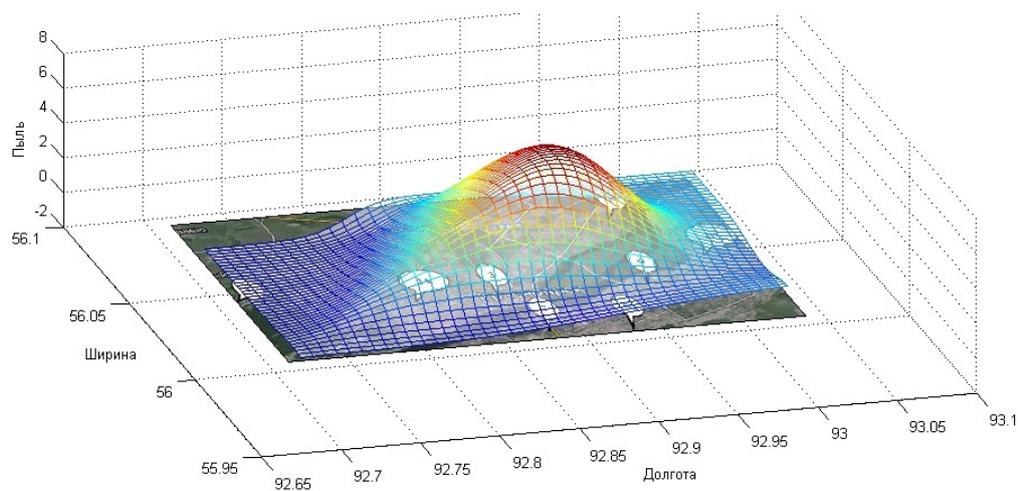


Рисунок 4.9 — Пространственное распределение пыли в г. Красноярск
06.04.2004 в 13.00 (расчетные данные)

Результаты расчетов согласуются с данными натуральных наблюдений. Это позволяет сделать вывод, что данная модель в общей мере отражает экологическую обстановку в городе.

В ходе моделирования, установлено что изменение экологической обстановки проходит циклично. Так, например, наибольший уровень CO достигает в 13.00, наименьший в 7.00.

Указанная особенность позволяет отследить выбросы примесей в атмосферу при помощи следующего алгоритма: если измеренное значение отличается от рассчитанной модели, в большую сторону, на заданный порог (задаваемый экспертом), тогда можно считать, что в рассматриваемой точке зарегистрирован выброс примесей в атмосферу.

Проведен сравнительный анализ содержания CO в атмосферном воздухе г. Красноярска между 2004 и 2012 годами. Среднее содержание CO в атмосферном воздухе отражены на рисунках 4.10 и 4.11

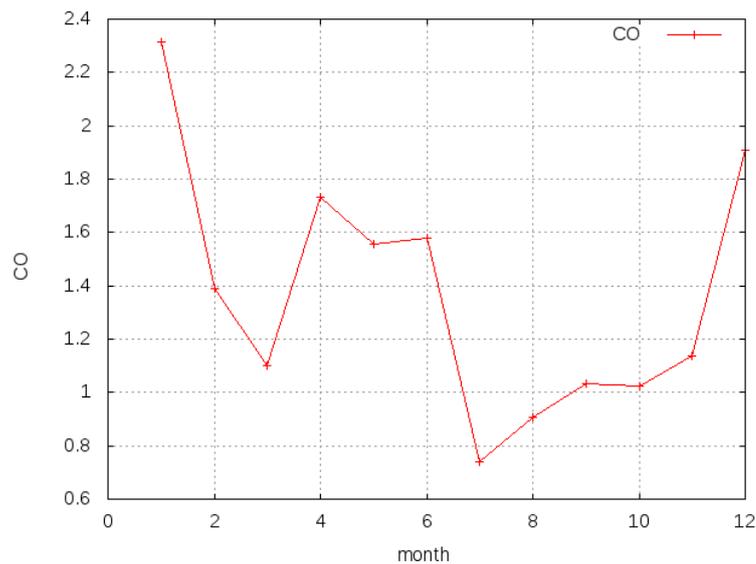


Рисунок 4.10 — Среднее содержание CO в атмосферном воздухе в городе Красноярске в 2004 году, по месяцам

