Б. А. ГОРЛАЧ,
В. Г. ШАХОВ

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
ПОСТРОЕНИЕ МОДЕЛЕЙ И ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ

РЕКОМЕНДОВАНО
к изданию редакционно-издательским советом
Самарского государственного аэрокосмического университета
им. академика С. П. Королева в качестве учебного пособия
для студентов вузов, обучающихся по программам высшего
образования в областях: «Инженерное дело, технологии
и технические науки» и «Науки об обществе»

САНКТ-ПЕТЕРБУРГ · МОСКВА · КРАСНОДАР
2016

ББК 22.16я73
Г 69

Горлач Б. А., Шахов В. Г.

Г 69 Математическое моделирование. Построение моделей и численная реализация: Учебное пособие. — СПб.: Издательство «Лань», 2016. — 292 с. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-2168-8

Излагаются, сопровождаемые примерами решения задач, основы математического моделирования и численной реализации математических моделей в объеме минимально необходимых знаний для инженерного образования. Приведены примеры решения конкретных задач.

Пособие предназначено для студентов, обучающихся по инженерным и экономическим направлениям подготовки.

ББК 22.16я73

Обложка
Е. А. ВЛАСОВА

© Издательство «Лань», 2016
© Б. А. Горлач, В. Г. Шахов, 2016
© Издательство «Лань»,
художественное оформление, 2016

ПРЕДИСЛОВИЕ

Учебное пособие предназначено для студентов инженерных и экономических специальностей и включает в себя 6 глав.

1. **Моделирование объектов**

Приведены основные понятия и определения предмета, дана классификация моделей по различным признакам. Описаны основные и производные единицы измерения.

2. **Математическое и физическое моделирование**

Раскрыты понятия подобия математических и физических объектов. Рассмотрены степенные комплексы и приведены методы установления их подобия. Описаны способы установления масштабов и составления масштабных уравнений. Раскрыта суть π-теоремы. Описаны критерии подобия и модельный эксперимент. Приведены методы определения критериев подобия путем анализа размерностей. Разобраны решения задач, сгруппированные в три темы семинарских занятий.

3. **Численные методы**

Описаны методы численной реализации решений некоторых математических моделей: итерационные методы определения корней нелинейных алгебраических уравнений, методы хорд и касательных; методы пробных точек сокращения интервалов унимодальности функций, в частности, метод золотого сечения; симплексный метод и метод градиентного спуска определения экстремумов функций нескольких переменных; методы трапеций и Симпсона (Simpson Thomas (1710–1761) — английский математик) численного интегрирования; методы Эйлера (Euler Leonhard (1707–1783) — швейцарский, немецкий и российский математик и механик), Рунге–Ку́тты (Runge Karl David (1856–1927), Kutta Martin Wilgelm (1867–1944) — немецкие физики и математики) и конечных

разностей интегрирования дифференциальных уравнений. Описаны методы оценки точности получаемых решений. Дано задание на выполнение этапа курсовой работы.

4. **Теория корреляции**

Обсуждаются возможности детерминированного, корреляционного и стохастического подходов к построению зависимостей между различными величинами. Рассмотрен метод наименьших квадратов, аппроксимации экспериментальных данных и построения уравнений регрессии. Рассмотрены различные характеристики уравнений линейной и нелинейной регрессии. Рассмотрен пример построения уравнений регрессии по статистическим данным. Дано задание на выполнение этапа курсовой работы.

5. **Динамика случайных процессов**

Рассмотрены подходы к построению, сглаживанию и анализу временных рядов, к формированию трендов, сезонных и случайных составляющих временных рядов. Дано задание на выполнение этапа курсовой работы.

6. **Прогнозирование**

Описаны подходы к построению аппроксимирующих функций, методы проверки их адекватности реальным процессам, рассмотрены различные подходы к составлению прогноза протекания процессов и явлений. Дано задание на выполнение этапа курсовой работы.

Первые три главы описывают подходы к построению и реализации решения детерминированных математических моделей, в последних трех главах обсуждаются вопросы, связанные с моделированием случайных процессов.

Знакомство с материалом пособия предполагает наличие у читателей начальных знаний по линейной алгебре, математическому анализу, математической статистике, физике, в том числе теоретической механике, электротехнике, экономике на уровне понимания основных закономерностей рассматриваемых в этих предметах процессов и явлений.

ВВЕДЕНИЕ

Никто не может сказать, когда зародилась математика. Но можно предположить, что она началась со счета, с момента, когда человек (или человекообразный) сообразил, что один сорванный им плод и еще один — это два плода; один предмет и еще один предмет — это два предмета.

В переходе от конкретного плода к абстракции в виде произвольных предметов даже при счете до двух заключается смысл создания первой математической модели. От конкретного действия с плодами человек перешел к действию над абстрактными объектами, приводящему к новому объекту. Простейшая математическая модель:

$$\text{один} + \text{один} = \text{два}.$$

Может быть после двух для первобытного человека начиналось понятие «много». По мере развития умственных способностей человека появилась необходимость расширения числа цифр. Возможно, в ход пошли пальцы руки, которые позволяли считать до десяти. С гениальной догадкой человека ввести для обозначения «ничего» цифру «нуль» была сформирована система счета, которая позволила создавать математические модели численного соотношения между предметами.

Развитие человеческого мышления ставило новые задачи. Появились разнообразные действия над числами: вычитание, деление, умножение, возведение в степень, сравнение, дифференцирование, интегрирование и т. д.

Возможно, предположение о таком процессе развития математики позволило составителям энциклопедии дать такое определение: «Математика — это наука, занимающаяся построением и изучением абстрактных количественных моделей».

Возникает вопрос: что представляет собой математическая модель, то есть то, с помощью чего можно судить о реальных предметах, процессах, явлениях?

Мало кто из образованных людей может сомневаться в том, что современный уровень развития математики позволяет строить математические модели любых объектов, процессов и явлений, для которых можно установить закономерности их строения и изменения.

Математика в своем развитии зашла так далеко, что стала сама себя моделировать. Один из величайших математиков прошедшего столетия Д. Гильберт (Hilbert David (1862–1943) — немецкий математик) считал, что математика должна быть свободной от предметной области и «ее понятия должны быть связаны только требованиями быть непротиворечивыми и соответствовать понятиям, введенным ранее посредством точных определений».

Отметим, что такой подход к построению математических моделей оказался весьма плодотворным. Характерный пример: обобщение математических формул на многомерные пространства впоследствии нашли конкретное применение, например, в экономике, где координаты n-мерных векторов представляют собой совокупность предметов, действий и т. д.

Закономерности, описываемые математическими соотношениями, удивительно точно отражают явления и процессы природы и взаимоотношения в обществе. Гармония математических моделей, их сходство для разных процессов и явлений подтверждает философский постулат о единстве мира. Г. Галилей (Galileo Galilei (1564–1642) — итальянский физик, механик, астроном, философ и математик) более трех столетий назад отметил это. В назидание молодежи он писал: «Перед вами открытая книга (имелась в виду природа). Прочесть эту книгу сможет тот, кто изучит язык, на котором она написана. А написана она на языке математики».

Как отмечено выше, математика развивается и самостоятельно, следуя потребностям людей в их созидательной деятельности. Соответственно, со временем в процессе познания мира совершенствуются и математические модели. Характерный пример. Ньютон (Sir Isaak Newton (1642–1727) — ан-

глийский физик, механик, математик и астроном), установив в свое время связь между ускорением, силой и массой тела (второй закон Ньютона), опирался на закономерности, которые можно было выявить на уровне познания мира XVII века.

Но пришло время Эйнштейна (Albert Einstein (1879–1955) — немецкий ученый, основатель современной теоретической физики) и закономерности эти подверглись серьезной ревизии. С. Маршак (Самуил Яковлевич (1887–1964) — русский советский поэт, переводчик и драматург) по этому поводу съязвил:

Был этот мир
сплошною тьмой окутан.
«Да будет свет»! —
и вот явился Ньютон.
Но сатана не долго
ждал реванша:
пришел Эйнштейн
и стало все, как раньше.

Отметим, что кроме математических создаются другие модели, имитирующие состояние и поведение реальных объектов. Достаточно сослаться, например, на произведения искусства. Картины художников, являющиеся моделями реального мира, живописно передают фиксированные картины различных явлений. Есть попытки художников передать движение. Например, Коровин (Константин Алексеевич (1861–1939) — русский художник-импрессионист) в картине «Бульвар Капуцинок» путем «размазывания» фигур людей в направлении движения добился такого зрительного ощущения.

Наверное, глубина проникновения человека в неосязаемые им явления (микромир, космос) толкнула художников в мир абстракционизма, а глубинные переживания с видениями непонятных явлений — к сюрреализму. А разве музыкальные произведения не являются абстрактными образами ощущений человека? А словесные модели? А сами слова? Разве они не позволяют восстановить в памяти реальные образы?

Модель — *это материальный или мысленно представляемый объект, который в процессе изучения предмета,*

явления или процесса замещает объект-оригинал, сохраняя те его характерные особенности, которые важны для исследования.

Человек живет в мире моделей. И потому закономерностям их создания необходимо учиться.

Предлагаемое пособие ставит целью на ограниченном числе примеров продемонстрировать читателям некоторые подходы к построению математических моделей, а также некоторых методов реализации решений этих моделей (некоторые численные методы) и анализа результатов.

Приобрести и развить навыки построения и анализа математических моделей можно только постоянным обучением с размышлениями. Еще в древние времена великий китайский мыслитель Конфуций (Кун-цзы (551–479 до н. э.) — древнекитайский мыслитель) отмечал: «Учение без размышления бесполезно, но и размышление без учения опасно».

Пособие не претендует на полноту. В нем рассматриваются лишь некоторые из существующих подходов к построению моделей. Количество созданных человеком конкретных математических моделей необозримо и в каждой модели имеются характерные только для нее особенности. Практически во всех опубликованных книгах по математическому моделированию принципы их создания обсуждаются на многочисленных конкретных примерах. Абстрактных, отвлеченных от конкретного смысла, моделей не существует.

Глава 1

МОДЕЛИРОВАНИЕ ОБЪЕКТОВ

§ 1.1. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Любые два объекта O_1 и O_2 всегда в чем-то сходны, а в чем-то различны. При переходе от O_1 к O_2 может иметь место, например, существенное различие и несущественное сходство, или наоборот, существенное сходство и несущественные различия. Эти два крайние случая представляют особый интерес.

В случае существенного сходства возможно изучение большого числа важнейших свойств O_1 с помощью O_2. В случае существенного различия имеет место бóльшая простота изучения свойств объекта O_1 (очень немногих), совпадающих со свойствами объекта O_2. Простота обусловлена тем, что совпадающие свойства могут рассматриваться обособленно от других.

Замещение объекта O_1 объектом O_2 для изучения или фиксации важнейших свойств объекта O_1 с помощью объекта O_2 называется *моделированием* объекта O_1 объектом O_2. Замещаемый (моделируемый) объект O_1 называется *оригиналом* или *натурой*, а замещающий (моделирующий) объект O_2 — *моделью*.

В общем случае процесс моделирования состоит из нескольких этапов.

1. Формулировка цели исследования, постановка задачи, определение свойств оригинала, подлежащих исследованию.

2. Констатация затруднительности или невозможности исследования оригинала в натуре.

3. Формулировка гипотез и предположений, облегчающих задачу построения модели (например, сплошность среды в задачах прочности, гидро- и аэродинамики, ее однородность, идеализация геометрических характеристик объектов, их физических свойств, внешних воздействий и т. д.).

4. Выбор модели, достаточно хорошо фиксирующей (проявляющей) существенные свойства оригинала и легко подлежащей исследованию, отбрасывание факторов и эффектов, несущественных для исследователя.

5. Формирование особенностей и закономерностей, которым подчиняется объект и которые интересуют исследователя.

6. Выбор принципов построения модели (вариационных, аналогий, сходственность с оригиналом, согласованность с законами природы, полнота и точность их отражения в модели, адекватность модели, . . .).

7. Разработка алгоритмов и программ реализации модели. Исследование модели в соответствии с поставленной задачей.

8. Перенос результатов исследований на оригинал.

9. Проверка и анализ результатов.

Основными задачами моделирования являются выбор модели (п. 4), ее исследование (п. 7) и перенос результатов исследований модели на оригинал (п. 8). Эти задачи решаются с помощью достаточно общих методов.

Перечисленные этапы создания и анализа математических моделей не отличаются конкретностью, а отражают лишь последовательность действий при их построении. Успех математического моделирования зависит во многом (и даже в большей части) от квалификации исследователя, глубины познаний в соответствующей конкретной области его интересов.

§ 1.2. КЛАССИФИКАЦИЯ МОДЕЛЕЙ

Важнейшими признаками моделей, которые могут быть положены в основу их классификации, являются следующие.

1. Закон функционирования и характерные особенности выражения свойств и отношений оригинала.

2. Основания для преобразования свойств и отношений модели в свойства и отношения оригинала.

По первому признаку модели можно разделить на:

а) логические;

б) материальные.

Логические модели делятся на:

— образные;

— знаковые;

— образно-знаковые.

Материальные модели делятся на:

— функциональные;

— геометрические;

— функционально-геометрические.

Логические модели функционируют по законам логики, а материальные — по объективным законам природы.

Образные (иконические) модели выражают свойства оригинала с помощью наглядных (чувственных) образов, имеющих прообраз среди элементов оригинала или объектов материального мира.

Знаковые (символические) модели выражают свойства оригинала с помощью условных знаков или символов (уравнения, формулы и т. п.).

Образно-знаковые модели обладают признаками образных и знаковых моделей (схемы, графики, графы и т. п.).

Функциональные модели отражают только функциональные свойства оригинала.

Геометрические модели отражают только пространственные свойства оригинала.

Функционально-геометрические модели отражают и функциональные и пространственные свойства оригинала.

В зависимости от физической однородности с оригиналом функциональные и функционально-геометрические модели можно разделить на физические и формальные.

По второму признаку различают:

а) условные модели;

б) аналогичные модели;

в) математические модели.

Условные модели выражают свойства и отношения оригинала на основании принятого условия или соглашения. Сходство с оригиналом при этом может совершенно отсутствовать.

К условным моделям относят все знаковые и образно-знаковые модели.

Аналогичные модели обладают сходством с оригиналом, достаточным для перехода к оригиналу на основании умозаключения по аналогии.

Математические модели обеспечивают переход к оригиналу, фиксацию и исследование его свойств с помощью математических методов. Эти модели делят на:

— расчетные;

— соответственные.

Расчетные модели выражают свойства оригинала с помощью математических представлений.

В соответственных моделях переменные величины связаны с соответствующими величинами оригинала определенными математическими зависимостями.

Математические модели имеют признаки условных моделей и могут обладать признаками аналогичных.

Важнейшей разновидностью соответственных моделей являются подобные модели, в которых переменные величины пропорциональны соответствующим переменным оригинала. Они могут быть как логическими, так и материальными.

Подобные материальные модели разделяются на:

— аналоговые (непрерывные);

— цифровые (дискретные);

— аналогово-цифровые.

Одним из признаков математических моделей является их определенность. По этому признаку модели делятся на:

— детерминированные — модели, в которых имеет место взаимно-однозначное соответствие между переменными, характеризующими рассматриваемые объекты, процессы и явления;

— стохастические — модели, в которых связь между переменными носит случайный характер.

Перечислим основные требования, предъявляемые к математическим моделям.

1. Адекватность — способность правдоподобно с заданной точностью отражать свойства моделируемых объектов.

2. Точность — количественная характеристика, определяемая степенью совпадения числовых характеристик модели и реального объекта.

3. Экономичность — определяется затратами времени на создание модели и на реализацию решения.

Конечно, к требованиям можно отнести другие показатели, носящие зачастую субъективный характер, такие как обозримость, удобство использования, наглядность, универсальность и др.

§ 1.3. УСЛОВНОЕ МОДЕЛИРОВАНИЕ

Условное моделирование – это замещение оригинала условной моделью, представляющей оригинал только благодаря определенной договоренности о смысле, приписываемом этой модели. К условным моделям относятся все знаковые модели.

Знак (символ) – это искусственный образ, чисто условно обозначающий вполне определенный объект и, как правило, не имеющий с этим объектом сходства.

Строгих правил создания знаковых моделей нет, т. к. формирование таких моделей всегда имеет поисковый (эвристический) характер.

Основными требованиями при создании знаковых моделей являются:

— необходимость (невозможность использования уже имеющихся символов);

— простота (простые, при прочих равных условиях, символы предпочтительнее);

— наглядность (наличие общих черт с оригиналом);

— индивидуальность (отличие от других символов);

— однозначность (недопустимость обозначения одним символом различных объектов);

— единообразие (использование одинаковых символов при моделировании однородных объектов);

— определенность (указание четких условий использования данного символа);

— традиционность (учет имеющихся традиций в этой области).

Условными являются и образно-знаковые модели. Они отличаются наглядностью и могут обладать определенным сходством с оригиналом. Примерами могут служить структурные схемы, направленные графы.

К знаковым и образно-знаковым моделям относят все математические формы выражения количественных отношений между постоянными и переменными величинами. Поэтому условное моделирование входит составной частью в исследования, применяющие математические методы решения задач.

Практически приходится иметь дело с математическими описаниями материальных объектов, являющимися условными логическими моделями количественных отношений между размерами или числовыми значениями физических величин.

§ 1.4. РАЗМЕРЫ И ЧИСЛОВЫЕ ЗНАЧЕНИЯ

В общем случае физическая величина X определяет некоторое свойство материальных объектов, допускающее количественное выражение.

Количественное выражение физической величины X (в конкретном материальном объекте x) называют *размером* этой физической величины. Например, L — длина вообще, l — длина конкретного объекта.

Для определения размера x физической величины X требуется его измерить, т. е. сравнить с размером $\{x\}$ той же физической величины другого объекта, принятого за единицу. В результате измерения устанавливается *числовое значение* $\overset{\circ}{x}$ величины x размера $\{x\}$. Так что $x = \overset{\circ}{x}\,\{x\}$.

Символы x, $\overset{\circ}{x}$ и $\{x\}$ моделируют размер, числовое значение и единицу физической величины X. Таким образом, размер выражается через его числовое значение и единицу. Например, в записи «длина стола = 1,5 м» 1,5 — числовое значение, м — единица измерения.

Размер x не зависит от единицы ее измерения $\{x\}$. От $\{x\}$ зависит только числовое значение $\overset{\circ}{x}$ размера x. Размер — ха-

рактеристика объекта, инвариантная по отношению к выбору единицы измерения.

Каждый материальный объект обладает несколькими свойствами, допускающими количественное выражение (линейные размеры, масса, температура, ...). Каждое свойство характеризуется определенной физической величиной своего размера.

Между различными свойствами объективно существуют конкретные связи. Они обусловливают определенные соотношения между размерами физических величин, которые можно выразить в виде формул. Поэтому можно выделить произвольно единицы некоторых физических величин и через них выразить единицы других физических величин.

§ 1.5. ОСНОВНЫЕ И ПРОИЗВОДНЫЕ ЕДИНИЦЫ ИЗМЕРЕНИЯ

Физические величины, размеры единиц которых выбираются произвольно, называются *основными*. Единицы измерения всех остальных физических величин выражаются через основные и называются *производными*.

Совокупность основных и производных единиц составляют *систему единиц измерения* (СИ). В СИ основные величины: длина, масса, время, сила электрического тока, сила света, количество вещества, температура. Основные единицы, соответственно: метр (м), килограмм (кг), секунда (с), ампер (А), кандела (кд), моль (моль), кельвин (К).

Пусть физические величины X_1, X_2, ... выбраны как независимые основные величины. Для некоторой физической величины Y требуется установить производную единицу $\{y\}$.

Для решения задачи выбирается объект, в котором размеры величин Y, X_1, X_2, ... связаны уравнением:

$$y = kF(x_1, x_2, \ldots), \tag{1.1}$$

где k — коэффициент пропорциональности.

Иначе, т. к. $y = \overset{\circ}{y}\ \{y\}$, $x_1 = \overset{\circ}{x}_1\ \{x_1\}$, $x_2 = \overset{\circ}{x}_2\ \{x_2\}$, ..., можно записать

$$\overset{\circ}{y}\ \{y\} = kF(\overset{\circ}{x}_1\ \{x_1\}, \overset{\circ}{x}_2\ \{x_2\}, \ldots). \tag{1.2}$$

Положим в этом соотношении, что все числовые значения равны единице (в математическом смысле): $\overset{\circ}{y}= 1$, $\overset{\circ}{x}_1= 1$, $\overset{\circ}{x}_2= 1, \ldots$.

Тогда производную единицу измерения можно выразить через основные:

$$\{y\} = kF(\{x_1\}, \{x_2\}, \ldots). \tag{1.3}$$

Для обеспечения «согласованности» (идентичности) выражений (1.3) и (1.1), связывающих размеры, должна выполняться зависимость:

$$\overset{\circ}{y}= kF(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \ldots). \tag{1.4}$$

Подставив (1.3) и (1.4) в (1.2), получим:

$$kF(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \ldots)kF(\{x_1\}, \{x_2\}, \ldots) = kF(\overset{\circ}{x}_1\ \{x_1\}, \overset{\circ}{x}_2\ \{x_2\}, \ldots).$$

Это уравнение равносильно системе

$$\begin{aligned} &k^2 = k \\ &F(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \ldots)F(\{x_1\}, \{x_2\}, \ldots) = F(\overset{\circ}{x}_1\ \{x_1\}, \overset{\circ}{x}_2\ \{x_2\}, \ldots). \end{aligned} \tag{1.5}$$

Первое из уравнений этой системы имеет два вещественных корня: 0 и 1. Первый корень должен быть отброшен в данном случае. Значит $k = 1$.

Второе уравнение системы позволяет сделать некоторые заключения о функции F. По смыслу соотношения (1.1), функция F должна быть непрерывной. Единственной непрерывной функцией, удовлетворяющей второму уравнению системы (1.5), является функция вида

$$F(x_1, x_2, \ldots) = x_1^{\alpha_1} \cdot x_2^{\alpha_2}, \tag{1.6}$$

где α_1, α_2 — некоторые числа.

Такая функция называется *степенным комплексом*. Таким образом, *для установления производных единиц измерения пригодны только физические формулы в виде степенных комплексов с постоянным коэффициентом, равным 1.*

Степенной комплекс, выбранный для установления производной единицы измерения, называется *определяющим уравнением* для этой единицы.

Пример. Требуется установить размерность силы. Спрашивается, как ввести силу? В физике используют формулы для сил (F — инерции; F_t — тяготения; F_k — Кулона):

$$F = ma; \qquad F_t = k\frac{m_1 m_2}{r^2}; \qquad F_k = \mu\frac{q_1 q_2}{r^2}.$$

Все эти формулы представляют собой степенные комплексы, но коэффициент пропорциональности равен единице только в первой из них. Что касается второй и третьей формул, то в них коэффициенты пропорциональности (k и μ) — размерные. Следовательно, единицу измерения силы необходимо установить по первой формуле.

Если $\{m\} = M$, $\{a\} = L/T^2$, то $\{F\} = ML/T^2$. В частности, для системы СИ, где $M =$ кг, $L =$ м, $T =$ с: $\{F\} =$ кг·м/c^2.

§ 1.6. РАЗМЕРНЫЕ И БЕЗРАЗМЕРНЫЕ ВЕЛИЧИНЫ

Основные и производные единицы обозначается либо символом $\{x\}$ соответствующей величины X, либо специальным символом, представляющим ее сокращенное название.

Производные единицы при их конкретизации обозначаются либо специальным символом, представляющим ее *сокращенное название* (в равенстве $\{F\}{=}H$ $\{F\}$ — размерность силы; H — символ (сокращение от Ньютон)), либо через *основные единицы* (в равенстве $\{F\}{=}MLT^{-2}$ $\{F\}$ — размерность; M, L, T — символы).

Формулы размерности производных единиц не дают представления о конкретных размерах единиц, входящих в них. Так, в выражении $\{F\} = MLT^{-2}$ масса может быть представлена и в килограммах (кг), и в граммах (г); длина в метрах (м), километрах (км), сантиметрах (см); время в секундах (с), часах (ч) и т. д.

Размерность — *это символическое выражение единицы величины через основные единицы, показывающее на соотношение между их размерами без указания самих размеров.*

Понятие размерности относится к единице величины, однако его распространяют и на саму величину. Это приводит к понятиям размерных и безразмерных величин.

Величина Y называется *безразмерной,* если ее размерность равна 1. Для безразмерной величины

$$\{y\} = \{x_1\}^0\{x_2\}^0 ... = 1.$$

Все, что не безразмерно — размерно. Формула размерности любой физической величины однозначно определяется выбором основных единиц измерения и определяющим уравнением. Одна и та же формула размерности может соответствовать различным физическим величинам, например, {работа} = = {энергия} = ML^2T^{-2}.

§ 1.7. АНАЛОГИЯ. АНАЛОГИЧНОЕ МОДЕЛИРОВАНИЕ

Аналогия — *это сходство различных объектов по некоторым признакам.*

Если не все свойства оригинала и модели совпадают, говорят о *частичной аналогии* двух объектов.

Объекты, сходные по соответствующим признакам, называют *ана́логами*, а эти их признаки — *сходственными*.

Основное значение аналогии — возможность переноса свойств с одного объекта на другой на основании умозаключения по аналогии.

Умозаключение по аналогии основано на предположении существования тождественности в различном и выполняется следующим образом.

1. Установлено, что объект O_1 (оригинал) обладает основными свойствами C_0, C_1, C_2, ..., C_N и дополнительными свойствами C'_1, C'_2, ..., C'_{n_1}.

2. Установлено, что объект O_2 (модель) обладает основными свойствами C_1, C_2, ..., C_N и дополнительными C''_1, C''_2, ..., C''_{n_2}.

3. Возможно, что объект O_2 обладает также свойством C_0, как и объект O_1, но это не установлено.

Если среди свойств серии C''_k $(k = \overline{1, n_2})$ есть хотя бы одно свойство, несовместное со свойством C_0 (противоположно ему), то сходство объектов по свойствам C_1, C_2, ..., C_N не имеет никакого значения для построения аналогии — аналогию провести нельзя.

Умозаключение по аналогии имеет гипотетический характер. Оно может привести и к истинному, и к ложному выводу. Поэтому суждения, полученные по аналогии, как правило, нуждаются в специальной проверке.

Вероятность правильности этих суждений тем больше, чем:

— сильнее связь между свойствами C_1, C_2, ..., C_N объектов O_1 и O_2;

— слабее связи между C_k $(k = \overline{1, N})$, с одной стороны, и C'_i $(i = \overline{1, n_1})$ и C''_j $(j = \overline{1, n_2})$, с другой стороны (основных с дополнительными);

— больше N (число основных свойств) и меньше n_1 и n_2 (число дополнительных свойств).

Умозаключение по аналогии имеет и доказательную силу, если общие свойства объектов C_1, C_2, ..., C_N обусловливают свойство C_0. Такое обстоятельство является основой аналогичного моделирования.

Существенное значение аналогии заключается в возможности использовать ее для строгих выводов и доказательств. Она позволяет перейти к важному понятию *подобия*, при котором возможен строгий пересчет данных модели O_2 в данные оригинала O_1. Такой переход обеспечивает важнейший вид аналогии — *математическую аналогию*, при которой сходство объектов следует из сходства их математических описаний.

Наиболее полная математическая аналогия имеет место, если объекты описываются сходственными функциями и уравнениями.

Сходственные функции различаются только аргументами и ненулевыми постоянными.

Сходственные переменные — переменные, входящие под знак сходственных функций одинаковым образом.

Аналогично определяются и *сходственные постоянные*.

Сходственные уравнения получаются приравниванием друг другу сходственных функций.

Например, в двух функциональных зависимостях $y = kx_1x_2^2$ и $z = \frac{1}{2}mv^2$ наблюдается следующее соответствие:

$$y \sim z; \quad k \sim \frac{1}{2}; \quad x_1 \sim m; \quad x_2 \sim v.$$

Аналогичное моделирование – *это замещение оригинала аналогичной моделью, т. е. моделью, обладающей сходством с оригиналом, достаточным для переноса ее свойств и отношений на свойства и отношения оригинала на основании умозаключения по аналогии.*

Аналогичное моделирование обычно используют при сравнительно слабой изученности оригинала, когда имеющиеся сведения о его свойствах имеют лишь качественный характер.

Пример. Регулятор системы автоматического управления (САУ), состоящий из управляющего объекта (УО) и регулятора (Р), автоматически подбирает переменную y, обращающую функцию $F(x; y)$ в нуль (рис. 1.1,а).

Так реализуется решение уравнения $F(x; y) = 0$ относительно y: $y = y(x)$. Если связь $y = y(x)$ однозначна, то получается единственное положение равновесия: устойчивое, или

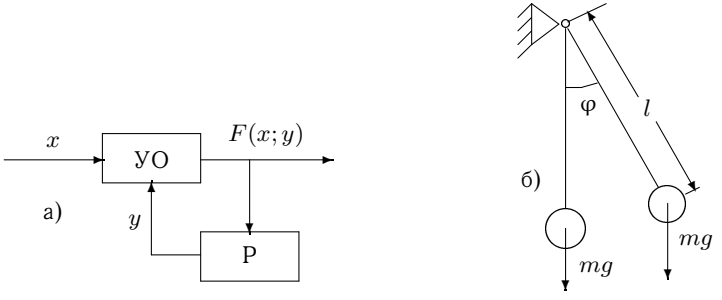


Рис. 1.1
Модели: а) – регулятор САУ; б) – маятник

неустойчивое. Если решение уравнения не единственное, то неединственным будет и положение равновесия.

Для иллюстрации возможных положений равновесия (устойчивости/неустойчивости) обратимся к модели физического маятника, показанного на рис. 1.1,б.

Из механики известно, что равновесие маятника определяется уравнением $mgl\sin\varphi = 0$. Корнями этого уравнения являются значения $\varphi = k\pi$ $(k \in Z)$, из которых устойчивыми будут корни уравнения с четными значениями k (маятник находится в крайнем нижнем положении); неустойчивыми — с нечетными значениями k (крайнее верхнее положение маятника).

Аналогия модели и оригинала приводит к умозаключению, что в нелинейной системе автоматического управления (САУ) с несколькими положениями равновесия устойчивые и неустойчивые положения чередуются. Между двумя соседствующими устойчивыми положениями равновесия существует одно неустойчивое положение и наоборот. В реальных же моделях справедливо лишь второе заключение, то есть что между двумя неустойчивыми положениями существует одно устойчивое.

Неполная аналогия модели и оригинала объясняется тем, что у реального маятника устойчивые положения могут быть только колебательными, а неустойчивые непериодическими (апериодическими).

Строя модели, необходимо обращать внимание на их неполную адекватность реальным системам, а результаты модельных исследований проверять, сопоставляя их с результатами натурных экспериментов.

§ 1.8. РЕЗЮМЕ

Для исследования реальных объектов, как правило, прибегают к предварительному изучению их свойств на моделях — объектах, поведение или характер которых по основным признакам идентичны оригиналу. Важнейшим видом моделей, широко используемым для описания реальных объектов, являются математические модели.

Богатый опыт и установленные закономерности создания математических моделей, разработанные эффективные методы реализации их решения, в том числе с использованием вычислительной техники, делают математическое моделирование неотъемлемой частью прогрессирующего развития человеческой цивилизации.

Большую роль в создании адекватных математических моделей играет корректное введение размерностей различных величин, выделение из всей их совокупности тех величин, через размерности которых могут быть выражены размерности других величин.

Соотношения между размерностями перекликаются с соотношениями между самими величинами, представляющими (описывающими) реальные процессы.

§ 1.9. ВОПРОСЫ

1. Что такое моделирование объектов, систем и процессов?
2. Что такое оригинал? Модель?
3. Что собой представляют логические модели? Материальные модели?
4. К какому типу моделей относятся математические модели?
5. Что такое условное моделирование?
6. Какие требования предъявляются к знаковым моделям?
7. Что такое размер (размерность) величины? Числовое значение величины?
8. Какие величины относят к основным? К производным?
9. Перечислите основные величины системы СИ.
10. Что представляет собой степенной комплекс?
11. Какое уравнение называют «определяющим» для установления производных единиц измерения?
12. Какие величины называют размерными? Безразмерными?
13. Что такое аналогия? Аналогичное моделирование? По каким признакам можно сделать умозаключение по аналогии?

Глава 2

МАТЕМАТИЧЕСКОЕ И ФИЗИЧЕСКОЕ МОДЕЛИРОВАНИЕ

§ 2.1. ПОДОБИЕ

Особое значение имеют *подобные модели,* обеспечивающие перенос данных с модели на оригинал на основании их подобия.

Подобие — *это полная математическая аналогия при наличии пропорциональности между соответствующими величинами, сохраняющаяся при всех их возможных значениях.*

Математическое описание объекта может иметь различные формы. В простейшем случае — это явная функция, однозначно выражающая зависимую переменную y через n ее независимых переменных x_i $(i = \overline{1,n})$:

$$y = y(x_1, x_2, \ldots, x_n) = y(x_i).$$

Другое математическое представление зависимости между переменными

$$F(y; x_1, x_2, \ldots, x_n) = F(y; x_i) = 0$$

использует неявное задание функции y через аргументы x_i $(i = \overline{1,n})$.

Две приведенные записи являются алгебраическими зависимостями.

Если в функциональную зависимость включаются производные от функции и ее аргументов, то математическая модель представляется в виде различного рода дифференциальных уравнений. В более общем случае вместо производных функций и аргументов могут стоять более общие операторы.

Два объекта подобны, если:

1) они имеют сходственные математические описания ($i = \overline{1,n}$):

$$F(y_1, x_{1i}) = 0, \qquad F(y_2, x_{2i}) = 0;$$

2) сходственные переменные, содержащиеся в математических описаниях подобных моделей, связаны постоянными коэффициентами пропорциональности, которые называются *масштабами* или *константами подобия*:

$$m_y = \frac{y_1}{y_2}, \quad m_{xi} = \frac{x_{1i}}{x_{2i}}.$$

Если можно ввести масштабы, то сходственные уравнения и функции, а также содержащиеся в них сходственные переменные, называются *подобными*.

Частными случаями подобия являются *геометрическое, физическое* и *временное*.

§ 2.2. ПОДОБИЕ СТЕПЕННЫХ КОМПЛЕКСОВ

Степенные комплексы (1.6) можно подразделить на *простые* и *составные*. Простые комплексы, в отличие от составных, нельзя образовать из других степенных комплексов.

Рассмотрим функцию y, представляющую собой степенной комплекс, а именно, произведение некоторой постоянной a на степенной комплекс ее аргументов x_i ($i = \overline{1,n}$):

$$y = F(x_1, x_2, \ldots, x_n) = a x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} = a \prod_{i=1}^{n} x_i^{\alpha_i}. \qquad (2.1)$$

Безразмерные степенные комплексы, определяющие подобие различных объектов, называют *критериями подобия*.

Сформулируем без доказательства свойства степенных комплексов.

1. Число простых степенных комплексов, образованных из некоторых независимых величин, не может превзойти числа этих величин.

Если в степенной комплекс, стоящий в правой части равенства (2.1), входит n независимых величин x_1, x_2, ..., x_n, то число l простых степенных комплексов

$$k_p = \prod_{j=1}^{m} x_j^{\beta_j} \ \ (p = \overline{1,l}: \ \ m \le n), \tag{2.2}$$

согласно сформулированному свойству, не может быть больше n $(l \le n)$.

2. Любую функцию некоторых величин x_i $(i = \overline{1,n})$ можно представить в виде функции степенных комплексов (2.2) этих величин:

$$y = F(x_1, x_2, \ldots, x_n) = \prod_{p=1}^{l} k_p^{\gamma_p}. \tag{2.3}$$

3. Любую безразмерную функцию размерных величин можно представить в виде функции безразмерных степенных комплексов, образованных из этих величин.

4. Любую размерную функцию размерных величин можно представить в виде произведения размерного степенного комплекса, составленного из этих величин, и безразмерной функции этих величин.

Докажем справедливость последнего свойства.

Если функция $y = F(x_1, \ldots, x_n)$ размерная, ее размерность определяется размерностью величин, от которых она зависит. Поэтому существует размерный степенной комплекс $k = x_1^{\beta_1} x_2^{\beta_2} \cdots x_n^{\beta_n}$, составленный из этих величин, с размерностью, совпадающей с размерностью этой функции.

Пусть размерная функция y является размерным степенным комплексом с размерностью, совпадающей с размерностью F : $\{k\} = \{F\}$.

Умножим и разделим y на k так, что

$y = k\Phi(x_1, \ldots, x_n)$ и $\Phi(x_1, \ldots, x_n) = F(x_1, \ldots, x_n)/k$.

Если

$\{k\} = \{F\}$, то $\{\Phi\} = 1$.

Например, для функции $y = x_1^2 x_2$ $\Phi = \dfrac{y}{x_1^2 x_2}$. То есть $k = x_1^{-2} x_2^{-1}$.

Рассмотрим две функции y_1 и y_2, представленных в виде степенных комплексов n переменных x_i, которые будем считать *основными соотношениями*:

$$y_1 = a_1 \prod_{i=1}^{n} x_{1i}^{\alpha_i}; \tag{2.4}$$

$$y_2 = a_2 \prod_{i=1}^{n} x_{2i}^{\alpha_i}. \tag{2.5}$$

Для того, чтобы считать объекты, описываемые уравнениями (2.4) и (2.5), подобными, введем масштабы (*дополнительные соотношения*):

$$m_y = y_1/y_2; \qquad m_{xi} = x_{1i}/x_{2i}. \tag{2.6}$$

Дополнительные соотношения должны быть такими, чтобы вся система была непротиворечивой. Требование *непротиворечивости* является необходимым условием подобия. При непротиворечивости системы можно:

1) либо задавшись значениями x_{1i} определить y_1 двумя путями;

2) либо задавшись значениями x_{2i} двумя путями определить y_2.

В первом случае (рис. 2.1):

— прямым путем из (2.4) находим y_1;

— окольным путем из (2.6) находим x_{2i};

— зная x_{2i} из (2.5) вычисляем y_2;

— с помощью (2.6) находим y_1 через y_2.

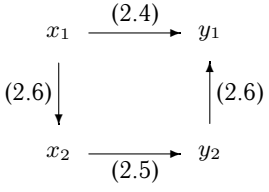


Рис. 2.1
Схема установления непротиворечивости

Система непротиворечива, если обоими путями получается одно и то же значение y_1.

Так, для прямого пути

$$y_1' = a_1 \prod_{i=1}^{N} x_{1i}^{\alpha_i}.$$

Тогда как по окольному пути

$$x_{2i} = \frac{x_{1i}}{m_{xi}},$$

что дает

$$y_2 = a_2 \prod_{i=1}^{N} x_{2i}^{\alpha_i} = a_2 \prod_{i=1}^{N} \left(\frac{x_{1i}}{m_{xi}} \right)^{\alpha_i}.$$

Далее, с учетом $y_1'' = m_y y_2$ сравниваются два результата и проверяется выполнение равенства $y_1' = y_1''$.

Для получения условия непротиворечивости продолжим преобразования, начиная с подстановки в условие, вытекающее из (2.6), соотношения (2.5):

$$y_1 = m_y a_2 \prod_{i=1}^{N} \left(\frac{x_{1i}}{m_{xi}} \right)^{\alpha_i} = m_y a_2 \prod_{i=1}^{N} \left(\frac{1}{m_{xi}} \right)^{\alpha_i} \prod_{i=1}^{N} (x_{1i})^{\alpha_i} =$$

$$= \frac{m_y a_2}{a_1} \prod_{i=1}^{N} \left(\frac{1}{m_{xi}} \right)^{\alpha_i} a_1 \prod_{i=1}^{N} (x_{1i})^{\alpha_i} = \frac{m_y a_2}{a_1} \prod_{i=1}^{N} \left(\frac{1}{m_{xi}} \right)^{\alpha_i} y_1.$$

Сравнение коэффициентов при y_1 в правой и левой частях выписанного равенства позволяет получить *условие непротиворечивости*, выраженное через масштабы:

$$1 = \frac{m_y a_2}{a_1} \prod_{i=1}^{N} \left(\frac{1}{m_{xi}} \right)^{\alpha_i}. \tag{2.7}$$

Аналогично (или возведением обеих частей равенства (2.7) в степень –1) можно получить условие непротиворечивости во

втором случае. Оно будет таким же, только перевернутым. То есть

$$\frac{a_1}{m_y a_2}\prod_{i=1}^{N}(m_{xi})^{\alpha_i}=1. \tag{2.8}$$

Уравнения (2.7) и (2.8) называются *масштабными*, а изложенный метод доказательства непротиворечивости моделей — *методом подстановки*.

Таким образом, ввиду непротиворечивости, можно, благодаря масштабам, преобразовать (2.4) в (2.5) и наоборот. При этом устанавливается условие непротиворечивости в виде безразмерного степенного комплекса, численное значение которого равно единице (2.7).

Так как m_y и m_{xi} связаны уравнением (2.7) (условием непротиворечивости), то этот набор масштабов не независимый. Некоторые из величин мы можем выбирать произвольно, а другие не можем.

В теории моделирования безразмерные степенные комплексы принято обозначать буквами π или Π.

К безразмерному виду, аналогичному условию непротиворечивости, можно привести уравнения (2.4) и (2.5):

$$\Pi_1=\frac{1}{y_1}a_1\prod_{i=1}^{N}x_{1i}^{\alpha_i}=1;\quad \Pi_2=\frac{1}{y_2}a_2\prod_{i=1}^{N}x_{2i}^{\alpha_i}=1.$$

Условие непротиворечивости (2.8) вытекает из соотношения $\Pi_1/\Pi_2=1$:

$$\frac{\Pi_1}{\Pi_2}=\frac{a_1\prod_{i=1}^{N}x_{1i}^{\alpha_i}}{y_1}\Bigg/\frac{a_2\prod_{i=1}^{N}x_{2i}^{\alpha_i}}{y_2}=$$

$$=\frac{a_1\prod_{i=1}^{N}(x_{1i}/x_{2i})^{\alpha_i}}{a_2(y_1/y_2)}=\frac{a_1\prod_{i=1}^{N}m_{xi}^{\alpha_i}}{a_2 m_y}=1.$$

Доказательство непротиворечивости математических моделей с использованием отношений безразмерных степенных комплексов называется *способом критериев*.

Вывод: для подобия степенных комплексов необходимо их сходство, связь сходственных величин через масштабы, непротиворечивость выражений сходственных степенных комплексов и уравнений, выражающих масштабы сходственных переменных.

§ 2.3. ПОДОБНОЕ МОДЕЛИРОВАНИЕ

Подобное моделирование — *это частный случай математического моделирования, представляющий собой замещение оригинала подобной ему моделью.*

Подобная модель — *это модель, свойства, значения и параметры переменных которой пропорциональны свойствам, значениям и параметрам переменных оригинала.*

Для подобного моделирования оригинала необходимо располагать рядом исходных данных. В общем случае к ним относятся:

1) математическое описание оригинала;

2) пределы изменения фигурирующих в математическом описании переменных;

3) условия однозначности решения соответствующих уравнений;

4) задание функциональных зависимостей известными функциями.

Процесс подобного моделирования в общем случае разбивается на ряд этапов.

1. Из числа существующих выбирают или создают специальный объект — модель с математическим описанием, сходственным с описанием оригинала.

2. Определяют критерии подобия для оригинала и модели (Π и Π_M).

3. Составляют в общем виде масштабные уравнения ($\Pi/\Pi_M = 1$).

4. Вводят масштабы сходственных переменных и придают масштабным уравнениям окончательный вид $\left(m_y^{\alpha_y} m_{x_i}^{\beta_i} m_{tj}^{\varepsilon_{jr}} = c\right)$.

5. Анализируют системы масштабных уравнений (зависимые уравнения исключаются; при установлении противоречивости делается вывод о невозможности подобия).

6. Выбирают конкретные числовые значения масштабов с учетом реальных предельных значений сходственных переменных.

7. Устанавливают условия однозначности модели, подобные условиям однозначности оригинала.

8. Рассчитывают функциональные зависимости модели, подобные функциональным зависимостям оригинала.

§ 2.4. ФИЗИЧЕСКОЕ МОДЕЛИРОВАНИЕ. МОДЕЛЬНЫЙ ЭКСПЕРИМЕНТ

В зависимости от характера модели, подвергаемой исследованию, модельный эксперимент может быть *реальным* или *мыслимым (виртуальным)*; другими словами, *материальным* или *логическим*.

При постановке задачи мысленный эксперимент всегда предшествует реальному. В общем случае изучение материальных объектов заключается в установлении необходимых опытных фактов и составлении уравнений, описывающих объекты. В ходе выполнения этой работы встречаются три основные трудности:

1) трудность, или даже невозможность, опытного исследования натуры;

2) сложность самой натуры и следующая отсюда трудность составления ее математической модели;

3) сложность математического описания натуры и трудность решения соответствующих уравнений.

Сложности математического описания модели и исследования на ее базе реального объекта можно избежать, прибегая к физическому моделированию.

Физическая модель — это материальная функциональная модель, которая может быть аналогичной или подобной оригиналу. Последнее предпочтительней, т. к. дает возможность строгого переноса данных экспериментального исследования

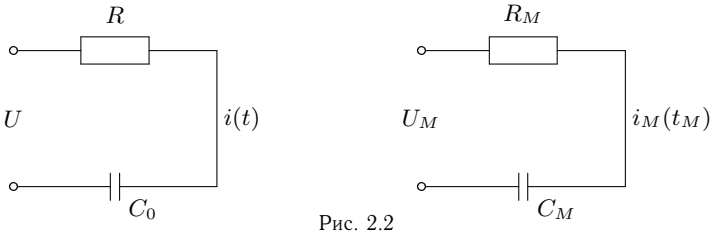


Рис. 2.2
Физический оригинал (слева) и модель

модели на натуру. При переносе данных на объект могут возникнуть новые затруднения, вызванные отсутствием математического описания объекта, что делает невозможным установление критериев подобия и масштабов.

Вспоминая условия подобия (раздел § 2.3), можно заключить, что *подобное моделирование возможно только тогда, когда известны математические описания оригинала и модели.* В действительности при определенных условиях, подобное моделирование возможно и при отсутствии математических описаний оригинала и модели.

Исследуем, например, прохождение тока через последовательно соединенные резистор и конденсатор (рис. 2.2). Допустим, что получить уравнения, описывающие оригинал (левый рисунок), мы не умеем (электротехнику не учили), нужных для натурного эксперимента деталей нет.

Но мы можем создать физически однородную с оригиналом модель (правый рисунок), собранную из того, что есть.

Математические описания этих двух различных, но физически однородных объектов, неизвестны, но есть основание предполагать, что объекты сходны.

Таким образом, первое условие подобия (сходственность математических описаний) может быть обеспечено физической однородностью оригинала и модели. В других случаях может понадобиться геометрическое подобие.

Не зная математических описаний оригинала и модели, мы можем перечислить все сходственные величины (протекающие

в них токи i, i_M, напряжения U, U_M, сопротивления R, R_M, емкости C, C_M, моменты времени t, t_M).

Установив физический смысл объекта и модели, можно ввести масштабы m_i, m_U, m_R, m_C, m_t. Таким образом, второе условие подобия (связь сходственных переменных масштабами) достигается установлением полного перечня величин, фигурирующих в математических описаниях оригинала и модели. Интересующая нас зависимость $i = i(t)$ при прочих равных условиях определяется начальным значением тока $i_0 = i(0)$ — это условие выполнения требования однозначности.

Начальные условия для модели должны быть подобны:

$$i_M(0) = \frac{i(m_t \cdot t)}{m_i}.$$

Таким образом, четвертое условие подобия (сходственность условий однозначности решений сходственных уравнений и пропорциональность сходственных переменных) обеспечивается физической однородностью оригинала и модели.

Может потребоваться и геометрическое подобие, заключающееся в данном случае в подобии графиков функций $i(t)$ и $U(t)$. Искомый закон изменения $i = i(t)$ зависит от формы напряжения U.

Форма напряжения модели $U_M(t_M)$ должна быть подобной оригиналу:

$$U_M(t_M) = \frac{U(m_t \cdot t_M)}{m_U}.$$

Таким образом, пятое условие подобия (подобие известных функций) осуществляется, если известны масштабы.

Осталось только третье условие подобия, т. е. выбор масштабов в соответствии с масштабными уравнениями. Есть два способа их получения: подстановка и критерии подобия.

Первый требует знаний уравнений оригинала и модели, следовательно, он не подходит. Остается способ критериев подобия.

Перечень величин, фигурирующих в описании оригинала и модели, анализ их размерности и так называемая π-теорема

(см. следующий раздел) позволяют найти критерии подобия и масштабные уравнения без использования математических моделей.

Задача моделирования сводится к анализу размерности и π-теореме при составлении математических описаний расчетных моделей на основании данных, полученных путем экспериментального исследования натуры.

§ 2.5. ОПРЕДЕЛЯЮЩИЕ ЗАВИСИМОСТИ. π-ТЕОРЕМА

Состояние любого материального объекта в конкретных условиях характеризуют определяющие величины и объективно существующая зависимость, связывающая эти величины.

Определяющие величины x_1, ..., x_N — это различные размерные или безразмерные переменные и размерные постоянные.

Определяющая зависимость выражается в общем случае некоторым конечным уравнением, которое, кроме определяющих величин, может содержать безразмерные постоянные, отличные от единицы. В дальнейшем их не будем рассматривать, чтобы не усложнять зависимости.

При любой системе единиц измерения с q основными единицами η_1, η_2, ..., η_q все определяющие величины разделяются на три группы:

Первая группа содержит k размерных постоянных величин x_1, ..., x_k с независимыми размерностями:

$$\{x_i\} = \prod_{j=1}^{q} \eta_j^{a_{ij}} \quad (i = \overline{1,k}). \tag{2.9}$$

Независимыми размерностями называют такие, которые не могут быть представлены в виде степенных комплексов размерностей других величин. То есть ни одна из вышеприведенных зависимостей (2.9) не может быть представлена в виде степенного комплекса, составленного из других. Следовательно, степенные комплексы, выражающие размерность величин $\{x_1\}$, ..., $\{x_k\}$ — простые.

Число k величин первой группы ограничено числом q основных единиц, то есть $k \leq q$.

Вторая группа содержит μ постоянных величин: $x_{k+1} = X_1$, ..., $x_{k+\mu} = X_\mu$ с зависимыми размерностями ϑ_j ($j = \overline{1, p}$):

$$\{X_i\} = \prod_{j=1}^{p} \vartheta_j^{\beta_{ij}} \quad (i = \overline{1, \mu}). \tag{2.10}$$

Зависимость размерностей означает, что каждая из них может быть представлена в виде степенного комплекса, образованного из размерностей величин первой группы:

$$\{\vartheta_j\} = \prod_{i=1}^{q} \eta_i^{\gamma_{ij}} \quad (j = \overline{1, \mu}). \tag{2.11}$$

Это уже будут составные степенные комплексы по отношению к комплексам первой группы.

Пусть определяющая зависимость, т. е. зависимость, связывающая определяющие величины, представляется в виде

$$F(x_1, \ldots, x_k; X_1, \ldots, X_\mu) = 0. \tag{2.12}$$

В общем случае

$$\{F\} \neq 1.$$

Любая размерная функция может быть представлена в виде произведения степенного комплекса и соответствующей безразмерной функции (раздел § 2.2, свойство 4):

$$F = k\Phi, \qquad k \neq 0.$$

После сокращения на k уравнение (2.12) примет безразмерную форму:

$$\Phi(x_1, \ldots, x_k; X_1, \ldots, X_\mu) = 0; \tag{2.13}$$

$$\{\Phi\} = 1.$$

Определяющие величины и безразмерные постоянные под знаком безразмерной функции Φ объединяются в безразмерные степенные комплексы общего вида

$$\pi_r = \frac{X_r}{\prod\limits_{i=1}^{k} x_i^{\lambda_{ir}}}. \tag{2.14}$$

Пусть простыми из этих комплексов будут π_1, π_2, $\ldots \pi_p$ $(p \le k)$. Тогда зависимость для Φ можно представить в виде

$$\Phi = \Psi\left(\pi_1, \pi_2, \ldots \pi_p\right). \tag{2.15}$$

Это уравнение, выражающее зависимость между определяющими величинами через критерии подобия, называется *критериальным уравнением*.

Определяющее уравнение (2.12) выражает в неявной форме одну из определяющих величин через другие как через независимые аргументы. Но критериальное уравнение может и не выражать один из критериев подобия через другие, т. к. хотя они и простые, но все они могут зависеть от одной и той же определяющей величины.

Выражение (2.14) содержит безразмерные степенные комплексы, образованные всеми размерными определяющими величинами. Это означает, что определяющие величины всегда разделяются на величины с зависимыми и величины с независимыми размерностями. Получить критерии подобия в виде (2.14) и выделить из них простые (2.15) можно только располагая выражением определяющей зависимости (2.12).

Полученные результаты обобщаются π-теоремой.

Теорема. *Любое уравнение* (2.12), *содержащее величины с независимыми размерностями:* $\{x_i\}$ $(i = \overline{1,k})$ *и с зависимыми размерностями:* $\{X_j\}$ $(j = \overline{1,l})$, *можно привести к уравнению* (2.15) для условных безразмерных критериев π_r $(r \le k)$.

Из вывода π-теоремы и приведенных рассуждений вытекает **утверждение**: *если среди определяющих* $k + l$ *величин*

k величин имеют независимую размерностью и l — зависимую, то всегда можно составить p простых степенных комплексов, причем $p \le k$.

Пример. Для математического маятника (рис. 2.3) необходимо установить связь между периодом его колебания T, длиной нити маятника l, массой m подвешенного груза и ускорением свободного падения g, если размерностями характеристик маятника являются:

$$\{T\} = T, \quad \{l\} = L, \quad \{m\} = M, \quad \{g\} = LT^{-2}.$$

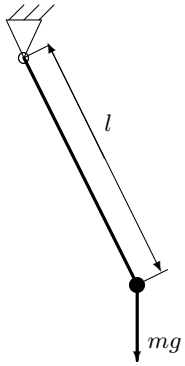


Рис. 2.3
Математический маятник

Это определяющие величины и их размерности. В задаче число определяющих величин $n = 4$ (T, l, m, g), число независимых величин $k = 3$ (T, l, m).

При рассмотрении размерностей определяющих величин устанавливаем, что размерность M входит лишь в $[m]$. Это означает, что ни один степенной комплекс, содержащий m, не будет безразмерным. Значит, либо m должна быть исключена из перечня определяющих величин, либо к этому перечню необходимо добавить некоторую величину, причем размерность этой величины должна содержать M.

