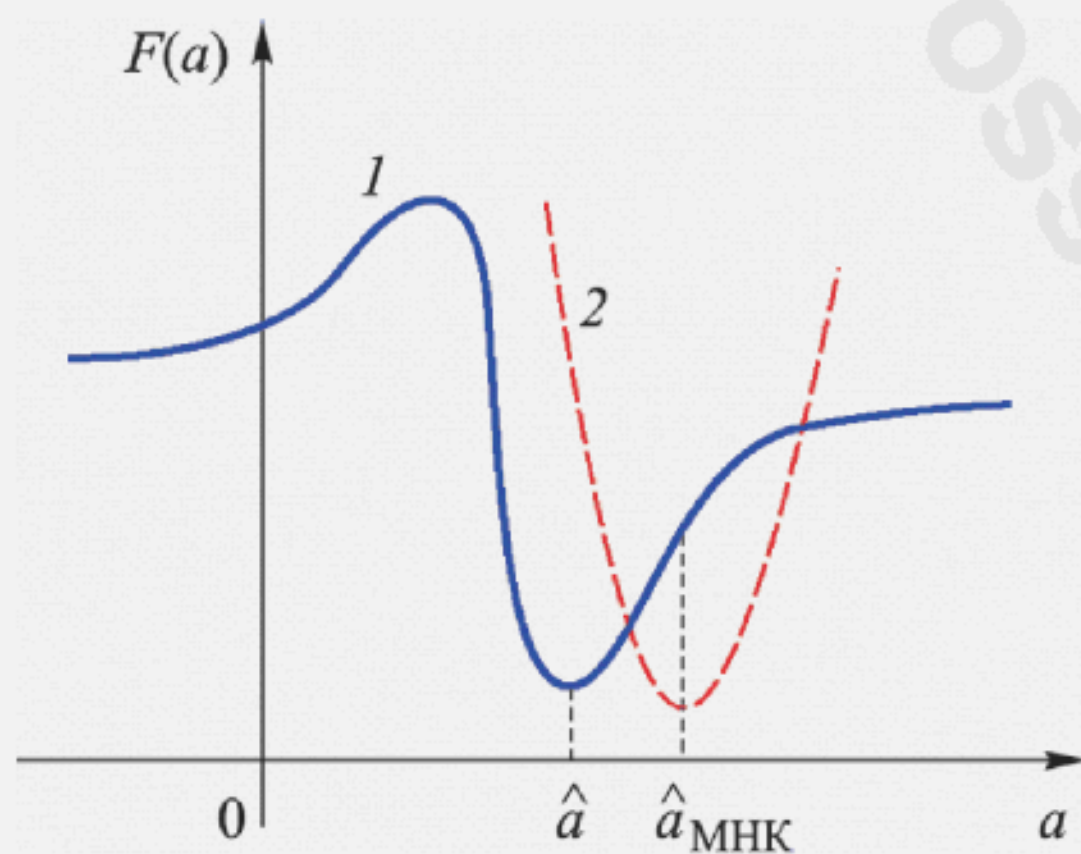


А.А. Грешилов

# КОНФЛЮЭНТНЫЙ АНАЛИЗ

КАК УЧЕСТЬ  
ПОГРЕШНОСТЬ  
ВСЕХ СЛУЧАЙНЫХ  
ВЕЛИЧИН



УДК 621.372.542

ББК 22.18

Г81

Рецензенты

*Б.И. Махартин*, доктор технических наук, профессор Московского государственного технического университета имени Н.Э. Баумана

*П.Н. Антонюк*, кандидат технических наук, доцент Московского государственного университета имени М.В. Ломоносова

**Грешилов А.А.**

**Г81** Конфлюэнтный анализ. Как учесть погрешность всех случайных величин: учебное пособие. – М.: Университетская книга. – 2020. – 136 с.

ISBN 978-5-89699-321-8

Рассматривается класс задач, в которых обрабатываются данные наблюдений и измерений. Показано, что в ходе измерений с необходимостью возникают погрешности. Если они не учитываются, то данные измерений и наблюдений не отражают реальные процессы и явления, а содержат существенные искажения и отклонения. Для их предотвращения используется метод наименьших квадратов, когда в матрицах системы элементы считаются детерминированной величиной. Если все величины, участвующие в решении, являются случайными, то применяются методы конфлюэнтного анализа. Предлагаемые алгоритмы являются основой решения систем уравнений, в которых случайные величины присутствуют в матрицах системы, а также применяются при определении характеристик систем передачи данных и при учете влияния этих характеристик на передаваемую и обрабатываемую информацию. В доступной форме рассмотрены алгоритмы учета погрешностей в исходных данных (элементах матриц систем алгебраических уравнений и правых частях уравнений), приводятся методы оценки свободных параметров различных функций известного вида; примеры, поясняющие особенности алгоритмов, а также примеры, имеющие прикладное значение, включая идентификацию источников радиоактивных изотопов криптона и ксенона, определение координат источников излучения и др. Использован математический аппарат, соответствующий программе технических направлений и специальностей вузов.

Для научных работников и специалистов, занимающихся цифровой и аналоговой обработкой информации. Может использоваться в учебном процессе со студентами, магистрантами и аспирантами высших учебных заведений инженерно-технического профиля.

УДК 621.372.542

ББК 22.18

ISBN 978-5-89699-321-8

© Грешилов А.А., 2020

© Университетская книга, 2020

# ОГЛАВЛЕНИЕ

|  |           |
|--|-----------|
| Введение.....  | 7         |
| <b>Глава 1. Инженерное образование и конфлюэнтный анализ.....</b>                                    | <b>12</b> |
| <b>Глава 2. Методы регрессионного и конфлюэнтного анализа</b>  |           |
| 2.1. Понятие регрессии; основные определения.....  | 19        |
| 2.2. Линейная регрессия $\eta$ на $\xi$ и $\xi$ на $\eta$ .....                                      | 23        |
| 2.3. Регрессионный парадокс.....   | 24        |
| 2.4. Ортогональная регрессия.....  | 26        |
| 2.5. Метод наименьших квадратов. Оценка свободных параметров функций, линейных по параметрам.....    | 28        |
| 2.6. Оценка параметров моделей с помощью функции правдоподобия.....                                  | 34        |
| 2.7. Байесовский подход к оцениванию параметров моделей.....   | 41        |
| 2.8. Интервальная оценка линии регрессии и прогнозируемых значений функции.....                      | 42        |
| <b>Глава 3. Активный и пассивный эксперименты</b>  |           |
| 3.1. Оценивание параметров функции известного вида в пассивном эксперименте.....                     | 46        |
| 3.2. Анализ систем в активном эксперименте.....  | 52        |
| <b>Глава 4. Примеры других методов в конфлюэнтном анализе</b>  |           |
| 4.1. Оценки параметров функции известного вида с учетом ошибок в значениях функций и аргументов..... | 56        |
| 4.2. Оценки свободных параметров в уравнении прямой линии.....                                       | 60        |
| 4.3. О единственности оценок параметров. Состоятельность оценок и алгоритм их получения.....         | 63        |
| 4.4. Оценка параметров многомерной линейной модели.....  | 70        |
| 4.5. Оценка параметров полиномиальной зависимости.....   | 72        |
| 4.6. Оценка значений параметров в сигноме (многомерном полиноме).....                                | 74        |
| 4.7. Построение сглаживающих кубических сплайнов.....  | 77        |
| 4.8. Непосредственное определение оценок параметров экспоненты и функции Гаусса.....                 | 81        |

|  |            |
|--|------------|
| <b>Глава 5. Определение координат источника излучения путем объединения всей информации по пеленгам.....</b> | <b>84</b>  |
| <b>Глава 6. Идентификация источников радиоактивных благородных газов</b>                                     |            |
| 6.1. Источники радиоактивных благородных газов.....  | 89         |
| 6.2. Трудности идентификации источников изотопов криптона и ксенона.....                                     | 97         |
| 6.3. Алгоритм идентификации источников изотопов криптона и ксенона.....                                      | 101        |
| <b>Глава 7. Планирование экспериментов и некорректные задачи</b>   |            |
| 7.1. Планирование экспериментов в конфлюэнтном анализе.....  | 109        |
| 7.2. Некорректные задачи в конфлюэнтном анализе.....   | 115        |
| Литература.....  | 120        |
| <b>Приложение 1. Описание программы «Регрессия» (инструкция для пользователя).....</b>                       | <b>122</b> |
| <b>Приложение 2. Вкладка «Метод наименьших квадратов».....</b>   | <b>129</b> |

Мы как карлики на плечах  
гигантов, и поэтому можем  
видеть больше и дальше, чем они.

*Исаак Ньютон*

## ВВЕДЕНИЕ

Не мы хватаем идею, идея хватает  
и гонит нас на арену, чтобы мы, как  
невольники-гладиаторы, сражались за нее.

*Геррих Гейне*

Данная книга является расширенным и дополненным изданием книги А. А. Грешилова «Метод наименьших квадратов и элементы конфлюэнтного анализа» (Москва, Изд-во МГТУ им. Н. Э. Баумана, 1990). За эти годы произошел грандиозный скачок в развитии персональных компьютеров и программного обеспечения к ним, что позволило решать самые сложные практические задачи, в частности, задачи регрессионного и конфлюэнтного анализа. Любые измерения и наблюдения приводят к случайным величинам. Случайная величина – это не одно число, а функция на множестве реализаций. В процессе измерений невозможно получить детерминированную величину (число): каждый результат содержит неизвестную погрешность (ошибку)<sup>1</sup>. Основным математическим аппаратом теории обработки информации и сигналов является аппарат статистических методов, используемый для анализа и синтеза технических устройств и протекающих процессов, выделения полезного сигнала на фоне помех, измерения статистических характеристик пото-

---

<sup>1</sup> В словаре русского языка С. И. Ожегова погрешность определяется как ошибка. В данной книге оба слова будут применяться в тексте в качестве синонимов. В английском языке ситуация обстоит так же.

ков сигналов, обеспечения надлежащей разрешающей способности в спектрометрических и радиометрических измерениях и т. п. Для создания современной техники, приборов и оборудования требуются большие материальные затраты, поэтому крайне важно увеличить информационность расчетных методов и экспериментов. Последние достижения в области прикладной и фундаментальной математики позволяют с большой эффективностью решать задачи, стоящие перед исследователями.

В процессе анализа сложных (больших) систем возникают проблемы оценки погрешности, вносимой в расчеты при замене неизвестных вероятностных характеристик (параметров) статистическими. Следовательно, в каждом конкретном случае требуется обосновать значения таких показателей, как объемы выборок, границы доверительных интервалов, уровни значимости (доверия). В связи с этим при численных расчетах возникают проблемы выбора аппроксимирующей системы с учетом требуемой точности получаемых результатов и допустимой сложности вычислений, проверки адекватности используемой математической модели изучаемому реальному процессу и оценки коэффициентов чувствительности моделей относительно вариации их параметров.

С точки зрения математической статистики данные проблемы сводятся к задачам аппроксимации исходных данных, разработке методов оценок параметров однофакторных и многофакторных моделей, решению интегральных и линейных алгебраических уравнений и т. п. При этом приходится оперировать с данными, содержащими неопределенность из-за погрешностей измерения, случайных флуктуации или операций усреднения, группировки и округления. Практическая польза от алгоритмов, реализующих перечисленные задачи, определяется тем, насколько адекватна модель реальному процессу и в какой степени учитывается неопределенность исходных данных независимо от ее природы.

На рубеже 1803–1805-х годов математики К. Гаусс и А. Лежандр независимо друг от друга заложили основы метода наименьших квадратов (МНК), ставшего впоследствии базой линейного регрессионного анализа. Но в классическом методе наименьших квадратов невозможно учесть неопределенность всей исходной информации, что

ведет к получению неверных (смещенных) оценок и, как следствие, к неверным практическим выводам. Учесть влияние всех случайных величин позволяет конфлюэнтный анализ – совокупность методов математической статистики, относящихся к анализу априори постулируемых (заданных) функциональных связей между количественными (случайными и неслучайными) переменными в условиях, когда наблюдаются не сами переменные, а случайные величины, образуемые добавкой случайной погрешности измерений. Если анализируется зависимость одной переменной от остальных, то это задача регрессионного анализа. Естественно, что математические процедуры над случайными величинами более сложные, чем те же процедуры над детерминированными переменными: для случайных величин необходимо *применять их закон распределения*.

С учетом изложенного выше можно сделать следующие заключения:

- реальные задачи содержат не одну, а несколько статистических величин с соизмеримыми дисперсиями, которые необходимо учитывать в процессе обработки исходной информации. При этом каждый исследователь должен показать, как будут отличаться результаты при повторном проведении того же исследования с использованием аналогичных исходных данных, для чего во всех задачах необходимо найти интервальные оценки решения. Тогда можно будет оценить качество получаемого решения и сделать выводы о том, обеспечивает ли исходная информация получение надежного решения, удовлетворяет ли поставленным условиям математический алгоритм решения или требуется повысить точность исходных данных либо искать другие способы решения задачи;

- наличие погрешности в исходных данных приводит к тому, что встречающиеся в практике задачи, решаемые с помощью интегральных уравнений и систем плохо обусловленных алгебраических уравнений, некорректны, т. е. бесконечно малым приращением в исходных данных могут отвечать сколь угодно большие изменения в решении.

Цель данной книги – ознакомить инженеров и исследователей, работающих в различных отраслях, с имеющимися в настоящее время способами учета погрешностей в исходных данных как в коррект-

ных, так и в некорректных (условно корректных) задачах, а также показать способы получения устойчивого решения некорректных задач и интервальных оценок решения.

В главе 1 на простейших примерах введено понятие *конфлюэнтного* анализа. На простом примере – определение параметров прямой линии – рассмотрен метод наименьших квадратов, в котором учитываются погрешности всех исходных данных (в элементах матрицы и в правой части). Рассмотрены регрессионный парадокс и ортогональная регрессия для метода наименьших квадратов.

Глава 2 посвящена методу *регрессионного* анализа, являющегося предтечей конфлюэнтного анализа. Здесь предполагается, что погрешность имеет место только в правых частях уравнений. Исследованы методы определения точечных и интервальных оценок искомых параметров. Рассмотрены физические причины некорректности задач, обусловленность линейных систем, метод наименьших квадратов, оценка параметров моделей с помощью функции правдоподобия, теорема Крамера – Рао, байесовский подход к оцениванию параметров моделей.

В главе 3 приведена общая постановка задачи учета погрешностей и в значениях функции, и в значениях аргументов. В разных работах эта задача имеет различные названия. В данной монографии применен термин «конфлюэнтный анализ», под которым понимается набор алгоритмов, позволяющих с позиций математической статистики находить оценки свободных параметров функции заданного вида (в том числе интегральных уравнений и систем линейных алгебраических уравнений) с учетом погрешностей значений как функции, так и аргументов.

Дано описание пассивной схемы эксперимента в конфлюэнтном анализе, приведена общая постановка задачи оценки параметров функции известного вида. Здесь же рассмотрены вопросы о единственности и несмещенности получаемых оценок и методом статистических испытаний исследованы свойства получаемых оценок.

Изложен метод получения оценок в активной схеме эксперимента: показано, что в этом случае задача сводится к регрессионной.

В главе 4 методы конфлюэнтного анализа конкретизированы для прямой линии, линейных функций и систем линейных алгебраиче-



ных, так и в некорректных (условно корректных) задачах, а также показать способы получения устойчивого решения некорректных задач и интервальных оценок решения.

В главе 1 на простейших примерах введено понятие *конфлюэнтного* анализа. На простом примере – определение параметров прямой линии – рассмотрен метод наименьших квадратов, в котором учитываются погрешности всех исходных данных (в элементах матрицы и в правой части). Рассмотрены регрессионный парадокс и ортогональная регрессия для метода наименьших квадратов.

Глава 2 посвящена методу *регрессионного* анализа, являющегося предтечей конфлюэнтного анализа. Здесь предполагается, что погрешность имеет место только в правых частях уравнений. Исследованы методы определения точечных и интервальных оценок искомых параметров. Рассмотрены физические причины некорректности задач, обусловленность линейных систем, метод наименьших квадратов, оценка параметров моделей с помощью функции правдоподобия, теорема Крамера – Рао, байесовский подход к оцениванию параметров моделей.

В главе 3 приведена общая постановка задачи учета погрешностей и в значениях функции, и в значениях аргументов. В разных работах эта задача имеет различные названия. В данной монографии применен термин «конфлюэнтный анализ», под которым понимается набор алгоритмов, позволяющих с позиций математической статистики находить оценки свободных параметров функции заданного вида (в том числе интегральных уравнений и систем линейных алгебраических уравнений) с учетом погрешностей значений как функции, так и аргументов.

Дано описание пассивной схемы эксперимента в конфлюэнтном анализе, приведена общая постановка задачи оценки параметров функции известного вида. Здесь же рассмотрены вопросы о единственности и несмещенности получаемых оценок и методом статистических испытаний исследованы свойства получаемых оценок.

Изложен метод получения оценок в активной схеме эксперимента: показано, что в этом случае задача сводится к регрессионной.

В главе 4 методы конфлюэнтного анализа конкретизированы для прямой линии, линейных функций и систем линейных алгебраиче-

# Глава 1

## ИНЖЕНЕРНОЕ ОБРАЗОВАНИЕ И КОНФЛЮЭНТНЫЙ АНАЛИЗ

Нужно много учиться,  
чтобы немного знать.  
*Шарль Монтескье*

Целью инженерного образования является подготовка специалистов, которые проводят различные измерения и наблюдения. Другими словами, они оперируют случайными величинами. Любые экспериментальные значения содержат ошибки, присутствие которых может серьезно повлиять на результаты анализа. Однако во многих случаях учет всех погрешностей исходных данных не проводится, исследователи даже не подозревают об имеющихся алгоритмах. Учет все погрешности исходных данных позволяет конфлюэнтный анализ. Но во многих учебниках по математической статистике конфлюэнтный анализ даже не упоминается.

Понятие «конфлюэнтный анализ» появилось в начале 30-х годов прошлого столетия. Но оказалось, что оценка параметров даже прямой линии приводит к системам нелинейных уравнений, решать которые в то время не могли. К тому же укоренились некоторые ошибочные положения: например, что методом наименьших квадратов всегда можно получить несмещенные оценки (!) искомых параметров.

Будем считать, что вид оцениваемых функций априори известен, но надлежит найти точечные и интервальные оценки свободных параметров этих функций, по которым можно определить точечные и интервальные оценки самих функций.

Конфлюэнтный анализ позволяет учитывать влияние всех случайных величин, участвующих в расчете. Применение конфлюэнт-

ного анализа рассмотрим на примере оценки параметров прямой линии методом наименьших квадратов (МНК), широко применяемым в различных приложениях.

Рассмотрим уравнение прямой  $Y = aX + b$ , когда необходимо определить точечные оценки параметров  $a$  и  $b$  и их дисперсии. Если мы имеем дело с детерминированными величинами, то записи  $Y = aX + b$  и  $X = Y/a - b/a$  эквивалентны. Если мы оперируем случайными величинами, то при применении метода наименьших квадратов записи  $Y = aX + b$  и  $X = Y/a - b/a$  при одних и тех же исходных данных приведут к разным результатам. В первой записи  $Y$  – случайная величина, а  $X$  – детерминированная (регрессия  $Y$  на  $X$ , рис. 1.1); для второй записи – все наоборот (регрессия  $X$  на  $Y$ , рис. 1.2).

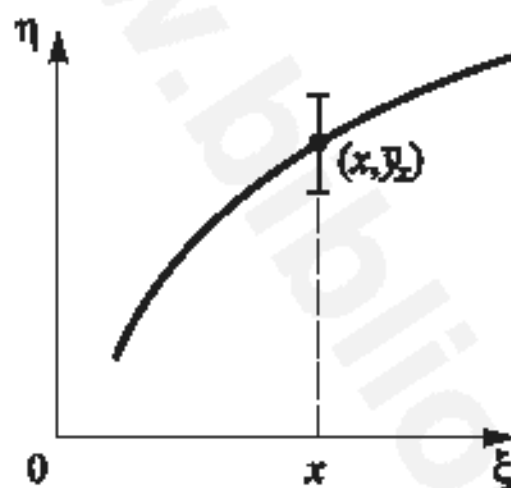


Рис. 1.1. Регрессия  $Y$  на  $X$

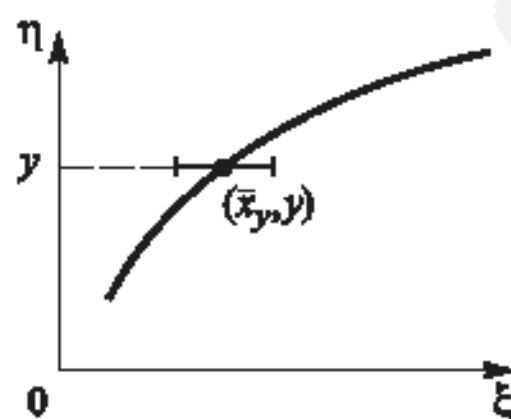


Рис. 1.2. Регрессия  $X$  на  $Y$

Если по одним и тем же результатам наблюдений найти параметры  $a$  и  $b$  и построить в одной системе координат полученные линии  $X$  на  $Y$  и  $Y$  на  $X$ , можно получить разные линии. Это явление называют *регрессионным парадоксом*, хотя никакого парадокса нет: здесь речь идет о разных случайных величинах. На рис. 1.3 приведены три линии регрессии:  $X$  на  $Y$  (1);  $Y$  на  $X$  (2) и ортогональная (3).

ного анализа рассмотрим на примере оценки параметров прямой линии методом наименьших квадратов (МНК), широко применяемым в различных приложениях.

Рассмотрим уравнение прямой  $Y = aX + b$ , когда необходимо определить точечные оценки параметров  $a$  и  $b$  и их дисперсии. Если мы имеем дело с детерминированными величинами, то записи  $Y = aX + b$  и  $X = Y/a - b/a$  эквивалентны. Если мы оперируем случайными величинами, то при применении метода наименьших квадратов записи  $Y = aX + b$  и  $X = Y/a - b/a$  при одних и тех же исходных данных приведут к разным результатам. В первой записи  $Y$  – случайная величина, а  $X$  – детерминированная (регрессия  $Y$  на  $X$ , рис. 1.1); для второй записи – все наоборот (регрессия  $X$  на  $Y$ , рис. 1.2).

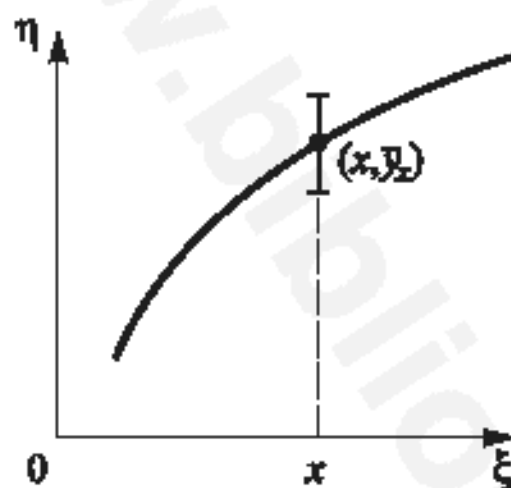


Рис. 1.1. Регрессия  $Y$  на  $X$

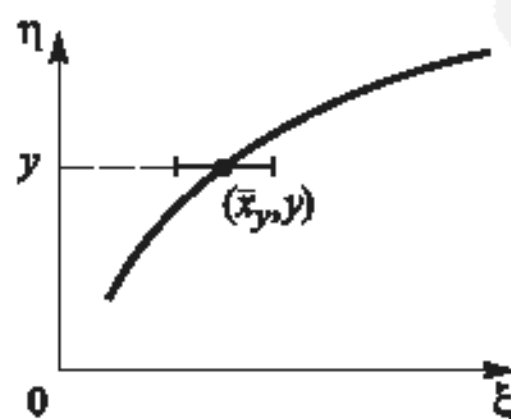


Рис. 1.2. Регрессия  $X$  на  $Y$

Если по одним и тем же результатам наблюдений найти параметры  $a$  и  $b$  и построить в одной системе координат полученные линии  $X$  на  $Y$  и  $Y$  на  $X$ , можно получить разные линии. Это явление называют *регрессионным парадоксом*, хотя никакого парадокса нет: здесь речь идет о разных случайных величинах. На рис. 1.3 приведены три линии регрессии:  $X$  на  $Y$  (1);  $Y$  на  $X$  (2) и ортогональная (3).

ного анализа рассмотрим на примере оценки параметров прямой линии методом наименьших квадратов (МНК), широко применяемым в различных приложениях.

Рассмотрим уравнение прямой  $Y = aX + b$ , когда необходимо определить точечные оценки параметров  $a$  и  $b$  и их дисперсии. Если мы имеем дело с детерминированными величинами, то записи  $Y = aX + b$  и  $X = Y/a - b/a$  эквивалентны. Если мы оперируем случайными величинами, то при применении метода наименьших квадратов записи  $Y = aX + b$  и  $X = Y/a - b/a$  при одних и тех же исходных данных приведут к разным результатам. В первой записи  $Y$  – случайная величина, а  $X$  – детерминированная (регрессия  $Y$  на  $X$ , рис. 1.1); для второй записи – все наоборот (регрессия  $X$  на  $Y$ , рис. 1.2).

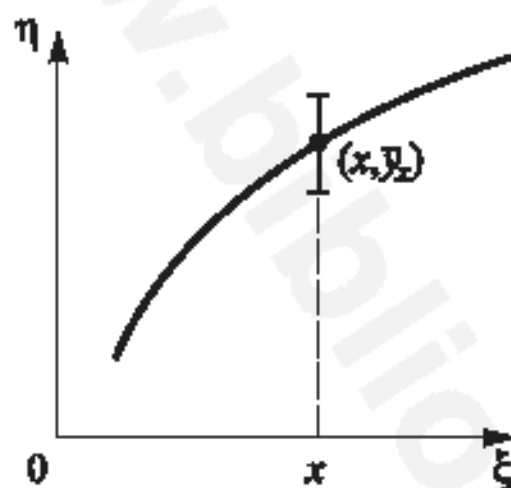


Рис. 1.1. Регрессия  $Y$  на  $X$

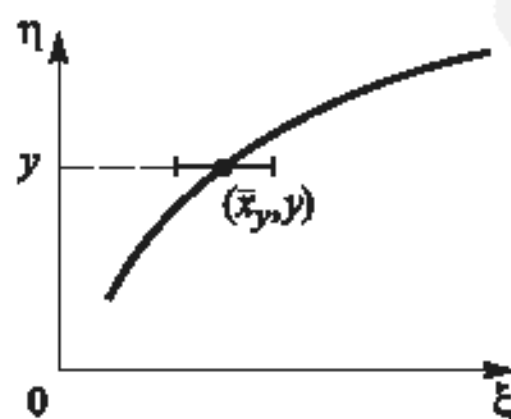


Рис. 1.2. Регрессия  $X$  на  $Y$

Если по одним и тем же результатам наблюдений найти параметры  $a$  и  $b$  и построить в одной системе координат полученные линии  $X$  на  $Y$  и  $Y$  на  $X$ , можно получить разные линии. Это явление называют *регрессионным парадоксом*, хотя никакого парадокса нет: здесь речь идет о разных случайных величинах. На рис. 1.3 приведены три линии регрессии:  $X$  на  $Y$  (1);  $Y$  на  $X$  (2) и ортогональная (3).

ного анализа рассмотрим на примере оценки параметров прямой линии методом наименьших квадратов (МНК), широко применяемым в различных приложениях.

Рассмотрим уравнение прямой  $Y = aX + b$ , когда необходимо определить точечные оценки параметров  $a$  и  $b$  и их дисперсии. Если мы имеем дело с детерминированными величинами, то записи  $Y = aX + b$  и  $X = Y/a - b/a$  эквивалентны. Если мы оперируем случайными величинами, то при применении метода наименьших квадратов записи  $Y = aX + b$  и  $X = Y/a - b/a$  при одних и тех же исходных данных приведут к разным результатам. В первой записи  $Y$  – случайная величина, а  $X$  – детерминированная (регрессия  $Y$  на  $X$ , рис. 1.1); для второй записи – все наоборот (регрессия  $X$  на  $Y$ , рис. 1.2).

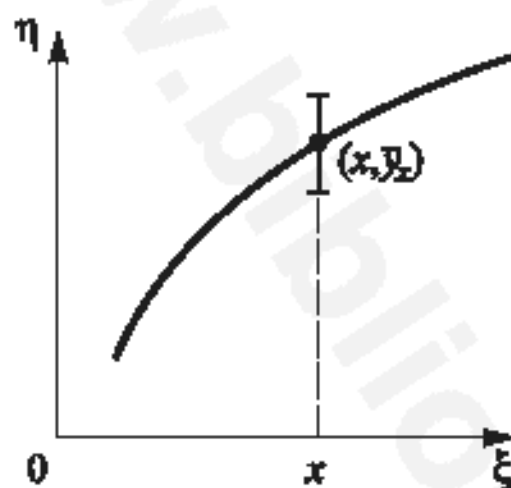


Рис. 1.1. Регрессия  $Y$  на  $X$

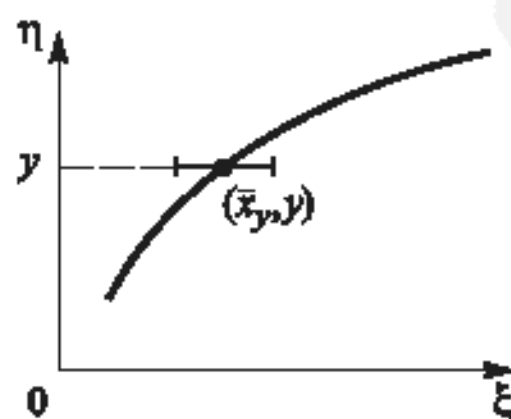


Рис. 1.2. Регрессия  $X$  на  $Y$

Если по одним и тем же результатам наблюдений найти параметры  $a$  и  $b$  и построить в одной системе координат полученные линии  $X$  на  $Y$  и  $Y$  на  $X$ , можно получить разные линии. Это явление называют *регрессионным парадоксом*, хотя никакого парадокса нет: здесь речь идет о разных случайных величинах. На рис. 1.3 приведены три линии регрессии:  $X$  на  $Y$  (1);  $Y$  на  $X$  (2) и ортогональная (3).

ного анализа рассмотрим на примере оценки параметров прямой линии методом наименьших квадратов (МНК), широко применяемым в различных приложениях.

Рассмотрим уравнение прямой  $Y = aX + b$ , когда необходимо определить точечные оценки параметров  $a$  и  $b$  и их дисперсии. Если мы имеем дело с детерминированными величинами, то записи  $Y = aX + b$  и  $X = Y/a - b/a$  эквивалентны. Если мы оперируем случайными величинами, то при применении метода наименьших квадратов записи  $Y = aX + b$  и  $X = Y/a - b/a$  при одних и тех же исходных данных приведут к разным результатам. В первой записи  $Y$  – случайная величина, а  $X$  – детерминированная (регрессия  $Y$  на  $X$ , рис. 1.1); для второй записи – все наоборот (регрессия  $X$  на  $Y$ , рис. 1.2).

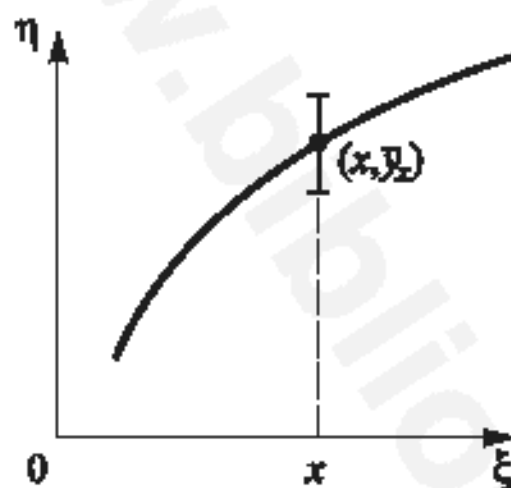


Рис. 1.1. Регрессия  $Y$  на  $X$

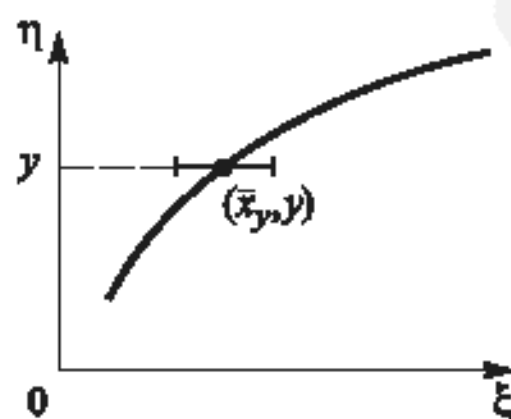


Рис. 1.2. Регрессия  $X$  на  $Y$

Если по одним и тем же результатам наблюдений найти параметры  $a$  и  $b$  и построить в одной системе координат полученные линии  $X$  на  $Y$  и  $Y$  на  $X$ , можно получить разные линии. Это явление называют *регрессионным парадоксом*, хотя никакого парадокса нет: здесь речь идет о разных случайных величинах. На рис. 1.3 приведены три линии регрессии:  $X$  на  $Y$  (1);  $Y$  на  $X$  (2) и ортогональная (3).

ного анализа рассмотрим на примере оценки параметров прямой линии методом наименьших квадратов (МНК), широко применяемым в различных приложениях.

Рассмотрим уравнение прямой  $Y = aX + b$ , когда необходимо определить точечные оценки параметров  $a$  и  $b$  и их дисперсии. Если мы имеем дело с детерминированными величинами, то записи  $Y = aX + b$  и  $X = Y/a - b/a$  эквивалентны. Если мы оперируем случайными величинами, то при применении метода наименьших квадратов записи  $Y = aX + b$  и  $X = Y/a - b/a$  при одних и тех же исходных данных приведут к разным результатам. В первой записи  $Y$  – случайная величина, а  $X$  – детерминированная (регрессия  $Y$  на  $X$ , рис. 1.1); для второй записи – все наоборот (регрессия  $X$  на  $Y$ , рис. 1.2).

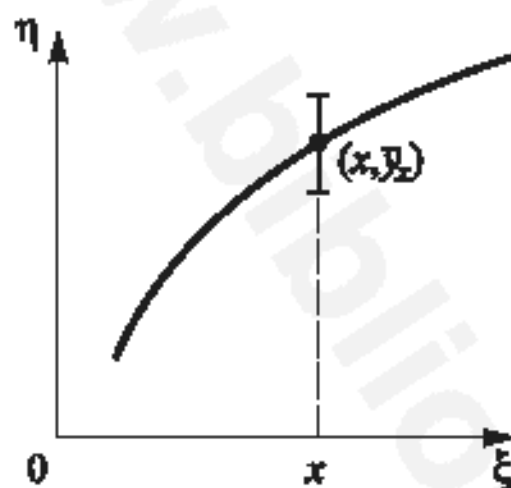


Рис. 1.1. Регрессия  $Y$  на  $X$

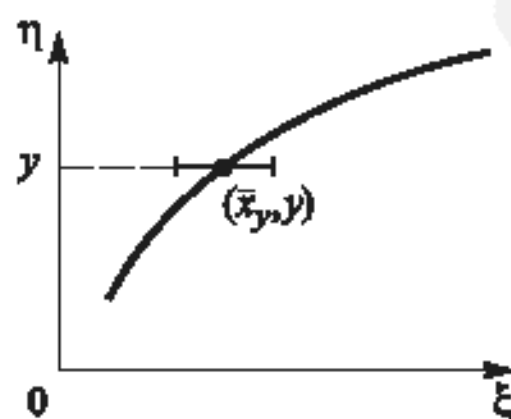


Рис. 1.2. Регрессия  $X$  на  $Y$

Если по одним и тем же результатам наблюдений найти параметры  $a$  и  $b$  и построить в одной системе координат полученные линии  $X$  на  $Y$  и  $Y$  на  $X$ , можно получить разные линии. Это явление называют *регрессионным парадоксом*, хотя никакого парадокса нет: здесь речь идет о разных случайных величинах. На рис. 1.3 приведены три линии регрессии:  $X$  на  $Y$  (1);  $Y$  на  $X$  (2) и ортогональная (3).



## Глава 2

# МЕТОДЫ РЕГРЕССИОННОГО И КОНФЛЮЭНТНОГО АНАЛИЗА

Не разгрызая ореха, ядра не отгадаешь.  
*Русская народная пословица*

### 2.1. Понятие регрессии; основные определения

При определении вида функции плотности вероятности (ФПВ) распознаваемых классов или разделяющих (решающих) функций в качестве исходных данных используют выборки, полученные в результате конкретных экспериментов. И, как любые экспериментальные значения, исходные данные содержат ошибки, присутствие которых может серьезно повлиять на результаты анализа. Поскольку рассматриваемые алгоритмы учета погрешностей в значениях признаков будут применяться к различным функциям: условным плотностям вероятности, разделяющим (решающим) функциям, – измеренные значения признаков обозначим переменной  $X$ , а значения оцениваемых функций –  $y(x)$ . Будем считать, что общий вид оцениваемых функций априори известен, но надлежит найти точные и интервальные оценки свободных параметров этих функций, по которым можно определить точечные и интервальные оценки самих функций.

Модели, позволяющие учитывать ошибки в значениях функций и аргументов, рассмотрены в [1, 3, 4–10, 15–20]. Наиболее часто ставится задача определения оценок  $\theta$  параметров модели  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ ; где  $\varepsilon_i$  – случайная ошибка, имеющая нормальное распределение с параметрами  $M(\varepsilon_i) = 0$ ,  $D(\varepsilon_i) = \sigma_i^2$ ,  $D(\varepsilon_i, \varepsilon_j) = 0$ ,  $i, j = 1, \dots, n$ . Данная постановка является классической регрессионной задачей, которая решается методом максимума правдоподобия (ММП) или методом наименьших квадратов (МНК).

Если дисперсия наблюдаемых значений  $\sigma_i^2 = \sigma^2$ ,  $i = 1, \dots, n$ , и значение  $\sigma^2$  не задано, то оценку параметра  $\sigma$  можно найти по формуле  $\hat{\sigma}^2 = \frac{s}{n-p}$ , где  $s = \sum_{i=1}^n [y_i - f(x_i, \hat{\theta})]^2$ ;  $p$  – число параметров;  $n$  – число точек наблюдений.

Регрессионный анализ предполагает, что переменные  $X$  являются детерминированными – измерены с погрешностью много меньшей, чем погрешность значений функции. На практике это требование очень часто невыполняется, поэтому возникает необходимость учета погрешностей аргумента  $X$ . Если погрешности аргумента не учитываются, то оценки искомых параметров будут смещенными.

Рассмотрим пассивный эксперимент (наблюдаются текущие значения) и определение оценок параметров функции  $\eta = f(\xi, \theta)$ , где  $\theta$  – неизвестный вектор параметров. В процессе измерений значения  $\eta$  и  $\xi$  не наблюдают, а получают набор значений  $x_i$  и  $y_i$ , определяемых как  $y_i = \eta_i + \varepsilon_i$ ,  $x_i = \xi_i + \delta_i$ ,  $i = 1, \dots, n$ , где  $\varepsilon_i$  и  $\delta_i$  – ошибки измеренных значений функции и аргумента.

Предположим, что ошибки измерений  $\varepsilon_i$  и  $\delta_i$  – нормально распределенные случайные величины с нулевыми средними значениями, дисперсиями  $\sigma^2(y_i)$  и  $\sigma^2(x_i)$  соответственно и коэффициентом корреляции  $\rho_i = 0$ .