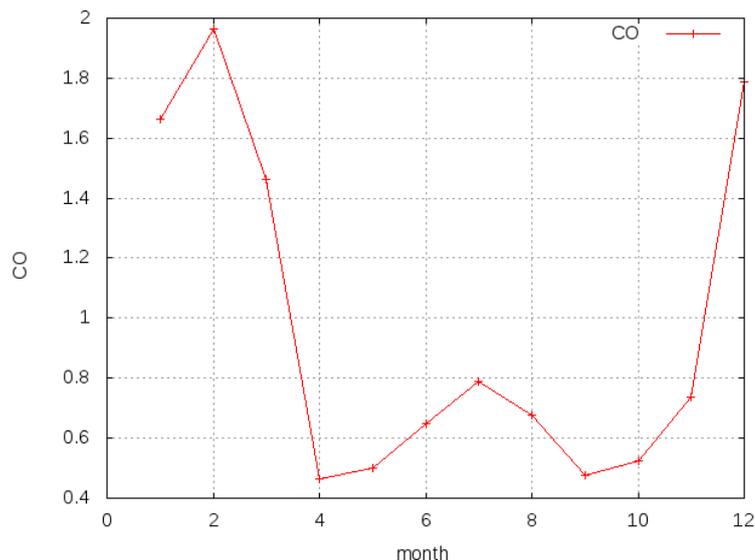


Рисунок 4.11 — Среднее содержание CO в атмосферном воздухе в городе Красноярске в 2012 году, по месяцам

На рисунках 4.10 и 4.11 ось абсцисс означает номер месяца, ось ординат - усредненный уровень содержания CO в месяце.

Как можно увидеть из рисунков 4.10 и 4.11, средний уровень содержания CO в атмосферном воздухе в 2012 году в сравнении с 2004 годом снизился на 28,8 %.

В то же время был проведен подобный сравнительный анализ содержания формальдегида в атмосферном воздухе г. Красноярска между 2004 и 2012 годами. Распределение содержания формальдегида в атмосферном воздухе по месяцам отражены на рисунках 4.12 и 4.13

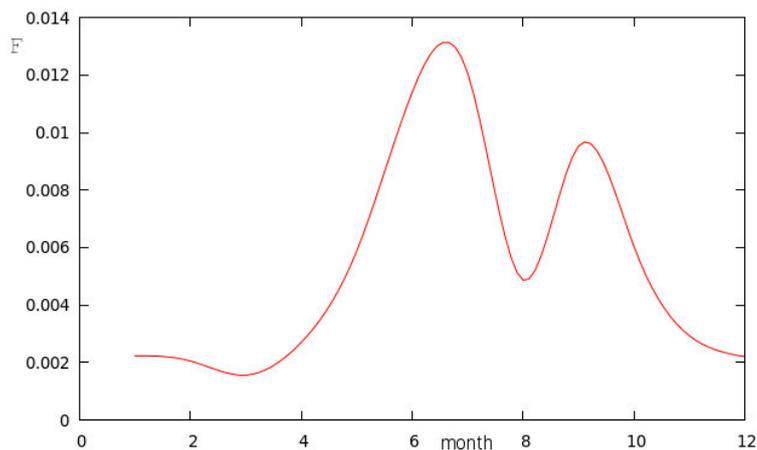


Рисунок 4.12 — Среднее содержание формальдегида в атмосферном воздухе в городе Красноярске в 2004 году, по месяцам

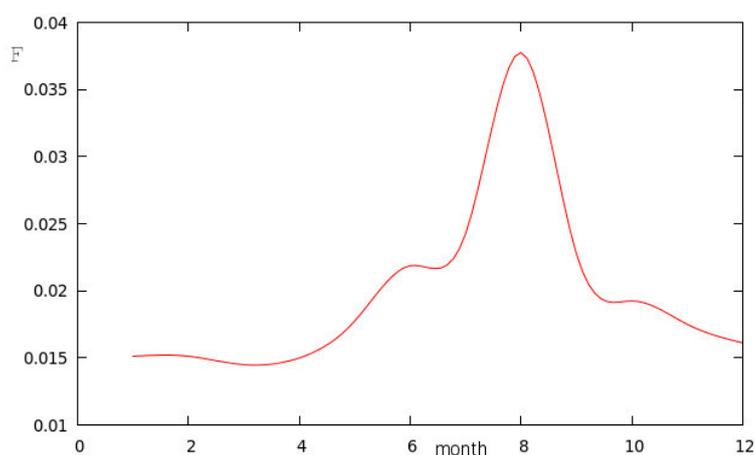


Рисунок 4.13 — Среднее содержание формальдегида в атмосферном воздухе в городе Красноярске в 2012 году, по месяцам

Наблюдается увеличение значения содержания формальдегида в атмосферном воздухе города Красноярска в 3,7 раз.