В рассматриваемой задаче имеет место первый случай, т. е. m должна быть исключена из рассмотрения.

Составляем новый перечень:

$$T, l, g \to n = 3; \quad [T] = T, \quad [l] = L, \quad [g] = LT^{-2};$$

$$k = 2; \to p \le n - k = 3 - 2 = 1.$$

Безразмерный степенной комплекс, составленный из определяющих величин, примет вид

$$\pi = T g^{1/2} l^{-1/2} = T\sqrt{\frac{g}{l}}.$$

Зная g и l и определяя из эксперимента T, можно вычислить π. Тогда $T = \pi\sqrt{\dfrac{l}{g}}$.

Из аналитического решения (и из соответствующего эксперимента) следует, что в качестве степенного комплекса π следует выбрать величину 2π – параметр длины окружности. Поэтому

$$T = 2\pi\sqrt{\frac{l}{g}}.$$

§ 2.6. РЕЗЮМЕ

Перенесение свойств модели на оригинал оправдано только в том случае, когда между этими двумя объектами существует подобие — полная математическая аналогия между интересующими исследователя величинами.

Подобие между объектами имеет место, если, во-первых, эти объекты описываются сходственными математическими соотношениями, во-вторых, сходственные переменные связаны постоянными константами подобия, называемыми масштабами.

В случае, когда математические модели имеют вид степенных комплексов, их непротиворечивость оригиналу можно установить, используя масштабные уравнения (метод подстановки) или опираясь на способ критериев подобия.

Важное место в моделировании и построении адекватных моделей занимает π-теорема. На основании этой теоремы можно привлечь к описанию процессов такие величины, функциональная зависимость между которыми может быть установлена путем согласования их размерностей.

§ 2.7. ВОПРОСЫ

1. Что такое математическое моделирование объектов, систем и процессов?
2. Какие математические модели называют подобными?
3. Что собой представляют масштабы в подобных моделях?

4. Перечислите основные свойства степенных комплексов.
5. Что такое масштабное уравнение?
6. Как установить непротиворечивость математических моделей?
7. В чем заключается суть метода подстановки установления непротиворечивости моделей? Способа критериев подобия?
8. Какие исходные данные требуются для осуществления подобного моделирования?
9. Перечислите основные этапы подобного моделирования.
10. Что представляет собой физическая модель?
11. Что такое определяющая величина? Определяющее уравнение?
12. Какие размерности называются зависимыми? Независимыми?
13. Что такое критериальное уравнение?
14. Сформулируйте π-теорему и поясните ее смысл.
15. Какое соотношение имеет место между количеством простых степенных комплексов и количеством основных определяющих величин?

§ 2.8. УПРАЖНЕНИЯ К РАЗДЕЛУ 2.1

ТЕМА 2.1
ПОДОБИЕ. СТЕПЕННЫЕ КОМПЛЕКСЫ

СЕМИНАРСКОЕ ЗАНЯТИЕ

Вопросы

1. Что такое математическое моделирование объектов, систем и процессов?
2. Какие математические модели называют подобными?
3. Что представляет собой степенной комплекс?
4. Перечислите основные свойства степенных комплексов.

5. Какое уравнение называют «определяющим» для установления производных единиц измерения?
6. Какие величины называют размерными? Безразмерными?
7. Что такое аналогия? аналогичное моделирование? По каким признакам можно сделать умозаключение по аналогии?

Задачи

1. Найти сумму бесконечного ряда:

$$S = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots, \tag{2.16}$$

используя аналогию между известным представлением коэффициентов уравнения:

$$a_0 - a_1 x + a_2 x^2 - a_3 x^3 + \ldots = 0, \tag{2.17}$$

имеющего положительные и вещественные корни, в частности,

$$a_1 = a_0 \left(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \ldots \right) \tag{2.18}$$

и разложением синуса в степенной ряд:

$$\sin t = \frac{t}{1!} - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \ldots. \tag{2.19}$$

Р е ш е н и е.

Впервые поставленная задача была решена Эйлером методом аналогии.

Полагая в (2.17) $x = y^2$, придем к уравнению:

$$a_0 - a_1 y^2 + a_2 y^4 - a_3 y^6 + \ldots = 0, \tag{2.20}$$

имеющему корни $\pm y_1,\ \pm y_2,\ \ldots,\ \pm y_n,\ \ldots$.

Коэффициенты a_1 и a_0 уравнения (2.20) на основании (2.18) связаны соотношением:

$$a_1 = a_0 \left(\frac{1}{y_1^2} + \frac{1}{y_2^2} + \frac{1}{y_3^2} + \ldots \right). \tag{2.21}$$

Приравняем в (2.19) $\sin t$ нулю:

$$\sin t = \frac{t}{1!} - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} + \ldots = 0. \qquad (2.22)$$

Функция $\sin t$ имеет период π. Поэтому корни последнего уравнения известны: $\pm\pi$, $\pm 2\pi$, $\ldots$, $\pm n\pi$, $\ldots$.

Эти же корни будут удовлетворять вытекающему из (2.22) уравнению ($t \neq 0$ в общем случае):

$$\frac{1}{1!} - \frac{t^2}{3!} + \frac{t^4}{5!} - \frac{t^6}{7!} + \ldots = 0. \qquad (2.23)$$

Между уравнениями (2.20) и (2.23) существует аналогия. Эти уравнения совпадают, если

$$y = t; \quad a_0 = \frac{1}{1!} = 1; \quad a_1 = \frac{1}{3!} = \frac{1}{6}; \quad \ldots.$$

Поэтому на основании (2.21) запишем соотношение между коэффициентами, стоящими при первом и нулевом слагаемыми разложения (2.19):

$$\frac{1}{3!} = 1 \cdot \left[\frac{1}{\pi^2} + \frac{1}{(2\pi)^2} + \frac{1}{(3\pi)^2} + \ldots \right]$$

или

$$\frac{\pi^2}{6} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots.$$

Сравнивая последнее выражение с (2.16) найдем искомую сумму:

$$S = \frac{\pi^2}{6}.$$

2. Из $N = 3$ величин x_1, x_2, x_3 образовано $n = 4$ степенных комплекса:

$$k_1 = x_1 x_2; \quad k_2 = x_1 x_2^2 x_3; \quad k_3 = x_1^2 x_2^3 x_3; \quad k_4 = x_1^3 x_2^4 x_3. \quad (2.24)$$

Требуется определить простые комплексы.

Р е ш е н и е.

Из первого равенства найдем: $x_1 = \dfrac{k_1}{x_2}$. Подставляя полученное значение в остальные равенства (2.24), после преобразования получим:

$$k_2 = k_1 x_2 x_3; \quad k_3 = k_1^2 x_2 x_3; \quad k_4 = k_1^3 x_2 x_3. \tag{2.25}$$

Определим из первого равенства $x_2 x_3 = \dfrac{k_2}{k_1}$ и подставим полученное выражение в два оставшиеся равенства:

$$k_3 = k_1 k_2; \quad k_4 = k_1^2 k_2.$$

Так как комплексы k_3 и k_4 удалось выразить через два комплекса k_1 и k_2, то они не являются независимыми. Следовательно, из четырех заданных комплексов (2.24) только два являются независимыми. Это базисные комплексы

$$k_1 = x_1 x_2; \quad k_2 = x_1 x_2^2 x_3.$$

Соотношения (2.25) позволяют получить и другие варианты выражений степенных комплексов k_1, k_2, k_3 и k_4 через любые два из них, выбранных в качестве базисных. В качестве самостоятельной работы предлагается получить, например, соотношения $k_2 = k_1^{-1} k_3$, $k_4 = k_1 k_3$, где за базисные выбраны комплексы k_1 и k_3.

3. Угол φ сдвига фазы тока I относительно ЭДС с частотой ω в контуре, образованном последовательным соединением индуктивности $\mathcal{L}$, емкости C и активного сопротивления R, выражается безразмерной функцией:

$$\varphi = \operatorname{arctg} \frac{\omega\mathcal{L} - (\omega C)^{-1}}{R}. \tag{2.26}$$

Требуется:

1) представить записанное выражение в виде степенного комплекса $\varphi = \varphi(k_1; k_2; \ldots)$;

2) убедиться в том, что аргумент является безразмерной величиной.

Р е ш е н и е.

Два степенных комплекса $k_1 = \omega\mathcal{L}R^{-1}$ и $k_2 = \omega^{-1}C^{-1}R^{-1}$ линейно независимы (не могут быть выражены один через другой) хотя бы потому, что переменные $\mathcal{L}$ и C входят только в один из комплексов. Поэтому

$$\varphi = \operatorname{arctg}(k_1 - k_2).$$

Убедимся в том, что k_1 и k_2 безразмерные.

Если использовать обозначения размерностей:

L — длина;

M — масса;

T — время;

I — сила электрического тока,

то размерности входящих в (2.26) величин можно представить в виде

$$\{\varphi\} = 1; \quad \{\omega\} = T^{-1}; \quad \{\mathcal{L}\} = L^2MT^{-2}I^{-2};$$

$$\{C\} = T^4I^2L^{-2}M^{-1}; \quad \{R\} = L^2MT^{-3}I^{-2}.$$

Определим размерности степенных комплексов:

$$\{k_1\} = \frac{\{\omega\}\cdot\{\mathcal{L}\}}{\{R\}} = \frac{T^{-1}\cdot L^2MT^{-2}I^{-2}}{L^2MT^{-3}I^{-2}} = 1;$$

$$\{k_2\} = \frac{1}{\{\omega\}\cdot\{C\}\cdot\{R\}} = \frac{1}{T^{-1}\cdot T^4I^2L^{-2}M^{-1}\cdot L^2MT^{-3}I^{-2}} = 1.$$

То есть оба степенных комплекса — безразмерные величины.

4. Модуль суммарного сигнала на выходе приемника РЛС определяется выражением

$$|U_{\text{вых}}| = \sqrt{U^2 + U_{\text{ш}}^2 + 2UU_{\text{ш}}\cos\Delta\varphi}, \qquad (2.27)$$

где U, $U_{\text{ш}}$ — амплитуды сигнала и шума; $\Delta\varphi = \varphi_{\text{вых}} - \varphi_{\text{ш}}$ — разность фаз суммарного и шумового сигналов.

Размерности входящих в (2.27) величин известны:

$$\{U\} = \{U_{\text{вых}}\} = \{U_{\text{ш}}\} = B; \qquad \{\varphi_{\text{вых}}\} = \{\varphi_{\text{ш}}\} = 1.$$

Требуется выразить размерный степенной комплекс (2.27) через безразмерный.

Р е ш е н и е.

Требуемое представление должно быть сведено к зависимости между размерным $F(x_1; x_2; \ldots)$ и безразмерным $f(x_1; x_2; \ldots)$ степенными комплексами:

$$F(x_1; x_2; \ldots) = kf(x_1; x_2; \ldots),$$

где k — некоторый размерный комплекс. В случае, если $f(x_1; x_2; \ldots)$ — безразмерная функция, размерность k должна совпадать с размерностью функции $F(x_1; x_2; \ldots)$.

Задача может быть решена несколькими способами:

$$|U_{\text{вых}}| = U\sqrt{1 + \frac{U_{\text{ш}}^2}{U^2} + 2\frac{U_{\text{ш}}}{U}\cos\Delta\varphi} =$$

$$= \sqrt{UU_{\text{ш}}}\sqrt{\frac{U}{U_{\text{ш}}} + \frac{U_{\text{ш}}}{U} + \cos\Delta\varphi} =$$

$$= U_{\text{ш}}\sqrt{\frac{U^2}{U_{\text{ш}}^2} + 1 + \frac{U}{U_{\text{ш}}}\cos\Delta\varphi}.$$

Во всех трех равенствах суммы под знаком корня — безразмерные величины, а множители перед корнями имеют размерность B.

5. Представить функцию $y = \dfrac{x_2}{x_1^2 + x_2^2 x_3 \sin(x_4 x_5)}$ в виде степенного комплекса.

В каких случаях степенные функции теряют независимость?

Р е ш е н и е.

1). Пусть $k_1 = x_1^2$; $k_2 = x_2$; $k_3 = x_2^2 x_3$; $k_4 = x_4 x_5$.

Тогда

$$y = \frac{k_2}{k_1 + k_3 \sin(k_4)}.$$

2). Пусть $k_1 = x_2$; $k_2 = x_1^2$; $k_3 = x_3$; $k_4 = x_4 x_5$.
Тогда

$$y = \frac{k_1}{k_2 + k_1^2 k_3 \sin(k_4)}.$$

И другие варианты.

Степенные комплексы потеряют независимость в случае, если хотя бы один из них может быть представлен в виде произведения различных степеней других комплексов. Например, если $x_3 = x_1 x_2$, то в случае первого варианта k_3 становится зависимым, так как может быть выражен через k_1 и k_2. Действительно, $k_3 = x_2^2 x_3 = x_1 x_2^3 = \sqrt{k_1} k_2^3$.

§ 2.9. УПРАЖНЕНИЯ К РАЗДЕЛАМ 2.2, 2.3

ТЕМА 2.2
МАСШТАБНЫЕ УРАВНЕНИЯ
СЕМИНАРСКОЕ ЗАНЯТИЕ

Вопросы

1. Что собой представляют масштабы в подобных моделях?
2. Что такое масштабное уравнение?
3. Как установить непротиворечивость в математических моделях?
4. В чем заключается суть метода подстановки установления непротиворечивости моделей? Способа критериев подобия?
5. Какие исходные данные требуются для осуществления подобного моделирования?
6. Перечислите основные этапы подобного моделирования.
7. Что такое определяющая величина? Определяющее уравнение?

Задачи

1. Установить подобие функций $y_1 = x_1^2$ и $y_2 = 8x_2^2$.
Р е ш е н и е.

Введем критерии подобия (безразмерные степенные комплексы):

$$\Pi_1 = \frac{x_1^2}{y_1}; \qquad \Pi_2 = \frac{8x_2^2}{y_2}$$

и масштабы:

$$m_y = \frac{y_1}{y_2}; \qquad m_x = \frac{x_1}{x_2}.$$

Исходные функции будут подобными, если их критерии подобия равны. То есть

$$\frac{\Pi_1}{\Pi_2} = \frac{x_1^2}{y_1} \cdot \frac{y_2}{8x_2^2} = \frac{(x_1/x_2)^2}{8y_1/y_2} = \frac{m_x^2}{8m_y} = 1.$$

Записанное отношение будет равно единице, если масштабные коэффициенты $m_x = x_1/x_2$ и $m_y = y_1/y_2$, являющиеся целыми числами (или числами, обратными целым), будут равны $m_x = 4$ и $m_y = 2$.

Тогда из отношения $\dfrac{y_1}{y_2} = \dfrac{x_1^2}{8x_2^2}$ $\left(\dfrac{\Pi_1}{\Pi_2} = 1\right)$ следует $m_y = m_x^2/8$.

На рис. 2.4 слева изображены две заданные в условии задачи функции, представленные одной параболой, но в координатах, имеющих различный масштаб. На том же рисунке справа — функции изображены в одинаковых координатных осях.

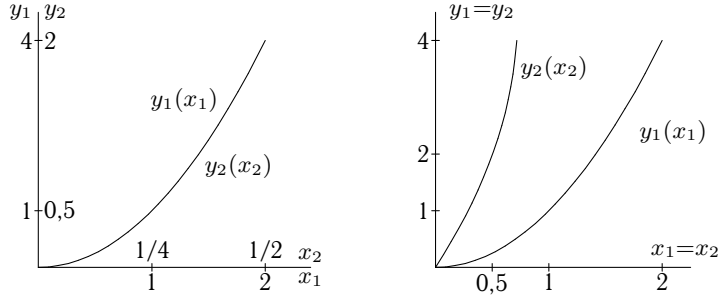


Рис. 2.4
Результаты решения задачи 1

2. Заданы два сходственных уравнения:

$$a_{11}x_{11}\sin(a_{12}y_1) - a_{13}x_{12}^2 = 0; \tag{2.28}$$

$$a_{21}x_{21}\sin(a_{22}y_2) - a_{23}x_{22}^2 = 0. \tag{2.29}$$

Требуется построить масштабные уравнения.

Р е ш е н и е.
Введем масштабные коэффициенты:

$$m_y = \frac{y_1}{y_2}; \qquad m_{x1} = \frac{x_{11}}{x_{21}}; \qquad m_{x2} = \frac{x_{12}}{x_{22}}. \tag{2.30}$$

Прямое решение уравнения (2.28) имеет вид

$$y_1 = \frac{1}{a_{12}}\arcsin\frac{a_{13}x_{12}^2}{a_{11}x_{11}}. \tag{2.31}$$

Заменим в (2.29) переменные y_2, x_{21} и x_{22} выраженными из (2.30) переменными y_1, x_{11} и x_{12} :

$$a_{21}\frac{x_{11}}{m_{x1}}\sin\left(a_{22}\frac{y_1}{m_y}\right) - a_{23}\frac{x_{12}^2}{m_{x2}^2} = 0.$$

Решением этого уравнения, являющимся «окольным» решением уравнения (2.28), является функция:

$$y_1^* = \frac{m_y}{a_{22}}\arcsin\left(\frac{a_{23}m_{x1}}{a_{21}m_{x2}^2}\cdot\frac{x_{12}^2}{x_{11}}\right). \tag{2.32}$$

Для совпадения решений y_1 и y_1^* необходимо выполнение условий:

$$\frac{1}{a_{12}} = \frac{m_y}{a_{22}}; \qquad \frac{a_{13}}{a_{11}} = \frac{a_{23}m_{x1}}{a_{21}m_{x2}^2},$$

или

$$\frac{a_{12}m_y}{a_{22}} = 1; \qquad \frac{a_{13}a_{21}m_{x2}^2}{a_{11}a_{23}m_{x1}} = 1. \tag{2.33}$$

Последние соотношения являются искомыми масштабными уравнениями. Правда, из второго уравнения (2.33) масштабы m_{x1} и m_{x2} определяются неоднозначно.

Рассмотрим еще один вариант решения задачи.

Для вывода условия равенства прямого y_1 (2.31) и окольного y_1^* (2.32) решений умножим (2.29) на $a_{11}m_{x_1}/a_{21}$:

$$a_{11}x_{11}\sin\left(\frac{a_{22}}{m_y}y_1\right) - \frac{a_{23}a_{11}m_{x_1}}{m_{x_2}^2 a_{21}}x_{12}^2 = 0. \qquad (2.34)$$

Размерности сходственных коэффициентов (2.34) и (2.28) равны.

Для того, чтобы решения (2.28) и (2.29) были тождественными, необходимо выполнить условия:

$$a_{12} = \frac{a_{22}}{m_y}; \qquad a_{13} = \frac{a_{23}a_{11}m_{x_1}}{m_{x_2}^2 a_{21}},$$

равносильные условиям (2.33).

Если в частном случае размерности сходственных коэффициентов уравнений (2.31) и (2.32) одинаковы, то приравнивая их, получим:

$$a_{12} = \frac{a_{22}}{m_y}; \qquad a_{13} = \frac{a_{23}}{m_{x_2}^2}; \qquad a_{11} = \frac{a_{21}}{m_{x_1}}.$$

Такая сходственная система уравнений определяет все три масштаба однозначно.

3. Вывести критерии подобия и записать критериальные уравнения для сходственных уравнений задачи 2.

Р е ш е н и е.

Приведем уравнения (2.28) и (2.29) к безразмерному виду:

$$\sin(a_{12}y_1) - \frac{a_{13}x_{12}^2}{a_{11}x_{11}} = 0;$$

$$\sin(a_{22}y_2) - \frac{a_{23}x_{22}^2}{a_{21}x_{21}} = 0$$

и затем к критериальному виду

$$\sin\Pi_{11} - \Pi_{12} = 0; \qquad (2.35)$$

$$\sin\Pi_{21} - \Pi_{22} = 0. \qquad (2.36)$$

В этих уравнениях введены критериальные коэффициенты:

$$\Pi_{11} = a_{12}y_1; \qquad \Pi_{12} = \frac{a_{13}x_{12}^2}{a_{11}x_{11}};$$
$$\Pi_{21} = a_{22}y_2; \qquad \Pi_{22} = \frac{a_{23}x_{23}^2}{a_{21}x_{21}}.$$

Вводя масштабы для сходственных переменных:

$$m_y = \frac{y_1}{y_2}; \qquad m_{x_1} = \frac{x_{11}}{x_{21}}; \qquad m_{x_2} = \frac{x_{12}}{x_{22}},$$

составим масштабные уравнения:

$$\frac{\Pi_{11}}{\Pi_{21}} = \frac{a_{12}y_1}{a_{22}y_2} = \frac{a_{13}m_y}{a_{22}} = 1;$$

$$\frac{\Pi_{12}}{\Pi_{22}} = \frac{a_{13}x_{12}^2 \cdot a_{21}x_{21}}{a_{11}x_{11} \cdot a_{23}x_{22}^2} = \frac{a_{13}a_{21}m_{x_2}^2}{a_{23}a_{11}m_{x_1}} = 1.$$

При выполнении этих масштабных уравнений критериальные уравнения (2.35) и (2.36) становятся тождественными.

Последовательность преобразования сходственных уравнений с применением критерия подобия следующая.

1. Сходственные уравнения приводятся к безразмерному виду.

2. Определяются критерии подобия.

3. Отношения сходственных критериев приравниваются к единице. В результате получаются масштабные уравнения.

Формы масштабных уравнений аналогичны формам соответствующих критериев подобия.

Отметим, что способ подстановки нагляден и естественен, хотя и несколько сложен. Способ критериев подобия носит более формальный характер (по сравнению со способам подстановки), но значительно проще в практическом применении.

4. Имеются два генератора переменного тока. Напряжения u_1 и u_2 в этих генераторах выражаются через время t зависимостями:

$$u_1 = 100\sin[2\pi(t_1/4)] \quad \text{и} \quad u_2 = 10\sin[2\pi(t_2/2)].$$

Требуется получить выражения для масштабов m_u и m_t.

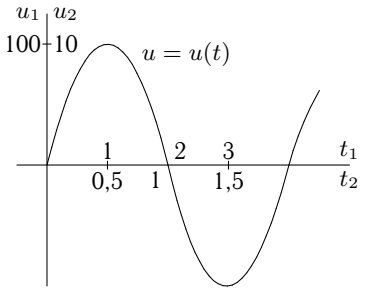


Рис. 2.5
Напряжения в двух генераторах

Р е ш е н и е.

Ясно, что физическое подобие имеет место.

Масштаб $m_t = t_1/t_2$ соответствует отношению периодов колебаний напряжений T_1/T_2.

Так как периоды $T_1 = 4$ и $T_2 = 2$, то

$$m_t = \frac{T_1}{T_2} = \frac{t_1/4}{t_2/2} = \frac{4}{2} = 2.$$

Или $t_1 = 2t_2$.

При соблюдении этого масштаба для времени координаты по оси t будут совпадать, и синусы в выражениях для напряжений будут равны (рис. 2.5). В этом случае масштаб для напряжений

$$m_u = \frac{u_1}{u_2} = \frac{100\sin[2\pi(t_2/2)]}{10\sin[2\pi(t_2/2)]} = 10.$$

5. Для системы уравнений

$$\begin{cases} y_1 = a_1x_1 + a_2x_1^2 + a_3x_1^3; \\ y_2 = b_1x_2 + b_2x_2^2 + b_3x_2^3 \end{cases} \tag{2.37}$$

получить масштабные уравнения и записать условия, обеспечивающие подобие объектов.

Р е ш е н и е.

Приведем уравнения к безразмерному виду

$$\begin{cases} \dfrac{a_1x_1}{y_1} + \dfrac{a_2x_1^2}{y_1} + \dfrac{a_3x_1^3}{y_1} = 1; \\ \dfrac{b_1x_2}{y_2} + \dfrac{b_2x_2^2}{y_2} + \dfrac{b_3x_2^3}{y_2} = 1. \end{cases}$$

Вводя критерии подобия:

$$\Pi_{11} = \frac{a_1x_1}{y_1}; \quad \Pi_{12} = \frac{a_2x_1^2}{y_1}; \quad \Pi_{13} = \frac{a_3x_1^3}{y_1};$$

$$\Pi_{21} = \frac{b_1x_2}{y_2}; \quad \Pi_{22} = \frac{b_2x_2^2}{y_2}; \quad \Pi_{23} = \frac{b_3x_2^3}{y_2},$$

перепишем заданные уравнения:

$$\begin{cases} \Pi_{11} + \Pi_{12} + \Pi_{13} = 1; \\ \Pi_{21} + \Pi_{22} + \Pi_{23} = 1. \end{cases}$$

Запишем необходимые условия для подобия объектов (равенство единице отношений соответствующих критериев), вводя масштабные коэффициенты ($m_x = x_1/x_2$; $m_y = y_1/y_2$) :

$$\frac{\Pi_{11}}{\Pi_{21}} = \frac{a_1x_1y_2}{b_1x_2y_1} = \frac{a_1m_x}{b_1m_y} = 1 \longrightarrow m_y = \frac{a_1}{b_1}m_x;$$

$$\frac{\Pi_{12}}{\Pi_{22}} = \frac{a_2x_1^2y_2}{b_2x_2^2y_1} = \frac{a_2m_x^2}{b_2m_y} = 1 \longrightarrow m_y = \frac{a_2}{b_2}m_x^2;$$

$$\frac{\Pi_{13}}{\Pi_{23}} = \frac{a_3x_1^3y_2}{b_3x_2^3y_1} = \frac{a_3m_x^3}{b_3m_y} = 1 \longrightarrow m_y = \frac{a_3}{b_3}m_x^3.$$

Эта система масштабных уравнений позволяет записать дополнительные условия для коэффициентов (приравниваем правые части последних равенств и исключаем из них масштабы):

$$\frac{a_1}{b_1} = \frac{a_2^2b_3}{b_2^2a_3}$$

и выражения для масштабов:

$$m_x = \frac{a_2 b_3}{b_2 a_3}; \qquad m_y = \frac{a_1}{b_1} m_x = \frac{a_1 a_2 b_3}{b_1 b_2^2 a_3} = \frac{a_1^2 b_2}{a_2^2 b_1}.$$

6. Заданы два сходственных уравнения ($a > 0$, $b > 0$):

$$y = ax_1 + bx_2;$$
$$\varphi = -\varphi_1 + 2\varphi_2.$$

Приведем уравнения к безразмерному виду

$$1 = \frac{ax_1}{y} + \frac{bx_2}{y};$$

$$1 = -\frac{\varphi_1}{\varphi} + \frac{2\varphi_2}{\varphi}.$$

Вводя критерии подобия:

$$\Pi_{11} = \frac{ax_1}{y}; \quad \Pi_{12} = \frac{bx_2}{y};$$
$$\Pi_{21} = -\frac{\varphi_1}{\varphi}; \quad \Pi_{22} = \frac{2\varphi_2}{\varphi}$$

и масштабы:

$$m = \frac{y}{\varphi}; \quad m_1 = \frac{-x_1}{\varphi_1} = \frac{x_1}{-\varphi_1}; \quad m_2 = \frac{x_2}{\varphi_2}, \tag{2.38}$$

составим масштабные уравнения:

$$\frac{\Pi_{11}}{\Pi_{12}} = -\frac{ax_1\varphi}{\varphi_1 y} = \frac{am_1}{m} = 1; \quad \frac{\Pi_{21}}{\Pi_{22}} = \frac{bx_2\varphi}{2\varphi_2 y} = \frac{bm_2}{2m} = 1.$$

Или

$$m_1 = \frac{m}{a}; \qquad m_2 = \frac{2m}{b}.$$

К такому же результату можно прийти, вводя отличные от (2.38) масштабные коэффициенты, например:

$$m = \frac{-y}{\varphi} = \frac{y}{-\varphi}; \qquad m_1 = \frac{x_1}{\varphi_1}; \qquad m_2 = \frac{-x_2}{\varphi_2} = \frac{x_2}{-\varphi_2}.$$

7. Дана система масштабных уравнений:

$$\begin{cases} m_1 m_2^{-2} m_3 &= 2; \\ m_1^2 m_2 m_3^{-1} &= 2,5; \\ m^4 m_2^{-3} m_3 &= 10. \end{cases}$$

Требуется:

1) доказать непротиворечивость системы;

2) выделить независимые уравнения.

Р е ш е н и е.

Для линеаризации входящих в систему уравнений прологарифмируем их:

$$\begin{cases} \lg m_1 - 2\lg m_2 + \lg m_3 &= \lg 2; \\ 2\lg m_1 + \lg m_2 - \lg m_3 &= \lg 2,5; \\ 4\lg m_1 - 3\lg m_2 + \lg m_3 &= 1. \end{cases}$$

Вычислим определитель коэффициентов последней системы уравнений:

$$\Delta = \begin{vmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 4 & -3 & 1 \end{vmatrix} \quad (-2N_1 - N_2) \quad = \begin{vmatrix} 1 & -2 & 1 \\ 2 & 1 & -1 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

Определитель матрицы коэффициентов системы равен нулю, а ее ранг, как видно из последнего равенства, равен 2. По теореме Кронекера-Капелли, система будет иметь решения (бесчисленное множество в случае равенства нулю определителя), если ранг ее расширенной матрицы будет равен рангу матрицы коэффициентов.

Ранг расширенной матрицы равен двум. Это следует из того, что сумма умноженного на два первого уравнения системы и второго уравнения дает третье уравнение. То есть третье уравнение линейно зависит от первых двух.

Принимая за базовые переменные $\lg m_2$ и $\lg m_3$, выразим их из первых двух уравнений через $\lg m_1$:

$$\lg m_2 = 3\lg m_1 - \lg 5;$$

$$\lg m_3 = 5\lg m_1 - \lg 12,5.$$

Если масштабных уравнений больше, чем масштабов, прибегают к искусственному приему, рассмотренному в следующем примере.

8. Задана система трех уравнений с двумя неизвестными — масштабами:

$$m_1 m_2^{-1} = 6; \quad m_1^2 m_2^{-1} = 1,5; \quad m_1^3 m_2^{-1} = 3/8.$$

Для получения решения системы уравнений заменяем в каком-либо из них одну из переменных, например m_2, на новую переменную $\overline{m}_2$. В этом случае исходная система двух неизвестных становится системой с тремя неизвестными:

$$m_1 m_2^{-1} = 6; \quad m_1^2 m_2^{-1} = 1,5; \quad m_1^3 \overline{m}_2^{-1} = 3/8.$$

Решением последней системы являются корни:

$$m_1 = \frac{1}{4}; \quad m_2 = \frac{1}{24}; \quad \overline{m}_2 = \frac{1}{24}.$$

Равенство $m_2 = \overline{m}_2$ говорит о том, что исходная система непротиворечива.

Полученное решение дает возможность однозначно определить масштабные коэффициенты.

З а д а н и е.

Получите решение для масштабных коэффициентов предыдущего примера, используя описанный прием (в последнем, например, уравнении заменить m_3 на $\overline{m}_3$).

§ 2.10. УПРАЖНЕНИЯ К РАЗДЕЛАМ 2.4, 2.5

ТЕМА 2.3
π-ТЕОРЕМА

СЕМИНАРСКОЕ ЗАНЯТИЕ

Вопросы

1. Что представляет собой физическая модель?
2. Что такое определяющая величина? Определяющее уравнение?

3. Какие размерности называются зависимыми? Независимыми?
4. Что такое критериальное уравнение?
5. Сформулируйте π-теорему и поясните ее смысл.
6. Какое соотношение имеет место между количеством простых степенных комплексов и количеством основных определяющих величин?

Задачи

1. Осуществить подобное моделирование объекта, описываемого соотношением:

$$\Phi = \int_0^t \varphi dt, \tag{2.39}$$

где Φ, φ — угловые величины, измеряемые в радианах; t — время в секундах.

1). В качестве объекта модели выберем математический аналог — генератор линейно изменяющегося напряжения, — описываемый сходственным с (2.39) уравнением:

$$U = -2\int_0^\tau u d\tau, \tag{2.40}$$

где U — напряжение в вольтах; t — время в секундах.

Вводя два оператора дифференцирования

$$D = \frac{d}{dt}, \qquad D = \frac{d}{d\tau}$$

и продифференцировав соотношения (2.39) и (2.40) по t и τ, соответственно, перейдем к двум дифференциальным уравнениям первого порядка:

$$D\Phi = \varphi; \qquad DU = -2u \tag{2.41}$$

с нулевыми начальными условиями:

$$\Phi(0) = 0; \qquad U(0) = 0.$$

Множитель при Φ в первом уравнении (2.41) имеет размерность $1/c$.

2). Приведем дифференциальные уравнения (2.41) к безразмерному виду и введем критерии подобия:

$$1 = \frac{\varphi}{D\Phi} = \Pi_1; \qquad 1 = -\frac{2u}{DU} = \Pi_2.$$

3). Составим масштабное уравнение (в рассматриваемой задаче оно одно):

$$\frac{\Pi_1}{\Pi_2} = -\frac{\varphi DU}{2uD\Phi} = 1.$$

Из последнего равенства следует, что в одной из двух пар сходственных переменных φ и u или Φ и U переменные должны иметь разные знаки.

4). Введя масштабы

$$m_1 = \frac{\varphi}{-u}; \qquad m_2 = \frac{\Phi}{U}; \quad m_t = \frac{t}{\tau},$$

запишем масштабное уравнение в окончательном виде

$$\frac{m_1 m_t}{2m_2} = 1.$$

Это уравнение единственное, поэтому не может быть противоречивым.

Отметим, что нулевые начальные условия являются подобными условиями однозначности при любых масштабах.

Частные случаи

Согласно полученному масштабному уравнению можно, например, принять $m_1 = 2B^{-1}$; $m_2 = 10B^{-1}$. Тогда $m_t = 10$.

При таком масштабе времени процессы, проходящие в модели, будут аналогичны процессам, проходящим в оригинале, но будут протекать в 10 раз быстрее.

Для расчета функции $u = u(\tau)$, подобной функции $\varphi = \varphi(t)$, имеем:

$$u = -\frac{\varphi}{m_1} = -\frac{\varphi(t)}{m_1} = -\frac{\varphi(\tau m_t)}{m_1}.$$

Или окончательно:

$$u = -\frac{1}{2}\varphi(10\tau).$$

Если, в частности, $\varphi = 10\sin t$, то

$$u = -5\sin(10\tau).$$

2. Явление прохождения электрического тока i через последовательно соединенные резистор R и конденсатор C (рис. 2.6) при включении источника напряжения $u = const$ в момент времени t описывается величинами i, u, R, C, t, связанными функциональной зависимостью:

$$i = \frac{u}{R}\exp\left(-\frac{t}{RC}\right). \tag{2.42}$$

Переходный процесс в RC-цепи определяется $N = 5$ величинами i, u, R, C, t. Определяющее выражение, связывающее эти величины, представляется в виде конечного уравнения:

$$F(i; u; R; C; t) = \frac{u}{R}\exp\left(-\frac{t}{RC}\right) - i = 0. \tag{2.43}$$

Или в безразмерной форме:

$$\Phi(i; u; R; C; t) = \frac{u}{Ri}\exp\left(-\frac{t}{RC}\right) - 1 = 0. \tag{2.44}$$

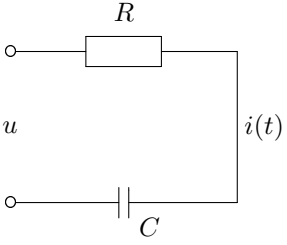


Рис. 2.6
RC-цепь

Введя обозначения:

$$\Pi_1 = \frac{u}{Ri}; \qquad \Pi_2 = \frac{t}{RC}, \tag{2.45}$$

запишем (2.44) в виде критериального уравнения:

$$\Psi(\Pi_1; \Pi_2) = \Pi_1 \exp(-\Pi_2) - 1 = 0. \tag{2.46}$$

Если определяющее уравнение (2.43) не известно, а известен лишь перечень определяющих величин i, u, R, C, t, то можно найти *условные критерии подобия* π.

Размерности определяющих величин в системе СИ таковы:

$$\{i\}=I; \quad \{u\}=L^2MT^{-3}I^{-1}; \quad \{R\}=L^2MT^{-3}I^{-2};$$

$$\{C\}=L^{-2}M^{-1}T^4I^2; \quad \{t\}=T.$$

Анализируя эти размерности, можно установить, что $\{u\}$ и $\{t\}$ могут быть выражены через $\{i\}$, $\{R\}$ и $\{C\}$:

$$\{u\} = \{R\}\{i\}; \qquad \{t\} = \{R\}\{C\}. \tag{2.47}$$

То есть из $N = 5$ входящих в определяющее уравнение величин i, u, R, C и t, имеют независимую размерность $n = 3$: i, R, C, а зависимую размерность $m = N - n = 2$: u и t.

Зависимость между i, u, R, C и t может быть представлена *условно критериальными уравнениями,* связывающими два условных критерия подобия π_1 и π_2, образованных из $N = 5$ исходных величин.

Соотношения (2.47) позволяют записать эти критерии:

$$\pi_1 = \frac{u}{Ri}; \qquad \pi_2 = \frac{t}{RC}. \tag{2.48}$$

В рассмотренной задаче отсутствуют безразмерные постоянные. Поэтому, согласно (2.48) и (2.45), $\Pi_1 = \pi_1$ и Π_2, а условно критериальное уравнение можно представить в виде, совпадающем с (2.46):

$$\Psi(\pi_1; \pi_2) = \pi_1 \exp(-\pi_2) - 1 = 0. \tag{2.49}$$

3. Требуется определить расход жидкости q, вытекающей в единицу времени через круглое отверстие диаметром d из проектируемого резервуара больших размеров. Установившееся струйное течение жидкости плотности ρ происходит под напором столба жидкости высотой h.

Р е ш е н и е.

Используем физическое подобное моделирование. Предполагаем, что процесс вытекания жидкости может быть описан четырьмя заданными величинами q, d, ρ, h. Размерности этих величин в системе СИ:

$$\{q\} = LMT^{-3}; \quad \{\rho\} = ML^{-3}; \quad \{h\} = L; \quad \{d\} = L.$$

Для плоской задачи массовый расход ρdV, где V — скорость. Поэтому

$$\{q\} = \frac{M}{L^2} \cdot \frac{L}{T} \cdot L = \frac{M}{LT} = Ml^{-1}T^{-1}.$$

Анализ этих размерностей приводит к выводу о том, что заданных исходных величин d, ρ и h недостаточно для определения расхода q — в размерность q входит время t, которое не входит ни в одну из размерностей заданных величин. Таким образом, перечень заданных величин не является исчерпывающим для определения q.

Известно, что струйное течение жидкости характеризуется ее инерционностью и весомостью. Показателями этих величин являются ускорение силы тяжести g с размерностью $\{g\} = LT^{-2}$ и упомянутая ранее в заданных величинах плотность ρ.

В размерность $\{g\}$ входит время T, поэтому введение g должно решить задачу формирования функциональной зависимости.

Пять заданных величин ($N = 5$) имеют $k = 3$ независимых размерностей.

Скорость $v = \sqrt{gh}$, Поэтому $q = \rho dV$ и

$$\{q\} = \{\rho\}\{g\}^{1/2}\{h\}^{1/2} = ML^{-3} \cdot L^{1/2}T^{-1} \cdot L^{3/2} = ML^{-1}T^{-1};$$

$$\{d\} = \{h\} = L.$$

Число критериев подобия $m = N - k = 5 - 3 = 2$:

$$\pi_1 = \frac{q}{\rho g^{3/2} h^{5/2}}; \quad \pi_2 = \frac{d}{h}.$$

Критерии подобия оригинала:

$$\pi_{O1} = \frac{q_O}{\rho_O g^{3/2} h_O^{5/2}}; \quad \pi_{O2} = \frac{d_O}{h_O}.$$

Критерии подобия модели:

$$\pi_{M1} = \frac{q_M}{\rho_M g^{3/2} h_M^{5/2}}; \quad \pi_{M2} = \frac{B_M}{h_M}.$$

Масштабные уравнения:

$$\frac{\pi_{O1}}{\pi_{M1}} = \frac{q_O \rho_M h_M^{5/2}}{q_M \rho_O h_O^{5/2}} = 1;$$

$$\frac{\pi_{O2}}{\pi_{M2}} = \frac{d_O h_M}{d_M h_O} = 1.$$

Введя масштабы:

$$m_q = \frac{q_O}{q_M}; \quad m_\rho = \frac{\rho_O}{\rho_M}; \quad m_h = \frac{h_O}{h_M}; \quad m_d = \frac{d_O}{d_M},$$

приведем масштабные уравнения к виду

$$\frac{m_q}{m_\rho m_h^{5/2}} = 1; \quad \frac{m_d}{m_h} = 1.$$

Из последних уравнений видно, что два масштабных коэффициента могут быть выбраны произвольно. Тогда третий и четвертый из них определятся через два выбранных однозначно.

Выбрав, например, m_h, можно для стока жидкости построить резервуар уменьшенного диаметра d, но с увеличенной высотой $H_M > h_M = h_O/m_h$.

Выбрав для модели жидкость плотностью ρ_M, можно определить масштабы: $m_\rho = \rho_O/\rho_M, \quad m_q = m_\rho m_h^{5/2}$.

После подстановки в модель постоянного напора h_M экспериментально определяют ρ_M и рассчитывают $q = q_O = m_q q_M$.

4. Путем физического подобного моделирования требуется установить, какую частоту ω малых свободных колебаний будет иметь маятник, имеющий длину l и массу m.

Р е ш е н и е.

Движение маятника обусловлено движением силы тяжести. Поэтому среди определяющих величин в математической модели маятника должно присутствовать ускорение силы тяжести g.

Полагаем, что в искомую модель входят $N = 4$ исходные величины:

– ускорение силы тяжести g с размерностью $\{g\} = LT^{-2}$;
– частота колебаний ω с размерностью $\{\omega\} = T^{-1}$;
– длина нити маятника l с размерностью $\{l\} = L$;
– масса маятника m с размерностью $\{m\} = M$.

Для размерности частоты при заданных других размерностях возможна только одна комбинация исходных величин:

$$\{\omega\} = \{g\}^{1/2}\{l\}^{-1/2} = \left(LT^{-2}\right)^{1/2} \cdot L^{-1/2} = T^{-1}$$

и единственный критерий подобия:

$$\pi = \frac{\omega}{g^{1/2} l^{-1/2}}.$$

Получить критерий подобия, содержащий массу m маятника, невозможно. Это подтверждает известное положение физики: частота колебаний математического маятника не зависит от его массы.

Критерии подобия для оригинала и модели:

$$\pi_O = \frac{\omega_O}{g^{1/2} l_O^{-1/2}}; \quad \pi_M = \frac{\omega_M}{g^{1/2} l_M^{-1/2}}.$$

Масштабное уравнение

$$\frac{\pi_O}{\pi_M} = \frac{\omega_O g^{1/2} l_O^{1/2}}{\omega_M g^{1/2} l_M^{1/2}} = \frac{\omega_O l_O^{1/2}}{\omega_M l_M^{1/2}} = 1.$$

Отсюда:

$$\omega_O = \omega_M \sqrt{\frac{l_M}{l_O}}.$$

Зная l_M и измерив ω_M, по $l = l_O$ находим $\omega = \omega_O$.

5. Консольно заделанная балка прямоугольного сечения нагружена на конце сосредоточенной силой. Установить зависимость между силой и перемещением конца балки (прогибом).

Р е ш е н и е.

Введем обозначения:

l, b, δ, w — длина, ширина, толщина и прогиб балки. Размерности всех этих величин равны ($\{l\} = \{b\} = \{\delta\} = \{w\} = L$). P — сила. Ее размерность $\{P\} = MLT^{-2}$.

Первого взгляда на размерности пяти приведенных величин достаточно, чтобы сделать вывод о невозможности построить по ним математическую модель: только в размерность силы входят L и T. Вместе с тем ясно, что прогиб балки зависит от материала, из которого балка сделана. Из механики деформирования известно, что упругие свойства материала характеризует модуль упругости E, размерность которого $\{E\} = ML^{-1}T^{-2}$.

Для построения искомой модели обратимся к здравому смыслу, опыту и интуиции. С уверенностью можно сказать, что прогиб консоли будет увеличиваться с увеличением длины l и силы P; будет уменьшаться с увеличением площади поперечного сечения $F = \delta b$ (размерность $\{F\} = L^2$) и модуля упругости E. Следовательно, при записи выражения для w в правой части формулы в числителе должны стоять l и P в некоторых степенях, в знаменателе — F и E. При этом степени размерностей P и E должны быть одинаковыми для того, чтобы в их отношении $\left(\frac{\{P\}}{\{E\}}\right)^{\alpha} = \left(\frac{ML/T^2}{M/LT^2}\right)^{\alpha} = L^{2\alpha}$ исчезли размерности M и T.

Учитывая сказанное, запишем математическую модель изгиба балки в виде степенного комплекса

$$w = \frac{l^{\gamma} P^{\alpha}}{F^{\beta} E^{\alpha}}.$$

Представим эту формулу в виде соотношения для размерностей входящих в нее величин:

$$L = \frac{L^{\gamma} L^{2\alpha}}{L^{2\beta}} = L^{\gamma + 2\alpha - 2\beta}.$$

Соотношение будет выполняться, в частности, при $\alpha = \beta = \gamma = 1$. В этом случае искомая формула:

$$w = \frac{lP}{FE}.$$

Именно в таком виде формула получается в курсе «Сопротивление материалов».

6. (Заимствована из журнала: Изв. вузов, серия «Энергетика», 1972. С. 67–70.)

Для скорейшего достижения однородности расплава в плавящей печи после внесения в нее небольшой порции расплавляющей составляющей производится принудительное перемешивание смеси потоком газа. Процесс перемешивания характеризуется $N = 6$ определяющими величинами:

— временем перемешивания t;
— скоростью газа v;
— расходом массы газа в единицу времени q;
— вязкостью расплава ν;
— массы расплава m;
— характерным размером ванны для расплава l.

Для рассматриваемого проекта известны величины (оригинал):

$$v_O; \quad q_O; \quad \nu_O; \quad m_O; \quad l_O.$$

Требуется найти время перемешивания τ_O.

Р е ш е н и е.

Аналитический расчет или непосредственное экспериментальное исследование оригинала не представляются возможными.

Экспериментальное исследование выполнялось на лабораторной модели, геометрически подобной оригиналу, при нормальной температуре. Рабочее тело представляло собой воду, глицерин или водный раствор глицерина, перемешиваемыми потоком воздуха. В качестве перемешивающей составляющей использовались чернила.

Размерности определяющих величин в системе СИ:

$$\{t\} = T; \quad \{v\} = LT^{-1}; \quad \{q\} = MT^{-1};$$

$$\{\nu\} = L^2T^{-1}; \quad \{G\} = LMT^{-2}; \quad \{l\} = L.$$

Составляем исчерпывающие зависимость между размерностями величин:

$$\{t\} = \{l\}\{v\}^{-1}; \quad \{\nu\} = \{l\}\{v\}; \quad \{G\} = \{q\}\{v\}.$$

Так как число соотношений $k = 3$, то $m = N - k = 6 - 3 = 3$. Поэтому составляем $m = 3$ критерия подобия:

$$\pi_1 = \frac{tv}{l}; \quad \pi_2 = \frac{\nu}{lv}; \quad \pi_3 = \frac{G}{qv}.$$

Критерии подобия оригинала:

$$\pi_{O1} = \frac{t_O v_O}{l_O}; \quad \pi_{O2} = \frac{\nu_O}{l_O v_O}; \quad \pi_{O3} = \frac{G_O}{q_O v_O}$$

и модели:

$$\pi_{M1} = \frac{t_M v_M}{l_M}; \quad \pi_{M2} = \frac{\nu_M}{l_M v_M}; \quad \pi_{M3} = \frac{G_M}{q_M v_M}.$$

По заданным l_O, v_O, ν_O, G_O и q_O находятся π_{O2} и π_{O3}. Затем выбирают l_M, v_M, ν_M, G_M и q_M так, чтобы $\pi_{M2} = \pi_{O2}$ и $\pi_{M3} = \pi_{O3}$. На экспериментальной установке выполняют эксперимент и задавая l_M, v_M, ν_M, G_M и q_M измеряют t_M.

Из равенства

$$\pi_{O1} = \frac{t_O v_O}{l_O} = \frac{t_M v_M}{l_M} = \pi_{M1}$$

находят

$$t_O = \frac{\pi_{M1} l_O}{v_O}.$$

7. Моделируется опорная ферма линии электропередач с исходными характеристиками:

— удельный вес материала фермы γ;
— модуль Юнга E;
— коэффициент Пуассона μ;
— характерный размер l;
— внешняя нагрузка (сила) F;
— напряжения в материале фермы σ.

Задача моделирования состоит в определении напряжения в произвольном элементе конструкции фермы.

Требуется показать, что моделирование невозможно без привлечения дополнительных данных или установок.

Р е ш е н и е.

Запишем размерности величин в СИ:

$$\{l\} = L; \quad \{F\} = LMT^{-2}; \quad \{E\} = L^{-1}MT^{-2}; \quad \{\mu\} = 1;$$

$$\{\gamma\} = L^{-2}MT^{-2}; \quad \{\sigma\} = L^{-1}MT^{-2}.$$

Из $N = 6$ определяющих величин независимыми размерностями обладают только две $(k = 2)$: l и F.

Так как $m = N - k = 6 - 2 = 4$, то записываем четыре критерия подобия:

$$\pi_1 = \frac{El^2}{F}; \quad \pi_2 = \frac{\gamma l^3}{F}; \quad \pi_3 = \frac{\sigma l^2}{F}; \quad \pi_4 = 1.$$

Если модель и оригинал выполнены из одного материала, то

$$\mu_O = \mu_M; \quad E_O = E_M.$$

В этом случае для $\pi_1 = const$ не выполнимо условие:

$$l^2/F = const. \tag{2.50}$$

Это означает, что при моделировании необходимо внешнюю нагрузку изменять пропорционально квадрату линейного

размера. При этом согласно выражению для π_3 должно быть $\sigma = const$. Это в свою очередь означает, что для подобных состояний напряжения в сходственных точках оригинала и модели должны быть одинаковыми.

Однако согласно выражению для π_2 при условии (2.50) должно быть $\gamma l = const$, что невозможно, т. к. $\gamma_O = \gamma_M$.

8. Одной из проблем, решаемых инженерами в пору становления космонавтики, была проблема определения таких параметров спускаемых аппаратов, которые позволили бы после удара космического аппарата о землю сохранить находящуюся в нем аппаратуру. Подход к решению задачи с позиций механики деформируемого твердого тела приводил к системам более чем 50 нелинейных уравнений и соотношений, не поддающихся решению с привлечением существовавших в то время средств вычислительной техники.

Эксперимент, имитирующий в начальном приближении поставленную задачу, сводился с сжатию полусферической оболочки и установлению зависимости между сжимающей силой P и вызываемым этой силой перемещениям вершины оболочки H. Оболочка имела радиус кривизны R, толщину δ, модуль упругости материала E.

Эксперимент показал, что зависимость $P = P(H)$ силы от перемещения линейная, несмотря на конечность и неупругость деформаций.

Запишем размерности величин в СИ:

$$\{R\} = \{\delta\} = \{H\} = L; \quad \{P\} = LMT^{-2}; \quad \{E\} = L^{-1}MT^{-2}.$$

Положения π-теоремы позволили получить соотношение (k — безразмерный коэффициент пропорциональности):

$$P = k\frac{E\delta^3}{R}H.$$

Установленная зависимость дала возможность определить обобщенную толщину δ, которая при допустимом перемещении вершины оболочки обеспечивала сохранить оборудования, находящегося выше зоны деформирования.

§ 2.11. ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Заданы два сходственных уравнения

$$D_1 y_1 + a y_1 = b_1 D_1 x_1 + b_0 x_1;$$

$$D_2 y_2 + \alpha y_2 = \beta_1 D_2 x_2 + \beta_0 x_2,$$

где

$$y_1 = y_1(t_1); \quad x_1 = x_1(t_1); \quad D_1 = \frac{d}{dt_1};$$

$$y_2 = y_2(t_2); \quad x_2 = x_2(t_2); \quad D_2 = \frac{d}{dt_2}.$$

Требуется построить масштабные уравнения.

2. Заданы два сходственных уравнения:

$$F(y_1, t_1) = y_1^2 - 2t_1 y_1 - 4 = 0;$$

$$F(y_2, t_2) = y_2^2 - 4t_2 y_2 - 9 = 0.$$

Требуется выбрать условия однозначности:

при $t_{11} = 4$ и $t_{12} = 8$;
при $t_{21} = 3$ и $t_{22} = 6$

так, чтобы два объекта, описываемых исходными уравнениями, были подобны.

3. Заданы два сходственных уравнения:

$$a_2 D_1^2 y_1 + a_1 D_1 y_1 + a_0 y_1 = 0;$$

$$b_2 D_2^2 y_2 + b_1 D_2 y_2 + b_0 y_2 = 0,$$

где $D_1 = \dfrac{d}{dt_1}$; $D_2 = \dfrac{d}{dt_2}$ — операторы дифференцирования по переменным t_1 и t_2; $y_1 = y_1(t_1)$; $y_2 = y_2(t_2)$ и начальные условия (условия однозначности):

$$y_{10} = y_1(0); \quad D_1 y_{10} = y_1'(0);$$
$$y_{20} = y_2(0); \quad D_2 y_{20} = y_2'(0).$$

Требуется получить выражения для масштабных коэффициентов m_t и m_y при условии, что исходные уравнения и начальные условия описывают подобные объекты.

4. Задана система дифференциальных уравнений:

$$\begin{cases} a_1 \dfrac{\partial^2 y_1}{\partial t_1^2} + a_2 \dfrac{\partial^2 y_1}{\partial t_1 \partial x_1} + a_3 \dfrac{\partial y_1}{\partial t_1} = 0; \\ b_2 \dfrac{\partial^2 y_2}{\partial t_2^2} + b_2 \dfrac{\partial^2 y_2}{\partial t_2 \partial x_2} + b_3 \dfrac{\partial y_2}{\partial t_2} = 0 \end{cases}$$

с известными начальными условиями:

при $t_1 = 0$: $\quad y_1 = 0$ и $\dfrac{\partial y_1}{\partial t_1} = -a_4 x_1 + a_5$;

при $t_2 = 0$: $\quad y_2 = 0$ и $\dfrac{\partial y_2}{\partial t_2} = -b_4 x_2 + b_5$.

Требуется получить масштабы m_x, m_y, m_t и дополнительные условия для коэффициентов уравнений.