Начнем с простейшего случая. Рассмотрим две случайные величины  $\xi$  и  $\eta$ , которые имеют совместное распределение непрерывного типа с плотностью вероятности  $f(x, y)$ . В процессе наблюдений случайная величина  $\xi$  приняла значения  $x_i$ , а случайная величина  $\eta$  – значения  $y_i$ ,  $i = 1, 2, \dots, n$ . Будем рассматривать  $\xi$  как независимую переменную. Тогда фиксированному значению  $\xi = x$  соответствует распределение вероятностей зависимой переменной  $\eta$  с плотностью вероятности

$$f(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy},$$

где  $f_1(x)$  – плотность вероятности частного распределения случайной величины  $\xi$ .

Для условного распределения вероятностей переменной  $\eta$  можно найти его числовые характеристики: среднее значение, моду, меди-

Если дисперсия наблюдаемых значений  $\sigma_i^2 = \sigma^2, i = 1, \dots, n$ , и значение  $\sigma^2$  не задано, то оценку параметра  $\sigma$  можно найти по формуле  $\hat{\sigma}^2 = \frac{s}{n-p}$ , где  $s = \sum_{i=1}^n [y_i - f(x_i, \hat{\theta})]^2$ ;  $p$  – число параметров;  $n$  – число точек наблюдений.

Регрессионный анализ предполагает, что переменные  $X$  являются детерминированными – измерены с погрешностью много меньшей, чем погрешность значений функции. На практике это требование очень часто невыполняется, поэтому возникает необходимость учета погрешностей аргумента  $X$ . Если погрешности аргумента не учитываются, то оценки искомых параметров будут смещенными.

Рассмотрим пассивный эксперимент (наблюдаются текущие значения) и определение оценок параметров функции  $\eta = f(\xi, \theta)$ , где  $\theta$  – неизвестный вектор параметров. В процессе измерений значения  $\eta$  и  $\xi$  не наблюдают, а получают набор значений  $x_i$  и  $y_i$ , определяемых как  $y_i = \eta_i + \varepsilon_i, x_i = \xi_i + \delta_i, i = 1, \dots, n$ , где  $\varepsilon_i$  и  $\delta_i$  – ошибки измеренных значений функции и аргумента.

Предположим, что ошибки измерений  $\varepsilon_i$  и  $\delta_i$  – нормально распределенные случайные величины с нулевыми средними значениями, дисперсиями  $\sigma^2(y_i)$  и  $\sigma^2(x_i)$  соответственно и коэффициентом корреляции  $\rho_i = 0$ .

Начнем с простейшего случая. Рассмотрим две случайные величины  $\xi$  и  $\eta$ , которые имеют совместное распределение непрерывного типа с плотностью вероятности  $f(x, y)$ . В процессе наблюдений случайная величина  $\xi$  приняла значения  $x_i$ , а случайная величина  $\eta$  – значения  $y_i, i = 1, 2, \dots, n$ . Будем рассматривать  $\xi$  как независимую переменную. Тогда фиксированному значению  $\xi = x$  соответствует распределение вероятностей зависимой переменной  $\eta$  с плотностью вероятности

$$f(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy},$$

где  $f_1(x)$  – плотность вероятности частного распределения случайной величины  $\xi$ .

Для условного распределения вероятностей переменной  $\eta$  можно найти его числовые характеристики: среднее значение, моду, меди-

Если дисперсия наблюдаемых значений  $\sigma_i^2 = \sigma^2$ ,  $i = 1, \dots, n$ , и значение  $\sigma^2$  не задано, то оценку параметра  $\sigma$  можно найти по формуле

$\hat{\sigma}^2 = \frac{s}{n-p}$ , где  $s = \sum_{i=1}^n [y_i - f(x_i, \hat{\theta})]^2$ ;  $p$  – число параметров;  $n$  – число точек наблюдений.

Регрессионный анализ предполагает, что переменные  $X$  являются детерминированными – измерены с погрешностью много меньшей, чем погрешность значений функции. На практике это требование очень часто невыполняется, поэтому возникает необходимость учета погрешностей аргумента  $X$ . Если погрешности аргумента не учитываются, то оценки искомых параметров будут смещенными.

Рассмотрим пассивный эксперимент (наблюдаются текущие значения) и определение оценок параметров функции  $\eta = f(\xi, \theta)$ , где  $\theta$  – неизвестный вектор параметров. В процессе измерений значения  $\eta$  и  $\xi$  не наблюдают, а получают набор значений  $x_i$  и  $y_i$ , определяемых как  $y_i = \eta_i + \varepsilon_i$ ,  $x_i = \xi_i + \delta_i$ ,  $i = 1, \dots, n$ , где  $\varepsilon_i$  и  $\delta_i$  – ошибки измеренных значений функции и аргумента.

Предположим, что ошибки измерений  $\varepsilon_i$  и  $\delta_i$  – нормально распределенные случайные величины с нулевыми средними значениями, дисперсиями  $\sigma^2(y_i)$  и  $\sigma^2(x_i)$  соответственно и коэффициентом корреляции  $\rho_i = 0$ .

Начнем с простейшего случая. Рассмотрим две случайные величины  $\xi$  и  $\eta$ , которые имеют совместное распределение непрерывного типа с плотностью вероятности  $f(x, y)$ . В процессе наблюдений случайная величина  $\xi$  приняла значения  $x_i$ , а случайная величина  $\eta$  – значения  $y_i$ ,  $i = 1, 2, \dots, n$ . Будем рассматривать  $\xi$  как независимую переменную. Тогда фиксированному значению  $\xi = x$  соответствует распределение вероятностей зависимой переменной  $\eta$  с плотностью вероятности

$$f(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dy},$$

где  $f_1(x)$  – плотность вероятности частного распределения случайной величины  $\xi$ .

Для условного распределения вероятностей переменной  $\eta$  можно найти его числовые характеристики: среднее значение, моду, меди-

Заметим, что использование функции правдоподобия приводит к МНК только в случае нормального закона распределения случайной величины (зависимой переменной). Например, если случайная величина  $\eta$  подчиняется распределению Лапласа

$$f(x) = \frac{1}{2\beta} \exp \left\{ - \frac{\left| y - \sum_i \alpha_i \varphi_i(x) \right|}{\beta} \right\},$$

то функция правдоподобия для независимых наблюдений имеет вид

$$L(y|\alpha) = \frac{1}{2^n \beta^n} \prod_{j=1}^n \exp \left\{ - \frac{\left| y_j - \sum_i \alpha_i \varphi_i(x_j) \right|}{\beta} \right\},$$

а функционал, аналогичный функционалу (2.3),

$$F = \frac{1}{\beta} \sum_{j=1}^n \left| y_j - \sum_i \alpha_i \varphi_i(x_j) \right|. \quad (2.4)$$

Функционалы (2.3) и (2.4) показывают, что то или иное определение «возможно лучшего» зависит от законов распределения: исходных данных.

Здесь мы рассмотрели функционалы для регрессии  $\eta$  на  $\xi$ . Аналогично можно получить выражения для функционалов, определяющих регрессию  $\xi$  на  $\eta$ . В обоих случаях подчеркнем тот факт, что независимые переменные принимают фиксированные (детерминированные, неслучайные) значения. Случайным остается только значение зависимой переменной.

## 2.2. Линейная регрессия $\eta$ на $\xi$ и $\xi$ на $\eta$

Получим уравнения линий регрессии в этом широко распространенном случае. Пусть уравнение регрессии  $\eta$  на  $\xi$  имеет вид

$$\bar{y}_x = \theta_1 + \theta_2 x.$$

Используя определение линейной средней квадратической регрессии, найдем оценки  $\theta_1$  и  $\theta_2$  и выразим уравнение регрессии через данные наблюдений. Для оценки  $\theta_1$  и  $\theta_2$  продифференцируем

$$F = \sum_i (y_i - \theta_1 - \theta_2 x_i)^2 \quad (2.5)$$

по  $\theta_1$  и  $\theta_2$  и приравняем производные к нулю. После преобразований получим оценки

$$\hat{\theta}_2 = \rho \frac{\sigma_y}{\sigma_x}; \quad \hat{\theta}_1 = \bar{y} - \hat{\theta}_2 \bar{x},$$

где  $\hat{\theta}_2$  – коэффициент регрессии  $\eta$  на  $\xi$ ;  $\bar{x}$ ,  $\bar{y}$  – соответственно средние значения  $\{x_i\}$  и  $\{y_i\}$ ;  $\rho$  – коэффициент корреляции случайных величин  $\xi$  и  $\eta$ ;  $\sigma_x$  и  $\sigma_y$  – соответственно средние квадратические отклонения вариационных рядов  $x$  и  $y$  (не путать с величиной в (2.3) – дисперсией нормального закона распределения случайной величины  $\eta$ ). Отсюда прямая средней квадратической регрессии  $\eta$  на  $\xi$  имеет вид

$$\frac{\bar{y}_x - \bar{y}}{\sigma_y} = \rho \frac{x - \bar{x}}{\sigma_x}. \quad (2.6)$$

Подставив найденные оценки  $\hat{\theta}_1$  и  $\hat{\theta}_2$  в (2.5), найдем минимальное значение функционала  $F$ :

$$F_{\min} = \sigma_y^2(1 - \rho^2).$$

Выражение (2.5) определяет сумму квадратов расстояний по вертикали (вдоль оси  $OY$ ) между точками  $(x_i, y_i)$  и прямой  $\bar{y}_x = \theta_1 + \theta_2 x$ .

В случае регрессии  $\xi$  на  $\eta$  надо найти оценки параметров  $\alpha$  и  $\beta$ , при которых

$$M \left[ (\xi - \alpha - \beta\eta)^2 \right] \rightarrow \min_{\alpha, \beta},$$

или

$$\Phi = \sum_i (x_i - \alpha - \beta y_i)^2 \rightarrow \min_{\alpha, \beta},$$

где  $y_i$  фиксированы. Здесь расстояние между точками  $(y_i, z_i)$  и прямой  $\bar{x}_y = \alpha + \beta y$  измеряется по горизонтали.

Уравнение прямой средней квадратической регрессии  $\xi$  на  $\eta$  имеет вид

$$\frac{\bar{x}_y - \bar{x}}{\sigma_x} = \rho \frac{y - \bar{y}}{\sigma_y}. \quad (2.7)$$

Коэффициент регрессии  $\hat{\beta} = \rho \frac{\sigma_x}{\sigma_y}$  и  $\Phi_{\min} = \sigma_x^2(1 - \rho^2)$ .

### 2.3. Регрессионный парадокс

По одним и тем же исходным данным (если не учитывать строго их погрешности) согласно уравнениям (2.6) и (2.7) можно получить различные прямые, а следовательно, и различные выводы. Прямые

регрессии совпадают только при  $\rho \pm 1$ ; при  $\rho = 0$  (когда  $\xi$  и  $\eta$  – независимые случайные величины) имеем соответственно уравнения

$$\bar{y}_x = \bar{y}; \quad \bar{x}_y = \bar{x}$$

– две взаимно перпендикулярные прямые.

Если  $\bar{x} = \bar{y} = 0$ ,  $\sigma_x = \sigma_y = 1$ , то линия регрессии  $\eta$  на  $\xi$  описывается уравнением  $\bar{y}_x = \rho x$ ; а линия регрессии  $\xi$  на  $\eta$  – уравнением  $y = \bar{x}_y / \rho$ .

В качестве примера рассмотрим линии регрессии, характеризующие зависимость роста сыновей от роста отцов. Пусть переменная  $\xi$  характеризует рост отца, а переменная  $\eta$  – рост сына (рис. 2.1).

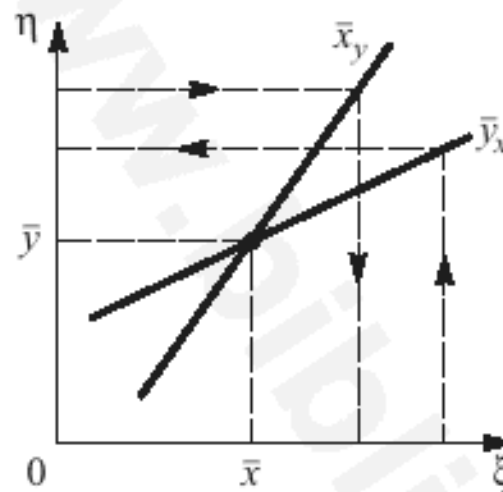


Рис. 2.1. Регрессионный парадокс

Уравнение регрессии  $\eta$  на  $\xi$  имеет вид

$$\frac{\bar{y}_x - \bar{y}}{\sigma_y} = \rho \frac{x - \bar{x}}{\sigma_x}; \quad |\rho| \leq 1; \quad \rho > 0.$$

Положим  $\sigma_y = \sigma_x = 1$  и перенесем начало координат в центр тяжести, в точку  $(\bar{x}, \bar{y})$ , т.е. в уравнениях регрессии надо принять  $\bar{x} = \bar{y} = 0$ . Получим

$$\bar{y}_x = \rho x < x.$$

Это означает, что в среднем сыновья высоких отцов не так высоки, как их отцы. Рост сыновей имеет тенденцию к усреднению.

Согласно регрессии  $\xi$  на  $\eta$

$$\frac{\bar{x}_y - \bar{x}}{\sigma_x} = \rho \frac{y - \bar{y}}{\sigma_y}$$

при сделанных предположениях

$$\bar{x}_y = \rho y < y.$$

Это означает, что в среднем отцы высоких детей не так высоки, как их дети. Первое поколение имеет тенденцию к усреднению, при-

чем оба этих утверждения имеют место одновременно – налицо регрессионный парадокс. Чтобы устранить этот парадокс, необходимо при определении уравнения регрессии учесть случайный характер как переменной  $\xi$ , так и переменной  $\eta$ .

#### 2.4. Ортогональная регрессия

Учесть одновременно случайный характер переменных  $\xi$  и  $\eta$  позволяет ортогональная регрессия. Пусть нам известны набор случайных точек  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ , и закон распределения каждого наблюдения. Неопределенность каждой точки теперь, в отличие от регрессий  $\eta$  на  $\xi$  и  $\xi$  на  $\eta$ , имеет вид некоторой области, форма которой зависит от закона распределения результатов наблюдений. И линия регрессии должна возможно лучшим образом пройти по областям неопределенности (рис. 2.2).

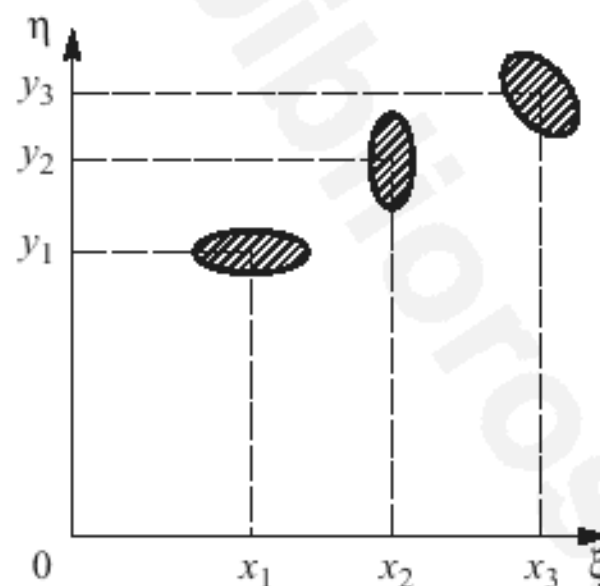


Рис. 2.2. Области неопределенности

Пусть результаты наблюдений  $x_i$  и  $y_i$  подчиняются нормальным законам распределения соответственно с математическими ожиданиями  $\xi_i$  и  $\eta_i$ , дисперсиями  $\sigma^2(x_i)$  и  $\sigma^2(y_i)$  и коэффициентами корреляции  $\rho_i$ . Тогда функция правдоподобия имеет вид

$$L(x, y | \theta) = \prod_{i=1}^n \frac{1}{2\pi\sigma(x_i)\sigma(y_i)\sqrt{1-\rho_i^2}} \times \exp \left\{ -\frac{1}{2(1-\rho_i^2)} \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} - 2\rho_i \frac{(x_i - \xi_i)(y_i - \eta_i)}{\sigma(x_i)\sigma(y_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right] \right\}, \quad (2.8)$$



а соответствующий минимизируемый функционал

$$F = \frac{1}{2} \sum_{i=1}^n \frac{1}{1 - \rho_i^2} \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} - 2\rho_i \frac{(x_i - \xi_i)(y_i - \eta_i)}{\sigma(x_i)\sigma(y_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right]. \quad (2.9)$$

При  $\rho = 0$

$$F = \frac{1}{2} \sum_{i=1}^n \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right] \rightarrow \min. \quad (2.10)$$

Геометрически (при  $\sigma_x = \sigma_y = 1$ ) это означает, что минимизируется сумма квадратов расстояний между точками  $(x_i, y_i)$  и соответствующими точками линии регрессии  $(\xi_i, \eta_i)$ .

Однако из функционала  $F$  нельзя найти оценки параметров  $\hat{\theta}_j$ , так как решить систему уравнений

$$\frac{\partial F}{\partial \theta_j} = \sum_{i=1}^n \frac{y_i - \eta_i}{\sigma^2(y_i)} \frac{\partial \eta_i}{\partial \theta_j} \Big|_{\theta_j = \hat{\theta}_j} = 0, \quad j = 1, 2, \dots, m, \quad (2.11)$$

нельзя, поскольку неизвестны истинные значения  $\xi_i$ , входящие в выражения для  $\eta_i$ .

Доопределим условие задачи. В качестве оценок  $\hat{\xi}_i$  возьмем те значения  $\xi_i$ , которые обратят в нуль частные производные  $\frac{\partial F}{\partial \xi_i}$ ,  $i = 1, 2, \dots, n$ :

$$\frac{\partial F}{\partial \xi_i} = \frac{x_i - \xi_i}{\sigma^2(x_i)} + \frac{y_i - \eta_i}{\sigma^2(y_i)} \frac{\partial \eta_i}{\partial \xi_i} = 0. \quad (2.12)$$

Условие (2.12) при  $\sigma(x_i) = \sigma(y_i) = 1$  есть условие ортогональности векторов  $\{(x_i - \hat{\xi}_i); (y_i - \hat{\eta}_i)\}$  и  $\left\{1; \frac{\partial \eta_i}{\partial \xi_i}\right\}$ . Причем последний вектор направлен по касательной к кривой  $\eta = f(\xi, \theta)$  или вдоль прямой при линейной регрессии (см. рис. 1.4). Условия (2.11), (2.12) определяют третью линию регрессии – ортогональную (см. рис. 1.3), не совпадающую ни с  $\eta$  на  $\xi$ , ни с  $\xi$  на  $\eta$ , но одновременно учитывающую случайный характер и  $\xi$ , и  $\eta$ .

Приведенные результаты справедливы для пассивного эксперимента, когда просто наблюдаются текущие значения  $x_i$  и соответствующие им значения  $y_i$ . К результатам активного эксперимента, когда заранее устанавливаются значения  $x_i$ , в общем случае выводы пассивного эксперимента неприменимы.

### 2.5. Метод наименьших квадратов. Оценка свободных параметров функций, линейных по параметрам

Примером функций, линейных по параметрам  $\theta$ , могут быть функции  $\eta = \sum_{i=0}^k \theta_i x^i$ ,  $\eta = \sum_{i=0}^k \theta_i f_i(x)$  и т.д. Они часто используются для описания разделяющих функций, приближенного описания законов распределения и т.п.

Здесь независимая переменная  $x$  (или в общем случае  $\vec{x}$ ) измеряется без ошибок (детерминирована), т.е. рассматривается регрессия  $Y$  на  $X$ . Результаты наблюдений за случайной величиной  $\eta$  дают значения  $y$ , соответствующие значениям  $x$ . Каждому значению  $x_i$  может отвечать либо единственное значение  $y_i$ , либо множество значений  $y_{i1}, y_{i2}, \dots, y_{im}$ .

Будем считать, что каждое наблюдение  $y_i$  содержит ошибку  $\epsilon_i$ .

Пусть  $n$ -мерный вектор наблюдений (вектор откликов)  $\vec{y} = \{y_1, y_2, \dots, y_n\}^T$  порождается моделью  $\vec{y} = \vec{x}\vec{\theta} + \vec{\epsilon}$ , где  $\vec{x}$  – известная матрица измерений (матрица плана), размером  $n \times p$ , построенная либо по значениям вектора  $\vec{x}$ , либо по значениям  $f_i(x)$ ,  $i = 1, 2, \dots, k$ ;  $\vec{\theta} = \{\theta_1, \theta_2, \dots, \theta_p\}$  – неизвестный  $p$ -мерный вектор параметров  $\theta$ ;  $\vec{\epsilon}$  –  $n$ -мерный случайный вектор ошибок, удовлетворяющий условию  $M(\vec{\epsilon}) = 0$ ,  $D(\vec{\epsilon}) = M(\vec{\epsilon}^T \vec{\epsilon}) = \sigma^2 I$ . Здесь  $\sigma^2$  – неизвестный скалярный параметр;  $I$  – единичная матрица. Линейные модели наблюдений могут быть полного и неполного ранга. Если ранг матрицы измерений  $\text{rank } \vec{x}$  равен  $p$ , то модель наблюдений называется моделью полного ранга, если  $\text{rank } \vec{x} < p$ , – неполного ранга. Другими словами, если число точек наблюдений меньше числа искомых параметров  $\vec{\theta}$ , то имеет место модель наблюдений неполного ранга. В этом случае часть параметров  $\vec{\theta}$  являются свободными. Модель неполного ранга здесь не рассматривается.

А. Простейший случай:  $\sigma^2 = 1$ .

Согласно МНК будем искать точку минимума квадратичной формы

$$Q(\vec{\theta}) = \vec{\epsilon}^T \vec{\epsilon} = (\vec{y} - \vec{x}\vec{\theta})^T (\vec{y} - \vec{x}\vec{\theta}) = \vec{y}^T \vec{y} - 2\vec{\theta}^T \vec{x}^T \vec{y} + \vec{\theta}^T \vec{x}^T \vec{x} \vec{\theta}.$$

Дифференцируем  $Q(\vec{\theta})$  по  $\vec{\theta}$ :

$$-2\vec{x}^T \vec{y} + 2\vec{x}^T \vec{x} \vec{\theta} = 0$$

– система нормальных уравнений.

Отсюда

$$\vec{\theta} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}.$$

Минимальное значение  $S$ , которое имеет  $Q(\vec{\theta})$  при  $\vec{\theta} = \hat{\vec{\theta}}$ ,

$$\begin{aligned} S &= \vec{y}^T \vec{y} - 2 \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{y} + \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{x} \times \left[ \vec{x}^T \vec{x} \right]^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - 2 \vec{y}^T \vec{x} \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right] + \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}; \\ S &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} \hat{\vec{\theta}}. \end{aligned}$$

Здесь использованы соотношения  $(AB)^T = B^T A^T$ ;  $(AB)^{-1} = B^{-1} A^{-1}$ .

Итак, метод оценивания параметров (МНК) не требует никаких предположений о нормальном законе распределения исходных данных при получении точечных оценок, но они становятся необходимыми при построении доверительных интервалов и для проверки гипотез о параметрах  $\vec{\theta}$ .

Рассмотрим свойства оценок МНК.

**Теорема.** Оценка  $\hat{\vec{\theta}}$  является несмещенной оценкой параметра  $\vec{\theta}$ .

Доказательство: Найдем математическое ожидание оценок

$$M[\hat{\vec{\theta}}] = M \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} \right] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T M[\vec{y}].$$

Поскольку переменные  $x$  – детерминированные величины,

$$M[\vec{y}] = \vec{x} \vec{\theta}; \quad M[\hat{\vec{\theta}}] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} \vec{\theta} = \vec{\theta}$$

– несмещенная оценка. Но если какая-то переменная  $x$  – случайная величина, оценки  $\hat{\vec{\theta}}$  будут смещенными.

**Теорема (Гаусса – Маркова).** Среди класса оценок  $\theta^*$  величины  $\theta$ , которые: а) являются несмещенными оценками и б) представляют собой линейные комбинации исходных данных  $\vec{y}$ , с помощью метода наименьших квадратов можно найти такую оценку  $\hat{\vec{\theta}}$ , что  $D(\hat{\vec{\theta}}) \leq D(\theta^*)$ , т.е.  $\hat{\vec{\theta}}$  – наиболее точная оценка  $\theta$  из всех возможных, принадлежащих данному классу (эффективная оценка). Без доказательства.

Получим формулу для оценки  $\vec{D}(\hat{\vec{\theta}})$ . Пусть дисперсия  $D(\varepsilon_j) = \sigma^2$ ,  $j = 1, \dots, n$ , т.е. дисперсия  $D(\vec{y}) = \sigma^2 I$ . В общем случае  $D(\vec{W} \vec{y}) = \vec{W} D(\vec{y}) \vec{W}^T$ . При  $\hat{\vec{\theta}} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}$

$$\begin{aligned} D[\hat{\vec{\theta}}] &= (\vec{x}^T \vec{x})^{-1} \vec{x}^T D(\vec{y}) \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \right]^T = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \sigma^2 I \vec{x} (\vec{x}^T \vec{x})^{-1} = \\ &= \sigma^2 (\vec{x}^T \vec{x})^{-1} \vec{x}^T I \vec{x} (\vec{x}^T \vec{x})^{-1} = \sigma^2 (\vec{x}^T \vec{x})^{-1}. \end{aligned}$$

Отсюда

$$\vec{\theta} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}.$$

Минимальное значение  $S$ , которое имеет  $Q(\vec{\theta})$  при  $\vec{\theta} = \hat{\vec{\theta}}$ ,

$$\begin{aligned} S &= \vec{y}^T \vec{y} - 2 \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{y} + \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{x} \times \left[ \vec{x}^T \vec{x} \right]^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - 2 \vec{y}^T \vec{x} \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right] + \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}; \\ S &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} \hat{\vec{\theta}}. \end{aligned}$$

Здесь использованы соотношения  $(AB)^T = B^T A^T$ ;  $(AB)^{-1} = B^{-1} A^{-1}$ .

Итак, метод оценивания параметров (МНК) не требует никаких предположений о нормальном законе распределения исходных данных при получении точечных оценок, но они становятся необходимыми при построении доверительных интервалов и для проверки гипотез о параметрах  $\vec{\theta}$ .

Рассмотрим свойства оценок МНК.

**Теорема.** Оценка  $\hat{\vec{\theta}}$  является несмещенной оценкой параметра  $\vec{\theta}$ .

Доказательство: Найдем математическое ожидание оценок

$$M[\hat{\vec{\theta}}] = M \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} \right] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T M[\vec{y}].$$

Поскольку переменные  $x$  – детерминированные величины,

$$M[\vec{y}] = \vec{x} \vec{\theta}; \quad M[\hat{\vec{\theta}}] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} \vec{\theta} = \vec{\theta}$$

– несмещенная оценка. Но если какая-то переменная  $x$  – случайная величина, оценки  $\hat{\vec{\theta}}$  будут смещенными.

**Теорема (Гаусса – Маркова).** Среди класса оценок  $\theta^*$  величины  $\theta$ , которые: а) являются несмещенными оценками и б) представляют собой линейные комбинации исходных данных  $\vec{y}$ , с помощью метода наименьших квадратов можно найти такую оценку  $\hat{\vec{\theta}}$ , что  $D(\hat{\vec{\theta}}) \leq D(\theta^*)$ , т.е.  $\hat{\vec{\theta}}$  – наиболее точная оценка  $\theta$  из всех возможных, принадлежащих данному классу (эффективная оценка). Без доказательства.

Получим формулу для оценки  $\vec{D}(\hat{\vec{\theta}})$ . Пусть дисперсия  $D(\varepsilon_j) = \sigma^2$ ,  $j = 1, \dots, n$ , т.е. дисперсия  $D(\vec{y}) = \sigma^2 I$ . В общем случае  $D(\vec{W} \vec{y}) = \vec{W} D(\vec{y}) \vec{W}^T$ . При  $\hat{\vec{\theta}} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}$

$$\begin{aligned} D[\hat{\vec{\theta}}] &= (\vec{x}^T \vec{x})^{-1} \vec{x}^T D(\vec{y}) \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \right]^T = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \sigma^2 I \vec{x} (\vec{x}^T \vec{x})^{-1} = \\ &= \sigma^2 (\vec{x}^T \vec{x})^{-1} \vec{x}^T I \vec{x} (\vec{x}^T \vec{x})^{-1} = \sigma^2 (\vec{x}^T \vec{x})^{-1}. \end{aligned}$$

Отсюда

$$\vec{\theta} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}.$$

Минимальное значение  $S$ , которое имеет  $Q(\vec{\theta})$  при  $\vec{\theta} = \hat{\vec{\theta}}$ ,

$$\begin{aligned} S &= \vec{y}^T \vec{y} - 2 \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{y} + \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{x} \times \left[ \vec{x}^T \vec{x} \right]^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - 2 \vec{y}^T \vec{x} \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right] + \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}; \\ S &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} \hat{\vec{\theta}}. \end{aligned}$$

Здесь использованы соотношения  $(AB)^T = B^T A^T$ ;  $(AB)^{-1} = B^{-1} A^{-1}$ .

Итак, метод оценивания параметров (МНК) не требует никаких предположений о нормальном законе распределения исходных данных при получении точечных оценок, но они становятся необходимыми при построении доверительных интервалов и для проверки гипотез о параметрах  $\vec{\theta}$ .

Рассмотрим свойства оценок МНК.

**Теорема.** Оценка  $\hat{\vec{\theta}}$  является несмещенной оценкой параметра  $\vec{\theta}$ .

Доказательство: Найдем математическое ожидание оценок

$$M[\hat{\vec{\theta}}] = M \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} \right] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T M[\vec{y}].$$

Поскольку переменные  $x$  – детерминированные величины,

$$M[\vec{y}] = \vec{x} \vec{\theta}; \quad M[\hat{\vec{\theta}}] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} \vec{\theta} = \vec{\theta}$$

– несмещенная оценка. Но если какая-то переменная  $x$  – случайная величина, оценки  $\hat{\vec{\theta}}$  будут смещенными.

**Теорема (Гаусса – Маркова).** Среди класса оценок  $\theta^*$  величины  $\theta$ , которые: а) являются несмещенными оценками и б) представляют собой линейные комбинации исходных данных  $\vec{y}$ , с помощью метода наименьших квадратов можно найти такую оценку  $\hat{\vec{\theta}}$ , что  $D(\hat{\vec{\theta}}) \leq D(\theta^*)$ , т.е.  $\hat{\vec{\theta}}$  – наиболее точная оценка  $\theta$  из всех возможных, принадлежащих данному классу (эффективная оценка). Без доказательства.

Получим формулу для оценки  $\vec{D}(\hat{\vec{\theta}})$ . Пусть дисперсия  $D(\varepsilon_j) = \sigma^2$ ,  $j = 1, \dots, n$ , т.е. дисперсия  $D(\vec{y}) = \sigma^2 I$ . В общем случае  $D(\vec{W} \vec{y}) = \vec{W} D(\vec{y}) \vec{W}^T$ . При  $\hat{\vec{\theta}} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}$

$$\begin{aligned} D[\hat{\vec{\theta}}] &= (\vec{x}^T \vec{x})^{-1} \vec{x}^T D(\vec{y}) \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \right]^T = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \sigma^2 I \vec{x} (\vec{x}^T \vec{x})^{-1} = \\ &= \sigma^2 (\vec{x}^T \vec{x})^{-1} \vec{x}^T I \vec{x} (\vec{x}^T \vec{x})^{-1} = \sigma^2 (\vec{x}^T \vec{x})^{-1}. \end{aligned}$$

Отсюда

$$\vec{\theta} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}.$$

Минимальное значение  $S$ , которое имеет  $Q(\vec{\theta})$  при  $\vec{\theta} = \hat{\vec{\theta}}$ ,

$$\begin{aligned} S &= \vec{y}^T \vec{y} - 2 \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{y} + \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{x} \times \left[ \vec{x}^T \vec{x} \right]^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - 2 \vec{y}^T \vec{x} \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right] + \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}; \\ S &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} \hat{\vec{\theta}}. \end{aligned}$$

Здесь использованы соотношения  $(AB)^T = B^T A^T$ ;  $(AB)^{-1} = B^{-1} A^{-1}$ .

Итак, метод оценивания параметров (МНК) не требует никаких предположений о нормальном законе распределения исходных данных при получении точечных оценок, но они становятся необходимыми при построении доверительных интервалов и для проверки гипотез о параметрах  $\vec{\theta}$ .

Рассмотрим свойства оценок МНК.

**Теорема.** Оценка  $\hat{\vec{\theta}}$  является несмещенной оценкой параметра  $\vec{\theta}$ .

Доказательство: Найдем математическое ожидание оценок

$$M[\hat{\vec{\theta}}] = M \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} \right] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T M[\vec{y}].$$

Поскольку переменные  $x$  – детерминированные величины,

$$M[\vec{y}] = \vec{x} \vec{\theta}; \quad M[\hat{\vec{\theta}}] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} \vec{\theta} = \vec{\theta}$$

– несмещенная оценка. Но если какая-то переменная  $x$  – случайная величина, оценки  $\hat{\vec{\theta}}$  будут смещенными.

**Теорема (Гаусса – Маркова).** Среди класса оценок  $\theta^*$  величины  $\theta$ , которые: а) являются несмещенными оценками и б) представляют собой линейные комбинации исходных данных  $\vec{y}$ , с помощью метода наименьших квадратов можно найти такую оценку  $\hat{\vec{\theta}}$ , что  $D(\hat{\vec{\theta}}) \leq D(\theta^*)$ , т.е.  $\hat{\vec{\theta}}$  – наиболее точная оценка  $\theta$  из всех возможных, принадлежащих данному классу (эффективная оценка). Без доказательства.

Получим формулу для оценки  $\vec{D}(\hat{\vec{\theta}})$ . Пусть дисперсия  $D(\varepsilon_j) = \sigma^2$ ,  $j = 1, \dots, n$ , т.е. дисперсия  $D(\vec{y}) = \sigma^2 I$ . В общем случае  $D(\vec{W} \vec{y}) = \vec{W} D(\vec{y}) \vec{W}^T$ . При  $\hat{\vec{\theta}} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}$

$$\begin{aligned} D[\hat{\vec{\theta}}] &= (\vec{x}^T \vec{x})^{-1} \vec{x}^T D(\vec{y}) \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \right]^T = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \sigma^2 I \vec{x} (\vec{x}^T \vec{x})^{-1} = \\ &= \sigma^2 (\vec{x}^T \vec{x})^{-1} \vec{x}^T I \vec{x} (\vec{x}^T \vec{x})^{-1} = \sigma^2 (\vec{x}^T \vec{x})^{-1}. \end{aligned}$$

Отсюда

$$\vec{\theta} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}.$$

Минимальное значение  $S$ , которое имеет  $Q(\vec{\theta})$  при  $\vec{\theta} = \hat{\vec{\theta}}$ ,

$$\begin{aligned} S &= \vec{y}^T \vec{y} - 2 \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{y} + \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right]^T \vec{x}^T \vec{x} \times \left[ \vec{x}^T \vec{x} \right]^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - 2 \vec{y}^T \vec{x} \left[ (\vec{x}^T \vec{x})^{-1} (\vec{x}^T \vec{y}) \right] + \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} = \\ &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}; \\ S &= \vec{y}^T \vec{y} - \vec{y}^T \vec{x} \hat{\vec{\theta}}. \end{aligned}$$

Здесь использованы соотношения  $(AB)^T = B^T A^T$ ;  $(AB)^{-1} = B^{-1} A^{-1}$ .

Итак, метод оценивания параметров (МНК) не требует никаких предположений о нормальном законе распределения исходных данных при получении точечных оценок, но они становятся необходимыми при построении доверительных интервалов и для проверки гипотез о параметрах  $\vec{\theta}$ .

Рассмотрим свойства оценок МНК.

**Теорема.** Оценка  $\hat{\vec{\theta}}$  является несмещенной оценкой параметра  $\vec{\theta}$ .

Доказательство: Найдем математическое ожидание оценок

$$M[\hat{\vec{\theta}}] = M \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y} \right] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T M[\vec{y}].$$

Поскольку переменные  $x$  – детерминированные величины,

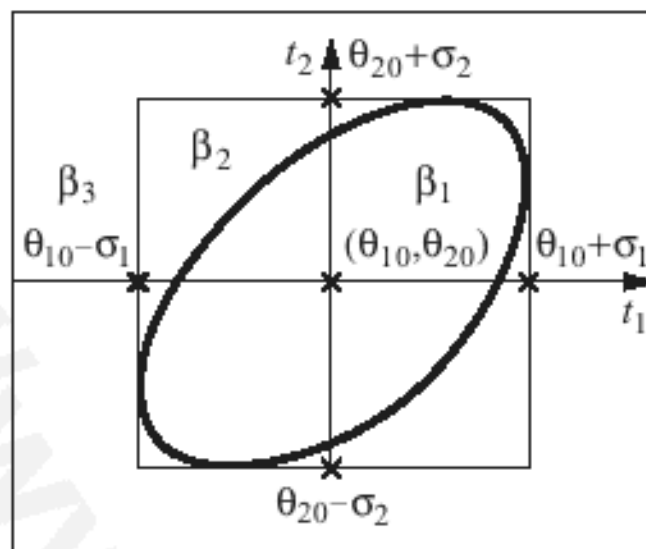
$$M[\vec{y}] = \vec{x} \vec{\theta}; \quad M[\hat{\vec{\theta}}] = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{x} \vec{\theta} = \vec{\theta}$$

– несмещенная оценка. Но если какая-то переменная  $x$  – случайная величина, оценки  $\hat{\vec{\theta}}$  будут смещенными.

**Теорема (Гаусса – Маркова).** Среди класса оценок  $\theta^*$  величины  $\theta$ , которые: а) являются несмещенными оценками и б) представляют собой линейные комбинации исходных данных  $\vec{y}$ , с помощью метода наименьших квадратов можно найти такую оценку  $\hat{\vec{\theta}}$ , что  $D(\hat{\vec{\theta}}) \leq D(\theta^*)$ , т.е.  $\hat{\vec{\theta}}$  – наиболее точная оценка  $\theta$  из всех возможных, принадлежащих данному классу (эффективная оценка). Без доказательства.

Получим формулу для оценки  $\vec{D}(\hat{\vec{\theta}})$ . Пусть дисперсия  $D(\varepsilon_j) = \sigma^2$ ,  $j = 1, \dots, n$ , т.е. дисперсия  $D(\vec{y}) = \sigma^2 I$ . В общем случае  $D(\vec{W} \vec{y}) = \vec{W} D(\vec{y}) \vec{W}^T$ . При  $\hat{\vec{\theta}} = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \vec{y}$

$$\begin{aligned} D[\hat{\vec{\theta}}] &= (\vec{x}^T \vec{x})^{-1} \vec{x}^T D(\vec{y}) \left[ (\vec{x}^T \vec{x})^{-1} \vec{x}^T \right]^T = (\vec{x}^T \vec{x})^{-1} \vec{x}^T \sigma^2 I \vec{x} (\vec{x}^T \vec{x})^{-1} = \\ &= \sigma^2 (\vec{x}^T \vec{x})^{-1} \vec{x}^T I \vec{x} (\vec{x}^T \vec{x})^{-1} = \sigma^2 (\vec{x}^T \vec{x})^{-1}. \end{aligned}$$



**Рис. 2.6.** Доверительные области для нормальных оценок  $\theta$  параметра  $t$  при  $k_{\beta} = 1$  и  $\rho = 0,5$ :  
 $\beta_1, \beta_2, \beta_3$  – соответственно вероятностное содержание для эллиптической области, для описанного прямоугольника, для горизонтальной полосы соответственно

Приведенные результаты и рисунки позволяют пользователю по своему усмотрению выбирать области рассеяния реализуемой величины как в виде эллипсоидов, так и в виде прямоугольников.