4.4 Результаты применения разработанных методов для управления котлоагрегатами ТЭЦ

Результаты диссертационной работы используются при создании системы автоматизации котлоагрегатов ТЭЦ-1. Был получен акт о применении на производстве результатов диссертационной работы (Приложение А).

Внедрение адаптивной системы управления позволит более рационально вести процесс сжигания угля в котлоагрегате и корректировать режимы введения процесса (температура питательной воды, температура воздуха и т.д.), при известных контролируемых неуправляемых параметрах (характеристики топлива), с целью достижения заданных значений температуры, давления перегретого (острого) пара и др. Проведенные экспериментальные расчеты и построенные модели показали, что применение разработанных алгоритмов управления позволит повысить эффективность сжигания угля в котлоагрегате, так как основные характеристики приближены к значению, требуемому инструкцией по эксплуатации котла БКЗ-320-140 и режимной картой котла. В связи с чем уменьшаются потери от механического недожога и улучшаются экологические показатели предприятия.

4.5 Выводы по главе

Глава посвящена применению на практике методов моделирования, основанных на непараметрической статистике. Рассмотрена задача восстановления полей распределения различных примесей, в воздухе города Красноярска.

В главе рассмотрены существующие алгоритмы восстановления полей распределения различных примесей в воздухе и их недостатки, обоснована необходимость в разработке новых методов восстановления полей.

Предложен способ восстановления полей распределения различных при-

месей в воздухе, на основе методов непараметрической статистики. Предложенный способ работает в условиях непараметрической неопределенности, что важно для задач с экологической направленностью. Это обосновано тем, что восстановить параметрическую структуру экологического процесса представляет существенные трудности ввиду обилия влияющих факторов.

Представлены результаты восстановления полей распределения для разных примесей в разное время.

Результаты моделирования соответствуют результатам натурных наблюдений, что позволяет удостовериться в работоспособности предложенных моделей.

В главе проведен сравнительный анализ содержания различных примесей между 2004 и 2012 годами. Сравнение производится по количеству примесей СО и формальдегида. В результате сравнения было выявлено, что объем СО в воздухе уменьшился, в то время как объем формальдегида наоборот - увеличился.

В конце главы описаны результаты применения разработанных методов для управления котлоагрегатами ТЭЦ.

Заключение

Разработан новый алгоритм нахождения существенных переменных. Алгоритм позволяет отсеять переменные, не влияющие или мало влияющие на процесс, и строить модель только по существенным переменным. Это позволяет строить модель процесса в пространстве меньшего количества переменных. Предложенный алгоритм основан на методах непараметрической статистики.

Разработана модификация алгоритма моделирования, объединяющая параметрическое моделирование и индикаторную функцию, основанную на методах непараметрической статистики. Данная модификация позволяет моделировать многомерные процессы, с зависимыми входными переменными.

Модифицирован алгоритм непараметрического дуального управления. Модифицированный алгоритм является адаптивным, что позволяет ему не только управлять процессом, но и изучать его. Модифицированный алгоритм был применен для управления многомерными H -процессами. В ходе управления было продемонстрировано, что модифицированный алгоритм непараметрического дуального управления успешно приводит многомерный H -процесс к заданию.

Построена модель экологической обстановки в г. Красноярске, позволяющая восстановить поля различных примесей в воздухе города. В отличие от существующих моделей экологической обстановки, предложенная модель построена при помощи методов непараметрической статистики. Построенная модель позволяет выявить выбросы примесей в атмосферу. Было проведено сравнение содержания различных примесей в воздухе между 2004 и 2011 годами.