5. Заданы два объекта, описываемые уравнениями:

$$y' + \left(a + \frac{v'}{v}\right) y = bx;$$

$$y_1' + \alpha y_1 = \beta x_1,$$

где a, b, α, β — положительные константы;

$y = y(t)$, $x = x(t)$, $v = v(t) > 0$, $y_1 = y_1(t_1)$, $x_1 = x_1(t_1)$;

$y(0) = y_1(0) = 0$ — начальные условия.

Составить условия подобия, введя предварительно новые переменные для обеспечения сходственности исходных уравнений.

6. Заданы прямоугольники $OA_1B_1C_1$ и $OA_2B_2C_2$, изображенные на рис. 2.7. Требуется записать условия их подобия в аналитическом виде.

7. Явление подъема жидкости в капиллярной трубке, опущенной одним концом в жидкость, характеризуют величины:

h — высота подъема жидкости;

r — внутренний радиус трубки;

ρ — плотность жидкости;

g — ускорение силы тяжести;

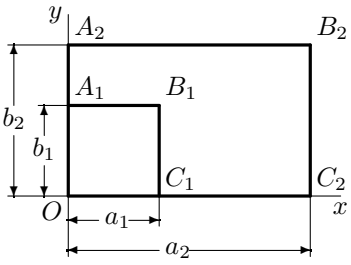


Рис. 2.7
Подобие прямоугольников

σ — коэффициент поверхностного натяжения;
ϕ — угол наклона трубки к поверхности жидкости.

Записать определяющее уравнение, опираясь на размерности

$$[h] = L; \quad [r] = L; \quad [\rho] = L^{-3}M; \quad [g] = LT^{-2};$$

$$[\sigma] = MT^{-2}; \quad [\phi] = 1.$$

8. Доказать, что если в качестве модели взять объект (2.40) задачи 1 темы 2.3, то можно, изучая процессы, происходящие с моделью, утверждать, что они будут подобны процессам, происходящим в оригинале.

Получить формулу для u при масштабах $m_2 = m_t = 10$ для частного случая $\varphi = 10 \sin t$.

Глава 3

ЧИСЛЕННЫЕ МЕТОДЫ

Определение корней, экстремумов функций, координат точек перегиба и других характеристик многих даже элементарных функций представляет собой непростую задачу.

Математические модели, описывающие реальные процессы и явления, как правило, не позволяют определить точные значения упомянутых характеристик. А получение приближенных значений требует необходимых знаний в области специально разработанных для этой цели численных методов. Описанию некоторых из этих методов посвящена глава.

Отметим, что многие задачи исследования характеристик функций (например, одной переменной) сводятся к нахождению корней алгебраических уравнений. В частности, для выявления точек, где функция достигает экстремальных значений, достаточно взять производную от этой функции и, приравняв ее нулю, свести задачу к нахождению корней полученного уравнения. Абсциссы точек перегиба определяются из уравнения, получаемого приравниванием нулю второй производной.

Поэтому задача нахождения корней алгебраических уравнений является ключевой в проблеме отыскания характеристик функций. Существуют методы поиска экстремальных точек непосредственно по поведению самих функций. Этим методам для функций одной переменной посвящены первые разделы главы.

§ 3.1. ПРИБЛИЖЕНИЯ И ОШИБКИ

В результате реализации различных численных методов, неточности эксперимента, округления чисел и по другим причинам появляются ошибки, отличающие численные значения, характеризующие модель и оригинал. Каждое измеряемое значение некоторой величины в общем случае является лишь приближенным значением этой величины.

Если $\tilde{a}$ есть приближенное значение числа a, то:
$a - \tilde{a}$ называется *истинной погрешностью* числа a;
$(a - \tilde{a})/a$ — его *относительной погрешностью*.

Если в реальных задачах значение a остается неизвестным, то неизвестна и истинная погрешность. Тем не менее, часто можно указать удовлетворяющую исследователя граничную величину истинной погрешности — положительное число Δa, для которого выполняется условие:

$$|a - \tilde{a}| \le \Delta a, \quad \text{или} \quad \tilde{a} - \Delta a \le a \le \tilde{a} + \Delta a. \tag{3.1}$$

Величину Δa называют *абсолютной погрешностью*, а

$$\delta a = \frac{\Delta a}{a} \approx \frac{\Delta a}{\tilde{a}} \tag{3.2}$$

относительной погрешностью.

Пример 1. Если число $\tilde{a} = 0{,}32$ получено путем «обрыва» измеряемой величины с удержанием двух знаков после запятой, то в качестве абсолютной погрешности можно принять число $\Delta a = 10^{-2}$. Тогда из неравенства (3.1) следует:

$$0{,}32 - 0{,}01 \le a \le 0{,}32 + 0{,}01, \quad \text{или} \quad 0{,}31 \le a \le 0{,}33.$$

Для функции n переменных $f(x_1, \ldots, x_n)$ часто требуется знание абсолютной погрешности:

$$\Delta f \ge |f(a_1, \ldots, a_n) - f(x_1, \ldots, x_n)|$$

при условии, что величины $\Delta a_i \quad (i = \overline{1, n})$ заданы.

Если $f(x_1, \ldots, x_n)$ имеет непрерывные частные производные f'_{x_i} по всем x_i, то приращение функции

$$\Delta f = \sum_{i=1}^{n} \Delta a_i \left| f'_{x_i}(x_1, \ldots, x_n) \right|_{x_i = a_i}. \tag{3.3}$$

Пример 2. Определить абсолютную погрешность функции $f(x,y) = x/y$ в точке $M(5,-2)$ в случае, если приращения аргументов ограничены величинами $\Delta x = 0{,}1$, $\Delta y = 0{,}2$.

Р е ш е н и е. Найдем частные производные функции в точке M:

$$f'_x = \frac{1}{y}\bigg|_{x=5,y=-2} = -\frac{1}{2}, \quad f'_y = -\frac{x}{y^2}\bigg|_{x=5,y=-2} = -\frac{5}{4}.$$

Абсолютная погрешность:

$$\Delta f\big|_M = \left|f'_x \Delta x + f'_y \Delta y\right|_{x=5,y=-2} = \left|-\frac{1}{2}\cdot 0{,}1 - \frac{5}{4}\cdot 0{,}2\right| = 0{,}6.$$

§ 3.2. ИТЕРАЦИОННЫЕ МЕТОДЫ

Многие задачи исследования функций одной переменной сводятся к необходимости определения аргумента из условия равенства двух функций.

$$f(x) = \varphi(x). \tag{3.4}$$

Функции в общем случае могут быть нелинейными, но при рассмотрении методов их решения, описанных ниже, предполагается, что обе функции, входящие в (3.4), непрерывны по крайней мере в окрестности того значения $x = x_*$, которое является решением (корнем уравнения). Отметим, что решению x_* уравнения соответствует на графике двух функций $y = f(x)$ и $y = \varphi(x)$ точка их пересечения.

Далеко не всегда подобные уравнения позволяют найти точное значение решения (или решений, если оно не одно).

Одним из способов решения (3.4) является *метод последовательных приближений* (*итераций*), заключающийся в следующем. В исходном, нулевом приближении аргументу x, стоящему в правой (или левой) части уравнения, придается некоторое конкретное значение $x_{(0)}$. По формуле $y = \varphi(x)$ находят то значение $y_{(0)}$, которое соответствует значению $x = x_{(0)}$. По найденному $y_{(0)}$ из уравнения $x = f^{(-1)}(y)$ (f^{-1} — функция, обратная к f) определяется новое значение $x_{(1)}$ первого приближения. Затем последовательно из *рекуррентных* зависимостей находим:

$$f(x_{(2)}) = \varphi(x_{(1)}), \quad f(x_{(3)}) = \varphi(x_{(2)}), \ldots$$
$$\ldots, f(x_{(n)}) = \varphi(x_{(n-1)}).$$

Процесс приближений ведется до достижения требуемой *точности вычисления* x. Точность может быть задана некоторой величиной ε и при выполнении условия

$$\frac{|x_{(n)} - x_{(n-1)}|}{\frac{1}{2}(|x_{(n)}| + |x_{(n-1)}|)} \le \varepsilon \tag{3.5}$$

$x_{(n)}$ считается равным искомому значению корня уравнения (3.4).

Пример. Две кривые (рис. 3.1) заданы уравнениями $y = \sqrt{x}$ и $y = 1/x$. Найти с точностью до 5% координаты точки пересечения этих кривых.

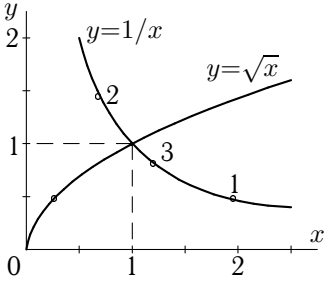


Рис. 3.1
Пример

Р е ш е н и е. Абсцисса точки пересечения кривых определится из равенства двух функций:

$$\sqrt{x} = \frac{1}{x}.$$

Полученное нелинейное уравнение, в котором $x \in (0; +\infty)$, перепишем в виде

$$x = \frac{1}{\sqrt{x}} \tag{3.6}$$

и применим для его решения итерационный метод. В исходном, нулевом приближении принимаем $x_{(0)} = 0{,}25$. Подставляя это значение x в левую часть (3.6), найдем значение x в

первом приближении:

$$x_{(1)} = \frac{1}{\sqrt{x_{(0)}}} = \frac{1}{\sqrt{0{,}25}} = 2.$$

Отношение приращения полученных значений переменной x к ее полусумме (относительная ошибка вычисления при условии, что в качестве решения уравнения будет принято значение $x_{(1)} = 2$) составит величину

$$\Delta_1 = \frac{|x_{(1)} - x_{(0)}|}{\frac{1}{2}(|x_{(1)}| + |x_{(0)}|)} = \frac{2 - 0{,}25}{\frac{1}{2}(2 + 0{,}25)} \approx 1{,}55,$$

то есть 155%, что не удовлетворяет требуемой условием задачи точности вычисления.

Переходим ко второму приближению. Для этого значение $x_{(1)} = 2$ подставляем в левую часть уравнения (3.6) и находим значение x во втором приближении:

$$x_{(2)} = \frac{1}{\sqrt{x_{(1)}}} = \frac{1}{\sqrt{2}} \approx 0{,}707.$$

Относительная ошибка второго приближения

$$\Delta_2 = \frac{|x_{(2)} - x_{(1)}|}{\frac{1}{2}(|x_{(2)}| + |x_{(1)}|)} = \frac{|0{,}707 - 2|}{\frac{1}{2}(0{,}707 + 2)} \approx 0{,}96$$

также не удовлетворяет требуемой условием задачи точности вычисления.

Результаты дальнейших приближений, полученных по рекуррентной формуле

$$x_{(i)} = \frac{1}{\sqrt{x_{(i-1)}}},$$

представлены в виде табл. 3.1.

Относительная ошибка вычисления сходящейся последовательности приведенных значений x в случае, если принять $x_* = x_{(7)} = 1{,}011$ не превышает величины 5% ($3{,}3 < 5$).

Таблица 3.1
Последовательные приближения

i	0	1	2	3	4	5	6	7
x_i	0,25	2	0,707	1,189	0,917	1,044	0,978	1,011
$\Delta_i\%$	—	155	96	50,5	26	13	6,5	3,3

Принимая $x_* = 1{,}011$, найдем $y_* = \dfrac{1}{x_*} = \dfrac{1}{1{,}011} \approx 0{,}989$. Значение y_* можно найти и по формуле $y = \sqrt{x}$:

$$y_* = \sqrt{x_*} = \sqrt{1{,}011} \approx 1{,}005.$$

Из-за того, что корень уравнения x_* определен приближенно, значения y_*, найденные по двум формулам, несколько отличаются друг от друга. Точное значение $y_* = 1$ находится между двумя определенными выше значениями этой величины. Аналогично, точное значение $x_* = 1$ находится между двумя последующими приближениями x. В процессе приближений переменные колеблются, постепенно «стягиваясь» к точному решению — точке (1;1).

§ 3.3. ПАУТИННАЯ МОДЕЛЬ

Описанный итерационный процесс хорошо интерпретируется на математической модели рынка.

Пусть предложение s некоторого товара на рынке описывается зависимостью от цены p:

$$s = p^2,$$

а спрос d на этот же товар зависимостью:

$$d = \frac{1}{p}.$$

На рис. 3.2 представлены кривые, соответствующие записанным функциям. Напомним, что в экономической литературе принято цену, являющуюся аргументом функций предложения и спроса, откладывать на графике по оси ординат.

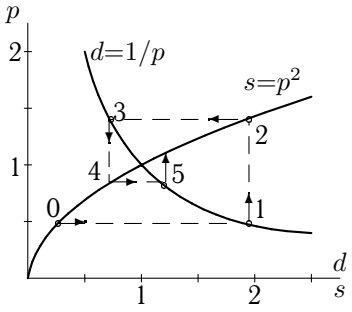


Рис. 3.2
Паутинная модель

Поэтому упомянутые функции, являясь обратными по отношению к функциям рассмотренного выше примера, изображаются на графиках рис. 3.2 такими же, как на рис. 3.1, кривыми. Будем, кроме того, считать, что p — это величина, равная отношению цены к некоторому ее среднему значению. Эта же оговорка справедлива по отношению к функциям d и s.

Обратимся к рыночным отношениям, вернувшись к условию рассматриваемой задачи. Пусть в начальной точке 0 рассматриваемого процесса регулирования рыночных отношений продавец выставил товар в количестве $s_{(0)} = 0{,}25$ единиц по цене $p_{(0)} = \sqrt{s_{(0)}} = \sqrt{0{,}25} = 0{,}5$. Координаты начальной точки $0(s_{(0)}; p_{(0)}) = 0(0{,}25; 0{,}5)$.

Цена $p_{(0)} = 0{,}5$ оказалась достаточно низкой для рыночных отношений. Этот вывод можно сделать по тому, что при цене $p_{(0)} = 0{,}5$ спрос на товар $d_1 = \dfrac{1}{p_{(0)}} = \dfrac{1}{0{,}5} = 2$ (точка 1 кривой $d(p)$) намного превышает количество представленных на рынок товаров $s_{(0)} = 0{,}25$. Отношения на рынке стали соответствовать точке 1 кривой спроса с координатами: $(d_{(1)}; p_{(0)}) = (2; 0{,}5)$.

Превышение спроса над предложением позволяет предпринимателю поднять цены на товары до уровня, где цена соответствует кривой предложения: $p_{(1)} = \sqrt{d_{(1)}} = \sqrt{s_{(2)}} = \sqrt{2} \approx 1{,}414$. Рыночные отношения переходят в точку 2, где $s_{(2)} = d_{(1)} = 2$, с координатами $(s_{(2)}; p_{(1)}) = (2; 1{,}414)$.

Следующим шагом саморегулирования рыночных отношений является переход в точку 3 с координатами $\left(d_{(3)} = \dfrac{1}{p_{(1)}} = \right.$ $\left. = \dfrac{1}{1{,}414} \approx 0{,}707\right)$ $(d_{(3)}; p_{(1)}) = (0{,}707; 1{,}414)$.

Полученные и последующие значения координат точек приведены в таблице.

$0(p_0; s_0) \Longrightarrow (0{,}5; 0{,}25)$	$1(p_0; d_1) \Longrightarrow (0{,}5; 2)$
$2(p_1; s_2) \Longrightarrow (1{,}414; 2)$	$3(p_1; d_3) \Longrightarrow (1{,}414; 0{,}707)$
$4(p_2; s_4) \Longrightarrow (0{,}841; 0{,}707)$	$5(p_2; d_5) \Longrightarrow (0{,}841; 1{,}189)$
$6(p_3; s_6) \Longrightarrow (1{,}891; 1{,}189)$	$7(p_3; d_7) \Longrightarrow (1{,}091; 0{,}917)$
$8(p_4; s_8) \Longrightarrow (0{,}958; 0{,}917)$	$9(p_4; d_9) \Longrightarrow (0{,}958; 1{,}044)$
$10(p_5; s_{10}) \Longrightarrow (1{,}022; 1{,}044)$	$11(p_5; d_{11}) \Longrightarrow (1{,}022; 0{,}979)$
$12(p_6; s_{12}) \Longrightarrow (0{,}989; 0{,}979)$	$13(p_7; d_{13}) \Longrightarrow (0{,}989; 1{,}011)$
$14(p_7; s_{14}) \Longrightarrow (1{,}005; 1{,}011)$	$15(p_7; d_{15}) \Longrightarrow (1{,}005; 0{,}995)$

Из рис. 3.2 видно, что последовательные шаги на рынке образуют сходящуюся к точке пересечения кривых прямоугольную спираль, своим видом напоминающую паутину. Отсюда модель получила свое название.

Паутинная модель экономики рынка, как видно из сравнения ее решения с рассмотренным выше примером, представляет собой итерационный метод решения уравнений.

При использовании этой модели следует иметь в виду, что сходимость итерационного процесса зависит от вида функций, точка пересечения графиков функций которых определяется.

Во многом сходимость процесса зависит от выбора начальной точки и направления движения к точке пересечения. Так, попытка прийти к искомому решению путем движения на паутинной модели по часовой стрелке (фиксация в нулевом приближении переменной $d_{(0)} = s_{(0)}$ и определения по $d_{(0)}$ переменной $p_{(1)} = \dfrac{1}{d_{(0)}} = \dfrac{1}{0{,}25} = 4$) приведет к тому, что итерационный процесс определения решения уравнения будет расходиться.

§ 3.4. ВЫБОР И ИСКЛЮЧЕНИЕ ИНТЕРВАЛОВ

Пусть задана непрерывная и дифференцируемая функция $y = f(x)$ и известно, что эта функция имеет единственный экстремум на интервале $[a, b]$. Такие функции называют *унимодальными*.

Для унимодальных функций сравнение значений функции на концах интервала $[a; b]$ и значений функций $y_1 = f(x_1)$ и $y_2 = f(x_2)$ внутри этого интервала позволяет уменьшить интервал изменения переменной, внутри которого находится экстремум.

Теорема. *Пусть функция $y = f(x)$ унимодальна на интервале $[a, b]$ и достигает минимума (максимума) в некоторой точке $x_* \in [a, b]$. Тогда из сравнения значений функции в двух точках x_1 и x_2 $(a < x_1 < x_2 < b)$ следует:*

1. если $y_1 > y_2$, то точка x_* не лежит в интервале $[a; x_1)$ $\big((x_2; b]$ для максимума$\big)$, а лежит в интервале $[x_1; b)$ $\big([a; x_2)$ для максимума$\big)$;

2. если $y_1 < y_2$, то точка x_* лежит (не лежит для максимума) в интервале $[a; x_2)$ $\big((x_1; b]$ для максимума$\big)$, и не лежит в интервале $[x_2; b)$ $\big([a; x_1)]$ для максимума$\big)$;

3. если $y_1 = y_2$, то точка x_* лежит в интервале $[x_1; x_2]$.

Точки $x_1, x_2, \ldots$, выбранные произвольно или следуя какому-либо алгоритму внутри интервала $[a; b]$, называют *пробными точками*.

На рис. 3.3,а и б изображены интервалы с минимальными значениями функций, соответствующие пунктам 1 и 2 теоремы.

Читателям предлагается изобразить аналогичные графики функций, имеющих максимум на $[a; b]$, а также график функций с максимумом и минимумом, соответствующий пункту 3 теоремы.

Используя положения теоремы, можно последовательно сужать интервал нахождения экстремума до тех пор, пока значения функции в двух внутренних точках станут достаточно близкими и, следовательно, примерно равными экстремальному значению.

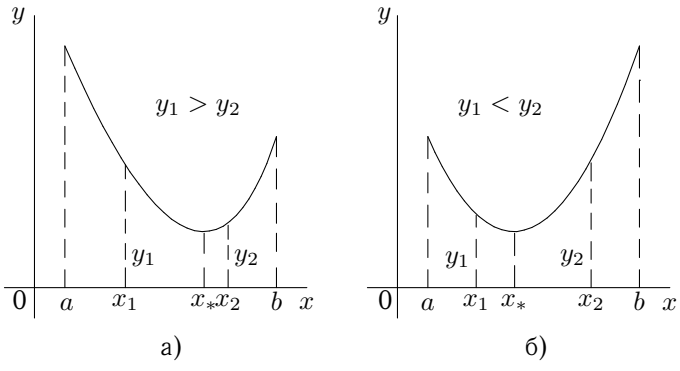


Рис. 3.3
К методу исключения интервалов

Суть описываемого метода, называемого методом *исключения интервалов,* заключается в том, что на каждом приближении исключается из рассмотрения тот интервал изменения функции, на котором, согласно теореме, отсутствует экстремум. На оставшемся интервале выбирается дополнительная точка (удобно ее выбирать в середине одного из оставшихся подынтервалов), и снова производится сравнение значений функции во внутренних точках интервала.

Метод исключения интервалов удобен тем, что при его использовании не привлекаются производные функции (использование производных в численном анализе может привести к существенному снижению точности вычислений).

При формулировке теоремы этого раздела предполагалось, что интервал $[a, b]$, на котором находится один экстремум функции, найден. Оказывается, что изложенный метод исключения интервалов может быть привлечен и к установлению границ начального интервала.

Опишем процедуру установления такого интервала на примере функции, имеющей минимум справа от некоторой фиксированной координаты $x = x_0$. Для случая максимума все рассуждения аналогичны и будут отличаться лишь знаками неравенств.

Заметим сразу, что при принятых условиях функция, по крайней мере в окрестности точки x_0, справа от нее убывает.

Выберем справа от точки x_0 две пробные точки x_1 и $x_2 > x_1$ и сравним значения функции в этих точках. Как и в изложенной выше теореме, при этом могут иметь место три ситуации:

1. если $y_1 > y_2$, то минимума на интервале $[x_0; x_1]$ нет;

2. если $y_1 < y_2$, то минимум находится на интервале $[x_0; x_2]$;

3. если $y_1 = y_2$, то минимум находится на интервале $[x_1; x_2]$.

В случаях 2 и 3 указанные концы интервалов принимают за границы начального интервала $[a, b]$ и далее решают задачу отыскания минимума функции на этом интервале, опираясь на теорему 1.

Для ситуации пункта 1 исключают из рассмотрения интервал $[x_0; x_1)$, выбирают следующую пробную точку $x_3 > x_2$ и для трех точек x_1, x_2 и x_3 проводят анализ, следуя пунктам 1–3. И так до тех пор, пока не будет найден интервал с минимумом функции.

Значения координаты каждой последующей пробной точки удобно выбирать таким образом, чтобы следовать определенной логике, которая необходима при составлении программ для ЭВМ. Проще всего выбирать значения координат, отстоящих друг от друга на равные расстояния. Есть и другие алгоритмы. Можно, например, координаты каждой последующей пробной точки x_{n+1} определять по координатам предыдущей точки x_n и некоторым постоянным положительным параметрам a и d:

$$x_{n+1} = x_n + a^n d. \tag{3.7}$$

При $a > 1$ прирост координат пробных точек на каждой итерации увеличивается, при $a < 1$ уменьшается.

Пример. Определить интервал, на котором функция $y = (100 - x)^2$ принимает минимальное значение.

Р е ш е н и е. Будем определять координаты пробных точек интервала по формуле (3.7), для чего примем $x_0 = 50$;

$a = 2$, $d = 5$. Тогда $x_1 = x_0 + 2^0 \cdot d = 50 + 5 = 55$; $x_2 = x_1 + 2^1 \cdot d = 55 + 2 \cdot 5 = 65$.

Найдем значения функции в точках x_0, x_1, x_2:

$y_0 = f(50) = (100 - 50)^2 = 2500$;

$y_1 = f(55) = (100 - 55)^2 = 2025 < y_0$;

$y_2 = f(65) = (100 - 65)^2 = 1225 < y_1$.

Сравнение значений функций, согласно изложенной выше схеме, указывает на то, что для выявления интервала, содержащего минимум функции, необходимо рассмотреть следующую пробную точку. Значения ее координат:

$$x_3 = x_2 + 2^2 d = 65 + 4 \cdot 5 = 85;$$

$$y_3 = f(85) = (100 - 85)^2 = 225 < y_2.$$

Для $x_4 = x_3 + 2^3 d = 85 + 8 \cdot 5 = 125$:

$$y_4 = f(125) = (100 - 125)^2 = 625 > y_3.$$

В точке x_4 функция возросла по сравнению с точкой x_3, поэтому минимум следует искать на интервале $[x_2; x_4] = [65; 125]$.

Определение интервала нахождения минимума функции состоялось. Для построения алгоритма уменьшения интервала рассмотрим метод, основанный на понятии золотого сечения.

Предварительно, возвращаясь к примеру, заметим, что в точке $x = 100$ $f(100) = 0 < y_3$. Поэтому для дальнейших поисков точки минимума функции можно ограничиться интервалом $[a; b] = [85; 125]$.

§ 3.5. МЕТОД ЗОЛОТОГО СЕЧЕНИЯ

Пусть известно, что функция $y = f(x)$ — унимодальна на интервале $[a; b]$, то есть имеет на этом интервале единственный экстремум. Для удобства его отыскания введем безразмерную координату $\omega = \dfrac{x - a}{b - a}$, которая изменяется в пределах от нуля до единицы. Действительно, $\omega_a = 0$ (при $x = a$) и $\omega_b = 1$ (при $x = b$).

При введении новой переменной $y = f(x)$ становится сложной функцией от ω, так как $x = a + (b - a)\omega$. Будем считать, что $y = g(\omega)$.

На интервале единичной длины (рис. 3.4) расположим две пробные точки, отстоящие от граничных точек интервала $\omega_a = 0$ и $\omega_b = 1$ на расстоянии ζ: $\omega_1 = \zeta$ и $\omega_2 = 1 - \zeta$.

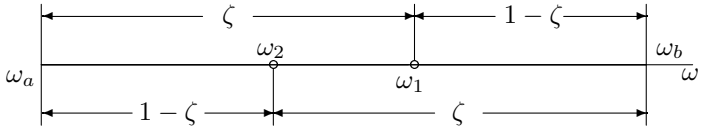


Рис. 3.4
Метод золотого сечения

При таком симметричном расположении пробных точек длина отрезка, остающегося после исключения подынтервала длины $(1 - \zeta)$, равна ζ и не зависит от того, какой именно подынтервал исключается из рассмотрения.

Остающийся отрезок содержит одну пробную точку, расположенную на расстоянии $(1 - \zeta)$ от границы отрезка, противоположной отброшенному подынтервалу.

Для того чтобы симметрия поискового процесса сохранялась, расстояние $(1 - \zeta)$ должно составлять ζ-ю часть длины оставшегося подынтервала, длины ζ. То есть должна иметь место пропорция:

$$\frac{\omega_2 - \omega_a}{\omega_1 - \omega_a} = \frac{\omega_1 - \omega_a}{\omega_b - \omega_a}, \quad \text{или} \quad \frac{1 - \zeta}{\zeta} = \frac{\zeta}{1} = \zeta, \tag{3.8}$$

или

$$\zeta^2 + \zeta - 1 = 0.$$

Решение уравнения:

$$\zeta = (-1 \pm \sqrt{5})/2.$$

Отрицательное значение корня лишено смысла, а положительное: $\zeta \approx 0,618$.

Итак, координаты пробных точек: $\omega_1 = \zeta \approx 0,618$, $\omega_2 = 1 - \zeta = \zeta^2 \approx 0,382$.

Следующая пробная точка ω_3 размещается на отрезке $[\omega_2, \omega_1]$ на расстоянии, составляющем ζ-ю часть этого отрезка и откладываемом от одной из его граничных точек ω_1 или ω_2. При этом сохраняем пропорции, имевшие место для пробных точек ω_1 и ω_2. При соотношении $\omega_3 - \omega_2 > \omega_1 - \omega_3$ отношение большего отрезка во всему отрезку:

$$\frac{\omega_3 - \omega_2}{\omega_1 - \omega_2} = \frac{\omega_1 - \omega_a}{\omega_b - \omega_a} = \zeta.$$

Отсюда следует

$$\omega_3 - \omega_2 = \zeta(\omega_1 - \omega_2) = \zeta(\zeta - \zeta^2) = \zeta^2(1 - \zeta) = \zeta^3.$$

Длина меньшего отрезка $[\omega_3, \omega_1]$ определится из соотношения

$$\frac{\omega_1 - \omega_3}{\omega_3 - \omega_2} = \zeta,$$

отсюда при $\omega_3 - \omega_2 = \zeta^3$ получаем

$$\omega_1 - \omega_3 = \zeta^4.$$

Таким образом, на любом подынтервале разбиваемом пробной точкой на отрезки:

1) отношение длины меньшего отрезка к длине большего отрезка равна ζ;

2) длина каждого предыдущего подынтервала отличается от «вложенного в него» последующего подынтервала множителем ζ.

Если подынтервал $[\omega_{n-1}, \omega_{n-2}]$, $(n = 1, 2, \ldots)$ точкой ω_n делится в отношении ζ, то при $\omega_n - \omega_{n-2} > \omega_{n-1} - \omega_n$ на некотором этапе разбиения отрезка получаем соотношения:

$$\omega_n - \omega_{n-2} = \zeta^n, \qquad \omega_{n-1} - \omega_n = \zeta^{n+1}.$$

Схема поиска требуемых значений функции, в частности ее экстремумов, основанная на изложенном методе симметричного последовательного разделения интервалов в отношении ζ, носит название *метода золотого сечения*.

Пример. Найти минимальное значение функции $y = (100 - x)^2$, для которой в предыдущем разделе установлен начальный интервал $[85; 125]$ изменения x, внутри которого располагается экстремум.

Р е ш е н и е.

1. Переходим к безразмерной координате $\omega \in [0, 1]$:

$$\omega = \frac{x - a}{b - a} = \frac{x - 85}{125 - 85} = \frac{1}{40}(x - 85).$$

Отсюда

$$x = 85 + 40\omega$$

и

$$y = f(x) = (100 - x)^2 = (100 - 85 - 40\omega)^2 = \\ = (15 - 40\omega)^2 = g(\omega).$$

2. Организуем, согласно методу золотого сечения, алгоритм дробления отрезка $\omega \in [0, 1]$: $\omega_1 = \zeta = 0,618$; $\omega_2 = \zeta^2 = 0,382$.

На рис. 3.5 нанесены эти и последующие точки уменьшаемых от приближения к приближению интервалов. Каждый последующий интервал выделен дополнительной прямой, параллельной оси ω.

3. Определяем значения $y = g(\omega)$ в точках ω_1 и ω_2 и сравниваем их значения. При этом имеем в виду, что значения функции y на концах исходного интервала известны (были определены в предыдущем разделе $y_a = g(0) = 225$, $y_b = g(1) = 625$), и известно, что функция имеет минимум на интервале:

$y_1 = g(\omega_1) = (15 - 40 \cdot 0{,}618)^2 \approx 94{,}48$;

$y_2 = g(\omega_2) = (15 - 40 \cdot 0{,}372)^2 \approx 0{,}01 < y_1$.

Согласно пункту 2 теоремы предыдущего раздела об интервалах, приходим к выводу о том, что минимум функции лежит в интервале $[0; \omega_1] = [0; 0{,}618]$.

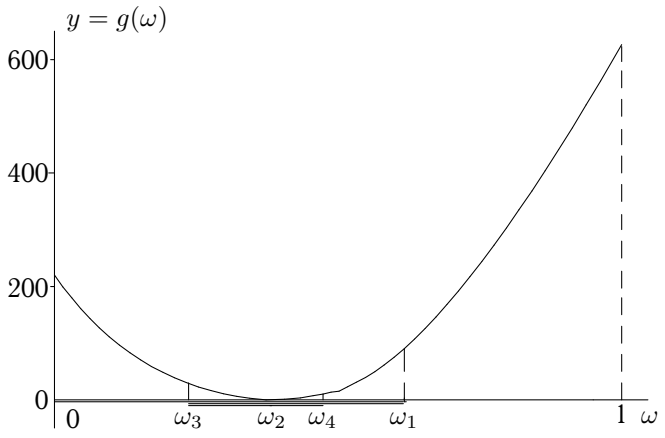


Рис. 3.5
Итерации метода золотого сечения

4. Определяем координату пробной точки на следующем приближении:

$$\omega_3 = \zeta^3 = 0{,}618^3 \approx 0{,}236$$

и находим

$$y_3 = g(\omega_3) = (15 - 40 \cdot 0{,}236)^2 \approx 30{,}91 > y_2.$$

Поэтому минимум находится на интервале $[\omega_3; \omega_1] = [0{,}236; 0{,}618]$.

5. Переходим к следующему приближению:

$$\zeta^4 = 0{,}618^4 \approx 0{,}146$$

и находим

$$\omega_4 = \omega_1 - \zeta^4 = 0{,}618 - 0{,}146 = 0{,}472.$$

Последнее вычисление требует разъяснений.

При изложении сути метода золотого сечения было сказано, что на каждом приближении новые точки разбиения выбираются как ζ-я доля сформированного к рассматриваемому приближению интервала. Величина остающегося на n-ном приближении интервала $\omega_{n-1} - \omega_n = \zeta^n$.

На рассматриваемом приближении установлено, что подынтервал $[\omega_4;\omega_1]$, который следует исключить из интервала $[\omega_3;\omega_1]$, отсчитывается со стороны координаты ω_1 (рис. 3.5). Поэтому подынтервал следующего приближения получился из подынтервала предыдущего путем вычитания из координаты ω_1 подынтервала рассматриваемого приближения, равного ζ^4.

$y_4 = g(\omega_4) = (15 - 40 \cdot 0{,}472)^2 \approx 15{,}05 > y_2$.

Минимум находится на подынтервале $[\omega_3;\omega_4] = [0{,}236;0{,}472]$.

И так далее. Процесс приближений сходится к найденной из аналитического решения точке $\omega = 0{,}375$, где $y = y_{\min} = 0$.

Основным преимуществом описанного метода перед рассматриваемыми ниже методами является то, что при его использовании нет необходимости вычислять производные функции.

Метод золотого сечения позволяет облечь в определенные правила назначение промежуточных пробных точек и таким образом определить алгоритм их нахождения на каждом шаге вычисления экстремума функции.

Ясно, что можно предложить сколь угодно много способов (методов) выбора промежуточных точек. Простейшим из них, но не самым эффективным, считается метод половинного деления. В этом методе интервал изменения аргумента на каждом последующем шаге отыскания экстремума функции делится пробной точкой пополам.

§ 3.6. МЕТОД ХОРД

Одним из самых эффективных численных методов определения корней уравнений является метод хорд и касательных.

Считаем, что для заданной непрерывной и дифференцируемой функции $y = f(x)$ выделен промежуток $[a;b]$, на котором функция имеет единственный корень.

Рассмотрим сначала *метод хорд*.

Определив значения функции на концах промежутка $[a,b]$ $\big(y_a = f(a),\ \ y_b = f(b)\big)$, запишем уравнение хорды — прямой AB (рис. 3.6,а), проходящей через точки графика функции,

соответствующие концам промежутка:

$$y - y_a = \frac{y_b - y_a}{b - a}(x - a). \tag{3.9}$$

Найдем координату x_1 точки пересечения этой прямой с осью абсцисс. Из уравнения (3.9) при $y = 0$ получим:

$$x_1 = a - \frac{b - a}{y_b - y_a} y_a.$$

Координата x_1 пересечения хорды AB с осью абсцисс является приближенным значением корня на первой итерации, которая на этом и заканчивается. Чтобы уточнить значение корня, необходимо продолжить итерационный процесс.

Для решения задачи на второй итерации определим на графике (рис. 3.6,а) функции ординату точки A_1, соответствующую x_1: $y_1 = f(x_1)$. Через точки A_1 и B проводим новую хорду, находим координату x_2 ее пересечения с осью абсцисс. По ней определяем ординату $y_2 = f(x_2)$ точки A_2 и т. д.

Значение корня уравнения на n-ной итерации определится по формуле

$$x_n = x_{n-1} - \frac{b - x_{n-1}}{y_b - y_{n-1}} y_{n-1}. \tag{3.10}$$

Процесс итераций завершается, когда разность приближенных значений корней $(x_n - x_{n-1})$ становится меньше заданной точности решения задачи. В этом случае считаем, что корень уравнения $f(x) = 0$: $x_0 \approx x_n$.

Процесс последовательных приближений к точному решению x_0 рассмотрен на примере выпуклой на отрезке $[a, b]$ функции. В этом случае приближение к x_0 шло от точки A к точке B.

Если функция вогнута (рис. 3.6,б), то приближение к точному решению будет идти справа налево. При этом значение корня уравнения на n-ной итерации:

$$x_n = x_{n-1} - \frac{x_{n-1} - a}{y_{n-1} - y_a} y_{n-1}.$$

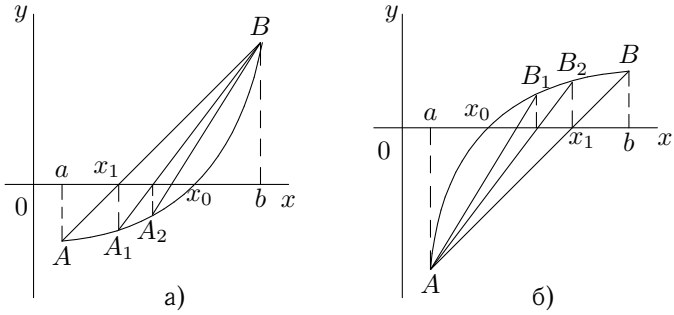


Рис. 3.6
Метод хорд

В случае, если функция внутри промежутка $[a, b]$ выпукло-вогнутая, направления приближений к точному решению могут в процессе итераций изменяться.

Читателю предлагается изобразить график такой функции и убедиться в справедливости сказанного.

§ 3.7. МЕТОД КАСАТЕЛЬНЫХ

Еще одним итерационным методом отыскания корней уравнения непрерывной и дифференцируемой функции является *метод касательных*. Он заключается в следующем.

Через точку A графика функции $y = f(x)$, соответствующую началу интервала на рис. 3.7,а (точку B — конца интервала на рис. 3.7,б), проводится касательная к кривой.

Уравнение касательной к графику функции — это уравнение прямой, проходящей через заданные точки $A(a, y_a)$ или $B(b, y_b)$ в заданных направлениях, определяемых угловыми коэффициентами (тангенсами углов наклона касательных к оси абсцисс) $k_A = y'(a)$ или $k_B = y'(b)$:

$$y - y_a = y'(a)(x - a) \quad \text{или} \quad y - y_b = y'(b)(x - b). \qquad (3.11)$$

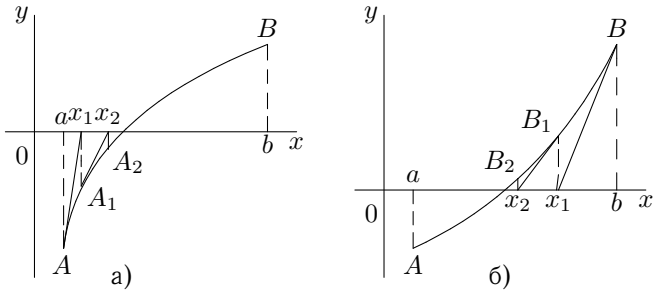


Рис. 3.7
Метод касательных

Находим точку пересечения выбранной прямой (3.11) с осью абсцисс, приравнивая нулю ординату y:

$$x_1 = a - \frac{y_a}{y'(a)} \quad \text{или} \quad x_1 = b - \frac{y_b}{y'(b)}. \tag{3.12}$$

Найденное по одной из этих формул значение координаты x_1 является приближенным значением искомого корня на первой итерации. Для его более точного определения переходим ко второй итерации. Для этого находим координату $y_1 = f(x_1)$ точки A_1 (или B_1), проводим через эту точку касательную к графику функции, находим координату x_2 ее пересечения с осью абсцисс, являющуюся значением корня на второй итерации и так далее.

На n-ной итерации значение корня определится из формулы:

$$x_n = x_{n-1} - \frac{y_{n-1}}{y'(x_{n-1})}. \tag{3.13}$$

Точность вычисления, как и в методе хорд, можно определить по разности значений корней, найденных в двух последовательных приближениях.

Отметим важную особенность метода касательных. Из рис. 3.7 видно, что начальные касательные к графикам функций проводились в тех граничных точках промежутка $[a, b]$, где выпуклость графика функции направлена в сторону оси

абсцисс. Если этого условия не придерживаться и провести касательную к графику функции на другом конце промежутка (в точке B для рис. 3.7,а или в точке A для рис. 3.7,б), то такая касательная может пересечь ось абсцисс далеко за пределами исследуемого интервала. Если продолжить итерационный процесс при таком выборе начальных точек, то он, как правило, не будет сходиться к точному решению.

Для функций, которые изменяют внутри промежутка $[a, b]$ выпуклость так, что на всем промежутке кривая графика функции направлена вогнутостью к оси абсцисс (на оси располагается точка перегиба), применение метода касательных для отыскания корней вообще неприемлемо.

Предлагаем читателю проанализировать описанные ситуации на графиках функций.

Вспоминая связь второй производной с выпуклостью функции, можно записать правило выбора начальной точки для организации итерационного процесса по методу касательных.

Правило. *Начальная точка для организации итерационного процесса по методу касательных выбирается таким образом, чтобы в этой точке знаки функции и второй производной совпадали.*

§ 3.8. МЕТОД ХОРД И КАСАТЕЛЬНЫХ

На практике целесообразно применять методы хорд и касательных одновременно, чередуя их в процессе итераций.

Это во многих случаях дает возможность приближаться к точному решению сразу с двух сторон числовой оси.

Пример. Найти экстремальное значение функции

$$g(x) = \frac{1}{2}x^4 + x^3 + x^2 + x + \frac{3}{2},$$

если известно, что оно находится на промежутке $[-1{,}5; 0]$.

Р е ш е н и е. Найдем производную заданной функции и приравняем ее нулю (для удобства записи обозначим эту производную через y):

$$y = g'(x) = f(x) = 2x^3 + 3x^2 + 2x + 1 = 0.$$

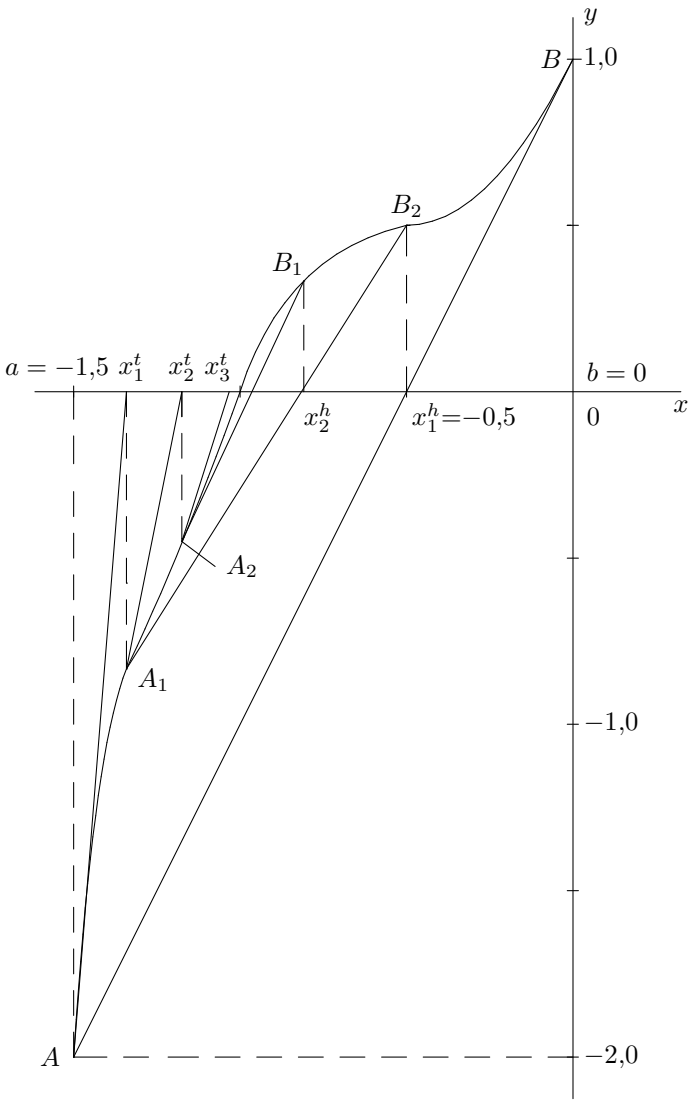


Рис. 3.8
Пример: метод хорд и касательных

Таким образом, задача отыскания экстремального значения функции свелась к задаче определения корня алгебраического уравнения на заданном интервале изменения функции.

Для приближенного анализа характера поведения функции y определим ее первую и вторую производные:

$$y' = 6x(x+1) + 2; \qquad y'' = 6(2x+1).$$

Первая производная функции положительна на промежутке $[-1{,}5; 0]$, причем $y'(-1{,}5) = 6{,}5;\ y'(0) = 2$. По поведению производной устанавливаем, что функция монотонно возрастает на рассматриваемом промежутке и ее производная нигде не обращается в нуль. То есть точек экстремума функция не имеет.

Что касается исходной функции $g(x)$, то производная от y является для нее второй производной. А так как $g''(x) = y'$ положительна на всем интервале $[-1{,}5;\ 0]$, то функция $g(x)$ имеет на этом интервале минимум.

Вторая производная $y''(x)$ обращается в нуль в точке $x{=}{-}0{,}5 \in [-1{,}5; 0]$. Кривая в этой точке имеет перегиб, где она переходит от выпуклости слева от точки $\big(y''(-1{,}5) = -12 < 0\big)$ к вогнутости — справа от точки $\big(y''(0) = 6 > 0\big)$.

Знаки функции на концах заданного промежутка: $\big(y(-1{,}5) = -2 < 0;\ \ y(0) = 1 > 0\big)$ совпадают с соответствующими знаками вторых производных. Поэтому метод касательных можно применить для нахождения корня, опираясь при этом на любую из граничных точек функции или попеременно на обе. Будем проводить касательные к графику функции с левой его стороны.

П е р в о е п р и б л и ж е н и е.

Метод хорд. Через точки A и B проводим хорду (3.9) и по формуле (3.10) определяем координату точки ее пересечения с осью абсцисс:

$$x_1^h = a - \frac{b-a}{y(b)-y(a)}y(a) = -1{,}5 - \frac{0-(-1{,}5)}{1-(-2)} \cdot (-2) = -0{,}5.$$

Метод касательных. Через точку A проводим касательную (3.11) к графику функции и по формуле (3.12) определяем

координату точки ее пересечения с осью абсцисс:

$$x_1^t = a - \frac{y(a)}{y'(a)} = -1{,}5 - \frac{-1}{6{,}5} \approx -1{,}346.$$

Интервал $[-1{,}346; -0{,}5]$ является новым, полученным на первой итерации промежутком, внутри которого находится корень уравнения. Этот интервал, как и следовало ожидать, меньше первоначального $\big([-1{,}5; 0]\big)$, но не обеспечивает достаточную точность определения корня. Поэтому переходим ко второму приближению.

Второе приближение.

Находим ординаты точек графика функции, соответствующих концам найденного в первом приближении промежутка, содержащего корень уравнения: $y(x_1^t) = y(-1{,}346) = 2(-1{,}346)^3 + 3(-1{,}346)^2 + 2(-1{,}346) + 1 \approx -0{,}83$; $y(x_1^h) = y(-0{,}5) = 0{,}5$.

Таким образом, установлены координаты двух новых точек на графике функции: $A_1(-1{,}346; -0{,}83)$ и $B_1(-0{,}5; 0{,}5)$.

Метод хорд. Через точки A_1 и B_1 проводим хорду и по формуле (3.10) определяем координату точки ее пересечения с осью абсцисс:

$$x_2^h = x_1^t - \frac{x_1^h - x_1^t}{y(x_1^h) - y(x_1^t)} y(x_1^t) \approx -0{,}81.$$

Метод касательных. Через точку A_1 проводим касательную к графику функции и по формуле (3.13) определяем координату точки ее пересечения с осью абсцисс:

$$x_2^t = x_1^t - \frac{y(x_1^t)}{y'(x_1^t)} \approx -1{,}17.$$

Таким образом, промежуток, в котором находится корень, снова уменьшился и стал равным $[-1{,}17; -0{,}81]$.

После третьего приближения получим интервал $[-1{,}035; -0{,}985]$, в который заключен корень исходного уравнения.

Точное решение уравнения дает значение корня $x_0 = -1$. Разность между значением любой точки, взятой из полученного в третьем приближении интервала, и точным решением

отличается от последнего вторым знаком после запятой, что вполне приемлемо для многих реальных задач.

Относительная точность решения обычно определяется по формуле:

$$\varepsilon_n = \frac{\big||x_n^{\max}| - |x_n^{\min}|\big|}{\big(|x_n^{\max}| + |x_n^{\min}|\big)/2}.$$

После третьего ($n = 3$) приближения:

$$\varepsilon_3 = \frac{\big||-0{,}985| - |-1{,}035|\big|}{\big(|-0{,}985| + |-1{,}035|\big)/2} = \frac{0{,}05}{1{,}01} \approx 0{,}049,$$

то есть ошибка не превышает 5%.

Приняв корень уравнения $x_* = -1$ и подставляя его значение в исходное уравнение, получим

$$g_{\min} = g(-1) = 1.$$

§ 3.9. СИМПЛЕКСНЫЙ МЕТОД ПОИСКА ЭКСТРЕМУМА

Перейдем к рассмотрению методов приближенного вычисления экстремумов функций нескольких переменных. Для простоты и наглядности изложения материала ограничимся функцией двух переменных.

Одним из простейших методов поиска экстремума является метод симплексов (от английского *simple* — простой).

Регулярный симплекс в двухмерном пространстве представляет собой равносторонний треугольник (в трехмерном пространстве — правильный тетраэдр).

Симплексный метод заключается в следующем.

Предположим, что некоторая непрерывная и дифференцируемая функция двух переменных $u = f(x, y)$ имеет локальный экстремум в некоторой точке $M_*(x_*, y_*)$.

Рассмотрим на координатной плоскости xy проекцию области функции $u = f(x, y)$, прилегающую к точке M_*. На рис. 3.9 показаны симплексы, координаты которых соответствуют разобранному в этом разделе примеру (ниже).

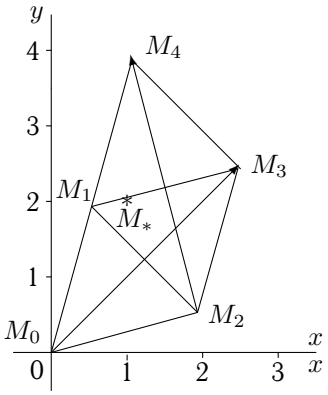


Рис. 3.9
Симплексный метод

В произвольно выбранной точке M_0 области определения функции строится исходный равносторонний треугольник $M_0M_1M_2$ — первый симплекс. О способе построения симплексов будет сказано ниже.

Определяются значения функции $u = f(x, y)$ в каждой из вершин построенного треугольника. При определении локального минимума (максимума) для точки вершины, в которой функция имеет наибольшее (наименьшее) из трех вершин значение функции (пусть это будет M_0) определяется симметричная ей относительно прямой M_1M_2 точка M_3. Точки M_1, M_2 и M_3 образуют новый симплекс-треугольник.

Далее находится значение функции в точке M_3, выявляется вершина треугольника $M_1M_2M_3$, имеющая наибольшее (наименьшее) значение функции. Определяется точка M_4, симметричная найденной точке с максимальным (минимальным) значением функции (на рисунке M_3) нового симплекс-треугольника. И так далее до тех пор, пока очередной треугольник не накроет точку минимума (максимума) функции.

Критерием того, что треугольник накрыл точку экстремума, является увеличение (уменьшение для максимума) значения функции в точках следующих треугольников.

Опишем один из способов построения симплекс-треугольников.

Исходная точка M_0 выбирается произвольно (с учетом опыта и интуиции расчетчика). Координаты точек $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ можно определить, например, по формулам:

$$\begin{cases} x_1 = x_0 + p; \\ y_1 = y_0 + q; \end{cases} \qquad \begin{cases} x_2 = x_0 + q; \\ y_2 = y_0 + p. \end{cases} \tag{3.14}$$

Параметры p и q выбраны так, чтобы точки M_0, M_1 и M_2 являлись вершинами равностороннего треугольника, в котором сторона M_1M_2 является отрезком прямой, отсекающей на координатных осях равные отрезки. Характерный размер треугольника определяется параметром $d > 0$. Перечисленные свойства приводят к формулам:

$$p = \frac{\sqrt{3}-1}{2\sqrt{2}}d; \qquad q = \frac{\sqrt{3}+1}{2\sqrt{2}}d. \tag{3.15}$$

При $d = 1$ стороны симплекс-треугольников $l = \sqrt{p^2+q^2} = 1$. С увеличением d стороны треугольника увеличиваются и уменьшается, в общем, количество приближений, необходимых для достижения конечной цели — накрытия точки экстремума. Но при этом уменьшается точность решения. При уменьшении d наблюдается обратная картина.

Записанные для определения p и q формулы (3.15) являются частным (при n=2) случаем общих для n-мерного пространства соотношений:

$$p = \left(\frac{\sqrt{n+1}-1}{n\sqrt{2}}\right)d; \quad q = \left(\frac{\sqrt{n+1}+n-1}{n\sqrt{2}}\right)d. \tag{3.16}$$

В аналитической геометрии показано, что координаты точки M_3, симметричной вершине M_0 равностороннего треугольника относительно стороны M_1M_2, определяются по формулам:

$$x_3 = x_1 + x_2 - x_0; \quad y_3 = y_1 + y_2 - y_0. \tag{3.17}$$

Пример. Найти минимальное значение функции:

$$u = (1-x)^2 + (2-y)^2,$$

если известно, что точка (0,0) входит в область локального минимума этой функции.

Р е ш е н и е. Примем точку $M_0(0,0)$ за исходную для построения начального симплекса и задаемся величиной $d = 2$.

По формулам (3.15) (для $n = 2$) находим:

$$p = \frac{\sqrt{3}-1}{2\sqrt{2}}2 \approx 0{,}518; \quad q = \frac{\sqrt{3}+1}{2\sqrt{2}}2 \approx 1{,}932.$$

Используя эти параметры, вычислим координаты двух других вершин симплекса в соответствии с (3.14):

$$x_1 = x_0 + p = 0 + 0{,}518 = 0{,}518; \;\; y_1 = y_0 + q = 0 + 1{,}932 = 1{,}932;$$

$$x_2 = x_0 + q = 0 + 1{,}932 = 1{,}932; \;\; y_2 = y_0 + p = 0 + 0{,}518 = 0{,}518.$$

Найдем значения функции в вершинах исходного симплекс-треугольника:

$$u(M_0) = u(0;0) = (1-0)^2 + (2-0)^2 = 5;$$
$$u(M_1) = u(0{,}518;1{,}932) = (1-0{,}518)^2 + (2-1{,}932)^2 \approx 0{,}237;$$
$$u(M_2) = u(1{,}932;0{,}518) = (1-1{,}932)^2 + (2-0{,}518)^2 \approx 3{,}066.$$

Так как наибольшее из полученных значений функция имеет в точке M_0, то вершину M_3 следующего симплекс-треугольника выберем симметричной M_0 относительно прямой M_1M_2.