## 2.6. Оценка параметров моделей с помощью функции правдоподобия

Одним из наиболее мощных методов оценки свободных параметров функций известного вида (в данном случае функций распределения и плотностей вероятности) является метод максимума правдоподобия (ММП).

Пусть случайная величина  $X$  имеет плотность вероятности  $f(X, \vec{\theta})$ , где  $\vec{\theta}$  – неизвестный вектор параметров этого распределения. Проведем независимые выборки и получим реализации  $x_1, x_2, \dots, x_n$ . Дифференциал вероятности совместной плотности распределения результатов наблюдений имеет вид

$$\begin{aligned} L(x_1, x_2, \dots, x_n; \vec{\theta}) dx_1, dx_2, \dots, dx_n &= \\ &= f(x_1, \vec{\theta}) f(x_2, \vec{\theta}) \dots f(x_n, \vec{\theta}) dx_1, dx_2, \dots, dx_n. \end{aligned}$$

Совместная плотность вероятности  $L(X, \vec{\theta}) = \prod_{i=1}^n f(x_i, \vec{\theta})$  называется *функцией правдоподобия* для выборки  $x_1, x_2, \dots, x_n$ .



Аналогично под  $f(x_i, \vec{\theta})$  можно понимать вероятность того, что дискретная случайная величина  $X$  принимает значение  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ , т.е.

$$f(x_i, \vec{\theta}) = P(X = x_i).$$

Можно найти такое число (оценку)  $\hat{\theta}$ , которое обращает функцию правдоподобия в максимум;  $\hat{\theta}$  и будет точечной оценкой параметра  $\vec{\theta}$ , т.е. для вычисления оценки  $\hat{\theta}$  надо решить уравнение  $\frac{\partial L(X, \theta)}{\partial \theta} = 0$ .

Будем считать, что оценки  $\hat{\theta}$  удовлетворяют следующим условиям (условиям регулярности):

1) пусть  $\hat{\theta} = \theta(x_1, x_2, \dots, x_n)$  – однозначная функция от  $x_1, x_2, \dots, x_n$  непрерывная и имеющая непрерывные частные производные  $\frac{\partial \theta}{\partial x_i}$  во всех точках  $X = (x_1, x_2, \dots, x_n) \in R_n$  и операторы  $\frac{\partial}{\partial \theta}$  и  $\int$  коммутируют;

2) область значений  $X$  зависит от параметра  $\theta$ .

Предположим, что математическое ожидание оценки  $M[\hat{\theta}] = \theta + b(\theta)$ , где  $b(\theta)$  – смещение оценки  $\hat{\theta}$ .

**Теорема.** В каждом случае регулярной оценки непрерывного типа квадрат среднего квадратического отклонения оценки  $\hat{\theta}$  от истинного значения  $\theta$  удовлетворяет неравенству

$$D(\hat{\theta}) = M[(\hat{\theta} - \theta)^2] \geq \frac{\left(1 + \frac{db}{d\theta}\right)^2}{\int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial \theta}\right)^2 L(x, \theta) dx}.$$

**Доказательство.** Имеем

$$\int L(\hat{\theta}|\theta) d\hat{\theta} = 1.$$

Дифференцируем это выражение по  $\theta$ :

$$\int \frac{\partial L}{\partial \theta} d\hat{\theta} = 0 = \int \frac{\partial \ln L}{\partial \theta} L(\hat{\theta}|\theta) d\hat{\theta}, \text{ так как } \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}.$$

Имеем

$$M(\hat{\theta}) = \int \hat{\theta} \cdot L(\hat{\theta}|\theta) d\hat{\theta} = \theta + b(\theta);$$

Аналогично под  $f(x_i, \vec{\theta})$  можно понимать вероятность того, что дискретная случайная величина  $X$  принимает значение  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ , т.е.

$$f(x_i, \vec{\theta}) = P(X = x_i).$$

Можно найти такое число (оценку)  $\hat{\theta}$ , которое обращает функцию правдоподобия в максимум;  $\hat{\theta}$  и будет точечной оценкой параметра  $\vec{\theta}$ , т.е. для вычисления оценки  $\hat{\theta}$  надо решить уравнение  $\frac{\partial L(X, \theta)}{\partial \theta} = 0$ .

Будем считать, что оценки  $\hat{\theta}$  удовлетворяют следующим условиям (условиям регулярности):

1) пусть  $\hat{\theta} = \theta(x_1, x_2, \dots, x_n)$  – однозначная функция от  $x_1, x_2, \dots, x_n$  непрерывная и имеющая непрерывные частные производные  $\frac{\partial \theta}{\partial x_i}$  во всех точках  $X = (x_1, x_2, \dots, x_n) \in R_n$  и операторы  $\frac{\partial}{\partial \theta}$  и  $\int$  коммутируют;

2) область значений  $X$  зависит от параметра  $\theta$ .

Предположим, что математическое ожидание оценки  $M[\hat{\theta}] = \theta + b(\theta)$ , где  $b(\theta)$  – смещение оценки  $\hat{\theta}$ .

**Теорема.** В каждом случае регулярной оценки непрерывного типа квадрат среднего квадратического отклонения оценки  $\hat{\theta}$  от истинного значения  $\theta$  удовлетворяет неравенству

$$D(\hat{\theta}) = M[(\hat{\theta} - \theta)^2] \geq \frac{\left(1 + \frac{db}{d\theta}\right)^2}{\int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial \theta}\right)^2 L(x, \theta) dx}.$$

**Доказательство.** Имеем

$$\int L(\hat{\theta}|\theta) d\hat{\theta} = 1.$$

Дифференцируем это выражение по  $\theta$ :

$$\int \frac{\partial L}{\partial \theta} d\hat{\theta} = 0 = \int \frac{\partial \ln L}{\partial \theta} L(\hat{\theta}|\theta) d\hat{\theta}, \text{ так как } \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}.$$

Имеем

$$M(\hat{\theta}) = \int \hat{\theta} \cdot L(\hat{\theta}|\theta) d\hat{\theta} = \theta + b(\theta);$$

Аналогично под  $f(x_i, \vec{\theta})$  можно понимать вероятность того, что дискретная случайная величина  $X$  принимает значение  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ , т.е.

$$f(x_i, \vec{\theta}) = P(X = x_i).$$

Можно найти такое число (оценку)  $\hat{\theta}$ , которое обращает функцию правдоподобия в максимум;  $\hat{\theta}$  и будет точечной оценкой параметра  $\vec{\theta}$ , т.е. для вычисления оценки  $\hat{\theta}$  надо решить уравнение  $\frac{\partial L(X, \theta)}{\partial \theta} = 0$ .

Будем считать, что оценки  $\hat{\theta}$  удовлетворяют следующим условиям (условиям регулярности):

1) пусть  $\hat{\theta} = \theta(x_1, x_2, \dots, x_n)$  – однозначная функция от  $x_1, x_2, \dots, x_n$  непрерывная и имеющая непрерывные частные производные  $\frac{\partial \theta}{\partial x_i}$  во всех точках  $X = (x_1, x_2, \dots, x_n) \in R_n$  и операторы  $\frac{\partial}{\partial \theta}$  и  $\int$  коммутируют;

2) область значений  $X$  зависит от параметра  $\theta$ .

Предположим, что математическое ожидание оценки  $M[\hat{\theta}] = \theta + b(\theta)$ , где  $b(\theta)$  – смещение оценки  $\hat{\theta}$ .

**Теорема.** В каждом случае регулярной оценки непрерывного типа квадрат среднего квадратического отклонения оценки  $\hat{\theta}$  от истинного значения  $\theta$  удовлетворяет неравенству

$$D(\hat{\theta}) = M[(\hat{\theta} - \theta)^2] \geq \frac{\left(1 + \frac{db}{d\theta}\right)^2}{\int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial \theta}\right)^2 L(x, \theta) dx}.$$

**Доказательство.** Имеем

$$\int L(\hat{\theta}|\theta) d\hat{\theta} = 1.$$

Дифференцируем это выражение по  $\theta$ :

$$\int \frac{\partial L}{\partial \theta} d\hat{\theta} = 0 = \int \frac{\partial \ln L}{\partial \theta} L(\hat{\theta}|\theta) d\hat{\theta}, \text{ так как } \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}.$$

Имеем

$$M(\hat{\theta}) = \int \hat{\theta} \cdot L(\hat{\theta}|\theta) d\hat{\theta} = \theta + b(\theta);$$

Аналогично под  $f(x_i, \vec{\theta})$  можно понимать вероятность того, что дискретная случайная величина  $X$  принимает значение  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ , т.е.

$$f(x_i, \vec{\theta}) = P(X = x_i).$$

Можно найти такое число (оценку)  $\hat{\theta}$ , которое обращает функцию правдоподобия в максимум;  $\hat{\theta}$  и будет точечной оценкой параметра  $\vec{\theta}$ , т.е. для вычисления оценки  $\hat{\theta}$  надо решить уравнение  $\frac{\partial L(X, \theta)}{\partial \theta} = 0$ .

Будем считать, что оценки  $\hat{\theta}$  удовлетворяют следующим условиям (условиям регулярности):

1) пусть  $\hat{\theta} = \theta(x_1, x_2, \dots, x_n)$  – однозначная функция от  $x_1, x_2, \dots, x_n$  непрерывная и имеющая непрерывные частные производные  $\frac{\partial \theta}{\partial x_i}$  во всех точках  $X = (x_1, x_2, \dots, x_n) \in R_n$  и операторы  $\frac{\partial}{\partial \theta}$  и  $\int$  коммутируют;

2) область значений  $X$  зависит от параметра  $\theta$ .

Предположим, что математическое ожидание оценки  $M[\hat{\theta}] = \theta + b(\theta)$ , где  $b(\theta)$  – смещение оценки  $\hat{\theta}$ .

**Теорема.** В каждом случае регулярной оценки непрерывного типа квадрат среднего квадратического отклонения оценки  $\hat{\theta}$  от истинного значения  $\theta$  удовлетворяет неравенству

$$D(\hat{\theta}) = M[(\hat{\theta} - \theta)^2] \geq \frac{\left(1 + \frac{db}{d\theta}\right)^2}{\int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial \theta}\right)^2 L(x, \theta) dx}.$$

**Доказательство.** Имеем

$$\int L(\hat{\theta}|\theta) d\hat{\theta} = 1.$$

Дифференцируем это выражение по  $\theta$ :

$$\int \frac{\partial L}{\partial \theta} d\hat{\theta} = 0 = \int \frac{\partial \ln L}{\partial \theta} L(\hat{\theta}|\theta) d\hat{\theta}, \text{ так как } \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}.$$

Имеем

$$M(\hat{\theta}) = \int \hat{\theta} \cdot L(\hat{\theta}|\theta) d\hat{\theta} = \theta + b(\theta);$$

Аналогично под  $f(x_i, \vec{\theta})$  можно понимать вероятность того, что дискретная случайная величина  $X$  принимает значение  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ , т.е.

$$f(x_i, \vec{\theta}) = P(X = x_i).$$

Можно найти такое число (оценку)  $\hat{\theta}$ , которое обращает функцию правдоподобия в максимум;  $\hat{\theta}$  и будет точечной оценкой параметра  $\vec{\theta}$ , т.е. для вычисления оценки  $\hat{\theta}$  надо решить уравнение  $\frac{\partial L(X, \theta)}{\partial \theta} = 0$ .

Будем считать, что оценки  $\hat{\theta}$  удовлетворяют следующим условиям (условиям регулярности):

1) пусть  $\hat{\theta} = \theta(x_1, x_2, \dots, x_n)$  – однозначная функция от  $x_1, x_2, \dots, x_n$  непрерывная и имеющая непрерывные частные производные  $\frac{\partial \theta}{\partial x_i}$  во всех точках  $X = (x_1, x_2, \dots, x_n) \in R_n$  и операторы  $\frac{\partial}{\partial \theta}$  и  $\int$  коммутируют;

2) область значений  $X$  зависит от параметра  $\theta$ .

Предположим, что математическое ожидание оценки  $M[\hat{\theta}] = \theta + b(\theta)$ , где  $b(\theta)$  – смещение оценки  $\hat{\theta}$ .

**Теорема.** В каждом случае регулярной оценки непрерывного типа квадрат среднего квадратического отклонения оценки  $\hat{\theta}$  от истинного значения  $\theta$  удовлетворяет неравенству

$$D(\hat{\theta}) = M[(\hat{\theta} - \theta)^2] \geq \frac{\left(1 + \frac{db}{d\theta}\right)^2}{\int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial \theta}\right)^2 L(x, \theta) dx}.$$

**Доказательство.** Имеем

$$\int L(\hat{\theta}|\theta) d\hat{\theta} = 1.$$

Дифференцируем это выражение по  $\theta$ :

$$\int \frac{\partial L}{\partial \theta} d\hat{\theta} = 0 = \int \frac{\partial \ln L}{\partial \theta} L(\hat{\theta}|\theta) d\hat{\theta}, \text{ так как } \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}.$$

Имеем

$$M(\hat{\theta}) = \int \hat{\theta} \cdot L(\hat{\theta}|\theta) d\hat{\theta} = \theta + b(\theta);$$

Аналогично под  $f(x_i, \vec{\theta})$  можно понимать вероятность того, что дискретная случайная величина  $X$  принимает значение  $y_i = f(x_i, \theta) + \varepsilon_i$ ,  $i = 1, \dots, n$ , т.е.

$$f(x_i, \vec{\theta}) = P(X = x_i).$$

Можно найти такое число (оценку)  $\hat{\theta}$ , которое обращает функцию правдоподобия в максимум;  $\hat{\theta}$  и будет точечной оценкой параметра  $\vec{\theta}$ , т.е. для вычисления оценки  $\hat{\theta}$  надо решить уравнение  $\frac{\partial L(X, \theta)}{\partial \theta} = 0$ .

Будем считать, что оценки  $\hat{\theta}$  удовлетворяют следующим условиям (условиям регулярности):

1) пусть  $\hat{\theta} = \theta(x_1, x_2, \dots, x_n)$  – однозначная функция от  $x_1, x_2, \dots, x_n$  непрерывная и имеющая непрерывные частные производные  $\frac{\partial \theta}{\partial x_i}$  во всех точках  $X = (x_1, x_2, \dots, x_n) \in R_n$  и операторы  $\frac{\partial}{\partial \theta}$  и  $\int$  коммутируют;

2) область значений  $X$  зависит от параметра  $\theta$ .

Предположим, что математическое ожидание оценки  $M[\hat{\theta}] = \theta + b(\theta)$ , где  $b(\theta)$  – смещение оценки  $\hat{\theta}$ .

**Теорема.** В каждом случае регулярной оценки непрерывного типа квадрат среднего квадратического отклонения оценки  $\hat{\theta}$  от истинного значения  $\theta$  удовлетворяет неравенству

$$D(\hat{\theta}) = M[(\hat{\theta} - \theta)^2] \geq \frac{\left(1 + \frac{db}{d\theta}\right)^2}{\int_{-\infty}^{\infty} \left(\frac{\partial \ln L}{\partial \theta}\right)^2 L(x, \theta) dx}.$$

**Доказательство.** Имеем

$$\int L(\hat{\theta}|\theta) d\hat{\theta} = 1.$$

Дифференцируем это выражение по  $\theta$ :

$$\int \frac{\partial L}{\partial \theta} d\hat{\theta} = 0 = \int \frac{\partial \ln L}{\partial \theta} L(\hat{\theta}|\theta) d\hat{\theta}, \text{ так как } \frac{\partial \ln L}{\partial \theta} = \frac{1}{L} \frac{\partial L}{\partial \theta}.$$

Имеем

$$M(\hat{\theta}) = \int \hat{\theta} \cdot L(\hat{\theta}|\theta) d\hat{\theta} = \theta + b(\theta);$$

2) полученный интервал может иметь очень широкие и нецентральные границы для первоначальной переменной;

3) если функция правдоподобия имеет несколько максимумов, то эта процедура дает несколько интервалов; интерпретация такого утверждения в качестве доверительного интервала для  $\theta$  до некоторой степени сомнительна: можно было бы найти один непрерывный интервал с тем же самым вероятностным содержанием.

Не следует забывать, что интервальная оценка указывает интервал, который с вероятностью  $\beta$  накрывает истинное значение  $\theta$ . Но можно указать, строго говоря, множество интервалов, имеющих вероятностное содержание  $\beta$ , изменяя значения  $\theta$  под кривой плотности распределения  $f(\hat{\theta}|\theta)$ ;

4) когда предложенный метод дает неопределенные (или бесконечные) доверительные интервалы, вероятнее всего, используемое преобразование недостаточно хорошо по своей природе. Возможно, что для этой задачи требуется более сложная интерпретация, чем интерпретация ее в виде единственной оценки ММП и доверительного интервала.

## 2.7. Байесовский подход к оцениванию параметров моделей

В байесовском подходе неизвестный параметр  $\theta$  – не детерминированная (постоянная) величина, а случайная величина, для которой априори известно распределение  $\pi(\theta)$  возможных значений  $\theta$ .

После наблюдения  $X$  получим апостериорное распределение

$$\pi(\theta|X) = \frac{\left[ \prod_{i=1}^n f(x_i|\theta) \right] \pi(\theta)}{\int \left[ \prod_{i=1}^n f(x_i|\theta) \right] \pi(\theta) d\theta}.$$

Байесовский доверительный интервал  $(\theta_H, \theta_B)$  с вероятностным содержанием  $\beta$  определяется как

$$\int_{\theta_H}^{\theta_B} \pi(\theta|X) d\theta = \beta.$$

Согласно этому определению доверительный интервал соответствует доле  $\beta$  полной веры наблюдается в параметр  $\theta$ , т.е. на языке пари шанс на выигрыш составляет  $\beta/(1-\beta)$ .

Полное апостериорное распределение  $\pi(\theta|X)$  как бы заменяется дискретным двухточечным распределением:

$$P(\theta_{\alpha} \leq \theta \leq \theta_{\beta}) = \beta; \quad P\{\theta \notin [\theta_{\alpha}, \theta_{\beta}]\} = 1 - \beta.$$

Как и в классическом варианте, интервал определяется неоднозначно. В обоих случаях не указывается, какое значение внутри интервала является наиболее вероятным. Поэтому желательно дополнительно провести точечную оценку по максимуму  $\pi(\theta|X)$ .

Замечание. Для нелинейных функций оценку свободных параметров часто проводят после линеаризации заданной функции, например,  $\eta = Ae^{\lambda x} \Rightarrow \ln \eta = \ln A + \lambda x$  или  $\eta_1 = \theta_1 + \theta_2 x$ , где  $\eta_1 = \ln \eta$ ,  $\theta_1 = \ln A$ ,  $\theta_2 = \lambda$ . Но здесь возникают две проблемы: единственности оценок и изменения законов распределения исходных данных, что повлияет на величину интервальных оценок.

### 2.8. Интервальная оценка линии регрессии и прогнозируемых значений функции

Пусть для функции  $\eta = f(\vec{x}, \vec{\theta})$  получены оценки  $\vec{\hat{\theta}}$  и их дисперсии  $D(\vec{\hat{\theta}})$ . Это означает, что по результатам наблюдений можно построить целое семейство линий. Для модели  $\eta = f(\vec{x}, \vec{\theta})$  оценкой для линии регрессии  $\eta$  будет  $\hat{y} = f(\vec{x}, \vec{\hat{\theta}})$ . Надо теперь определить доверительные интервалы для значений  $\hat{y}$ , которые с заданной доверительной вероятностью  $\beta$  накроют истинные значения  $\eta$ . Кроме того, нужно найти доверительные интервалы для прогнозов, которые накроют новое наблюдение  $y$ , соответствующее заданному значению  $x$ .

Рассмотрим эти оценки для случая прямой линии  $\eta = \theta_0 + \theta_1(x - \bar{x})$ , где  $\bar{x}$  – среднее значение переменной (числа)  $x$ . Тогда дисперсия

$$D(\hat{y}) = D(\hat{\theta}_0) + D[\hat{\theta}_1(x - \bar{x})] = D(\hat{\theta}_0) + (x - \bar{x})^2 D(\hat{\theta}_1).$$

Величина  $\frac{|\hat{y} - \eta|}{\sqrt{D(\hat{y})}}$  подчиняется распределению Стьюдента с  $n - p$

степенями свободы; при заданном  $\beta$  квантиль распределения Стьюдента равна  $t_{\beta}$  и интервальная оценка имеет вид



$$P \left[ \hat{y} - t_\gamma \sqrt{D(\hat{y})} < \eta < \hat{y} + t_\gamma \sqrt{D(\hat{y})} \right] = \beta.$$

В общем случае для линии регрессии, описываемой уравнением  $\eta(x, \theta) = f(x, \theta)$ , дисперсия наблюдаемых значений  $\hat{y}$  определяется выражением

$$D(\hat{y}) = \sum_{j=1}^N \left( \frac{\partial f}{\partial \theta_j} \right)^2 D(\hat{\theta}_j) + 2 \sum_{i>j}^N \frac{\partial f}{\partial \theta_j} \frac{\partial f}{\partial \theta_i} D(\hat{\theta}_j, \hat{\theta}_i),$$

где  $N$  – размерность вектора  $\theta$ ;  $D(\hat{\theta}_j, \hat{\theta}_i)$  – корреляционный момент случайных величин  $\hat{\theta}_j$  и  $\hat{\theta}_i$ .

Для линейных функций  $\eta(x, \theta) = \sum_{j=1}^N \theta_j f_j(x)$  дисперсия наблюдаемых значений  $\hat{y}$  в точке  $x$  согласно той же формуле может быть записана в виде

$$D(\hat{y}) = f^T(x) D(\hat{\theta}) f(x).$$

Например, для прямой линии  $\eta(x, \theta) = \theta_0 + \theta_1(x - \bar{x})$  вектор  $f(x) = \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix}$  и

$$D(\hat{y}) = \begin{pmatrix} 1 & x - \bar{x} \end{pmatrix} \begin{pmatrix} D(\hat{\theta}_0) & 0 \\ 0 & D(\hat{\theta}_1) \end{pmatrix} \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix} = D(\hat{\theta}_0) + (x - \bar{x})^2 D(\hat{\theta}_1).$$

Задав серию значений  $x$  и вычислив соответствующие им значения  $\sigma^2(\hat{y}) = D(\hat{y})$ , найдем интервальную оценку линии. Для прямой линии интервальную оценку можно получить аналитически [17]. Интервальные границы в случае прямой линии для  $M(\eta | \xi_0)$  будут

$$\hat{\eta} - t_{1-\alpha/2} \sigma(\eta) \leq M(\eta | \xi_0) < \hat{\eta} + t_{1-\alpha/2} \sigma(\eta),$$

где  $\sigma^2(\hat{\eta}) = (1, \xi_0) D(\hat{\theta}) (1, \xi_0)^T$ ,  $\alpha$  – уровень значимости.

Огибающая семейства всех возможных прямых будет кривой второго порядка, уравнение которой получим из последней формулы

$$\sum_{i=1}^n (y_i - \theta_1 - \theta_2 x_i)^2 \leq C_v (1 + \theta_2^2).$$

Принимая за начало отсчета наблюдаемые средние значения  $x$  и  $y$ , перепишем это неравенство в виде [17]

$$\theta_2^2 (S_x^2 - C) - 2\theta_2 S_{xy} + \theta_1^2 \leq C_v - S_y^2,$$

$$P \left[ \hat{y} - t_\gamma \sqrt{D(\hat{y})} < \eta < \hat{y} + t_\gamma \sqrt{D(\hat{y})} \right] = \beta.$$

В общем случае для линии регрессии, описываемой уравнением  $\eta(x, \theta) = f(x, \theta)$ , дисперсия наблюдаемых значений  $\hat{y}$  определяется выражением

$$D(\hat{y}) = \sum_{j=1}^N \left( \frac{\partial f}{\partial \theta_j} \right)^2 D(\hat{\theta}_j) + 2 \sum_{i>j}^N \frac{\partial f}{\partial \theta_j} \frac{\partial f}{\partial \theta_i} D(\hat{\theta}_j, \hat{\theta}_i),$$

где  $N$  – размерность вектора  $\theta$ ;  $D(\hat{\theta}_j, \hat{\theta}_i)$  – корреляционный момент случайных величин  $\hat{\theta}_j$  и  $\hat{\theta}_i$ .

Для линейных функций  $\eta(x, \theta) = \sum_{j=1}^N \theta_j f_j(x)$  дисперсия наблюдаемых значений  $\hat{y}$  в точке  $x$  согласно той же формуле может быть записана в виде

$$D(\hat{y}) = f^T(x) D(\hat{\theta}) f(x).$$

Например, для прямой линии  $\eta(x, \theta) = \theta_0 + \theta_1(x - \bar{x})$  вектор  $f(x) = \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix}$  и

$$D(\hat{y}) = \begin{pmatrix} 1 & x - \bar{x} \end{pmatrix} \begin{pmatrix} D(\hat{\theta}_0) & 0 \\ 0 & D(\hat{\theta}_1) \end{pmatrix} \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix} = D(\hat{\theta}_0) + (x - \bar{x})^2 D(\hat{\theta}_1).$$

Задав серию значений  $x$  и вычислив соответствующие им значения  $\sigma^2(\hat{y}) = D(\hat{y})$ , найдем интервальную оценку линии. Для прямой линии интервальную оценку можно получить аналитически [17]. Интервальные границы в случае прямой линии для  $M(\eta | \xi_0)$  будут

$$\hat{\eta} - t_{1-\alpha/2} \sigma(\eta) \leq M(\eta | \xi_0) < \hat{\eta} + t_{1-\alpha/2} \sigma(\eta),$$

где  $\sigma^2(\hat{\eta}) = (1, \xi_0) D(\hat{\theta}) (1, \xi_0)^T$ ,  $\alpha$  – уровень значимости.

Огибающая семейства всех возможных прямых будет кривой второго порядка, уравнение которой получим из последней формулы

$$\sum_{i=1}^n (y_i - \theta_1 - \theta_2 x_i)^2 \leq C_v (1 + \theta_2^2).$$

Принимая за начало отсчета наблюдаемые средние значения  $x$  и  $y$ , перепишем это неравенство в виде [17]

$$\theta_2^2 (S_x^2 - C) - 2\theta_2 S_{xy} + \theta_1^2 \leq C_v - S_y^2,$$

$$P \left[ \hat{y} - t_\gamma \sqrt{D(\hat{y})} < \eta < \hat{y} + t_\gamma \sqrt{D(\hat{y})} \right] = \beta.$$

В общем случае для линии регрессии, описываемой уравнением  $\eta(x, \theta) = f(x, \theta)$ , дисперсия наблюдаемых значений  $\hat{y}$  определяется выражением

$$D(\hat{y}) = \sum_{j=1}^N \left( \frac{\partial f}{\partial \theta_j} \right)^2 D(\hat{\theta}_j) + 2 \sum_{i>j}^N \frac{\partial f}{\partial \theta_j} \frac{\partial f}{\partial \theta_i} D(\hat{\theta}_j, \hat{\theta}_i),$$

где  $N$  – размерность вектора  $\theta$ ;  $D(\hat{\theta}_j, \hat{\theta}_i)$  – корреляционный момент случайных величин  $\hat{\theta}_j$  и  $\hat{\theta}_i$ .

Для линейных функций  $\eta(x, \theta) = \sum_{j=1}^N \theta_j f_j(x)$  дисперсия наблюдаемых значений  $\hat{y}$  в точке  $x$  согласно той же формуле может быть записана в виде

$$D(\hat{y}) = f^T(x) D(\hat{\theta}) f(x).$$

Например, для прямой линии  $\eta(x, \theta) = \theta_0 + \theta_1(x - \bar{x})$  вектор  $f(x) = \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix}$  и

$$D(\hat{y}) = \begin{pmatrix} 1 & x - \bar{x} \end{pmatrix} \begin{pmatrix} D(\hat{\theta}_0) & 0 \\ 0 & D(\hat{\theta}_1) \end{pmatrix} \begin{pmatrix} 1 \\ x - \bar{x} \end{pmatrix} = D(\hat{\theta}_0) + (x - \bar{x})^2 D(\hat{\theta}_1).$$

Задав серию значений  $x$  и вычислив соответствующие им значения  $\sigma^2(\hat{y}) = D(\hat{y})$ , найдем интервальную оценку линии. Для прямой линии интервальную оценку можно получить аналитически [17]. Интервальные границы в случае прямой линии для  $M(\eta | \xi_0)$  будут

$$\hat{\eta} - t_{1-\alpha/2} \sigma(\eta) \leq M(\eta | \xi_0) < \hat{\eta} + t_{1-\alpha/2} \sigma(\eta),$$

где  $\sigma^2(\hat{\eta}) = (1, \xi_0) D(\hat{\theta}) (1, \xi_0)^T$ ,  $\alpha$  – уровень значимости.

Огибающая семейства всех возможных прямых будет кривой второго порядка, уравнение которой получим из последней формулы

$$\sum_{i=1}^n (y_i - \theta_1 - \theta_2 x_i)^2 \leq C_v (1 + \theta_2^2).$$

Принимая за начало отсчета наблюдаемые средние значения  $x$  и  $y$ , перепишем это неравенство в виде [17]

$$\theta_2^2 (S_x^2 - C) - 2\theta_2 S_{xy} + \theta_1^2 \leq C_v - S_y^2,$$

## Глава 3

# АКТИВНЫЙ И ПАССИВНЫЙ ЭКСПЕРИМЕНТЫ

И одной матки детки, да не равны.  
Русская народная пословица

### 3.1. Оценивание параметров функции известного вида в пассивном эксперименте

Ниже будут рассмотрены алгоритмы, позволяющие находить оценки свободных параметров элементарных функций при наличии погрешностей и в значениях аргумента, и в значениях функции. Эти алгоритмы объединены под названием «конфлюэнтный анализ» (*confluence* – слияние, *confluent* – сливающийся). Когда и значения функции, и значения аргумента – случайные величины, процедуры оценки свободных параметров функции зависят от того, как проводится эксперимент: пассивно считывались значения аргумента и функции или значения аргумента устанавливались заранее [1, 3, 7, 15, 16, 18, 20].

Пассивный эксперимент наиболее часто встречается в процессах получения оценок параметров моделей систем во многих отраслях науки и техники. Здесь требуется найти интервальную оценку параметра  $\vec{\theta}$  функции  $\vec{\eta} = f(\vec{\xi}, \vec{\theta})$ , когда точные значения  $\vec{\eta}$  и  $\vec{\xi}$  наблюдать нельзя, но можно наблюдать значения случайных величин  $\vec{Y}$  и  $\vec{X}$ , определяемые следующим образом:

$$x_i = \xi_i + \delta_i; \quad y_i = \eta_i + \epsilon_i; \quad i = 1, 2, \dots, n,$$

где  $\delta_i$  и  $\epsilon_i$  – соответственно ошибки значений переменных и функции (случайные величины).

В активном эксперименте можно заставить  $X$  принимать любые значения ( $X$  называется контролируемой переменной), на которые накладывается помеха  $\delta$ :  $\xi = X + \delta$ . Значения  $Y$  есть результат влияния случайной ошибки на истинное значение  $\eta$ :  $Y = \eta + \epsilon$ .

Структурное соотношение между наблюдаемыми значениями  $x_i$  и  $y_i$  примет вид  $Y = f(X + \delta, \theta) + \varepsilon$  и является обычным уравнением регрессии, поскольку теперь  $X$  не случайная величина.

Рассмотрим пассивную схему эксперимента. Пусть имеем статистический ряд экспериментальных значений  $\{x_i\} \in X$  и соответствующий им ряд значений функций  $\{y_i\} \in Y$ ;  $i = 1, 2, \dots, n$ ;  $n \geq m$ , где  $m$  – число оцениваемых параметров  $\vec{\theta}$ . Будем считать, что переменные  $x_i$  и  $y_i$  не детерминированы, но являются выборками из генеральных совокупностей  $X$  и  $Y$  с известными плотностями функции вероятностей. Переменные  $x_i = \xi_i + \delta_i$  и  $y_i = \eta_i + \varepsilon_i$  могут быть статистически как зависимы, так и независимы; могут быть как коррелированы, так и не коррелированы.

В основном мы будем иметь дело с выборками из независимых наблюдений из одного и того же распределения. Пусть  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  – соответственно плотности вероятности случайных величин  $x_i$  и  $y_i$ , если  $x_i$  и  $y_i$  непрерывны, и соответственно вероятности значений  $x_i$  и  $y_i$ , если распределения  $x_i$  и  $y_i$  дискретны; для  $x_i$ ,  $y_i$  распределения  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  могут быть как одномерными, так и многомерными.

Найдем выражение для совместной плотности вероятности экспериментальных данных при условии, что значения  $\varepsilon_i$  и  $\eta_i$  связаны функциональной зависимостью, но их погрешности  $\delta_i$  и  $\varepsilon_i$  являются независимыми при переходе от одной точки к другой. Тогда совместная плотность вероятности получить одновременно значения  $x_i$  и  $y_i$

$$P_i = f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Совместная плотность вероятности получить  $n$  статистически независимых точек  $(x_i, y_i)$

$$L(x, y | \vec{\theta}) = \prod_{i=1}^n f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Аналогично можно вывести формулы совместной плотности вероятности для зависимых или коррелированных экспериментальных точек. Эти формулы могут быть взяты из книг по теории вероятностей.

Для нас важен тот факт, что в выражения для совместной плотности вероятности входят математические ожидания экспериментальных данных, экспериментальные значения и оцениваемые па-

Структурное соотношение между наблюдаемыми значениями  $x_i$  и  $y_i$  примет вид  $Y = f(X + \delta, \theta) + \varepsilon$  и является обычным уравнением регрессии, поскольку теперь  $X$  не случайная величина.

Рассмотрим пассивную схему эксперимента. Пусть имеем статистический ряд экспериментальных значений  $\{x_i\} \in X$  и соответствующий им ряд значений функций  $\{y_i\} \in Y$ ;  $i = 1, 2, \dots, n$ ;  $n \geq m$ , где  $m$  – число оцениваемых параметров  $\vec{\theta}$ . Будем считать, что переменные  $x_i$  и  $y_i$  не детерминированы, но являются выборками из генеральных совокупностей  $X$  и  $Y$  с известными плотностями функции вероятностей. Переменные  $x_i = \xi_i + \delta_i$  и  $y_i = \eta_i + \varepsilon_i$  могут быть статистически как зависимы, так и независимы; могут быть как коррелированы, так и не коррелированы.

В основном мы будем иметь дело с выборками из независимых наблюдений из одного и того же распределения. Пусть  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  – соответственно плотности вероятности случайных величин  $x_i$  и  $y_i$ , если  $x_i$  и  $y_i$  непрерывны, и соответственно вероятности значений  $x_i$  и  $y_i$ , если распределения  $x_i$  и  $y_i$  дискретны; для  $x_i$ ,  $y_i$  распределения  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  могут быть как одномерными, так и многомерными.

Найдем выражение для совместной плотности вероятности экспериментальных данных при условии, что значения  $\varepsilon_i$  и  $\eta_i$  связаны функциональной зависимостью, но их погрешности  $\delta_i$  и  $\varepsilon_i$  являются независимыми при переходе от одной точки к другой. Тогда совместная плотность вероятности получить одновременно значения  $x_i$  и  $y_i$

$$P_i = f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Совместная плотность вероятности получить  $n$  статистически независимых точек  $(x_i, y_i)$

$$L(x, y | \vec{\theta}) = \prod_{i=1}^n f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Аналогично можно вывести формулы совместной плотности вероятности для зависимых или коррелированных экспериментальных точек. Эти формулы могут быть взяты из книг по теории вероятностей.

Для нас важен тот факт, что в выражения для совместной плотности вероятности входят математические ожидания экспериментальных данных, экспериментальные значения и оцениваемые па-

Структурное соотношение между наблюдаемыми значениями  $x_i$  и  $y_i$  примет вид  $Y = f(X + \delta, \theta) + \varepsilon$  и является обычным уравнением регрессии, поскольку теперь  $X$  не случайная величина.

Рассмотрим пассивную схему эксперимента. Пусть имеем статистический ряд экспериментальных значений  $\{x_i\} \in X$  и соответствующий им ряд значений функций  $\{y_i\} \in Y$ ;  $i = 1, 2, \dots, n$ ;  $n \geq m$ , где  $m$  – число оцениваемых параметров  $\vec{\theta}$ . Будем считать, что переменные  $x_i$  и  $y_i$  не детерминированы, но являются выборками из генеральных совокупностей  $X$  и  $Y$  с известными плотностями функции вероятностей. Переменные  $x_i = \xi_i + \delta_i$  и  $y_i = \eta_i + \varepsilon_i$  могут быть статистически как зависимы, так и независимы; могут быть как коррелированы, так и не коррелированы.

В основном мы будем иметь дело с выборками из независимых наблюдений из одного и того же распределения. Пусть  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  – соответственно плотности вероятности случайных величин  $x_i$  и  $y_i$ , если  $x_i$  и  $y_i$  непрерывны, и соответственно вероятности значений  $x_i$  и  $y_i$ , если распределения  $x_i$  и  $y_i$  дискретны; для  $x_i$ ,  $y_i$  распределения  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  могут быть как одномерными, так и многомерными.

Найдем выражение для совместной плотности вероятности экспериментальных данных при условии, что значения  $\varepsilon_i$  и  $\eta_i$  связаны функциональной зависимостью, но их погрешности  $\delta_i$  и  $\varepsilon_i$  являются независимыми при переходе от одной точки к другой. Тогда совместная плотность вероятности получить одновременно значения  $x_i$  и  $y_i$

$$P_i = f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Совместная плотность вероятности получить  $n$  статистически независимых точек  $(x_i, y_i)$

$$L(x, y | \vec{\theta}) = \prod_{i=1}^n f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Аналогично можно вывести формулы совместной плотности вероятности для зависимых или коррелированных экспериментальных точек. Эти формулы могут быть взяты из книг по теории вероятностей.

Для нас важен тот факт, что в выражения для совместной плотности вероятности входят математические ожидания экспериментальных данных, экспериментальные значения и оцениваемые па-

Структурное соотношение между наблюдаемыми значениями  $x_i$  и  $y_i$  примет вид  $Y = f(X + \delta, \theta) + \varepsilon$  и является обычным уравнением регрессии, поскольку теперь  $X$  не случайная величина.

Рассмотрим пассивную схему эксперимента. Пусть имеем статистический ряд экспериментальных значений  $\{x_i\} \in X$  и соответствующий им ряд значений функций  $\{y_i\} \in Y$ ;  $i = 1, 2, \dots, n$ ;  $n \geq m$ , где  $m$  – число оцениваемых параметров  $\vec{\theta}$ . Будем считать, что переменные  $x_i$  и  $y_i$  не детерминированы, но являются выборками из генеральных совокупностей  $X$  и  $Y$  с известными плотностями функции вероятностей. Переменные  $x_i = \xi_i + \delta_i$  и  $y_i = \eta_i + \varepsilon_i$  могут быть статистически как зависимы, так и независимы; могут быть как коррелированы, так и не коррелированы.

В основном мы будем иметь дело с выборками из независимых наблюдений из одного и того же распределения. Пусть  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  – соответственно плотности вероятности случайных величин  $x_i$  и  $y_i$ , если  $x_i$  и  $y_i$  непрерывны, и соответственно вероятности значений  $x_i$  и  $y_i$ , если распределения  $x_i$  и  $y_i$  дискретны; для  $x_i$ ,  $y_i$  распределения  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  могут быть как одномерными, так и многомерными.