Список литературы

1. Антонов А.В. Системный анализ : учебник / А.В. Антонов. – Москва : Высшая школа, 2004. – 454 с.
2. Медведев, А.В. Анализ данных в задаче идентификации / А.В. Медведев // Минск: Сборник научных статей Международной конференции компьютерного анализа данных и моделирования ком. Том 2\\, БГУ, 1995. – 201-207 с.
3. Введение в математическое моделирование: учебное пособие / под ред. П.В.Трусова. Москва: Логос, 2005. – 440 с.
4. Советов Б.Я, Яковлев С.А. Моделирование систем: учебник для вузов / Б.Я. Советов, С.А. Яковлев. – Москва: Высшая школа, 2001. – 343 с.
5. Фельдбаум А.А. Основы теории оптимальных автоматических систем/ А.А. Фельдбаум . – Москва: Физматгиз, 1963г. – 552 с.
6. Надарая Э.А. Непараметрические оценки плотности вероятности и кривой регрессии / Э.А. Надарая. Тбил. ун-т, 1983. – 194 с.
7. Медведев А.В. Непараметрические системы адаптации. / А.В. Медведев – Новосибирск: Наука, 1983. – 174 с.
8. Цыпкин Я.З. Адаптация и обучение в автоматических системах/ Я.З. Цыпкин. – Москва: Наука, 1968г. – 400 с.
9. Теория систем и системный анализ: учебное пособие / под ред. А.Н.Тырсына. Челябинск: Знания, 2002 – 128 с.
10. Медведев А.В. Н-модели для безынерционных систем с запаздыванием/ А.В. Медведев \\Красноярск: Вестник СибГау. – 2012. – №5. – С 35 – 42.
11. Эйкхофф П. Основы идентификации систем управления/ П. Эйкхофф. – Москва : Мир, 1975 – 683с.
12. Француз А.Г. О влиянии корреляции между признаками на информативность для распознавания образов / А.Г. Француз // Изв. АН СССР. Сер. Техн. Кибернетика. 1964. - № 4. - С. 68-77.

13. Минский М., Пейперт С. Перцептроны / Минский М., Пейперт С. - Москва: Мир, 1971. — 261 с.
14. Загоруйко Н.Г. Когнитивный анализ данных. - Новосибирск: Гео, 2013.— 186 с.
15. Абдушукуров А.А. Статистика неполных данных (асимптотическая теория оценивания для неклассических моделей) / А.А. Абдушукуров. – Ташкент: Университет, 2009. – 271 с.
16. Агафонов Е.Д. Об исследовании непараметрических оценок производной кривой регрессии / Е.Д. Агафонов, Н.А. Медведева // Сб. научных трудов «Информатика и системы управления», Красноярск: изд-во КГТУ. – 1996. – с. 176-182.
17. Адаптивные модели сложных технологических процессов / Лапко А.В., Медведев А.В., Николаев А.Г. и др. // Адаптивные системы и их приложения. – Новосибирск: Наука, 1978. – С. 143-158.
18. Вероятность и математическая статистика: Энциклопедия / Под.ред. Ю.В. Прохорова. - М.: Большая Российская энциклопедия, 2003. – 912с.
19. Загоруйко Н.Г. Прикладные методы анализа данных и знаний / Н.Г. Загоруйко. – Новосибирск: Издательство ИМ СО РАН, 1999. – 264 с.
20. Калман Р.Е. Идентификация систем с шумами. Успехи математических наук / Р.Е. Калман. – М.: «Наука», 1985. – 244 с.
21. Катковник В.Я. Непараметрическая идентификация и сглаживание данных / В.Я. Катковник. – М.: Наука, 1985. – 336 с.
22. Климов Г.П. Теория вероятностей и математическая статистика / Г.П. Климов. – М.: Изд-во МГУ, 1983. – 328 с.
23. Колемаев В.А. Теория вероятностей и математическая статистика / В.А. Колемаев, В.Н. Калинина. – М.: Инфра-М, 1977. – 302 с.
24. Корн Г. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М.: Наука. 1968. – 720 с.