По формулам (3.17) находим:

$$x_3 = x_1 + x_2 - x_0 = 0{,}518 + 1{,}932 - 0 = 2{,}450;$$
$$y_3 = y_1 + y_2 - y_0 = 1{,}932 + 0{,}518 - 0 = 2{,}450.$$

В полученной точке M_3 значение функции

$$u(M_3) = u(2{,}45;2{,}45) = (1-2{,}45)^2 + (2-2{,}45)^2 \approx 2{,}303.$$

Сравнивая значения функции в трех вершинах нового симплекс-треугольника $M_1M_2M_3$, устанавливаем, что ее

максимальное значение имеет вершина M_2. При построении следующего симплекс-треугольника симметрично именно этой точке относительно прямой M_1M_3 следует искать новую вершину M_4 с координатами:

$$x_4 = x_1 + x_3 - x_2 = 0{,}518 + 2{,}45 - 1{,}932 = 1{,}036;$$
$$y_4 = y_1 + y_3 - y_2 = 1{,}932 + 2{,}45 - 0{,}518 = 3{,}864.$$

Значения функции в этой точке:

$$u(M_4) = u(1{,}036; 3{,}864) = (1 - 1{,}036)^2 + (2 - 3{,}864)^2 \approx 3{,}475.$$

Это значение больше, чем значения функции в любой из вершин симплекс-треугольника $M_1M_2M_3$. Такая ситуация говорит о том, что точка экстремума находится внутри указанного треугольника и продолжение приближений лишено смысла.

В общем, если процесс приближений был направлен на отыскание минимального значения функции, за ее значение можно принять наименьшее из значений функции в вершинах треугольника. То есть положить $u_{\min} \approx u(M_1) = 0{,}232$.

Точное минимальное значение функции, найденное аналитически $u_{\min} = u(1; 2) = 0$.

Замечание. Параметр d, отвечающий за размер симплекс-треугольника, можно изменять в процессе приближений. В частности, при приближении к точке экстремума этот параметр, как правило, уменьшают для достижения большей точности вычислений.

Изменяют этот параметр и в ситуациях, когда симплекс-треугольники выстраиваются в «циклические кольца», замыкающиеся вокруг точки экстремума.

Симплексный метод отыскания экстремумов функции нескольких переменных (ФНП) отличается простотой и не требует вычисления производных. Как и методы пробных точек для функций одной переменной, этот метод позволяет накрывать и уменьшать область определения функции, стягивая ее к стационарной точке. Симплексный метод обладает простотой и наглядностью, но зачастую, не отличается быстротой сходимости и устойчивостью решения — итерационный

процесс может не сходиться (образуются циклические кольца из симплекс-треугольников).

Лучшие результаты дают методы, использующие производные от функции (частные производные для функций нескольких переменных).

Наиболее распространенным из таких методов является *метод градиентного спуска,* называемый часто *методом Коши* (Cauchy Augustin-Louis (1789–1857) — великий французский математик и механик).

§ 3.10. МЕТОД ГРАДИЕНТНОГО СПУСКА

Известно, что вектор-градиент функции u (∇u) — это вектор, направленный в сторону ее возрастания, ортогональный линии уровня функции $u = f(x, y)$ и, кроме того, лежит в плоскости, касательной к поверхности, изображающей функцию.

Существуют методы градиентного спуска, основанные на использовании частных производных первого порядка, второго порядка и т. д. Как и любой численный метод, этот метод использует последовательные приближения (итерации) к точному решению. Остановимся на описании лишь градиентного метода первого порядка.

Пусть $\boldsymbol{X}_k = (x_k,\ y_k)$ — вектор, характеризующий положение точки M_k на плоскости координат (x, y) на k-той итерации (рис. 3.10,а).

Будем строить итерационный процесс в предположении, что в следующее положение M_{k+1}, характеризуемое вектором $\boldsymbol{X}_{k+1}$, точка переводится вектором градиентом $\boldsymbol{\nabla} u(X_k)$, умноженным на некоторый корректирующий множитель α_k. Тогда вектор, определяющий положение M_{k+1}, точки экстремума на $(k + 1)$ итерации, определится по рекуррентной формуле:

$$\boldsymbol{X}_{k+1} = \boldsymbol{X}_k + \alpha_k \boldsymbol{\nabla} u(X_k). \tag{3.18}$$

Величина множителя α_k в каждом приближении подсчитывается, исходя из требования экстремальности функции.

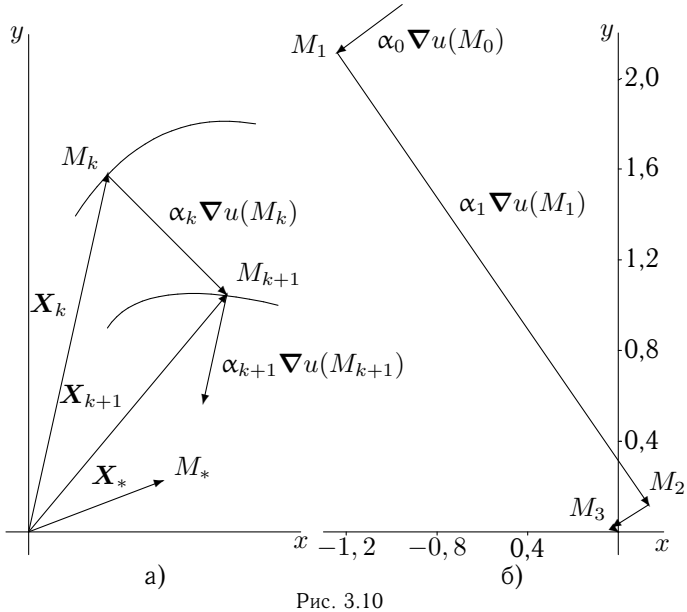


Рис. 3.10
Итерации метода градиентного спуска

Рассмотрим организацию итерационного цикла метода градиентного спуска на примере.

Пример. Найти минимальное значение функции:

$$u = 8x^2 + 4xy + 5y^2,$$

если известно, что точка $M_0(10; 10)$ лежит в окрестности точки локального минимума.

Р е ш е н и е.

Первое приближение.

Выбирая точку M_0 за начальную (точка находится за пределами рис. 3.10,б), найдем в ней значение функции:

$$u(M_0) = 8 \cdot 10^2 + 4 \cdot 10 \cdot 10 + 5 \cdot 10^2 = 1700.$$

1. Найдем первые производные функции, представляя их в виде совокупности двух (по размерности пространства) координат вектора:

$$\begin{pmatrix} u'_x \\ u'_y \end{pmatrix} = \begin{pmatrix} 16x + 4y \\ 4x + 10y \end{pmatrix}.$$

Подставляя в эти соотношения значения координат точки M_0, найдем совокупность координат вектора градиента:

$$\nabla u(M_0) = \begin{pmatrix} u'_x(M_0) \\ u'_y(M_0) \end{pmatrix} = \begin{pmatrix} 16 \cdot 10 + 4 \cdot 10 \\ 4 \cdot 10 + 10 \cdot 10 \end{pmatrix} = \begin{pmatrix} 200 \\ 140 \end{pmatrix}.$$

2. Подставим полученные выражения для векторов в формулу (3.18):

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \alpha_0 \begin{pmatrix} u'_x(M_0) \\ u'_y(M_0) \end{pmatrix},$$

или

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 10 + 200\alpha_0 \\ 10 + 140\alpha_0 \end{pmatrix}.$$

3. Найдем значения функции в первом приближении с точностью до множителя α_0:

$$u(x_1, y_1) = 8(10 + 200\alpha_0)^2 + \\ + 4(10 + 200\alpha_0)(10 + 140\alpha_0) + 5(10 + 140\alpha_0)^2.$$

4. Определим α_0 из условия минимума функции $u(x_1, y_1)$ одной переменной α_0. Для этого производную от функции по переменной приравниваем нулю $\left(\dfrac{\partial u(x_1, y_1)}{\partial \alpha_0} = 0\right)$:

$$8 \cdot 2 \cdot 200(10 + 200\alpha_0) + \\ + 4 \cdot 200(10 + 140\alpha_0) + 4 \cdot 140(10 + 200\alpha_0) + \\ + 5 \cdot 2 \cdot 140(10 + 140\alpha_0) = 0.$$

Отсюда:

$$\alpha_0 = -\frac{596}{10600} \approx -0{,}056.$$

5. Найдем координаты независимых переменных точки первого приближения:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix} - 0{,}056 \begin{pmatrix} 200 \\ 140 \end{pmatrix} = \\ = \begin{pmatrix} -1{,}20 \\ 2{,}16 \end{pmatrix} \longrightarrow M_1(-1{,}20;\ 2{,}16).$$

6. Определим значение функции в точке M_1:

$$u(M_1) = 24{,}23.$$

На этом первое приближение заканчивается. Второе приближение. Выполним все действия в последовательности первого приближения.

1. $\nabla u(M_1) = \begin{pmatrix} u'_x(M_1) \\ u'_y(M_1) \end{pmatrix} = \\ = \begin{pmatrix} 16 \cdot (-1{,}20) + 4 \cdot 2{,}16 \\ 4 \cdot (-1{,}20) + 10 \cdot 2{,}16 \end{pmatrix} = \begin{pmatrix} -10{,}56 \\ 16{,}80 \end{pmatrix}.$

2. $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \alpha_1 \begin{pmatrix} u'_x(M_1) \\ u'_y(M_1) \end{pmatrix} = \\ = \begin{pmatrix} -1{,}20 \\ 2{,}16 \end{pmatrix} + \alpha_1 \begin{pmatrix} -10{,}56 \\ 16.8 \end{pmatrix}.$

3. $u(x_2, y_2) = 8(-1{,}20 - 10{,}56\alpha_1)^2 + \\ + 4(-1{,}20 - 10{,}56\alpha_1)(2{,}16 + 16{,}8\alpha_1) + 5(2{,}16 + 16{,}8\alpha_1)^2.$

4. Возьмем производную от полученного выражения по α_1 и приравняем ее нулю:

$$8 \cdot 2 \cdot (-10{,}56)(-1{,}20 - 10{,}56\alpha_1) + \\ + 4 \cdot (-10{,}56)(2{,}16 + 16{,}8\alpha_1) + \\ + 4 \cdot 16{,}8(-1{,}20 - 10{,}56\alpha_1) + \\ + 5 \cdot 2 \cdot 16{,}8(2{,}16 + 16{,}8\alpha_1) = 0.$$

Отсюда:

$$\alpha_1 \approx -0{,}152.$$

5. $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0{,}144 \\ 0{,}145 \end{pmatrix} \longrightarrow M_2(0{,}144;\ 0{,}145).$

6. $u(M_2) = 0{,}354.$

Второе приближение закончено.

После выполнения вычислений в следующих приближениях получим:

$$M_3(-0{,}018;\ 0{,}031); \quad u(M_3) = 0{,}005.$$
$$M_4(0{,}002;\ 0{,}002); \quad u(M_4) = 0{,}000\ldots.$$

Точное решение дает $u_{min} = u(0, 0, 0) = 0$.

Как видно из хода решения задачи, сходимость метода градиентного спуска, по крайней мере для рассмотренной функции, очень хорошая. Уже после второго шага значение функции уменьшилось по сравнению с нулевой точкой на три порядка, а после третьего шага практически не отличалось от нуля — локального минимума функции.

§ 3.11. ПРИБЛИЖЕННОЕ ВЫЧИСЛЕНИЕ ИНТЕГРАЛОВ. ФОРМУЛЫ ПРЯМОУГОЛЬНИКОВ И ТРАПЕЦИЙ

Использование понятия интегральной суммы позволяет получить формулы для приближенного вычисления определенных интегралов.

Пусть требуется вычислить приближенно интеграл от функции $y = f(x)$ на отрезке $x \in [a; b]$ (рис. 3.11).

Разобьем $[a; b]$ на n в общем неравных подынтервалов и в произвольных точках $\xi_k \in [x_{k-1}; x_k]$, принадлежащих каждому из n подынтервалов, определим значения функции $y_k = f(\xi_k)$.

Заменяя интеграл интегральной суммой, придем к приближенному равенству:

$$S = \int_a^b f(x)dx \approx \sum_{k=1}^{n} f(\xi_k)\Delta x_k = \sum_{k=1}^{n} S_k = S_{\text{пр}},$$

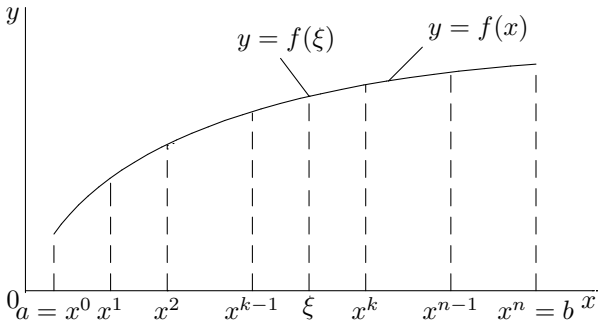


Рис. 3.11
Интегральная сумма

называемому *формулой прямоугольников* приближенного вычисления определенных интегралов.

В случае равенства всех интервалов $\Delta x_k = \Delta x = \dfrac{b-a}{n}$ и $f(\xi_k) = y_k$ формула прямоугольников представляется в виде

$$S_{\text{пр}} = \frac{b-a}{n} \sum_{k=1}^{n} y_k.$$

Если

$$f(\xi_k) = \min_{[x_{k-1};x_k]} f(x) \quad (k = \overline{1,n}), \quad \text{то} \quad S_{\text{пр}} = S_{\min}.$$

Если

$$f(\xi_k) = \max_{[x_{k-1};x_k]} f(x) \quad (k = \overline{1,n}), \quad \text{то} \quad S_{\text{пр}} = S_{\max}.$$

При этом справедливо неравенство:

$$S_{\min} \le \int_a^b f(x)dx \le S_{\max}.$$

Можно показать, что максимальная ошибка, которая может отличать точное значение интеграла от его приближенного значения, найденного с использованием формулы прямоугольников, не превышает

$$\Delta = \frac{(b-a)^2}{2n} \max_{[a;b]} |f'(x)|.$$

Если заменить криволинейные участки графика функции на подынтервалах стягивающими их хордами, то криволинейные трапеции на участках $[x_{k-1}; x_k]$ заменяются на прямолинейные и их площади определятся по формуле:

$$\Delta S_k = \frac{1}{2}(y_{k-1} + y_k)\Delta x_k.$$

Суммарная площадь элементарных трапеций:

$$S_{\text{тр}} = \sum_{k=1}^{n} = \frac{1}{2}(y_0 + y_1)\Delta x_1 + \frac{1}{2}(y_1 + y_2)\Delta x_2 + \cdots$$
$$\cdots + \frac{1}{2}(y_{n-1} + y_n)\Delta x_n.$$

При равенстве подынтервалов *формула трапеций* определит приближенное значение интеграла:

$$\int_a^b f(x)dx \approx S_{\text{тр}} = \frac{b-a}{n}\left(\frac{y_0 + y_1}{2} + y_1 + y_2 + \cdots + y_{n-1}\right).$$

Наибольшая ошибка отличия точного значения интеграла от значения, полученного по формуле трапеций, не превышает

$$\Delta = \frac{(b-a)^3}{12n^2}\max_{[a,b]}|f''(x)|.$$

§ 3.12. ФОРМУЛА СИМПСОНА ПРИБЛИЖЕННОГО ВЫЧИСЛЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

Формулы прямоугольников и трапеций не всегда дают достаточную точность расчетов при ограниченном количестве подынтервалов. В случае нелинейности подынтегральной функции большую точность дает замена участков кривых не прямыми линиями, как в прямоугольниках (горизонтальные прямые) или трапециях (наклонные прямые), а параболами $y = ax^2 + bx + c$.

Для получения формулы численного интегрирования, основанной на замене реальной кривой на квадратную параболу,

докажем справедливость соотношения:

$$\int_{x_0}^{x_2}(ax^2+bx+c)dx=\frac{x_2-x_0}{6}(y_0+4y_1+y_2), \qquad (3.19)$$

в котором y_0 и y_2 — значения функции на концах интервала интегрирования $[x_0;x_2]$, y_1 — на его середине. С целью упрощения доказательства в качестве интервала $[x_0;x_2]$ рассмотрим отрезок $[-h;h]$, симметричный относительно начала координат. В этом случае $x_0=-h$, $x_1=0$, $x_2=h$, $y_0=ah^2-bh+c$, $y_1=c$, $y_2=ah^2+bh+c$, и правая часть (3.19) преобразуется к виду

$$\frac{x_2-x_0}{6}(y_0+4y_1+y_2)=$$
$$=\frac{h+h}{6}(ah^2-bh+c+4c+ah^2+bh+c)=\frac{2h}{3}(ah^2+3c).$$

Вычислим левую часть (3.19), используя формулу Ньютона–Лейбница (Leibniz Gottfried Wilhelm (1646–1716) — немецкий философ, логик, математик и механик):

$$\int_{-h}^{h}(ax^2+bx+c)dx=\left(\frac{ax^3}{3}+\frac{bx^2}{2}+cx\right)\Bigg|_{-h}^{h}=\frac{2h}{3}(ah^2+3c).$$

Совпадение последних двух результатов говорит о справедливости соотношения (3.19).

Перейдем к непосредственному выводу формулы приближенного вычисления определенных интегралов путем параболической аппроксимации подынтегральной функции.

Разобьем $[a;b]$ на $2n$ равных (для простоты и обозримости формулы) подынтервалов. Пусть в граничных точках подынтервалов для координат $x_0,x_1,\ldots,x_{2n-1},x_{2n}$ найдены значения подынтегральной функции $y_0,y_1,\ldots,y_{2n-1},y_{2n}$.

Заменим кривые, описываемые подынтегральной функцией $y=f(x)$, на каждом подынтервале, содержащем по три точки (x_0,x_1,x_2), $(x_2,x_3,x_4),\ldots,(x_{2n-2},x_{2n-1},x_{2n})$, параболами, имеющими в перечисленных точках ординаты $(y_0,y_1,y_2),(y_2,y_3,y_4),\ldots,(y_{2n-2},y_{2n-1},y_{2n})$.

Найдем площади соответствующих параболических трапеций, используя формулу (3.19):

$$S_1 = \frac{b-a}{6n}(y_0 + 4y_1 + y_2),$$
$$S_2 = \frac{b-a}{6n}(y_2 + 4y_3 + y_4),$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$S_n = \frac{b-a}{6n}(y_{2n-2} + 4y_{2n-1} + y_{2n}).$$

Суммируя записанные площади и приводя в получаемой сумме подобные члены, приходим к *формуле Симпсона* приближенного вычисления определенных интегралов:

$$S_{\text{сим}} = \frac{b-a}{6n}[y_0+y_{2n}+4(y_1+y_3+\ldots+y_{2n-1})+2(y_2+y_4+\ldots+y_{2n-2})].$$

Подчеркнем, что формула Симпсона предполагает разбивку интервала интегрирования на четное число подынтервалов. В ней n — половина общего числа подынтервалов.

Погрешность формулы Симпсона не превышает величины

$$\Delta = \frac{(b-a)^5}{2880n^4} \max_{[a,b]} |f^{(4)}(x)|.$$

§ 3.13. ПРИМЕР ЧИСЛЕННОГО ОПРЕДЕЛЕНИЯ ПЛОЩАДИ

Найти площадь криволинейной трапеции, ограниченной графиком функции: $y = \dfrac{1}{\sqrt{x}}$ и прямыми: $y = 0$, $x = a = 1$, $x = b = 11$ методами:

а) прямоугольников;

б) трапеций;

в) Симпсона;

г) аналитическим с использованием формулы Ньютона–Лейбница.

Построить график заданной функции с разбиением отрезка $[a; b]$ на $n = 10$ интервалов и график функции $s = \displaystyle\int_0^x \frac{1}{\sqrt{t}}\,dt$ на отрезке $x \in [a; b]$.

Р е ш е н и е. Составим таблицу разбиения отрезка интегрирования на $n = 10$ равных участков с

$$\Delta x = \frac{b-a}{n} = \frac{11-1}{10} = 1.$$

Во второй строке таблицы представлены значения

$$y_k = f(x_k) = \frac{1}{\sqrt{x_k}} \quad (k = \overline{1;11}),$$

увеличенные в 10 раз. В табличных числовых данных количество значащих цифр ограничено тремя с целью компактности таблицы.

Таблица 3.2

Данные для численных методов

x_k	1	2	3	4	5	6	7	8	9	10	11
$10y_k$	10	7,07	5,77	5,00	4,47	4,08	3,78	3,54	3,33	3,16	3,02
s_k	0	0,82	1,46	2,00	2,48	2,90	3,30	3,66	4,00	4,32	4,64

а). Используя формулу прямоугольников с высотами, представляющими собой левые значения функции на концах подынтервалов, найдем приближенное значение площади криволинейной трапеции, в виде прямоугольников, очерченных на рис. 3.12 сплошными линиями:

$$S^{-}_{\text{пр}} = \sum_{k=1}^{n} y_k \Delta x = \\ = (1 + 0{,}707 + 0{,}577 + 0{,}5 + \ldots + 0{,}333 + 0{,}316) \cdot 1 = \\ = 5{,}020.$$

Та же формула прямоугольников, но с подстановкой в нее высот, равных правым значениям функции на концах подынтервалов, даст значение интеграла, равного площади ограниченных пунктиром прямоугольников:

$$S^{+}_{\text{пр}} = \sum_{k=2}^{n+1} y_k \Delta x = \\ = (0{,}707 + 0{,}577 + 0{,}5 + \ldots + 0{,}316 + 0{,}302) \cdot 1 = \\ = 4{,}322.$$

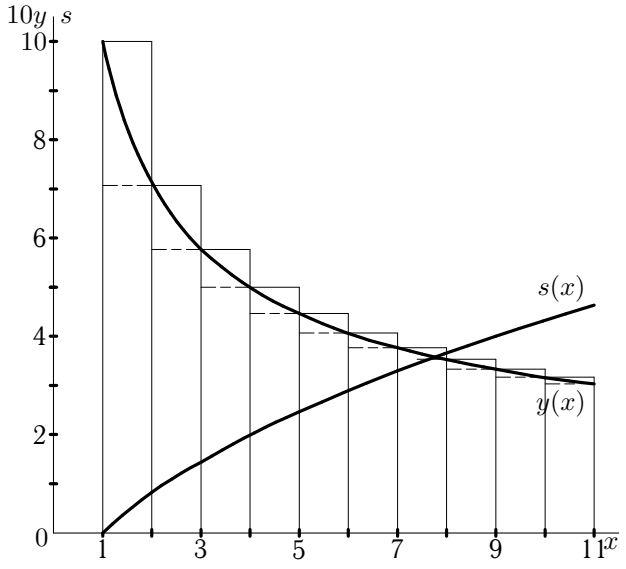


Рис. 3.12
Метод прямоугольников и первообразная s

б). По формуле трапеций получим:

$$S_{\text{тр}} = \left(\frac{y_1 + y_{n+1}}{2} + \sum_{k=2}^{n} y_k \right) \Delta x = \\ = \left(\frac{1 + 0{,}302}{2} + 0{,}707 + \ldots + 0{,}316 \right) \cdot 1 = 4{,}671.$$

Следует обратить внимание на очевидное равенство:

$$S_{\text{тр}} = \frac{1}{2}(S^{-}_{\text{пр}} + S^{+}_{\text{пр}}).$$

в). По формуле Симпсона ($n = 5$ — количество спаренных подынтервалов):

$$S_{\text{сим}} = \frac{b-a}{6n}\left[y_1 + y_{2n+1} + 4\sum_{k=1}^{n} y_{2k} + 2\sum_{k=2}^{n} y_{2k-1}\right] =$$

$$= \frac{11-1}{6\cdot 5}[1 + 0{,}302 + 4(0{,}707 + 0{,}5 + 0{,}408 + 0{,}354 + 0{,}316) +$$

$$+ 2(0{,}577 + 0{,}447 + 0{,}378 + 0{,}333)] = 4{,}637.$$

г). Определим точное значение интеграла, являющегося табличным:

$$S = \int_1^{11} \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_1^{11} = 2(\sqrt{11} - \sqrt{1}) \approx 4{,}633.$$

Найдем относительные ошибки определения площадей различными использованными методами численного интегрирования, сравнивая их с точным значением площади, полученным методом Ньютона–Лейбница:

$$\Delta_{\text{пр}}^{-} = \frac{S_{\text{пр}}^{+} - S}{S} = \frac{5{,}020 - 4{,}633}{4{,}633} \approx 0{,}08.$$

Аналогично: $\Delta_{\text{пр}}^{+} \approx 0{,}07 = 7 \cdot 10^{-2}$; $\Delta_{\text{тр}} \approx 8 \cdot 10^{-3}$; $\Delta_{\text{сим}} \approx 9 \cdot 10^{-4}$.

Найдем функцию s, заданную в виде интеграла с переменным верхним пределом от заданной функции. Это по своей сути — первообразная функции y с равном нулю значением при $x = 1$:

$$s = \int_1^{x} \frac{dt}{\sqrt{t}} = 2\sqrt{t}\Big|_1^{x} = 2(\sqrt{x} - 1).$$

Значения функции $s(x)$ для координат границ интервалов интегрирования приведены в последней строке табл. 3.2.

Графики первообразной $s(x)$ и заданной функции $y(x)$ показаны на рис. 3.12.

§ 3.14. РАЗНОСТНЫЕ УРАВНЕНИЯ

В экономических задачах часто приходится иметь дело не с непрерывно изменяющимися функциями, а с величинами, изменяющимися скачкообразно за конечные промежутки времени. В этом случае удобно рассматривать не производные от функции (предельные величины), а приращения функций, то есть конечные разности между двумя ее значениями.

Пусть для двух достаточно близких значений x_i и x_{i+1} независимой переменной x известны зависящие от них значения y_i и y_{i+1} переменной y.

Конечной разностью первого порядка называется величина

$$\Delta y_i = y_{i+1} - y_i.$$

Разность двух первых разностей образует *конечную разность второго порядка* и обозначается

$$\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i = y_{i+2} - y_{i+1} - (y_{i+1} - y_i) = y_{i+2} - 2y_{i+1} + y_i.$$

Для вычисления конечной разности третьего порядка

$$\Delta^3 y_i = \Delta^2 y_{i+1} - \Delta^2 y_i = y_{i+3} - 3y_{i+2} + 3y_{i+1} - y_i$$

требуется значения функции в четырех точках: x_{i+3}, x_{i+2}, x_{i+1} и x_i. Вычисления конечных разностей удобно осуществлять, занося их значения в таблицу (табл. 3.3).

Конечные разности позволяют заменить производные различных порядков их приближенными выражениями. Так, если приращения независимой переменной не изменяются при переходе от одной точки к другой ($\Delta x_i = x_{i+1} - x_i = h$), то в точке с координатой $x = x_i$:

$$\left.\frac{dy}{dx}\right|_{x=x_i} \approx \left.\frac{\Delta y}{\Delta x}\right|_{x=x_i} = \frac{1}{h}\,(y_{i+1} - y_i);$$

$$\left.\frac{d^2y}{dx^2}\right|_{x=x_i} \approx \left.\frac{\Delta^2 y}{\Delta x^2}\right|_{x=x_i} = \frac{1}{h^2}\,(y_{i+2} - 2y_{i+1} + y_i); \quad \ldots$$

Конечные разности позволяют перейти от дифференциальных уравнений n-ного порядка к *разностным уравнениям*

$$F\,(n, y_i, y_{i+1}, \ldots, y_{i+n}) = 0. \tag{3.20}$$

Таблица 3.3
Конечные разности

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
x_0	y_0				
		Δy_0			
x_1	y_1		$\Delta^2 y_0$		
		Δy_1		$\Delta^3 y_0$	
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$
		Δy_2		$\Delta^3 y_1$	
x_3	y_3		$\Delta^2 y_2$		
		Δy_3			
x_4	y_4				

В (3.20) ($\imath = \overline{0, m}$), m — количество интервалов, на которые разбивается область изменения переменной x.

Многие разностные уравнения можно записать в виде *линейных разностных уравнений*:

$$A_0 y_\imath + A_1 y_{\imath+1} + \ldots + A_n y_{\imath+n} = f(x_\imath) \quad (\imath = \overline{1, m}), \qquad (3.21)$$

где A_j ($j = \overline{1, n}$) — постоянные коэффициенты. Решения линейных разностных уравнений во многом аналогичны соответствующим решениям линейных дифференциальных уравнений (ДУ). Продемонстрируем сказанное на примере разностного уравнения второго порядка:

$$y_{\imath+2} + p y_{\imath+1} + q y_\imath = f(x_\imath) \quad (\imath = \overline{1, m}). \qquad (3.22)$$

Общее решение $y_\imath$ этого уравнения равно сумме общего решения $y_{\imath 0}$ однородного уравнения (при $f(x_\imath) = 0$) и какого-либо частного решения $y_{\imath *}$ неоднородного уравнения.

Для нахождения общего решения однородного уравнения составляем характеристическое уравнение:

$$k^2 + pk + q = 0. \qquad (3.23)$$

Уравнение имеет два корня k_1 и k_2.

Как и при решении ДУ возможны три случая:
1) если корни действительные и различные, то

$$y_{i0} = C_1 k_1^i + C_2 k_2^i; \tag{3.24}$$

2) если корни действительные и равные ($k_1 = k_2 = k$), то

$$y_{i0} = (C_1 + iC_2)k^i;$$

3) если корни комплексно сопряженные $k_{1,2} = r(\cos\alpha \pm i\sin\alpha)$, то

$$y_{i0} = r^i(C_1\cos\alpha + C_2 i\sin\alpha).$$

В последнем выражении $i = \sqrt{-1}$ — «мнимая единица». Ее следует отличать от используемого в предыдущих записях индекса $i = 1, 2, \ldots$.

Отметим, что, в отличие от частных решений однородных ДУ, в частные решения разностных уравнений входят не функции e^{kx}, а сами решения k характеристического уравнения в степени i.

Частное решение неоднородного разностного уравнения зависит, как и в случае ДУ, от вида функции $y_i(x_i)$ и корней характеристического уравнения.

Пример. Найти решение разностного уравнения

$$y_{i+2} - 9y_{i+1} + 20y_i = 30 \cdot 2^i \tag{3.25}$$

при условии, что интервал изменения независимой переменной x разбит на пять подынтервалов ($i = \overline{0,5}$) и известны значения функции y на концах интервала (граничные условия):

$$y(x_0) = 0, \qquad y(x_5) = 1343. \tag{3.26}$$

Р е ш е н и е. Рассмотрим однородное уравнение, построенное на основе заданного разностного:

$$y_{i+2,0} - 9y_{i+1,0} + 20y_{i,0} = 0. \tag{3.27}$$

Для определения решения этого уравнения составим характеристическое уравнение:

$$k^2 - 9k + 20 = 0.$$

Корни уравнения $k_1 = 4$ и $k_2 = 5$ — действительные и различные. Поэтому решение однородного уравнения (3.27) представляем в виде (3.24):

$$y_{i0} = C_1 k_1^i + C_2 k_2^i = 4^i C_1 + 5^i C_2. \tag{3.28}$$

Частное решение неоднородного уравнения будем искать в виде функции, определяемой правой частью уравнения (3.25):

$$y_{i*} = A \cdot 2^i.$$

Множитель A найдем, подставив выбранное частное решение в исходное разностное уравнение. Слагаемые левой части уравнения:

$$y_{i+2} = A \cdot 2^{i+2} = 4A \cdot 2^i; \quad y_{i+1} = A \cdot 2^{i+1} = 2A \cdot 2^i.$$

Поэтому из (3.25) получим:

$$A \cdot 2^i(4 - 9 \cdot 2 + 20) = 30 \cdot 2^i, \quad \to \quad 6A = 30, \quad \to \quad A = 5.$$

Таким образом,

$$y_{i*} = 5 \cdot 2^i.$$

Общее решение разностного уравнения складывается из общего решения однородного уравнения (3.28) и частного решения неоднородного уравнения. Поэтому

$$y_i = 4^i C_1 + 5^i C_2 + 5 \cdot 2^i. \tag{3.29}$$

Решение разностного уравнения записано с точностью до постоянных C_1 и C_2. Для их определения используем граничные условия (3.26). Из них следует:
при $i = 0$ задано $y_0 = y(x_0) = 0$:

$$0 = 4^0 C_1 + 5^0 C_2 + 5 \cdot 2^0, \text{ или } C_1 + C_2 + 5 = 0;$$

при $i = 5$ задано $y_5 = y(x_5) = 1343$:

$$1343 = 4^5 C_1 + 5^5 C_2 + 5 \cdot 2^5, \text{ или } 1024C_1 + 3125C_2 + 5 - 1183 = 0.$$

Решая полученную систему уравнений, получим:

$$C_1 = -8, \qquad C_2 = 3.$$

Таким образом, решение разностного уравнения, удовлетворяющее заданным граничным условиям:

$$y_i = -8 \cdot 4^i + 3 \cdot 5^i + 5 \cdot 2^i. \tag{3.30}$$

В табл. 3.4 приведены значения функции y_i для шести значений номеров i граничных точек подынтервалов изменения переменной x.

Таблица 3.4
Результаты решения примера

i	0	1	2	3	4	5
y_i	0	-7	-33	-97	-93	1343

§ 3.15. ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ КОШИ

К настоящему времени разработано большое количество достаточно эффективных методов численного решения ДУ. Остановимся на двух из них — методах, основанных на замене операций дифференцирования конечными приращениями функций и аргументов. Методы реализуют решение задачи Коши — ДУ с заданными начальными условиями.

Простейшим из методов численного интегрирования ДУ является метод Эйлера. Суть этого метода заключается в следующем.

Пусть задано ДУ уравнение в виде

$$y' = f(x, y).$$

Требуется найти решение этого уравнения на интервале $x \in [a; b]$ при заданном граничном условии $y(a) = y(x_0) = y_0$ (рис. 3.13,а).

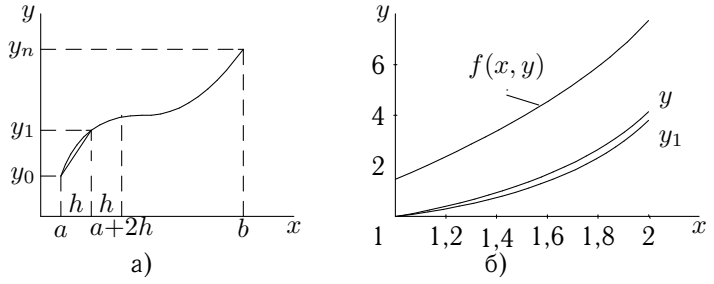


Рис. 3.13
Иллюстрация метода Эйлера и пример

Разобьем интервал $[a;b]$ на n равных (для простоты) подынтервалов с шагом $h=(b-a)/n$.

На участке $[a;a+h)$ дифференциальное уравнение можно заменить приближенным равенством

$$\frac{y_1-y_0}{h}\approx f(x_0,y_0).$$

Здесь $x_0=a,\quad y_0=y(a)$ — параметры граничного условия. Если x_0 и y_0 известны, то значение функции $f(x_0,y_0)$ может быть подсчитано. Тогда из приведенного приближенного равенства следует:

$$y_1\approx y_0+hf(x_0,y_0).$$

Это приближенное значение искомой переменной в точке с координатой $x_1=a+h$.

Определяя далее $f(x_1,y_1)$, из приближенного равенства:

$$\frac{y_2-y_1}{h}\approx f(x_1,y_1)$$

находим

$$y_2\approx y_1+hf(x_1,y_1).$$

И так далее. Для вычисления значения функции в $(k+1)$-й точке имеем рекуррентную формулу:

$$y_{k+1}\approx y_k+hf(x_k,y_k).$$

Метод Эйлера выгодно отличается от других методов численного интегрирования ДУ своей простотой, но не большой точностью. Поэтому на практике большее распространение получил другой метод решения ДУ — *метод Рунге–Кутты,* который имеет несколько разновидностей.

В методе Рунге–Кутты *третьего порядка* вычисление искомой переменной на $(k+1)$-м шаге интегрирования осуществляется с использованием формулы:

$$y_{k+1} = y_k + \frac{1}{6}(p_1 + 4p_2 + p_3),$$

где

$$\begin{aligned} p_1 &= hf(x_k,\ y_k), \\ p_2 &= hf\left(x_k + \frac{1}{2}h,\ y_k + \frac{1}{2}p_1\right), \\ p_3 &= hf(x_k + h,\ y_k - p_1 + 2p_2). \end{aligned}$$

Пример. Решить численно методами Эйлера и Рунге–Кутты ДУ

$$y' = \frac{2x}{x^2+1}y + x\sqrt{x^2+1}$$

на отрезке $x \in [1, 2]$ с шагом $h = 0,2$, если $y(1) = 0$.

Таблица 3.5
Результаты решения примера

x	$f(x,y)$	y_1	y_2	y
1.0	1,4142	0	0	0
1.2	2,2294	0,2828	0,368	0,3608
1.4	3,2655	0,7286	0,910	0,9065
1.6	4,5312	1,3817	1,60	1,6824
1.8	6,0281	2,2879	2,64	2,7345
2.0	7,7595	3,8388	4,00	4,1093

Результаты вычислений, выполненных с использованием вышеприведенных формул, сведены в таблицу и показаны на двух нижних графиках (рис. 3.13,б). На графиках и в

табл. 3.5 y_1 относится к решению методом Эйлера; y_2 — методом Рунге–Кутты y (без индекса) — аналитическим методом. Решение, полученное методом Рунге–Кутты y_2, практически сливается с графиком теоретической кривой y и поэтому не выделено на рисунке. В правом столбце таблицы приведены результаты аналитического решения уравнения, которое имеет вид:

$$y = (\sqrt{x^2+1} - \sqrt{2})(x^2+1).$$

§ 3.16. РЕЗЮМЕ

Математические модели, описывающие реальные процессы, часто представляются уравнениями, решение которых методами математического анализа затруднен или невозможен. В таких случаях прибегают к численным методам.

Наиболее распространенным методом численного анализа является метод последовательных приближений (итераций). Этот метод, в частности, используется для решения нелинейных алгебраических уравнений при нахождении их корней.

Существует большое количество модификаций итерационных методов. К ним относятся следующие.

Преобразование исследуемого уравнения типа $x = f(x)$ к последовательности рекуррентных соотношений

$$x_n = f(x_{n-1}),$$

в которых по заданному в начальном приближении значению x_0 находится значение x_1, по x_1 находится x_2 и т. д. Для сходящегося процесса последовательных приближений можно с любой точностью определить корень уравнения x_*, приняв его равным x_n. Ошибка ε в вычислении корня обычно оценивается по относительному изменению искомой величины (разности двух соседних итераций):

$$\varepsilon = \frac{x_n - x_{n-1}}{x_n}.$$

Разновидностью итерационного метода в экономике, используемого для определения точки равновесия между спросом и предложением, является паутинная модель, отслеживающая реальные колебания цен на рынке.

Методы пробных точек (половинного деления, золотого сечения) позволяют в уравнениях, связывающих функцию с независимой переменной $y = f(x)$, искусственно сужать интервал изменения x, стягивая его к точке, где функция имеет характерную особенность (экстремум, равенство нулю, ...).

Среди численных методов отыскания корней уравнений широкое распространение нашли методы хорд и касательных. Совместное их использование позволяет достаточно быстро (по количеству итераций) стягивать интервал изменения аргумента к искомой точке.

Метод касательных применим для дифференцируемых функций. Если функция дифференцируема и определение ее производной не затруднительно, то метод хорд и касательных позволяет определить стационарную точку функции — точку, где ее производная обращается в нуль.

Для функции нескольких переменных также разработано большое количество численных методов. В частности, симплексный метод позволяет «накрывать» область изменения независимых переменных многогранниками (треугольниками на плоскости). Многогранники можно уменьшать до таких размеров, что в их внутренней области искомая функция будет сколь угодно мало отличаться от ее экстремального (или другого интересующего исследователя) значения.

Эффективным методом отыскания экстремумов дифференцируемых функций нескольких переменных является метод градиентов. В качестве шагов, позволяющих при решении задачи перейти от произвольно заданной точки функции нескольких переменных к искомой, выбирается вектор, сонаправленный градиенту. Модуль этого вектора определяется из требования скорейшего приближения решения к требуемому.

Численное интегрирование сводится к разбиению области интегрирования на подобласти представляющие собой криволинейные трапеции. Точность методов численного интегрирования во многом зависит от точности аналитического описания криволинейной границы. Применение для аппроксимации криволинейных границ кусочно-постоянных функций, линейных и нелинейных (как в методе Симпсона) зависимостей

приводят к различной точности получаемых решений при одинаковом количестве подынтервалов области интегрирования.

При численном дифференцировании, в том числе при использовании численных методов для решения дифференциальных уравнений, операторы дифференцирования заменяются конечными разностями. Здесь, как и при численном интегрировании, точность решения зависит от характера зависимости, используемой для аппроксимации кривой на подынтервалах интегрирования. Методы Рунге–Кутты нелинейной аппроксимации дают бóльшую точность, чем метод Эйлера линейной аппроксимации.

Численные методы требуют, как правило, больших затрат времени. Поэтому они становятся эффективными только при использовании ЭВМ, для которых в настоящее время создана обширная библиотека стандартных программ по численным методам.

§ 3.17. ВОПРОСЫ

1. В чем состоят достоинства и недостатки численных методов?
2. Как свести задачи отыскания экстремума и точек перегиба функций к определению корней уравнений?
3. Какие функции называются унимодальными?
4. Какие ограничения должны быть наложены на функцию при отыскании начального интервала существования ее корня?
5. Какие ограничения накладываются на функцию в задаче уменьшения интервалов существования ее корней?
6. Сформулируйте условия существования (отсутствия) корней на рассматриваемом интервале.
7. Какие условия должны быть выполнены, чтобы на рассматриваемом интервале существовал экстремум функции?
8. Поясните на графике функции, как выбирается интервал, который можно исключить при отыскании корней уравнений.

9. Что такое пробная точка?
10. Поясните на графике функции смысл теоремы об исключении интервалов при определении корней.
11. Какие три ситуации могут возникнуть при выборе начального интервала существования корней уравнения? точек экстремума?
12. Что такое золотое сечение?
13. Из каких соображений выбирается положение пробных точек в методе золотого сечения?
14. В чем смысл введения и как вводится безразмерная координата, в частности, в методе золотого сечения?
15. Какими соотношениями связывается безразмерная координата с размерной?
16. Как связан размер остающегося для рассмотрения интервала в методе золотого сечения с параметрами координат пробных точек на n-й итерации?
17. Как определить точность вычислений в численных методах?
18. В чем заключается метод хорд?
19. Какие ограничения следует наложить на функцию при использовании метода хорд отыскания корней уравнений?
20. Влияет ли выпуклость функции на организацию итерационного процесса метода хорд?
21. Опишите аналитически задачу приближенного определения корня уравнения методом хорд на n-м приближении.
22. Изобразите на графике функции последовательность отыскания корня методом хорд.
23. В чем заключается метод касательных определения корней уравнений?
24. Какие ограничения накладываются на функцию при определении ее корней методом касательных?
25. Как определить начальную точку итерационного процесса метода касательных? Какой для этой точки должна

быть связь между значением функции и ее выпуклостью?

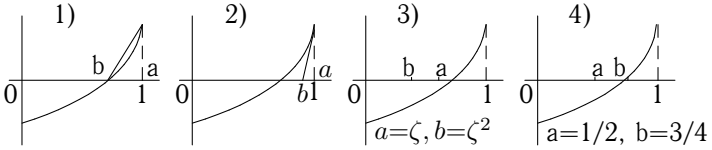
26. В каких случаях решение, получаемое методом касательных, не будет сходиться к точному решению? Поясните на графике.
27. Опишите аналитически задачу приближенного определения корня уравнения методом касательных на n-м приближении.
28. Изобразите на графике функции процесс отыскания корня уравнения методом хорд и касательных.
29. В чем смысл симплексного метода отыскания экстремума функции нескольких переменных?
30. Что собой представляют симплексы для двух- и трехмерных областей?
31. Как организуется процесс последовательных приближений в симплексном методе? Поясните на рисунке.
32. Как выбрать направление продвижения к очередному симплексу при отыскании минимума функции? Максимума?
33. Каким параметром и как можно регулировать скорость и точность решения в симплексном методе?
34. Как строится новый симплекс-треугольник?
35. Каков критерий накрытия симплексом точки экстремума?
36. В чем смысл метода градиентного спуска?
37. Как определяется направление вектора градиента по отношению к точке экстремума для выпуклых и вогнутых функций? Каким способом можно регулировать это направление?
38. Опишите последовательность определения параметров на итерации в методе градиентного спуска первого порядка.
39. Как строится итерационный процесс в методе градиентного спуска?

40. Из каких соображений выбирается множитель, корректирующий величину вектора градиента на итерациях?
41. Покажите на графике линий равного уровня последовательные положения вектора градиента в определении точки экстремума функции двух переменных.
42. В чем суть применения метода прямоугольников при вычислении определенных интегралов?
43. Можно ли с использованием формулы прямоугольников получить точное значение интеграла? Если да, то в каких случаях?
44. В чем суть метода трапеций? Объяснение сопроводите графическим построением.
45. Поясните, почему совпали результаты приближенного вычисления интеграла по формуле трапеций с полусуммой наименьшей и наибольшей площадей, вычисленных по формуле прямоугольников в примере раздела § 3.11? Будут ли такие совпадения иметь место для произвольных подынтегральных функций?
46. Какая аппроксимация подынтегральной функции осуществляется при выводе формулы Симпсона? Изобразите рисунок, поясняющий разбиение интервала интегрирования на подынтервалы, соответствующее формуле Симпсона.
47. Могут ли результаты вычисления определенных интегралов по формулам трапеций или прямоугольников быть точнее результатов, полученных по формуле Симпсона? Если да, то в каких случаях?
48. В чем суть метода Эйлера численного интегрирования ДУ? Пояснения сопроводите графическим построением.
49. В чем суть метода Рунге–Кутты решения ДУ? В чем состоят его достоинства и недостатки по сравнению с методом Эйлера?
50. Можно ли с помощью метода Эйлера получить более точный результат решения ДУ, чем по методу Рунге–Кутты? Если да, то в каких случаях?

51. Что собой представляют разностные уравнения?
52. Как осуществить переход от дифференциальных уравнений к разностным?

Выберите правильные ответы

1. Расставьте номера названий методов численного определения корней уравнений, в порядке следования номеров поясняющих эти методы рисунков. При неверном построении рисунка первого приближения в ответе поставьте цифру 5. На рисунках: $[0;1]$ — интервал, на котором находится корень; a и b координаты точки x, используемые в первой итерации; $\zeta = (\sqrt{5}-1)/2$.



1. Метод хорд.
2. Метод золотого сечения.
3. Метод касательных.
4. Метод половинного деления.
5. Ни один из пунктов 1–4 не соответствует рисункам.

2. Перечислите соотношения и утверждения, справедливые для метода золотого сечения, если $\omega \in [0,1]$ – точка интервала, где находится корень уравнения.

1. $\zeta^2 + \zeta - 1 = 0$. 2. $\zeta < \omega$. 3. $0 < \zeta < 1$.
4. $\zeta = (\sqrt{5}-1)/2$ – точка экстремума.
5. Среди ответов 1–4 нет правильных.

3. Расставьте номера ответов (правая колонка) в порядке следования вопросов (левая колонка) для задачи отыскания

методом исключения интервалов минимума функции $f(x_*)$ на определенном интервале $[a,b]$, если $a < x_1 < x_2 < b$, $x_* \in \in [a,b]$. В местах отсутствия правильных ответов поставьте цифру 5.

$1.x_* \in [x_1, b)$.	$1.f(x_1) > f(x_2)$.
$2.x_* \in (a, x_2]$.	$2.f(x_1) < f(x_2)$.
$3.x_* \in [x_1, x_2]$.	$3.f(x_1) < f(b)$.
$4.x_* \in (x_2, b)$.	$4.f(x_1) = f(x_2)$.

4. Перечислите соотношения, используемые при приближенном определении корней уравнения на интервале $[a,b]$: a) методом хорд; b) методом касательных.

1. $y - y_a = \dfrac{y_b - y_a}{b-a}(x-a)$. 2. $x_n = x_{n-1} - \dfrac{b - x_{n-1}}{y_b - y_{n-1}} y_{n-1}$.

3. $y - y_a = y'(a)(x-a)$. 4. $x_n = x_{n-1} - \dfrac{y_{n-1}}{y'(x_{n-1})}$.

5. Среди соотношений 1–4 нет требуемых.

5. Перечислите номера обязательных требований к расположению начальной точки метода касательных определения корней уравнений.

1. Для выпуклого графика функции — правая точка.
2. Для вогнутого графика функции — левая точка.
3. Там, где совпадают знаки первой и второй производных.
4. Там, где совпадают знаки функции и второй производной.
5. Среди ответов 1–4 нет правильных.

6. Перечислите правильные утверждения, справедливые для функции $u = u(x,y)$, имеющей в точке A локальный минимум, единственный во всей рассматриваемой области определения функции.

1. $u(X) > u(A)$ $(X \in (x;y) \neq A)$.
2. $u''_{xx}(A) < 0$; $|\nabla u|_A = 0$.
3. В симплексном методе треугольник $M_{n-1}M_nM_{n+1}$ «накрывает» точку A, если $u(M_{n+2}) > u(M_{n-1})$, где M_{n+2} — вершина треугольника $M_nM_{n+1}M_{n+2}$.

4. В методе градиентного спуска вектор градиента совпадает по направлению с касательной к линии уровня функции u.

5. Среди ответов 1–4 нет правильных.

7. Расставите номера ответов (правая колонка) в порядке следования соответствующих им вопросов (левая колонка), касающихся методов численного интегрирования ($h = (b-a)/n$). В местах отсутствия полного соответствия поставьте цифру 5.

1.Прямоугольников S^{-}_{np}.	1. $\frac{h}{6}\big(y_0 + y_{2n} + 2(y_2 + \ldots + y_{2n-2}) + 4(y_1 + \ldots + y_{2n-1})\big)$.
2.Прямоугольников S^{+}_{np}.	2. $h\left(\frac{y_0 + y_n}{2} + y_1 + y_2 + \ldots + y_{n-1}\right)$.
3.Трапеций.	3. $h(y_1 + y_2 + \ldots + y_n)$.
4.Симпсона.	4. $h(y_0 + y_1 + \ldots + y_{n-1})$.

О Т В Е Т Ы
на вопросы тестов

1. 5324; **2**. 5; **3**. 1245; **4**. *а*) 1, *b*) 3; **5**. 4; **6**. 123; **7**. 4321.

§ 3.18. РАСЧЕТНАЯ РАБОТА, ЧАСТИ 1–5

В заданиях на выполнение курсовой работы параметр k определить по формуле $k = 1 + \frac{N}{n}$, где N — последняя цифра номера группы, в которой числится студент; n — порядковый номер студента в списке группы.

Часть 1

Составить математическую модель произвольного реального физического процесса, описываемого степенным комплексом, включающую в себя не менее трех независимых размерностей. Проверить правильность построения модели с использованием π-теоремы.

Часть 2

Задана функция $g(x) = \dfrac{k}{x} + (k+1)x + k^2$.

1. Установить один интервал унимодальности функции.
2. Найти локальный экстремум на этом интервале:
   2.1. аналитическим методом;
   2.2. методом золотого сечения;
   2.3. методом хорд и касательных.

Часть 3

Приняв точку $M_0(k;\ k/2)$ за начальную, найти минимальное значение функции $u = \dfrac{(x-2k)^2}{16} + \dfrac{(y-k)^2}{9} - k^2$:

1. симплексным методом;
2. методом градиентного спуска.

Часть 4

Задана функция $f(x) = \dfrac{1}{\sqrt{k+x}}$.

Количество подынтервалов на заданных ниже интервалах при численном решении задач частей 4 и 5 равно 10.

Вычислить интеграл $\int_k^{k+10} f(x)dx$:

1. используя метод замены переменной (при получении первообразной) и формулу Ньютона–Лейбница (аналитическое решение);
2. методом прямоугольников, определяя площади по левым значениям функции на подынтервалах $S^-_{\text{пр}}$ и по правым — $S^+_{\text{пр}}$;
3. методом трапеций;
4. методом Симпсона ($2n = 10$);
5. определить функцию $s = \int_k^x f(x)dx$;

6. изобразить графики функции, ее первообразной и пояснить их смысл.

Часть 5

Получить решение ДУ $y' = y^{-1}f(x)$ на интервале $[N; N+10]$ при условии $y(N) = 1$:

1. аналитическим методом разделения переменных;
2. методом Эйлера;
3. методом Рунге–Кутты.
4. Представить полученные результаты в виде таблиц и графиков.

Выражение для функции $f(x)$ задано в части 4 задания.

Приближенные вычисления проводить или с точностью до третьей значащей цифры, или ограничиться 5-ю итерациями.

Дать оценку точности полученных результатов.

Глава 4

ТЕОРИЯ КОРРЕЛЯЦИИ

Стохастический подход к построению математических моделей широко используется при анализе реальных процессов. Без него нельзя обойтись в тех случаях, когда изменение показателей не носит детерминированного характера, а может считаться случайным процессом.

Анализу стохастических закономерностей различных процессов предшествует получение статистических данных. Для этого проводят *статистические испытания* — специально организованный эксперимент над натурным объектом или его моделью с учетом воздействий на них случайных возмущений.

Моделирование реальных явлений с учетом внешних, в том числе случайных воздействий называется *стохастическим моделированием*.

При стохастическом моделировании возмущения представляют в виде набора конечного числа случайных величин с известными законами их распределения.

Статистические испытания процессов с конкретными внешними воздействиями называют *реализацией случайного процесса в данной модели*.

При анализе одномерных статистических совокупностей используются полученные в статистическом испытании вариационные ряды, выборочные законы распределения, выборочные характеристики (статистики). Для исследования многомерных статистических совокупностей привлекают более сложные соотношения, в том числе корреляционные зависимости и уравнения регрессии — понятия, о которых идет речь в данной главе.