Найдем выражение для совместной плотности вероятности экспериментальных данных при условии, что значения  $\varepsilon_i$  и  $\eta_i$  связаны функциональной зависимостью, но их погрешности  $\delta_i$  и  $\varepsilon_i$  являются независимыми при переходе от одной точки к другой. Тогда совместная плотность вероятности получить одновременно значения  $x_i$  и  $y_i$

$$P_i = f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Совместная плотность вероятности получить  $n$  статистически независимых точек  $(x_i, y_i)$

$$L(x, y | \vec{\theta}) = \prod_{i=1}^n f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Аналогично можно вывести формулы совместной плотности вероятности для зависимых или коррелированных экспериментальных точек. Эти формулы могут быть взяты из книг по теории вероятностей.

Для нас важен тот факт, что в выражения для совместной плотности вероятности входят математические ожидания экспериментальных данных, экспериментальные значения и оцениваемые па-



Структурное соотношение между наблюдаемыми значениями  $x_i$  и  $y_i$  примет вид  $Y = f(X + \delta, \theta) + \varepsilon$  и является обычным уравнением регрессии, поскольку теперь  $X$  не случайная величина.

Рассмотрим пассивную схему эксперимента. Пусть имеем статистический ряд экспериментальных значений  $\{x_i\} \in X$  и соответствующий им ряд значений функций  $\{y_i\} \in Y$ ;  $i = 1, 2, \dots, n$ ;  $n \geq m$ , где  $m$  – число оцениваемых параметров  $\vec{\theta}$ . Будем считать, что переменные  $x_i$  и  $y_i$  не детерминированы, но являются выборками из генеральных совокупностей  $X$  и  $Y$  с известными плотностями функции вероятностей. Переменные  $x_i = \xi_i + \delta_i$  и  $y_i = \eta_i + \varepsilon_i$  могут быть статистически как зависимы, так и независимы; могут быть как коррелированы, так и не коррелированы.

В основном мы будем иметь дело с выборками из независимых наблюдений из одного и того же распределения. Пусть  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  – соответственно плотности вероятности случайных величин  $x_i$  и  $y_i$ , если  $x_i$  и  $y_i$  непрерывны, и соответственно вероятности значений  $x_i$  и  $y_i$ , если распределения  $x_i$  и  $y_i$  дискретны; для  $x_i$ ,  $y_i$  распределения  $f_1(x_i | \theta)$  и  $f_2(y_i | \theta)$  могут быть как одномерными, так и многомерными.

Найдем выражение для совместной плотности вероятности экспериментальных данных при условии, что значения  $\varepsilon_i$  и  $\eta_i$  связаны функциональной зависимостью, но их погрешности  $\delta_i$  и  $\varepsilon_i$  являются независимыми при переходе от одной точки к другой. Тогда совместная плотность вероятности получить одновременно значения  $x_i$  и  $y_i$

$$P_i = f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Совместная плотность вероятности получить  $n$  статистически независимых точек  $(x_i, y_i)$

$$L(x, y | \vec{\theta}) = \prod_{i=1}^n f_1(x_i | \vec{\theta}) \cdot f_2(y_i | \vec{\theta}).$$

Аналогично можно вывести формулы совместной плотности вероятности для зависимых или коррелированных экспериментальных точек. Эти формулы могут быть взяты из книг по теории вероятностей.

Для нас важен тот факт, что в выражения для совместной плотности вероятности входят математические ожидания экспериментальных данных, экспериментальные значения и оцениваемые па-

В случае, когда функция  $\eta$  описывает плотность нормального закона распределения, возникают трудности нахождения оценок параметров, так как функционал (3.2) в данном случае имеет сложный вид. Поэтому оценки параметров и истинные значения  $\xi_i$  определяют итерационными методами, предварительно определив «хорошее» нулевое приближение [4–7].

### 3.2. Анализ систем в активном эксперименте

Пусть проводится эксперимент, ставящий своей целью определение оценок свободных параметров  $\theta$  функции  $\eta = f(\xi, \theta)$ , описывающей некоторое распределение. Причем мы можем заставить  $X$  принимать любые заданные значения и затем измерить соответствующие им значения  $Y$ , которые можно рассматривать как результат влияния случайной ошибки  $\epsilon$  на истинное значение  $\eta$ :  $Y = \eta + \epsilon$ .

Переменная  $X$  называется контролируемой переменной, ее значение фиксировано заранее. Однако и на значения  $X$  может накладываться случайная ошибка  $\delta$ :  $\xi = X + \delta$ , т.е. неизвестное истинное значение  $\xi$  есть случайная величина. Пусть ошибки  $\delta$  имеют нулевое среднее значение. Тогда среднее значение  $\xi$  будет равно  $X$ , коэффициент корреляции между  $\xi$  и  $\delta$  равен  $+1$  и структурное соотношение имеет вид

$$Y = f(X + \delta, \theta) + \epsilon. \quad (3.7)$$

Поскольку  $X$  не является случайной величиной, то ни  $\epsilon$ , ни  $\delta$  не коррелированы с  $X$ . Таким образом, выражение (3.7) представляет собой обычное уравнение регрессии, к которому без всяких изменений можно применять методы регрессионного анализа, например МНК. Закон распределения случайной величины  $f(X + \delta, \theta)$  в выражении (3.7) может быть рассчитан по формулам переноса ошибок. Затем, в соответствии с полученным законом распределения случайной величины  $f(X + \delta, \theta)$ , с помощью ММП можно составить функционал, координаты точки минимума которого будут определять оценки искоемых параметров  $\theta$ .

Если распределение случайной величины  $f(X + \delta, \theta)$  подчиняется нормальному закону, то для оценки параметров  $\theta$  получим функционал метода наименьших квадратов, но дисперсия случайной вели-

чины  $f(X + \delta, \theta)$  будет зависеть от значений  $\theta$ , т.е. соответствующие уравнения для определения оценок  $\theta$  будут нелинейными.

Если контролируемая переменная  $X$  сама является случайной величиной (т.е. определяется с помощью некоторого процесса случайного выбора), полученные выше выводы останутся в силе, когда  $\delta$  и  $\varepsilon$  некоррелированы с  $X$ . Предположение относительно  $\varepsilon$  обычно выполняется, но относительно  $\delta$  ситуация сложнее. Например, требование некоррелированности  $\delta$  и  $X$  в данном случае означает, что с большими и малыми значениями  $X$  не должна быть связана тенденция к большим или меньшим ошибкам в определении истинного значения. Выполнение этого условия может быть проверено только эмпирическим путем.

Заметим, что при неизвестной дисперсии  $\sigma^2(\delta)$  или ее оценки в активной схеме эксперимента не идентифицируемы даже свободные параметры кубической параболы.

Рассмотрим в качестве примера задачу оценки свободных параметров  $\theta$  в уравнении прямой линии

$$\eta = \theta_1 + \theta_2 \xi.$$

Выразим  $\eta$  и  $\xi$  через  $X$  и  $Y$ , получим структурное соотношение

$$y_i = \theta_1 + \theta_2 x_i + \theta_2 \delta_i + \varepsilon_i.$$

Пусть независимые случайные величины  $\delta$  и  $\varepsilon$  подчиняются нормальному закону распределения с нулевыми математическими ожиданиями и известными дисперсиями  $\sigma^2(x_i)$  и  $\sigma^2(y_i)$ . Функции плотности вероятности случайных величин  $\theta_2 \delta_i + \varepsilon_i$  будут иметь математическое ожидание

$$E[\theta_2 \delta_i + \varepsilon_i] = 0$$

и дисперсию

$$\sigma^2[\theta_2 \delta_i + \varepsilon_i] = \theta_2^2 \sigma^2(x_i) + \sigma^2(y_i); \quad i = 1, 2, \dots, n.$$

В данном случае функционал МНК, точка минимума которого определяет оценки параметров  $\theta$ , имеет вид

$$F = \sum \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma^2(y_i) + \theta_2^2 \sigma^2(x_i)}. \quad (3.8)$$

Рассуждая аналогично, получим функционал метода наименьших квадратов для оценки параметров  $\theta$  линейных функций:

чины  $f(X + \delta, \theta)$  будет зависеть от значений  $\theta$ , т.е. соответствующие уравнения для определения оценок  $\theta$  будут нелинейными.

Если контролируемая переменная  $X$  сама является случайной величиной (т.е. определяется с помощью некоторого процесса случайного выбора), полученные выше выводы останутся в силе, когда  $\delta$  и  $\varepsilon$  некоррелированы с  $X$ . Предположение относительно  $\varepsilon$  обычно выполняется, но относительно  $\delta$  ситуация сложнее. Например, требование некоррелированности  $\delta$  и  $X$  в данном случае означает, что с большими и малыми значениями  $X$  не должна быть связана тенденция к большим или меньшим ошибкам в определении истинного значения. Выполнение этого условия может быть проверено только эмпирическим путем.

Заметим, что при неизвестной дисперсии  $\sigma^2(\delta)$  или ее оценки в активной схеме эксперимента не идентифицируемы даже свободные параметры кубической параболы.

Рассмотрим в качестве примера задачу оценки свободных параметров  $\theta$  в уравнении прямой линии

$$\eta = \theta_1 + \theta_2 \xi.$$

Выразим  $\eta$  и  $\xi$  через  $X$  и  $Y$ , получим структурное соотношение

$$y_i = \theta_1 + \theta_2 x_i + \theta_2 \delta_i + \varepsilon_i.$$

Пусть независимые случайные величины  $\delta$  и  $\varepsilon$  подчиняются нормальному закону распределения с нулевыми математическими ожиданиями и известными дисперсиями  $\sigma^2(x_i)$  и  $\sigma^2(y_i)$ . Функции плотности вероятности случайных величин  $\theta_2 \delta_i + \varepsilon_i$  будут иметь математическое ожидание

$$E[\theta_2 \delta_i + \varepsilon_i] = 0$$

и дисперсию

$$\sigma^2[\theta_2 \delta_i + \varepsilon_i] = \theta_2^2 \sigma^2(x_i) + \sigma^2(y_i); \quad i = 1, 2, \dots, n.$$

В данном случае функционал МНК, точка минимума которого определяет оценки параметров  $\theta$ , имеет вид

$$F = \sum \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma^2(y_i) + \theta_2^2 \sigma^2(x_i)}. \quad (3.8)$$

Рассуждая аналогично, получим функционал метода наименьших квадратов для оценки параметров  $\theta$  линейных функций:

чины  $f(X + \delta, \theta)$  будет зависеть от значений  $\theta$ , т.е. соответствующие уравнения для определения оценок  $\theta$  будут нелинейными.

Если контролируемая переменная  $X$  сама является случайной величиной (т.е. определяется с помощью некоторого процесса случайного выбора), полученные выше выводы останутся в силе, когда  $\delta$  и  $\varepsilon$  некоррелированы с  $X$ . Предположение относительно  $\varepsilon$  обычно выполняется, но относительно  $\delta$  ситуация сложнее. Например, требование некоррелированности  $\delta$  и  $X$  в данном случае означает, что с большими и малыми значениями  $X$  не должна быть связана тенденция к большим или меньшим ошибкам в определении истинного значения. Выполнение этого условия может быть проверено только эмпирическим путем.

Заметим, что при неизвестной дисперсии  $\sigma^2(\delta)$  или ее оценки в активной схеме эксперимента не идентифицируемы даже свободные параметры кубической параболы.

Рассмотрим в качестве примера задачу оценки свободных параметров  $\theta$  в уравнении прямой линии

$$\eta = \theta_1 + \theta_2 \xi.$$

Выразим  $\eta$  и  $\xi$  через  $X$  и  $Y$ , получим структурное соотношение

$$y_i = \theta_1 + \theta_2 x_i + \theta_2 \delta_i + \varepsilon_i.$$

Пусть независимые случайные величины  $\delta$  и  $\varepsilon$  подчиняются нормальному закону распределения с нулевыми математическими ожиданиями и известными дисперсиями  $\sigma^2(x_i)$  и  $\sigma^2(y_i)$ . Функции плотности вероятности случайных величин  $\theta_2 \delta_i + \varepsilon_i$  будут иметь математическое ожидание

$$E[\theta_2 \delta_i + \varepsilon_i] = 0$$

и дисперсию

$$\sigma^2[\theta_2 \delta_i + \varepsilon_i] = \theta_2^2 \sigma^2(x_i) + \sigma^2(y_i); \quad i = 1, 2, \dots, n.$$

В данном случае функционал МНК, точка минимума которого определяет оценки параметров  $\theta$ , имеет вид

$$F = \sum \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma^2(y_i) + \theta_2^2 \sigma^2(x_i)}. \quad (3.8)$$

Рассуждая аналогично, получим функционал метода наименьших квадратов для оценки параметров  $\theta$  линейных функций:

## Глава 4

### ПРИМЕРЫ ДРУГИХ МЕТОДОВ В КОНФЛЮЭНТНОМ АНАЛИЗЕ

Электронные мозги будут  
ошибаться гораздо точнее.  
*Габриэль Лауб*

#### 4.1. Оценки параметров функции известного вида с учетом ошибок в значениях функций и аргументов

Предположим, что ошибки измерений  $\varepsilon_i$  и  $\delta_i$  – нормально распределенные случайные величины с нулевыми средними значениями, с дисперсиями  $\sigma^2(y_i)$  и  $\sigma^2(x_i)$  соответственно и коэффициентом корреляции  $\rho_i = 0$ .

Рассмотрим несколько других подходов к решению данной задачи.

I. Один из подходов подразумевает использование вместо истинных значений  $\xi_i$  значения наблюдаемых  $x_i$  при оценивании параметров МНК. То есть при игнорировании ошибки измерения  $\delta_i$  получаем следующую модель:

$$y = f(x, \theta) + \text{ошибка.}$$

В основном, как показано в [5, 10], этот подход дает несостоятельную оценку с большим асимптотическим смещением.

II. В работе [17] рассмотрен случай, когда  $x$  – случайная величина, независимая от  $\varepsilon_i$  и  $\delta_i$ , с характеристиками  $E(\xi) = \mu$ ,  $\text{cov}(\xi) = \Sigma_\xi$ . значения  $\xi_i$  аппроксимируются результатами измерений  $x_i$  и при нормальном законе распределения случайной величины  $\xi_i$

$$E(\xi(x)) = x\Lambda + \mu(I - \Lambda),$$

где  $\Lambda = (\Sigma_\xi + \sigma^2(x)I)^{-1} \Sigma_\xi$ ;  $I$  – единичная матрица.

Получаем модель

$$y = f[x\Lambda + \mu(I - \Lambda), \theta] + \text{ошибка.}$$

Когда  $\mu$  и  $\Lambda$  неизвестны, то, используя исходную выборку  $x_i, i = 1, \dots, n$ , определяем оценки  $\mu, \Lambda$  и в итоге получаем модель

$$y = f\left[x \cdot \hat{\Lambda} + \hat{\mu}(I - \hat{\Lambda}), \hat{\theta}\right] + \text{ошибка} = f(\hat{\xi}, \hat{\theta}) + \text{ошибка}. \quad (4.1)$$

Оценки параметров модели (4.1) могут быть определены МНК. Например, для прямой линии  $\eta = a\xi + b$ , имеем

$$\hat{\xi} = x \frac{\sigma_x^2}{\sigma_x^2 + \sigma_0^2} + \mu \left(1 - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_0^2}\right),$$

где  $\sigma_0^2$  – дисперсия ошибки  $\delta_i$ .

Минимизируемый функционал примет вид

$$F = \frac{\sum_{i=1}^n [y_i - \eta(\xi_i, \theta)]^2}{\sigma^2(y)} = \frac{\sum_{i=1}^n [(y_i - b - a\mu)(\sigma_x^2 + \sigma_0^2) - (x_i - \mu)a\sigma^2 x]^2}{(\sigma_x^2 + \sigma_0^2)^2 \cdot \sigma^2(y)}.$$

В случае, когда функция  $f(\xi, \theta)$  нелинейна по параметрам, используют разложение  $f(\xi, \theta)$  в ряд Тейлора в окрестности точки  $\hat{\xi}$  [10, 17]. Например, для нормального закона распределения с параметрами  $\theta_1$  и  $\theta_2$ , используя формулу  $\eta = f(\xi) + f'(\xi)(x - \xi)$  получим

$$\eta \approx \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} + \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} \times \left[-\frac{x - \theta_1}{\theta_2^2}\right]_{x=\hat{\xi}} (x - \hat{\xi}) = \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} \left[1 + \frac{\theta_1 - \hat{\xi}}{\theta_2^2} (x - \hat{\xi})\right].$$

Оценки параметров  $\theta_1$  и  $\theta_2$  найдем из условия минимума функционала

$$F = \frac{\sum_{i=1}^n [y_i - \eta(\xi_i, \theta)]^2}{\sigma^2(y)}.$$

III. Поставленная задача может быть решена итеративным методом наименьших квадратов с уточняемыми весами [10, 14, 19].

Предполагается, что  $x_i$  является выборкой из некоторой генеральной совокупности с функцией плотности распределения  $f(x_i, \theta)$ . При этом первые моменты функции  $f(x_i, \theta)$  известны и конечны. В этом случае для построения минимизируемого функционала  $F$  необходимо знать вид функции плотности распределения  $f(x_i, \theta)$  или ее моменты.

Когда  $\mu$  и  $\Lambda$  неизвестны, то, используя исходную выборку  $x_i, i = 1, \dots, n$ , определяем оценки  $\mu, \Lambda$  и в итоге получаем модель

$$y = f\left[x \cdot \hat{\Lambda} + \hat{\mu}(I - \hat{\Lambda}), \hat{\theta}\right] + \text{ошибка} = f(\hat{\xi}, \hat{\theta}) + \text{ошибка}. \quad (4.1)$$

Оценки параметров модели (4.1) могут быть определены МНК. Например, для прямой линии  $\eta = a\xi + b$ , имеем

$$\hat{\xi} = x \frac{\sigma_x^2}{\sigma_x^2 + \sigma_0^2} + \mu \left(1 - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_0^2}\right),$$

где  $\sigma_0^2$  – дисперсия ошибки  $\delta_i$ .

Минимизируемый функционал примет вид

$$F = \frac{\sum_{i=1}^n [y_i - \eta(\xi_i, \theta)]^2}{\sigma^2(y)} = \frac{\sum_{i=1}^n [(y_i - b - a\mu)(\sigma_x^2 + \sigma_0^2) - (x_i - \mu)a\sigma^2 x]^2}{(\sigma_x^2 + \sigma_0^2)^2 \cdot \sigma^2(y)}.$$

В случае, когда функция  $f(\xi, \theta)$  нелинейна по параметрам, используют разложение  $f(\xi, \theta)$  в ряд Тейлора в окрестности точки  $\hat{\xi}$  [10, 17]. Например, для нормального закона распределения с параметрами  $\theta_1$  и  $\theta_2$ , используя формулу  $\eta = f(\hat{\xi}) + f'(\hat{\xi})(x - \hat{\xi})$  получим

$$\eta \approx \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} + \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} \times \left[-\frac{x - \theta_1}{\theta_2^2}\right]_{x=\hat{\xi}} (x - \hat{\xi}) = \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} \left[1 + \frac{\theta_1 - \hat{\xi}}{\theta_2^2} (x - \hat{\xi})\right].$$

Оценки параметров  $\theta_1$  и  $\theta_2$  найдем из условия минимума функционала

$$F = \frac{\sum_{i=1}^n [y_i - \eta(\xi_i, \theta)]^2}{\sigma^2(y)}.$$

III. Поставленная задача может быть решена итеративным методом наименьших квадратов с уточняемыми весами [10, 14, 19].

Предполагается, что  $x_i$  является выборкой из некоторой генеральной совокупности с функцией плотности распределения  $f(x_i, \theta)$ . При этом первые моменты функции  $f(x_i, \theta)$  известны и конечны. В этом случае для построения минимизируемого функционала  $F$  необходимо знать вид функции плотности распределения  $f(x_i, \theta)$  или ее моменты.



Когда  $\mu$  и  $\Lambda$  неизвестны, то, используя исходную выборку  $x_i, i = 1, \dots, n$ , определяем оценки  $\mu, \Lambda$  и в итоге получаем модель

$$y = f\left[x \cdot \hat{\Lambda} + \hat{\mu}(I - \hat{\Lambda}), \hat{\theta}\right] + \text{ошибка} = f(\hat{\xi}, \hat{\theta}) + \text{ошибка}. \quad (4.1)$$

Оценки параметров модели (4.1) могут быть определены МНК. Например, для прямой линии  $\eta = a\xi + b$ , имеем

$$\hat{\xi} = x \frac{\sigma_x^2}{\sigma_x^2 + \sigma_0^2} + \mu \left(1 - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_0^2}\right),$$

где  $\sigma_0^2$  – дисперсия ошибки  $\delta_i$ .

Минимизируемый функционал примет вид

$$F = \frac{\sum_{i=1}^n [y_i - \eta(\xi_i, \theta)]^2}{\sigma^2(y)} = \frac{\sum_{i=1}^n [(y_i - b - a\mu)(\sigma_x^2 + \sigma_0^2) - (x_i - \mu)a\sigma^2 x]^2}{(\sigma_x^2 + \sigma_0^2)^2 \cdot \sigma^2(y)}.$$

В случае, когда функция  $f(\xi, \theta)$  нелинейна по параметрам, используют разложение  $f(\xi, \theta)$  в ряд Тейлора в окрестности точки  $\hat{\xi}$  [10, 17]. Например, для нормального закона распределения с параметрами  $\theta_1$  и  $\theta_2$ , используя формулу  $\eta = f(\xi) + f'(\xi)(x - \xi)$  получим

$$\eta \approx \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} + \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} \times \left[-\frac{x - \theta_1}{\theta_2^2}\right]_{x=\hat{\xi}} (x - \hat{\xi}) = \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x - \theta_1)^2}{2\theta_2^2}\right]_{x=\hat{\xi}} \left[1 + \frac{\theta_1 - \hat{\xi}}{\theta_2^2} (x - \hat{\xi})\right].$$

Оценки параметров  $\theta_1$  и  $\theta_2$  найдем из условия минимума функционала

$$F = \frac{\sum_{i=1}^n [y_i - \eta(\xi_i, \theta)]^2}{\sigma^2(y)}.$$

III. Поставленная задача может быть решена итеративным методом наименьших квадратов с уточняемыми весами [10, 14, 19].

Предполагается, что  $x_i$  является выборкой из некоторой генеральной совокупности с функцией плотности распределения  $f(x_i, \theta)$ . При этом первые моменты функции  $f(x_i, \theta)$  известны и конечны. В этом случае для построения минимизируемого функционала  $F$  необходимо знать вид функции плотности распределения  $f(x_i, \theta)$  или ее моменты.

Дисперсии оценок определяют по формулам

$$D(\hat{\xi}_i) = \left[ (a; I) \cdot \Sigma^{-1} \cdot (a; I)^T \right]^{-1}$$

или

$$D(\hat{\xi}_i) = \left[ a^2 \sigma_\delta^2 + \sigma_\varepsilon^2 \right]^{-1}.$$

Далее уточняют оценки параметров  $a$  и  $b$  при новых значениях аргументов  $\xi_i$ . Затем снова определяют оценки  $\xi_i$ .

Итерационный процесс продолжают до тех пор, пока относительные изменения параметров на соседних итерациях не будут меньше некоторой малой величины  $\gamma$ , т.е.

$$\max \left( \frac{\theta_j^s - \theta_j^{s-1}}{\theta_j^s} \right) \leq \gamma, \quad j = 1, 2.$$

Для линейных моделей мы получили функционал и формулу для расчета значений  $\xi_i$ , аналогичную приведенной в (3.6). Таким образом, в качестве единого подхода к задаче нахождения оценок параметров моделей с учетом погрешностей в значениях функции и аргументов могут быть использованы методы конфлюэнтного анализа, описанные в [4–9].

#### 4.2. Оценивание свободных параметров в уравнении прямой линии

Работы, посвященные оценке свободных параметров прямой линии с учетом погрешности по осям абсцисс и ординат, известны давно [15], и если какие-то результаты в конфлюэнтном анализе получены, то они получены для прямой линии. Родоначальником этих работ, по-видимому, следует считать Вальда [15].

Рассмотрим, каким образом определяются свободные параметры прямой линии  $\eta = \theta_1 + \theta_2 \xi$ , и на этом примере покажем основные особенности задач конфлюэнтного анализа.

При определении свободных параметров  $\theta_1$  и  $\theta_2$  уравнение прямой линии часто записывают в виде  $\eta = \theta'_1 + \theta_2 (\xi - \bar{\xi})$ . Оценки  $\hat{\theta}'_1$  и  $\hat{\theta}'_2$  статистически независимы, что несправедливо для оценок  $\hat{\theta}_1$  и  $\hat{\theta}_2$ . Оценки  $\hat{\theta}'_1$  и  $\hat{\theta}'_2$  параметров можно получить, не решая совместных систем связанных уравнений, как это приходится делать при другой форме записи модели.

Для нормально распределенных статистически независимых погрешностей экспериментальных точек  $(x_i, y_i)$  минимизируемый функционал имеет следующий вид:

$$F = \sum_{i=1}^n \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right].$$

Из условий

$$\left. \frac{\partial F}{\partial \xi_i} \right|_{\xi_i = \hat{\xi}_i} = 0, \quad i = 1, \dots, n,$$

получим

$$\hat{\xi}_i = \frac{[\sigma^2(y_i)x_i + \theta_2\sigma^2(x_i)(y_i - \theta_1)]}{[\sigma^2(y_i) + \theta_2^2\sigma^2(x_i)]}.$$

Тогда

$$F = \sum_{i=1}^n \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma^2(y_i) + \theta_2^2\sigma^2(x_i)} = \sum_{i=1}^n \omega_i (y_i - \theta_1 - \theta_2 x_i)^2,$$

где  $\omega^{-1} = \sigma^2(y_i) + \theta_2^2\sigma^2(x_i)$ . Согласно [20] введем обозначения:

$$\bar{x} = \frac{\sum_i \omega_i x_i}{\sum_i \omega_i}; \quad x'_i = x_i - \bar{x}; \quad \bar{y} = \frac{\sum_i \omega_i y_i}{\sum_i \omega_i}; \quad y'_i = y_i - \bar{y}.$$

Тогда

$$F = \sum_i \omega_i (\theta_2 x'_i - y'_i)^2; \quad \bar{y} = \theta_1 + \theta_2 \bar{x}.$$

Условие  $\partial F / \partial \theta_2 = 0$  дает

$$\sum_i \left[ \omega_i x'_i (\theta_2 x'_i - y'_i) - \omega_i^2 \theta_2 \sigma(x_i)^2 (\theta_2 x'_i - y'_i)^2 \right] = 0.$$

Это уравнение может быть решено итерациями. Но нетрудно убедиться, что не всякое нулевое приближение приведет к решению задачи, т.е. приходим к такому результату, какой был получен при непосредственном отыскании точки минимума функции  $F$ .

Поскольку переменная  $\theta_2$  входит в выражение для  $\omega_i$ , то равенство  $\frac{\partial F}{\partial \theta_2} = 0$  может быть представлено приближенно в виде полинома любой степени относительно  $\theta_2$  и в [20] для нахождения оценки  $\hat{\theta}_2$  получено уравнение

Для нормально распределенных статистически независимых погрешностей экспериментальных точек  $(x_i, y_i)$  минимизируемый функционал имеет следующий вид:

$$F = \sum_{i=1}^n \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right].$$

Из условий

$$\left. \frac{\partial F}{\partial \xi_i} \right|_{\xi_i = \hat{\xi}_i} = 0, \quad i = 1, \dots, n,$$

получим

$$\hat{\xi}_i = \frac{[\sigma^2(y_i)x_i + \theta_2\sigma^2(x_i)(y_i - \theta_1)]}{[\sigma^2(y_i) + \theta_2^2\sigma^2(x_i)]}.$$

Тогда

$$F = \sum_{i=1}^n \frac{(y_i - \theta_1 - \theta_2 x_i)^2}{\sigma^2(y_i) + \theta_2^2\sigma^2(x_i)} = \sum_{i=1}^n \omega_i (y_i - \theta_1 - \theta_2 x_i)^2,$$

где  $\omega^{-1} = \sigma^2(y_i) + \theta_2^2\sigma^2(x_i)$ . Согласно [20] введем обозначения:

$$\bar{x} = \frac{\sum_i \omega_i x_i}{\sum_i \omega_i}; \quad x'_i = x_i - \bar{x}; \quad \bar{y} = \frac{\sum_i \omega_i y_i}{\sum_i \omega_i}; \quad y'_i = y_i - \bar{y}.$$

Тогда

$$F = \sum_i \omega_i (\theta_2 x'_i - y'_i)^2; \quad \bar{y} = \theta_1 + \theta_2 \bar{x}.$$

Условие  $\partial F / \partial \theta_2 = 0$  дает

$$\sum_i \left[ \omega_i x'_i (\theta_2 x'_i - y'_i) - \omega_i^2 \theta_2 \sigma(x_i)^2 (\theta_2 x'_i - y'_i)^2 \right] = 0.$$

Это уравнение может быть решено итерациями. Но нетрудно убедиться, что не всякое нулевое приближение приведет к решению задачи, т.е. приходим к такому результату, какой был получен при непосредственном отыскании точки минимума функции  $F$ .

Поскольку переменная  $\theta_2$  входит в выражение для  $\omega_i$ , то равенство  $\frac{\partial F}{\partial \theta_2} = 0$  может быть представлено приближенно в виде полинома любой степени относительно  $\theta_2$  и в [20] для нахождения оценки  $\hat{\theta}_2$  получено уравнение

Пусть выбран уровень значимости  $\alpha$ . Тогда  $t = (\hat{\theta}_1 - \theta_1) / \sqrt{D(\hat{\theta}_1)}$  и  $100(1 - \alpha)\%$  -ный доверительный интервал для  $\theta_1$

$$\hat{\theta}_1 - t_{1-\alpha/2} \sqrt{D(\hat{\theta}_1)} \leq \theta_1 \leq \hat{\theta}_1 + t_{1-\alpha/2} \sqrt{D(\hat{\theta}_1)}.$$

Аналогично доверительный интервал для  $\theta_2$

$$\hat{\theta}_2 - t_{1-\alpha/2} \sqrt{D(\hat{\theta}_2)} \leq \theta_2 \leq \hat{\theta}_2 + t_{1-\alpha/2} \sqrt{D(\hat{\theta}_2)}.$$

Отсюда следует, что интервальные границы в случае прямой линии для  $M(\eta / \xi_0)$  будут

$$\hat{\eta} - t_{1-\alpha/2} \sigma(\hat{\eta}) \leq M(\eta / \xi_0) \leq \hat{\eta} + t_{1-\alpha/2} \sigma(\hat{\eta}).$$

где  $\sigma^2(\hat{\eta}) = (1, \xi_0) D(\hat{\theta}) (1, \xi_0)^T$ .

Огибающая семейства всех возможных прямых – кривая второго порядка, имеющая вид

$$\sum_{i=1}^n (y_i - \theta_1 - \theta_2 x_i)^2 \leq C_v (1 + \theta_2^2).$$

Принимая за начало отсчета наблюдаемые средние значения  $x$  и  $y$ , получим

$$\theta_2^2 (S_x^2 - C_v) - 2\theta_2 S_{xy} + \theta_1^2 \leq C_v - S_y^2,$$

где  $S_x^2$ ,  $S_y^2$  и  $S_{xy}$  – соответственно выборочные дисперсии  $x$ ,  $y$  и выборочная ковариация величин  $x$  и  $y$ .

Данное условие можно считать ограничением, которому удовлетворяет истинная прямая. Путем дифференцирования найдем огибающую семейства всех возможных прямых. Уравнение огибающей

$$\frac{(y - \hat{\theta}_2 x)^2}{C_v - b_1} - \frac{(\hat{\theta}_2 y + x)^2}{b_2 - C_v} = 1 + \hat{\theta}_2^2,$$

где  $b_1 = S_x^2 - S_{xy} / \hat{\theta}_2$ ;  $b_2 = S_x^2 + \hat{\theta}_2 S_{xy}$ ;  $\hat{\theta}_2 > 0$ ;  $b_2 > b_1$ .

Программа для получения оценок в конфлюэнтном анализе свободных параметров прямой линии и функций, сводящихся к прямой, приведена в [4, 8].

### 4.3. О единственности оценок параметров.

#### Состоятельность оценок и алгоритм их получения

Самый простой функционал (4.2) нелинеен по  $\theta$  даже в случае оценки параметров  $\theta$  прямой линии при одновременном учете погрешно-

стей и в значениях функции, и в значениях аргумента. Поэтому встает вопрос о единственности оценок искомых параметров, который тесно связан с выбором нулевого приближения оценок. Естественно взять в качестве нулевого приближения решение регрессионной задачи (например, с помощью МНК).

Из условия (4.2) следует, что для одной и той же задачи минимальное значение функционала в МНК больше минимального значения функционала (4.2), поскольку условие (3.6) – условие ортогональности вектора касательной к функции  $\eta = f(\xi, \theta)$  в точке, принимаемой за истинную, и вектора, проведенного через исходную точку  $(x_{i1}, \dots, x_{im}, y_i)$  и точку, принимаемую за истинную. Таким образом, минимизация функционала (3.5) при условии (3.6) приводит к минимизации суммы квадратов «наикратчайших расстояний» от точки  $(x_{i1}, \dots, x_{im}, y_i)$  до кривой, тогда как в МНК минимизируется сумма квадратов отклонений при фиксированных значениях  $x_{ij}$ .

Для конкретных видов функции  $\eta = f(\xi, \theta)$  в работах [4, 6, 7] получены условия существования единственного решения. Покажем существование единственного решения для линейных функций  $\eta = \sum_{j=1}^m \theta_j x_j + \theta_0$  в предположении, что 1) погрешности  $\epsilon$  и  $\delta$  статистически независимы и подчиняются нормальному закону распределения; 2) за нулевое приближение взято решение регрессионной задачи [5, 6]. Эти ограничения обусловлены тем, что с помощью преобразований координат функция Гаусса, описывающая нормальный закон распределения, сводится к уравнению прямой линии и оценки параметров этой прямой будут однозначно связаны с оценками параметров функции Гаусса.

**Теорема 1.** *Если погрешности  $\epsilon$  и  $\delta$  измеренных случайных величин  $X$  и  $Y$  статистически независимы и распределены по нормальному закону с известными числовыми характеристиками ФПВ и если в качестве нулевого приближения принято решение соответствующей регрессионной задачи, то достаточным условием единственности оценок параметров линейной функции  $\eta = \theta_0 + \sum_{j=1}^m \theta_j \xi_j$*

*является выполнение неравенства  $\bar{x}^2 > (1 - m(n)) \hat{\sigma}^2(y)$ . Здесь  $\hat{\sigma}^2(y)$  – оценка дисперсии значений  $y_i$  при равноточных измерениях.*

стей и в значениях функции, и в значениях аргумента. Поэтому встает вопрос о единственности оценок искомых параметров, который тесно связан с выбором нулевого приближения оценок. Естественно взять в качестве нулевого приближения решение регрессионной задачи (например, с помощью МНК).

Из условия (4.2) следует, что для одной и той же задачи минимальное значение функционала в МНК больше минимального значения функционала (4.2), поскольку условие (3.6) – условие ортогональности вектора касательной к функции  $\eta = f(\xi, \theta)$  в точке, принимаемой за истинную, и вектора, проведенного через исходную точку  $(x_{i1}, \dots, x_{im}, y_i)$  и точку, принимаемую за истинную. Таким образом, минимизация функционала (3.5) при условии (3.6) приводит к минимизации суммы квадратов «наикратчайших расстояний» от точки  $(x_{i1}, \dots, x_{im}, y_i)$  до кривой, тогда как в МНК минимизируется сумма квадратов отклонений при фиксированных значениях  $x_{ij}$ .

Для конкретных видов функции  $\eta = f(\xi, \theta)$  в работах [4, 6, 7] получены условия существования единственного решения. Покажем существование единственного решения для линейных функций  $\eta = \sum_{j=1}^m \theta_j x_j + \theta_0$  в предположении, что 1) погрешности  $\epsilon$  и  $\delta$  статистически независимы и подчиняются нормальному закону распределения; 2) за нулевое приближение взято решение регрессионной задачи [5, 6]. Эти ограничения обусловлены тем, что с помощью преобразований координат функция Гаусса, описывающая нормальный закон распределения, сводится к уравнению прямой линии и оценки параметров этой прямой будут однозначно связаны с оценками параметров функции Гаусса.

**Теорема 1.** *Если погрешности  $\epsilon$  и  $\delta$  измеренных случайных величин  $X$  и  $Y$  статистически независимы и распределены по нормальному закону с известными числовыми характеристиками ФПВ и если в качестве нулевого приближения принято решение соответствующей регрессионной задачи, то достаточным условием единственности оценок параметров линейной функции  $\eta = \theta_0 + \sum_{j=1}^m \theta_j \xi_j$*

*является выполнение неравенства  $\bar{x}^2 > (1 - m(n)) \hat{\sigma}^2(y)$ . Здесь  $\hat{\sigma}^2(y)$  – оценка дисперсии значений  $y_i$  при равноточных измерениях.*

стей и в значениях функции, и в значениях аргумента. Поэтому встает вопрос о единственности оценок искомых параметров, который тесно связан с выбором нулевого приближения оценок. Естественно взять в качестве нулевого приближения решение регрессионной задачи (например, с помощью МНК).

Из условия (4.2) следует, что для одной и той же задачи минимальное значение функционала в МНК больше минимального значения функционала (4.2), поскольку условие (3.6) – условие ортогональности вектора касательной к функции  $\eta = f(\xi, \theta)$  в точке, принимаемой за истинную, и вектора, проведенного через исходную точку  $(x_{i1}, \dots, x_{im}, y_i)$  и точку, принимаемую за истинную. Таким образом, минимизация функционала (3.5) при условии (3.6) приводит к минимизации суммы квадратов «наикратчайших расстояний» от точки  $(x_{i1}, \dots, x_{im}, y_i)$  до кривой, тогда как в МНК минимизируется сумма квадратов отклонений при фиксированных значениях  $x_{ij}$ .

Для конкретных видов функции  $\eta = f(\xi, \theta)$  в работах [4, 6, 7] получены условия существования единственного решения. Покажем существование единственного решения для линейных функций  $\eta = \sum_{j=1}^m \theta_j x_j + \theta_0$  в предположении, что 1) погрешности  $\epsilon$  и  $\delta$  статистически независимы и подчиняются нормальному закону распределения; 2) за нулевое приближение взято решение регрессионной задачи [5, 6]. Эти ограничения обусловлены тем, что с помощью преобразований координат функция Гаусса, описывающая нормальный закон распределения, сводится к уравнению прямой линии и оценки параметров этой прямой будут однозначно связаны с оценками параметров функции Гаусса.

**Теорема 1.** *Если погрешности  $\epsilon$  и  $\delta$  измеренных случайных величин  $X$  и  $Y$  статистически независимы и распределены по нормальному закону с известными числовыми характеристиками ФПВ и если в качестве нулевого приближения принято решение соответствующей регрессионной задачи, то достаточным условием единственности оценок параметров линейной функции  $\eta = \theta_0 + \sum_{j=1}^m \theta_j \xi_j$*

*является выполнение неравенства  $\bar{x}^2 > (1 - m(n)) \hat{\sigma}^2(y)$ . Здесь  $\hat{\sigma}^2(y)$  – оценка дисперсии значений  $y_i$  при равноточных измерениях.*



стей и в значениях функции, и в значениях аргумента. Поэтому встает вопрос о единственности оценок искомых параметров, который тесно связан с выбором нулевого приближения оценок. Естественно взять в качестве нулевого приближения решение регрессионной задачи (например, с помощью МНК).