25. Кошкин Г.М. Непараметрическая идентификация стохастических объектов / Г.М. Кошкин, И.Г. Пивен. – Хабаровск: РАН Дальневосточное отделение, 2009. – 336с.
26. Крамер Г. Стационарные случайные процессы. Свойства выборочных функций и их приложения/Г. Крамер, Лидбеттер М. – М.: Мир, 1969.–405с.
27. Лбов Г.С. Методы обработки разнотипных экспериментальных данных / Г.С. Лбов. – Новосибирск: Наука, Сиб. отд-ние, 1981. – 160 с.
28. Липцер Р.Ш. Статистика случайных процессов / Р.Ш. Липцер, А.Н. Ширяев. – М.: Наука, 1974. – 696 с.
29. Льюнг Л. Идентификация систем / Л.Льюнг. – М.: Наука, 1991. – 423 с.
30. Медведев А.В. Н-модели для безынерционных систем с запаздыванием // Вестник Сибирского государственного аэрокосмического университета имени академика М.Ф.Решетнева. – 2012. – №5 (45). – С. 84-89.
31. Медведев А.В. Адаптация в условиях непараметрической неопределенности / А.В. Медведев // Адаптивные системы и их приложения. – Новосибирск: Наука, 1978. – С. 4-34.
32. Медведев А.В. Анализ данных в задаче идентификации / А.В. Медведев // Компьютерный анализ данных моделирования. – Минск: БГУ, 1995. Т. 2. – С. 201-206.
33. Медведев А.В. Информатизация управления: учебное пособие / А.В. Медведев. – Красноярск: САА, 1995. – 80с.
34. Медведев А.В. Непараметрические системы адаптации / А.В. Медведев. – Новосибирск: Наука, 1983. – 173с.
35. Медведев А.В. Теория непараметрических систем. К-модели / А.В. Медведев // Вестник Сибирского государственного университета имени академика М.Ф. Решетнева. – 2011. – №3(36). – С. 57-62.
36. Медведев А.В. Теория непараметрических систем. Процессы / А.В. Медведев // Вестник Сибирского государственного аэрокосмического университета имени академика М.Ф.Решетнева. – 2010. – №3 (29). – С. 4-9.

37. Медведев А.В. Теория непараметрических систем. Управление – II / А.В. Медведев // Вестник Сибирского государственного университета имени академика М.Ф. Решетнева. – 2013. – №3(49). – С.85-90.
38. Медведев А.В. Элементы теории непараметрических систем управления / А.В. Медведев // Актуальные проблемы информатики, прикладной математики и механики. Часть 3, Информатика. – Новосибирск-Красноярск: СО РАН, 1996. – С. 87-112.
39. Медведев, А.В. Теория непараметрических систем. Моделирование / А.В. Медведев // Вестник сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева. – 2010.– №4(30). – С. 4-9.
40. Медведев, А.В. Теория непараметрических систем. Общий подход / А.В. Медведев // Вестник сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева. – 2008. – №4 (30). – С. 4-9.
41. Кендалл М. Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976.
42. Кибернетика ожидаемая и кибернетика неожиданная: сборник. - М: Наука, 1968
43. Мадельброт Б. Фрактальная геометрия природы / Ижев, ин-т компьютер. исслед.; Науч.-исслед. центр "Регулярная и хаотическая динамика". - М.; Ижевск, 2010.
44. Математическая теория оптимальных процессов / А.С. Понтрягин, В.Г. Болтянский, Р.В. Гамкрелидзе, Е.Ф. Мищенко. - М.: Наука 1961.
45. Моисеев Н.Н. Люди и кибернетика. - М.: Мол. гвардия, 1984.
46. Надарая Э.А. Замечания о непараметрических оценках плотности вероятности и кривой регрессии // Теория вероятностей и ее применение. - 1970. - Т. 15, вып. 1. - С. 139-142.