Задача математической статистики заключается не только в сборе статистических данных, но и в определении наличия, широты и интенсивности связей между показателями различных процессов, в создании математических моделей, описывающих эти процессы, с целью их анализа и выработки рекомендаций по оптимизации и управлению ими.

§ 4.1. СТАТИСТИЧЕСКАЯ И КОРРЕЛЯЦИОННАЯ ЗАВИСИМОСТИ

Рассмотрим зависимость случайной величины Y от одной случайной (в общем), или детерминированной (в частности) величины X.

Если каждому возможному значению X соответствует единственное возможное значение Y, то Y называют *функцией* аргумента X.

Строгая функциональная зависимость между двумя случайными величинами реализуется редко, так как обе величины могут быть подвержены воздействию случайных факторов. Пусть, например, Y зависит от случайных факторов U, V, W, а X — от случайных факторов P, Q, V. Так как среди влияющих факторов имеются общие (в данном случае V), то говорят, что X и Y связаны *статистической зависимостью*.

Статистической называется зависимость, при которой изменение одной из величин влечет за собой изменение *закона распределения* другой.

Если статистическая зависимость проявляется в том, что при изменении одной величины изменяется среднее значение (математическое ожидание) другой, то зависимость называется *корреляционной*.

Пример корреляционной зависимости.

Пусть Y — урожай зерна; X — количество внесенных удобрений. С одинаковых по площади участков земли при одинаковом количестве внесенных удобрений собирают различный урожай. В этом случае зависимость Y от X не является функциональной, так как на нее влияют различные факторы (условия): осадки, температура, квалификация комба́йнера и т. д. Вместе с тем средняя урожайность является функцией

от количества внесенных удобрений. Поэтому между Y и X существует корреляционная зависимость.

Между функциональной и корреляционной зависимостями случайных величин существует связь.

1. Если случайные величины X и Y функционально зависимы, то они коррелированы.

2. Если X и Y независимы, то они некоррелированы. Обратное утверждение в общем случае не верно. Однако в частном случае, когда X и Y распределены по нормальному закону, из некоррелированости этих величин следует их независимость.

Таким образом, корреляционная зависимость занимает промежуточное положение между функциональной зависимостью и независимостью. Поэтому корреляционную зависимость считают «слабой» зависимостью между случайными величинами.

§ 4.2. ИНТЕРПОЛИРОВАНИЕ ФУНКЦИЙ

Многие реальные процессы и явления (физические, экономические, социальные, ...) описываются функциональными зависимостями, предусматривающими строгую взаимную однозначность определения функции по значению аргумента. Экспериментальные данные, полученные по замерам процесса лишь в некоторых точках его протекания, обычно представляются в табличном виде. Если для значений x, не совпадающих с заданными, требуется определить соответствующие значения y, то возникает вопрос: как это сделать?

Замена совокупности дискретно заданных величин непрерывными функциями называется *апроксимацией*. Апроксимация, производимая для определения значений y по нетабличным значениям независимой переменной x для внутренних точек интервала ее изменения, есть задача *интерполирования*. Если апроксимация распространяется на значения, внешние по отношению к рассматриваемому интервалу значений x, то это – задача *экстраполирования*.

Пусть некоторой функцией $y = f(x)$ трем заданным значениям независимой переменной x, а именно (x_{k-1}, x_k, x_{k+1}), поставлены в соответствие три значения y: (y_{k-1}, y_k, y_{k+1}).

Простейший вариант решения задачи интерполирования сводится к тому, что для $x \in [x_{k-1}, x_k]$ функция y принимается равной или y_{k-1}, или y_k. В результате такой интерполяции получаем ступенчато-постоянную функцию, график которой отмечен на рис. 4.1 цифрой (1).

Наибольшее распространение получило *линейное* интерполирование, которое сводится к тому, что через табличные точки проводятся прямые.

Пусть точки (x_{k-1}, y_{k-1}) и (x_k, y_k) соединены прямой $y = ax + b$. Из условия прохождения прямой через две заданные точки (координаты этих точек должны обращать уравнение прямой в тождество) получаем систему двух уравнений с неизвестными a и b:

$$\begin{cases} y_{k-1} = ax_{k-1} + b; \\ y_k = ax_k + b. \end{cases}$$

Решая эту систему, определяем a и b, после чего уравнение прямой запишется в виде

$$y = (x - x_{k-1})\frac{y_k - y_{k-1}}{x_k - x_{k-1}} + y_{k-1}.$$

Чтобы получить уравнение прямой, проходящей через следующую пару точек, достаточно в полученном уравнении к

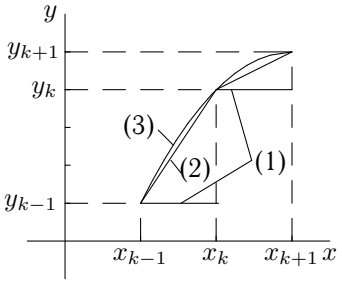


Рис. 4.1
Интерполирование

индексу k, который пробегает значения от 0 до n (n — номер последней точки), прибавить единицу.

На рис. 4.1 цифрой (2) помечена ломаная прямая, соответствующая линейной интерполяции.

Линейная интерполяция дает возможность описать функциональную зависимость, как правило, более точно, чем при кусочно-постоянной интерполяции. Тем не менее в некоторых задачах и она не дает достаточной точности.

Следующим шагом в повышении точности аналитического описания функции, заданной таблично, является *квадратичная интерполяция* кривой $y = ax^2 + bx + c$. Коэффициенты a, b и c находятся из условия, что парабола проходит через три точки $(x_{k-1}, y_{k-1}), (x_k, y_k)$ и (x_{k+1}, y_{k+1}). Подстановка этих координат в уравнение параболы дает возможность получить систему трех уравнений с тремя неизвестными a, b и c, из которой найти затем эти неизвестные. Из-за неоправданной громоздкости преобразований и получаемых соотношений они здесь не приводятся.

На рис. 4.1 цифрой (3) отмечена кривая квадратичной интерполяции.

Пример. Функция $y = f(x)$ задана таблицей:

x	1	2	3
y	0,5	2	2,5

Требуется осуществить ее интерполирование: (1) — ступенчатой функцией; (2) — кусочно-линейной; (3) — квадратичной.

Р е ш е н и е.

(1). Ступенчатая функция может быть представлена совокупностью двух прямых, параллельных оси x :

$$y = \begin{cases} 0{,}5, & \text{если } x \in [1,\ 2); \\ 2, & \text{если } x \in [2,\ 3]. \end{cases}$$

(2). Линейная интерполяция представляется зависимостью:

$$y = \begin{cases} 1{,}5x - 1, & \text{если } x \in [1,\ 2); \\ 0{,}5x + 1, & \text{если } x \in [2,\ 3]. \end{cases}$$

В справедливости последних соотношений легко убедиться. Они обращаются в тождества подстановкой координат табличных точек.

(3). Квадратичная интерполяция приводит к функции:

$$y = -0{,}5x^2 + 3x - 2,$$

которая обращается в тождество при табличных значениях координат.

Интерполирующие зависимости рассмотренного примера изображены графиками на рис. 4.1. Примеру соответствуют следующие данные: $x_{k-1} = 1$; $x_k = 2$; $x_{k+1} = 3$; $y_{k-1} = 0{,}5$; $y_k = 2$; $y_{k+1} = 2{,}5$.

Имеют право на использование и другие зависимости, например, интерполирование полиномами порядка выше 2 или трансцендентными функциями. Повышенную точность и «гладкость интерполяции» могут обеспечить *сплайн-функции*. При формировании таких функций в заданных точках обеспечивается непрерывность не только аппроксимирующей функции, но и ее производных.

Требование точности аппроксимации должно соответствовать точности статистического материала. Если исходные данные получены с точностью, например, до 5%, то применение аппроксимации полиномами степени выше второй, трансцендентными или сплайн-функциями лишено смысла.

Требуемую точность аппроксимации можно обеспечить не усложнением аппроксимирующих функций, а увеличением количества статистических данных.

§ 4.3. ПРИМЕР КОРРЕЛЯЦИОННОЙ ЗАВИСИМОСТИ

Пусть известны результаты экзаменов и время в часах, затраченное на подготовку к нему, среди опрошенных 100 студентов. После перехода к осредненным дискретным распределениям (за значения переменных приняты середины интервалов времени подготовки и оценок, проставляе-

мых преподавателем с точностью до десятых долей единицы) получено распределение в виде зависимости оценки $X = (2, 3, 4, 5)$ от затрат времени (в часах) на подготовку к экзамену $Y = (25, 30, 35, 40, 45, 50)$.

Результаты сведены в табл. 4.1.

Таблица 4.1
Корреляционная таблица примера

y_j / x_i	25	30	35	40	45	50	$\sum$	$\overline{y}_{xi}$
2	9	4	1	–	–	–	14	27,1
3	1	10	9	3	–	–	23	33,0
4	–	2	8	18	9	1	38	39,9
5	–	–	1	9	8	7	25	44,2
$\sum$	10	16	19	30	17	8	100	
$\overline{x}_{yj}$	2,1	2,88	3,47	4,20	4,47	4,88		

Таблицы такого типа со статистическим характером переменных называются *корреляционными*.

Сосредоточение показателей около диагонали таблицы говорит о том, что в данном испытании наблюдается тенденция увеличения оценки при росте времени, затрачиваемого студентами на подготовку к экзамену.

К построению аналитической зависимости между случайными переменными X и Y (построению математической модели задачи) можно подойти с разных позиций. Можно, например, каждому значению величины $x_i \in X$ $(i = \overline{2,5})$ поставить в соответствие среднее арифметическое $\overline{y}_{xi}$ времени, затраченного среднестатистическим студентом, получившим на экзамене оценку x_i.

Так, оценку $x_2 = 2$ получили 9 студентов, затративших на подготовку к экзамену в среднем 25 часов, 4 студента — 30 часов и 1 студент — 35 часов, поэтому при суммарном

количестве студентов, получивших на экзамене оценку 2, равном 14, получим:

$$\overline{y}_{x2} = (25 \cdot 9 + 30 \cdot 4 + 35 \cdot 1)/14 \approx 27{,}1.$$

Аналогично определяются средние арифметические затрат времени на подготовку к экзамену для получения оценок 3, 4 и 5:

$\overline{y}_{x3} = (25 \cdot 1 + 30 \cdot 10 + 35 \cdot 9 + 40 \cdot 3)/23 \approx 33{,}0.$
$\overline{y}_{x4} = (30 \cdot 2 + 35 \cdot 8 + 40 \cdot 18 + 45 \cdot 9 + 50 \cdot 1)/38 \approx 39{,}9.$
$\overline{y}_{x5} = (35 \cdot 1 + 40 \cdot 9 + 45 \cdot 8 + 50 \cdot 7)/25 \approx 44{,}2.$

Полученные данные занесены в последний столбец табл. 4.1 и в виде точек, обозначенных значком • , на рис. 4.2.

Из рисунка и таблицы видно, что с увеличением X усредненные по группам значения Y увеличиваются. При этом зависимость X от Y похожа на линейную и поэтому может быть описана уравнением прямой:

$$\overline{y} = ax + b. \tag{4.1}$$

Рассмотрим еще один подход к формированию математической модели, описывающей закономерность зависимости оценки от затрат времени на подготовку к экзамену.

По числу различных значений переменной y имеем шесть групп показателей y_j $(j = \overline{1,6})$ затрат времени на подготовку к экзамену. Для каждой из этих групп подсчитаем средние

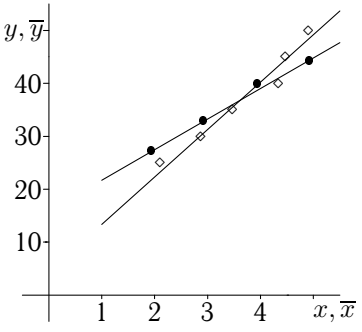


Рис. 4.2
К примеру

арифметические оценки $\overline{x}_{yj}$, полученные на экзамене при затратах на подготовку соответственно $y_1 = 25$ час, $y_2 = 30$ час, $\ldots, y_6 = 50$ час.

$\overline{x}_{y1} = (2 \cdot 9 + 3 \cdot 1)/10 \approx 2{,}10.$
$\overline{x}_{y2} = (2 \cdot 4 + 3 \cdot 10 + 4 \cdot 2)/16 \approx 2{,}88.$
$\overline{x}_{y3} = (2 \cdot 1 + 3 \cdot 9 + 4 \cdot 8 + 5 \cdot 1)/19 \approx 3{,}48.$
$\overline{x}_{y4} = (3 \cdot 3 + 4 \cdot 18 + 5 \cdot 9)/30 \approx 4{,}34.$
$\overline{x}_{y5} = (4 \cdot 9 + 5 \cdot 8)/17 \approx 4{,}47.$
$\overline{x}_{y6} = (4 \cdot 1 + 5 \cdot 7)/8 \approx 4{,}88.$

Полученные данные занесены в последнюю строку табл. 4.1 и отмечены значками $\diamond$ на рис. 4.2. Расположение точек подтверждает отмеченную ранее близкую к линейной зависимость переменных Y от X.

Точки на графике функции, полученные в первом варианте обработки исходных данных, не совпадают с точками второго варианта. Поэтому будем считать, что во втором варианте линейная зависимость описывается прямой

$$\overline{x} = cy + d. \tag{4.2}$$

Рассмотренный пример является иллюстрацией задачи обработки исходного статистического материала для установления вида зависимости между статистическими переменными. Оставляя пока открытым вопрос об определении коэффициентов в уравнениях прямых (4.1) и (4.2) (прямые изображены на рис. 4.2), перейдем к изложению задачи в общем виде.

§ 4.4. КОРРЕЛЯЦИОННАЯ ТАБЛИЦА

Пусть имеются n наблюдений пар случайных величин (X;Y) — случайного вектора. При этом $X = \{x_1, x_2, \ldots, x_i, \ldots, x_k\}$ и $Y = \{y_1, y_2, \ldots, y_j, \ldots, y_m\}$. Пусть каждая пара элементов $(x_i; y_j)$ встречается среди n таких пар с частотой n_{ij}.

Результаты наблюдений представлены в виде *корреляционной таблицы*.

Смысл элементов, входящих в таблицу, следующий.

Значения частот n_{ij} $(i = \overline{1,k} \quad j = \overline{1,m})$ показывают, сколько элементов совокупности n общего числа испытаний соответствует случаям, когда переменная X принимает

Таблица 4.2
Корреляционная таблица

Y / X	y_1	y_2	$\cdots$	y_j	$\cdots$	y_m	$\sum_{j=1}^{m} n_{ij}$	$\overline{y}_{xi}$
x_1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	n_{1m}	n_{x1}	$\overline{y}_{x1}$
x_2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	n_{2m}	n_{x2}	$\overline{y}_{x2}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		$\cdots$	$\cdots$
x_i	n_{i1}	n_{i2}	$\cdots$	n_{ij}	$\cdots$	n_{im}	n_{xi}	$\overline{y}_{xi}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		$\cdots$	$\cdots$
x_k	n_{k1}	n_{k2}	$\cdots$	n_{kj}	$\cdots$	n_{km}	n_{xk}	$\overline{y}_{xk}$
$\sum_{i=1}^{k} n_{ij}$	n_{y1}	n_{y2}	$\cdots$	n_{yj}	$\cdots$	n_{ym}	n	$\overline{y}$
$\overline{x}_{yj}$	$\overline{x}_{y1}$	$\overline{x}_{y2}$	$\cdots$	$\overline{x}_{yj}$	$\cdots$	$\overline{x}_{ym}$	$\overline{x}$	

значения x_i, а переменная Y — значения y_j. Сумма всех таких частот, занесенных в таблицу, равна n:

$$\sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij} = n.$$

Переменная X принимает k различных (в общем случае) значений x_1, x_2, …, x_i, …, x_k с частотами соответственно n_{x1}, n_{x2}, …, n_{xi} …, n_{xk}. Переменная Y — m различных значений y_1, y_2, …, y_j, …, y_m с частотами n_{y1}, n_{y2},…, n_{yj}, …, n_{ym}.

При этом

$$n_{xi} = \sum_{j=1}^{m} n_{ij}; \qquad n_{yj} = \sum_{i=1}^{k} n_{ij}.$$

Кроме того,

$$\sum_{i=1}^{k} n_{xi} = \sum_{j=1}^{m} n_{yj} = n.$$

Любая из частот n_{ij} может в частных случаях равняться нулю, но все частоты любой строки (столбца) одновременно равняться нулю не должны. Если последнее имеет место, следует соответствующую строку (столбец) вычеркнуть из таблицы, как не несущую никакой информации.

Совокупности пар чисел $(x_i; n_{xi})$ или пар чисел $(y_j; n_{yj})$ представляют собой эмпирические распределения одномерных случайных величин X и Y соответственно.

В последнем столбце таблицы приводятся *условные средние* $\overline{y}_{xi}$ случайной величины Y, соответствующие значению x_i случайной величины X:

$$\overline{y}_{xi} = \frac{1}{n_{xi}} \sum_{j=1}^{m} y_j n_{ij}.$$

В последней строке таблицы приводятся *условные средние* $\overline{x}_{yj}$ случайной величины X, соответствующие величине переменной Y, равной y_j:

$$\overline{x}_{yj} = \frac{1}{n_{yj}} \sum_{i=1}^{k} x_i n_{ij}.$$

По данным таблицы можно определить выборочные средние значений переменных X и Y:

$$\overline{x} = \frac{\sum\limits_{j=1}^{m} \overline{x}_{yj} n_{yj}}{\sum\limits_{j=1}^{m} n_{yj}} = \frac{1}{n} \sum_{j=1}^{m} \overline{x}_{yj} n_{yj};$$

$$\overline{y} = \frac{\sum\limits_{i=1}^{k} \overline{y}_{xi} n_{xi}}{\sum\limits_{i=1}^{k} n_{xi}} = \frac{1}{n} \sum_{i=1}^{k} \overline{y}_{xi} n_{xi}.$$

Если между переменной X, значения которой $\{x_1, x_2, \ldots, x_k\}$ приведены в первом столбце табл. 4.2, и переменной Y, представленной осредненными значениями $\{\overline{y}_{x1}, \overline{y}_{x1}, \ldots, \overline{y}_{xk}\}$ последнего столбца таблицы, просматривается функциональная (определенная) зависимость, то ее можно представить в виде функции:

$$\overline{y} = f(x). \tag{4.3}$$

График этой функции представляет собой кривую, вблизи которой располагаются точки $(x_i, \overline{y}_{xi})$.

В корреляционной таблице 4.2 переменные X и Y равноправны. Поэтому можно смоделировать функциональную зависимость:

$$\overline{x} = g(y). \tag{4.4}$$

График этой функции в общем случае представляет собой кривую, вблизи которой располагаются точки $(\overline{x}_{yj}, y_j)$, значения координат которых приведены в последней и первой строках таблицы.

Двухмерную выборку большого объема представляют в виде группированной корреляционной таблицы, в которой реализации случайных величин X и Y группируют по интервалам аналогично одномерному случаю. Пример такого случая представлен таблицей 4.3.

Таблица 4.3
Пример корреляционной таблицы

Y / X	[5;15]	(15;25]	(25;35]	n_{xi}
[10;20]	5	7	0	12
(20;30]	8	17	23	48
n_{yj}	13	24	23	60

Корреляционной зависимостью между двумя переменными называют функциональную зависимость (4.3) или (4.4) между одной из них и соответствующими средними арифметическими (точечными) другой.

Зависимость, определяемую уравнением (4.3), называют *корреляционной зависимостью Y на X*; (4.4) — *корреляционной зависимостью X на Y*. Соответствующие этим зависимостям уравнения (4.3) и (4.4) называются *уравнениями регрессии* соответственно Y на X и X на Y. Кривые этих функций называют *кривыми регрессии*.

Методами построения математических моделей, наилучшим образом соответствующих экспериментальным данным, и дальнейшим анализом этих соотношений занимается раздел математики, называемый *регрессионным анализом*.

Возвращаясь к корреляционному анализу, перечислим его основные задачи:

1) определение функциональной связи между случайными переменными, в том числе определение постоянных, входящих в эту связь (например, коэффициентов a, b, c, d уравнений (4.1), (4.2));

2) установление тесноты связи между одной случайной переменной, входящей в зависимость, и осредненными статистическими значениями другой случайной переменной.

§ 4.5. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

В практике математической обработки результатов экспериментов независимо от того, какие процессы они описывают (детерминированные или случайные), широкое распространение получил *метод наименьших квадратов*. Опишем суть этого метода.

Пусть имеется таблица исходных данных, представляющих собой k пар элементов (декартовых произведений) независимых величин $X = (x_1, \ldots, x_i, \ldots, x_k)$ и $Y = (y_1, y_2, \ldots, y_j, \ldots \ldots, y_m)$, например, совокупности величин, приведенных в табл. 4.2. Эти величины можно нанести на координатную плоскость декартовой ортогональной системы координат, как это было сделано на рис. 4.2.

Требуется построить такую функцию $y = f(x)$, которая с наибольшей точностью описывает представленную в виде опытных данных зависимость координат y_j от x_i $(j = \overline{1,m})$, $(i = \overline{1,k})$.

В методе наименьших квадратов точность описания экспериментальных данных аналитической зависимостью определяется требованием минимальности значений суммы квадратов отклонений ординат (или абсцисс) точек, построенных по опытным данным, от графика аналитически заданной функции.

Наиболее распространенными на практике аналитическими функциями для описания опытных данных являются линейная и квадратичная функции. Рассмотрим каждую из них.

§ 4.6. ЛИНЕЙНАЯ АППРОКСИМАЦИЯ

Пусть по характеру расположения заданных точек (x_i, y_i) $(i = \overline{1,k})$ для их аппроксимации исследователем выбрана линейная функция (прямая) $y = ax + b$. Точки с координатами (x_i, y_i) в общем случае не лежат на прямой и их координаты не будут обращать в тождество ее уравнение.

При подстановке абсцисс x_i точек в уравнение прямой получим соответствующие им (по уравнению прямой) ординаты $\tilde{y}_i = ax_i + b$, не совпадающие в общем с y_i. Найдем сумму квадратов отклонений значений функции $\tilde{y}_i$ от значений ординат y_i заданных точек:

$$S = \sum_{i=1}^{k} (\tilde{y}_i - y_i)^2 = \sum_{i=1}^{k} (ax_i + b - y_i)^2. \qquad (4.5)$$

Отметим, что x_i и y_i в записи (4.5) фиксированы (постоянны) для точек, а коэффициенты a и b варьируются (считаются изменяемыми). Именно от различных вариантов выбора этих коэффициентов зависит положение прямой на плоскости, и они влияют на величину суммарного отклонения прямой от экспериментальных точек.

Для определения минимума функции $S(a,b)$ двух варьируемых переменных a и b найдем и приравняем нулю производные от S по этим переменным:

$$\begin{cases} \dfrac{\partial S}{\partial a} = \displaystyle\sum_{i=1}^{k} 2(ax_i + b - y_i)x_i = 0, \\ \dfrac{\partial S}{\partial b} = \displaystyle\sum_{i=1}^{k} 2(ax_i + b - y_i) = 0. \end{cases}$$

Коэффициенты a и b внутри сумм не изменяются и их можно вынести за знак сумм. Учитывая, кроме того, равенство $\displaystyle\sum_{i=1}^{k} 1 = k$, получим систему двух уравнений для определения коэффициентов a и b, называемую *системой нормальных уравнений*:

$$\begin{cases} a\displaystyle\sum_{i=1}^{k} x_i^2 + b\sum_{i=1}^{k} x_i = \sum_{i=1}^{k} x_i y_i, \\ a\displaystyle\sum_{i=1}^{k} x_i + bk = \sum_{i=1}^{k} y_i. \end{cases} \tag{4.6}$$

Пример 1. Аппроксимировать зависимость между переменными y_i и x_i, приведенными в табл. 4.1, линейной функцией $y = ax + b$.

Р е ш е н и е. Полагая в формуле (4.6) $y_i = \overline{y}_{xi}$ (значения ординат берутся из последнего столбца табл. 4.1), для вычисления a и b заполним левую табл. 4.4.

Используя данные таблицы, запишем систему уравнений (4.6):

$$\begin{cases} 54a + 14b = 533{,}8; \\ 14a + 4b = 144{,}2. \end{cases}$$

Решая систему, найдем: $a = 5{,}82$; $b = 15{,}68$. Тогда уравнение искомой прямой:

$$y = 5{,}82x + 15{,}68.$$

Таблица 4.4

Линейная регрессия Y на X (слева) и X на Y (справа)

i	x_i	y_i	x_i^2	x_iy_i
1	2	27,1	4	54,2
2	3	33,0	9	99,0
3	4	39,9	16	159,6
4	5	44,2	25	221,0
$\sum$	14	144,2	54	533,8

j	y_j	x_j	y_j^2	x_jy_j
1	25	2,10	625	52,5
2	30	2,88	900	86,4
3	35	3,47	1225	121,5
4	40	4,20	1600	168,0
5	45	4,47	2025	201,1
6	50	4,88	2500	244,0
$\sum$	225	22,0	8875	873,5

На рис. 4.2 этой прямой соответствует линия с меньшим углом наклона к оси абсцисс.

В терминах предыдущего раздела полученная прямая является прямой регрессии Y на X. Для получения прямой регрессии X на Y необходимо минимизировать отклонения опытных данных от выбранной прямой $x = cy + d$ по абсциссам соответствующих точек. Такая минимизация, как нетрудно проверить, приведет к системе двух нормальных уравнений:

$$\begin{cases} c\sum\limits_{j=1}^{m} y_j^2 + d\sum\limits_{j=1}^{m} y_j = \sum\limits_{j=1}^{m} x_jy_j, \\ c\sum\limits_{j=1}^{m} y_j + dm = \sum\limits_{j=1}^{m} x_j. \end{cases} \tag{4.7}$$

Принимая за x_j величины $\overline{x}_{yj}$ последней строки табл. 4.1, заполним правую табл. 4.4. После определения элементов последней строки таблицы подставим их в систему уравнений (4.7):

$$\begin{cases} 8875c + 225d = 873{,}5; \\ 225c + 6d = 22{,}0. \end{cases}$$

Решениями системы являются значения $c = 0{,}111$; и $d = -0{,}496$, которые позволяют записать уравнения прямой регрессии X на Y в виде

$$x = 0{,}111y - 0{,}496.$$

Полученная прямая изображена на рис. 4.2. Эта прямая имеет больший по сравнению с первой прямой угол наклона к оси абсцисс.

Отметим, что систему уравнений (4.6) можно преобразовать к виду, во многих случаях более удобному для определения коэффициентов a и b. Заметим, что

$$\sum_{i=1}^{k} x_i = k\overline{x}; \qquad \sum_{i=1}^{k} y_i = k\overline{y},$$

где $\overline{x} = \dfrac{1}{k}\sum\limits_{i=1}^{k} x_i$, $\overline{y} = \dfrac{1}{k}\sum\limits_{i=1}^{k} y_i$ — средние арифметические значения x_i и y_i в пределах изменения $i = \overline{1,k}$. Кроме того,

$$\sum_{i=1}^{k} x_i^2 = \sum_{i=1}^{k}(x_i - \overline{x})^2 + 2\sum_{i=1}^{k} x_i\overline{x} - \sum_{i=1}^{k}\overline{x}^2 =$$

$$= \sum_{i=1}^{k}(x_i - \overline{x})^2 + k\overline{x}^2$$

и

$$\sum_{i=1}^{k} x_i y_i = \sum_{i=1}^{k}(x_i - \overline{x})(y_i - \overline{y}) + \sum_{i=1}^{k} x_i\overline{y} + \sum_{i=1}^{k} y_i\overline{x} - \sum_{i=1}^{k}\overline{xy} =$$

$$= \sum_{i=1}^{k}(x_i - \overline{x})(y_i - \overline{y}) + k\overline{xy}.$$

Подставим записанные соотношения во второе уравнение системы (4.6), а затем найденное для b выражение в первое уравнение. После преобразований получим:

$$a = \frac{\sum\limits_{i=1}^{k}(x_i - \overline{x})(y_i - \overline{y})}{\sum\limits_{i=1}^{k}(x_i - \overline{x})^2};$$

$$b = \overline{y} - a\overline{x}. \tag{4.8}$$

§ 4.7. КВАДРАТИЧНАЯ АППРОКСИМАЦИЯ

Далеко не всегда уравнения линейной аппроксимации достаточно корректно описывают экспериментальные данные. Рассмотрим еще один широко используемый на практике метод наименьших квадратов — метод квадратичной аппроксимации.

Пусть, как и в предыдущем разделе, экспериментальные данные зависимости между величинами $X=\{x_1, x_2, \dots x_i, \dots \dots x_k\}$ и $Y=\{y_1, y_2, \dots y_i, \dots y_k\}$ представлены в виде таблицы и нанесены на график декартовой ортогональной системы координат xy.

Пусть по виду расположения точек на графике можно предположить, что проводимая в окрестности расположения точек кривая может представлять собой параболу $y = ax^2 + bx + c$, имеющую три варьируемые переменные: a, b и c. Следуя схеме метода наименьших квадратов, изложенного в предыдущем разделе, составим разности ординат (отклонений) точек параболы $\tilde{y}_i = ax_i^2 + bx_i + c$ и опытных данных y_i $(i = \overline{1,k})$.

Запишем выражение для суммы квадратов полученных отклонений:

$$S = \sum_{i=1}^{k} (\tilde{y}_i - y_i)^2 = \sum_{i=1}^{k} (ax_i^2 + bx_i + c - y_i)^2.$$

Для определения минимального значения суммы приравняем нулю производные от S по варьируемым переменным a, b и c:

$$\begin{cases} \dfrac{\partial S}{\partial a} = \displaystyle\sum_{i=1}^{k} 2(ax_i^2 + bx_i + c - y_i)x_i^2 = 0, \\ \dfrac{\partial S}{\partial b} = \displaystyle\sum_{i=1}^{k} 2(ax_i^2 + bx_i + c - y_i)x_i = 0, \\ \dfrac{\partial S}{\partial c} = \displaystyle\sum_{i=1}^{k} 2(ax_i^2 + bx_i + c - y_i) = 0. \end{cases}$$

Отсюда получаем систему трех нормальных уравнений с тремя неизвестными — коэффициентами a, b и c — для квадратичной аппроксимации:

$$\begin{cases} a\sum\limits_{i=1}^{k} x_i^4 + b\sum\limits_{i=1}^{k} x_i^3 + c\sum\limits_{i=1}^{k} x_i^2 = \sum\limits_{i=1}^{k} x_i^2 y_i, \\ a\sum\limits_{i=1}^{k} x_i^3 + b\sum\limits_{i=1}^{k} x_i^2 + c\sum\limits_{i=1}^{k} x_i = \sum\limits_{i=1}^{k} x_i y_i, \\ a\sum\limits_{i=1}^{k} x_i^2 + b\sum\limits_{i=1}^{k} x_i + ck = \sum\limits_{i=1}^{k} y_i. \end{cases} \tag{4.9}$$

Сравнение последней системы уравнений с (4.6) говорит о существенном усложнении задачи и увеличении количества вычислений для определения параметров кривой второго порядка по сравнению с линейной аппроксимацией. Нетрудно убедиться в том, что аппроксимация опытных данных функциями, содержащими полиномы порядка, большего двух, а тем более аппроксимация произвольными функциями может привести к неразрешимым трудностям. Этим объясняется тот факт, что на практике редко встречаются основанные на использовании метода наименьших квадратов аппроксимации, отличные от линейной и параболической.

При невозможности описания опытных данных простейшими функциями можно использовать прием разбиения всего интервала аппроксимации на подынтервалы, на которых поведение функции можно считать подчиняющимся, например, линейной аппроксимации.

Замечание. В § 4.2 при рассмотрении задач аппроксимации говорилось о том, что линейные функции, как правило, удовлетворяют задаваемой точности решения. Задача построения кривых регрессии отличается от задачи «локальной аппроксимации» тем, что построение кривых регрессии преследует цель проведения дальнейшего анализа исследуемого процесса. Требуется установить общие (не локальные) закономерности и, зачастую, дать прогноз будущего характера процесса. Поэтому в задачах построения моделей регрессии линейные

функции далеко не всегда могут привести к удовлетворительным результатам.

Пример. Данные некоторого опыта представлены в виде табл. 4.5 пар значений $(x_i; y_i)$ переменных X и Y. Требуется подобрать аппроксимирующую функцию представленной зависимости.

Таблица 4.5
Исходные данные

x_i	0,5	1	1,5	2	2,5
y_i	0,8	1,9	4,9	8,8	13,9

Р е ш е н и е.

Нанесем исходные данные на координатную плоскость декартовой ортогональной системы координат xy (точки $\circ$ на рис. 4.3).

Данные для определения коэффициентов системы уравнений (4.9) и сами коэффициенты занесем в табл. 4.6.

Подставляя значения последней строки табл. 4.6 вместо соответствующих коэффициентов уравнений (4.9), получим

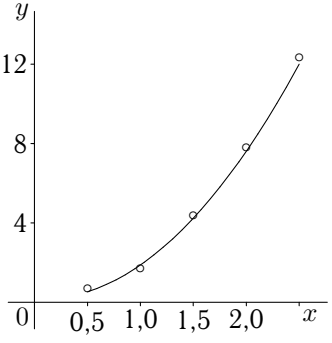


Рис. 4.3
Квадратичная аппроксимация

Таблица 4.6
Подсчет коэффициентов уравнений (4.9)

i	x_i	x_i^2	x_i^3	x_i^4	y_i	x_iy_i	$x_i^2y_i$
1	0,5	0,25	0,125	0,0625	0,8	0,4	0,2
2	1	1	1	1	1,9	1,9	1,9
3	1,5	2,25	3,375	5,0625	4,9	7,35	11,025
4	2	4	8	16	8,8	17,6	35,2
5	2,5	6,25	15,625	39,0625	13,9	34,75	86,875
$\sum_{i=1}^{5}$	7,5	13,75	28,125	61,1875	30,3	62,0	135,2

систему трех алгебраических уравнений:

$$\begin{cases} 61{,}1875a + 28{,}125b + 13{,}75c = 135{,}2; \\ 28{,}125a + 13{,}75b + 7{,}5c = 62; \\ 13{,}75a + 7{,}5b + 5c = 30{,}3. \end{cases}$$

Решениями системы будут значения: $a = 2{,}54$; $b = -1$; $c = 0{,}575$. Аппроксимирующая функция примет вид

$$y = 2{,}54x^2 - x + 0{,}575.$$

График функции показан на рис. 4.3.

Метод наименьших квадратов, описанный в двух последних разделах, может быть применен к аналитическому описанию экспериментальных данных независимо от того, какой характер (случайный или детерминированный) они носят. Рассмотрим особенности применения метода к описанию корреляционных зависимостей.

§ 4.8. УРАВНЕНИЯ ЛИНЕЙНОЙ РЕГРЕССИИ

Пусть имеются данные некоторых испытаний двух зависимых случайных величин $X = \{x_1, x_2, \ldots, x_i, \ldots, x_k\}$ и $Y = \{y_1, y_2, \ldots, y_j, \ldots, y_m\}$ (табл. 4.2) и пусть в результате обработки результатов этих испытаний получены условные средние арифметические значения $\overline{Y} = \{\overline{y}_{x1}, \overline{y}_{x2}, \ldots, \overline{y}_{xi}, \ldots, \overline{y}_{xk}\}$

случайной величины Y. Величины $\overline{y}_{xi}$ представлены в последнем столбце (табл. 4.2).

Пусть по данным указанной таблицы определены k точек $A_1(x_1, \overline{y}_{x1})$, $A_2(x_2, \overline{y}_{x2}), \ldots, A_i(x_i, \overline{y}_{xi}), \ldots, A_k(x_k, \overline{y}_{xk})$. По расположению этих точек предполагаем, что между случайными величинами существует линейная зависимость (рис. 4.4), которую опишем уравнением линейной регрессии (прямой регрессии):

$$y = ax + b. \tag{4.10}$$

Коэффициенты a и b найдем, используя метод наименьших квадратов. Для этого определим разности ординат точек $A_i(x_i, \overline{y}_{xi})$ и соответствующих им точек A'_i прямой (4.10), имеющих абсциссы, равные x_i:

$$y_i = ax_i + b.$$

Используя идеи методов наименьших квадратов и построения прямых регрессии Y на X, определим коэффициенты уравнения (4.10) так, чтобы сумма квадратов отклонений ординат точек опытных данных от их теоретических значений, соответствующих точкам прямой, была минимальной. При этом, в отличие от изложенного в § 4.4 метода наименьших квадратов, квадраты отклонений $(y_i - \overline{y}_{xi}) = (ax_i + b - \overline{y}_{xi})$ под знаком суммы будем умножать на относительные частоты $\overline{n}_{xi} = n_{xi}/n$ $\left(n_{xi} = \sum\limits_{j=1}^{m} n_{ij};\ \sum\limits_{i=1}^{k} n_{xi} = n,\ m \cdot k = n\right)$, о чем было

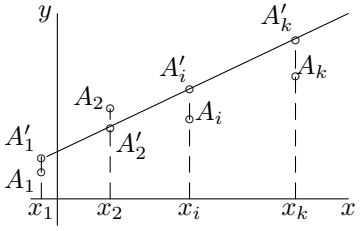


Рис. 4.4
Линейная регрессия Y на X

сказано в § 4.4). Такое умножение не изменит сути отыскания минимума суммы, так как сумма относительных частот равна единице.

Итак, указанную сумму представим в виде

$$S = \frac{1}{n}\sum_{i=1}^{k}(ax_i + b - \overline{y}_{xi})^2 n_{xi}. \tag{4.11}$$

Найдем и приравняем нулю производные от S по a и b:

$$\begin{cases} \dfrac{\partial S}{\partial a} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{k} 2(ax_i + b - \overline{y}_{xi})x_i n_{xi} = 0, \\ \dfrac{\partial S}{\partial b} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{k} 2(ax_i + b - \overline{y}_{xi})n_{xi} = 0. \end{cases}$$

Поделив оба уравнения на 2 и группируя слагаемые, придем к системе уравнений для определения a и b:

$$\begin{cases} a\dfrac{1}{n}\displaystyle\sum_{i=1}^{k} x_i^2 n_{xi} + b\frac{1}{n}\sum_{i=1}^{k} x_i n_{xi} = \frac{1}{n}\sum_{i=1}^{k} x_i \overline{y}_{xi} n_{xi}, \\ a\dfrac{1}{n}\displaystyle\sum_{i=1}^{k} x_i n_{xi} + b = \frac{1}{n}\sum_{i=1}^{k} \overline{y}_{xi} n_{xi}. \end{cases} \tag{4.12}$$

Коэффициент при b в последнем уравнении равен единице в силу справедливости отмеченного ранее равенства: $\dfrac{1}{n}\displaystyle\sum_{i=1}^{k} n_{xi} = 1$. Проанализируем другие коэффициенты уравнений (4.12).

Коэффициент при a в первом уравнении представляет собой выборочное среднее квадратов случайной величины X:

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_{xi}.$$

Аналогично запишем выражения для выборочных средних других величин, входящих в уравнения (4.12):

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{k}x_i n_{xi}.\quad \overline{y}=\frac{1}{n}\sum_{i=1}^{k}\overline{y}_{xi}n_{xi}.\quad \overline{xy}=\frac{1}{n}\sum_{i=1}^{k}x_i\overline{y}_{xi}n_{xi}.$$

С учетом последних соотношений перепишем систему уравнений (4.12):

$$\begin{cases} a\overline{x^2}+b\overline{x}=\overline{xy}, \\ a\overline{x}+b=\overline{y}. \end{cases} \tag{4.13}$$

Определим b из второго уравнения: $b=\overline{y}-a\overline{x}$. После подстановки этого выражения в уравнение (4.10) придем к уравнению прямой, проходящей через точку $(\overline{x},\overline{y})$ корреляционного графика:

$$y-\overline{y}=a(x-\overline{x}).$$

Коэффициент $a=\rho_{Y/X}$, называемый *коэффициентом регрессии Y на X*, находится при определенном выше выражении для b из первого уравнения (4.13):

$$\rho_{Y/X}=\frac{\overline{xy}-\overline{x}\cdot\overline{y}}{\overline{x^2}-\overline{x}^2}.$$

Числитель дроби $\overline{\sigma}_{XY}=\overline{xy}-\overline{x}\cdot\overline{y}$ называется *выборочной ковариацией*. Знаменатель представляет собой выборочную дисперсию $\overline{\sigma}_X^2$ случайной величины X. Действительно, $(\overline{n}_{xi}=n_{xi}/n)$:

$$\overline{\sigma}_X^2=\frac{1}{n}\sum_{i=1}^{k}(x_i-\overline{x})^2n_{xi}=\sum_{i=1}^{k}x_i^2\overline{n}_{xi}-2\overline{x}\sum_{i=1}^{k}\overline{x}_i\overline{n}_{xi}+\overline{x}^2\sum_{i=1}^{k}\overline{n}_{xi}=$$

$$=\overline{x^2}-2\overline{x}\cdot\overline{x}+\overline{x}^2=\overline{x^2}-\overline{x}^2.$$

Поэтому

$$\rho_{Y/X}=\frac{\overline{\sigma}_{XY}}{\overline{\sigma}_X^2}. \tag{4.14}$$

Итак, уравнение линейной регрессии Y на X имеет вид

$$y-\overline{y}=\rho_{Y/X}(x-\overline{x}). \tag{4.15}$$

Аналогичное уравнение линейной регрессии X на Y можно получить, строя зависимость между значениями $Y = \{y_1, y_2, \ldots, y_j, \ldots, y_m\}$ первой строки табл. 4.2 и значениями $\overline{X} = \{\overline{x}_{y1}, \overline{x}_{y2}, \ldots, \overline{x}_{yj}, \ldots, \overline{x}_{ym}\}$ последней строки этой таблицы:

$$x - \overline{x} = \rho_{X/Y}(y - \overline{y}). \tag{4.16}$$

Здесь

$$\rho_{X/Y} = \frac{\overline{\sigma}_{XY}}{\overline{\sigma}^2_Y} \tag{4.17}$$

коэффициент регрессии X на Y.

§ 4.9. КОЭФФИЦИЕНТ КОРРЕЛЯЦИИ И КОРРЕЛЯЦИОННЫЕ ОТНОШЕНИЯ

Выбором корреляционной зависимости не заканчивается задача описания связи двух случайных величин. Необходимо еще оценить корректность этой связи, то есть провести оценку рассеяния относительно аппроксимирующей кривой (в частном случае прямой) регрессии одной из переменных для различных значений другой. Рассмотрим дополнительно некоторые новые характеристики случайных явлений, описываемых двумя случайными величинами.

Коэффициентом корреляции случайных величин X и Y называется число r_{XY}, равное среднему геометрическому их коэффициентов регрессии и имеющее их знак. То есть

$$r_{XY} = r_{YX} = \pm\sqrt{\rho_{X/Y}\rho_{Y/X}}. \tag{4.18}$$

Поясним, что знаки у коэффициентов регрессии всегда одинаковы и зависят от знака числителя в формулах (4.14) или (4.17) (знаменатель всегда положителен).

С учетом упомянутых формул получим:

$$r_{XY} = \frac{\overline{\sigma}_{XY}}{\overline{\sigma}_X \overline{\sigma}_Y}. \tag{4.19}$$

Легко убедиться в справедливости формул, выражающих связь между коэффициентами регрессии и коэффициентом

корреляции:

$$\rho_{Y/X} = r_{XY}\frac{\sigma_Y}{\sigma_X}, \quad \rho_{X/Y} = r_{YX}\frac{\sigma_X}{\sigma_Y}. \tag{4.20}$$

Свойства коэффициента корреляции

1. *Абсолютная величина коэффициента корреляции не превосходит единицы*: $|r_{XY}| \le 1$.

2. *Выполнение условия* $r_{XY} = \pm 1$ *является необходимым и достаточным для того, чтобы* Y *и* X *были связаны линейной корреляционной зависимостью.*

3. *Если* $r_{XY} = 0$, *то случайные величины* X *и* Y *называются некоррелированными.*

Свойства 2 и 3 поясняют смысл крайних значений коэффициента корреляции: если он равен нулю, то линейная корреляционная связь отсутствует; если же он равен ± 1, то между X и Y существует линейная корреляционная зависимость.

Для нелинейной регрессии вместо соотношений (4.20) для оценки тесноты связи между переменными используется индекс корреляции:

$$\rho_{XY} = \sqrt{1 - \frac{\sum\limits_{i=1}^{n}(y_i - \tilde{y})^2}{\sum\limits_{i=1}^{n}(y_i - \overline{y})^2}}, \tag{4.21}$$

причем, $0 \le \rho_{XY} \le 1$.

Межгрупповая дисперсия δ_X^2 случайной величины X является дисперсией распределения групповых средних $\overline{x}_{yj}$ с относительными частотами их повторений n_{yj}. Зависимость между $\overline{x}_{yi}$ и y_i приведена в первой и последней строках табл. 4.2.

Аналогичная межгрупповая дисперсия δ_Y^2 может быть определена по зависимости групповых средних $\overline{y}_{xi}$ случайной величины Y от относительной частоты их повторений n_{xi} (по данным двух последних столбцов указанной таблицы).

Вычислить упомянутые дисперсии можно по формулам:

$$\delta_X^2 = \frac{1}{n}\sum_{j=1}^{m}(\overline{x}_{yj} - \overline{x})^2 n_{yj}, \quad \delta_Y^2 = \frac{1}{n}\sum_{i=1}^{k}(\overline{y}_{xi} - \overline{y})^2 n_{xi}. \tag{4.22}$$

Корреляционным отношением Y на X (X на Y) называется отношение межгруппового среднего квадратического отклонения случайной величины Y (X) к ее среднему квадратическому отклонению:

$$\eta_{YX} = \frac{\delta_Y}{\sigma_Y} \quad \left(\eta_{XY} = \frac{\delta_X}{\sigma_X}\right). \tag{4.23}$$

Свойства корреляционных отношений

1. *Корреляционные отношения неотрицательны и не превосходят единицы:* $0 \le \eta_{YX} \le 1, \quad 0 \le \eta_{XY} \le 1$.

2. *Равенство нулю корреляционного отношения* η_{YX} *является необходимым и достаточным условием того, что корреляционная связь Y на X отсутствует, а равенство нулю корреляционного отношения* η_{XY} *— необходимым и достаточным условием того, что отсутствует корреляционная связь X на Y.*

3. *Равенство единице корреляционных отношений* η_{YX} *и* η_{XY} *является необходимым и достаточным условием того, что между Y и X существует однозначная корреляционная зависимость.*

4. *Для одной и той же корреляционной таблицы, характеризующей связь случайных величин X и Y, оба корреляционных отношения не меньше абсолютной величины коэффициента корреляции.*

5. *Выполнение равенства* $\eta_{YX} = |r_{YX}| = |r_{XY}| = \eta_{XY}$ *является необходимым и достаточным условием того, что корреляционная зависимость между случайными величинами Y и X точно линейная.*

§ 4.10. ОЦЕНКА КАЧЕСТВА И АНАЛИЗ УРАВНЕНИЙ РЕГРЕССИИ

После построения уравнений регрессии следует дать оценку их качества для того, чтобы убедиться в возможности их использования для анализа описываемых ими реальных процессов.

Оценку качества построенных моделей дают средняя ошибка аппроксимации, коэффициент детерминации и другие характеристики.

Средняя ошибка аппроксимации $\overline{A}$ характеризует среднее отклонение расчетных значений $\tilde{y}_i$ изучаемой величины от ее фактических значений y_i $(i = \overline{1,n})$:

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \tilde{y}_i}{y_i}\right| \cdot 100\%. \qquad (4.24)$$

В практических задачах принимается, что предельное значение средней ошибки аппроксимации $\overline{A}$ не должно превышать 10%.

Долю дисперсии, обусловленную заменой реального процесса моделью регрессии, в общей дисперсии результативного признака характеризует *коэффициент (индекс) корреляции*:

$$R = r^2_{XY} = \frac{\sum\limits_{i=1}^{n}(\tilde{y}_i - \overline{y})^2}{\sum\limits_{i=1}^{n}(y_i - \overline{y})^2}. \qquad (4.25)$$

Оценить качество уравнений регрессии можно и путем проверки гипотезы H_0 о статистической незначимости уравнения регрессии и показателя тесноты связи.

Например, для использования критерия Фишера (Fisher Ronald Ailmer (1890–1962) — английский статистик) необходимо найти расчетное значение критерия, используя соотношения (4.19) или по формуле:

$$F = \frac{\sum\limits_{i=1}^{n}(\tilde{y}_i - \overline{y})^2(n-m-1)}{\sum\limits_{i=1}^{n}(y_i - \tilde{y})^2 m} = \frac{r^2_{XY}}{1 - r^2_{XY}}(n-2), \qquad (4.26)$$

где m — число параметров при переменных X в модели регрессии.

Если в результате расчета по формуле (4.26) величина F оказывается меньше единицы, то в качестве расчетного значения выбирается величина, обратная F.

Расчетное значение F сравнивается с табличным критическим значением F_{kp} критерия Фишера (Приложение 5) при заданном уровне значимости α и данных степенях свободы (числах единиц совокупностей m_1 — для величины, имеющей бóльшую дисперсию; m_2 — для величины с меньшей дисперсией).

Если $F > F_{kp}$, то гипотеза H_0 о случайной природе оцениваемых характеристик отклоняется и признается их статистическая значимость и надежность.

Если $F < F_{kp}$, то гипотеза H_0 не отклоняется и признается их статистическая незначимость, ненадежность уравнения регрессии.

Для расчета статистической значимости коэффициентов регрессии и корреляции рассчитываются t-критерий Стьюдента (Student (1876–1937) — псевдоним английского математика и статистика — William Sealy Gosset) и доверительные интервалы каждого из показателей. Выдвигается гипотеза H_0 о случайной природе показателей, т. е. о незначимом отклонении их от нулевого значения.

Предварительно определяются *случайные ошибки* параметров линейной регрессии:

$$m_{r_{XY}} = \sqrt{\frac{1 - r_{XY}^2}{n-2}}; \qquad m_b = \sqrt{\frac{\sum\limits_{i=1}^{n}(\tilde{y}_i - y_i)^2}{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2 (n-2)}};$$

$$m_a = \sqrt{\frac{\sum\limits_{i=1}^{n}(\tilde{y}_i - y_i)^2}{n-2} \cdot \frac{\sum\limits_{i=1}^{n} x_i^2}{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2}}. \qquad (4.27)$$

Далее находятся значения показателей критерия Стьюдента для каждого из параметров уравнения регрессии:

$$t_r = \frac{r_{XY}}{m_{r_{XY}}} = \sqrt{F}; \qquad t_a = \frac{a}{m_a}; \qquad t_b = \frac{b}{m_b}.$$

Если каждый из показателей t_r, t_a и t_b меньше взятых из таблицы Приложения 3 критических значений T_{kp} параметра Стьюдента, то гипотеза H_0 отклоняется, т. е. параметры r_{XY}, a и b не случайно отличаются от нуля и сформировались под воздействием систематически действующего фактора X. Если же перечисленные показатели больше T_{kp}, то гипотеза H_0 не отклоняется и признается случайной природа формирования r_{XY}, a и b.

После проверки корректности построенной модели регрессии можно приступать к анализу исследуемого процесса, заменяя его моделью регрессии.

§ 4.11. ПОСТРОЕНИЕ КОРРЕЛЯЦИОННЫХ СООТНОШЕНИЙ И УРАВНЕНИЙ РЕГРЕССИИ (ПРИМЕР)

Дано: Результаты испытания двух случайных величин $X = \{x_i\}$, $(i = \overline{1,10})$ и $Y = \{y_j\}$, $(j = \overline{1,6})$ представлены в виде корреляционной таблицы 4.7, в которой индекс i обозначает номер столбца, j — номер строки.

Таблица 4.7

Корреляционная таблица

$j \setminus i$		1	2	3	4	5	6	7	8	9	10
	$y_j \setminus x_i$	2,89	3,13	3,95	3,97	5,32	6,19	6,76	6,81	7,36	8,16
1	1,88	1	3								
2	2,62		4	6							
3	3,73				8	5					
4	3,94					3	8	6			
5	4,29						2	3	5		
6	4.61									4	2

Требуется:

1. Считая корреляционную зависимость линейной, составить уравнения регрессии Y на X и X на Y.

2. Определить коэффициент корреляции и корреляционные отношения.

3. Проанализировать результаты.

Р е ш е н и е:

П.1. Определим величины, необходимые для записи прямой регрессии Y на X.

Для удобства и наглядности определяемые величины представим в виде табл. 4.8. В первый и второй столбцы этой таблицы переписаны первая и вторая строки табл. 4.7.

В третий столбец таблицы записаны суммарные частоты $n_{xi} = \sum\limits_{j=1}^{6} n_{ji}$, с которыми повторяются величины y_j для рассматриваемой величины x_i (сумма частот, стоящих в столбцах табл. 4.7). Например, $n_{x1} = 1; \quad n_{x2} = 3 + 4 = 7; \ldots$.

Четвертый столбец состоит из относительных частот $\overline{n}_{xi} = \dfrac{1}{n}\sum\limits_{j=1}^{6} n_{ji}$, где $n = \sum\limits_{i=1}^{10} n_{xi} = \sum\limits_{j=1}^{6} n_{yj} = \sum\limits_{i=1}^{10}\sum\limits_{j=1}^{6} n_{ji} = 60$.

Средние групповые значения $\overline{y}_{xi} = \dfrac{1}{n_{xi}}\sum\limits_{j=1}^{6} y_j n_{ji}$ внесены в пятый столбец табл. 4.8.

Например, при $i = 2$ (второй столбец табл. 4.7):

$$\overline{y}_{x2} = \frac{1}{7}(y_1 n_{12} + y_2 n_{22} + \ldots + y_6 n_{62}) = \\ = \frac{1}{7}(1{,}88 \cdot 3 + 2{,}62 \cdot 4 + \ldots + 0) = 2{,}30.$$

Обозначения остальных вычисляемых величин приведены в первой строке табл. 4.8. В последней строке таблицы приведены суммарные значения представленных в ее столбцах величин.