Из условия (4.2) следует, что для одной и той же задачи минимальное значение функционала в МНК больше минимального значения функционала (4.2), поскольку условие (3.6) – условие ортогональности вектора касательной к функции  $\eta = f(\xi, \theta)$  в точке, принимаемой за истинную, и вектора, проведенного через исходную точку  $(x_{i1}, \dots, x_{im}, y_i)$  и точку, принимаемую за истинную. Таким образом, минимизация функционала (3.5) при условии (3.6) приводит к минимизации суммы квадратов «наикратчайших расстояний» от точки  $(x_{i1}, \dots, x_{im}, y_i)$  до кривой, тогда как в МНК минимизируется сумма квадратов отклонений при фиксированных значениях  $x_{ij}$ .

Для конкретных видов функции  $\eta = f(\xi, \theta)$  в работах [4, 6, 7] получены условия существования единственного решения. Покажем существование единственного решения для линейных функций  $\eta = \sum_{j=1}^m \theta_j x_j + \theta_0$  в предположении, что 1) погрешности  $\epsilon$  и  $\delta$  статистически независимы и подчиняются нормальному закону распределения; 2) за нулевое приближение взято решение регрессионной задачи [5, 6]. Эти ограничения обусловлены тем, что с помощью преобразований координат функция Гаусса, описывающая нормальный закон распределения, сводится к уравнению прямой линии и оценки параметров этой прямой будут однозначно связаны с оценками параметров функции Гаусса.

**Теорема 1.** *Если погрешности  $\epsilon$  и  $\delta$  измеренных случайных величин  $X$  и  $Y$  статистически независимы и распределены по нормальному закону с известными числовыми характеристиками ФПВ и если в качестве нулевого приближения принято решение соответствующей регрессионной задачи, то достаточным условием единственности оценок параметров линейной функции  $\eta = \theta_0 + \sum_{j=1}^m \theta_j \xi_j$*

*является выполнение неравенства  $\bar{x}^2 > (1 - m(n)) \hat{\sigma}^2(y)$ . Здесь  $\hat{\sigma}^2(y)$  – оценка дисперсии значений  $y_i$  при равноточных измерениях.*

стей и в значениях функции, и в значениях аргумента. Поэтому встает вопрос о единственности оценок искомых параметров, который тесно связан с выбором нулевого приближения оценок. Естественно взять в качестве нулевого приближения решение регрессионной задачи (например, с помощью МНК).

Из условия (4.2) следует, что для одной и той же задачи минимальное значение функционала в МНК больше минимального значения функционала (4.2), поскольку условие (3.6) – условие ортогональности вектора касательной к функции  $\eta = f(\xi, \theta)$  в точке, принимаемой за истинную, и вектора, проведенного через исходную точку  $(x_{i1}, \dots, x_{im}, y_i)$  и точку, принимаемую за истинную. Таким образом, минимизация функционала (3.5) при условии (3.6) приводит к минимизации суммы квадратов «наикратчайших расстояний» от точки  $(x_{i1}, \dots, x_{im}, y_i)$  до кривой, тогда как в МНК минимизируется сумма квадратов отклонений при фиксированных значениях  $x_{ij}$ .

Для конкретных видов функции  $\eta = f(\xi, \theta)$  в работах [4, 6, 7] получены условия существования единственного решения. Покажем существование единственного решения для линейных функций  $\eta = \sum_{j=1}^m \theta_j x_j + \theta_0$  в предположении, что 1) погрешности  $\epsilon$  и  $\delta$  статистически независимы и подчиняются нормальному закону распределения; 2) за нулевое приближение взято решение регрессионной задачи [5, 6]. Эти ограничения обусловлены тем, что с помощью преобразований координат функция Гаусса, описывающая нормальный закон распределения, сводится к уравнению прямой линии и оценки параметров этой прямой будут однозначно связаны с оценками параметров функции Гаусса.

**Теорема 1.** *Если погрешности  $\epsilon$  и  $\delta$  измеренных случайных величин  $X$  и  $Y$  статистически независимы и распределены по нормальному закону с известными числовыми характеристиками ФПВ и если в качестве нулевого приближения принято решение соответствующей регрессионной задачи, то достаточным условием единственности оценок параметров линейной функции  $\eta = \theta_0 + \sum_{j=1}^m \theta_j \xi_j$*

*является выполнение неравенства  $\bar{x}^2 > (1 - m(n)) \hat{\sigma}^2(y)$ . Здесь  $\hat{\sigma}^2(y)$  – оценка дисперсии значений  $y_i$  при равноточных измерениях.*

стей и в значениях функции, и в значениях аргумента. Поэтому встает вопрос о единственности оценок искомых параметров, который тесно связан с выбором нулевого приближения оценок. Естественно взять в качестве нулевого приближения решение регрессионной задачи (например, с помощью МНК).

Из условия (4.2) следует, что для одной и той же задачи минимальное значение функционала в МНК больше минимального значения функционала (4.2), поскольку условие (3.6) – условие ортогональности вектора касательной к функции  $\eta = f(\xi, \theta)$  в точке, принимаемой за истинную, и вектора, проведенного через исходную точку  $(x_{i1}, \dots, x_{im}, y_i)$  и точку, принимаемую за истинную. Таким образом, минимизация функционала (3.5) при условии (3.6) приводит к минимизации суммы квадратов «наикратчайших расстояний» от точки  $(x_{i1}, \dots, x_{im}, y_i)$  до кривой, тогда как в МНК минимизируется сумма квадратов отклонений при фиксированных значениях  $x_{ij}$ .

Для конкретных видов функции  $\eta = f(\xi, \theta)$  в работах [4, 6, 7] получены условия существования единственного решения. Покажем существование единственного решения для линейных функций  $\eta = \sum_{j=1}^m \theta_j x_j + \theta_0$  в предположении, что 1) погрешности  $\epsilon$  и  $\delta$  статистически независимы и подчиняются нормальному закону распределения; 2) за нулевое приближение взято решение регрессионной задачи [5, 6]. Эти ограничения обусловлены тем, что с помощью преобразований координат функция Гаусса, описывающая нормальный закон распределения, сводится к уравнению прямой линии и оценки параметров этой прямой будут однозначно связаны с оценками параметров функции Гаусса.

**Теорема 1.** *Если погрешности  $\epsilon$  и  $\delta$  измеренных случайных величин  $X$  и  $Y$  статистически независимы и распределены по нормальному закону с известными числовыми характеристиками ФПВ и если в качестве нулевого приближения принято решение соответствующей регрессионной задачи, то достаточным условием единственности оценок параметров линейной функции  $\eta = \theta_0 + \sum_{j=1}^m \theta_j \xi_j$*

*является выполнение неравенства  $\bar{x}^2 > (1 - m(n)) \hat{\sigma}^2(y)$ . Здесь  $\hat{\sigma}^2(y)$  – оценка дисперсии значений  $y_i$  при равноточных измерениях.*

Шаги 2 и 3 выполняют поочередно до тех пор, пока не будет выполнено хотя бы одно из трех условий:

1) на очередном шаге значение функционала (4.3) меньше заданного числа  $\epsilon_1$ ;

2) на соседних итерациях значения функционала  $F$  и оценок параметров  $\theta$  различаются несущественно, т.е.

$$\left| \frac{F_v - F_{v+1}}{F_v} \right| \leq \delta_1 \text{ и } \max_l \left| \frac{\theta_l^{(v)} - \theta_l^{(v+1)}}{\theta_l^{(v)}} \right| \leq \delta_2, \quad l = 1, \dots, k,$$

где  $\delta_1$  и  $\delta_2$  – заданные числа;

3) исчерпан заданный лимит итераций.

*Шаг 4.* После нахождения точечных оценок  $\hat{\theta}$  определяют дисперсии оценок.

Из приведенного алгоритма следует, что на первой итерации получается регрессионное решение (иными словами, решение любым из применявшихся ранее традиционных методов). На последующих шагах это решение уточняют. Такой переход от традиционных методов к предлагаемому имеет принципиальное значение в практических приложениях.

#### 4.4. Оценка параметров многомерной линейной модели

Линейные модели широко применяют в статистических задачах принятия решений, например, при нахождении разделяющих поверхностей в задачах распознавания образов. Рассмотрим применение описанного подхода к задаче определения оценок параметров линейной модели

$$\eta = f(\xi, \theta) = \sum_{i=1}^n \theta_i \xi_i$$

и к решению систем линейных алгебраических уравнений с погрешностями в матрице системы и в правой части уравнений, если ошибки измерений суть независимые нормально распределенные случайные величины с нулевыми средними значениями и известными дисперсиями  $\sigma^2(x_{ij})$  и  $\sigma^2(y_j)$ .

В этом случае функционал (3.5) будет иметь вид

$$F = \frac{1}{2} \sum_{j=1}^m \left\{ \sum_{i=1}^n \frac{(x_{ij} - \xi_{ij})^2}{\sigma^2(x_{ij})} + \frac{(y_j - \sum_i \theta_i \xi_{ij})^2}{\sigma^2(y_j)} \right\}, \quad (4.4)$$

а ограничение  $\bar{\xi}_i \in D_i$  можно записать следующим образом:

$$|x_{ij} - \bar{\xi}_{ij}| \leq 3\sigma(x_{ij}). \quad (4.5)$$

Здесь для упрощения вида функционала (4.4) ошибки измерений считают статистически независимыми.

Задача минимизации по  $\theta$  функционала (4.4) при фиксированных значениях  $\xi$  есть обычная задача регрессионного анализа. Затем следует непосредственная минимизация функционала (4.4) по  $\theta$ . Исследования показали, что для обеспечения минимального времени счета и большей точности лучше всего на первом шаге решать систему линейных уравнений при  $\xi_{ij} = x_{ij}$  методом Гаусса с выбором максимального элемента, а в дальнейшем непосредственно минимизировать по  $\theta$  функционал (4.4) методом сопряженных градиентов. Учитывая вид функционала (4.4), задачу минимизации по  $\theta$  легко свести к задаче поиска минимума квадратичной формы

$$G(\theta) = \frac{1}{2} \theta^T A \theta + a^T \theta,$$

которая решалась методом сопряженных градиентов. Здесь  $\theta$  – вектор искомых оценок параметров;  $A$  – матрица квадратичной формы, элементы которой вычисляются по формуле

$$A_{rp} = \sum_{j=1}^m \frac{1}{\sigma^2(y_j)} x_{rj} x_{pj}; \quad r = 1, 2, \dots, n; \quad p = 1, 2, \dots, n,$$

$x_{rj}$  – значение  $r$ -й переменной, полученное в  $j$ -м измерении;  $a^T$  – вектор, компоненты которого равны

$$a_r^T = -\sum_{j=1}^m \frac{y_j}{\sigma^2(y_j)} x_{rj}; \quad r = 1, 2, \dots, n.$$

Пересчет точных значений  $\xi$  на основании условий (3.6) сводится к решению  $m$  несвязанных между собой систем из  $n$  линейных уравнений вида

$$\sum_{r=1}^n \frac{\theta_r \theta_p}{\sigma^2(y_j)} \xi_{rj} + \frac{\xi_{pj}}{\sigma^2(x_{pj})} = \frac{x_{pj}}{\sigma^2(x_{pj})} + \frac{\theta_p y_j}{\sigma^2(y_j)}; \quad p = 1, 2, \dots, n; \quad j = 1, 2, \dots, m.$$

Полученные новые значения  $\hat{\xi}$  должны удовлетворять условию (4.5). В противном случае те  $\xi_{ij}$ , которые выходят за указанные границы, заменяют на значение ближайшей граничной точки. В связи с этим иногда можно ожидать увеличение функционала (4.3) на

новых точных значениях переменных  $\xi$  по сравнению с предыдущим шагом итерационного процесса. Это приводит к снижению скорости сходимости процесса и даже к возникновению колебаний. Для устранения таких нежелательных последствий после пересчета  $\xi$  те наборы  $\xi_j$ , на которых произошло увеличение соответствующих слагаемых функционала

$$F_j = \left\{ \sum_{i=1}^n \frac{(x_{ij} - \xi_{ij})^2}{\sigma^2(x_{ij})} + \frac{(y_j - \sum_i \theta_i \xi_{ij})^2}{\sigma^2(y_j)} \right\}$$

по сравнению с предыдущей итерацией, заменяют значениями  $\xi$  с предыдущего шага.

Применение описанного приема позволило в большинстве случаев резко увеличить скорость сходимости процесса (больше чем в 3 раза). Когда же первоначальный процесс не сходился (возникали колебания), решение достигалось за 12–17 итераций. Элементы дисперсионной матрицы ошибок (матрицы рассеяния) для оценок искомым параметров подсчитываются как элементы матрицы, обратной матрице  $N$  с элементами

$$N_{ij} = - \frac{\partial^2 F}{\partial \theta_i \partial \theta_j} \Big|_{\theta = \hat{\theta}},$$

где  $\hat{\theta}$  – полученные оценки параметров  $\theta$ .

Описанную выше задачу можно интерпретировать как задачу решения переопределенной системы из  $m$  линейных уравнений с  $n$  переменными ( $m > n$ ):  $\eta = \xi\theta$ , когда вместо значений  $h$  и  $x$  имеем соответственно  $Y = (y_1, \dots, y_m)^T$  – столбец значений функции;  $X$  – матрицу размером  $m \times n$ , элементы которой  $x_{ij}$  суть значения  $i$ -й переменной в  $j$ -м измерении;  $\theta = (\theta_1, \dots, \theta_n)^T$  – столбец неизвестных (искомых) параметров.

#### 4.5. Оценка параметров полиномиальной зависимости

Полиномом можно приближенно описать любую функцию или плотность вероятности. Рассмотрим задачу оценки параметров полинома. Требуется определить оценки  $\theta$  коэффициентов полинома степени  $m$

$$\eta = \sum_{i=0}^m \theta_i \xi^i$$

по совокупности значений аргумента и функции, содержащих погрешности измерений как в значениях функции, так и в значениях аргумента. Пусть исходные данные имеют вид

$$y_i = \eta_i + \varepsilon_i; \quad x_i = \xi_i + \delta_i; \quad i = 1, 2, \dots, n.$$

Предполагается, что ошибки измерений  $\vec{\varepsilon}$  и  $\vec{\delta}$  – нормально распределенные случайные величины с нулевыми средними значениями и с известными дисперсиями (соответственно  $\sigma^2(x_i)$  и  $|x_i - \xi_i| \leq 3\sigma(x_i)$ ), и известными коэффициентами корреляции  $\rho_i$  между измеренными значениями  $x_i$  и  $\theta$ . Тогда искомые оценки коэффициентов  $\hat{\theta}$  получают путем минимизации функционала

$$\eta = f(\xi, \theta) = \sum_{i=1}^k \theta_i \prod_{j=1}^m \xi_j^{\alpha_{ij}} \quad (4.6)$$

при условиях

$$\left. \frac{\partial F}{\partial \xi_i} \right|_{\xi_i = \hat{\xi}_i} = 0; \quad \sigma^2(x_{ij}). \quad (4.7)$$

Пересчет значений  $\hat{\xi}_i$  согласно условию (4.7) сводится к решению системы из  $m$  нелинейных уравнений вида

$$\begin{aligned} \frac{x_i - \xi_i}{\sigma^2(x_i)} - \rho_i \frac{y_i - \sum_{l=0}^m \theta_l \xi_i^l + (x_i - \xi_i) \sum_{l=0}^m l \theta_l \xi_i^{l-1} + (x_i - \xi_i)}{\sigma(x_i) \sigma(y_i)} + \\ + \frac{y_i - \sum_{l=0}^m \theta_l \xi_i^l + (x_i - \xi_i) \sum_{l=1}^m l \theta_l \xi_i^{l-1}}{\sigma^2(y_i)} = 0; \quad i = 1, \dots, n, \end{aligned} \quad (4.8)$$

на отрезке  $G(\theta) = \frac{1}{2} \theta A \theta + a^T \theta$ .

В процессе решений системы уравнений (4.8) может оказаться, что:

1) корень уравнения (4.8) принадлежит отрезку  $|x_i - \xi_i| \leq 3\sigma(x_i)$ . Значение корня находится методом хорд;

2) корень уравнения (4.8) не принадлежит отрезку  $|x_i - \xi_i| \leq 3\sigma(x_i)$ . Тогда с помощью алгоритма поиска глобального минимума функции находится наименьшее значение функции

$$\varphi(\xi_i) = \frac{x_i - \xi_i}{\sigma^2(x_i)} + \rho_i \frac{y_i - \sum_{l=0}^m \theta_l \xi_i^l + (x_i - \xi_i) \sum_{l=1}^m l \theta_l \xi_i^{l-1}}{\sigma(x_i)\sigma(y_i)} + \frac{y_i - \sum_{l=0}^m \theta_l \xi_i^l}{\sigma^2(y_i)} \sum_{l=1}^m l \theta_l \xi_i^{l-1},$$

принадлежащее отрезку  $|x_i - \xi_i| \leq 3\sigma(x_i)$ .

Поставленная задача минимизации функционала (4.6) решается по итерационной схеме, изложенной в п. 4.2. Дисперсии полученных оценок находились с помощью матрицы вторых производных функционала (4.6). Процедура нахождения оценок и их дисперсий аналогична ранее описанной в п. 2.5.

#### 4.6. Оценка значений параметров в сигноме (многомерном полиноме)

Рассмотрим задачу оценки параметров  $\theta$  сигнома (самый общий вид многомерного полинома)

$$\eta = f(\xi, \theta) = \sum_{i=1}^k \theta_i \prod_{j=1}^m \xi_j^{\alpha_j}$$

при условии, что все ошибки измерений – независимые нормально распределенные случайные величины с нулевыми средними и известными дисперсиями  $\sigma^2(x_{ij})$  и  $\sigma^2(y_i)$ . Функционал, позволяющий определить оценки параметров  $\theta$  сигнома, будет иметь вид

$$F = \frac{1}{2} \sum_{i=1}^n \left\{ \sum_{j=1}^m \frac{(x_{ij} - \xi_{ij})^2}{\sigma^2(x_{ij})} + \frac{[y_i - f(\xi_i; \theta)]^2}{\sigma^2(y_i)} \right\}. \quad (4.9)$$

Ограничение  $\xi_i \in D_i$  можно записать следующим образом:

$$|x_{ij} - \xi_{ij}| \leq 3\sigma(x_{ij}).$$

Чтобы упростить вид функционала (4.9), задачу минимизации по  $\theta$  сводим к задаче поиска минимума квадратичной формы:

$$G(\theta) = \frac{1}{2} \theta^T A \theta + a^T \theta.$$

Точку минимума  $G(\theta)$  определяли методом сопряженных градиентов. Элементы матрицы квадратичной формы  $A$  имеют вид

$$A_{rp} = \sum_{i=1}^n \frac{1}{\sigma^2(y_i)} \prod_{j=1}^m x_{ij}^{\alpha_{jr} + \alpha_{jp}}, \quad r = 1, 2, \dots, k; \quad p = 1, 2, \dots, k,$$



где  $x_{ij}$  – значение  $i$ -й переменной в  $j$ -м измерении;  $a_r^T$  – вектор, компоненты которого равны

$$a_r^T = -\sum_{l=1}^m \frac{y_l}{\sigma^2(y_l)} \prod_{j=1}^m x_{lj}^{\alpha_{jl}}, \quad r = 1, 2, \dots, k.$$

Пересчет точных значений  $\xi$  с учетом условий (1.4) сводится к решению  $n$  несвязанных между собой систем из  $m$  нелинейных уравнений вида

$$\frac{x_{ij} - \xi_{ij}}{\sigma^2(x_{ij})} + \frac{y_l - f(\vec{\xi}; \vec{\theta})}{\sigma^2(y_l)} \frac{\partial f}{\partial \xi_{ij}} = 0, \quad (4.10)$$

где

$$\frac{\partial f}{\partial \xi_{ij}} = \sum_{l=1}^k \theta_l^{\alpha_{jl}} x_{l1}^{\alpha_{l1}} \dots x_{lj}^{\alpha_{lj}} \dots x_{lm}^{\alpha_{lm}},$$

$\vec{\xi}_i = (\xi_{i1}, \dots, \xi_{im})$  – искомые значения переменных в  $i$ -м измерении.

Систему уравнений (4.10) решали методом итераций с использованием линеаризации исходных уравнений на каждом шаге. Такое упрощение вычислений возможно потому, что все расчеты ведутся в небольшой области пространства переменных. Таким образом, на каждой итерации в процессе решения системы (4.10) решают системы линейных уравнений:

$$\sum_{r=1}^m \frac{\partial \varphi_j(\xi_i)}{\partial \xi_{ir}} \Big|_{\xi_i = \xi_i^0} (\xi_{ir} - \xi_{ir}^0) = -\varphi_j(\xi_i^0), \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m,$$

где  $\varphi_j(\xi_i^0)$  – значение левой части (4.10) в точке  $\xi_i^0$  (точке разложения);

$$\frac{\partial \varphi_j}{\partial \xi_{ir}} = \frac{y_l - f(\xi_i; \theta)}{\sigma^2(y_l)} f''_{\xi_{ij}\xi_{ir}} - \frac{f'_{\xi_{ij}} f'_{\xi_{ir}}}{\sigma^2(y_l)}, \quad \text{если } r \neq j;$$

$$\frac{\partial \varphi_j}{\partial \xi_{ir}} = \frac{y_l - f(\xi_i; \theta)}{\sigma^2(y_l)} f''_{\xi_{ij}^2} - \frac{f_{\xi_{ij}}'^2}{\sigma^2(y_l)} - \frac{1}{\sigma^2(x_{ij})}, \quad \text{если } r = j.$$

Здесь

$$f''_{\xi_{ij}\xi_{ir}} = \sum_{l=1}^k \theta_l \alpha_{jl} \alpha_{rl} x_{l1}^{\alpha_{l1}} \dots x_{lj}^{\alpha_{lj}-1} \dots x_{lr}^{\alpha_{lr}-1} \dots x_{lm}^{\alpha_{lm}};$$

$$f''_{\xi_{ij}^2} = \sum_{l=1}^k \theta_l \alpha_{jl} (\alpha_{jl} - 1) x_{l1}^{\alpha_{l1}} \dots x_{lj}^{\alpha_{lj}-2} \dots x_{lr}^{\alpha_{lr}-1} \dots x_{lm}^{\alpha_{lm}}.$$

где  $x_{ij}$  – значение  $i$ -й переменной в  $j$ -м измерении;  $a_r^T$  – вектор, компоненты которого равны

$$a_r^T = -\sum_{l=1}^m \frac{y_l}{\sigma^2(y_l)} \prod_{j=1}^m x_{lj}^{\alpha_{jl}}, \quad r = 1, 2, \dots, k.$$

Пересчет точных значений  $\xi$  с учетом условий (1.4) сводится к решению  $n$  несвязанных между собой систем из  $m$  нелинейных уравнений вида

$$\frac{x_{ij} - \xi_{ij}}{\sigma^2(x_{ij})} + \frac{y_l - f(\vec{\xi}; \vec{\theta})}{\sigma^2(y_l)} \frac{\partial f}{\partial \xi_{ij}} = 0, \quad (4.10)$$

где

$$\frac{\partial f}{\partial \xi_{ij}} = \sum_{l=1}^k \theta_l^{\alpha_{jl}} x_{l1}^{\alpha_{l1}} \dots x_{lj}^{\alpha_{lj}} \dots x_{lm}^{\alpha_{lm}},$$

$\vec{\xi}_i = (\xi_{i1}, \dots, \xi_{im})$  – искомые значения переменных в  $i$ -м измерении.

Систему уравнений (4.10) решали методом итераций с использованием линеаризации исходных уравнений на каждом шаге. Такое упрощение вычислений возможно потому, что все расчеты ведутся в небольшой области пространства переменных. Таким образом, на каждой итерации в процессе решения системы (4.10) решают системы линейных уравнений:

$$\sum_{r=1}^m \frac{\partial \varphi_j(\xi_i)}{\partial \xi_{ir}} \Big|_{\xi_i = \xi_i^0} (\xi_{ir} - \xi_{ir}^0) = -\varphi_j(\xi_i^0), \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m,$$

где  $\varphi_j(\xi_i^0)$  – значение левой части (4.10) в точке  $\xi_i^0$  (точке разложения);

$$\frac{\partial \varphi_j}{\partial \xi_{ir}} = \frac{y_l - f(\xi_i; \theta)}{\sigma^2(y_l)} f''_{\xi_{ij}\xi_{ir}} - \frac{f'_{\xi_{ij}} f'_{\xi_{ir}}}{\sigma^2(y_l)}, \quad \text{если } r \neq j;$$

$$\frac{\partial \varphi_j}{\partial \xi_{ir}} = \frac{y_l - f(\xi_i; \theta)}{\sigma^2(y_l)} f''_{\xi_{ij}^2} - \frac{f_{\xi_{ij}}'^2}{\sigma^2(y_l)} - \frac{1}{\sigma^2(x_{ij})}, \quad \text{если } r = j.$$

Здесь

$$f''_{\xi_{ij}\xi_{ir}} = \sum_{l=1}^k \theta_l \alpha_{jl} \alpha_{rl} x_{l1}^{\alpha_{l1}} \dots x_{lj}^{\alpha_{lj}-1} \dots x_{lr}^{\alpha_{lr}-1} \dots x_{lm}^{\alpha_{lm}};$$

$$f''_{\xi_{ij}^2} = \sum_{l=1}^k \theta_l \alpha_{jl} (\alpha_{jl} - 1) x_{l1}^{\alpha_{l1}} \dots x_{lj}^{\alpha_{lj}-2} \dots x_{lr}^{\alpha_{lr}-1} \dots x_{lm}^{\alpha_{lm}}.$$

Результат решения регрессионной задачи:

$$\hat{\theta}_{1p} = 0,895, \quad \hat{\theta}_{2p} = -0,893, \quad \hat{\theta}_{3p} = 1,975.$$

Сумма квадратов невязок  $\sum_{i=1}^6 (y_i - f(x_i, \theta_p))^2$  равна 0,0232.

Результат решения конъюгентной задачи:

$$\theta_1 = 0,984, \quad \theta_2 = -0,979, \quad \theta_3 = 1,999.$$

Сумма квадратов невязок равна 0,0007. Матрица рассеяния оценок параметров  $\theta$  имеет вид

$$\begin{pmatrix} 0,00025 & -0,00022 & -0,00002 \\ -0,00022 & 0,00023 & -0,00004 \\ -0,00002 & -0,00004 & 0,00004 \end{pmatrix}.$$

Полученные результаты позволяют находить решения сложных систем нелинейных уравнений.

#### 4.7. Построение сглаживающих кубических сплайнов

Теория сплайн-функций дает эффективные методы решения задач сглаживания и дифференцирования таблично заданных функций [7]. Наибольшее применение получили кубические сплайны. В литературе рассматривалось использование сплайнов при наличии ошибок только в значениях функции. Вопрос применения сплайнов при погрешностях в значениях аргументов ранее не освещался, несмотря на важность этого случая.

Пусть функция  $y(x)$  представлена значениями  $\hat{y}_i = y(\hat{x}_i) + \delta_i$ ,  $\forall i = 1, \dots, n$ , определенными на сетке  $\Delta_1$ :  $\hat{x}_1 < \hat{x}_2 < \dots < \hat{x}_n$ , где  $\delta_i$  – погрешность  $i$ -го значения функции. При этом значения узлов сетки  $\Delta_1$  также содержат погрешность, т.е.  $\hat{x}_i = \xi_i + \varepsilon_i$ ,  $\forall i = 1, \dots, n$ , где  $\xi_i$  – точное значение  $i$ -го узла;  $\varepsilon_i$  – погрешность значения  $i$ -го узла. Требуется вычислить значения функции  $y(x)$  и (или) ее производной  $y'(x)$  на сетке  $\Delta_2$ :  $z_1 < z_2 < \dots < z_m$ , которая может и не совпадать с исходной сеткой  $\Delta_1$ .

Наличие погрешностей в узлах сетки  $\Delta_1$  и в значениях  $y_i$  не позволяет использовать для решения этой задачи интерполяционную функцию, которая в узлах  $\hat{x}_i$  совпадает с  $\hat{y}_i$ . Нельзя также применять теорию сглаживающих сплайн-функций в чистом виде, так как

она не учитывает погрешности в узлах сетки  $\Delta_1$ . В качестве сглаживающей функции возьмем такую функцию  $g(x)$ , непрерывную вместе с производными до второго порядка, которая дает минимум функционала

$$F = \alpha \int_{\xi_1}^{\xi_n} (g''(x))^2 dx + \frac{1}{2} \sum_{i=1}^n \left\{ p_i^{-1} (\hat{y}_i - g(\xi_i))^2 + q_i^{-1} (\hat{x}_i - \xi_i)^2 \right\},$$

где  $\alpha$  – коэффициент сглаживания;  $p_i$  и  $q_i$  – весовые множители;  $\xi_i$  – неизвестные точные значения узлов сетки  $\Delta_1$ .

Функционал  $F$  имеет такой вид при условии нормального закона распределения погрешностей значений  $\hat{y}$  и  $\hat{x}$ . При других законах распределения вид  $F$  будет иной. Например, для случаев, когда координаты  $x_i$  и  $y_i$  статистически независимы и подчиняются закону распределения Пуассона, функционал  $F$  имеет вид

$$F = \alpha \int_{\xi_1}^{\xi_n} (g''(x))^2 dx + \frac{1}{2} \sum_{i=1}^n \left\{ y_i \ln [g(\xi_i)] + x_i \ln \xi_i - g(\xi_i) - \xi_i \right\}.$$

Следуя идеям конфлюэнтного анализа, оценки неизвестных точных значений узлов сетки  $\Delta_1$  вычисляются исходя из условия

$$\frac{\partial F}{\partial \xi_i} = 0 \quad \forall i = 1, \dots, n.$$

Легко показать, что функцией, дающей минимум функционала  $F$  при последнем условии, является кубический сплайн, т.е. функция, удовлетворяющая условиям:

- 1) на каждом отрезке точной сетки  $[\xi_{i-1}, \xi_i]$   $g(x)$  – полином третьей степени;
- 2) функция  $g(x)$  и ее первые две производные непрерывны на всем отрезке  $[\xi_1, \xi_n]$ ;
- 3) функция  $g(x)$  удовлетворяет определенным граничным условиям.

Кубический сплайн, удовлетворяющий этим условиям, имеет вид

$$g(x) = m_{i-1} \frac{(\xi_i - x)^3}{6h_i} + m_i \frac{(x - \xi_{i-1})^3}{6h_i} + \left( \mu_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{\xi_i - x}{h_i} + \left( \mu_i - \frac{m_i h_i^2}{6} \right) \frac{x - \xi_{i-1}}{h_i},$$

где  $i$  выбирают из условия  $\xi_{i-1} \leq x \leq \xi_i$ ;  $h_i = \xi_i - \xi_{i-1}$ ;  $m_i$  и  $\mu_i$  – параметры сплайна, причем

она не учитывает погрешности в узлах сетки  $\Delta_1$ . В качестве сглаживающей функции возьмем такую функцию  $g(x)$ , непрерывную вместе с производными до второго порядка, которая дает минимум функционала

$$F = \alpha \int_{\xi_1}^{\xi_n} (g''(x))^2 dx + \frac{1}{2} \sum_{i=1}^n \left\{ p_i^{-1} (\hat{y}_i - g(\xi_i))^2 + q_i^{-1} (\hat{x}_i - \xi_i)^2 \right\},$$

где  $\alpha$  – коэффициент сглаживания;  $p_i$  и  $q_i$  – весовые множители;  $\xi_i$  – неизвестные точные значения узлов сетки  $\Delta_1$ .

Функционал  $F$  имеет такой вид при условии нормального закона распределения погрешностей значений  $\hat{y}$  и  $\hat{x}$ . При других законах распределения вид  $F$  будет иной. Например, для случаев, когда координаты  $x_i$  и  $y_i$  статистически независимы и подчиняются закону распределения Пуассона, функционал  $F$  имеет вид

$$F = \alpha \int_{\xi_1}^{\xi_n} (g''(x))^2 dx + \frac{1}{2} \sum_{i=1}^n \left\{ y_i \ln [g(\xi_i)] + x_i \ln \xi_i - g(\xi_i) - \xi_i \right\}.$$

Следуя идеям конфлюэнтного анализа, оценки неизвестных точных значений узлов сетки  $\Delta_1$  вычисляются исходя из условия

$$\frac{\partial F}{\partial \xi_i} = 0 \quad \forall i = 1, \dots, n.$$

Легко показать, что функцией, дающей минимум функционала  $F$  при последнем условии, является кубический сплайн, т.е. функция, удовлетворяющая условиям:

- 1) на каждом отрезке точной сетки  $[\xi_{i-1}, \xi_i]$   $g(x)$  – полином третьей степени;
- 2) функция  $g(x)$  и ее первые две производные непрерывны на всем отрезке  $[\xi_1, \xi_n]$ ;
- 3) функция  $g(x)$  удовлетворяет определенным граничным условиям.

Кубический сплайн, удовлетворяющий этим условиям, имеет вид

$$g(x) = m_{i-1} \frac{(\xi_i - x)^3}{6h_i} + m_i \frac{(x - \xi_{i-1})^3}{6h_i} + \left( \mu_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{\xi_i - x}{h_i} + \left( \mu_i - \frac{m_i h_i^2}{6} \right) \frac{x - \xi_{i-1}}{h_i},$$

где  $i$  выбирают из условия  $\xi_{i-1} \leq x \leq \xi_i$ ;  $h_i = \xi_i - \xi_{i-1}$ ;  $m_i$  и  $\mu_i$  – параметры сплайна, причем

она не учитывает погрешности в узлах сетки  $\Delta_1$ . В качестве сглаживающей функции возьмем такую функцию  $g(x)$ , непрерывную вместе с производными до второго порядка, которая дает минимум функционала

$$F = \alpha \int_{\xi_1}^{\xi_n} (g''(x))^2 dx + \frac{1}{2} \sum_{i=1}^n \left\{ p_i^{-1} (\hat{y}_i - g(\xi_i))^2 + q_i^{-1} (\hat{x}_i - \xi_i)^2 \right\},$$

где  $\alpha$  – коэффициент сглаживания;  $p_i$  и  $q_i$  – весовые множители;  $\xi_i$  – неизвестные точные значения узлов сетки  $\Delta_1$ .

Функционал  $F$  имеет такой вид при условии нормального закона распределения погрешностей значений  $\hat{y}$  и  $\hat{x}$ . При других законах распределения вид  $F$  будет иной. Например, для случаев, когда координаты  $x_i$  и  $y_i$  статистически независимы и подчиняются закону распределения Пуассона, функционал  $F$  имеет вид

$$F = \alpha \int_{\xi_1}^{\xi_n} (g''(x))^2 dx + \frac{1}{2} \sum_{i=1}^n \left\{ y_i \ln [g(\xi_i)] + x_i \ln \xi_i - g(\xi_i) - \xi_i \right\}.$$

Следуя идеям конфлюэнтного анализа, оценки неизвестных точных значений узлов сетки  $\Delta_1$  вычисляются исходя из условия

$$\frac{\partial F}{\partial \xi_i} = 0 \quad \forall i = 1, \dots, n.$$

Легко показать, что функцией, дающей минимум функционала  $F$  при последнем условии, является кубический сплайн, т.е. функция, удовлетворяющая условиям:

- 1) на каждом отрезке точной сетки  $[\xi_{i-1}, \xi_i]$   $g(x)$  – полином третьей степени;
- 2) функция  $g(x)$  и ее первые две производные непрерывны на всем отрезке  $[\xi_1, \xi_n]$ ;
- 3) функция  $g(x)$  удовлетворяет определенным граничным условиям.

Кубический сплайн, удовлетворяющий этим условиям, имеет вид

$$g(x) = m_{i-1} \frac{(\xi_i - x)^3}{6h_i} + m_i \frac{(x - \xi_{i-1})^3}{6h_i} + \left( \mu_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{\xi_i - x}{h_i} + \left( \mu_i - \frac{m_i h_i^2}{6} \right) \frac{x - \xi_{i-1}}{h_i},$$

где  $i$  выбирают из условия  $\xi_{i-1} \leq x \leq \xi_i$ ;  $h_i = \xi_i - \xi_{i-1}$ ;  $m_i$  и  $\mu_i$  – параметры сплайна, причем

Систему нелинейных уравнений можно решить аналитически. При этом предполагают, что новые значения  $\xi_i$  должны удовлетворять ограничениям  $|\tilde{x}_i - \xi_i| \leq 3q_i$ . После завершения итерационного процесса вычисляют значения сплайна и производной.

Из-за громоздкости вычислений корреляционные матрицы ошибок полученных оценок параметров не определяют, а вычисляют приближения для дисперсий параметров исходя из условия, что полученная сетка  $\Delta_\xi$  является точной. При этом определяют производные и поа, чтобы учесть сжатие интервалов при учете дополнительных условий.

Следует заметить, что при подготовке данных к обработке близко лежащие узлы сетки  $\Delta_1$  (т.е. такие, которые попадают в доверительный интервал соседних узлов) рекомендуется объединять в один с соответствующим пересчетом дисперсии и значения функции.

#### 4.8. Непосредственное определение оценок параметров экспоненты и функции Гаусса

В предыдущих параграфах показано, что оценки параметров экспоненты и функции Гаусса могут быть определены после соответствующих преобразований, приводящих исходную информацию к уравнению прямой линии. Подобный способ не всегда является желательным по следующим причинам:

1) искажаются функции плотностей вероятностей исходных данных, что влечет за собой изменение алгоритмов оценок;

2) не всегда преобразования обеспечивают однозначность физической интерпретации результатов. Так, при обработке статистических данных, подчиняющихся нормальному распределению, оценка дисперсии функции Гаусса в ряде работ определяется угловым коэффициентом в уравнении прямой линии, проведенной по точкам  $(x_i; \sqrt{\ln(y_{\max}/y_i)})$ ,  $i = 1, \dots, n$ . Здесь  $(x_i, y_i)$  – координаты  $i$ -й экспериментальной точки;  $y_{\max}$  – амплитуда гауссовского пика. Расчеты показывают, что ошибка в значении  $y_{\max}$  на  $n$  % для идеального гауссиана приводит к погрешности в определении полуширины гауссовского пика на  $2n$  %;

3) при обработке экспериментальных данных, поступающих с дискретных цифровых регистраторов, значение  $y_{\max}$  для функции

Гаусса может быть просто неизвестно и выполнить указанное преобразование невозможно.

Оценку параметров экспоненты  $\eta = Ae^{\lambda\xi}$  проводили в [5] путем определения точки минимума функционала

$$F = \frac{1}{2} \sum_{i=1}^n \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right]$$

при дополнительных условиях

$$\frac{x_i - \xi_i}{\sigma^2(x_i)} + \frac{y_i - \eta_i}{\sigma^2(y_i)} A \lambda e^{\lambda\xi_i} = 0.$$

В качестве оценок параметров  $(A, \sigma, \xi_0)$  функции Гаусса

$$\eta = \frac{A}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(\xi - \xi_0)^2}{2\sigma^2} \right]$$

принимали координаты точки минимума функционала  $F$ , а оценки истинных значений получали следующим образом:

$$\hat{\eta}_i = \frac{A}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(x_i - x_0)^2}{2\sigma^2} \right];$$

$$\hat{\xi}_i = x_0 + \sigma \sqrt{2 \ln \frac{A}{y_i \sqrt{2\pi}\sigma}}.$$

Эти условия позволили значительно упростить вычислительный процесс отыскания оценок.

В процессе вычисления оценок параметров функции Гаусса и экспоненты возникают определенные трудности: функционал  $F$  имеет сложный вид, содержит локальные экстремумы и континуумы, велико взаимодействие искомых параметров. При оценке неизвестных истинных координат экспериментальных точек в случае определения параметров функции Гаусса согласно приходится также сталкиваться с неоднозначностью решения.

Отыскание точки минимума функционала  $F$  при новых условиях проводилось методом сопряженных градиентов. Вычислительные эксперименты показали, что в практически реальных случаях итерационный процесс отыскания точки минимума функционала  $F$  сходится, когда нулевые приближения отличаются от точных значений до 40 %. Подобной точности нулевых приближений можно добиться-



Гаусса может быть просто неизвестно и выполнить указанное преобразование невозможно.