47. Надарая Э.А. Непараметрические оценки кривой регрессии // Некоторые вопросы теории вероятностных процессов.-Вып.5. - Тбилиси, 1965.- С.56-68.
48. Ильин В.А. Математический анализ / В.А Ильин, В.А. Садовничий, Бл. Х. Сендов. - М.: Наука. 1979. - 720 с.
49. Медведев А.В. Основы теории адаптивных систем: монография/ А.В. Медведев. – Красноярск: Сиб. гос. аэрокосмич. ун-т., 2015. – 526 с.
50. Параев Ю.А. Адаптация и оптимальное управление. Решение некоторых задач фильтрации динамических сигналов, формирующихся из случайных процессов гауссовского и пуассоновского типов // Адаптивные автоматические системы. - М.: Сов. радио, 1972. - С. 162-182.
51. Срагович В.Г. Адаптивное управление. -М.: Наука, 1976
52. Эльясберг П.Е. Измерительная информация: сколько ее нужно? Как ее обрабатывать? - М.:Наука, 1983.
53. Тасейко О.В., Сугак Е.В. Репрезентативность пунктов наблюдения при оценке качества воздуха в городской среде. - Современные проблемы науки и образования, 2014, № 6, с.1-11.
54. Комиссаров Ю.А. и др. Экологический мониторинг окружающей среды: Учебное пособие для вузов в 2 т. Т.1 / Ю.А. Комиссаров, Л.С. Гордеев, Ю.Д. Эдельштейн, Д.П. Вент; Под ред. П.Д. Саркисова.–М.: Химия, 2005.–365 с.
55. Антропов К.М., Казмер Ю.И., Вараксин А.Н. Описание пространственного распределения загрязнения атмосферного воздуха промышленного центра методом LandUseRegression (обзор) // Экологические системы и приборы. – 2010. – № 1. – С. 28-41.
56. Аналитический обзор состояния и загрязнения атмосферного воздуха <http://krasecology.ru/Air/LabReport>. (Дата обращения 14.08.2015).
57. Хардле В. Прикладная непараметрическая регрессия. – М.: Мир, 1993.

58. Васильев В. А., Добровидов А.В., Кошкин Г.М. Непараметрическое оценивание функционалов от распределений стационарных последовательностей. – М.: Наука, 2004..
59. Емельянов В. В., Курейчик В. В., Курейчик В. М. Теория и практика эволюционного моделирования. / Емельянов В. В., Курейчик В. В., Курейчик В. М. — Москва: Физматлит, 2003. — 432 с.
60. Курейчик В. М., Лебедев Б. К., Лебедев О. К. Поисковая адаптация: теория и практика./ Курейчик В. М., Лебедев Б. К., Лебедев О. К. — Москва: Физматлит, 2006. — 272 с.
61. Гладков Л. А., Курейчик В. В., Курейчик В. М. Генетические алгоритмы: Учебное пособие./ Гладков Л. А., Курейчик В. В., Курейчик В. М. — 2-е изд. — Москва: Физматлит, 2006. — 320 с.
62. Гладков Л. А., Курейчик В. В, Курейчик В. М. и др. Биоинспирированные методы в оптимизации: монография./ Гладков Л. А., Курейчик В. В, Курейчик В. М. — Москва: Физматлит, 2009. — 384 с.
63. Букатова И.Л. Эволюционное моделирование и его приложения./ Букатова И.Л. – Москва: Наука 1994. – 232 с.
64. Люггер Дж. Искусственный интеллект. Стратегия и методы решения сложных проблем. – Москва: Изд. дом «Вильямс», 2003. – 864 с.
65. Koza J.R. Genetic Programming.-Cambrige/MA: MIT press, 1994.-836 pp.
66. Скобцов Ю. А. Основы эволюционных вычислений / Ю.А. Скобцов. – Донецк: ДонНТУ, 2008. – 326 с.
67. Медведев А. В. Некоторые замечания к Н – моделям безынерционных процессов с запаздыванием // Вестник СибГАУ. 2014. № 2(54), С.24-34.
68. Рубан А. И. Методы анализа данных: Учебное пособие. Издание 2-е, испр. и доп. Красноярск: ИПЦ КГТУ, 2004. 319 с.
69. Методы классической и современной теории автоматического управления: Учебник в 5-и тт.; 2-е изд., перераб. И доп. Т.3: Синтез регуляторов