Таблица 4.8

Значения, необходимые для составления уравнения регрессии Y на X

i	x_i	n_{xi}	$\overline{n}_{xi}$	$\overline{y}_{xi}$	$\overline{y}_{xi}\overline{n}_{xi}$	$x_i\overline{n}_{xi}$	$x_i\overline{y}_{xi}\overline{n}_{xi}$	$x_i^2\overline{n}_{xi}$	$(\overline{y}_{xi}-\overline{y})^2\overline{n}_{xi}$
1	2,89	1	0,02	1,88	0,04	0,06	0,11	0,17	0,06
2	3,13	7	0,12	2,30	0,28	0,38	0,87	1,19	0,22
3	3,95	6	0,10	2,62	0,26	0,40	1,05	1,58	0,11
4	3,97	8	0,13	3,73	0,48	0,52	1,94	2,06	0,00
5	5,32	8	0,13	3,81	0,50	0,69	2,63	3,67	0,00
6	6,19	10	0,17	4,01	0,68	1,05	4,21	6,50	0,02
7	6,76	9	0,15	4,06	0,61	1,01	4,10	6,83	0,03
8	6,81	5	0,08	4,29	0,34	0,54	2,32	3,68	0,03
9	7,36	4	0,07	4,61	0,32	0,52	2,40	3,83	0,06
10	8,16	2	0,03	4,61	0,14	0,24	1,11	1,96	0,03
	$\sum_{i=1}^{10}$	60	1		3,65 $=\overline{y}$	5,41 $=\overline{x}$	20,74 $=\overline{xy}$	31,47 $=\overline{x^2}$	0,56 $=\delta_Y^2$

По данным таблицы вычисляем коэффициент регрессии Y на X

$$\rho_{Y/X} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{20{,}74 - 5{,}41\cdot 3{,}65}{31{,}47 - 5{,}41^2} = \\ = \frac{0{,}99}{2{,}2} \approx 0{,}45. \quad (4.28)$$

Подставляя полученное значение $\rho_{Y/X}$ и взятые из табл. 4.8 значения $\overline{x}$ и $\overline{y}$ в уравнение прямой регрессии (4.15)

$$y - \overline{y} = \rho_{Y/X}(x - \overline{x}),$$

получим

$$y - 3{,}65 = 0{,}45(x - 5{,}41).$$

После преобразования придем к уравнению прямой с угловым коэффициентом:

$$y = 0{,}45x + 1{,}22.$$

На рис. 4.5 показана эта прямая, помеченная знаком Y/X. Значками ∘ отмечены точки с координатами $(x_i; \overline{y}_{xi})$, расположение которых аппроксимирует описанная прямая.

Получим уравнение прямой регрессии X на Y. Для этого первые два столбца табл. 4.7 перепишем в табл. 4.9. Значения следующих столбцов приведены в первой строке этой таблицы. Они вычисляются аналогично подобным им значениям табл. 4.8.

Отметим, что приведенные в таблице значения $\overline{x}$, $\overline{y}$, и $\overline{xy}$ совпадают с соответствующими значениями табл. 4.8, хотя они подсчитаны по противоположным переменным (групповым вместо единичных и наоборот). Совпадение следовало ожидать, так как средние арифметические значения совокупности величин равны средним арифметическим их групповых значений.

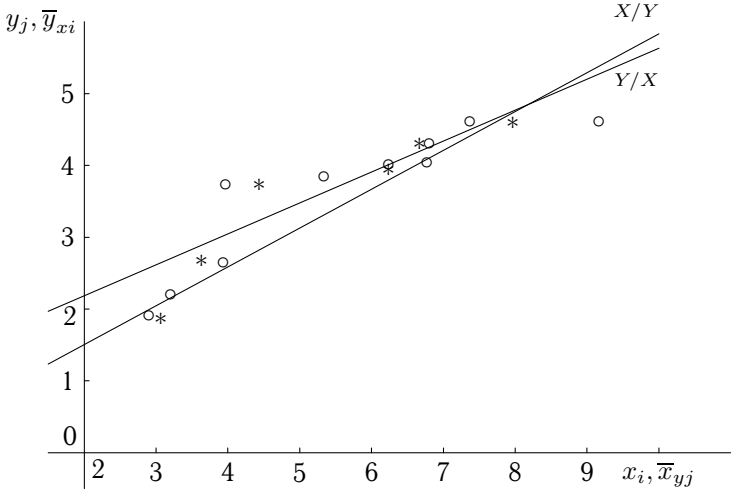


Рис. 4.5
Прямые регрессии примера

Таблица 4.9
Значения, необходимые для составления уравнения регрессии X на Y

j	y_j	n_{yj}	$\overline{n}_{yj}$	$\overline{x}_{yj}$	$\overline{x}_{yj}\overline{n}_{yj}$	$y_j\overline{n}_{yj}$	$\overline{x}_{yj}y_j\overline{n}_{yj}$	$y_j^2\overline{n}_{yj}$	$(\overline{x}_{yj}-\overline{x})^2\overline{n}_{yj}$
1	1,88	4	0,07	3,07	0,21	0,13	0,40	0,24	0,38
2	2,62	10	0,17	3,62	0,62	0,45	1,63	1,18	0,54
3	3,73	13	0,22	4,49	0,99	0,82	3,68	3,06	0,19
4	3,94	17	0,27	6,24	1,68	1,06	6,61	4,18	0,19
5	4,29	10	0,17	6,67	1,13	0,73	4,87	3,13	0,27
6	4,61	6	0,10	7,80	0,78	0,46	3,58	2,12	0,57
	$\sum_{j=1}^{6}$	60	1		5,41 $=\overline{x}$	3,65 $=\overline{y}$	20,74 $=\overline{xy}$	13,91 $=\overline{y^2}$	2,14 $=\delta_X^2$

По данным таблицы вычисляем коэффициент регрессии X на Y:

$$\rho_{X/Y} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\overline{y^2} - \overline{y}^2} = \frac{20{,}74 - 5{,}41\cdot 3{,}65}{13{,}91 - 3{,}65^2} = \frac{0{,}99}{0{,}59} = 1{,}68. \quad (4.29)$$

Подставляя полученное значение $\rho_{X/Y}$ и взятые из табл. 4.8 или табл. 4.9 значения $\overline{x}$ и $\overline{y}$ в уравнение прямой регрессии (4.16)

$$x - \overline{x} = \rho_{X/Y}(y - \overline{y}),$$

получим

$$x - 5{,}41 = 1{,}68(y - 3{,}65).$$

Выразив из этого уравнения y через x, придем к уравнению прямой с угловым коэффициентом:

$$y = 0{,}6x + 0{,}43.$$

Прямая, соответствующая этому уравнению, помечена на рис. 4.5 знаком X/Y. Значками $*$ отмечены точки с координатами $(\overline{x}_{yj}, y_j)$, расположение которых аппроксимирует описанная прямая.

П.3. Определяем коэффициент корреляции случайных величин X и Y (4.18):

$$r_{X/Y} = \pm\sqrt{\rho_{X/Y}\,\rho_{Y/X}} = +\sqrt{1{,}68\cdot 0{,}45} = 0{,}87.$$

Находим корреляционные отношения.

Предварительно, имея в виду, что знаменатели выражений (4.20) представляют собой дисперсии, найдем средние квадратичные отклонения случайных величин X и Y:

$$\sigma_X = \sqrt{\overline{x^2} - \overline{x}^2} = \sqrt{31{,}47 - 5{,}41^2} = \sqrt{2{,}2} = 1{,}48;$$

$$\sigma_Y = \sqrt{\overline{y^2} - \overline{y}^2} = \sqrt{13{,}91 - 3{,}65^2} = \sqrt{0{,}59} = 0{,}77.$$

Межгрупповые дисперсии случайных величин X и Y приведены в нижней строке последних столбцов таблиц 4.8 и 4.9. Имея эти данные, находим

$$\delta_X = \sqrt{\delta_X^2} = \sqrt{2{,}14} = 1{,}46; \quad \delta_Y = \sqrt{\delta_Y^2} = \sqrt{0{,}56} = 0{,}75.$$

По формулам (4.22) находим корреляционные отношения:

$$\eta_{XY} = \frac{\delta_X}{\sigma_X} = \frac{1{,}46}{1{,}48} = 0{,}99; \quad \eta_{YX} = \frac{\delta_Y}{\sigma_Y} = \frac{0{,}75}{0{,}77} = 0{,}97.$$

Анализ полученных коэффициентов корреляции и корреляционных отношений позволяет выдвинуть следующие гипотезы.

3.1. Существует достаточно тесная корреляционная связь между случайными величинами X и Y, так как корреляционные отношения близки к единицам. Особенно близка эта связь для корреляции X на Y ($\eta_{XY} = 0{,}99$). Близость связи можно было заметить по таблицам.

3.2. Неравенства $\eta_{XY} \neq r_{XY}$ и $\eta_{YX} \neq r_{XY}$ говорят о том, что предположение о линейности корреляционных отношений в общем-то не вполне корректно.

§ 4.12. ПЛОСКОСТИ РЕГРЕССИИ

Случайная величина может зависеть статистически не от одной, как это предполагалось в предыдущих разделах главы, а от нескольких случайных величин. Например, урожайность зависит от количества внесенных удобрений, количества выпавших осадков; масса человека — от его роста, объема груди, количества и качества потребляемой пищи и т. п.

Рассмотрим три величины, представленные своими значениями, определенными в каждом из $i = \overline{1,k}$ независимых испытаний. Средние значения одной из этих величин, $U = \{\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_i, \ldots, \overline{u}_k\}$, статистически зависят от двух независимых переменных $X = \{x_1, x_2, \ldots, x_i, \ldots, x_k\}$ и $Y = \{y_1, y_2, \ldots, y_i, \ldots, y_k\}$, имеющих случайный характер. Пусть произведено $n = \sum_{i=1}^{k} n_i$ наблюдений (n_i — количество наблюдений i-го испытания).

Результаты испытаний представлены в виде табл. 4.10.

Таблица 4.10
Результаты испытаний

U	$\overline{u}_1$	$\overline{u}_2$	$\ldots$	$\overline{u}_i$	$\ldots$	$\overline{u}_k$
X	x_1	x_2	$\ldots$	x_i	$\ldots$	x_k
Y	y_1	y_2	$\ldots$	y_i	$\ldots$	y_k
n	n_1	n_2	$\ldots$	n_i	$\ldots$	n_k

Простейшей корреляционной зависимостью U от X и Y является *линейная множественная корреляция*. Она определяется уравнением плоскости регрессии:

$$\tilde{u} = ax + by + c, \tag{4.30}$$

где a, b, c — некоторые постоянные (параметры) модели регрессии.

Задача состоит в том, чтобы найти такие значения параметров a, b и c уравнения (4.30), которые с точки зрения метода наименьших квадратов будет наилучшим образом соответствовать опытным данным.

Подставляя в уравнение (4.30) значения x_i и y_i $(i = \overline{1,k})$, получим значения $\tilde{u}_i$, в общем отличные от значений $\overline{u}_i$, приведенных в табл. 4.10.

Разность $\tilde{u}_i - \overline{u}_i = ax_i + by_i + c - \overline{u}_i$ характеризует отклонение опытных данных от их значений, определенных по функциональной зависимости (4.30).

Как и в случае составления уравнений прямой регрессии одномерных случайных величин, составляется сумма квадратов отклонений с учетом того, сколько раз наблюдалась каждая тройка значений $(x_i; y_i; u_i)$:

$$S = \frac{1}{n}\sum_{i=1}^{k}(ax_i + by_i + c - \overline{u}_i)^2 n_i.$$

Приравнивая нулю производные от S по параметрам a, b и c, можно прийти к системе трех *нормальных уравнений,* позволяющих определить эти параметры:

$$\begin{cases} a\sum\limits_{i=1}^{k} x_i^2 + b\sum\limits_{i=1}^{k} x_i y_i + c\sum\limits_{i=1}^{k} x_i = \sum\limits_{i=1}^{k} \overline{u}_i x_i; \\ a\sum\limits_{i=1}^{k} x_i y_i + b\sum\limits_{i=1}^{k} y_i^2 + c\sum\limits_{i=1}^{k} y_i = \sum\limits_{i=1}^{k} \overline{u}_i y_i; \\ a\sum\limits_{i=1}^{k} x_i + b\sum\limits_{i=1}^{k} y_i + ck = \sum\limits_{i=1}^{k} \overline{u}_i. \end{cases} \tag{4.31}$$

Если из уравнений исключить постоянную c (см. аналогичные преобразования для одномерной задачи в § 4.8), то можно прийти к уравнению линейной регрессии (уравнению плоскости, проходящей через точку с координатами $\overline{x}, \overline{y}, \overline{u}$) для двухмерной задачи:

$$u - \overline{u} = a(x - \overline{x}) + b(y - \overline{y}).$$

Здесь

$$a = \frac{r_{XU} - r_{XY} r_{YU}}{1 - r_{XY}^2} \cdot \frac{\sigma_U}{\sigma_X}; \quad b = \frac{r_{YU} - r_{XY} r_{XU}}{1 - r_{XY}^2} \cdot \frac{\sigma_U}{\sigma_Y};$$

r_{XY}, r_{XU}, r_{YU} — коэффициенты корреляции между парами случайных величин X и Y, X и U, Y и U соответственно:

$$r_{XY} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_X \sigma_Y}; \; r_{XU} = \frac{\overline{xu} - \overline{x} \cdot \overline{u}}{\sigma_X \sigma_U}; \; r_{YU} = \frac{\overline{yu} - \overline{y} \cdot \overline{u}}{\sigma_Y \sigma_U}; \tag{4.32}$$

$\overline{x}, \overline{y}, \overline{u}$ — выборочные средние случайных величин X, Y, U; $\sigma_X, \sigma_Y, \sigma_U$ — выборочные средние квадратичные отклонения этих величин.

Тесноту линейной корреляционной связи между переменными можно установить, определив *совокупный коэффициент корреляции:*

$$R = \sqrt{\frac{r_{XU}^2 + r_{YU}^2 - 2r_{XU}r_{YU}r_{XY}}{1 - r_{XY}^2}}. \tag{4.33}$$

Совокупный коэффициент корреляции определяет меру тесноты линейной корреляционной связи между случайными величинами X, Y и U. Приведем без доказательства его основные свойства.

1. *R изменяется от 0 до 1.*

2. *Если $R = 0$, то случайная величина U не связана со случайными величинами X и Y корреляционной зависимостью.*

3. *Если $R = 1$, то U связана с X и Y линейной корреляционной зависимостью.*

4. *Чем ближе R к нулю, тем менее тесна связь U с X и Y; чем ближе к единице, тем эта связь теснее.*

Наряду с исследованием связи случайной величины U одновременно с X и Y нередко требуется определить тесноту связи U с одной или другой случайными величинами в отдельности. Оценкой такой связи могут служить *частные коэффициенты корреляции*:

$$\begin{aligned} r_{XU(Y)} &= \frac{r_{XU} - r_{XY}r_{YU}}{\sqrt{(1 - r_{XY}^2)(1 - r_{YU}^2)}}, \\ r_{YU(X)} &= \frac{r_{YU} - r_{XY}r_{XU}}{\sqrt{(1 - r_{XY}^2)(1 - r_{XU}^2)}}. \end{aligned} \tag{4.34}$$

Первый из приведенных коэффициентов вычисляется в предположении, что случайная величина Y постоянна, второй — что постоянна X.

Свойства частных коэффициентов корреляции аналогичны свойствам коэффициентов корреляции между двумя случайными величинами. Каждый из них изменяется в пределах от -1 до 1 и обладает перечисленными в § 4.9 свойствами.

§ 4.13. ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ ПАРНОЙ РЕГРЕССИИ

Для семи ($n = 7$) железнодорожных станций ($i = \overline{1,7}$) по результатам сбора и обработки статистических данных получены сведения о расходах электрической и тепловой энергий (y — доля расходов на энергию в % от общих расходов) и о средней дневной заработной плате x одного работающего в некоторых денежных единицах (д. е.). Полученные данные приведены в табл. 4.11.

Таблица 4.11

Исходные данные примера 1

i	1	2	3	4	5	6	7
x	45,1	59,0	57,2	61,8	58,8	47,2	55,2
y	68,8	61,2	59,9	56,7	55,0	54,3	49,3

Т р е б у е т с я.

1. Рассчитать параметры зависимостей y от x:

а. линейной;

б. степенной;

в. показательной;

г. гиперболической.

2. Выбрать среди зависимостей наилучшую.

Р е ш е н и е.

1.а. Для расчета параметров линейной регрессии

$$\tilde{y} = ax + b$$

обратимся к вытекающим из системы нормальных уравнений соотношениям (4.9). Промежуточные и итоговые данные расчетов занесены в табл. 4.12.

$$a = \frac{\overline{yx} - \overline{y} \cdot \overline{x}}{\sigma_x^2} = \frac{3166 - 57{,}9 \cdot 54{,}9}{34{,}34} \approx -0{,}35;$$

$$b = \overline{y} - a\overline{x} = 57{,}9 - (-0{,}35) \cdot 54{,}9 \approx 76{,}9.$$

Таблица 4.12
Данные к линейной модели регрессии

i	y_i	x_i	$y_i x_i$	x_i^2	y_i^2	$\tilde{y}_i$	$y_i - \tilde{y}_i$	A_i
1	68,8	45,1	3103	2034	4733	61,3	7,5	10,9
2	61,2	59,0	3611	3481	3745	56,5	4,7	7,7
3	59,9	57,2	3426	3272	3588	57,1	2,8	4,7
4	56,7	61,8	3504	3819	3215	55,5	1,2	2,1
5	55,0	58,8	3234	3457	3025	56,5	-1,5	2,7
6	54,3	47,2	2563	2228	2948	60,5	-6,2	11,4
7	49,3	55,2	2721	3047	2430	57,8	-8,5	17,2
$\sum$	405,2	384,3	22162	21338	23686	405,2	0	56,7
$\sum/n$	57,9	54,9	3166	3048	3384			8,1
σ^2	32,9	34,34						
σ	5,74	5,86						

Уравнение линейной регрессии приобретает вид

$$\tilde{y} = 76{,}9 - 0{,}35x.$$

Из этого уравнения найдем приращения:

$$\Delta\tilde{y} = \tilde{y}(x + \Delta x) - \tilde{y} = -0{,}35\Delta x.$$

Из этой зависимости следует, что при увеличении средней зарплаты на 1 д.е. доля расходов на электроэнергию снижается на 0,35 процентных единицы.

Подсчитав линейный коэффициент парной корреляции

$$r_{xy} = a\frac{\sigma_x}{\sigma_y} = -0{,}35 \cdot \frac{5{,}86}{5{,}74} \approx -0{,}357,$$

делаем заключение о том, что связь между y и x обратная (знак минус) и умеренная (значение коэффициента меньше 0,5).

Величина коэффициента детерминации (4.25):

$$R = r_{xy}^2 = (-0{,}357)^2 \approx 0{,}127$$

указывает на то, что доля дисперсии, обусловленная введением модели линейной регрессии, в общем показателе дисперсии составляет 12,7%.

Подставляя в уравнение линейной регрессии фактические значения переменной x из третьего столбца табл. 4.12, получим расчетные значения $\tilde{y}$ расходов на энергию (седьмой столбец таблицы). После этого находим абсолютные величины ошибки аппроксимации $y_i - \tilde{y}_i$ (предпоследний столбец) и модули относительных ошибок $A_i = \left|\dfrac{y_i - \tilde{y}_i}{y_i}\right| \cdot 100\%$ (последний столбец).

Средняя ошибка аппроксимации (4.24):

$$\overline{A} = \frac{1}{n}\sum_{i=1} n\overline{A_i} = \frac{1}{7}56{,}7 \cdot 100\% = 8{,}1\%.$$

То есть в среднем расчетные значения отклоняются от фактических на 8,1%.

Определим значение критерия Фишера (4.26):

$$F_p = \frac{r_{xy}^2}{1 - r_{xy}^2}(n-2) = \frac{0{,}127}{0{,}873} \cdot 5 \approx 0{,}7.$$

Так как полученное значение меньше единицы, то в качестве критерия Фишера выбираем обратную величину: $F_p = \dfrac{1}{0{,}7} \approx 1{,}43$.

Это значение меньше значения $F_{kp} = 3{,}8$, найденного по таблице Приложения 5. Поэтому гипотеза H_0 о случайной природе выявленной зависимости и статистической незначимости параметров a и b уравнения линейной регрессии, а также показателя r_{xy} принимается.

1.б. Определению параметров степенной модели $y = bx^a$ предшествует процедура линеаризации уравнения с использованием ее логарифмирования:

$$\lg y = \lg b + a \lg x.$$

Вводя обозначения: $Y = \lg y$; $X = \lg x$; $B = \lg b$, перепишем степенную модель уравнения регрессии в линеаризованном виде:

$$Y = aX + B.$$

Промежуточные вычисления процедуры определения параметров a и B занесены в табл. 4.13.

Таблица 4.13

Данные к степенной модели регрессии

i	Y_i	X_i	Y_iX_i	Y_i^2	X_i^2	$\tilde{y}_i$	$y_i - \tilde{y}_i$	A_i
1	1,84	1,65	3,04	3,38	2,74	61,0	7,8	11,3
2	1,79	1,77	3,16	3,19	3,14	56,3	4,9	8,0
3	1,78	1,76	3,12	3,16	3,09	56,8	3,1	5,2
4	1,75	1,79	3,14	3,08	3,21	55,5	1,2	2,1
5	1,74	1,80	3,08	3,03	3,13	56,3	-1,3	2,4
6	1,73	1,67	2,90	3,01	2,80	60,2	-5,9	10,9
7	1,69	1,74	2,95	2,87	3,03	57,4	-8,1	16,4
$\sum$	12,32	12,16	21,40	21,71	21,14	403,5	1,7	56,3
$\sum/n$	1,76	1,74	3,06	3,10	3,02			8,0
σ^2	0,0018	0,0023						
σ	0,0425	0,0484						

Определив по формулам (4.9) параметры линеаризованной модели регрессии:

$$a = \frac{\overline{XY} - \overline{Y} \cdot \overline{X}}{\sigma_X^2} = \frac{3{,}06 - 1{,}76 \cdot 1{,}74}{0{,}0023} \approx -0{,}3;$$

$$B = \overline{Y} - a\overline{X} = 1{,}76 - (-0{,}3) \cdot 1{,}74 \approx 2{,}28,$$

запишем уравнение регрессии:

$$\tilde{Y} = 2{,}28 - 0{,}3X.$$

После потенцирования представим это уравнение в виде

$$\tilde{y} = 10^{2{,}28} \cdot x^{-0{,}3}.$$

По этой формуле получаем расчетные значения $\tilde{y}$, занесенные в седьмой столбец табл. 4.13.

Рассчитываем коэффициент парной корреляции для нелинейной регрессии (индекс корреляции) (4.21):

$$\rho_{XY} = \sqrt{1 - \frac{\sum\limits_{i=1}^{n}(y-\tilde{y})^2}{\sum\limits_{i=1}^{n}(y-\overline{y})^2}} = \sqrt{1 - \frac{28{,}3}{32{,}9}} \approx 0{,}376,$$

и среднюю ошибку аппроксимации: $\overline{A} = 8{,}0\%$.

Сравнивая полученные характеристики с аналогичными величинами линейной регрессии, убеждаемся в том, что они несколько лучше описывают связь между переменными.

1.в. Рассмотрим модель показательной кривой регрессии: $y = ba^x$.

Линеаризация этой модели повторяет действия со степенной моделью. Ее логарифмирование приводит к линейному уравнению:

$$Y = Ax + B,$$

где $Y = \lg y$; $A = \lg a$; $B = \lg b$.

Определим параметры A и B линеаризованного уравнения регрессии:

$$A = \frac{\overline{Y \cdot x} - \overline{Y} \cdot \overline{x}}{\sigma_x^2} = \frac{96{,}6 - 1{,}76 \cdot 54{,}9}{34{,}34} \approx -0{,}0023;$$

$$B = \overline{Y} - A\overline{x} = 1{,}76 - (-0{,}0023) \cdot 54{,}9 \approx 1{,}887.$$

Линеаризованное уравнение регрессии:

$$\tilde{Y} = 1{,}887 - 0{,}0023x$$

после потенцирования принимает вид:

$$\tilde{y} = 10^{1{,}887} \cdot 10^{-0{,}0023x} \approx 77{,}1 \cdot 0{,}995^x.$$

Таблица 4.14
Данные к показательной модели регрессии

i	Y_i	x_i	$Y_i x_i$	Y_i^2	x_i^2	$\tilde{y}_i$	$y_i - \tilde{y}_i$	A_i
1	1,84	45,1	82,9	3,38	2034	60,7	8,1	11,8
2	1,79	59,0	105,4	3,19	3481	56,4	4,8	7,8
3	1,78	57,2	101,7	3,16	3272	59,9	3,0	5,0
4	1,75	61,8	108,4	3,08	3819	55,5	1,2	2,1
5	1,74	58,8	102,3	3,03	3457	56,4	-1,4	2,5
6	1,73	47,2	81,9	3,01	2228	60,0	-5,7	10,5
7	1,69	55,2	93,4	2,87	3047	57,5	-8,2	16,6
$\sum$	12,32	384,3	676,0	21,71	21338	403,4	-1,8	56,3
$\sum/n$	1,76	54,9	96,6	3,10	3048			8,0
σ^2	0,0018	34,34						
σ	0,0425	5,86						

Тесноту связи оценим индексом корреляции:

$$\rho_{XY} = \sqrt{1 - \frac{\sum\limits_{i=1}^{n}(y-\tilde{y})^2}{\sum\limits_{i=1}^{n}(y-\overline{y})^2}} = \sqrt{1 - \frac{28{,}3}{32{,}9}} \approx 0{,}376.$$

Значение индекса ρ_{xy}, как и ошибка аппроксимации $\overline{A}$ совпадают с принятой точностью с соответствующими значениями степенной модели.

1.г. Уравнение гиперболы

$$y = \frac{a}{x} + b$$

линеаризуется заменой: $u = \dfrac{1}{x}$. Тогда

$$y = au + b.$$

В табл. 4.15 представлены промежуточные данные расчета параметров линеаризованной модели.

Таблица 4.15

Данные к показательной модели регрессии

i	y_i	$10^4 u_i$	$y_i u_i$	y_i^2	$10^6 u_i^2$	$\tilde{y}_i$	$y_i - \tilde{y}_i$	A_i
1	68,8	222	1,526	492	4733	61,8	7,0	10,2
2	61,2	169	1,037	287	3745	56,3	4,9	8,0
3	59,9	175	1,047	306	3588	56,9	3,0	5,0
4	56,7	162	0,916	262	3215	55,5	1,2	2,1
5	55,0	170	0,935	289	3025	56,4	-1,4	2,5
6	54,3	212	1,150	449	2948	60,8	-6,5	12,0
7	49,3	181	0,893	328	2430	57,5	-8,2	16,6
$\sum$	405,2	1291	7,506	2413	23686	405,2	0	56,5
$\sum/n$	57,9	184	1,072	345	3384			8,1
σ^2	32,9	21,5						
σ	5,74	5,86						

Определим параметры a и b линеаризованного уравнения регрессии:

$$a = \frac{\overline{y \cdot u} - \overline{y} \cdot \overline{u}}{\sigma_u^2} = \frac{1{,}072 - 57{,}9 \cdot 10^{-2} \cdot 184}{21{,}5 \cdot 10^4} \approx 1051;$$

$$b = \overline{y} - a\overline{u} = 57{,}9 - 1051 \cdot 184 \cdot 10^4 \approx 38{,}5.$$

Уравнение гиперболы:

$$\tilde{y} = 38{,}5 + 1051 \cdot \frac{1}{x}.$$

Индекс корреляции:

$$\rho_{XY} = \sqrt{1 - \frac{27{,}84}{32{,}92}} \approx 0{,}394.$$

Сравнение значений коэффициентов (индексов) корреляции для всех четырех рассмотренных моделей говорит о том,

что модель гиперболы наилучшим образом описывает заданную зависимость.

Что касается показателей $\overline{A}$, то они для всех моделей примерно равны 8% и не превосходят допустимого для практических задач уровня 10%.

2. Для лучшей модели аппроксимации (гиперболической функции) сделаем оценку по критерию Фишера:

$$F = \left(\frac{\rho_{XY}^2}{1-\rho_{XY}^2}(n-2)\right)^{-1} = \left(\frac{0{,}394^2}{1-0{,}394^2}\cdot 5\right)^{-1} \approx 1{,}09.$$

По таблице Приложения 5 для $m_1 = m_2 = 7$ и $\alpha = 0{,}05$ находим $F_{kp} = 3{,}8 > F$. Следовательно, гипотеза о статистической незначимости параметров уравнения регрессии принимается. Этот результат можно объяснить невысокой теснотой выявленной зависимости и небольшим числом наблюдений.

Линеаризованное уравнение регрессии:

$$\tilde{Y} = 1{,}887 - 0{,}0023x$$

после потенцирования принимает вид:

$$\tilde{y} = 10^{1{,}887}\cdot 10^{-0{,}0023x} \approx 77{,}1\cdot 0{,}995^x.$$

Тесноту связи оценим индексом корреляции:

$$\rho_{XY} = \sqrt{1-\frac{\sum\limits_{i=1}^{n}(y-\tilde{y})^2}{\sum\limits_{i=1}^{n}(y-\overline{y})^2}} = \sqrt{1-\frac{28{,}3}{32{,}9}} \approx 0{,}376.$$

Значение индекса ρ_{XY}, как и ошибка аппроксимации $\overline{A}$ совпадают с принятой точностью с соответствующими значениями степенной модели.

§ 4.14. МНОЖЕСТВЕННАЯ РЕГРЕССИЯ И КОРРЕЛЯЦИЯ

Обобщим понятие множественной регрессии на случай зависимости переменной $Y = \{y_1, y_2, \ldots, y_i, \ldots, y_n\}$ (*результативного признака*) от произвольного числа (m) независимых переменных (*факторов*) X_1, X_2,..., X_j, ..., X_m $(X_j = \{x_{j1}, x_{j2}, \ldots, x_{jn}\})$:

$$Y = F(X_1, X_2, \ldots, X_m). \tag{4.35}$$

В частности, уравнение множественной регрессии может опираться на использование линейных аппроксимирующих функций и функций, приводящихся к линейным.

В случае линейной аппроксимации зависимость (4.35) представляется суммой:

$$Y = a_0 + a_1X_1 + a_2X_2 + \ldots + a_mX_m + \varepsilon. \tag{4.36}$$

Вектор случайных компонент ε отличает статистические данные от данных, получаемых при аппроксимации. То есть $Y = \tilde{Y} + \varepsilon$.

При степенной аппроксимации:

$$Y = a_0X_1^{a_1}X_2^{a_2}\ldots X_m^{a_m} \cdot \varepsilon.$$

При экспоненциальной:

$$Y = \exp\left(a_0 + a_1X_1 + a_2X_2 + \ldots + a_mX_m + \varepsilon\right).$$

Степенная аппроксимирующая функция путем логарифмирования по любому основанию приводится к линейной, а экспоненциальная — путем логарифмирования по основанию e (для приведенного вида записи).

Для определения параметров a_j $(j = \overline{0, m})$ линейных или линеаризованных уравнения множественной регрессии применяют, как правило, метод наименьших квадратов, который для

линейных уравнений сводится к решению системы нормальных уравнений:

$$\begin{cases} \sum\limits_{i=1}^{n} y_i = na_0 + a_1 \sum\limits_{i=1}^{n} x_{i1} + a_2 \sum\limits_{i=1}^{n} x_{i2} + \ldots + a_m \sum\limits_{i=1}^{n} x_{im}; \\ \sum\limits_{i=1}^{n} y_i x_{i1} = a_0 \sum\limits_{i=1}^{n} x_{i1} + a_1 \sum\limits_{i=1}^{n} x_{i1}^2 + \\ \qquad + a_2 \sum\limits_{i=1}^{n} x_{i1}x_{i2} + \ldots + a_m \sum\limits_{i=1}^{n} x_{i1}x_{im}; \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ \sum\limits_{i=1}^{n} y_i x_{im} = a_0 \sum\limits_{i=1}^{n} x_{im} + a_1 \sum\limits_{i=1}^{n} x_{i1}x_{im} + \\ \qquad + a_2 \sum\limits_{i=1}^{n} x_{i1}x_{im} + \ldots + a_m \sum\limits_{i=1}^{n} x_{im}^2. \end{cases} \tag{4.37}$$

Еще один вид линейного уравнения множественной регрессии, отличный от (4.36), — это уравнение регрессии в *стандартизованном масштабе*:

$$\tau_i = \alpha_1 t_{i1} + \alpha_2 t_{i2} + \ldots + \alpha_m t_{im}, \tag{4.38}$$

где $\tau_i = \dfrac{y_i - \overline{Y}}{\sigma_Y}$; $t_{ij} = \dfrac{x_{ij} - \overline{X}_j}{\sigma_{X_j}}$ $(i = \overline{1,n})$ $(j = \overline{1,m})$ — стандартизованные переменные; α_j — стандартизованные коэффициенты регрессии. Эти коэффициенты могут быть определены из системы уравнений:

$$\begin{cases} r_{YX_1} = \alpha_1 + \alpha_2 r_{X_1X_2} + \alpha_3 r_{X_1X_3} + \ldots + \alpha_m r_{X_1X_m}; \\ r_{YX_2} = \alpha_1 r_{X_2X_1} + \alpha_2 + \alpha_3 r_{X_2X_3} + \ldots + \alpha_m r_{X_2X_m}; \\ \ldots \quad \ldots\ldots \quad \ldots\ldots \quad \ldots\ldots \quad \ldots\ldots \quad \ldots\ldots \quad \ldots \\ r_{YX_m} = \alpha_1 r_{X_mX_1} + \alpha_2 r_{X_mX_2} + \alpha_3 r_{X_mX_3} + \ldots + \alpha_m. \end{cases} \tag{4.39}$$

Здесь $r_{YX_i}, r_{X_iX_k}$ — коэффициенты корреляции между парами случайных величин (4.32).

В частности, для двухфакторной модели из системы уравнений (4.39) можно выразить искомые коэффициенты в явном виде:

$$\alpha_1 = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{1 - r^2_{X_1X_2}}; \quad \alpha_2 = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{1 - r^2_{X_1X_2}}. \qquad (4.40)$$

Связь между коэффициентами множественной регрессии a_j и стандартизованными коэффициентами α_j $(j = \overline{1,m})$ описывается соотношением:

$$a_j = \alpha_j \frac{\sigma_Y}{\sigma_{X_j}}. \qquad (4.41)$$

После определения коэффициентов $a_1 - a_m$ можно найти:

$$a_0 = \overline{Y} - \sum_{j=1}^{m} a_j \overline{X}_j. \qquad (4.42)$$

Частные коэффициенты корреляции, оценивающие влияние фактора X_i на результативный признак Y при неизменном уровне других факторов, можно определить по рекуррентной формуле:

$$r_{YX_i(X_1 \ldots X_{i-1} X_{i+1} \ldots X_m)} = \\ = \frac{r_{YX_i(X_1 \ldots X_{m-1})} - r_{YX_m(X_1 \ldots X_{m-1})} \cdot r_{X_i X_m (X_1 \ldots X_{m-1})}}{\sqrt{\left(1 - r^2_{YX_m(X_1 \ldots X_{m-1})}\right)\left(1 - r^2_{X_i X_m (X_1 \ldots X_{m-1})}\right)}}. \qquad (4.43)$$

Частные коэффициенты корреляции изменяются в пределах от -1 до 1.

Теснота совместного влияния факторов на результат оценивается *индексом множественной корреляции:*

$$R_{YX_1X_2 \ldots X_m} = \sqrt{\sum_{j=1}^{m} \alpha_j r_{YX_i(X_1 \ldots X_{i-1} X_{i+1} \ldots X_m)}}. \qquad (4.44)$$

Значение индекса $R_{YX_1X_2 \ldots X_m}$ изменяется в пределах от 0 до 1 и должно быть не меньше максимального коэффициента

корреляции между парами случайных величин:

$$R_{YX_1X_2\ldots X_m} \geq r_{YX_j}.$$

Наряду с корреляционными характеристиками, оценивающими влияние отдельных факторов на зависимую переменную, часто для этой цели используются статистические критерии Фишера, Стьюдента и другие. Например, *частный критерий Фишера (F-критерий)*:

$$F_{X_i} = \frac{R^2_{YX_1\ldots X_{i-1}X_iX_{i+1}\ldots X_m} - R^2_{YX_1\ldots X_{i-1}X_{i+1}\ldots X_m}}{1 - R^2_{YX_1\ldots X_{i-1}X_1X_{i+1}\ldots X_m}} \cdot \frac{n-m-1}{m}. \tag{4.45}$$

§ 4.15. РЕАЛИЗАЦИЯ МОДЕЛЕЙ МНОЖЕСТВЕННОЙ РЕГРЕССИИ

Методы корреляционно-регрессионного анализа, на базе которых строятся эконометрические модели, позволяют решать три основные задачи:

— определить связь между переменными $U = \{u_1, u_2, \ldots, u_m\}$ и совокупностью переменных $X = \{x_1, x_2, \ldots, x_k\}$, $Y = \{y_1, y_2, \ldots, y_m\}$ и т. д. (для парной регрессии между переменной U и одной переменной X);

— установление тесноты связи между U с одной стороны и X, Y и т. д. с другой стороны;

— анализ влияния отдельных факторных признаков (для многофакторной модели).

Рассмотрим, как реализуется решение эконометрических задач на конкретных примерах.

Пример 1. Имеются статистические данные о зависимости прибыли фирмы $U = \{u_i\}$ $(i = \overline{1,9})$ от затрат на производство продукции $X = \{x_i\}$ и от количества работающих на фирме $Y = \{y_i\}$ (табл. 4.16).

Табличные данные представляют собой двухфакторную статистическую модель. Будем аппроксимировать заданную зависимость линейной функцией:

$$\tilde{u} = ax + by + c.$$

Таблица 4.16
Исходные данные Примера 1

i	1	2	3	4	5	6	7	8	9
u_i	63	158	266	370	480	590	728	936	999
x_i	230	308	450	557	652	744	822	940	967
y_i	15	21	27	32	34	36	37	40	37

Параметры a, b и c модели определим методом наименьших квадратов из системы нормальных уравнений (4.31):

$$\begin{cases} 4133 \cdot 10^3 a + 193 \cdot 10^3 b + 5{,}668 \cdot 10^3 c = 3585 \cdot 10^3; \\ 193 \cdot 10^3 a + 9{,}21 \cdot 10^3 b + 279c = 162 \cdot 10^3; \\ 5868a + 279b + 9c = 4589. \end{cases} \tag{4.46}$$

Для определения коэффициентов системы представим результаты промежуточных вычислений в виде табл. 4.17 и 4.18. В таблицах все величины после их вычисления округлены до трех значащих цифр.

Решая систему, найдем:

$$a \approx 1{,}83; \qquad b \approx -19{,}5; \qquad c \approx -42{,}4,$$

что позволяет записать аппроксимирующую функцию в виде

$$\tilde{u} = 1{,}83x - 19{,}5y - 42{,}4.$$

Для определения тесноты связи между переменными определим предварительно средние квадратичные отклонения заданных величин:

$$\sigma_U = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(u_i - \overline{u})^2} = \sqrt{\frac{1}{9}878 \cdot 10^3} \approx 312;$$

$$\sigma_X = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(x_i - \overline{x})^2} = \sqrt{\frac{1}{9}561 \cdot 10^3} \approx 250;$$

Таблица 4.17

Вычисляемые параметры

i	u_i	x_i	y_i	$10^{-3}x_i^2$	$10^{-1}y_i^2$	$10^{-2}x_iy_i$	$10^{-3}u_ix_i$	$10^{-2}u_iy_i$
1	63	230	15	53	23	35	14	9
2	158	308	21	95	44	65	49	33
3	266	450	27	203	73	122	120	72
4	370	557	32	310	102	178	206	118
5	480	652	34	425	116	222	313	163
6	590	744	36	553	130	268	439	212
7	728	822	37	676	137	304	598	269
8	936	940	40	884	160	376	880	374
9	999	967	37	935	137	358	966	370
$\sum$	4589	5868	279	4134	921	1926	3585	1622
$\sum/9$	510	630	31			214	398	180

Таблица 4.18

Вычисляемые параметры (продолжение)

i	$u_i-\overline{u}$	$x_i-\overline{x}$	$y_i-\overline{y}$	$10^{-3}(u_i-\overline{u})^2$	$10^{-3}(x_i-\overline{x})^2$	$(y_i-\overline{y})^2$
1	-447	-400	-16	200	160	256
2	-352	-322	-10	124	104	100
3	-244	-180	-4	60	32	16
4	-140	-73	1	20	5	1
5	-30	22	3	1	0	9
6	80	114	5	6	13	25
7	218	192	6	47	37	36
8	426	310	8	181	96	81
9	489	337	6	239	114	36
$\sum$	0	0	0	878	561	560
$\sqrt{\sum/9}$				312	250	7,9

$$\sigma_Y = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(y_i - \overline{y})^2} = \sqrt{\frac{1}{9}560} \approx 7{,}9$$

и затем парные коэффициенты корреляции (4.32):

$$r_{XY} = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\sigma_X \sigma_Y} = \frac{214\cdot 10^2 - 630\cdot 31}{250\cdot 7{,}9} \approx 0{,}95;$$

$$r_{XU} = \frac{\overline{xu} - \overline{x}\cdot\overline{u}}{\sigma_X \sigma_U} = \frac{398\cdot 10^3 - 510\cdot 630}{312\cdot 250} \approx 0{,}99;$$

$$r_{YU} = \frac{\overline{yu} - \overline{y}\cdot\overline{u}}{\sigma_Y \sigma_U} = \frac{180\cdot 10^2 - 510\cdot 31}{312\cdot 7{,}9} \approx 0{,}90.$$

По формуле (4.33) определим совокупный коэффициент корреляции:

$$R = \sqrt{\frac{r_{XU}^2 + r_{YU}^2 - 2r_{XU}r_{YU}r_{XY}}{1 - r_{XY}^2}} = \\ = \sqrt{\frac{0{,}99^2 + 0{,}9^2 - 2\cdot 0{,}99\cdot 0{,}9\cdot 0{,}95}{1 - 0{,}95^2}} \approx 0{,}998.$$

Расчетное значение R очень близко к единице, больше любого парного коэффициента корреляции. Это говорит о том, что взаимная связь исследуемых показателей процесса является очень тесной, практически не стохастической, а функциональной. Совместное влияние расходов и количества исполнителей на 99,8% объясняет изменения прибыли этих фирм.

Теснота связи между результативным (эндогенным) признаком и одним из так называемых факторных (экзогенных) признаков устанавливается в многофакторных моделях частными коэффициентами корреляции (4.34). Определим эти коэффициенты:

$$r_{XU(Y)} = \frac{r_{XU} - r_{XY}r_{YU}}{\sqrt{(1 - r_{XY}^2)(1 - r_{YU}^2)}} = \\ = \frac{0{,}99 - 0{,}95\cdot 0{,}9}{\sqrt{(1 - 0{,}95^2)(1 - 0{,}9^2)}} \approx 0{,}993;$$

$$r_{YU(X)} = \frac{r_{YU} - r_{XY} r_{XU}}{\sqrt{(1 - r_{XY}^2)(1 - r_{XU}^2)}} = \\ = \frac{0{,}9 - 0{,}95 \cdot 0{,}99}{\sqrt{(1 - 0{,}95^2)(1 - 0{,}99^2)}} \approx -0{,}92.$$

Значения обоих коэффициентов по абсолютной величине достаточно близки к единице и поэтому связь между прибылью и количеством работающих без учета расходов, а также между прибылью и расходами без учета количества работающих можно считать достаточно тесной.

На практике для всех оценочных коэффициентов влияния одних показателей на другие связь считается тесной, если соответствующие коэффициенты корреляции превышают 0,8.

Пример 2. Имеются данные обработки статистического материала, представляющего собой зависимость душевого дохода y от заработной платы x_1 и возраста работающих x_2.

В табл. 4.19 представлены результаты обработки данных, полученных по $n = 30$ наблюдениям: $\overline{(\cdots)}$ — средние арифметические; σ_i — средние квадратичные отклонения и r_{yx_i} — линейные коэффициенты парной корреляции.

Таблица 4.19

Исходные данные Примера 2

	$\overline{(\cdots)}$	σ_i	r_{yx_i}
y	868	114	
x_1	549	59	0,84
x_2	33,	0,58	-0,21

$r_{x_1x_2} = -0{,}116$

Требуется:

1. Построить уравнение множественной регрессии в стандартизованной и естественной формах. Сравнить коэффициенты уравнения и сделать соответствующие выводы.

2. Рассчитать коэффициенты множественной и частной корреляции, сравнить их с коэффициентами парной кор-

реляции, пояснить смысл различия между значениями этих коэффициентов.

3. Рассчитать частные значения критерия Фишера и с его помощью проанализировать исходные данные.

Р е ш е н и е.

1. Используя соотношения (4.40), определим коэффициенты стандартизованного уравнения двухфакторной регрессии:

$$\alpha_1 = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{1 - r^2_{X_1X_2}} = \frac{0{,}84 - (-0{,}21) \cdot (-0{,}116)}{1 - (-0{,}116)^2} \approx 0{,}827;$$

$$\alpha_2 = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{1 - r^2_{X_1X_2}} = \frac{-0{,}21 - (0{,}84) \cdot (-0{,}116)}{1 - (-0{,}116)^2} \approx -0{,}114.$$

Найденные коэффициенты позволяют записать стандартизованной уравнение регрессии (4.38) в виде

$$\tau = 0{,}827t_1 - 0{,}114t_2.$$

Перейдем, следуя формулам (4.41), от коэффициентов α_j к a_j:

$$a_1 = \alpha_1 \frac{\sigma_Y}{\sigma_{X_1}} = 0{,}827 \frac{114}{59} \approx 1{,}61.$$

$$a_2 = \alpha_2 \frac{\sigma_Y}{\sigma_{X_2}} = -0{,}114 \frac{114}{0{,}58} \approx -22{,}5.$$

Значение коэффициента a_0 найдем из соотношения (4.42):

$$a_0 = \overline{Y} - a_1\overline{X}_1 - a_2\overline{X}_2 = \\ = 868 - 1{,}61 \cdot 549 - (-22{,}5 \cdot 33{,}5) \approx -735.$$

Найденные коэффициенты a_0, a_1 и a_2 позволяют записать уравнение линейной двухфакторной регрессии (4.36):

$$y = -735 + 1{,}61x_1 - 22{,}5x_2.$$

Сравниваем по модулю коэффициенты α_1 и α_2:

$$|\alpha_1| = 0{,}827 > |\alpha_2| = 0{,}114.$$

Неравенство позволяет сделать вывод о том, что сила влияния заработной платы, характеризуемой коэффициентом α_1, на доход гораздо сильнее влияния среднего возраста, характеризуемого коэффициентом α_2.

2. Воспользовавшись формулой (4.43), найдем линейные коэффициенты частной корреляции:

$$r_{YX_1(X_2)} = \frac{r_{YX_1} - r_{YX_2} \cdot r_{X_1X_2}}{\sqrt{\left(1 - r^2_{YX_2}\right)\left(1 - r^2_{X_1X_2}\right)}} = \\ = \frac{0{,}84 - (-0{,}21) \cdot (-0{,}116)}{\sqrt{\left(1 - (-0{,}21)^2\right)\left(1 - (-0{,}116)^2\right)}} \approx 0{,}84.$$

$$r_{YX_2(X_1)} = \frac{r_{YX_2} - r_{YX_1} \cdot r_{X_1X_2}}{\sqrt{\left(1 - r^2_{YX_1}\right)\left(1 - r^2_{X_1X_2}\right)}} = \\ = \frac{-0{,}21 - 0{,}84 \cdot (-0{,}116)}{\sqrt{\left(1 - (0{,}84)^2\right)\left(1 - (-0{,}116)^2\right)}} \approx -0{,}21.$$

$$r_{X_1X_2(Y)} = \frac{r_{X_1X_2} - r_{YX_1} \cdot r_{YX_2}}{\sqrt{\left(1 - r^2_{YX_1}\right)\left(1 - r^2_{YX_2}\right)}} = \\ = \frac{-0{,}116 - 0{,}84 \cdot (-0{,}21)}{\sqrt{\left(1 - (0{,}84)^2\right)\left(1 - (-0{,}21)^2\right)}} \approx 0{,}114.$$

В пределах заданной точности вычислений найденные значения коэффициентов не отличаются от заданных в табл. 4.16 коэффициентов парной корреляции. Это говорит о том, что выводы о тесноте и направлении связи на основе коэффициентов парной регрессии и частной корреляции совпадают.

Расчет линейного коэффициента множественной корреляции выполним по формуле (4.44):

$$R_{YX_1X_2} = \sqrt{\alpha_1 r_{YX_1} + \alpha_2 r_{YX_2}} = \\ = \sqrt{0{,}827 \cdot 0{,}84 + (-0{,}114) \cdot (-0{,}21)} \approx \\ \approx \sqrt{0{,}72} \approx 0{,}85.$$

Значение $R_{YX_1X_2} \approx 0{,}85 > 0{,}8$ говорит о тесной зависимости Y от совместного действия факторов X_1 и X_2.

Цифра, стоящая под корнем последнего выражения $R^2_{YX_1X_2} \approx 0{,}72$, говорит о том, что 72% душевого дохода определяется заработной платой и возрастом. Прочие факторы, не включенные в модель, составляют 28% доходов.

3. Критерии Фишера для факторов F_{X_1} и F_{X_2} оценивают статистическую значимость присутствия факторов X_1 и X_2 в уравнении множественной регрессии и оценивают целесообразность включения в уравнение одного фактора после другого. То есть F_{X_1} оценивает целесообразность включения в уравнение фактора X_1 после того, как в нее включен фактор X_2. И наоборот, F_{X_2} указывает на целесообразность включения в модель фактора X_2 после того, как в нее включен фактор X_1.

Определим критерий Фишера по формуле (4.45):

$$F_{X_1} = \frac{R^2_{YX_1X_2} - r^2_{YX_2}}{1} - R^2_{YX_1X_2} \cdot \frac{n-m-1}{1} = \\ = \frac{0{,}85^2 - (-0{,}21)^2}{1 - 0{,}85^2} \cdot \frac{30-2-1}{1} \approx 65.$$

Для уровня значимости $\alpha = 0{,}05$ табличное значение критерия Фишера $F_{kp} = 4{,}21 < F_{X_1} = 65$.

Тот факт, что табличное значение критерия Фишера оказалось меньше расчетного, говорит о целесообразности включения в модель фактора X_1 после включения в нее фактора X_2. Гипотезу H_0 о несущественности прироста R^2_y за счет включения дополнительного фактора X_1 отклоняем.

Целесообразность включения в модель фактора X_2 после включения в нее фактора X_1 проверяет критерий Фишера для фактора X_2 :

$$F_{X_2} = \frac{R^2_{YX_1X_2} - r^2_{YX_1}}{1} - R^2_{YX_1X_2} \cdot \frac{n-m-1}{1} = \\ = \frac{0{,}85^2 - (-0{,}84)^2}{1 - 0{,}85^2} \cdot \frac{30-2-1}{1} \approx 0{,}06.$$

Столь малое расчетное значение критерия Фишера (намного меньше $F_{kp} = 4{,}21$) говорит о нецелесообразности включения в модель фактора X_2 после того, как в нее включен фактор X_1.

§ 4.16. СИСТЕМЫ ВЗАИМОСВЯЗАННЫХ УРАВНЕНИЙ

Сложные многофакторные процессы могут описываться *системами взаимосвязанных уравнений*. Среди таких систем выделяют несколько характерных.

• *Система независимых уравнений*. Характерной особенностью этой системы является то, что каждая зависимая переменная из совокупности $Y = \{y_1, y_2, \ldots, y_n\}$ рассматривается как функция одной и той же совокупности независимых переменных (факторов) $X = \{x_1, x_2, \ldots, x_n\}$:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n; \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n; \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n. \end{cases} \tag{4.47}$$

• *Система рекурсивных уравнений* — зависимая переменная Y одного уравнения выступает как фактор в других уравнениях:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n; \\ y_2 = b_{21}y_1 + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_; \\ y_3 = b_{31}y_1 + b_{32}y_2 + a_{31}x_1 + \ldots + a_{3n}x_n; \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ y_m = b_{m1}y_1 + \ldots + b_{m,m-1}y_{m-1} + \\ \qquad\qquad + a_{m1}x_1 + \ldots + a_{mn}x_n. \end{cases} \tag{4.48}$$

• *Система взаимосвязанных (совместных) уравнений* или *структурная форма модели* — одни и те же зависимые переменные входят в левую часть одних уравнений и в правую часть — других:

$$\begin{cases} y_1 = b_{12}y_2 + b_{13}y_3 + \ldots + b_{1m}y_m + \\ \qquad\qquad + a_{11}x_1 + \ldots + a_{1n}x_n; \\ y_2 = b_{21}y_1 + b_{23}y_3 + \ldots + b_{2m}y_m + a_{21}x_1 + \\ \qquad\qquad + a_{22}x_2 + \ldots + a_{2n}x_; \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ y_m = b_{m1}y_1 + b_{m2}y_2 + \ldots + b_{m,m-1}y_{m-1} + \\ \qquad\qquad + a_{m1}x_1 + \ldots + a_{mn}x_n. \end{cases} \tag{4.49}$$

В этих системах уравнений (4.47)–(4.49) кроме эндогенных (зависимых) переменных Y, определяемых внутри системы, и экзогенных (независимых) переменных X, определяемых вне системы или задаваемых, различают *предопределенные переменные* — экзогенные и *лаговые эндогенные* (включенные в систему за время, предшествующее текущему моменту).

Для решения систем уравнений (4.48)–(4.49) они предварительно приводятся к виду (*приведенная форма уравнений*):

$$\begin{cases} \hat{y}_1 = c_{11}x_1 + c_{12}x_2 + \ldots + c_{1n}x_n; \\ \hat{y}_2 = c_{21}x_1 + c_{22}x_2 + \ldots + c_{2n}x_n; \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ \hat{y}_m = c_{m1}x_1 + c_{m2}x_2 + \ldots + c_{mn}x_n. \end{cases} \tag{4.50}$$

Перед решением систем уравнений необходимо произвести их *идентификацию* — процедуру, связанную с оценкой параметров a_{ij}, b_{ij} и c_{ij} с целью выяснения их достаточности для описания рассматриваемого процесса.

Введем обозначения:

M — число эндогенных переменных в уравнении;

N — число предопределенных переменных, отсутствующих в рассматриваемом уравнении, но присутствующих в системе.

Необходимое условие идентификации:

$N + 1 = M$ — уравнение идентифицируемо;

$N + 1 < M$ — уравнение неидентифицируемо;

$N + 1 > M$ — уравнение сверхидентифицируемо.