Оценку параметров экспоненты  $\eta = Ae^{\lambda\xi}$  проводили в [5] путем определения точки минимума функционала

$$F = \frac{1}{2} \sum_{i=1}^n \left[ \frac{(x_i - \xi_i)^2}{\sigma^2(x_i)} + \frac{(y_i - \eta_i)^2}{\sigma^2(y_i)} \right]$$

при дополнительных условиях

$$\frac{x_i - \xi_i}{\sigma^2(x_i)} + \frac{y_i - \eta_i}{\sigma^2(y_i)} A \lambda e^{\lambda\xi_i} = 0.$$

В качестве оценок параметров  $(A, \sigma, \xi_0)$  функции Гаусса

$$\eta = \frac{A}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(\xi - \xi_0)^2}{2\sigma^2} \right]$$

принимали координаты точки минимума функционала  $F$ , а оценки истинных значений получали следующим образом:

$$\hat{\eta}_i = \frac{A}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(x_i - x_0)^2}{2\sigma^2} \right];$$

$$\hat{\xi}_i = x_0 + \sigma \sqrt{2 \ln \frac{A}{y_i \sqrt{2\pi}\sigma}}.$$

Эти условия позволили значительно упростить вычислительный процесс отыскания оценок.

В процессе вычисления оценок параметров функции Гаусса и экспоненты возникают определенные трудности: функционал  $F$  имеет сложный вид, содержит локальные экстремумы и континуумы, велико взаимодействие искомых параметров. При оценке неизвестных истинных координат экспериментальных точек в случае определения параметров функции Гаусса согласно приходится также сталкиваться с неоднозначностью решения.

Отыскание точки минимума функционала  $F$  при новых условиях проводилось методом сопряженных градиентов. Вычислительные эксперименты показали, что в практически реальных случаях итерационный процесс отыскания точки минимума функционала  $F$  сходится, когда нулевые приближения отличаются от точных значений до 40 %. Подобной точности нулевых приближений можно добиться-

## Глава 5

# ОПРЕДЕЛЕНИЕ КООРДИНАТ ИСТОЧНИКА ИЗЛУЧЕНИЯ ПУТЕМ ОБЪЕДИНЕНИЯ ВСЕЙ ИНФОРМАЦИИ ПО ПЕЛЕНГАМ

Иди наперед, а оглядывайся назад.  
*Русская народная пословица*

Воспользуемся результатами, полученными в главе 1. Пусть нам известен набор пеленгов и их погрешностей от источников излучения, зарегистрированных разными пеленгаторами. Координаты регистраторов сигналов и погрешности этих координат также известны. Будем пользоваться для наглядности декартовой системой координат. Формулы перехода к другим системам координат известны [5]. Азимутальный и угломестный пеленги определяют координаты направляющего вектора прямой в пространстве, проходящей через точку с известными координатами  $M_1(x_0, y_0, z_0)$  (регистратор сигналов) и через точку с неизвестными координатами  $x, y$  и  $z$  (источник излучения). Каноническое уравнение прямой в пространстве, проходящей через точку  $M_1(x_0, y_0, z_0)$  параллельно вектору  $\vec{S} = l\vec{i} + m\vec{j} + n\vec{k}$ , имеет вид

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}.$$

Координаты  $l, m$  и  $n$  определяют по известным пеленгам.

Уравнение этой же прямой может быть записано как пересечение двух плоскостей

$$\begin{cases} \frac{x - x_0}{l} = \frac{y - y_0}{m}; \\ \frac{y - y_0}{m} = \frac{z - z_0}{n}. \end{cases}$$

Уравнение прямой на плоскости (например, на плоскости  $XU$ )

$$\frac{x - x_0}{l} = \frac{y - y_0}{m}.$$

Для прямой на плоскости введем  $l = \cos \alpha$ ;  $m = \sin \alpha$ . Тогда

$$x \sin \alpha - y \cos \alpha = x_0 \sin \alpha - y_0 \cos \alpha$$

или

$$x \operatorname{tg} \alpha - y = b, \text{ где } b = x_0 \operatorname{tg} \alpha - y_0.$$

В данном уравнении прямой две случайные величины:  $\operatorname{tg} \alpha$  и  $b$ .

При известных дисперсиях  $\sigma^2(x_0)$ ;  $\sigma^2(y_0)$ ;  $\sigma^2(\alpha)$  получим дисперсию

$$D(b) = D(x_0 \operatorname{tg} \alpha - y_0) = \left( \frac{x_0}{\cos^2 \alpha} \right)^2 \sigma^2(\alpha) + \operatorname{tg}^2 \alpha \cdot \sigma^2(x_0) + \sigma^2(y_0);$$

$$D(\operatorname{tg} \alpha) = \left( \frac{1}{\cos^4 \alpha} \right) \sigma^2(\alpha).$$

Собрав данные о пеленгах с  $N$  разных пеленгаторов, получим систему уравнений

$$x \operatorname{tg} \alpha_i - y = b_i, \quad i = 1, \dots, N, \quad (5.1)$$

в которой нужно определить координаты  $(x, y)$  источника излучений.

Метод наименьших квадратов применить нельзя, так как он применим только в случае, если в левой части уравнения нет случайных величин. Записав одно уравнение дважды с разными правыми частями  $b_i$  и  $\operatorname{tg} \alpha_i$ , получим две прямые – разные линии регрессии  $x$  на  $y$  и  $y$  на  $x$ , каждая из которых приведет к смещенным оценкам координат источника излучения.

Для определения координат источника радиоизлучения (ИРИ) необходимо найти точку пересечения линий ортогональной регрессии от каждого пеленгатора, т.е. решить систему уравнений (5.1). С этой целью найдем точку минимума функционала (см. главу 1)

$$F = \frac{1}{2} \sum_{i=1}^N \frac{(b_i - x \operatorname{tg}^2 \alpha_i + y)^2}{D(b_i) + x^2 D(\operatorname{tg} \alpha_i)}, \quad (5.2)$$

Которая определяет точечные оценки координат ИРИ. Чтобы определить координаты источника, нужно решить систему двух нелинейных уравнений с двумя неизвестными  $x$  и  $y$ :

$$\frac{\partial F}{\partial x} = 0; \quad \frac{\partial F}{\partial y} = 0. \quad (5.3)$$

Уравнение прямой на плоскости (например, на плоскости  $XU$ )

$$\frac{x - x_0}{l} = \frac{y - y_0}{m}.$$

Для прямой на плоскости введем  $l = \cos \alpha$ ;  $m = \sin \alpha$ . Тогда

$$x \sin \alpha - y \cos \alpha = x_0 \sin \alpha - y_0 \cos \alpha$$

или

$$x \operatorname{tg} \alpha - y = b, \text{ где } b = x_0 \operatorname{tg} \alpha - y_0.$$

В данном уравнении прямой две случайные величины:  $\operatorname{tg} \alpha$  и  $b$ .

При известных дисперсиях  $\sigma^2(x_0)$ ;  $\sigma^2(y_0)$ ;  $\sigma^2(\alpha)$  получим дисперсию

$$D(b) = D(x_0 \operatorname{tg} \alpha - y_0) = \left( \frac{x_0}{\cos^2 \alpha} \right)^2 \sigma^2(\alpha) + \operatorname{tg}^2 \alpha \cdot \sigma^2(x_0) + \sigma^2(y_0);$$

$$D(\operatorname{tg} \alpha) = \left( \frac{1}{\cos^4 \alpha} \right) \sigma^2(\alpha).$$

Собрав данные о пеленгах с  $N$  разных пеленгаторов, получим систему уравнений

$$x \operatorname{tg} \alpha_i - y = b_i, \quad i = 1, \dots, N, \quad (5.1)$$

в которой нужно определить координаты  $(x, y)$  источника излучений.

Метод наименьших квадратов применить нельзя, так как он применим только в случае, если в левой части уравнения нет случайных величин. Записав одно уравнение дважды с разными правыми частями  $b_i$  и  $\operatorname{tg} \alpha_i$ , получим две прямые – разные линии регрессии  $x$  на  $y$  и  $y$  на  $x$ , каждая из которых приведет к смещенным оценкам координат источника излучения.

Для определения координат источника радиоизлучения (ИРИ) необходимо найти точку пересечения линий ортогональной регрессии от каждого пеленгатора, т.е. решить систему уравнений (5.1). С этой целью найдем точку минимума функционала (см. главу 1)

$$F = \frac{1}{2} \sum_{i=1}^N \frac{(b_i - x \operatorname{tg}^2 \alpha_i + y)^2}{D(b_i) + x^2 D(\operatorname{tg} \alpha_i)}, \quad (5.2)$$

Которая определяет точечные оценки координат ИРИ. Чтобы определить координаты источника, нужно решить систему двух нелинейных уравнений с двумя неизвестными  $x$  и  $y$ :

$$\frac{\partial F}{\partial x} = 0; \quad \frac{\partial F}{\partial y} = 0. \quad (5.3)$$

Уравнение прямой на плоскости (например, на плоскости  $XU$ )

$$\frac{x - x_0}{l} = \frac{y - y_0}{m}.$$

Для прямой на плоскости введем  $l = \cos \alpha$ ;  $m = \sin \alpha$ . Тогда

$$x \sin \alpha - y \cos \alpha = x_0 \sin \alpha - y_0 \cos \alpha$$

или

$$x \operatorname{tg} \alpha - y = b, \text{ где } b = x_0 \operatorname{tg} \alpha - y_0.$$

В данном уравнении прямой две случайные величины:  $\operatorname{tg} \alpha$  и  $b$ .

При известных дисперсиях  $\sigma^2(x_0)$ ;  $\sigma^2(y_0)$ ;  $\sigma^2(\alpha)$  получим дисперсию

$$D(b) = D(x_0 \operatorname{tg} \alpha - y_0) = \left( \frac{x_0}{\cos^2 \alpha} \right)^2 \sigma^2(\alpha) + \operatorname{tg}^2 \alpha \cdot \sigma^2(x_0) + \sigma^2(y_0);$$

$$D(\operatorname{tg} \alpha) = \left( \frac{1}{\cos^4 \alpha} \right) \sigma^2(\alpha).$$

Собрав данные о пеленгах с  $N$  разных пеленгаторов, получим систему уравнений

$$x \operatorname{tg} \alpha_i - y = b_i, \quad i = 1, \dots, N, \quad (5.1)$$

в которой нужно определить координаты  $(x, y)$  источника излучений.

Метод наименьших квадратов применить нельзя, так как он применим только в случае, если в левой части уравнения нет случайных величин. Записав одно уравнение дважды с разными правыми частями  $b_i$  и  $\operatorname{tg} \alpha_i$ , получим две прямые – разные линии регрессии  $x$  на  $y$  и  $y$  на  $x$ , каждая из которых приведет к смещенным оценкам координат источника излучения.

Для определения координат источника радиоизлучения (ИРИ) необходимо найти точку пересечения линий ортогональной регрессии от каждого пеленгатора, т.е. решить систему уравнений (5.1). С этой целью найдем точку минимума функционала (см. главу 1)

$$F = \frac{1}{2} \sum_{i=1}^N \frac{(b_i - x \operatorname{tg}^2 \alpha_i + y)^2}{D(b_i) + x^2 D(\operatorname{tg} \alpha_i)}, \quad (5.2)$$

Которая определяет точечные оценки координат ИРИ. Чтобы определить координаты источника, нужно решить систему двух нелинейных уравнений с двумя неизвестными  $x$  и  $y$ :

$$\frac{\partial F}{\partial x} = 0; \quad \frac{\partial F}{\partial y} = 0. \quad (5.3)$$

Уравнение прямой на плоскости (например, на плоскости  $XU$ )

$$\frac{x - x_0}{l} = \frac{y - y_0}{m}.$$

Для прямой на плоскости введем  $l = \cos \alpha$ ;  $m = \sin \alpha$ . Тогда

$$x \sin \alpha - y \cos \alpha = x_0 \sin \alpha - y_0 \cos \alpha$$

или

$$x \operatorname{tg} \alpha - y = b, \text{ где } b = x_0 \operatorname{tg} \alpha - y_0.$$

В данном уравнении прямой две случайные величины:  $\operatorname{tg} \alpha$  и  $b$ .

При известных дисперсиях  $\sigma^2(x_0)$ ;  $\sigma^2(y_0)$ ;  $\sigma^2(\alpha)$  получим дисперсию

$$D(b) = D(x_0 \operatorname{tg} \alpha - y_0) = \left( \frac{x_0}{\cos^2 \alpha} \right)^2 \sigma^2(\alpha) + \operatorname{tg}^2 \alpha \cdot \sigma^2(x_0) + \sigma^2(y_0);$$

$$D(\operatorname{tg} \alpha) = \left( \frac{1}{\cos^4 \alpha} \right) \sigma^2(\alpha).$$

Собрав данные о пеленгах с  $N$  разных пеленгаторов, получим систему уравнений

$$x \operatorname{tg} \alpha_i - y = b_i, \quad i = 1, \dots, N, \quad (5.1)$$

в которой нужно определить координаты  $(x, y)$  источника излучений.

Метод наименьших квадратов применить нельзя, так как он применим только в случае, если в левой части уравнения нет случайных величин. Записав одно уравнение дважды с разными правыми частями  $b_i$  и  $\operatorname{tg} \alpha_i$ , получим две прямые – разные линии регрессии  $x$  на  $y$  и  $y$  на  $x$ , каждая из которых приведет к смещенным оценкам координат источника излучения.

Для определения координат источника радиоизлучения (ИРИ) необходимо найти точку пересечения линий ортогональной регрессии от каждого пеленгатора, т.е. решить систему уравнений (5.1). С этой целью найдем точку минимума функционала (см. главу 1)

$$F = \frac{1}{2} \sum_{i=1}^N \frac{(b_i - x \operatorname{tg}^2 \alpha_i + y)^2}{D(b_i) + x^2 D(\operatorname{tg} \alpha_i)}, \quad (5.2)$$

Которая определяет точечные оценки координат ИРИ. Чтобы определить координаты источника, нужно решить систему двух нелинейных уравнений с двумя неизвестными  $x$  и  $y$ :

$$\frac{\partial F}{\partial x} = 0; \quad \frac{\partial F}{\partial y} = 0. \quad (5.3)$$

## Глава 6

# ИДЕНТИФИКАЦИЯ ИСТОЧНИКОВ РАДИОАКТИВНЫХ БЛАГОРОДНЫХ ГАЗОВ

Машины должны работать. Люди должны думать.  
*Девиз компании «IBM»*

### 6.1. Источники радиоактивных благородных газов

При функционировании ядерных реакторов, импульсных источников нейтронов на основе делящихся тяжелых ядер, при мгновенном делении тяжелых ядер и переработке ядерного горючего (отработанного), ядерных взрывов – почти во всех этих случаях в атмосферу попадают радиоактивные изотопы благородных (инертных) газов – продукты деления изотопов тяжелых ядер (урана и плутония). Для оценки экологической ситуации и в других целях важно знать вид источника радиоактивных газов, это позволяет предсказать структуру ионизирующих излучений и более четко определить негативные последствия радиационного загрязнения [5]. Часто в данных случаях необходимо решать системы уравнений, в которых элементы систем имеют значительно большие погрешности, чем элементы правой части. Поэтому здесь необходимы методы конфлюэнтного анализа.

Выделим следующие процессы образования радиоактивных изотопов – продуктов деления тяжелых ядер:

1) мгновенное деление тяжелых ядер, которое приводит к независимому образованию различных осколков, и последующий распад осколков деления по изобарным цепочкам, который может происходить как без сепарации образующихся продуктов деления от предшествующих им по цепочкам распада, так и с сепарацией;

2) непрерывное в течение длительного времени деление тяжелых ядер и накопление продуктов деления без сепарации от предшествующих им по изобарным цепочкам;

3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.



3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14 МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.

3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14 МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.

3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14 МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.

3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14 МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.

3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.

3) сепарация изотопов инертных газов криптона и ксенона от предшествующих им изотопов по цепочке распада и последующий их распад (без подпитки).

Рассмотрим два источника образования радиоактивных благородных газов: мгновенное деление и накопление в ядерном реакторе. Для этих источников характерны все три указанных процесса образования газообразных продуктов деления. Формулы для расчета активности продуктов деления тяжелых ядер в этих случаях известны [5].

Активность радиоизотопов благородных газов (криптона и ксенона) зависит от типа делящихся ядер ( $U^{235}$ ,  $Pu^{239}$  и др.) и от энергии нейтронов, вызывающих деление (тепловых, нейтронов спектра деления, нейтронов с энергией 14 МэВ). Это позволяет использовать изотопы криптона и ксенона для идентификации источника деления (мгновенное деление, ядерный реактор), вида делящихся материалов (тяжелых ядер) и для определения числа делений нейтронами различных энергетических групп.

Изотопы криптона и ксенона появляются независимо в процессе деления тяжелых ядер и в результате последующих радиоактивных превращений по изобарным цепочкам распада. Практическую ценность для идентификации источников продуктов деления представляют изотопы, которые могут быть зарегистрированы. Теоретические результаты показывают, что в течение 24 ч после мгновенного деления можно обнаружить следующие изотопы:  $Kr^{83m}$ ,  $Kr^{85m}$ ,  $Kr^{85}$ ,  $Kr^{87}$ ,  $Kr^{88}$ ,  $Xe^{133m}$ ,  $Xe^{133}$ ,  $Xe^{135m}$ ,  $Xe^{135}$ ,  $Xe^{138}$  (табл. 6.1). Таблиц 6.1 рассчитана для случая деления урана – 235 тепловыми нейтронами.

В этой же таблице видно, какие изотопы определяют активность в пробе в конкретные моменты времени, что позволяет регистрировать активность каждого изотопа в пробе гамма-спектрометром. Однако анализ источника по этим изотопам криптона и ксенона затруднен тем, что они не являются начальными в цепочках распада (рис. 6.1).

При выходе в атмосферу радиоактивных изотопов инертных газов происходит сепарация (отделение) изотопов криптона и ксенона от предшествующих им по цепочке распада. Процесс сепарации, происходящий в результате взаимодействия элементов – членов цепочки (обладающих различными физическими и химическими свойствами) с окружающей средой, довольно сложен.

$$a_k^{\text{gen}}(\theta, \eta, \lambda, t, t_0, T) = \begin{cases} a_k^{\text{B}}(\theta, \eta, \lambda, t, T) & \text{при } t \leq T; \\ a_k(\theta, \eta, \lambda, t) & \text{при } T < t < t_0; \\ a_k^{\text{I}}(\theta, \eta, \lambda, t, t_0) & \text{при } t \geq t_0, \end{cases}$$

где  $T$  – время кампании реактора;  $t_0$  – время выдержки ТВЭЛа до разрушения (после этого момента происходит сепарация изотопов от предшественников).

### 6.2. Трудности идентификации источников изотопов криптона и ксенона

Для системы (6.2) можно получить решение  $\rho N_j$ ,  $j = 1, 2, \dots, m$ , и таким образом произвести идентификацию источника изотопов.

Однако при решении системы уравнений (6.2) возникает ряд проблем. Первая проблема уже упоминалась: изотопы криптона и ксенона не являются начальными в цепочках радиоактивных превращений; чтобы вычислить их активность, необходимо знать независимые выходы продуктов деления. Независимые выходы обычно содержат погрешность, более того, информация о независимых выходах некоторых продуктов деления при определенных видах деления вообще отсутствует.

Вторая проблема состоит в учете сепарации изотопов криптона и ксенона от предшественников по изобарной цепочке распада. Активность изотопов в отобранной пробе в значительной мере зависит от процессов сепарации. На рис. 6.2 представлена зависимость относительной активности изотопов  $\text{Kr}^{85\text{m}}$  и  $\text{Kr}^{88}$  (в сравнении с активностью изотопа  $\text{Xe}^{135}$ ). Верхняя и нижняя границы возможных значений относительной активности этих изотопов для различных видов деления при мгновенном делении в зависимости от времени и в предположении, что изотопы не сепарируются от предшествующих им, показаны сплошными (для  $\text{Kr}^{85\text{m}}$ ) и штриховыми (для  $\text{Kr}^{88}$ ) кривыми.

На рис. 6.2 приведены также измеренные относительные активности этих изотопов в образце (точки 1 и 2). Видно, что точки 1 и 2 не принадлежат возможным значениям относительной активности этих изотопов без сепарации.

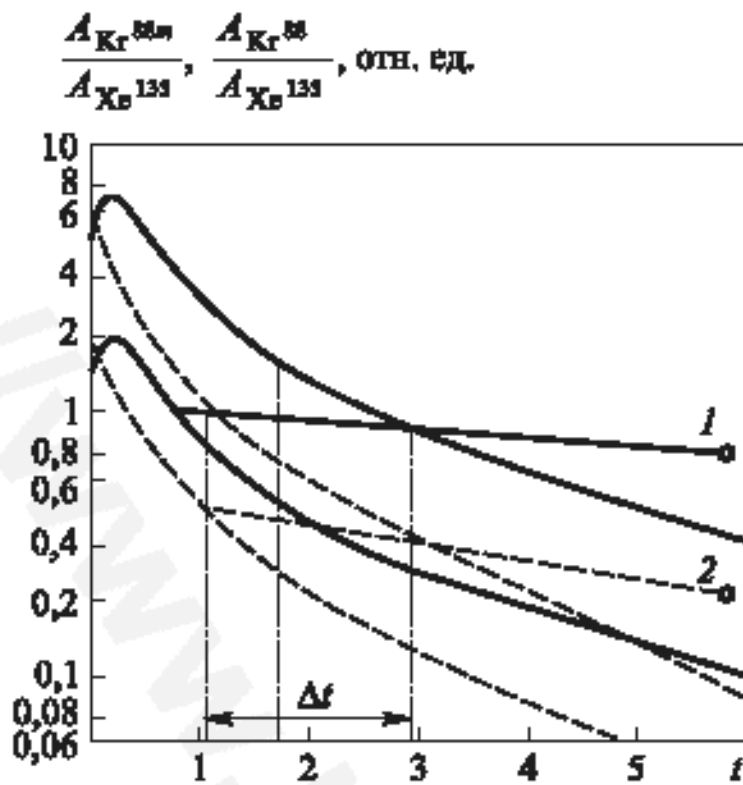


Рис. 6.2. Изменение относительной активности изотопов  $\text{Kr}^{85m}$  (сплошные линии) и  $\text{Kr}^{88}$  (штриховые линии) по отношению к активности  $\text{Xe}^{135}$  с учетом и без учета сепарации от предшествующих изотопов

Поэтому, чтобы оценить временной интервал  $\Delta t$  появления сепарации, рассчитаем относительные активности изотопов «в обратном времени», исходя из измеренных данных, в предположении, что отсутствует какое-либо влияние предшествующих изотопов на результаты измерения, т.е. активность изменяется по экспоненциальному закону (сплошные и штриховые прямые линии). Пересечение границ возможной относительной активности измеренных изотопов при мгновенном делении прямыми линиями показывает временной интервал  $\Delta t$ , когда, вероятно, произошло отделение изотопов от предшествующих им. Построив семейство таких кривых, можно качественно определить вид процесса деления, в результате которого появились изотопы криптона и ксенона. Так, по рис. 6.2 можно сделать вывод, что указанные изотопы образовались в процессе деления, которому соответствуют верхние сплошная и пунктирная кривые.

Влияние момента сепарации  $t_0$  на отношение активностей других изотопов видно из приведенной далее табл. 6.2. Следовательно, если неизвестны параметры процессов сепарации, невозможно строго вы-



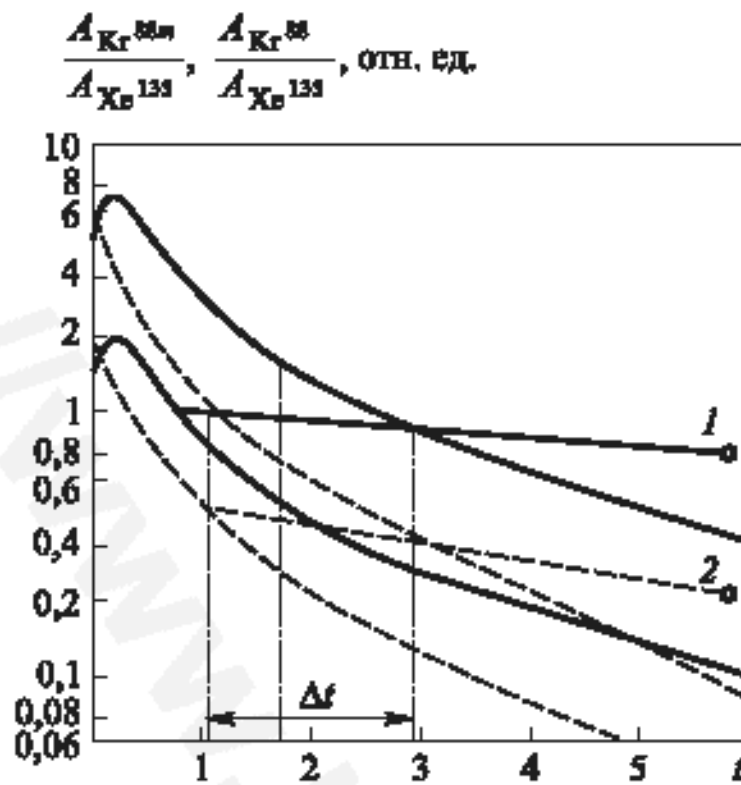


Рис. 6.2. Изменение относительной активности изотопов  $\text{Kr}^{85m}$  (сплошные линии) и  $\text{Kr}^{88}$  (штриховые линии) по отношению к активности  $\text{Xe}^{135}$  с учетом и без учета сепарации от предшествующих изотопов

Поэтому, чтобы оценить временной интервал  $\Delta t$  появления сепарации, рассчитаем относительные активности изотопов «в обратном времени», исходя из измеренных данных, в предположении, что отсутствует какое-либо влияние предшествующих изотопов на результаты измерения, т.е. активность изменяется по экспоненциальному закону (сплошные и штриховые прямые линии). Пересечение границ возможной относительной активности измеренных изотопов при мгновенном делении прямыми линиями показывает временной интервал  $\Delta t$ , когда, вероятно, произошло отделение изотопов от предшествующих им. Построив семейство таких кривых, можно качественно определить вид процесса деления, в результате которого появились изотопы криптона и ксенона. Так, по рис. 6.2 можно сделать вывод, что указанные изотопы образовались в процессе деления, которому соответствуют верхние сплошная и пунктирная кривые.

Влияние момента сепарации  $t_0$  на отношение активностей других изотопов видно из приведенной далее табл. 6.2. Следовательно, если неизвестны параметры процессов сепарации, невозможно строго вы-

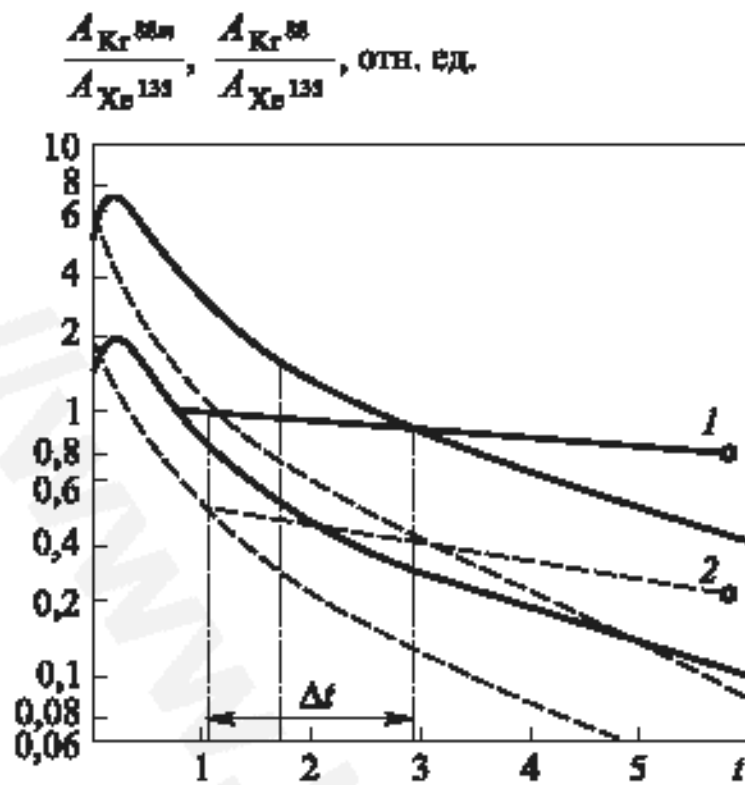


Рис. 6.2. Изменение относительной активности изотопов  $\text{Kr}^{85m}$  (сплошные линии) и  $\text{Kr}^{88}$  (штриховые линии) по отношению к активности  $\text{Xe}^{135}$  с учетом и без учета сепарации от предшествующих изотопов

Поэтому, чтобы оценить временной интервал  $\Delta t$  появления сепарации, рассчитаем относительные активности изотопов «в обратном времени», исходя из измеренных данных, в предположении, что отсутствует какое-либо влияние предшествующих изотопов на результаты измерения, т.е. активность изменяется по экспоненциальному закону (сплошные и штриховые прямые линии). Пересечение границ возможной относительной активности измеренных изотопов при мгновенном делении прямыми линиями показывает временной интервал  $\Delta t$ , когда, вероятно, произошло отделение изотопов от предшествующих им. Построив семейство таких кривых, можно качественно определить вид процесса деления, в результате которого появились изотопы криптона и ксенона. Так, по рис. 6.2 можно сделать вывод, что указанные изотопы образовались в процессе деления, которому соответствуют верхние сплошная и пунктирная кривые.

Влияние момента сепарации  $t_0$  на отношение активностей других изотопов видно из приведенной далее табл. 6.2. Следовательно, если неизвестны параметры процессов сепарации, невозможно строго вы-

Изображенные на рис. 6.3 кривые соответствуют делению  $U^{235}$  нейтронами спектра (1) и случаям реактора с кампанией  $T = 0,01$  года (2) и  $T = 3$  года (3) на момент  $0,1 \text{ ч} \leq t_0 \leq 5$  дней.

### 6.3. Алгоритм идентификации источников изотопов криптона и ксенона

Представленные результаты свидетельствуют о невозможности проведения процедуры идентификации по результатам наблюдений только двух изотопов, поскольку данный способ не позволяет достоверно определить процесс распада и надежно распознать источник изотопов. Следовательно, необходим надежный алгоритм идентификации источников изотопов, базирующийся на более подробном анализе физического процесса.

Рассмотрим пример такого алгоритма, основанного на конфлюэнтном анализе, который позволяет решить указанные проблемы.

Система (6.2) является линейной относительно неизвестных  $\rho N_j$ ,  $j = 1, 2, \dots, m$ , поэтому рассмотрим возможность ее решения в предположении, что время сепарации  $t_0$  известно, а элементы матрицы системы и правая часть – независимые случайные величины, распределенные по нормальному закону с математическими ожиданиями, равными нулю, и дисперсиями, равными соответственно  $\sigma^2(a_{ij}^{\text{true}})$  и  $\sigma^2(A_i^{\text{true}})$ :

$$\begin{aligned} a_{ij}^{t_0} &= a_{ij}^{\text{true}} \pm \varepsilon_{ij}, & \varepsilon_{ij} &\in N\left(0, \sigma^2(a_{ij}^{\text{true}})\right); \\ A_i &= A_i^{\text{true}} \pm \delta_i, & \delta_i &\in N\left(0, \sigma^2(A_i^{\text{true}})\right), \end{aligned}$$

где  $A_i^{\text{true}}$ ,  $a_{ij}^{\text{true}}$  – истинные значения активностей;  $N$  означает нормальное распределение.

Для учета погрешностей как в правой части, так и в элементах матрицы системы используем определение ортогональной регрессии [1]:

$$F_K = \frac{1}{2} \sum_{i=1}^n \left( \frac{\left( A_i - \rho \sum_{j=1}^m a_{ij}^{\text{true}} N_j \right)^2}{\sigma^2(A_i^{\text{true}})} + \sum_{j=1}^m \frac{(a_{ij} - a_{ij}^{\text{true}})^2}{\sigma^2(a_{ij}^{\text{true}})} \right). \quad (6.5)$$

В функционале (6.5) учитывается, что ошибки являются статистически независимыми величинами. В этом функционале наряду с неизвестным вектором  $\rho N_j$ ,  $j = 1, 2, \dots, m$ , неизвестными также

являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-

являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-

являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-

являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-

являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-



являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-

являются истинные значения вычисляемых активностей  $a_{ij}^{\text{true}}$ ,  $i = 1, 2, \dots, n$ ,  $j = 1, 2, \dots, m$ . Единственность и состоятельность оценок, получаемых методами конфлюэнтного анализа, доказаны в работе [4].

Здесь применен алгоритм минимизации функционала (6.5), структурная схема которого показана в п. 3.1.

На первом шаге алгоритма положим  $a_{ij}^{\text{true}} = a_{ij}$ , что сведет функционал (6.5) к функционалу метода наименьших квадратов, и получим первое приближение решения  $\rho N_j$ ,  $j = 1, 2, \dots, m$ . Далее используем какой-либо алгоритм решения плохо обусловленных систем линейных алгебраических уравнений с погрешностью в правой части для поиска оценки решения на каждой итерации, например метод регуляризации по числу итераций, описанный в работе [5].

Для получения оценок истинных значений  $a_{ij}^{\text{true}}$  при заданном значении  $t_0$  на каждом шаге получения оценок  $\hat{\rho N}_j$ ,  $j = 1, 2, \dots, m$ , используется условие

$$\left. \frac{\partial F_K}{\partial a_{ij}^{\text{true}}} \right|_{a_{ij}^{\text{true}} = \hat{a}_{ij}^{\text{true}}} = 0, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m,$$

что приводит к решению  $n$  систем линейных уравнений с  $m$  неизвестными следующего вида:

$$\sum_{r=1}^m \frac{\rho N_r \rho N_p}{\sigma^2(A_i^{\text{true}})} a_{ir}^{\text{true}} + \frac{a_{ip}^{\text{true}}}{\sigma^2(a_{ip}^{\text{true}})} = \frac{a_{ip}}{\sigma^2(a_{ip}^{\text{true}})} + \frac{\rho N_p A_i^{\text{true}}}{\sigma^2(A_i^{\text{true}})},$$

$$i = 1, 2, \dots, n, \quad p = 1, 2, \dots, m.$$

Полученные оценки значений  $\hat{a}_{ij}^{\text{true}}$  должны удовлетворять естественному условию (принадлежать области неопределенности измеренных значений  $a_{ij}$ ):

$$\left| a_{ij} - \hat{a}_{ij}^{\text{true}} \right| \leq 3\sigma(a_{ij}^{\text{true}}), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

Если это условие не выполняется, то  $a_{ij}^{\text{true}}$ ,  $j = 1, 2, \dots, m$ , которые не удовлетворяют этому неравенству, следует заменить на значения ближайших граничных точек. Из-за этого может происходить увеличение значений функционала на новых точных значениях переменных по сравнению с предыдущим шагом итерационного процесса, что приводит к снижению скорости схо-

## Глава 7

# ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ И НЕКОРРЕКТНЫЕ ЗАДАЧИ

Чудеса иногда случаются, но над этим приходится очень много работать.

*Хаим Вейцман*

### 7.1. Планирование экспериментов в конфлюэнтном анализе

Планированием эксперимента начали заниматься в 20–30-е годы XX века на базе работ Р. Фишера. В настоящее время можно выделить два основных направления: планирование экстремальных экспериментов и планирование экспериментов по выяснению механизма явлений. В первом случае цель эксперимента заключается в выборе таких значений параметров, при которых «выход» эксперимента будет наилучшим. Во втором устанавливают связь между «выходом» эксперимента и величинами, влияющими на него; находят математическую модель процесса, или, другими словами, выясняют вид поверхности отклика.

Что же дает планирование эксперимента?

1. Прежде всего, новые процедуры оказываются эффективнее ранее применявшихся.

Как показывает имеющийся опыт, планирование экспериментов с помощью методов математической статистики позволяет в 3...8 раз более эффективно использовать средства (время, деньги, материалы).

2. Планирование эксперимента существенным образом изменило мировоззрение, отношение к стратегии эксперимента:

а) математическая статистика внесла в эксперимент концепцию случая, стало возможным количественно оценивать влияние случайных факторов;

б) планирование эксперимента дает четкую логическую схему проведения всех операций. Исследователь знает когда, что и как надо делать и как надо интерпретировать результаты наблюдений. Так, в экстремальных задачах сначала находят линейное приближение для малого участка поверхности отклика, затем движутся по градиенту, потом переходят к описанию поверхности отклика полиномом второго, третьего и более высокого порядка. Далее проводится исследование полученного уравнения или нескольких уравнений, если речь идет не об одном параметре оптимизации. Такая четкая последовательность операций позволяет избежать тех ловушек, в которые нередко попадают из-за господствующих теоретических или эмпирических предрассудков. Этот выигрыш трудно выразить в какой-либо количественной мере.

Рассмотрим, почему же нет этих преимуществ в пассивном эксперименте.

Во-первых, исследователь обрабатывает методами многомерного регрессионного анализа (методом наименьших квадратов) данные, полученные в результате пассивного наблюдения за протеканием технологического процесса.

В активном эксперименте исследователь активно вмешивается в эксперимент, заранее разрабатывает его стратегию, хотя результаты эксперимента обрабатывает с помощью того же многомерного регрессионного анализа.

Многомерный регрессионный анализ очень чувствителен к нарушению исходных предпосылок. В частности, априори должна быть выбрана математическая модель (степень полинома и число независимых переменных). В пассивном эксперименте выбирают большее число переменных, считая, что часть из них в процессе обработки будет отнесена по  $t$ -критерию к незначимым. Но оценки коэффициентов регрессии в этом случае сильно коррелированы и оказываются смещенными.

В активном эксперименте отсеивание проводится на стадии линейного приближения, когда матрица независимых переменных ортогональна.

Вторая предпосылка: независимые переменные измеряются без ошибок. Эта предпосылка постоянно нарушается, но в активном экспе-

б) планирование эксперимента дает четкую логическую схему проведения всех операций. Исследователь знает когда, что и как надо делать и как надо интерпретировать результаты наблюдений. Так, в экстремальных задачах сначала находят линейное приближение для малого участка поверхности отклика, затем движутся по градиенту, потом переходят к описанию поверхности отклика полиномом второго, третьего и более высокого порядка. Далее проводится исследование полученного уравнения или нескольких уравнений, если речь идет не об одном параметре оптимизации. Такая четкая последовательность операций позволяет избежать тех ловушек, в которые нередко попадают из-за господствующих теоретических или эмпирических предрассудков. Этот выигрыш трудно выразить в какой-либо количественной мере.

Рассмотрим, почему же нет этих преимуществ в пассивном эксперименте.

Во-первых, исследователь обрабатывает методами многомерного регрессионного анализа (методом наименьших квадратов) данные, полученные в результате пассивного наблюдения за протеканием технологического процесса.

В активном эксперименте исследователь активно вмешивается в эксперимент, заранее разрабатывает его стратегию, хотя результаты эксперимента обрабатывает с помощью того же многомерного регрессионного анализа.

Многомерный регрессионный анализ очень чувствителен к нарушению исходных предпосылок. В частности, априори должна быть выбрана математическая модель (степень полинома и число независимых переменных). В пассивном эксперименте выбирают большее число переменных, считая, что часть из них в процессе обработки будет отнесена по  $t$ -критерию к незначимым. Но оценки коэффициентов регрессии в этом случае сильно коррелированы и оказываются смещенными.

В активном эксперименте отсеивание проводится на стадии линейного приближения, когда матрица независимых переменных ортогональна.