- систем автоматического управления / Под ред. К.А. Пупкова и Н.Д. Егупова. – М.: Издательство МГТУ им. Н.Э. Баумана, 2004.-616 с.
70. Объективности ради// Еженедельная газета научного сообщества "Поиск". Москва. Российская академия наук, Издательство "Поиск". № 39 от 30 сентября 2011 г.
71. Айвазян А., Енюков И. С., Мешалкин Л.Д. Прикладная статистика, исследование зависимостей. М.: Финансы и статистика, 1985.
72. Hausdorff. F. Dimension und ausseres Mass. *Matematische Annalen*. 1919. 79. P. 157-179
73. Розенблатт Ф. Принципы нейродинамики. Перцептроны и теория механизмов мозга. Москва: Мир, 1965. 480 с.
74. Айвазян С.А. Прикладная статистика: Основы моделирования и первичная обработка данных / С.А. Айвазян, И.С. Енюков. – М.: Финансы и статистика, 1983. – 471 с.
75. Апраушева Н.Н. Использование непараметрических оценок в регрессионном анализе / Н.Н. Апраушева, В.Д. Конаков // Заводск. лаб. – 1973. – № 5. – С. 556-569.
76. Аффифи А. Статистический анализ: подход с использованием ЭВМ / А. Аффифи, С. Эйзен. – М.: Мир, 1982 – 488 с.
77. Бойко Р. Непараметрические H-модели процесса нагревания / Р. Бойко, Я. Демченко // Труды международной конференции «Applied methods of statistical analysis. Simulations and statistical inference». – 2011. – P.224-236.
78. Большев Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. – М.: Наука. Главная редакция физико-математической литературы, 1983. – 416 с.
79. Боровков А.А. Математическая статистика / А.А. Боровков. – М.: Наука, 1984. – 472 с.
80. Боровков А.А. Математическая статистика. Оценка параметров. Проверка гипотез / А.А. Боровков. – М.: Наука, 1984. – 472 с.

81. Вапник В.Н. Восстановление зависимостей по эмпирическим данным / В.Н. Вапник. – М.: Наука, 1979. – 448 с.
82. Воронов А.А. Основы теории автоматического регулирования и управления. – Учеб. пособие для вузов. М., «Высш. школа», 1977, 519 с.
83. Воронов А.А. Основы теории автоматического управления: автоматическое регулирование непрерывных линейных систем / А.А. Воронов. – М.: Энергия, 1980. – 312 с.
84. Вулих Б.З. Введение в функциональный анализ / Б.З. Вулих. – М.: Наука, 1967. – 416 с.
85. Гнеденко Б.В. Курс теории вероятностей / Б.В. Гнеденко. – М.: Наука, 1988. – 446 с.
86. Гусаров В.М. Статистика / В.М. Гусаров. – М.: ЮНИТИ, 2003. – 463 с.
87. Деврой Л. Непараметрическое оценивание плотности / Л. Деврой, Л. Дьерфи. – М.: Мир, 1988. – 408 с.
88. Диалоговая система оптимизации и принятия решений для управления производственным комплексом с непрерывным характером технологического процесса. В.И.Волков, Б.В.Казаков, А.В.Медведев. Препринт ВЦ СО АН РАН СССР №14, 1985, г.Красноярск, 33с.
89. Ефимов А.С. Решение задачи кластеризации методом конкурентного обучения при неполных статистических данных / А.С. Ефимов // Вестник Нижегородского университета им. Н.И.Лобачевского. – 2010. – №1. – С. 220-225.
90. Ball D. Statistical analysis with missing data / D. Ball // Journal of Marketing Research. – 2003. – Т.40. №3. – Р.374.
91. Eddy W.F. Optimum kernel estimators of the mode / W.F. Eddy // Ann. Math.Statist. – 1980. – V. 8. – P. 870-882.
92. Efroimovich S.Yu. Nonparametric curve estimation. Methods, theory and application / S.Yu. Efroimovich S.Yu. – Berlin, New-York: Springer-Verlag, 1999.