Достаточное условие идентификации: определитель матрицы, составленной из коэффициентов при переменных, отсутствующих в исследуемом уравнении, не равен нулю и ранг матрицы не меньше числа эндогенных переменных системы без единицы.

Для решения систем уравнений (4.47)–(4.50) обычно используется метод наименьших квадратов и различные его модификации. Так, для решения идентифицируемого уравнения применяют *косвенный метод наименьших квадратов*. Согласно этому методу:

— составляется приведенная форма математической модели процесса (4.50) и для каждого из уравнений системы обычным методом наименьших квадратов определяются значения требуемых параметров;
— переходят от приведенной формы модели (4.50) к структурной форме (4.49), получая тем самым численные оценки структурных параметров a_{ij} и b_{ij}.

Для решения сверхидентифицируемых систем уравнений применяют *двухшаговый метод наименьших квадратов.* Последовательность выполняемых в этом методе действий:
— составляется приведенная форма модели и обычным методом наименьших квадратов определяются значения параметров, входящих в систему уравнений;
— выделяются эндогенные переменные в правых частях структурных уравнений (4.49), параметры этих уравнений определяют двухшаговым методом наименьших квадратов, затем по соответствующим уравнениям приведенной формы (4.50) находят расчетные значения параметров;
— обычным методом наименьших квадратов определяются параметры структурных уравнений, при этом в качестве исходных данных используются значения предопределенных переменных и расчетные значения эндогенных переменных, стоящих в правых частях структурных уравнений.

§ 4.17. ПРИМЕР РЕШЕНИЯ СИСТЕМ УРАВНЕНИЙ

1. Оценить на идентификацию структурную форму модели (СФ) (4.49):

$$\begin{cases} y_1 = b_{13}y_3 + a_{11}x_1 + a_{13}x_3; \\ y_2 = b_{21}y_1 + b_{23}y_3 + a_{22}x_2 + \ldots + a_{2n}x; \\ y_3 = b_{32}y_2 + a_{31}x_1 + a_{33}x_3. \end{cases}$$

2. Исходя из приведенной формы уравнений (ПФ) (4.50):

$$\begin{cases} y_1 = 2x_1 + 4x_2 + 10x_3; \\ y_2 = 3x_1 - 6x_2 + 2x_3; \\ y_3 = -5x_1 + 8x_2 + 5x_3, \end{cases} \tag{4.51}$$

найти структурные коэффициенты модели.

Р е ш е н и е. **1.** Модель включает в себя три эндогенные $Y = \{y_1; y_2; y_3\}$ и три экзогенные $X = \{x_1; x_2; x_3\}$ переменные.

Проверим каждое уравнение на необходимое (НУ) и достаточное (ДУ) условия идентификации.

Первое уравнение. НУ: уравнение содержит $M = 2$ эндогенных переменных (y_1, y_3); в нем отсутствует $N = 1$ экзогенная переменная (x_2).

Так как $N + 1 = M$ (1+1=2), т. е. равенство строго выполнено, то уравнение точно идентифицируемо.

ДУ: в первом уравнении отсутствуют y_2 и x_2. Построим матрицу коэффициентов, стоящих при этих переменных в двух оставшихся уравнениях системы:

Уравнение	y_2	x_2
второе	-1	a_{22}
третье	b_{32}	0

Определитель матрицы выписанных коэффициентов ($a_{22} \neq \neq 0$, $b_{32} \neq 0$):

$$\Delta = \begin{vmatrix} -1 & a_{22} \\ b_{32} & 0 \end{vmatrix} = -1 \cdot 0 - a_{22} \cdot b_{32} \neq 0.$$

Определитель матрицы не равен нулю, ранг матрицы равен 2. Следовательно, выполняется достаточное условие и первое уравнение точно идентифицируемо.

Второе уравнение. НУ: уравнение содержит $M = 3$ эндогенных переменных (y_1, y_2, y_3); в нем отсутствуют $N = 2$ экзогенных переменных (x_1, x_3).

Так как $N + 1 = M$ (2+1=3), т. е. равенство строго выполнено, то уравнение точно идентифицируемо.

ДУ: во втором уравнении отсутствуют x_1 и x_3. Построим матрицу коэффициентов, стоящих при этих переменных в двух оставшихся уравнениях системы:

Уравнение	x_1	x_3
первое	a_{11}	a_{13}
третье	a_{31}	a_{33}

Определитель матрицы выписанных коэффициентов:

$$\Delta = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11} \cdot a_{33} - a_{13} \cdot b_{31} \neq 0.$$

Определитель матрицы не равен нулю, ранг матрицы равен 2. Следовательно, выполняется достаточное условие и первое уравнение точно идентифицировано.

Третье уравнение. НУ: уравнение содержит $M = 2$ эндогенных переменных (y_2, y_3); в нем отсутствует $N = 1$ экзогенная переменная (x_2).

Так как $N + 1 = M$ (1+1=2), т. е. равенство строго выполнено, то уравнение точно идентифицируемо.

ДУ: в третьем уравнении отсутствуют y_1 и x_2. Построим матрицу коэффициентов, стоящих при этих переменных в двух оставшихся уравнениях системы:

Уравнение	y_1	x_2
первое	-1	0
третье	b_{21}	a_{22}

Определитель матрицы выписанных коэффициентов:

$$\Delta = \begin{vmatrix} -1 & 0 \\ b_{21} & a_{22} \end{vmatrix} = -1 \cdot a_{22} - 0 \cdot b_{21} \neq 0.$$

Определитель матрицы не равен нулю, ранг матрицы равен 2. Следовательно, выполняется достаточное условие и третье уравнение точно идентифицировано.

Для всех трех уравнений условия идентифицируемости выполнены точно, следовательно, система уравнения точно идентифицируема и для ее решения можно применить косвенный метод наименьших квадратов.

2. Вычислим структурные коэффициенты модели.

Из третьего уравнения приведенной формы (система уравнений (4.51)) выразим переменную x_2, так как эта переменная не содержится в первом уравнении (4.51) СФ:

$$x_2 = \frac{1}{8}(y_3 + 5x_1 - 5x_3). \qquad (4.52)$$

Выражение содержит переменные y_3, x_1 и x_3, которые присутствуют в первом уравнении СФ (4.51). Подставим полученное для x_2 выражение в первое уравнение (4.51) ПФ:

$$y_1 = 0{,}5y_3 + 4{,}5x_1 + 7{,}5x_3. \tag{4.53}$$

Во втором уравнении (4.51) отсутствуют переменные x_1 и x_3. Структурные параметры этого уравнения определим в два этапа.

Первый этап. Выразим x_1, например, из первого уравнения (4.51):

$$x_1 = 0{,}5y_1 - 2x_2 - 5x_3.$$

Подстановка полученного выражения во второе уравнение (4.51) ПФ не решит задачу до конца, так как в x_1 присутствует x_3, которого нет в СФ. Поэтому выразим x_3 из третьего уравнения (4.51):

$$x_3 = \frac{1}{5}(y_3 + 5x_1 - 8x_2) \tag{4.54}$$

и подставим в выражение для x_1. После преобразования получим:

$$x_1 = \frac{1}{12}(y_1 - 2y_3 + 12x_2). \tag{4.55}$$

Второй этап. Чтобы выразить x_3 через искомые y_1, y_3 и x_2, заменим в (4.54) значение x_1 на полученное для него из первого уравнения (4.51) ПФ выражение. Тогда

$$x_3 = 0{,}2y_3 + 0{,}5y_1 - 3{,}6x_2 - 5x_3.$$

Отсюда найдем:

$$x_3 = 0{,}033y_3 + 0{,}083y_1 - 0{,}6x_2. \tag{4.56}$$

Подставим полученные для x_1 и x_3 выражения во второе уравнение ПФ (4.51). После преобразования получим:

$$y_2 = 0{,}416y_1 - 0{,}434y_3 - 4{,}2x_2. \tag{4.57}$$

Это второе уравнение СФ.

Из второго уравнения (4.51) ПФ выразим x_2, так как этой переменной нет в третьем уравнении СФ:

$$x_2 = -0{,}167y_2 + 0{,}5x_1 + 0{,}333x_3$$

и подставим полученное выражение в третье уравнение (4.51) ПФ. Тем самым получим третье уравнение СФ:

$$y_3 = -1{,}336 - x_1 + 7{,}664x_3. \tag{4.58}$$

Таким образом, три уравнения (4.53), (4.57) и (4.58) образуют систему уравнений СФ:

$$\begin{cases} y_1 = 0{,}5y_3 + 4{,}5x_1 + 7{,}5x_3; \\ y_2 = 0{,}416y_1 - 0{,}434y_3 - 4{,}2x_2; \\ y_3 = -1{,}336y_2 - x_1 + 7{,}664x_3. \end{cases}$$

§ 4.18. РЕЗЮМЕ

Уравнение регрессии, связывающее две случайные величины, является аналитической моделью, позволяющей в допустимых пределах разброса случайных величин проанализировать их поведение. В простейшей модели линейной регрессии присутствуют два параметра (уравнение прямой на плоскости):

$$y = ax + b.$$

Параметры a и b можно определить из условия наименьшего отклонения значений случайной величины от прямой регрессии. В частности, метод наименьших квадратов приводит к формулам для определения параметров прямой регрессии:

$$a = \frac{\sum\limits_{i=1}^{k}(x_i - \overline{x})(y_i - \overline{y})}{\sum\limits_{i=1}^{k}(x_i - \overline{x})^2};$$

$$b = \overline{y} - a\overline{x}.$$

Уравнение прямой регрессии после введения коэффициента регрессии Y на X:

$$\rho_{Y/X} = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \overline{x}^2} = \frac{\overline{\sigma}_{XY}}{\overline{\sigma}_Y^2}$$

приводится к виду

$$x - \overline{x} = \rho_{X/Y}(y - \overline{y}).$$

Два коэффициента регрессии Y на X и X на Y позволяют определить коэффициент корреляции:

$$r_{XY} = \pm\sqrt{\rho_{Y/X}\,\rho_{X/Y}}.$$

Коэффициент корреляции изменяется от -1 до 1. Если $|r_{XY}| = 1$, то между случайными величинами X и Y имеется строго линейная корреляционная зависимость. При $r_{XY} = 0$ корреляционной зависимости не существует.

При установлении зависимости между несколькими случайными величинами уравнения прямой регрессии становится уравнениями плоскости (гиперплоскости) регрессии. При этом выражения для корреляционных коэффициентов, выражающих тесноту связи между случайными величинами, намного усложняются (§ 4.14).

Если соотношения нелинейны, то для построения уравнений регрессии используются функции: квадратичные, логарифмические, экспоненциальные ..., в том числе и в пространствах различных размерностей. Сложные многофакторные случайные процессы описываются системами уравнений различной сложности.

Ошибки, возникающие при замене связи между случайными величинами на уравнения регрессии, могут быть оценены с привлечением различных критериев (Стьюдента, Фишера, ...).

§ 4.19. ВОПРОСЫ

1. Чем отличается функциональная зависимость от корреляционной?
2. Какая информация заносится в корреляционные таблицы?

3. Что такое корреляционная зависимость Y на X? X на Y? Чем отличаются эти зависимости?
4. Что такое уравнение регрессии? Какие основные виды уравнений регрессии применяются на практике?
5. Чем отличается уравнение регрессии Y на X от уравнения регрессии X на Y?
6. Что такое метод наименьших квадратов?
7. Что положено в основу определения коэффициентов аппроксимирующих функций в методе наименьших квадратов?
8. Воспроизведите запись условия минимума квадратов отклонений точек графика от прямой.
9. Воспроизведите систему уравнений для определения коэффициентов, определяющих расположение аппроксимирующих прямых в методе наименьших квадратов.
10. В чем смысл квадратичной аппроксимации в методе наименьших квадратов?
11. Как формируется система уравнений для определения коэффициентов уравнения квадратичной аппроксимации?
12. Чем объяснить тот факт, что в методе наименьших квадратов наибольшее распространение получила линейная аппроксимация?
13. Каков смысл коэффициента регрессии Y на X? X на Y?
14. Запишите линейное уравнение регрессии как уравнение прямой, проходящей через заданную точку с заданным угловым коэффициентом. Поясните смысл координат точки и углового коэффициента.
15. В чем сходство и в чем различие между прямыми регрессии Y на X и X на Y?
16. Что такое коэффициент корреляции? Как он выражается через коэффициенты регрессии?
17. Как определяется знак коэффициента корреляции?
18. Что такое корреляционные отношения?

19. Перечислите основные свойства коэффициента корреляции? Корреляционного отношения?
20. В каких пределах могут изменяться коэффициенты корреляции? Корреляционные отношения?
21. О чем говорит близость к нулю коэффициента корреляции? Корреляционного отношения?
22. О чем говорит близость к единице коэффициента корреляции? Корреляционного отношения?

§ 4.20. ЗАДАНИЕ НА РАСЧЕТНУЮ РАБОТУ, ЧАСТЬ 6

Задана исходная корреляционная таблица:

Таблица 4.20
корреляционная

$y_{0j} \setminus x_{0i}$	0,7	1,9	2,5	2,7	2,9	3,1	3,3	3,8	4,1	5,5
2,5	1	2								
3,0		1	4							
3,2				5	4					
3,5					3	6	4			
3,9						2	3	4		
4,7									3	2

Требуется:

1. Пересчитать значения элементов первой строки и первого столбца таблицы по формулам

$$x_i = (1+k)x_{0i} + k; \quad y_j = (1+k)y_{0j} + k,$$

где k определяется в описании первой части расчетной работы.

2. Считая корреляционную зависимость линейной, составить уравнения регрессии Y на X и X на Y.

3. Определить коэффициент корреляции и корреляционные отношения.

4. Построить прямые регрессии на поле статистических данных.

5. Проанализировать результаты.

Глава 5

ДИНАМИКА СЛУЧАЙНЫХ ПРОЦЕССОВ

§ 5.1. ВРЕМЕННЫЕ РЯДЫ

Динамическим рядом или *рядом динамики* называют ряд наблюдаемых параметров (в том числе статистических), отражающий изменение одной рассматриваемой величины от последовательно возрастающих или убывающих значений другой (независимой) величины.

Если независимая случайная величина представляет собой время, то динамический ряд называют *временны́м рядом*.

Числовые значения показателей динамических рядов называются *уровнями* этих рядов.

Если в качестве зависимых величин во временных рядах выступают определенные фиксированные моменты времени, то ряды называются *моментными* (пример: табл. 5.1), если промежутки времени — *интервальными* (пример: табл. 5.2).

Длиной временного ряда называют время от начального момента наблюдения за процессом до его окончания. Длина рядов, представленных таблицами 5.1 и 5.2, равна четырем месяцам.

Таблица 5.1

Численность работников предприятия

Число. Месяц	1.1	1.2	1.3	1.4	30.4
Численность	823	846	864	940	1021

Таблица 5.2
Средняя зарплата работников предприятия

Месяц	Январь	Февраль	Март	Апрель
Зарплата	1862	1954	2031	2128

Иногда под длиной временного ряда понимают количество рассматриваемых в нем уровней. В этом понимании длина ряда табл. 5.1 равна 5, ряда табл. 5.2 — 4.

Ряды, охватывающие достаточно большой промежуток времени, такой что по ряду можно судить о тенденциях изменения зависимой случайной величины, называют *трендом* (английское «trend» — тенденция).

Для выявления тренда (тенденции развития) рассматриваемого процесса используют аппарат математической статистики и регрессионный анализ.

Во временных рядах последовательные значения уровней случайных величин, как правило, зависят от предыдущих и от истории их изменения. Поэтому использование для анализа таких рядов аппарата теории вероятностей и математической статистики требует осторожности.

Пусть имеется временной ряд, состоящий из совокупности n уровней:

$$Y = \{y_1;\ y_2;\ \ldots;\ y_n\}.$$

Во временных рядах в общем случае можно выделить четыре элемента ($t = \overline{1, n}$):

1) тренд с составляющими U_t;

2) сезонные компоненты V_t;

3) циклические компоненты C_t;

4) случайные компоненты ε_t.

Тренд, сезонная и циклическая компоненты представляют собой устойчивые, регулярные изменения процесса за достаточно продолжительный отрезок времени. Что касается четвертой компоненты, то ее поведение нельзя предсказать.

Она имеет случайный, нерегулярный характер и обладает свойствами:

- случайность колебаний уровней;
- подчинение нормальному закону распределения;
- равенство нулю математического ожидания;
- независимость уровней случайной последовательности.

Трендовые модели считаются адекватными реальному процессу, если их случайная компонента ε_t удовлетворяет всем четырем перечисленным свойствам. Если хотя бы одно из свойств не выполняется, модель считается неадекватной.

Исследование трендовой модели на адекватность может осуществляться различными способами, рассмотренными ниже.

§ 5.2. ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

Предварительный анализ временных рядов заключается в выявлении случайных выбросов (аномалий) и наличия тренда (закономерностей в поведении показателей ряда).

Аномальным уровнем называется такой уровень временного ряда, который не отвечает закономерностям протекания процесса и который оказывает существенное влияние на основные характеристики ряда, в том числе на его трендовую составляющую. Причинами появления аномальных уровней могут быть ошибки первого и второго родов.

Ошибками первого рода считаются выбросы показателей временного ряда, обусловленные техническими причинами (неверными показаниями приборов, протокольными описками, некорректной группировкой элементов ряда и т. д.).

Ошибки второго рода возникают во временном ряде из-за воздействия на процесс факторов, носящих объективный характер, но появляющихся эпизодически (резкая смена погоды, природные катаклизмы, эпидемии болезней и т. д.).

Если ошибки первого рода можно устранить, то ошибки второго рода устранению не подлежат.

Для выявления аномальных уровней во временных рядах разработаны специальные методы анализа статистического материала. Рассмотрим некоторые из них.

5.2.1. Метод Ирвина

Согласно методу Ирвина для выявления аномальных уровней во временных рядах используют относительные показатели λ_t $(t = \overline{1,n})$, определяемые по формуле:

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, \tag{5.1}$$

где

$$\sigma_y = \frac{1}{\sqrt{n-1}} \sqrt{\sum_{t=1}^{n} (y_t - \overline{y})^2} \tag{5.2}$$

— среднее квадратичное отклонение временного ряда;

$$\overline{y} = \frac{1}{n} \sum_{t=1}^{n} y_t \tag{5.3}$$

— его среднее значение.

Расчетные значения λ_2, λ_3, ..., λ_n сравниваются с табличными значениями критерия Ирвина (приведены для уровня значимости $\alpha = 0{,}05$ в таблице Приложения 6). Расчетное значение λ_t $(t = \overline{2,n})$ считается аномальным для некоторого t, если оно превышает соответствующее значение, взятое из указанной таблицы.

Пример. Для ряда распределения среднемесячных (t — номер месяца) объемов (y_t) продаж, представленного в табл. 5.3, методом Ирвина выявить аномальные уровни.

Таблица 5.3

Объем продаж и результаты расчета по методу Ирвина

t	1	2	3	4	5	6	7	8	9	10	11	12
y_t	2,6	2,8	3,6	3,9	4,4	5,2	4,9	5,8	3,6	3,1	2,5	3,0
$\lvert y_{t-1} - y_t \rvert$	—	0,2	0,8	0,3	0,5	0,6	0,3	0,9	1,2	0,5	0,6	0,5
$10 \cdot \lambda_t$	—	1,1	7,4	2,8	4,6	5,6	2,8	8,3	11	4,6	5,6	4,6

Р е ш е н и е.

Определим среднее арифметическое уровней заданного ряда по формуле (5.3):

$$\overline{y} = \frac{1}{n}\sum_{t=1}^{n} y_t = \frac{1}{12}(2{,}6 + 2{,}8 + \ldots + 3{,}0) \approx 3{,}78$$

и среднее квадратичное отклонение по формуле (5.2):

$$\sigma_y = \frac{1}{\sqrt{n-1}}\sqrt{\sum_{t=1}^{n}(y_t - \overline{y})^2} =$$

$$= \frac{1}{\sqrt{11}}\sqrt{(2{,}6 - 3{,}78)^2 + \ldots + (3{,}0 - 3{,}78)^2} \approx 1{,}08.$$

Для каждого уровня ряда, начиная со второго, по формуле (5.1) определяем значения λ_t $(t = \overline{2,12})$ и заносим их в последнюю строку табл. 5.3.

Для n=12 уровней ряда из таблице Приложения 6 путем линейной аппроксимации определим значение λ_α:

$$\lambda_\alpha(n{=}12) = \lambda_{kp} = \lambda_\alpha(10) - \frac{\lambda_\alpha(12) - \lambda_\alpha(20)}{12 - 10} \approx 1{,}46.$$

Сравнивая это значение критерия Ирвина со значениями λ_t, приведенными в последней строке табл. 5.3, убеждаемся в том, что критическое значение критерия Ирвина больше любого из табличных значений этого параметра. Следовательно, в исходном ряде отсутствуют аномальные уровни.

После выявления аномальных уровней выясняют причину их появления во временном ряду. Если точно установлено, что аномальный уровень обусловлен ошибкой первого рода, то ошибка устраняется. Устранить ошибку в таком аномальном уровне можно заменой его либо на среднее арифметическое двух соседних граничащих с ним уровней (линейная аппроксимация), либо на величину уровня, получаемого путем более сложной аппроксимации показателей временного ряда.

Аппроксимирующая кривая является трендовой моделью рассматриваемого процесса.

Для проверки *наличия тренда* в сформированных временных рядах используют методы, рассмотренные в следующих разделах.

5.2.2. Метод проверки разностей средних уровней

Согласно этому методу определения наличия тренда в исходном временном ряде следует осуществить четыре последовательных этапа.

Первый этап. Уровни y_1, y_2, ..., y_n разбиваются на две части с примерно равным количеством уровней: n_1 и n_2 ($n_1 + n_2 = n$).

Второй этап. Для каждой из двух частей полученных уровней вычисляются их средние значения:

$$\overline{y}_1 = \frac{1}{n_1}\sum_{t=1}^{n_1} y_t; \qquad \overline{y}_2 = \frac{1}{n_2}\sum_{t=1}^{n_2} y_t$$

и дисперсии:

$$\sigma^2_{y_1} = \frac{1}{n_1 - 1}\sum_{t=1}^{n_1}(y_t - \overline{y}_1)^2; \quad \sigma^2_{y_2} = \frac{1}{n_2 - 1}\sum_{t=1}^{n_2}(y_t - \overline{y}_2)^2.$$

Третий этап. Проверяется однородность (степень равенства) дисперсий обеих частей ряда. Для этого используется F-критерий Фишера:

$$F = \begin{cases} \dfrac{\sigma^2_{y_1}}{\sigma^2_{y2}}, & \text{если} \quad \sigma^2_{y_1} > \sigma^2_{y_2}, \\ \dfrac{\sigma^2_{y_2}}{\sigma^2_{y1}}, & \text{если} \quad \sigma^2_{y_2} < \sigma^2_{y_1}. \end{cases}$$

Найденное значение F сравнивается с табличным (критическим) значением F_{kp} при заданном уровнем значимости, равным 10% ($\alpha = 0{,}1$) в задачах подобного типа.

Если расчетное значение критерия Фишера больше критического, то гипотеза H_0 о равенстве дисперсий отклоняется и делается вывод о том, что рассматриваемый метод не позволяет сделать вывод о наличии тренда во временном ряде.

Таблица 5.4
Два подуровня исходной задачи

t	1	2	3	4	5	6
y_t	2,6	2,8	3,6	3,9	4,4	5,2
t	7	8	9	10	11	12
y_t	4,9	5,8	3,6	3,1	2,5	3,0

Если же $F \le F_{kp}$, то гипотеза H_0 принимается и можно переходить к четвертому этапу.

Четвертый этап. Проверяется гипотеза H_0 об отсутствии тренда с использованием t-критерия Стьюдента. Для этого определяется расчетное значение критерия Стьюдента:

$$t = \frac{|\overline{y}_1 - \overline{y}_2|}{\sqrt{(n_1-1)\overline{\sigma}_1^2 + (n_2-1)\overline{\sigma}_2^2}} \sqrt{\frac{n_1 n_2 (n_1+n_2-2)}{n_1+n_2}}.$$

Если расчетное значение t меньше критического значения T_{kp} критерия Стьюдента (таблица Приложения 3), то гипотеза H_0 принимается на заданном уровне значимости α, т. е. тренда нет. В противном случае тренд имеется.

Заметим, что критическое значение берется для числа степеней свободы $k = n_1 + n_2 - 2$. Метод применим только для рядов с монотонно изменяющимися величинами.

Пример. Для ряда распределения среднемесячных (t — номер месяца) объемов (y_t) продаж товаров некоторой фирмы, представленного в табл. 5.3, методом проверки разностей средних уровней определить наличие тренда.

Р е ш е н и е. Первый этап. Согласно изложенному в этом разделе методу разобьем исходный ряд, представленный табл. 5.3, на два ряда (табл. 5.4) с одинаковым количеством уровней $n_1 = n_2 = 6$ ($n_1 + n_2 = 12$).

Второй этап. Для каждой из выделенных частей ряда находим средние арифметические:

$$\overline{y}_1 = \frac{1}{n_1}\sum_{t=1}^{n_1} y_t = \frac{1}{6}(2{,}6 + \ldots + 5{,}2) \approx 3{,}75;$$

$$\overline{y}_2 = \frac{1}{n_2}\sum_{t=1}^{n_2} y_t = \frac{1}{6}(4{,}9 + \ldots + 3{,}0) \approx 3{,}82$$

и дисперсии:

$$\sigma^2_{y_1} = \frac{1}{n_1 - 1}\sum_{t=1}^{n_1}(y_t - \overline{y}_1)^2 = \\ = \frac{1}{5}\left((2{,}6 - 3{,}75)^2 + \ldots + (5{,}2 - 3{,}75)^2\right) \approx 0{,}96;$$

$$\sigma^2_{y_2} = \frac{1}{n_2 - 1}\sum_{t=1}^{n_2}(y_t - \overline{y}_2)^2 = \\ = \frac{1}{5}\left((4{,}9 - 3{,}82)^2 + \ldots + (3{,}0 - 3{,}82)^2\right) \approx 1{,}57.$$

Третий этап. Определяем критерий Фишера, равный отношению большей дисперсии к меньшей:

$$F = \frac{\sigma^2_{y_2}}{\sigma^2_{y_1}} = \frac{1{,}57}{0{,}96} \approx 1{,}64.$$

По таблице Приложения 5 критических значений критерия Фишера для уровня значимости $\alpha = 0{,}05$ и двух степеней свободы $m_1 = n_1 - 1 = 6 - 1 = 5$ и $m_2 = n_2 - 1 = 6 - 1 = 5$ определяем: $F_{kp} = 5{,}5$. Это величина больше расчетного значения $f = 1{,}64$, поэтому гипотеза о равенстве дисперсий двух частей одного исходного ряда принимается.

Четвертый этап по изложенной методике можно проводить только для рядов с монотонно изменяющимися уровнями. Этому требованию заданный ряд не удовлетворяет. Поэтому для него невозможно дать однозначное заключение о наличии тренда.

5.2.3. Метод Фостера–Стюарта

Как и метод проверки разностей средних уровней, метод Фостера–Стюарта состоит из четырех этапов.

Первый этап. Сравнивается каждый последующий уровень исходного временного ряда, начиная со второго, со всеми

предыдущими. При этом определяются две числовые последовательности ($t = \overline{2, n}$):

$$k_t^+ = \begin{cases} 1, & \text{если } y_t > y_i \text{ для всех } i = \overline{1, t-1}; \\ 0, & \text{в противном случае,} \end{cases} \tag{5.4}$$

$$k_t^- = \begin{cases} 1, & \text{если } y_t < y_i \text{ для всех } i = \overline{1, t-1}; \\ 0, & \text{в противном случае.} \end{cases} \tag{5.5}$$

Второй этап. Определяются величины:

$$s = \sum_{t=2}^{n}(k_t^+ + k_t^-); \quad d = \sum_{t=2}^{n}(k_t^+ - k_t^-). \tag{5.6}$$

Величина s характеризует изменение временного ряда и принимает значения от 0 (все уровни ряда равны между собой) до $n-1$ (ряд монотонный).

Величина d характеризует изменение дисперсии уровней временного ряда и меняется от $-(n-1)$ (монотонно убывающий ряд) до $+(n-1)$ (монотонно возрастающий ряд).

Третий этап. С использованием расчетных значений t-критерия Стьюдента проверяется гипотеза о том, можно ли считать случайными отклонения двух величин:

1) средней

$$t_s = \frac{|s - \mu|}{\sigma_s}; \tag{5.7}$$

2) дисперсии

$$t_d = \frac{|d - 0|}{\sigma_d}. \tag{5.8}$$

Здесь μ — математическое ожидание величины s, полученной в предположении, что в заданном исходном ряде уровни расположены в случайном порядке (располагаются с привлечением алгоритма получения случайных чисел);

$$\sigma_s = \sqrt{2 \ln n - 3{,}4253}, \quad \sigma_d = \sqrt{2 \ln n - 0{,}8456} \tag{5.9}$$

— средние квадратичные отклонения для величин s и d.

Величины μ, σ_s и σ_d задаются в виде таблицы, фрагмент которой представлен (табл. 5.5).

Таблица 5.5
Математическое ожидание и дисперсии

n	10	20	30	40
μ	3,858	5,195	5,990	6,557
σ_s	1,088	1,677	1,882	2,019
σ_d	1,964	2,279	2,447	2,561

Четвертый этап. Расчетные значения t_s и t_d сравниваются с табличными значениями T_{kp} критерия Стьюдента при заданном уровне значимости α. Если расчетное значение меньше табличного, то гипотеза об отсутствии соответствующего тренда принимается; в противном случае тренд есть. Если, например, $t_s > T_{kp}$, а $t_d < T_{kp}$, то для рассматриваемого ряда тренд для средних величин уровней ряда имеется, а для дисперсии уровней его нет.

Пример. Для ряда, представленного двумя первыми строками табл. 5.3, методом Фостера–Стюарта установить факт существования тренда ряда и его дисперсии.

Р е ш е н и е. Первый этап. В приведенным к виду табл. 5.7 заданном ряде сравниваем последующие значения уровней y_t ряда со всеми предыдущими. Используя соотношения (5.4)–(5.5), получаем значения k_t^+ и k_t^-, которые занесены в последние две строки табл. 5.7.

Второй этап. По формулам (5.6) вычисляем величины s и d, характеризующие изменение временного ряда:

$$s = \sum_{t=2}^{12}(k_t^+ + k_t^-) = 7; \qquad d = \sum_{t=2}^{12}(k_t^+ - k_t^-) = 5.$$

Третий этап. Используя данные табл. 5.5 путем их линейной аппроксимации на значение $n = 12$ находим математическое ожидание

$$\mu_{12} = \frac{\mu_{20} - \mu_{10}}{n_2 - n_1} \cdot 2 + \mu_{10} = \frac{5{,}195 - 3{,}858}{20 - 10} + 3{,}858 \approx 4{,}11$$

Таблица 5.6
Объем продаж в условных единицах

t	1	2	3	4	5	6	7	8	9	10	11	12
y_t	2,6	2,8	3,6	3,9	4,4	5,2	4,9	5,8	3,6	3,1	2,5	3,0
k_t^+	—	1	1	1	1	1	0	1	0	0	0	0
k_t^-	—	0	0	0	0	0	0	0	0	0	1	0

и дисперсии: $\sigma_s \approx 1{,}24$; $\sigma_d \approx 2{,}03$ (такие же результаты для дисперсии получаются при их подсчете по формулам (5.9)).

Используя соотношения (5.7) и (5.8) определяем:

$$t_s = \frac{|s-\mu|}{\sigma_s} = \frac{7-4{,}11}{1{,}24} \approx 3{,}69; \quad t_d = \frac{|d-0|}{\sigma_d} == \frac{6-0}{2{,}03} \approx 2{,}46.$$

Задаваясь уровнем значимости $\alpha = 0{,}05$, для числа степеней свободы $k = m - 2 = 12 - 2 = 10$ по таблице Приложения 3 находим критическое значение критерия Стьюдента: $T_{kp} = 2{,}23$.

Так как $t_s > T_{kp}$ $(3{,}69 > 2{,}23)$ и $t_d > T_{kp}$ $(2{,}46 > 2{,}23)$, то гипотеза об отсутствии тенденции как в среднем значении ряда, так и в дисперсии отвергается. Следовательно, в рассмотренном временном ряде присутствует тренд и имеется тенденция в изменении дисперсии.

§ 5.3. СГЛАЖИВАНИЕ ВРЕМЕННЫХ РЯДОВ

Колебания временных рядов, вызванные случайными воздействиями, зачастую скрывают характер и тенденции протекания процессов. Поэтому для более четкого выявления сути рассматриваемых процессов их исходные вариационные ряды желательно сгладить.

Рассмотрим два метода сглаживания.

5.3.1. Метод скользящей средней

Для временного ряда с исходными значениями n уровней $y_1, y_2, \ldots, y_t, \ldots, y_n$ задается начальный интервал сглаживания $y_1, y_2, \ldots, y_m$ $(m < n)$.

Если у ряда необходимо сохранить достаточно слабые колебания, то число m выбранных для сглаживания уровней должно быть мало, в противном случае m увеличивают. Для удобства вычислений рекомендуется число m задавать нечетным.

После выбора начального интервала сглаживания для его m уровней вычисляется средняя арифметическая $\overline{y}_{(m+1)/2}$ — это первый сглаженный уровень, соответствующий значению аргумента $t=(m+1)/2$. Далее рассматривается второй интервал сглаживания смещением ряда вправо на один уровень и находится его средняя арифметическая $\overline{y}_{(m+3)/2}$. И так далее. На шаге t средняя арифметическая интервала сглаживания находится по формуле $\left(p=\dfrac{m-1}{2}<t\right)$:

$$\overline{y}_t=\frac{1}{m}\sum_{i=t-p}^{t+p}y_i. \tag{5.10}$$

После «прогонки» описанной последовательности от начала до конца временного ряда получаются $(n-m+1)$ сглаженных уровней ряда. При этом первые p и последние p значений ряда остаются несглаженными.

Пример. Методом скользящей средней произвести выравнивание ряда, представленного двумя первыми строками табл. 5.7.

Таблица 5.7

Выравнивание ряда различными методами

t	1	2	3	4	5	6	7	8	9	10	11	12
y_t	2,6	2,8	3,6	3,9	4,4	5,2	4,9	5,8	3,6	3,1	2,5	3,0
$\overline{y}_t^n$	–	–	3,4	4,0	5,0	5,8	5,8	5,1	4,0	3,6	–	–
$\overline{y}_t^э$	2,9	2,9	3,0	3,2	3,3	3,6	3,8	4,1	4,0	3,9	3,7	3,6

Р е ш е н и е.

Принимая $m=5$, найдем $p=\dfrac{m-1}{2}=2$ и затем определим по формуле (5.10) новые, осредненные по пяти точкам

значения уровней, начиная с третьего:

$$\overline{y}_3 = \frac{1}{5}\sum_{i=3-2}^{3+2} y_i = \frac{1}{5}(y_1 + \ldots + y_5) = \frac{1}{5}(2{,}6 + \ldots + 4{,}4) \approx 3{,}4;$$

$$\overline{y}_4 = \frac{1}{5}(y_2 + \ldots + y_6) = \frac{1}{5}(2{,}8 + \ldots + 5{,}2) \approx 4{,}0;$$

. . . . . .

$$\overline{y}_{10} = \frac{1}{5}(y_8 + \ldots + y_{12}) = \frac{1}{5}(5{,}8 + \ldots + 3{,}0) \approx 3{,}6.$$

Вычисленные значения осредненных уровней внесены в табл. 5.7 в строчку, обозначенную $\overline{y}_t^n$. Первые и последние два значения уровня при описанном методе осреднения не подсчитываются. Они могут быть приняты равными исходным уровням или подсчитываются с использованием дополнительных предположений [8].

Отличительной особенностью метода скользящей средней сглаживания временных рядов является равное «участие» всех уровней интервала сглаживания в подсчете средней арифметической. Однако в большинстве реальных процессов влияние предшествующих событий на последующие ослабевает со временем. Это обстоятельство учитывается, в частности, в методе экспоненциального сглаживания, рассмотренном в следующем разделе.

5.3.2. Метод экспоненциального сглаживания

Отличительной особенностью этого метода является то, что для нахождения сглаженного уровня в нем используются значения только предшествующих уровней ряда, умноженные на их веса. Веса уровней уменьшаются по мере их удаления от момента времени, для которого определяется сглаженное значение уровня.

Сглаженные значения временного ряда $\overline{y}_t$ $(t = \overline{1,n})$ определятся по формуле:

$$\overline{y}_t = \alpha y_t + (1 - \alpha)\overline{y}_{t-1}, \tag{5.11}$$

где $\alpha \in [0;1]$ — параметр сглаживания. Величина $1 - \alpha$ называется *коэффициентом дисконтирования*.

Подставляя в рекуррентное соотношение (5.11) последовательно все значения $\overline{y}_t$ от нулевого до $(n-1)$-го, придем к следующей формуле определения среднего значения уровня сглаженного экспоненциальным методом ряда:

$$\overline{y}_t = (1-\alpha)^t y_0 + \alpha \sum_{i=0}^{t-1} (1-\alpha)^i y_{t-i}.$$

Здесь $\overline{y}_0$ — нулевая, не входящая в сглаженный ряд, фиктивная величина. Обычно она принимается или равной значению y_1 исходного временного ряда, или среднему арифметическому его трех первых значений:

$$\overline{y}_0 = \frac{1}{3}(y_1 + y_2 + y_3).$$

В практических задачах обработки временных рядов величину параметра сглаживания ограничивают отрезком $\alpha \in \in [0{,}1;\ 0{,}3]$. Р. Браун предлагает определять величину α в зависимости от длины выравниваемого ряда:

$$\alpha = \frac{2}{n+1}.$$

Если сглаженные описанным методом значения правого конца ряда существенно отличаются от соответствующих значений исходного ряда, то необходимо перейти к новому параметру α.

В методе экспоненциального сглаживания начальные и конечные уровни получаемого ряда не теряются.

Пример. Методом экспоненциального сглаживания выровнять ряд, представленный двумя первыми строками табл. 5.7.

Р е ш е н и е.

Предварительно по формуле Брауна определим параметр сглаживания:

$$\alpha = \frac{2}{n+1} = \frac{2}{12+1} \approx 0{,}15; \quad 1-\alpha = 0{,}85$$

и найдем начальное значение:

$$\overline{y}_0 = \frac{1}{3}(y_1 + y_2 + y_3) = \frac{1}{3}(2{,}6 + 2{,}8 + 3{,}6) = 3{,}0.$$

По рекуррентной формуле (5.11) последовательно находим:

$$\overline{y}_1 = \alpha y_1 + (1-\alpha)\overline{y}_0 = 0{,}15 \cdot 2{,}6 + 0{,}85 \cdot 3{,}0 = 2{,}9;$$

$$\overline{y}_2 = \alpha y_2 + (1-\alpha)\overline{y}_1 = 0{,}15 \cdot 2{,}8 + 0{,}85 \cdot 2{,}9 = 2{,}9;$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\overline{y}_{12} = \alpha y_{12} + (1-\alpha)\overline{y}_{11} = 0{,}15 \cdot 3{,}0 + 0{,}85 \cdot 3{,}7 = 3{,}6.$$

Все рассчитанные значения уровней сглаженного экспоненциальным методом ряда внесены в табл. 5.7 в строку, обозначенную $\overline{y}_t^{\,э}$.

Следует обратить внимание на существенные расхождения сглаженных уровней одного и того же исходного ряда, полученных двумя разными методами. Даже средние арифметические значения у полученных рядов отличаются друг от друга и от исходного ряда. Отмеченную особенность необходимо иметь в виду при сглаживании рядов, описывающих реальные процессы. Есть опасность неоправданного искажения исходных данных.

§ 5.4. ХАРАКТЕРИСТИКИ ВРЕМЕННОГО РЯДА

Важной характеристикой интервального временного ряда с заданными уровнями y_1, y_2, $\ldots$, y_n является их среднее арифметическое, называемое *средним уровнем ряда*:

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \qquad (5.12)$$

Если ряд имеет неравноотстоящие по времени уровни, то его средний уровень, называемый иногда *средней хронологической* может быть вычислен по формуле:

$$\overline{y} = \frac{\sum\limits_{i=1}^{n} y_i t_i}{\sum\limits_{i=1}^{n} t_i}. \tag{5.13}$$

Для моментного ряда с равноотстоящими уровнями формула для подсчета средней хронологической может быть приведена к виду, напоминающему формулу трапеций приближенного вычисления определенного интеграла:

$$\overline{y} = \frac{1}{n-1}\left(\frac{1}{2}(y_1 + y_n) + \sum_{i=2}^{n-1} y_i\right). \tag{5.14}$$

Формула для этой же характеристики, но для ряда с разноотстоящими по времени уровнями имеет вид

$$\overline{y} = \frac{1}{2\sum\limits_{i=1}^{n} t_i}\left(\sum_{i=1}^{n-1}(y_i + y_{i+1})t_i\right). \tag{5.15}$$

Здесь t_i — половина времени протекания i-го и $(i+1)$-го уровней.

При статистическом анализе временных рядов возникает необходимость оценить зависимость значения уровня в рассматриваемый момент времени от его значений в предшествующие моменты времени (происходящие с запаздыванием). Эти моменты времени могут быть сдвинуты по отношению к рассматриваемому уровню на единицу, два и т. д.

Зависимость значения рассматриваемого уровня временного ряда от предыдущих называется *автокорреляцией* во временном ряду.

Числовой характеристикой зависимости значений уровней ряда от предыдущих его значений является взаимная корреляционная функция между исходным рядом с уровнями $y_i(t)$

и этим же рядом, но сдвинутым во времени на величину τ с уровнями $y_i(t+\tau)$.

Такая числовая характеристика называется *корреляционной функцией*. Она характеризует внутреннюю структуру временного ряда и зависит от множества коэффициентов автокорреляции, которые определяются по формуле:

$$r_\tau = \frac{Y_{t;t+\tau}}{\sqrt{Y_{t;t} Y_{t+\tau;t+\tau}}}, \tag{5.16}$$

где

$$Y_{t;t} = (n-n_\tau) \sum_{i=1}^{n-n_\tau} y_i^2 - \left(\sum_{i=1}^{n-n_\tau} y_i \right)^2;$$

$$Y_{t;t+\tau} = (n-n_\tau) \sum_{i=1}^{n-n_\tau} y_i y_{i+n_\tau} - \sum_{i=1}^{n-n_\tau} y_i \sum_{i=1}^{n-n_\tau} y_{i+n_\tau};$$

$$Y_{t+\tau;t+\tau} = (n-n_\tau) \sum_{i=1}^{n-n_\tau} y_{i+n_\tau}^2 - \left(\sum_{i=1}^{n-n_\tau} y_{i+n_\tau} \right)^2.$$

Задаваясь различными значениями сдвигов $n_\tau = 1, 2, 3, \ldots$, можно по приведенной формуле получить различные значения коэффициентов корреляции: r_1, r_2, r_3, ... В практических задачах эти коэффициенты рекомендуется вычислять в количествах от $n/4$ до $n/3$.

График автокорреляционной функции называется *коррелограммой* и показывает величину запаздывания, с которым изменение уровня y_i связывается на его последующих значениях. Сдвиг τ, которому соответствует наибольшее по абсолютной величине значение r_τ, называется *временным лагом*.

Часто вместо формулы (5.16) используют упрощенную, менее точную формулу:

$$r_\tau = \frac{\dfrac{1}{n-n_\tau} \displaystyle\sum_{i=1}^{n-n_\tau} (y_i - \overline{y})(y_{i+n_\tau} - \overline{y})}{\dfrac{1}{n} \displaystyle\sum_{i=1}^{n-n_\tau} (y_i - \overline{y})^2}, \tag{5.17}$$

где $\overline{y}$ — средний уровень ряда (5.12).

Таблица 5.8
Временной ряд $y(t)$

t	1	2	3	4	5	6	7	8	9	10	11	12
y_t	2,6	2,8	3,6	3,9	4,4	5,2	4,9	5,8	3,6	3,1	2,5	3,0

Пример.

Определить автокорреляционную функцию и коэффициенты автокорреляции для временного ряда, заданного в табл. 5.8.

Р е ш е н и е.

Заданный ряд содержит $n = 12$ уровней ($t = \overline{1,12}$).

Согласно приведенным выше рекомендациям определим значения сдвигов τ исходного ряда для построения новых рядов:

$$\tau_1 = \frac{n}{4} = \frac{12}{4} = 3; \qquad \tau_2 = \frac{n}{4} = \frac{12}{3} = 4.$$

Других целых значений τ на интервале $\tau \in [3;4]$ нет.

Построим новые ряды, сдвинутые по отношению к исходному на 3 единицы (табл. 5.9) и 4 единицы (табл. 5.10). В указанные таблицы занесены и некоторые результаты промежуточных вычислений.

В табл. 5.9 значения уровней y_t сдвинуты по сравнению со значениями уровней табл. 5.8 на $\tau = 3$ единицы вперед, а в табл. 5.10 — на $\tau = 4$ единицы.

Определим для τ=3 входящие в формулу (5.16) величины:

$$Y_{t,t} = (n-3)\sum_{i=i}^{n+3} y_i^2 - \left(\sum_{i=i}^{n+3} y_i^2\right)^2 = \\ = 9 \cdot 159{,}78 - (36{,}8)^2 = 83{,}78;$$

$$Y_{t,t+3} = (n-3)\sum_{i=i}^{n+3} y_i y_{i+3} - \sum_{i=i}^{n+3} y_i \sum_{i=i}^{n+3} y_{i+3} = \\ = 9 \cdot 145{,}02 - 36{,}8 \cdot 36{,}4 = -34{,}34;$$

Таблица 5.9
Ряд со сдвигом $\tau = 3$

t	y_t	y_{t+3}	y_t^2	$y_t y_{t+3}$	y_{t+3}^2
1	2,6	3,9	6,76	10,14	15,21
2	2,8	4,4	7,84	12,32	19,36
3	3,6	5,2	12,96	18,72	27,04
4	3,9	4,9	15,21	19,11	24,01
5	4,4	5,8	19,36	25,52	33,64
6	5,2	3,6	27,04	18,72	12,96
7	4,9	3,1	24,01	15,19	9,61
8	5,8	2,5	33,64	14,50	6,25
9	3,6	3,0	12,96	10,80	9,00
$\sum$	36,8	36,4	159,78	145,02	157,08

Таблица 5.10
Ряд со сдвигом $\tau = 4$

t	y_t	y_{t+4}	y_t^2	$y_t y_{t+4}$	y_{t+4}^2
1	2,6	4,4	6,76	11,44	19,36
2	2,8	5,2	7,84	14,56	27,04
3	3,6	4,9	12,96	17,64	24,01
4	3,9	5,8	15,21	22,62	33,64
5	4,4	3,6	19,36	15,84	12,96
6	5,2	3,1	27,04	16,12	9,61
7	4,9	2,5	24,01	12,25	6,25
8	5,8	3,0	33,64	17,04	9,00
$\sum$	32,5	32,9	146,82	127,87	141,87

$$Y_{t+3,t+3} = (n-3)\sum_{i=i}^{n+3} y_{i+3}^2 - \left(\sum_{i=i}^{n+3} y_{i+3}\right)^2 = \\ = 9 \cdot 157{,}08 - (36{,}4)^2 = 88{,}76.$$

Теперь по формуле (5.16) найдем коэффициент корреляции для $\tau = 3$:

$$r_3 = \frac{Y_{t;t+3}}{\sqrt{Y_{t;t} Y_{t+3;t+3}}} = \frac{-33{,}34}{\sqrt{83{,}78 \cdot 88{,}76}} \approx -0{,}398.$$

Аналогично, используя данные табл. 5.10, определим:

$$r_4 \approx -0{,}587.$$

Так как в рекомендуемом диапазоне изменения τ $|r_4| > |r_3|$, то величина сдвига $\tau = 4$ является временным лагом.

§ 5.5. РАСЧЕТ ПОКАЗАТЕЛЕЙ ДИНАМИКИ ВРЕМЕННОГО РЯДА

Расчет показателей динамики процессов проводится на базе соответствующих временных рядов с заданными уровнями $y_1, y_2, \ldots, y_n$. Для этого предварительно преобразуют абсолютные показатели уровня в относительные, принимая за базовый некоторый фиксированный уровень, например y_1, или относя каждый последующий уровень к предыдущему. В первом случае вновь полученные относительные значения уровней называются *базисными*; во втором — *цепными*.

На практике используют различные показатели динамики развития (скорости изменения уровней) процессов. Рассмотрим некоторые из них.

Абсолютный прирост определяется приращением величин уровней ряда (под y_i $(i = \overline{1,n})$ в дальнейшем понимаются относительные базисные или цепные показатели уровней ряда):

$$\Delta y_i = y_i - y_{i-k}. \tag{5.18}$$

Индекс k в этом выражении может принимать любые значения от 1 до $n-1$ в зависимости от цели исследования. При $k = 1$ получаются цепные показатели; при $k = i - 1$ — базисные.

Средний абсолютный прирост — это величина, определяемая отношением:

$$\overline{\Delta y_i} = \frac{y_i - y_{i-k}}{k}. \tag{5.19}$$

В частности средний абсолютный прирост за весь период времени протекания процесса

$$\overline{\Delta y} = \frac{y_n - y_1}{n-1} \tag{5.20}$$

определяет среднюю скорость изменения показателей ряда.

Для определения относительной скорости изменения показателей ряда используются приведенные ниже относительные показатели.

Коэффициент роста для i-го периода:

$$K_{i(p)} = \frac{y_i}{y_{i-k}}. \tag{5.21}$$

При $K_{i(p)} > 0$ уровень повышается; при $K_{i(p)} = 0$ не изменяется и при $K_{i(p)} < 0$ понижается со временем.

Коэффициент прироста:

$$K_{i(np)} = \frac{y_i - y_{i-k}}{y_{i-k}}. \tag{5.22}$$

Средняя скорость изменения какого-либо показателя ряда за рассматриваемый период времени может быть определена как средняя геометрическая этого показателя. Например, *средний темп роста* определяется формулой:

$$\overline{T}_{(p)} = \sqrt[n-1]{T_{1(p)} \cdot T_{2(p)} \cdot \ldots \cdot T_{n(p)}} = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100\%. \tag{5.23}$$

Последняя формула в силу того, что в ней для определения среднего темпа используются только начальное y_1 и конечное y_n значения уровней ряда, нечувствительна к колебаниям уровня во внутренних точках этого ряда. Более корректной в этом отношении является формула, в которой фигурирует отношение сглаженных по уровню тренда показателей $\overline{y}_1$ и $\overline{y}_n$:

$$\overline{T}_{(p)} = \sqrt[n-1]{\frac{\overline{y}_n}{\overline{y}_1}} \cdot 100\%. \tag{5.24}$$

§ 5.6. ТРЕНД-СЕЗОННЫЕ ПРОЦЕССЫ И ИХ АНАЛИЗ

Сезонными колебаниями называют регулярные, периодические наступления внутригодовых подъемов и спадов показателей временных рядов. Если рассматриваемый процесс подвержен периодическим колебаниям, имеющим определенный и не меняющийся период, равный некоторому годовому промежутку времени, то составленный по его выборочным данным ряд называется *тренд-сезонным*.

Ниже рассматриваются тренд-сезонные временные ряды, уровни которых $Y = \{y_1, y_2, \ldots, y_n\}$ складываются из тренда $U = \{u_1, u_2, \ldots, u_n\}$, сезонной волны $V = \{v_1, v_2, \ldots, v_n\}$ и случайной компоненты $E = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$:

$$Y_t = U_t + V_t + E_t, \tag{5.25}$$

где индекс t указывает на то, что слагаемые формулы (5.25) принадлежат подынтервалу t всего интервала времени T протекания процесса ($t \in T$).

Перечислим основные задачи исследования временных рядов.

1. Выявление в исходном ряде тренда.
2. Выявление наличия сезонных колебаний.
3. Фильтрация (отделение друг от друга) компонент ряда.
4. Анализ динамики сезонной волны.
5. Анализ сезонных колебаний.
6. Прогнозирование тренд-сезонных процессов.

Относительно тренда U_t предполагается, что это гладкая функция, степень гладкости которой заранее неизвестна, но представляющая собой полином как можно меньшей степени.

Выявление наличия во временном ряду сезонных колебания сводится к проверке на случайность ряда, оставшегося после выделения тренда.

Под фильтрацией понимается выделение из исходного ряда с уровнями y_t его составляющих u_t, v_t и ε_t.

Сезонная компонента (волна) v_t имеет период T_0 (для годового ряда с месячными данными $T_0 = 12$; с квартальными $T_0 = 4\ \ldots$), так что $v_{t+T_0} = v_t$. Кроме того, весь период T времени процесса должен без остатка делиться на T_0, что равносильно выражению:

$$T = mT_0,$$

где m — натуральное число. Если T_0 — число месяцев или кварталов в году, то m — число лет протекания процесса, описываемого временным рядом с уровнями y_t.

Для объективного суждения о процессе необходимо более детальное исследование (анализ) динамики сезонной волны, которое сводится к решению на каждом временном интервале трех связанных между собой задач анализа:

1) динамики амплитуды волны;
2) динамики точек экстремума;
3) изменений формы волны.

Часто исходные данные тренд-сезонного временного ряда представляются в виде матрицы:

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1j} & \cdots & y_{1T_0} \\ y_{21} & y_{22} & \cdots & y_{2j} & \cdots & y_{2T_0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{i1} & y_{i2} & \cdots & y_{ij} & \cdots & y_{iT_0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mj} & \cdots & y_{mT_0} \end{pmatrix};$$

$$(i = \overline{1, m}; \quad j = \overline{1, T_0}).$$

Индексы i, j и t связаны зависимостями:

$$i = \left[\frac{t}{T_0}\right] + 1; \qquad j = t - (i-1) \cdot T_0. \tag{5.26}$$

Квадратные скобки [ ] обозначают целую часть заключенного между ними числа.

После введения двухиндексных обозначений зависимость (5.25) запишем в виде

$$y_{ij} = u_{ij} + v_{ij} + \varepsilon_{ij}. \tag{5.27}$$

При исследовании временных рядов, представляемых в виде суммы (5.25), операции сглаживания и фильтрации имеют некоторые особенности.