Вторая предпосылка: независимые переменные измеряются без ошибок. Эта предпосылка постоянно нарушается, но в активном экспе-

б) планирование эксперимента дает четкую логическую схему проведения всех операций. Исследователь знает когда, что и как надо делать и как надо интерпретировать результаты наблюдений. Так, в экстремальных задачах сначала находят линейное приближение для малого участка поверхности отклика, затем движутся по градиенту, потом переходят к описанию поверхности отклика полиномом второго, третьего и более высокого порядка. Далее проводится исследование полученного уравнения или нескольких уравнений, если речь идет не об одном параметре оптимизации. Такая четкая последовательность операций позволяет избежать тех ловушек, в которые нередко попадают из-за господствующих теоретических или эмпирических предрассудков. Этот выигрыш трудно выразить в какой-либо количественной мере.

Рассмотрим, почему же нет этих преимуществ в пассивном эксперименте.

Во-первых, исследователь обрабатывает методами многомерного регрессионного анализа (методом наименьших квадратов) данные, полученные в результате пассивного наблюдения за протеканием технологического процесса.

В активном эксперименте исследователь активно вмешивается в эксперимент, заранее разрабатывает его стратегию, хотя результаты эксперимента обрабатывает с помощью того же многомерного регрессионного анализа.

Многомерный регрессионный анализ очень чувствителен к нарушению исходных предпосылок. В частности, априори должна быть выбрана математическая модель (степень полинома и число независимых переменных). В пассивном эксперименте выбирают большее число переменных, считая, что часть из них в процессе обработки будет отнесена по  $t$ -критерию к незначимым. Но оценки коэффициентов регрессии в этом случае сильно коррелированы и оказываются смещенными.

В активном эксперименте отсеивание проводится на стадии линейного приближения, когда матрица независимых переменных ортогональна.

Вторая предпосылка: независимые переменные измеряются без ошибок. Эта предпосылка постоянно нарушается, но в активном экспе-

б) планирование эксперимента дает четкую логическую схему проведения всех операций. Исследователь знает когда, что и как надо делать и как надо интерпретировать результаты наблюдений. Так, в экстремальных задачах сначала находят линейное приближение для малого участка поверхности отклика, затем движутся по градиенту, потом переходят к описанию поверхности отклика полиномом второго, третьего и более высокого порядка. Далее проводится исследование полученного уравнения или нескольких уравнений, если речь идет не об одном параметре оптимизации. Такая четкая последовательность операций позволяет избежать тех ловушек, в которые нередко попадают из-за господствующих теоретических или эмпирических предрассудков. Этот выигрыш трудно выразить в какой-либо количественной мере.

Рассмотрим, почему же нет этих преимуществ в пассивном эксперименте.

Во-первых, исследователь обрабатывает методами многомерного регрессионного анализа (методом наименьших квадратов) данные, полученные в результате пассивного наблюдения за протеканием технологического процесса.

В активном эксперименте исследователь активно вмешивается в эксперимент, заранее разрабатывает его стратегию, хотя результаты эксперимента обрабатывает с помощью того же многомерного регрессионного анализа.

Многомерный регрессионный анализ очень чувствителен к нарушению исходных предпосылок. В частности, априори должна быть выбрана математическая модель (степень полинома и число независимых переменных). В пассивном эксперименте выбирают большее число переменных, считая, что часть из них в процессе обработки будет отнесена по  $t$ -критерию к незначимым. Но оценки коэффициентов регрессии в этом случае сильно коррелированы и оказываются смещенными.

В активном эксперименте отсеивание проводится на стадии линейного приближения, когда матрица независимых переменных ортогональна.

Вторая предпосылка: независимые переменные измеряются без ошибок. Эта предпосылка постоянно нарушается, но в активном экспе-

б) планирование эксперимента дает четкую логическую схему проведения всех операций. Исследователь знает когда, что и как надо делать и как надо интерпретировать результаты наблюдений. Так, в экстремальных задачах сначала находят линейное приближение для малого участка поверхности отклика, затем движутся по градиенту, потом переходят к описанию поверхности отклика полиномом второго, третьего и более высокого порядка. Далее проводится исследование полученного уравнения или нескольких уравнений, если речь идет не об одном параметре оптимизации. Такая четкая последовательность операций позволяет избежать тех ловушек, в которые нередко попадают из-за господствующих теоретических или эмпирических предрассудков. Этот выигрыш трудно выразить в какой-либо количественной мере.

Рассмотрим, почему же нет этих преимуществ в пассивном эксперименте.

Во-первых, исследователь обрабатывает методами многомерного регрессионного анализа (методом наименьших квадратов) данные, полученные в результате пассивного наблюдения за протеканием технологического процесса.

В активном эксперименте исследователь активно вмешивается в эксперимент, заранее разрабатывает его стратегию, хотя результаты эксперимента обрабатывает с помощью того же многомерного регрессионного анализа.

Многомерный регрессионный анализ очень чувствителен к нарушению исходных предпосылок. В частности, априори должна быть выбрана математическая модель (степень полинома и число независимых переменных). В пассивном эксперименте выбирают большее число переменных, считая, что часть из них в процессе обработки будет отнесена по  $t$ -критерию к незначимым. Но оценки коэффициентов регрессии в этом случае сильно коррелированы и оказываются смещенными.

В активном эксперименте отсеивание проводится на стадии линейного приближения, когда матрица независимых переменных ортогональна.

Вторая предпосылка: независимые переменные измеряются без ошибок. Эта предпосылка постоянно нарушается, но в активном экспе-



где  $\alpha^* = (v - n) / \lceil (v - 1)n \rceil$  выбрано из условия максимального уменьшения определителя корреляционной матрицы ошибки.

Шаги 2–4 с заменой  $\varepsilon_0$  и  $\varepsilon_1$  повторяют до тех пор, пока итерационная процедура не исчерпает лимита итераций либо не будет достигнута заданная точность оценок. В результате такой процедуры найдется  $(m_0 + 1)$ -я точка плана. Аналогично получают следующие точки оптимального плана.

## 7.2. Некорректные задачи в конфлюэнтном анализе

Итерационный алгоритм решения конфлюэнтной задачи состоит из двух основных элементов: решение системы линейных алгебраических уравнений для получения оценки и решение методом линеаризации систем нелинейных алгебраических уравнений для определения оценок истинных значений аргументов  $\xi$ . На обоих этапах возможна ситуация, когда бесконечно малым приращением в исходных данных могут отвечать сколь угодно большие изменения в решении. Такие системы называют плохообусловленными, а задачи – некорректными. Приращения в исходных данных могут быть вызваны как ошибками в измерениях, так и округлением величин в процессе расчетов на ЭВМ. В плохообусловленных системах строгое математическое решение может не соответствовать физической постановке задачи. Критической величиной, которая определяет физическую надежность строгого математического решения, является отношение наибольшего собственного значения симметричной матрицы  $A^T A$  ( $A$  – матрица системы алгебраических уравнений) к наименьшему. Квадратный корень этого отношения показывает увеличение помех в направлении, соответствующем наименьшему собственному значению.

В качестве другой характеристики обусловленности системы с квадратной матрицей  $A$  размером  $n \times n$  вводят  $N$ -обуславливающие числа,

$$N = \frac{1}{n} \|A\| \cdot \|A^{-1}\|,$$

где  $\|A\|$  и  $\|A^{-1}\|$  – соответственно нормы матрицы  $A$  и обратной матрицы  $A^{-1}$ .

Наилучшими обусловленными матрицами являются ортогональные, для которых  $N$ -обуславливающие числа равны единице.

В процессе обработки реальных экспериментов и аппроксимации исходных данных функциями различных видов с большим числом оцениваемых параметров в решении появляются осцилляции. Они возникают не только из-за погрешностей эксперимента, но и в результате неадекватного представления исследуемого явления выбранными функциями. С одной стороны, желательно описать как можно точнее изучаемое явление большим числом параметров, с другой стороны, увеличение размерности задачи ухудшает обусловленность систем и задача становится некорректной. Обусловленность систем линейных алгебраических уравнений еще больше ухудшается, когда мы имеем дело с коррелированными исходными случайными величинами.

При анализе и синтезе систем выделяют также класс некорректных задач. Задачу считают поставленной корректно, если ее решение удовлетворяет требованиям (условиям) Адамара, которые для операторного уравнения  $Y = A\theta$  формулируются следующим образом:

- 1) решение  $\theta$  существует для любого  $Y \in Q_A \subseteq F$ ;
- 2) решение  $\theta$  единственно в пространстве  $U$ ;
- 3) решение  $\theta$  непрерывно зависит от  $Y$ , т.е. если приращения  $\Delta Y$  стремятся к нулю, то приращение  $\Delta\theta$  также стремится к нулю.

Другими словами, задача некорректна, если определитель линейного оператора  $A$  равен нулю. Здесь  $\theta$  является элементом метрического пространства  $U$ , а  $Y$  – элементом метрического пространства  $F$ . Область определения оператора  $A$ , действующего из  $U$  в  $F$ , есть  $Q_A \subseteq U$ , область его значений  $Q_a = A(D_a) \subseteq F'$ .

Иногда плохо обусловленные и некорректные задачи называют просто некорректными. Этим термином мы будем пользоваться в дальнейшем.

Таким образом, чтобы корректно учитывать реальную экспериментальную информацию, необходимо иметь метод решения некорректных задач. Для этой цели применяется метод регуляризации А.Н. Тихонова и большая серия методов, развитых на его основе.

Строго говоря, в задачах, перечисленных в главе 2, мы должны были бы применять на каждой итерации метод регуляризации

Наилучшими обусловленными матрицами являются ортогональные, для которых  $N$ -обуславливающие числа равны единице.

В процессе обработки реальных экспериментов и аппроксимации исходных данных функциями различных видов с большим числом оцениваемых параметров в решении появляются осцилляции. Они возникают не только из-за погрешностей эксперимента, но и в результате неадекватного представления исследуемого явления выбранными функциями. С одной стороны, желательно описать как можно точнее изучаемое явление большим числом параметров, с другой стороны, увеличение размерности задачи ухудшает обусловленность систем и задача становится некорректной. Обусловленность систем линейных алгебраических уравнений еще больше ухудшается, когда мы имеем дело с коррелированными исходными случайными величинами.

При анализе и синтезе систем выделяют также класс некорректных задач. Задачу считают поставленной корректно, если ее решение удовлетворяет требованиям (условиям) Адамара, которые для операторного уравнения  $Y = A\theta$  формулируются следующим образом:

- 1) решение  $\theta$  существует для любого  $Y \in Q_A \subseteq F$ ;
- 2) решение  $\theta$  единственно в пространстве  $U$ ;
- 3) решение  $\theta$  непрерывно зависит от  $Y$ , т.е. если приращения  $\Delta Y$  стремятся к нулю, то приращение  $\Delta\theta$  также стремится к нулю.

Другими словами, задача некорректна, если определитель линейного оператора  $A$  равен нулю. Здесь  $\theta$  является элементом метрического пространства  $U$ , а  $Y$  – элементом метрического пространства  $F$ . Область определения оператора  $A$ , действующего из  $U$  в  $F$ , есть  $Q_A \subseteq U$ , область его значений  $Q_a = A(D_a) \subseteq F'$ .

Иногда плохо обусловленные и некорректные задачи называют просто некорректными. Этим термином мы будем пользоваться в дальнейшем.

Таким образом, чтобы корректно учитывать реальную экспериментальную информацию, необходимо иметь метод решения некорректных задач. Для этой цели применяется метод регуляризации А.Н. Тихонова и большая серия методов, развитых на его основе.

Строго говоря, в задачах, перечисленных в главе 2, мы должны были бы применять на каждой итерации метод регуляризации

Наилучшими обусловленными матрицами являются ортогональные, для которых  $N$ -обуславливающие числа равны единице.

В процессе обработки реальных экспериментов и аппроксимации исходных данных функциями различных видов с большим числом оцениваемых параметров в решении появляются осцилляции. Они возникают не только из-за погрешностей эксперимента, но и в результате неадекватного представления исследуемого явления выбранными функциями. С одной стороны, желательно описать как можно точнее изучаемое явление большим числом параметров, с другой стороны, увеличение размерности задачи ухудшает обусловленность систем и задача становится некорректной. Обусловленность систем линейных алгебраических уравнений еще больше ухудшается, когда мы имеем дело с коррелированными исходными случайными величинами.

При анализе и синтезе систем выделяют также класс некорректных задач. Задачу считают поставленной корректно, если ее решение удовлетворяет требованиям (условиям) Адамара, которые для операторного уравнения  $Y = A\theta$  формулируются следующим образом:

- 1) решение  $\theta$  существует для любого  $Y \in Q_A \subseteq F$ ;
- 2) решение  $\theta$  единственно в пространстве  $U$ ;
- 3) решение  $\theta$  непрерывно зависит от  $Y$ , т.е. если приращения  $\Delta Y$  стремятся к нулю, то приращение  $\Delta\theta$  также стремится к нулю.

Другими словами, задача некорректна, если определитель линейного оператора  $A$  равен нулю. Здесь  $\theta$  является элементом метрического пространства  $U$ , а  $Y$  – элементом метрического пространства  $F$ . Область определения оператора  $A$ , действующего из  $U$  в  $F$ , есть  $Q_A \subseteq U$ , область его значений  $Q_a = A(D_a) \subseteq F'$ .

Иногда плохо обусловленные и некорректные задачи называют просто некорректными. Этим термином мы будем пользоваться в дальнейшем.

Таким образом, чтобы корректно учитывать реальную экспериментальную информацию, необходимо иметь метод решения некорректных задач. Для этой цели применяется метод регуляризации А.Н. Тихонова и большая серия методов, развитых на его основе.

Строго говоря, в задачах, перечисленных в главе 2, мы должны были бы применять на каждой итерации метод регуляризации

Наилучшими обусловленными матрицами являются ортогональные, для которых  $N$ -обуславливающие числа равны единице.

В процессе обработки реальных экспериментов и аппроксимации исходных данных функциями различных видов с большим числом оцениваемых параметров в решении появляются осцилляции. Они возникают не только из-за погрешностей эксперимента, но и в результате неадекватного представления исследуемого явления выбранными функциями. С одной стороны, желательно описать как можно точнее изучаемое явление большим числом параметров, с другой стороны, увеличение размерности задачи ухудшает обусловленность систем и задача становится некорректной. Обусловленность систем линейных алгебраических уравнений еще больше ухудшается, когда мы имеем дело с коррелированными исходными случайными величинами.

При анализе и синтезе систем выделяют также класс некорректных задач. Задачу считают поставленной корректно, если ее решение удовлетворяет требованиям (условиям) Адамара, которые для операторного уравнения  $Y = A\theta$  формулируются следующим образом:

- 1) решение  $\theta$  существует для любого  $Y \in Q_A \subseteq F$ ;
- 2) решение  $\theta$  единственно в пространстве  $U$ ;
- 3) решение  $\theta$  непрерывно зависит от  $Y$ , т.е. если приращения  $\Delta Y$  стремятся к нулю, то приращение  $\Delta\theta$  также стремится к нулю.

Другими словами, задача некорректна, если определитель линейного оператора  $A$  равен нулю. Здесь  $\theta$  является элементом метрического пространства  $U$ , а  $Y$  – элементом метрического пространства  $F$ . Область определения оператора  $A$ , действующего из  $U$  в  $F$ , есть  $Q_A \subseteq U$ , область его значений  $Q_a = A(D_a) \subseteq F'$ .

Иногда плохо обусловленные и некорректные задачи называют просто некорректными. Этим термином мы будем пользоваться в дальнейшем.

Таким образом, чтобы корректно учитывать реальную экспериментальную информацию, необходимо иметь метод решения некорректных задач. Для этой цели применяется метод регуляризации А.Н. Тихонова и большая серия методов, развитых на его основе.

Строго говоря, в задачах, перечисленных в главе 2, мы должны были бы применять на каждой итерации метод регуляризации

## ЛИТЕРАТУРА

1. *Бородюк В.И., Вошинин А.П.* Ошибки регистрации независимых переменных в задачах множественной регрессии // Заводская лаборатория. – 1973. – Т. 39, № 7. – С. 217–222.
2. *Браунли К.А.* Статистическая теория и методология в науке и технике / Под ред. Л.Н. Большева. – М.: Наука, 1977. – 407 с.
3. *Гончарский А.В., Леонов А.С., Ягола А.Г.* Некоторое обобщение принципа невязки для случая оператора, заданного с ошибкой // Докл. АН СССР. – 1972. – Т. 203, № 6. – С. 1238–1239.
4. *Грешилов А.А.* Математические методы принятия решений: учеб. пособие (с расчетными программами на оптическом диске). – 2-е изд., испр. и доп. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2014. – 647 с.
5. *Грешилов А.А.* Некорректные задачи цифровой обработки информации и сигналов. – Изд. 2-е, доп – М.: Университетская книга; Логос, 2009. – 360 с.
6. *Грешилов А.А.* Анализ и синтез стохастических систем. Параметрические модели и конфлюэнтный анализ. – М.: Радио и связь, 1990. – 320с.
7. *Грешилов А.А.* Метод наименьших квадратов и элементы конфлюэнтного анализа – М.: Изд-во МГТУ им. Н.Э. Баумана, 1990. – 67 с.
8. *Грешилов А.А.* Компьютерные обучающие пособия для решения задач математической статистики и математического программирования. – Изд. 2-е., доп. (с расчетными программами на оптическом диске). – М.: Изд-во МГТУ им. Н.Э. Баумана, 2013. – 191 с.
9. *Грешилов А.А., Стакун В.А., Стакун А.А.* Статистические методы принятия решений с элементами конфлюэнтного анализа. – М.: Радио и связь, 1998. – 112 с.

10. *Жилинская Е.И., Товмаченко Н.Н., Федоров В.В.* Методы регрессионного анализа при наличии ошибки в предикторных переменных. – М.: Изд-во АН СССР, 1978. – 34 с.
11. *Круг Г.К., Кабанов В.А., Фомин Г.А., Фомина Г.С.* Планирование эксперимента в задачах нелинейного оценивания и распознавания образов. – М.: Наука, 1981. – 172 с.
12. *Леонов А.С.* Некоторые аспекты реализации регуляризующего алгоритма обобщенной невязки // Обработка и интерпретация физических экспериментов. Вып. 4. – М.: Изд-во МГУ, 1976. – С. 69–81.
13. *Морозов В.А.* Линейные и нелинейные некорректные задачи // Итоги науки и техники. Математический анализ. – Л. – М.: ВИНТИ, 1973. – 400 с.
14. *Успенский А.Б., Федоров В.В.* Вычислительные аспекты МНК при анализе и планировании регрессионных экспериментов. – М.: Изд-во Моск. ун-та, 1975. – 168 с.
15. *Frisch R.* Statistical confluence analysis by means of complete regression systems. – Oslo, 1934. – 192 p.
16. *Fuller W.A.* Measurement error models. – N. Y.: Wiley, 1987. – 440 p.
17. *Gleser L.J.* Improvements of the naive approach to estimation in nonlinear errors-in-variables regression models // Statist. Anal. Meas. Error Models and Appl.: Proc. AMS-IMS-SIAM. It Summer Res. Conf., Arcata, Calif., June 10 – 16, 1989. – Providence, 1990. – P. 99–114.
18. *Jefferys W.H.* Robust estimation when more then one variable per equation of condition has error // Biometrika. – 1990. – Vol. 77. – P. 597–607.
19. *Schater D.W.* Measurement error model estimation using iteratively weighted teast squares // Statist. Anal. Meas. Error Models and Appl.: Proc. AMS-IMS-SIAM. It Summer Res. Conf., Arcata, Calif., June 10 – 16, 1989. – Providence, 1990. – P. 129 – 138.
20. *Williamson J.H.* Least-squares fitting of straight line // Canad. J. Phys. – 1968. – Vol. 46, No. 16. – P. 1845.

**Описание программы «Регрессия»  
(инструкция для пользователя)**

Программа «Регрессия» предназначена для расчета параметров линейной регрессии  $Y$  на  $X$ ,  $X$  на  $Y$  и ортогональной регрессии, а также параметров прямой линии, параболы и полинома 3-й степени с помощью МНК. Главное окно программы представлено на рис. П1.1.

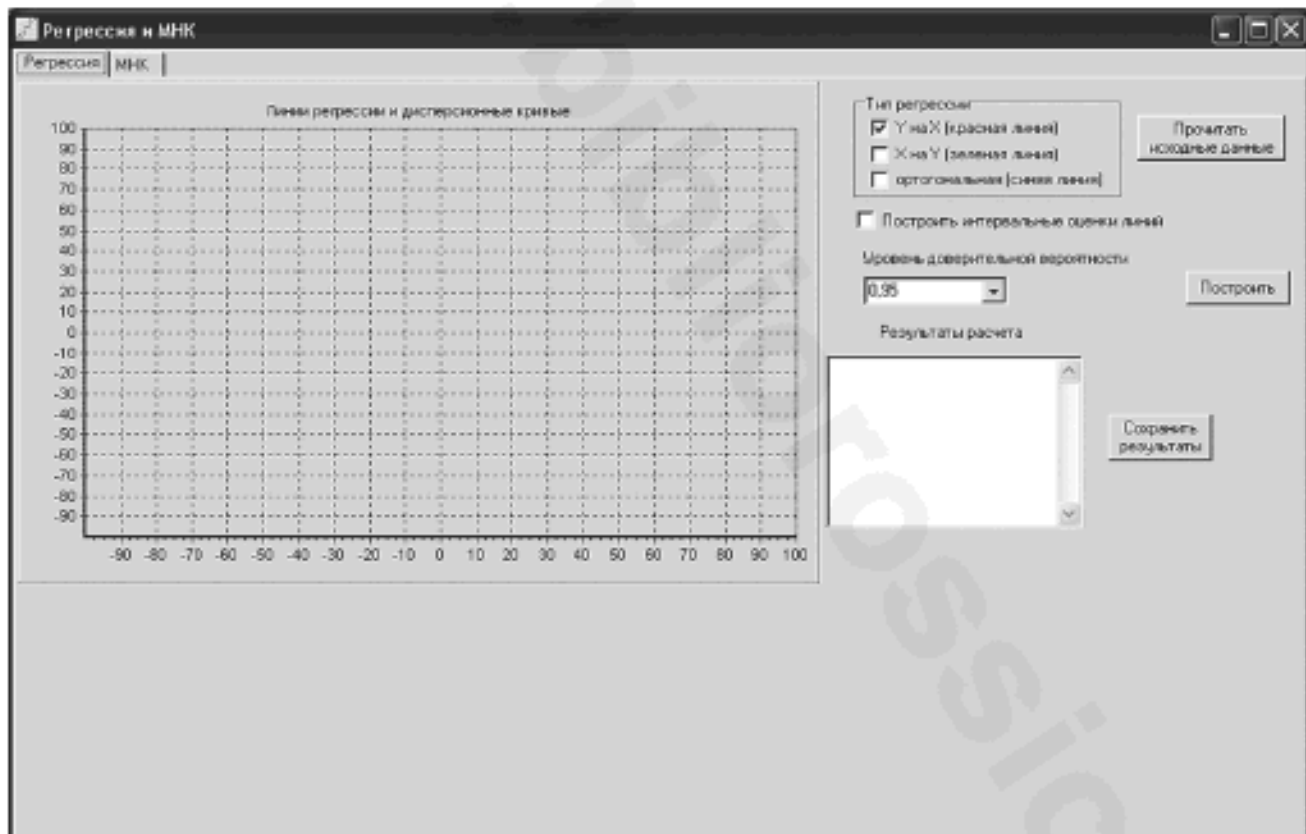


Рис. П1.1. Окно программы, вкладка «Регрессия»

В верхней части окна программы находятся две вкладки, которые позволяют переключаться между «режимами» регрессии и МНК.

Рассмотрим элементы, находящиеся на вкладке «Регрессия».

Для проведения расчетов необходимо загрузить в программу данные, для чего служит кнопка «Прочитать исходные данные». При ее нажатии появляется диалоговое окно выбора текстового файла с данными (рис. П1.2).



Текстовый файл можно создать в любом простом текстовом редакторе, например в Блокноте. Внимательно изучите формат файла исходных данных! В этом файле пять строк. В первой строке содержится число пар точек, по которым рассчитывается регрессия. Во второй строке находятся значения для  $X$ , в третьей – средние квадратические отклонения каждого измерения для  $X$ , в четвертой и пятой – значения и средние квадратические отклонения для  $Y$ . В строке между числами может быть любое число пробелов и символов табуляции, между строками с данными не должно быть пустых строк. Пример текстового файла представлен ниже:

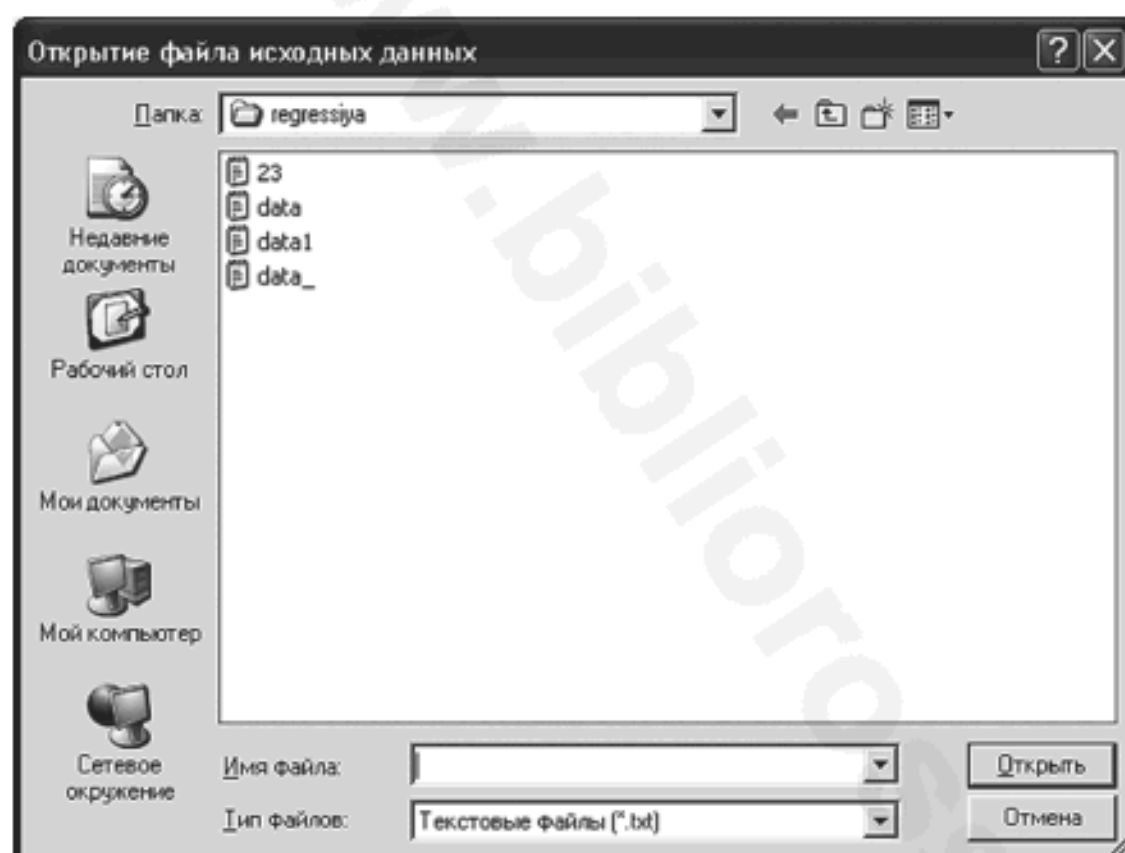


Рис. П1.2. Диалоговое окно выбора файла исходных данных

```

5
0 2 3 4 8
1.1 2.2 1.4 3.1 1.1
8 2 6 4 13
1 2 1 2 3

```

По данным одного файла строятся все три линии регрессии.

Следует обратить внимание, что при записи в файл дробных чисел разделителем целой и дробной частей является точка, а не запятая (см. третью строку приведенного примера текстового файла).

Текстовый файл можно создать в любом простом текстовом редакторе, например в Блокноте. Внимательно изучите формат файла исходных данных! В этом файле пять строк. В первой строке содержится число пар точек, по которым рассчитывается регрессия. Во второй строке находятся значения для  $X$ , в третьей – средние квадратические отклонения каждого измерения для  $X$ , в четвертой и пятой – значения и средние квадратические отклонения для  $Y$ . В строке между числами может быть любое число пробелов и символов табуляции, между строками с данными не должно быть пустых строк. Пример текстового файла представлен ниже:



Рис. П1.2. Диалоговое окно выбора файла исходных данных

```

5
0 2 3 4 8
1.1 2.2 1.4 3.1 1.1
8 2 6 4 13
1 2 1 2 3

```

По данным одного файла строятся все три линии регрессии.

Следует обратить внимание, что при записи в файл дробных чисел разделителем целой и дробной частей является точка, а не запятая (см. третью строку приведенного примера текстового файла).

Текстовый файл можно создать в любом простом текстовом редакторе, например в Блокноте. Внимательно изучите формат файла исходных данных! В этом файле пять строк. В первой строке содержится число пар точек, по которым рассчитывается регрессия. Во второй строке находятся значения для  $X$ , в третьей – средние квадратические отклонения каждого измерения для  $X$ , в четвертой и пятой – значения и средние квадратические отклонения для  $Y$ . В строке между числами может быть любое число пробелов и символов табуляции, между строками с данными не должно быть пустых строк. Пример текстового файла представлен ниже:



Рис. П1.2. Диалоговое окно выбора файла исходных данных

```

5
0 2 3 4 8
1.1 2.2 1.4 3.1 1.1
8 2 6 4 13
1 2 1 2 3

```

По данным одного файла строятся все три линии регрессии.

Следует обратить внимание, что при записи в файл дробных чисел разделителем целой и дробной частей является точка, а не запятая (см. третью строку приведенного примера текстового файла).

Текстовый файл можно создать в любом простом текстовом редакторе, например в Блокноте. Внимательно изучите формат файла исходных данных! В этом файле пять строк. В первой строке содержится число пар точек, по которым рассчитывается регрессия. Во второй строке находятся значения для  $X$ , в третьей – средние квадратические отклонения каждого измерения для  $X$ , в четвертой и пятой – значения и средние квадратические отклонения для  $Y$ . В строке между числами может быть любое число пробелов и символов табуляции, между строками с данными не должно быть пустых строк. Пример текстового файла представлен ниже:

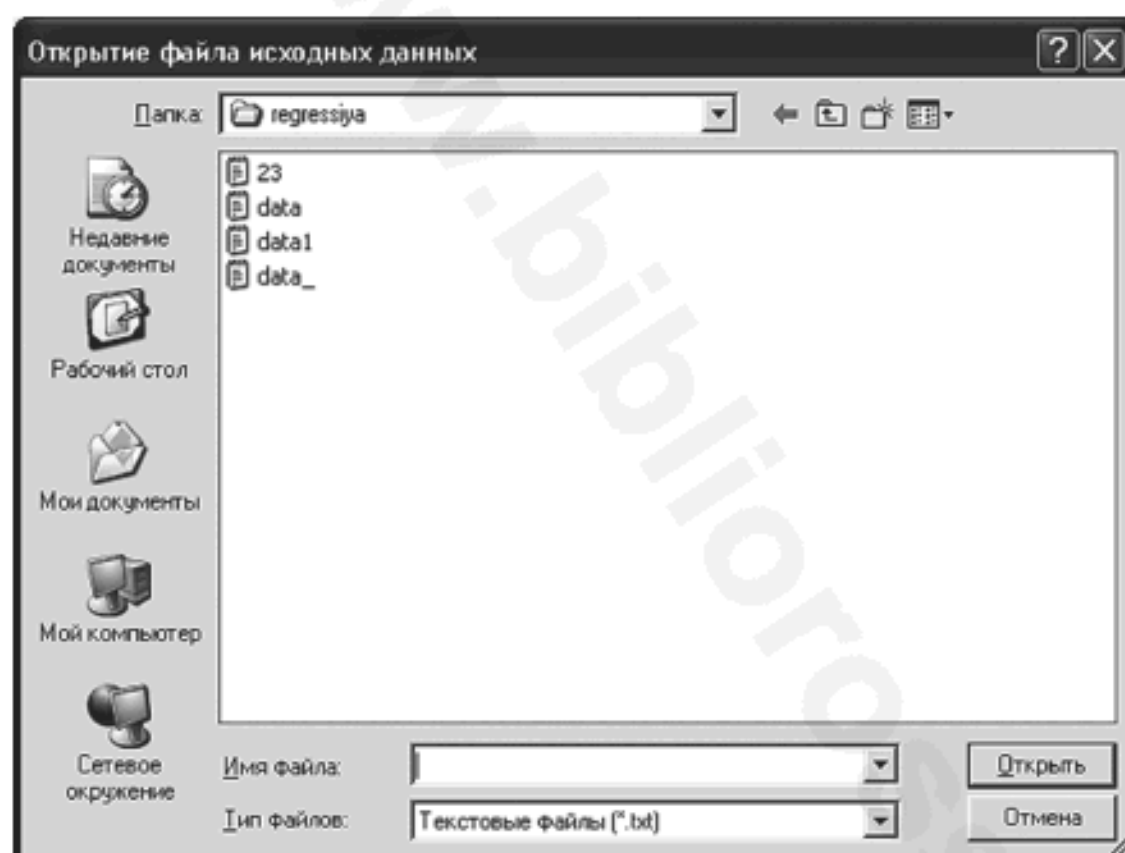


Рис. П1.2. Диалоговое окно выбора файла исходных данных

```

5
0 2 3 4 8
1.1 2.2 1.4 3.1 1.1
8 2 6 4 13
1 2 1 2 3

```

По данным одного файла строятся все три линии регрессии.

Следует обратить внимание, что при записи в файл дробных чисел разделителем целой и дробной частей является точка, а не запятая (см. третью строку приведенного примера текстового файла).

Текстовый файл можно создать в любом простом текстовом редакторе, например в Блокноте. Внимательно изучите формат файла исходных данных! В этом файле пять строк. В первой строке содержится число пар точек, по которым рассчитывается регрессия. Во второй строке находятся значения для  $X$ , в третьей – средние квадратические отклонения каждого измерения для  $X$ , в четвертой и пятой – значения и средние квадратические отклонения для  $Y$ . В строке между числами может быть любое число пробелов и символов табуляции, между строками с данными не должно быть пустых строк. Пример текстового файла представлен ниже:

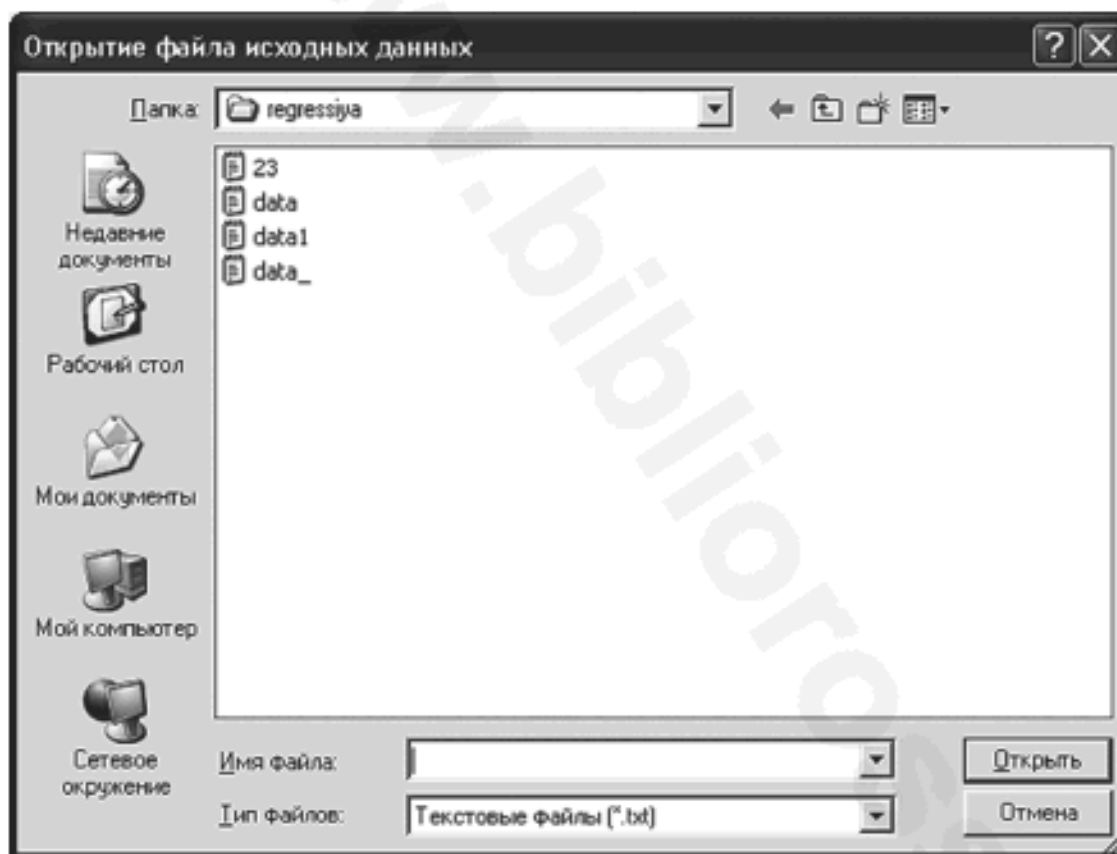


Рис. П1.2. Диалоговое окно выбора файла исходных данных

```

5
0 2 3 4 8
1.1 2.2 1.4 3.1 1.1
8 2 6 4 13
1 2 1 2 3

```

По данным одного файла строятся все три линии регрессии.

Следует обратить внимание, что при записи в файл дробных чисел разделителем целой и дробной частей является точка, а не запятая (см. третью строку приведенного примера текстового файла).