93. Gasser T. Kernel estimation of regression function / T. Gasser, H.G. Muller // Lect. Notes Math. – 1979. – V.757. – P. 23-68.
94. Keesman Karel J. System identification. An introduction / Karel J. Keesman. – London: Springer, 2011. – 351 p.
95. Kubat C.H. Taskin Bofy-fuzzy logic control for the Basic Oxygen Furnace (BOF) / C.H. Kubat // Robotics and Autonomous System. – 2004. – № 49. – P. 193- 205.
96. Marvin L. Brown. Data mining and the impact of missing data / Marvin L. Brown, John F. Kros // Industrial Management and Data Systems. – 2003. – T.103. № 8. – P. 611-621.
97. Medvedeva N.A. Nonparametrical Estimation of Statistical Characteristics in Problem of Modelling / N.A. Medvedeva // Proceedings of the International Conference «Computer Data Analysis and Modeling». – Minsk: BSU, 1995. – p.89 - 93.
98. Myrtyeit I. Analyzing data sets with missing data: an empirical evaluation of imputation methods and likelihood-based methods / I. Myrtyeit, E. Stensrud, U.H. Olsson // IEEE Transactions on Software Engineering, 2001. – T. 27. № 11. – P. 999.
99. Nemirovski A. Topics in non-parametric statistics / A. Nemirovski // Lectures on probability theory and statistics, in Ecole de Saint-Flour 1988. Lecture Notes in Math. – 2000. – P.85-277.
100. Pracasa Rao B.L.S. Nonparametric function estimation / B.L.S. Pracasa Rao. – Orlando: Academic Press, 1983. – 523 p.
101. Sarita Chauhan. Comparative study of BOF Steelmaking Process based on ANFIS and GRNN Model / Chauhan Sarita, Singh Mahendra, Vivek Kumar Meena // International Journal of Engineering and Innovative Technology (IJEIT). – 2013. – Vol. 2 (9). – P. 198-202.
102. Пугачев В.С. Теория вероятностей и математическая статистика В.С. Пугачев. – М.: Наука, 1979. – 496 с.

103. Райбман Н.С. Адаптивные модели в системах управления / Н.С Райбман, В.М. Чадеев. – М.: Сов. радио, 1966. – 159 с.
104. Растрингин Л.А. Адаптация сложных систем / Л.А. Растрингин. – Рига Зинатне, 1981. – 375 с.
105. Тихонов А.Н. Статистическая обработка результатов эксперимента / А.Н. Тихонов, М.В. Уфимцев. – М.: Изд-во Моск. ун-та, 1988. – 176 с.

Приложение А

Утверждаю
Главный инженер
ОАО «Красноярская ТЭЦ-1»
Окладников М.Е.
2016 г.


Акт

о применении на производстве результатов диссертационной работы
Михова Евгения Дмитриевича
на тему: «Непараметрические модели и алгоритмы управления Н-процессами»

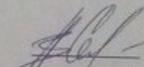
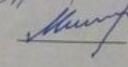
В диссертационной работе аспиранта кафедры «Информационные системы» ФГАОУ ВО «Сибирский Федеральный Университет» Михова Е.Д. разработаны модели и алгоритмы управления процессами в пространстве входных - выходных переменных, имеющих «трубчатую структуру», в условиях параметрической неопределенности. Актуальность проведения исследований связана с тем, что при автоматизации и управлении технологическими процессами априорной информации о них может быть недостаточно для построения параметрических моделей, также возможна ситуация при которой некоторые входные переменные процесса имеют стохастическую зависимость, что усложняет процесс моделирования и управления традиционными способами. В диссертации предложены непараметрические алгоритмы идентификации и управления, которые в некоторых случаях могут быть более эффективны для практики, чем используемые на сегодняшний день.

Результаты диссертационной работы, в частности, непараметрические модели и алгоритмы управления Н-процессами, могут быть использованы при разработке системы автоматизации котлоагрегатов ТЭЦ-1.

На основании исследования производственных данных февраля 2016 года построены непараметрические оценки вероятностных характеристик и регрессионных зависимостей между входными и выходными переменными процесса, идентификатор, характеризующий зависимость некоторых входных переменных, непараметрические модели и алгоритмы управления котлоагрегатом. Внедрение адаптивной системы управления позволит более рационально вести процесс сжигания угля в котлоагрегате и корректировать режимы ведения процесса (температура питательной воды, температура воздуха и т.д.), при известных контролируемых неуправляемых параметрах (характеристики топлива), с целью достижения заданных значений температуры, давления перегретого (острого) пара и др.. Проведенные экспериментальные расчеты и построенные модели показали, что применение разработанных алгоритмов управления позволит повысить эффективность сжигания угля в котлоагрегате, так как основные характеристики приближены к значению, требуемому инструкцией по эксплуатации котла БКЗ-320-140 и режимной картой котла. В связи с чем уменьшаются потери от механического недожога и улучшаются экологические показатели предприятия.

Полученные результаты будут использованы при разработке дополнений к технологическим инструкциям по эксплуатации котла БКЗ-320-140.

Начальник цеха ТАИ
Начальник ПТО

 Селиванов П.В.
 Машуков М.В.