Текстовый файл можно создать в любом простом текстовом редакторе, например в Блокноте. Внимательно изучите формат файла исходных данных! В этом файле пять строк. В первой строке содержится число пар точек, по которым рассчитывается регрессия. Во второй строке находятся значения для  $X$ , в третьей – средние квадратические отклонения каждого измерения для  $X$ , в четвертой и пятой – значения и средние квадратические отклонения для  $Y$ . В строке между числами может быть любое число пробелов и символов табуляции, между строками с данными не должно быть пустых строк. Пример текстового файла представлен ниже:

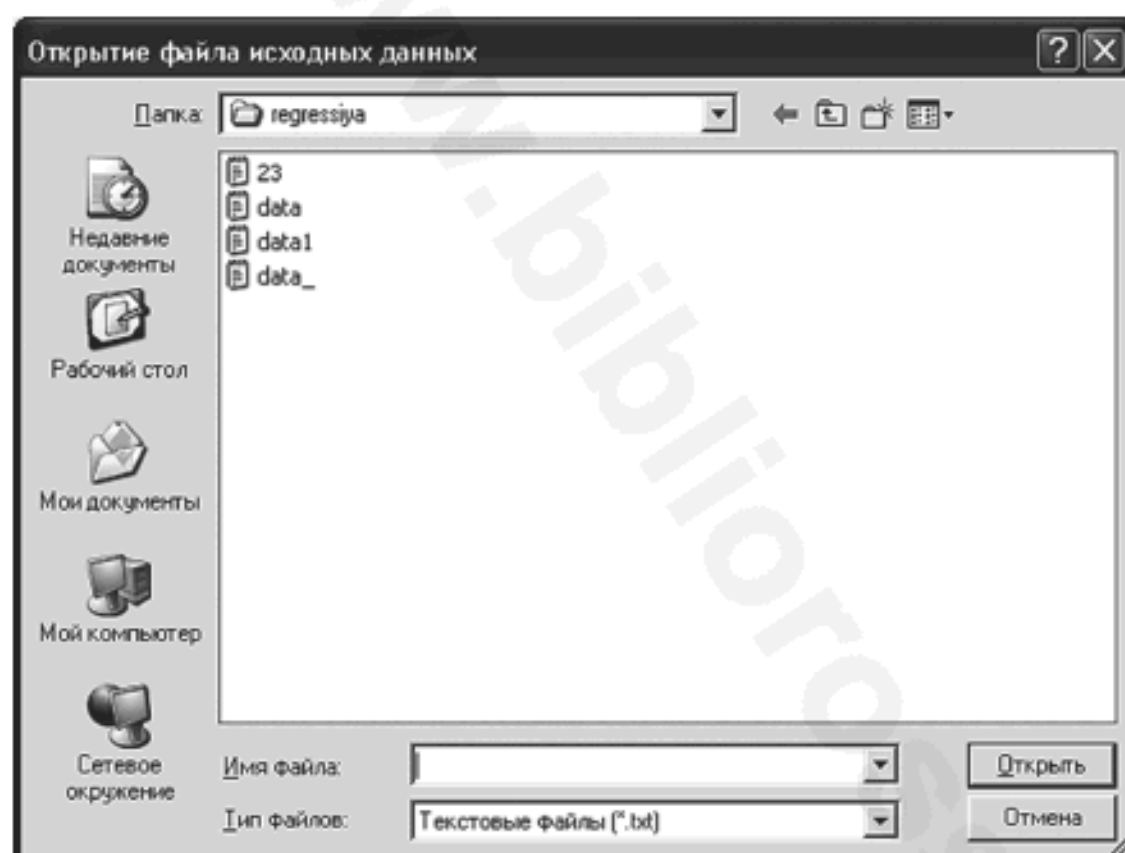


Рис. П1.2. Диалоговое окно выбора файла исходных данных

```

5
0 2 3 4 8
1.1 2.2 1.4 3.1 1.1
8 2 6 4 13
1 2 1 2 3

```

По данным одного файла строятся все три линии регрессии.

Следует обратить внимание, что при записи в файл дробных чисел разделителем целой и дробной частей является точка, а не запятая (см. третью строку приведенного примера текстового файла).

### Вкладка «Метод наименьших квадратов»

Вкладка «МНК» предназначена для расчетов параметров прямой линии, параболы и полинома 3-й степени методом наименьших квадратов. Также здесь предусмотрена возможность построения доверительных интервалов для указанных типов линий.

Внешний вид программы при активной вкладке «МНК» приведен на рис. П2.1.

Для загрузки исходных данных в программу служит кнопка «Прочитать исходные данные», при нажатии которой открывается стандартное диалоговое окно выбора файла, аналогичное приведенному на рис. П1.2. Внимательно изучите формат файла исходных данных!

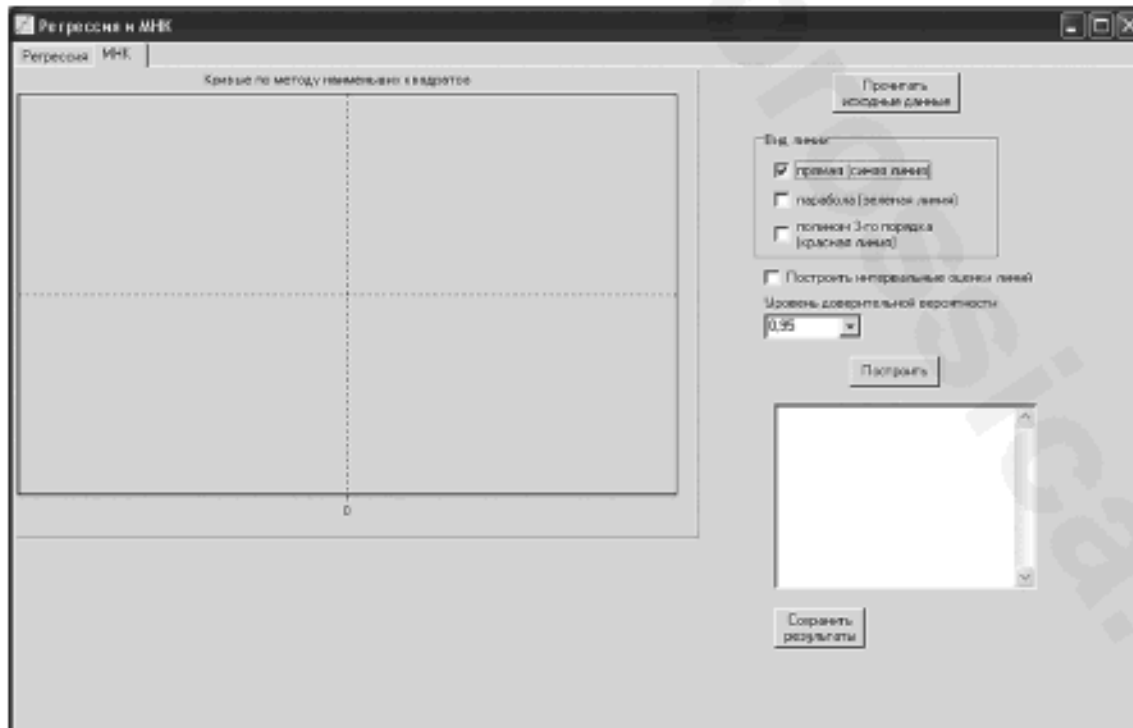


Рис. П2.1. Окно программы, вкладка «МНК»

Файл исходных данных должен состоять из трех строк: в первой строке указывается число точек, во второй перечислены значения  $X$ , в третьей –  $Y$ . Пример содержимого файла приведен ниже:

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

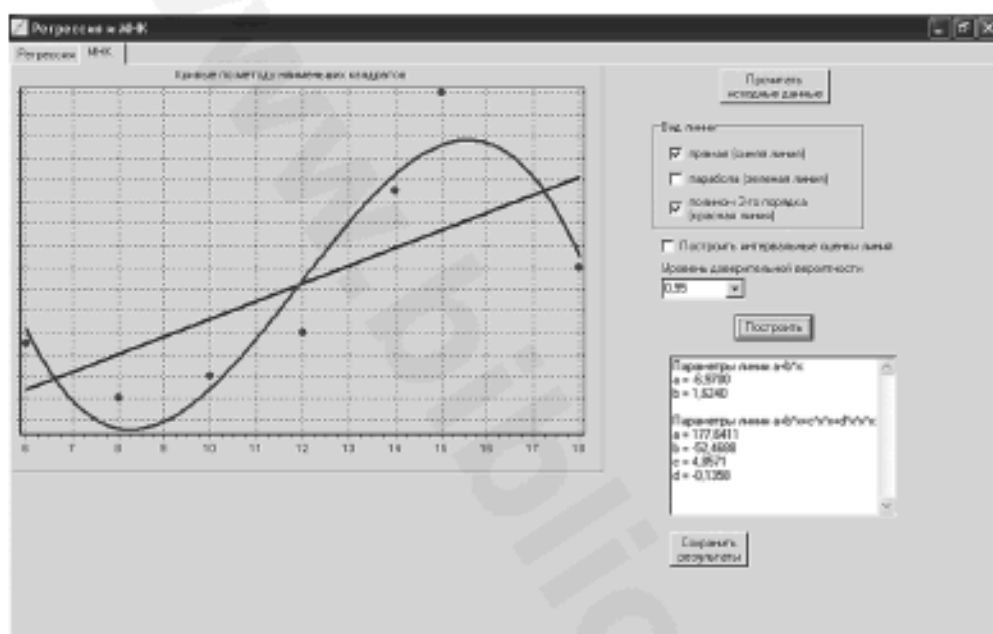


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

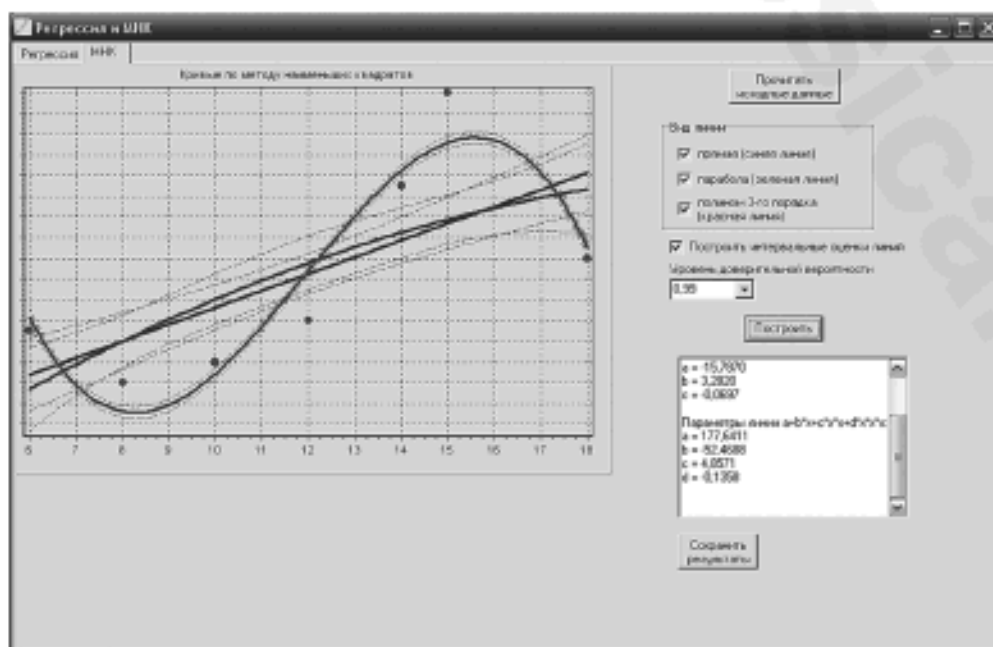


Рис. П2.3. Расчет параметров линий методом наименьших квадратов



7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

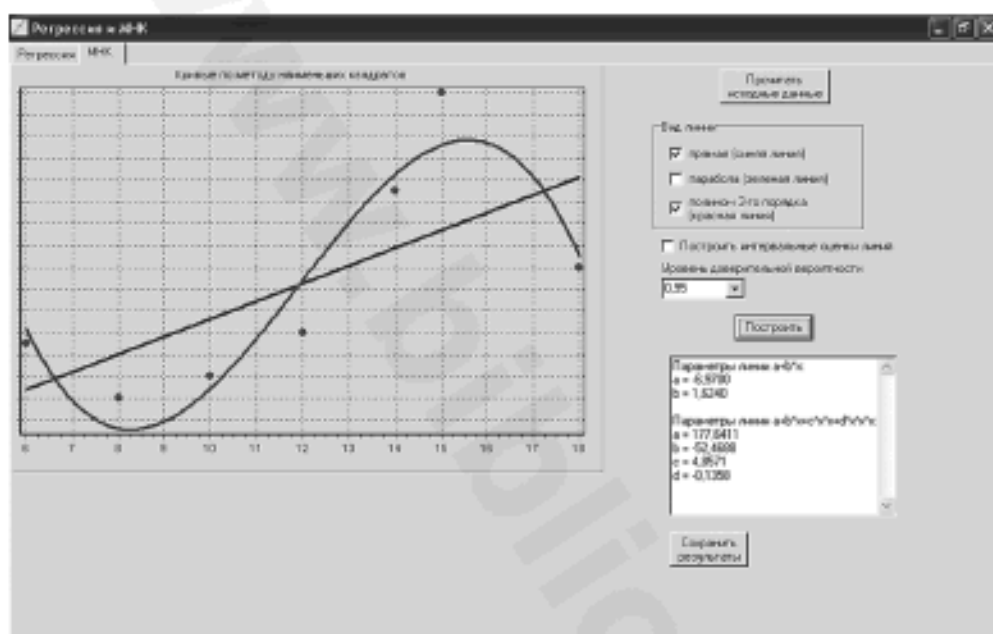


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

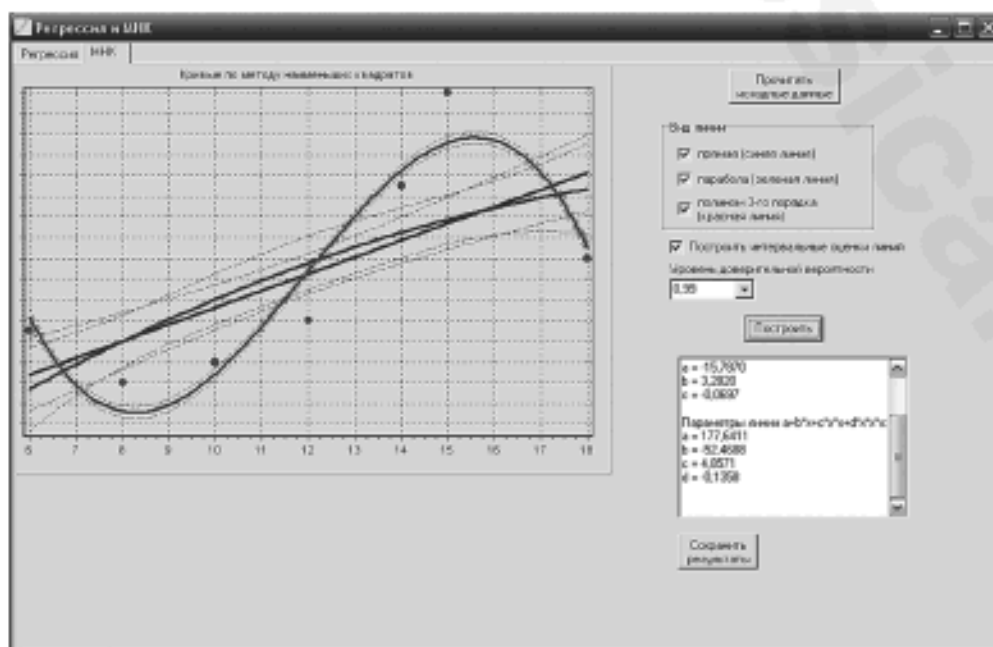


Рис. П2.3. Расчет параметров линий методом наименьших квадратов

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

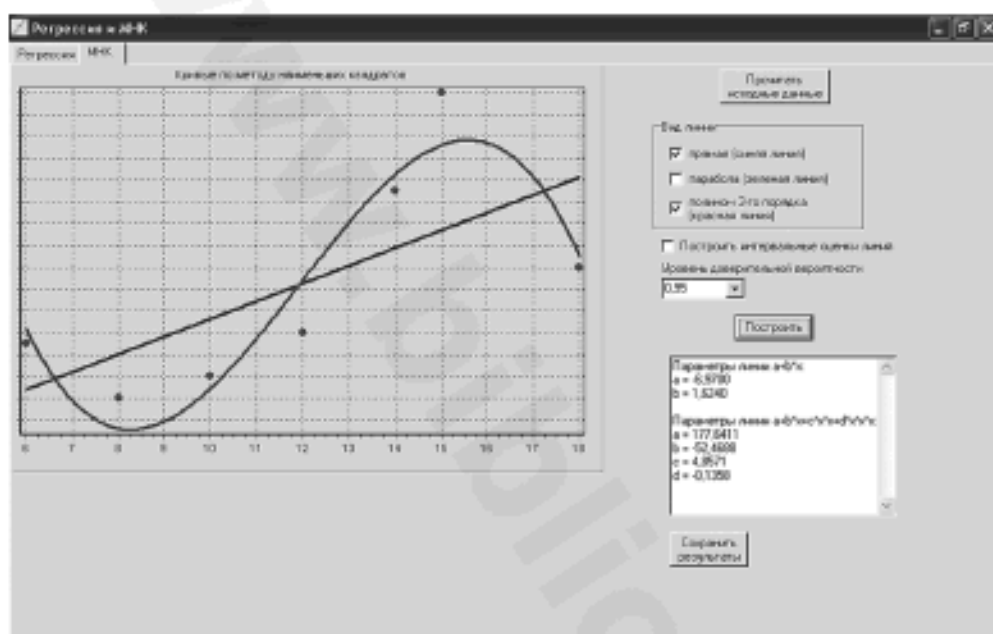


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

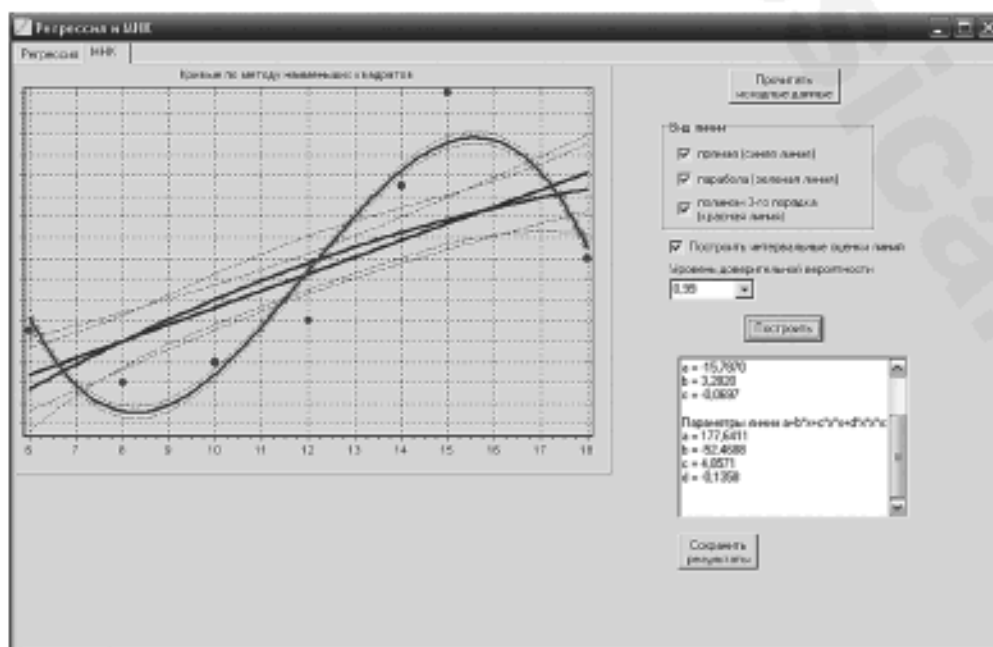


Рис. П2.3. Расчет параметров линий методом наименьших квадратов

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

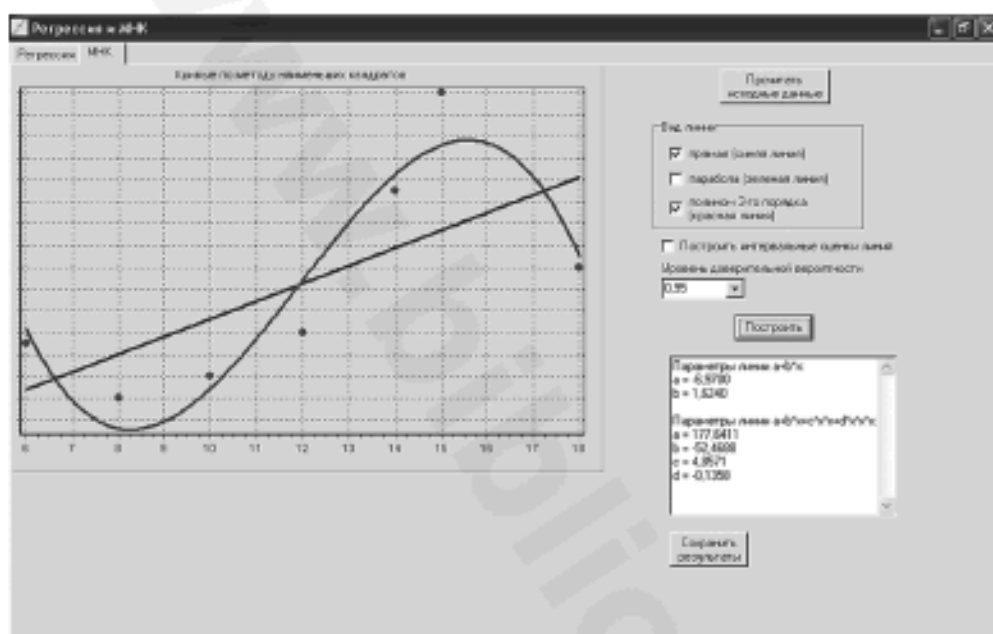


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

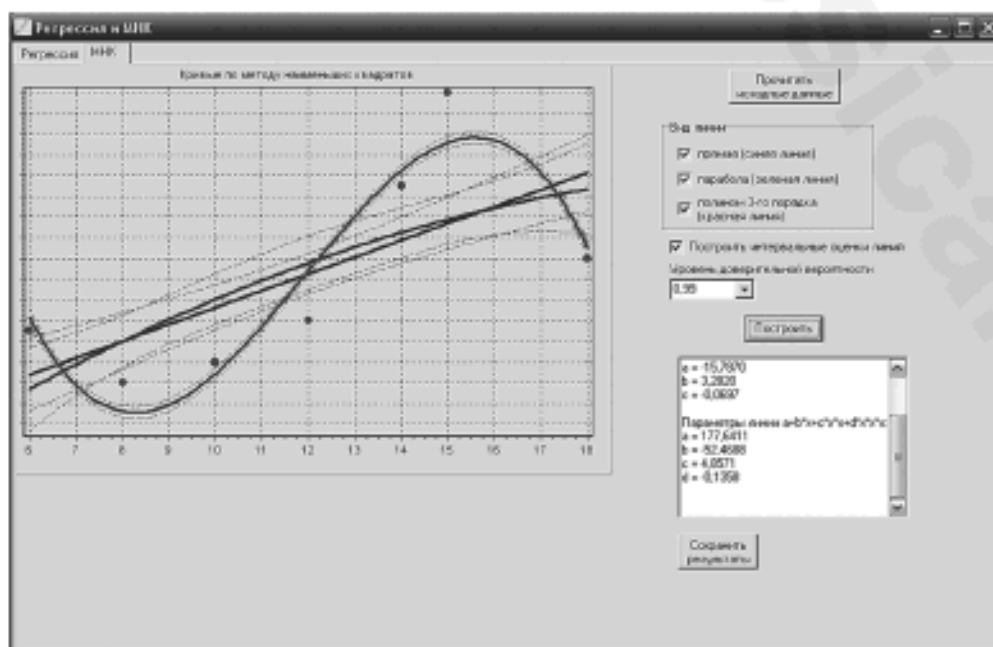


Рис. П2.3. Расчет параметров линий методом наименьших квадратов

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

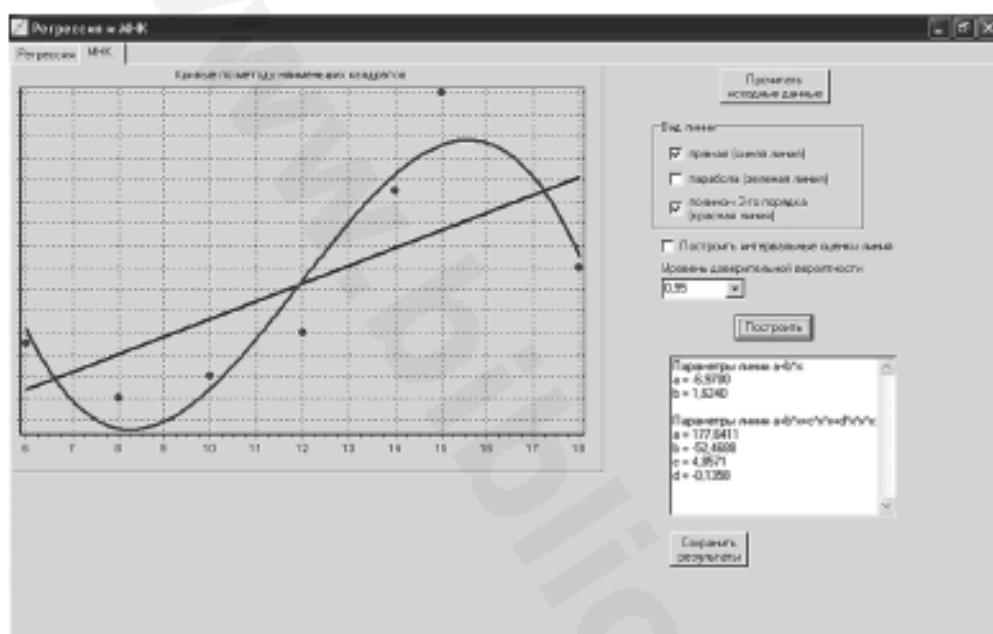


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

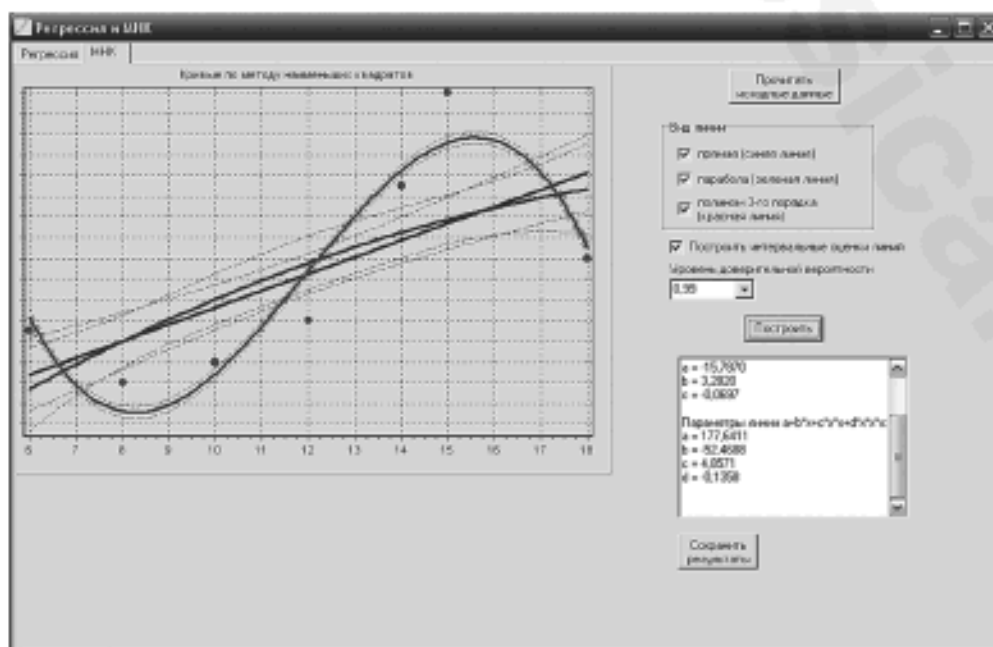


Рис. П2.3. Расчет параметров линий методом наименьших квадратов

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

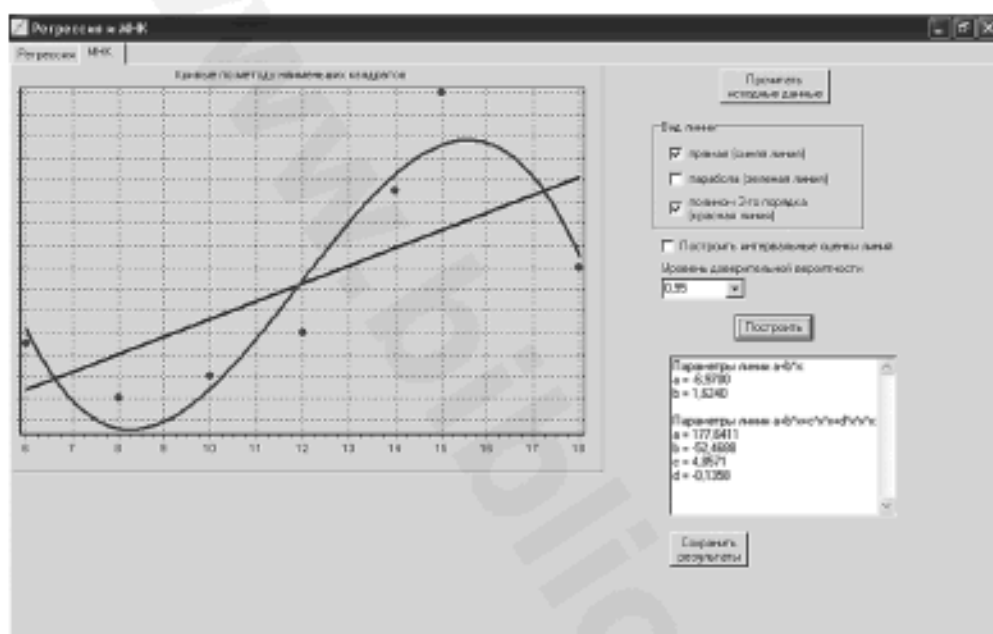


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

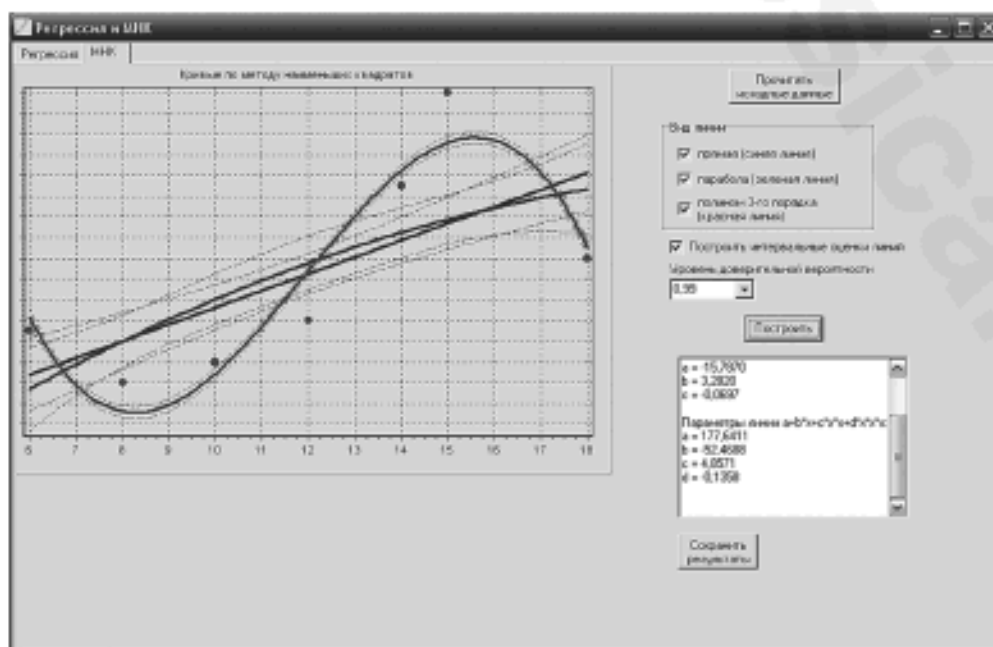


Рис. П2.3. Расчет параметров линий методом наименьших квадратов

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

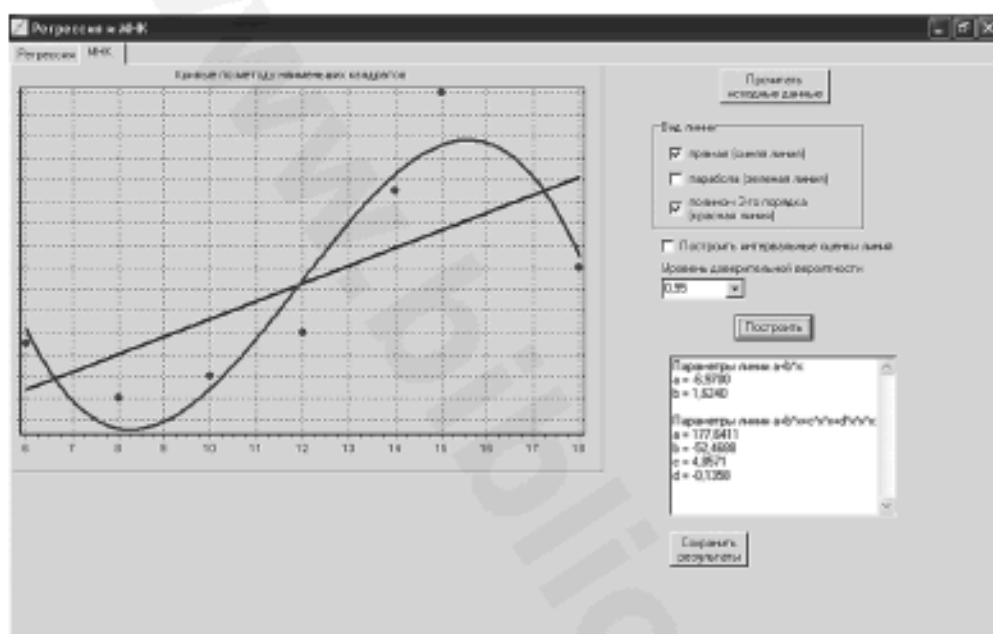


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

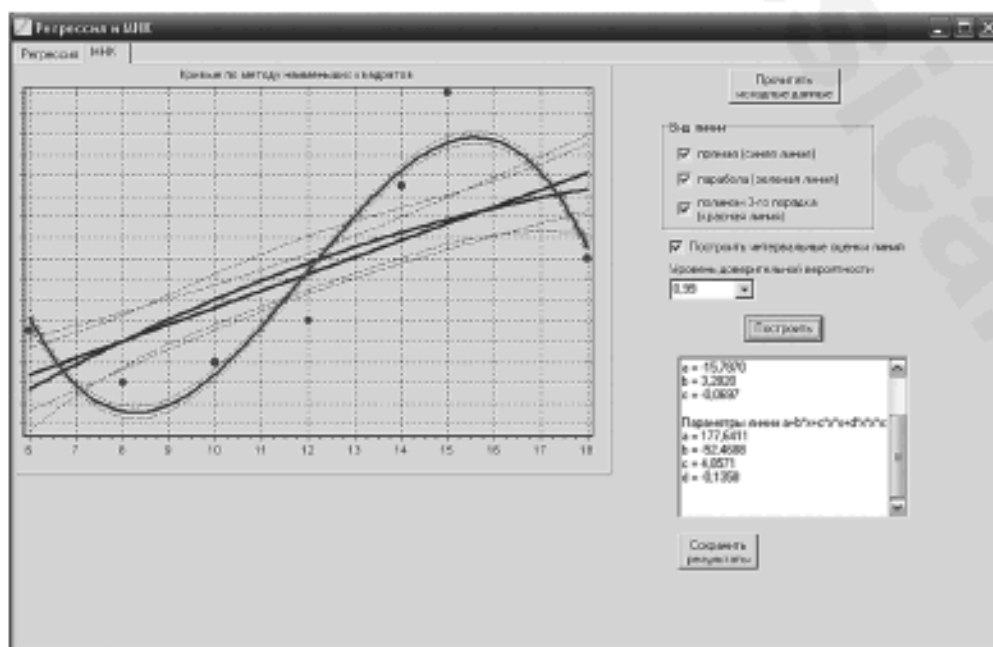


Рис. П2.3. Расчет параметров линий методом наименьших квадратов

7

8.1 6.2 9.9 12.1 14.2 14.9 18.2

1.9 6.9 3.9 8.1 21.2 30.1 13.8

После выбора файла следует в поле «Вид линии» отметить «галочками» те типы линий, параметры которых будут рассчитываться, и нажать кнопку «Построить». При этом будут отображены исходные точки и соответствующие линии, а в текстовом поле справа – параметры выбранных типов линий (рис. П2.2).

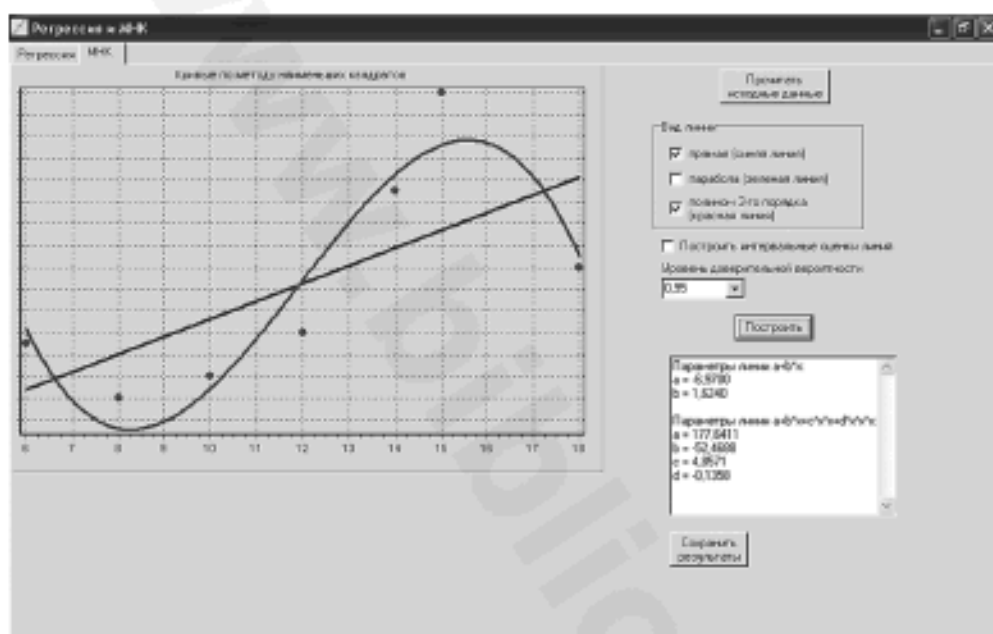


Рис. П2.2. Расчет параметров линий методом наименьших квадратов

Если отметить флажком параметр «Построить интервальные оценки линий», то при выводе будут отображаться интервальные оценки прямой линии, параболы или полинома 3-й степени (рис. П2.3).

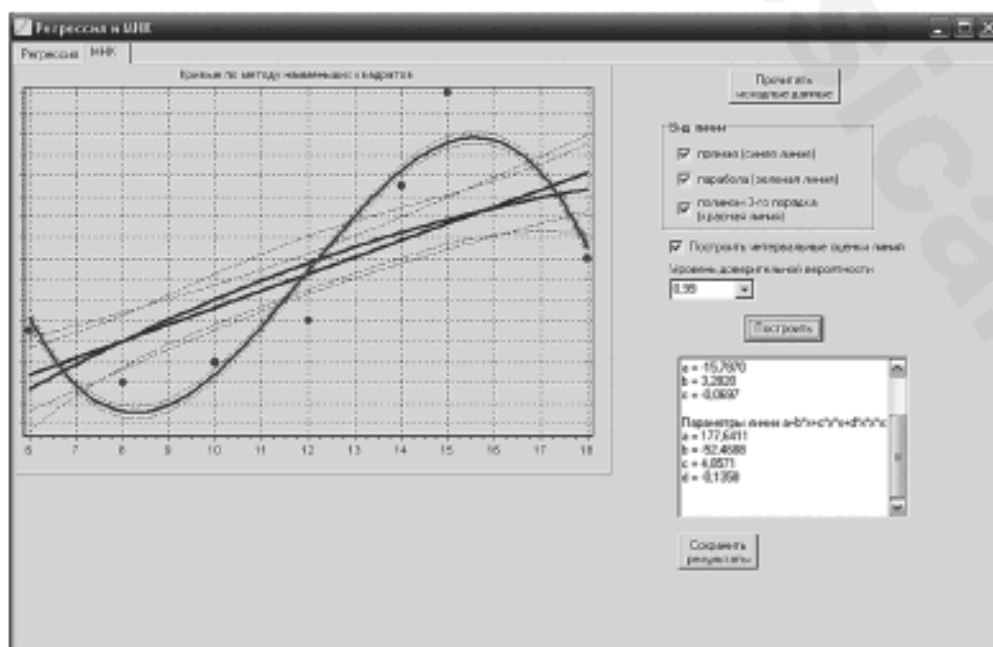


Рис. П2.3. Расчет параметров линий методом наименьших квадратов