Под *сглаживанием* тренд-сезонного временного ряда понимается процесс получения оценок суммы $\left(U_t + V_t\right)$, а под *фильтрацией* — суммы $\left(U_t + V_t + E_t\right)$. Методы сглаживания временных рядов рассмотрены выше в § 5.3. Что касается методов фильтрации, то они делятся на регрессионные, спектральные и итерационные. Рассмотрим для примера один из итерационных методов.

§ 5.7. МЕТОД ФИЛЬТРАЦИИ ЧЕТВЕРИКОВА

Идея методов фильтрации заключается в многократном применении *идентифицированной скользящей средней*:

$$\overline{y}_t = \frac{1}{T_0}\Big(\frac{1}{2}(y_{t-T_0/2} + y_{t+T_0/2}) + y_{t-T_0/2-1} + \ldots \\ \ldots + y_t + \ldots + y_{t-T_0/2+1}\Big) \quad (5.28)$$

и параллельной оценки сезонной компоненты на каждой итерации.

При использовании итерационных методов фильтрации, которые в применении к решению практических задач дают удовлетворительную «чистоту», возникает проблема потери части информации на концах интервалов.

Рассмотрим последовательность решения задачи фильтрации временного ряда на примере одного из распространенных итерационных методов — *метода Четверикова*.

1. Эмпирический ряд $Y = \{y_{11}; \ldots; y_{ij}; \ldots; y_{mT_0}\} = \{y_{ij}\}$ выравнивается скользящей средней (5.28) ($\overline{y}_t = \overline{y}_{ij}$) с периодом скольжения T_0, т. е. берутся $(T_0 + 1)$ членов ряда, из которых первый и последний имеют половинный вес: $\alpha_{-T_0/2} = \alpha_{T_0/2} = 1/2$. Выпадающие с обоих концов ряда уровни могут быть восстановлены, например, экстраполированием.

Использование формулы (5.28) приводит к получению значений выровненного ряда, который служит предварительной

оценкой тренда (оценкой в нулевом приближении):

$$Y' = \{\overline{y}_{ij}\}.$$

2. Определяются отклонения исходного ряда от выровненного:

$$L = Y - Y' \qquad (t = \overline{1,T}). \tag{5.29}$$

Для уровней ряда при двухиндексном (матричном) представлении эта формула принимает вид

$$l_{ij} = y_{ij} - \overline{y}_{ij} \qquad (i = \overline{1,m}; \quad j = \overline{1,T_0}). \tag{5.30}$$

3. Для каждого i-го года вычисляется среднее квадратичное отклонение σ_i, на которое затем делятся (нормируются) все внутригодовые (месячные, квартальные, ...) отклонения:

$$\tilde{l}_{ij} = \frac{l_{ij}}{\sigma_i}, \tag{5.31}$$

где

$$\sigma_i = \sqrt{\frac{1}{T_0 - 1}\left(\sum_{j=1}^{T_0} l_{ij}^2 - \frac{1}{T_0}\left(\sum_{j=1}^{T_0} l_{ij}\right)^2\right)}. \tag{5.32}$$

4. По найденным нормированным отклонениям вычисляется в первом приближении средняя нормированная сезонная волна $V^{(1)} = \{v_j^{(1)}\}$:

$$v_j^{(1)} = \frac{1}{m}\sum_{i=1}^{m} \tilde{l}_{ij}. \tag{5.33}$$

5. Значения уровней средней сезонной волны, полученные в первом приближении, умножаются на среднее квадратичное отклонение σ_i каждого i-того года и вычитаются из соответствующих значений исходного ряда. В результате получаются значения уровней тренда в первом приближении:

$$u_{ij}^{(1)} = y_{ij} - v_j^{(1)}\sigma_i. \tag{5.34}$$

6. Полученный сглаженный в первом приближении временной ряд, лишенный сезонной волны (в первом приближении), вновь сглаживается скользящей средней с использованием формул (5.29)–(5.34). В результате получается новая оценка тренда $U^{(2)}$ уже во втором приближении.

После этого по формуле, подобной (5.32), определяются отклонения значений уровней исходного ряда от ряда $U^{(2)}$:

$$l_{ij}^{(2)} = y_{ij} - u_{ij}^{(1)}.$$

Далее следует обработка ряда, следуя пунктам **2** и **3**.

Итерации продолжаются до выявления окончательной сезонной волны на некотором приближении.

Исключение окончательной сезонной волны производится после умножения средней сезонной волны на коэффициент напряженности сезонной волны:

$$k_i = \frac{\sum\limits_{j=1}^{T_0} l_{ij}^{(r)}\varepsilon_{ij}}{\sum\limits_{j=1}^{T_0}\varepsilon_{ij}}, \tag{5.35}$$

где $l_{ij}^{(r)}$ — выровненное значение ряда, ε_{ij} — случайная компонента, вычисляемая по формуле:

$$\varepsilon_{ij} = l_{ij}^{(r)} - v_{ij}^{(r)}. \tag{5.36}$$

§ 5.8. ПРИМЕР ФИЛЬТРАЦИИ РЯДА

Пусть временной ряд, характеризующий реализацию продукции фирмы в условных единицах, представлен в виде табл. 5.11, где в строках стоят показатели уровней ряда по месяцам, в столбцах — по годам.

Таблица 5.11
Исходный временной ряд

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1	76	68	90	97	108	115	114	117	112	108	93	96
2	85	81	99	107	112	120	124	127	118	122	110	106
3	94	90	114	117	131	131	137	137	124	127	107	105
4	88	102	116	125	133	136	139	138	131	130	120	108
5	92	80	109	120	140	144	144	145	136	134	124	121
6	108	102	126	135	150	153	152	150	145	141	131	124
7	119	115	128	143	155	156	152	158	155	148	139	140
8	108	116	133	148	161	166	164	167	167	157	147	148
9	122	125	148	157	169	173	177	176	164	164	152	142
10	132	125	153	164	174	173	179	178	170	169	158	149

Требуется выровнять ряд, выявить в нем тренд и сезонную волну, вычислить значения случайной компоненты.

Р е ш е н и е.

Первое приближение.

1. Проведем начальное выравнивание ряда, воспользовавшись формулой скользящей средней (5.28) и принимая период сглаживания $T_0 = 12$. Так как к операции сглаживания каждого уровня должны привлекаться $T_0/2 = 6$ уровней справа и шесть — слева от рассматриваемого, то упомянутая формула может быть применена лишь начиная с уровня $y_{1,7}$, то есть с уровня, находящегося в первой строке и седьмом столбце табл. 5.11. Вычислим сглаженное значение этого уровня:

$$\overline{y}_{1,7} = \frac{1}{12}\left(\frac{1}{2}(y_{1,1} + y_{2,1}) + y_{1,2} + y_{1,3} + \ldots + y_{1,12}\right) =$$
$$= \frac{1}{12}\left(\frac{1}{2}(76 + 85) + 68 + 90 + \ldots + 96\right) \approx 100;$$

$$\overline{y}_{1,8} = \frac{1}{12}\left(\frac{1}{2}(y_{1,2} + y_{2,2}) + y_{1,3} + y_{1,4} + \ldots + y_{2,1}\right) =$$
$$= \frac{1}{12}\left(\frac{1}{2}(68 + 81) + 90 + 97 + \ldots + 85\right) \approx 99;$$

и т. д.

$$\overline{y}_{10,6} = \frac{1}{12}\left(\frac{1}{2}(y_{9,12} + y_{10,12}) + y_{10,1} + y_{10,2} + \ldots + y_{10,11}\right) =$$
$$= \frac{1}{12}\left(\frac{1}{2}(142 + 149) + 132 + 125 + \ldots + 158\right) \approx 160.$$

Полученные величины образуют выровненный в первом приближении ряд и являются уровнями тренда $U^{(1)} = \{\overline{y}_{1,7}; \overline{y}_{1,8}; \ldots; \overline{y}_{10,6}\}$ в первом приближении. Особенности применяемого метода сглаживания не затрагивают первые шесть и последние шесть уровней исходного ряда. Поэтому в дальнейших таблицах эти уровни и соответствующие им величины отсутствуют.

Значения сглаженных уровней приведены в табл. 5.12.

На рис. 5.1 показаны точки, соответствующие значениям уровней заданного ряда по второму (○) и девятому (●) годам, а также прямые, полученные в результате выравнивания этих данных при использовании 13-точечной скользящей средней.

Полученные величины образуют выровненный в первом приближении ряд и являются уровнями тренда $U^{(1)} = \{\overline{y}_{1,7}; \overline{y}_{1,8}; \ldots; \overline{y}_{10,6}\}$ в первом приближении. Особенности применяемого метода сглаживания не затрагивают первые шесть и последние шесть уровней исходного ряда. Поэтому в дальнейших таблицах эти уровни и соответствующие им величины отсутствуют.

Таблица 5.12

Выровненный ряд нулевого приближения $\overline{y}_{ij}$

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1							100	99	101	102	102	103
2	103	104	105	106	106	107	107	108	109	110	111	112
3	113	114	115	115	116	115	115	115	116	116	117	117
4	117	117	118	118	118	119	119	118	117	116	116	117
5	117	118	119	119	119	120	122	123	125	126	127	128
6	129	129	130	130	131	131	133	134	134	135	135	136
7	136	136	137	138	138	139	139	139	139	140	140	141
8	142	143	143	144	145	146	147	148	149	150	150	151
9	152	153	153	153	154	154	154	154	154	155	155	156
10	156	157	158	159	159	160						

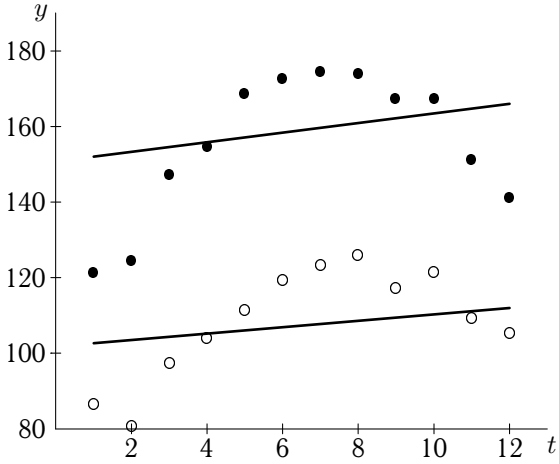


Рис. 5.1
Исходный ряд

Значения сглаженных уровней приведены в табл. 5.12.

2. Полученные в нулевом приближении уровни тренда (табл. 5.12) вычитаем из соответствующих значений исходного уровня, получая по формуле (5.30) значения остаточного

ряда:

$$l_{1,7} = y_{1,7} - \overline{y}_{1,7} = 114 - 100 = 14,$$
$$l_{1,8} = y_{1,8} - \overline{y}_{1,8} = 117 - 99 = 18,$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$
$$l_{10,6} = y_{10,6} - \overline{y}_{10,6} = 173 - 160 = 13.$$

Результаты вычислений занесены в табл. 5.13.

Таблица 5.13
Остаточный ряд l_{ij}

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1							14	18	11	6	-9	-7
2	-18	-23	-6	1	6	13	17	19	9	12	-1	-6
3	-19	-24	-1	2	15	16	22	22	8	9	-10	-12
4	-29	-15	-2	7	15	17	20	20	14	14	4	-9
5	-25	-38	-10	1	21	24	22	22	11	8	-3	-7
6	-21	-27	-4	5	19	22	19	16	9	6	-4	-12
7	-17	-21	-9	5	17	17	13	19	16	8	-1	-1
8	-34	-27	-10	4	16	20	17	19	18	7	-3	-3
9	-30	-28	-5	4	15	19	23	22	10	9	-3	-14
10	-24	-32	-5	5	15	13						

3. Проведем некоторые вычисления для определения средних квадратичных отклонений величин l_{ij}. Для всех рассматриваемых годовых промежутков ($i = \overline{1,10}$) получим данные, занесенные в табл. 5.13 (в первом и последнем годах ряд, как и в табл. 5.12, содержит данные по 6 месяцам, по остальным годам — по 12).

Для первого года:

$$\frac{1}{n_1-1}\sum_{j=7}^{12} l_{ij}^2 = \frac{1}{6-1}\left(14^2+18^2+\ldots+(-7)^2\right) = \frac{1}{5}\cdot 772 \approx 154.$$

Для второго года:

$$\frac{1}{n_2-1}\sum_{j=1}^{12} l_{ij}^2 = \frac{1}{12-1}\left((-18)^2+(-23)^2+\ldots+(-6)^2\right) = \\ = \frac{1}{11}\cdot 2007 \approx 183.$$

... ...

Для десятого года:

$$\frac{1}{n_{10}-1}\sum_{j=1}^{6} l_{ij}^2 = \frac{1}{6-1}\left((-24)^2+(-32)^2+\ldots+13^2\right) = \\ = \frac{1}{5}\cdot 2109 \approx 422.$$

Для каждого i-го года ($i=\overline{1,10}$) значения средних квадратичных отклонений вычисляются по формулам (5.30). Например,

$$\sigma_1^{(1)2} = \frac{1}{6-1}\left(\sum_{j=7}^{12} l_{1j}^2 - \left(\frac{1}{6}\sum_{j=7}^{12} l_{1j}\right)^2\right) = \\ = \frac{1}{5}\Big(14^2+18^2+\ldots+(-7)^2) - \\ - \frac{1}{6}(14+18+\ldots+(-7))^2\Big) = 116;$$

$\sigma_1^{(1)} = \sqrt{116} \approx 11.$

В табл. 5.14 приведены некоторые результаты промежуточных вспомогательных вычислений и округленные до целых чисел значения σ_j ($j=\overline{1,10}$).

Поделив каждое из значений строк табл. 5.13 на соответствующие значения σ_i из табл. 5.14, получим занесенные в табл. 5.15 значения $\overline{l}_{ij}$. Для сокращения записи данных таблицы они увеличены в 10 раз.

Таблица 5.14
Характеристики ряда в первом приближении

t	1	2	3	4	5	6	7	8	9	10
$\overline{l}_i^{(1)}$	5,3	1,9	2,3	2,2	4,7	2,3	3,8	2,0	1,8	-2,4
$\sigma_i^{(1)2}$	125	179	248	244	397	263	191	332	334	416
$\sigma_i^{(1)}$	11	13	16	16	20	16	14	18	18	20

4. В последней строке табл. 5.15 приведены значения предварительной средней сезонной волны $v_j^{(1)}$. Эти данные, как и все остальные, увеличены в 10 раз.

Число значащих элементов в каждом столбце табл. 5.15 составляет $m-1=10-1=9$. Поэтому в формуле (5.31) суммирование производится по 9 элементам для всех столбцов таблицы. Например, для первого столбца:

$$v_1^{(1)}=\frac{1}{m-1}\sum_{i=1}^{m-1}\tilde{l}_{i1}=\frac{1}{9}(-1{,}4-1{,}2-\ldots-1{,}2)=$$

$$=\frac{1}{9}(-12{,}2)\approx -1{,}4.$$

5. Значения средней сезонной волны в первом приближении $v_j^{(1)}$ $(j=\overline{1,12})$ (последняя строка табл. 5.15) умножаем на средние квадратичные отклонения каждого года в первом приближении $\sigma_i^{(1)}$ $(i=\overline{1,10})$ (последняя строка табл. 5.14). Полученные величины (табл. 5.16) вычитаем из исходного ряда (табл. 5.11).

В результате в первом приближении получаем временной ряд со значениями уровней:

$$u_{ij}^{(1)}=y_{ij}-v_j^{(1)}\sigma_i^{(1)},$$

лишенный предварительной сезонной волны (табл. 5.17).

Второе приближение.

6. Временной ряд, лишенный предварительной сезонной волны (табл. 5.17), сглаживаем с использованием простой

Таблица 5.15

Нормированный остаточный ряд $10 \cdot \tilde{l}_{ij}^{(1)}$, $10 \cdot \tilde{l}_{j}^{(1)}$ и $10 \cdot v_j^{(1)}$

	Месяц											
Год	1	2	3	4	5	6	7	8	9	10	11	12
1							13	15	10	5	-8	-6
2	-14	-17	-5	1	5	10	13	14	7	9	-1	-5
3	-12	-15	-6	13	10	10	14	14	5	6	-6	-8
4	-19	-10	-1	5	10	11	13	13	9	9	3	-6
5	-16	-19	5	1	11	12	11	11	6	4	-2	-4
6	-13	-17	-3	3	12	14	12	10	6	4	-3	-7
7	-12	-15	-7	4	12	13	9	14	12	6	-1	-1
8	-19	-15	-6	2	9	11	9	11	10	4	-2	-2
9	-17	-15	-3	2	8	11	13	12	6	5	-2	-8
10	-12	-15	-2	3	8	7						
$10\overline{l}_j$	-122	-139	-21	22	84	98	107	114	69	51	-21	-46
$10v_j^{(1)}$	-14	-15	-2	2	9	11	12	13	8	6	-2	-5

скользящей средней с интервалом сглаживания $m = 5$, используя для этого формулу, подобную (5.10):

$$u_{ij}^{(2)} = \frac{1}{m} \sum_{i=t-p}^{t+p} u_{ij}^{(1)},$$

где $p = (m-1)/2 = (5-1)/2 = 2$.

Чтобы использовать пятиинтервальное сглаживание, необходимо, чтобы справа и слева от рассматриваемой точки оставалось по $p = 2$ элемента. Первый сглаженный элемент нового ряда должен находиться на первой строке в девятом столбце табл. 5.17, так как упомянутая строка уже начинается с седьмого элемента первой строки.

Для первого рассматриваемого элемента имеем:

$$u_{19}^{(2)} = \frac{1}{5} \sum_{j=9-2}^{9+2} \overline{y}_{1j}^{(1)} = \frac{1}{5}(87 + 86 + 92 + 95 + 104) = 93.$$

Таблица 5.16

Значения сезонной волны $v_j^{(1)}\sigma_i^{(1)}$

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1							13	14	9	7	-2	-6
2	-18	-20	-3	3	12	14	16	17	10	8	-3	-7
3	-22	-24	-3	3	14	18	19	21	13	10	-3	-8
4	-22	-24	-3	3	14	18	19	21	13	10	-3	-8
5	-29	-30	-5	5	19	22	24	25	15	11	-5	-10
6	-22	-24	-3	3	14	18	19	21	13	10	-3	-8
7	-20	-21	-3	3	13	15	17	18	11	8	-3	-7
8	-25	-27	-4	4	16	20	22	23	14	11	-4	-9
9	-25	-27	-4	4	16	20	22	23	14	11	-4	-9
10	-28	-30	-4	4	18	22						

Таблица 5.17

Ряд без сезонной волны в первом приближении $u_{ij}^{(1)}$

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1							87	86	92	95	104	109
2	121	124	108	103	94	93	91	91	99	102	114	119
3	135	138	118	112	102	97	96	94	103	106	120	125
4	139	141	121	115	104	101	100	97	104	106	119	125
5	145	128	123	115	101	98	98	97	109	114	131	138
6	151	153	133	127	117	113	114	113	121	125	138	144
7	156	157	140	135	125	124	122	121	128	132	143	148
8	167	170	147	140	129	126	125	125	135	139	154	160
9	177	180	157	149	138	134	132	131	140	144	159	165
10	184	186	160	153	139	136						

Аналогично вычисляются все остальные значения нового ряда во втором приближении. Значения уровней этого ряда приведены в табл. 5.18.

Таблица 5.18

Оценка тренда $u_{ij}^{(2)}$ во втором приближении

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
1									93	97	104	111
2	113	113	111	105	98	95	94	95	100	105	114	122
3	123	123	119	112	103	98	97	97	102	108	117	124
4	127	126	122	115	106	102	100	100	103	108	118	127
5	130	129	125	115	105	100	101	103	110	118	129	138
6	141	141	136	129	121	117	116	117	122	128	137	144
7	147	147	143	136	129	126	124	126	129	134	144	152
8	155	155	151	143	139	129	128	130	136	143	153	162
9	166	165	160	152	142	137	135	136	141	148	159	168
10	171	169	164	155								

7. Вычисляем во втором приближении отклонения $l_{ij}^{(2)}$ тренда $u_{ij}^{(2)}$ (табл. 5.18) от исходного ряда y_{ij} (табл. 5.11).

В табл. 5.19 полученных результатов отсутствуют строки, соответствующие первому и десятому годам, так как в процессе сглаживания в этих строках осталось недостаточное для дальнейшего анализа количество элементов (всего по 4 элемента из 12).

8. Используя данные табл. 5.19, по аналогии с п. **3** вычисляем средние по каждому году значения $\sigma_i^{(2)2}$ и $\sigma_i^{(2)}$ во втором приближении. Полученные значения приведены в табл. 5.20.

9. Аналогично п. **3** составляем табл. 5.21 нормированного остаточного ряда, умножая все ее значения для удобства на 10.

На рис. 5.2 изображены выровненные значения сезонной волны: точками обозначены табличные значения; жирными линиями — сглаживающая функция; тонкой линией — аппроксимирующая парабола. Координаты точек получены следующим образом. Через точки 2–12 последней строки

табл. 5.21 проведена парабола $y = at^2 + bt + c$. Коэффициенты a, b и c определялись по методу наименьших квадратов с использованием данных для двух лет: второго и девятого (табл. 5.21). Эти данные позволили определить коэффициенты нормальной системы уравнений (4.9):

$$\begin{cases} a\sum\limits_{j=1}^{k} t_j^4 + b\sum\limits_{j=1}^{k} t_j^3 + c\sum\limits_{j=1}^{k} t_j^2 = \sum\limits_{j=1}^{k} t_j^2 v_j, \\ a\sum\limits_{j=1}^{k} t_j^3 + b\sum\limits_{j=1}^{k} t_j^2 + c\sum\limits_{j=1}^{k} t_j = \sum\limits_{j=1}^{k} t_j v_j, \\ a\sum\limits_{j=1}^{k} t_j^2 + b\sum\limits_{j=1}^{k} t_j + ck = \sum\limits_{j=1}^{k} v_j \end{cases}$$

и привести ее к виду

$$\begin{cases} 60709a + 6083b + 649c = 2467, \\ 6083a + 649b + 77c = 355, \\ 649a + 77b + 11c = 43. \end{cases}$$

Таблица 5.19

Отклонение $l_{ij}^{(2)}$ ряда во втором приближении

Год	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
2	-28	-32	-12	2	14	25	30	32	18	17	-4	-16
3	-29	-33	-5	5	28	39	40	40	22	19	-10	-22
4	-39	-34	-6	10	27	34	38	38	27	22	2	-19
5	-38	-39	-16	5	35	44	43	42	26	16	-5	-17
6	-33	-39	-10	6	29	36	36	33	23	13	-6	-20
7	-28	-32	-15	7	26	30	28	32	16	13	-6	-12
8	-47	-39	-18	5	22	37	36	37	31	14	-6	-14
9	-44	-40	-12	5	27	36	42	40	23	16	-7	-16

Таблица 5.20
Характеристики ряда во втором приближении

t	2	3	4	5	6	7	8	9
$\sigma_i^{(2)2}$	492	712	688	981	729	526	862	884
$\sigma_i^{(2)}$	22	27	26	31	27	23	29	30

Решение системы уравнений позволило найти коэффициенты $a = -1{,}1$, $b = 15{,}98$ и $c = -43{,}5$. То есть сезонная волна может быть описана уравнением:

$$\tilde{v} = -1{,}1t^2 + 15{,}98t - 43{,}5 \qquad (t = \overline{3;11}).$$

Первая точка ($t = 1$) не использовалась в аппроксимации. Кроме того, из рис. 5.2 видно, что значения $\tilde{v}$, соответствующие точкам $t = 2$ и $t = 12$, значительно отличаются от соответствующих точек на параболе и искажают суть процесса. Поэтому в качестве сезонной волны принята выделенная жирным линия, по которой между точками первой и второй, второй и третьей, 11 и 12 существует линейная зависимость, между остальными точками — параболическая.

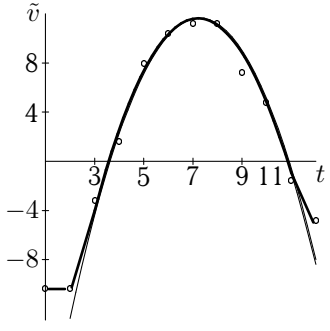


Рис. 5.2
Сезонная волна

Таблица 5.21
Нормированный остаточный ряд
$10\tilde{l}_{ij}^{(2)}$, $10\tilde{l}_{j}^{(2)}$, $10v_{j}^{(2)}$ и $10\tilde{v}_{j}$

	Месяц											
Год	1	2	3	4	5	6	7	8	9	10	11	12
2	-13	-15	-6	1	6	11	14	15	8	8	-2	-7
3	-11	-12	-2	2	10	12	15	15	8	7	-4	-8
4	-15	-9	-2	4	10	13	14	14	10	8	1	-7
5	-12	-16	-5	2	11	14	14	13	8	5	-2	-5
6	-12	-14	-4	2	11	13	13	12	9	5	-2	-7
7	-12	-14	-7	3	11	13	12	14	7	6	-3	-5
8	-16	-13	-6	2	8	13	12	13	11	5	-2	-5
9	-15	-13	-4	2	9	12	14	13	8	5	-2	-5
$10\bar{l}_{j}$	-105	-106	-35	16	77	101	108	109	68	49	-15	-51
$10v_{j}^{(2)}$	-13	-13	-4	2	10	13	14	14	9	6	-2	-6
$10\tilde{v}_{j}$	-13	-13	-1	5	9	12	13	12	10	6	0	-6

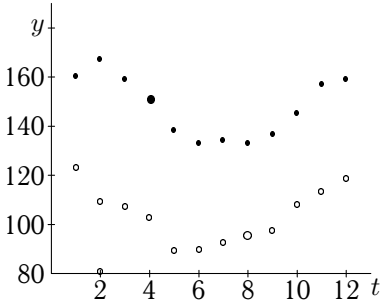


Рис. 5.3
Тренд

10. Вычисляем значения уровней тренда во втором приближении по аналогии с п. **5**, предварительно составляя табл. 5.22. Подчеркнем, что в этой таблице величины $v_{j}^{(2)}$ берутся из предпоследней строки табл. 5.21. В отличие от последней строки этой таблицы они содержат случайную составляющую.

В табл. 5.23 приведены значения уровней тренда в последнем приближении, а на рис. 5.3 показаны точки, соответствующие этим уровням для второго (○) и девятого (●) годов.

Если значения уровней сезонной волны в первом (табл. 5.16) и втором (табл. 5.22) приближениях достаточно близки, то дальнейших приближений для уточнения значений этих и других величин делать нецелесообразно. Хотя в указанных таблицах наблюдаются расхождения, результаты следующих приближений здесь не приведены, так как дальнейшие расчеты приводят к результатам, практически совпадающим с результатами второго приближения.

Отметим попутно, что в подобных задачах вычисление значений различных величин с бóльшей, чем принятая в примере, точностью лишено смысла, ибо сами исходные данные имеют разброс, выходящий за пределы этой точности.

Таблица 5.22

Значения сезонной волны $v_j^{(2)}\sigma_i^{(2)}$

	1	2	3	4	5	6	7	8	9	10	11	12
2	-29	-29	-9	4	22	29	31	31	20	13	-4	-13
3	-35	-35	-11	5	27	35	38	38	24	16	-5	-16
4	-34	-34	-10	5	26	34	36	36	23	16	-5	-16
5	-40	-40	-12	6	31	40	43	43	28	19	-6	-19
6	-35	-35	-11	5	27	35	38	38	24	16	-5	-16
7	-30	-30	-9	5	23	30	32	32	21	14	-5	-14
8	-38	-38	-12	6	29	38	41	41	26	17	-6	-17
9	-39	-39	-12	6	30	39	42	42	27	18	-6	-18

11. После завершения итерационного процесса по определению тренда можно выделить из временного ряда случайную компоненту (5.34):

$$\varepsilon_{ij} = l_{ij}^{(2)} - v_j^{(2)}\sigma_i^{(2)}.$$

Значения случайной компоненты приведены в табл. 5.24.

Таблица 5.23

Ряд без сезонной волны $u_{ij}^{(3)}$ (тренд)

	Месяц											
Год	1	2	3	4	5	6	7	8	9	10	11	12
2	124	110	108	103	90	91	93	96	98	109	114	119
3	133	125	125	112	104	96	99	99	100	111	112	121
4	122	136	126	120	107	102	103	102	113	114	125	124
5	132	120	119	115	114	104	101	102	108	115	130	140
6	143	137	137	130	123	118	114	112	121	125	136	140
7	149	145	137	138	132	126	120	126	134	134	144	154
8	146	154	145	142	132	128	123	126	141	140	153	165
9	161	164	160	151	139	134	135	134	137	146	158	160

Таблица 5.24

Случайная составляющая ряда $10\varepsilon_{ij}$

	Месяц											
Год	1	2	3	4	5	6	7	8	9	10	11	12
2	6	-34	1	-42	-56	-32	-19	14	-66	31	-22	-28
3	61	21	99	-26	80	44	8	25	-82	20	-78	-58
4	-58	-2	79	27	39	7	3	19	-21	56	41	-34
5	23	13	29	-37	74	40	-20	-11	-87	-35	-25	16
6	21	49	49	-16	50	14	-32	-45	-72	-4	-38	-38
7	19	-21	-41	6	55	6	-54	0	-98	-15	-42	18
8	-93	-13	-11	-31	-38	-1	-61	-33	-15	-43	-37	34
9	-50	-10	59	-34	3	14	-15	-17	-106	-29	-46	20

§ 5.9. РЕЗЮМЕ

В реальных задачах зачастую представляет интерес зависимость статистических данных (случайной величины Y), от времени (временные или динамические ряды).

Регрессионные модели, построенные на базе таких рядов, позволяют проводить анализ процессов. Для его осуществления из временного ряда необходимо выделить составляющие, ответственные за:

– тенденцию U_t изменения Y (тренд);

– сезонные изменения V_t, связанные с изменением времен года;

– циклические изменения C_t, которые могут зависеть от поставок партий товаров, сроками получения зарплаты, . . . ;

– случайные отклонения от закономерностей ε_t.

В результате исходный ряд разбивается на сумму временных рядов:

$$Y = U_t + V_t + C_t + \varepsilon_t.$$

Для выделения слагаемых представленного ряда необходимо предварительно проанализировать исходный статистический материал, избавить его от аномальных, неправдоподобных данных и вкравшихся в исходные данные ошибок. Для этого существуют основанные на статистических критериях (Стьюдента, Фишера,. . .) методы предварительного анализа временных рядов (Ирвина, Фостера–Стьюарта, проверки разностей средних уровней интервалов) (§ 5.2).

На втором этапе обработки статистического материала осуществляется сглаживание временных рядов, цель которого — избавиться от случайных выбросов в статистических исходных данных, не соответствующих закономерному поведению исследуемого процесса. Для сглаживания временных рядов используются методы простой и взвешенной скользящей средней, экспоненциального сглаживания и другие (раздел 5.3).

После определения необходимых параметров обработанного статического ряда (средних арифметических ряда и его уровней, показателей прироста. . .) и выделения указанных выше составляющих проводится анализ статистических данных.

§ 5.10. ВОПРОСЫ

1. Что представляет собой динамический ряд (д. р.)? Как измеряется его длина?
2. Перечислите и охарактеризуйте основные составляющие д. р.

3. Каким требованиям должны удовлетворять: тренд, сезонная, периодическая и случайная составляющие д. р.?
4. Охарактеризуйте ошибки первого и второго родов, приводящие к аномальным уровням д. р.
5. Опишите основные методы выявления аномальных уровней д. р.: Ирвина, проверки разностей средних уровней, Фостера–Стюарта.
6. Перечислите основные методы сглаживания д. р. Перечислите последовательность действий, предусмотренных этими методами.

§ 5.11. ЗАДАНИЕ НА РАСЧЕТНУЮ РАБОТУ, ЧАСТЬ 7

Задана таблица данных для составления исходного временного ряда y_{ij}^0 (i — номер строки, j — номер столбца) (табл. 5.25).

Таблица 5.25
Таблица данных

Год	Месяц									
	1	2	3	4	5	6	7	8	9	10
1	7	6	9	10	11	12	14	12	11	10
2	8	8	10	11	11	13	11	13	12	11
3	9	9	11	12	13	13	14	14	12	13
4	9	10	12	13	13	14	15	13	13	12
5	9	8	11	12	14	14	14	15	14	13
6	11	10	13	13	15	15	16	15	14	14
7	12	12	13	14	16	16	15	16	16	15
8	11	11	13	15	16	17	16	17	17	16
9	12	13	15	16	17	17	18	18	16	16
10	13	13	15	16	17	17	18	18	17	17

Требуется.

Составить на основе приведенной таблицы новую таблицу, элементы которой определить из соотношения

$$y_{ij} = (1 + k)y_{ji}^{0} + k,$$

где $k = 10N/n$ (N — последняя цифра номера учебной группы студента, n — порядковый номер студента в списке группы).

Полученную таблицу считать таблицей исходных данных и использовать их в нижеследующих пунктах.

1. Провести начальное выравнивания ряда методом центрированной скользящей средней (5.28), принимая период сглаживания $T_0 = 12$.
2. Получить остаточный ряд и предварительную сезонную волну.
3. Получить оценку тренда.
4. Повторить пункты 1–3 до достижения необходимой точности.
5. Получить окончательные значения тренда, сезонной волны и случайной компоненты.

Построить графики с нанесением на них точек, соответствующих табличным значениям:

— изменения тренда (для двух характерных лет);

— сезонной волны.

Глава 6

ПРОГНОЗИРОВАНИЕ

§ 6.1. ТРЕНДОВЫЕ МОДЕЛИ ОДНОМЕРНЫХ ПРОЦЕССОВ

Рассмотрим одномерные процессы, т. е. процессы, в которых рассматривается изменение одного их показателя во времени. Предполагается, что на изменение этого показателя влияют различные случайные внешние воздействия, которые затрудняют создание соответствующей однозначной детерминированной математической модели процесса. Описание таких процессов требует привлечения статистических методов.

Вопросы обработки данных, имеющих случайный характер, в том числе сглаживания временных рядов, их фильтрация, выделения тренда и т. д., рассматривались в предыдущей главе. В этой главе будет предполагаться, что из общего временного ряда выделен тренд и требуется аппроксимировать его некоторой функцией (ввести *кривую роста*), с помощью которой можно будет экстраполировать зависимость рассматриваемого показателя ряда от времени на несколько временных отрезков вперед, т. е. произвести прогноз изменения процесса.

Для аппроксимации тренда наибольшее распространение нашли полиномиальные, экспоненциальные и S-образные кривые роста. Остановимся на них более подробно.

Полиномиальные кривые роста описываются полиномом n-ного порядка:

$$y = a_o + a_1 t + a_2 t^2 + \ldots + a_n t^n, \qquad (6.1)$$

где a_i $(i = \overline{1,n})$ — постоянные коэффициенты, которые подбираются из условия наилучшего описания функцией $y(t)$ рассматриваемого процесса. Коэффициент a_1 называют *линейным приростом*; a_2 — *ускорением роста*; a_3 — *изменением ускорения роста* и т. д.

При $n = 0$ аппроксимирующая функция постоянна; при $n = 1$ — линейно зависит от времени; при $n = 2$ — представляет собой квадратную параболу и т. д.

Полином первой степени описывает процессы с постоянным законом роста. Приросты между всеми соседними уровнями показателя рассматриваемого процесса на одинаковых приращениях времени $\Delta t = 1$ одинаковы, т. е. $u_t = y_t - y_{t-1} = a_1 \Delta t = a_1$ для всех $t = \overline{2,n}$. Величины u_t называют *первыми приростами* функции $y(t)$.

Для полинома второй степени первые приросты будут иметь линейную зависимость от времени. Ряд из первых приростов (u_2, u_3, ...) на графике представляется прямой. *Вторые приросты,* которые являются первыми приростами первых приростов, т. е. $u_t^{(2)} = u_t - u_{t-1}$, для полинома второй степени равны постоянной и т. д.

Для полинома третьей степени первые приросты будут полиномами второй степени, вторые приросты — полиномами первой степени, а третьи приросты $u_t^{(3)} = u_t^{(2)} - u_{t-i}^{(2)} = a_3$ будут постоянными величинами.

Существует связь между приростами и производными функции. Производная полинома первой степени:

$$y' = \frac{dy}{dt} \approx \frac{\Delta y}{\Delta t} = \frac{y_t - y_{t-1}}{\Delta t} = \frac{u_t}{\Delta t}.$$

Отсюда, так как $y' = a_1$, получим при $\Delta t = 1$:

$$u_t = y' \Delta t = a_1.$$

Отметим важную особенность кривых роста, описываемых полиномами: *значения приростов для полиномов не зависит от значений самой функции.* Это свойство позволяет утверждать, что прогноз процессов, описываемых полиномами, не зависит от достигнутого уровня рассматриваемого показателя.

Экспоненциальные кривые роста могут описываться *простой показательной функцией*:

$$y = ka^t, \tag{6.2}$$

или *модифицированной показательной функцией*:

$$y = ka^t + b. \tag{6.3}$$

Здесь a, b, k — постоянные. В качестве a может использоваться основание натурального логарифма — число e.

В зависимости (6.2) постоянные a и k положительны. Причем, если $a < 1$ функция y убывает, при $a > 1$ возрастает во времени.

Ордината y (6.2) имеет постоянный темп прироста, т. е. для любого t отношение первого прироста к функции есть величина постоянная:

$$\frac{u_t}{y_t} = \frac{y_t - y_{t-1}}{y_t} = 1 - \frac{1}{a}.$$

Прологарифмируем (6.2) по любому основанию, например, по e :

$$\ln y = \ln k + t \ln a.$$

Полученное соотношение говорит о том, что логарифмы ординат простой экспоненты линейно зависят от времени.

В зависимости (6.3) $0 < a < 1$. Слагаемое $b > 0$ называется асимптотой функции, так как при росте t кривая (6.3) неограниченно (асимптотически) приближается к прямой $y = b$.

Первый прирост функции (6.3):

$$u_t = y_t - y_{t-1} = k(1 - a^{-1})a^t$$

зависит от времени, как простая экспонента. Следовательно, логарифмы первых приростов функции будут линейными функциями от времени (см. итог логарифмирования простой экспоненты).

Если взять отношение двух приростов рассматриваемой функции, то оно будет постоянной величиной. Действительно,

$$\frac{u_t}{u_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1} - y_{t-2}} = a.$$

Кривая Гомперца:

$$y = ka^{b^t}. \tag{6.4}$$

Здесь, как и в экспоненциальных функциях, $a > 0$, $0 < b < 1$, $k > 0$ — постоянные коэффициенты. Прямые $y = k$ и $y = 0$ являются горизонтальными асимптотами кривых, описываемых уравнением (6.4). Одна из таких кривых, а именно $y = 0{,}5^{\left(\sqrt{2}\right)^{-t}}$ изображена на рис. 6.1. Эта кривая по характеру поведения напоминает функцию распределения для нормального закона и часто называется *s-образной функцией*. Такие функции описывают многие процессы, в которых показатели растут сначала медленно, затем рост их ускоряется, а за ускорением следует замедление и стремление к некоторому пределу.

Логарифм функции Гомперца (например, по основанию a):

$$\log_a y = \log_a k + b^t$$

является модифицированной экспонентой (6.3). Поэтому отношения первых последовательных приростов логарифма этой функции есть величина постоянная.

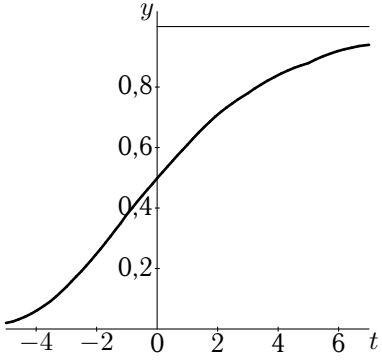


Рис. 6.1
Кривая Гомперца

Логистическая кривая, или *кривая Перла-Рида,* представляется в виде

$$y = \frac{k}{1 + ae^{-bt}}. \tag{6.5}$$

Существуют и другие виды логистических кривых, по виду напоминающие (6.5).

Если взять производную от логистической функции, то можно заметить, что скорость возрастания кривой в любой момент времени пропорциональна достигнутому уровню y и разности между предельным значением k функции и достигнутым уровнем. Логарифм отношения первого прироста функции к квадрату ее значения является линейной функцией времени.

График логистической кривой напоминает график функции Гомперца (рис. 6.1), но в отличие от последнего имеет точку симметрии, совпадающую с точкой перегиба.

§ 6.2. ВЫБОР АППРОКСИМИРУЮЩЕЙ КРИВОЙ

Пусть задан временной ряд $Y = \{y_1; y_2; \ldots; y_t; \ldots; y_n\}$. Требуется выбрать такую аппроксимирующую функцию (кривую роста), которая наилучшим образом описывает процесс, представленный временным рядом Y. Выбор можно осуществить, используя какой либо из ниже описанных методов.

6.2.1. Метод конечных разностей

Предположим, что наш выбор остановился на кривой, описываемой полиномиальной функцией. Наиболее распространенным методом определения коэффициентов аппроксимирующего полинома является *метод конечных разностей*, часто в экономической литературе называемый методом Тинтнера. Этот метод может быть использован для выбора кривой при двух условиях:

1) уровни заданного ряда состоят только из двух составляющих, а именно, тренда и случайной составляющей;
2) тренд является достаточно гладким.

Второе условие выдвигается при аппроксимации ряда любой (не только полиномиальной) дифференцируемой и непрерывной функцией.

На первом этапе использования метода конечных разностей вычисляются разности (приросты) до k-го порядка:

$$\begin{cases} u_t & = & y_t - y_{t-1}; \\ u_t^{(2)} & = & u_t - u_{t-1}; \\ \cdot \quad \cdot & & \cdot \quad \cdot \quad \cdot \quad \cdot \\ u_t^{(k)} & = & u_t^{(k-1)} - u_{t-1}^{(k-1)}. \end{cases}$$

После определения разностей, которые для практических задач обычно не превышают 3–4 порядков, определяют дисперсии исходного ряда:

$$\sigma_0^2 = \frac{1}{n-1}\left(\sum_{t=1}^{n} y_t^2 - \frac{1}{n}\left(\sum_{t=1}^{n} y_t\right)^2\right)$$

и каждого k-го разностного ряда:

$$\sigma_k^2 = \frac{\left(\sum\limits_{t=k+1}^{n} (u_t)^2\right)^k}{(n-k)C_{2k}^k},$$

где $C_{2k}^k = \dfrac{(2k)!}{(k!)^2}$ — число сочетаний из $2k$ элементов по k.

Далее сравниваются отклонения каждой последующей дисперсии от предыдущей $\left|\sigma_k^2 - \sigma_{k-1}^2\right|$.

Если для некоторого k отклонение дисперсии от предыдущей достаточно мало, т. е. значения двух последующих дисперсий оказались достаточно близкими, то степень аппроксимирующего полинома можно ограничить величиной $k-1$.

6.2.2. Метод характеристик прироста

Более универсальным, чем метод конечных разностей, является *метод характеристик прироста*.

При использовании этого метода исходный временной ряд Y предварительно сглаживается методом простой скользящей

средней. Например, для интервала сглаживания $m = 3$ сглаженные значения уровней ряда определяются как средние арифметические трех граничащих значений Y:

$$\overline{y}_t = \frac{1}{3}(y_{t-1} + y_t + y_{t+1}).$$

Чтобы не потерять первый и последний члены ряда, их можно сгладить по формулам:

$$\overline{y}_1 = \frac{1}{6}(5y_1 + 2y_2 - y_3);$$

$$\overline{y}_n = \frac{1}{6}(-y_{n-2} + 2y_{n-1} + 5y_n).$$

После этого вычисляются первые средние приросты сглаженного ряда:

$$\overline{u}_t = \frac{1}{2}(\overline{y}_{t+1} - \overline{y}_{t-1}), \qquad (t = \overline{2, n-1});$$

вторые средние приросты:

$$\overline{u}_t^{(2)} = \frac{1}{2}(\overline{u}_{t+1} - \overline{u}_{t-1})$$

и ряд величин, связанных с вычисленными сглаженными значениями уровней и средними приростами:

$$\frac{\overline{u}_t}{\overline{y}_t}, \quad \log \overline{u}_t, \quad \log \overline{u}_t \overline{y}_t, \quad \log \overline{u}_t \overline{y}_t^2.$$

По полученным величинам с использованием данных табл. 6.1 выбирается подходящий вид кривой роста, аппроксимирующей заданный ряд.

На практике обычно выбирают несколько подходящих кривых для дальнейшего исследования процессов.

После выбора типа кривой следует решить вопрос об определении параметров, которые полностью определяют аппроксимирующую функцию.

Таблица 6.1
Таблица для выбора вида кривых роста

Показатель	*Характер уровней*	*Вид кривой*
Первый средний прирост $\overline{u}_t$	Примерно одинаковы	Полином 1 порядка
То же	Изменяются линейно	Полином 2 порядка
Второй средний прирост $\overline{u}_t^{(2)}$	Изменяются линейно	Полином 3 порядка
$\overline{u}_t/\overline{y}_t$	Примерно одинаковы	Простая экспонента
$\log \overline{u}_t$	Изменяются линейно	Модифицированная экспонента
$\log(\overline{u}_t/\overline{y}_t)$	Изменяются линейно	Кривая Гомперца
$\log(\overline{u}_t/\overline{y}_t^2)$	Изменяются линейно	Логистическая кривая

В § 4.5 рассмотрен метод наименьших квадратов, который позволяет после решения системы нормальных уравнений определить коэффициенты полиномов, аппроксимирующих исследуемые процессы.

Параметры экспоненциальных и S-образных кривых находятся более сложными методами. Так выражение для простой экспоненты

$$y_t = ka^t$$

предварительно логарифмируют по любому основанию, получая линейную относительно переменной t зависимость:

$$\log y_t = \log k + t \log a.$$

Для определения логарифмов неизвестных параметров $\log k$ и $\log a$ составляют систему нормальных уравнений мето-

да наименьших квадратов. Решение этой системы позволяет определить логарифмы параметров, а затем и сами параметры.

При определении параметров кривых роста, имеющих асимптоты (кривая Гомперца, логистическая кривая), используют различные приближенные методы. Для частного случая, если значение k в выражениях (6.4)–(6.5) известно, то путем простых преобразований, связанных с изменением масштаба координат и перемещением их начала, выражения для кривых можно привести к экспоненциальному виду и затем провести логарифмирование, как отмечено выше.

При неизвестных значениях k используются другие приближенные методы для определения k и других параметров процесса.

§ 6.3. АДЕКВАТНОСТЬ ТРЕНДОВЫХ МОДЕЛЕЙ

Прежде, чем использовать построенную математическую модель реального процесса (аппроксимирующую функцию), необходимо убедиться в ее адекватности (соответствии модели реальному процессу). Конечно, о полной адекватности речь не идет (нельзя обеспечить полного соответствия модели реальному процессу), но по основным параметрам, наиболее существенным для процесса, адекватность должна быть обеспечена.

Трендовая модель $\tilde{Y} = \{\tilde{y_1}; \tilde{y_2}; \ldots; \tilde{y_n}\}$ временного ряда $Y = \{y_1; y_2; \ldots; y_n\}$ считается адекватной, если она правильно отражает основные свойства исходного временного ряда. Требование адекватности модели связывается с требованием того, чтобы остаточные компоненты $\varepsilon_t = y_t - \tilde{y}_t$ $(t = \overline{1,n})$ удовлетворяли свойствам случайной величины (§ 5.1). Для них должны выполняться условия:

— случайность характера уровней остаточной последовательности;

— соответствие распределения нормальному закону;

— равенство нулю математического ожидания;

— независимость друг от друга значений уровней.

Ниже показано, как можно проверить указанные свойства остаточной последовательности.

6.3.1. Проверка случайности характера уровней последовательности

Эта проверка должна подтвердить гипотезу о правильном выборе вида тренда. Для проверки используется совокупность разностей:

$$\varepsilon_t = y_t - \tilde{y}_t.$$

Характер разностей исследуется с помощью ряда критериев. Рассмотрим для примера один из них — *критерий серий*. Для использования этого критерия ряд из величин ε_t располагают в порядке возрастания их значений и находят медиану записанного ряда. *Медианой* ε_m называют срединное значение (значение, удаленное на одинаковое количество точек от крайних точек ряда) при нечетном n или среднее арифметическое пар срединных значений при четном n.

Сравнивается значения исходной последовательности ε_t с ε_m. При этом ставится знак «плюс», если $\varepsilon_t > \varepsilon_m$, и знак «минус» в противном случае. При равенстве соответствующее значение ε_t исключается из дальнейшего рассмотрения.

Последовательность подряд идущих плюсов (или минусов) называется *серией*. Чтобы последовательность ε_t была случайной, протяженность самой большой серии L_{max} не должна быть слишком большой, а количество серий ν слишком малым. Оценку этих величин можно осуществить следующим образом.

Последовательность признается случайной, если для 5%-го уровня значимости выполняются неравенства:

$$L_{max} < [3{,}3(\lg n + 1)];$$

$$\nu > \left[\frac{1}{2}\left(n + 1 - 1{,}96\sqrt{n-1}\right)\right]. \qquad (6.6)$$

Квадратные скобки обозначают целую часть числа.

Если хотя бы одно из неравенств (6.6) не выполняется, гипотеза о случайном характере отклонений уровней временного ряда от тренда отвергается и трендовая модель признается неадекватной.

6.3.2. Проверка соответствия распределения случайной компоненты нормальному закону

Указанная проверка осуществляется посредством показателей *асимметрии* α и *эксцесса* e:

$$\alpha = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^3}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^3}}; \qquad e = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^4}{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^2} - 3, \tag{6.7}$$

которые вместе со средними квадратичными отклонениями:

$$\sigma_\alpha = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}}; \qquad \sigma_e = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}} \tag{6.8}$$

входят в определяющие неравенства:

$$|\alpha| < 1{,}5\sigma_\alpha; \qquad \left|e + \frac{6}{n+1}\right| < 1{,}5\sigma_e. \tag{6.9}$$

Если оба строгих неравенства (6.9) выполняются, то гипотеза о нормальном характере распределении случайной компоненты принимается. В противном случае гипотеза отвергается и трендовая модель признается неадекватной.

При нормальном распределении показатели α и e рассматриваемой совокупности равны нулю.

6.3.3. Проверка равенства нулю математического ожидания случайной компоненты

Если с использованием неравенств (6.9) установлено, что случайная компонента распределена по нормальному закону, то следующим этапом определения адекватности трендовой модели является проверка равенства нулю математического ожидания этой случайной величины. Проверка осуществляется с использованием t-критерия Стьюдента:

$$t = \frac{\overline{\varepsilon} - 0}{\sigma_\varepsilon}\sqrt{n}, \tag{6.10}$$

где $\overline{\varepsilon}$ — среднее арифметическое значение остаточной последовательности $\{\varepsilon_t\}$; σ_ε — среднее квадратичное отклонение для этой последовательности.

Если расчетные значения критерия t меньше табличного значения T_{kp} критерия Стьюдента с заданным уровнем значимости α и числом степеней свободы $k = n-1$ (Приложение 3), то гипотеза о равенстве нулю математического ожидания остаточной последовательности принимается. В противном случае гипотеза отвергается и модель считается неадекватной.

6.3.4. Проверка независимости значений уровней случайной компоненты

Проверяется наличие или отсутствие существенной автокорреляции в остаточной последовательности. Наиболее распространенным критерием, по которому можно судить об этом, является d-критерий Дарбина–Уотсона:

$$d = \frac{\sum\limits_{t=2}^{n} \left(\varepsilon_t - \varepsilon_{t-1}\right)^2}{\sum\limits_{t=2}^{n} \varepsilon_t^2}. \qquad (6.11)$$

Отметим, что расчетное значение d в интервале $[2;4]$ неверно свидетельствует об отсутствии связи и вместо d в качестве расчетного критерия берется величина:

$$d' = 4 - d.$$

Расчетное значение критерия d (d') сравнивается с верхним $d_\text{В}$ и нижним $d_\text{Н}$ критическими значениями статистики Дарбина–Уотсона. Фрагмент таблицы этих значений для уровня значимости 5%, различного количества уровней n ряда и числа k определяемых параметров представлен в табл. Приложения 7.

Если $d > d_\text{В}$, то гипотеза о независимости уровней остаточной последовательности (об отсутствии автокорреляции) принимается. При $d < d_\text{Н}$ гипотеза отвергается, и трендовая модель считается неадекватной.

В случае $d_{\text{Н}} \leq d \leq d_{\text{В}}$ нет оснований ни для принятия гипотезы, ни для того, чтобы ее отвергнуть, и необходимы дополнительные исследования, например, по большему числу наблюдений.

Вывод об адекватности трендовой модели делается только в случае положительного результата *всех четырех условий* проверки свойств остаточной последовательности, изложенных в разделах 6.3.1–6.3.4.

§ 6.4. ТОЧНОСТЬ ТРЕНДОВЫХ МОДЕЛЕЙ

Точность трендовой модели можно определить по различным характеристикам, связанным с отклонением

$$\varepsilon_t = y_t - \tilde{y}_t$$

уровней исходного временного ряда y_t от значений $\tilde{y}_t$, полученных в результате аппроксимации тренда.

В качестве статистических оценок точности принимаются следующие характеристики.

Среднее квадратичное отклонение (k — число определяемых отклонение независимых параметров модели):

$$\sigma_\varepsilon = \sqrt{\frac{1}{n-k}\sum_{t=1}^{n}\varepsilon^2}. \tag{6.12}$$

Средняя относительная ошибка аппроксимации:

$$\overline{\varepsilon} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\varepsilon}{y_t}\right|. \tag{6.13}$$

Величины, подсчитанные по приведенным формулам, определяют в процентах относительную ошибку, обусловленную заменой исходного временного ряда трендовой моделью.

Значения перечисленных и других используемых для оценки точности временных рядов показателей позволяют выбрать

из нескольких адекватных трендовых моделей те, которые обладают большей точностью. Хотя могут иметь место случаи, когда по одному из показателей точности (6.12)–(6.13) оказывается более точной одна модель, по другому — другая.

§ 6.5. ПРИМЕР ИССЛЕДОВАНИЯ ТРЕНДОВОЙ МОДЕЛИ НА АДЕКВАТНОСТЬ И ТОЧНОСТЬ

Рассмотрим временной ряд, заданный в виде табл. 6.2.

Требуется:

1) подобрать функцию, аппроксимирующую этот ряд (тренд);

2) убедиться в адекватности трендовой модели;

3) оценить ее точность.

При решении поставленных задач будем для удобства некоторые промежуточные данные вычислений заносить в табл. 6.3. В последней строке этой таблицы приведены суммы занесенных в столбцы соответствующих величин.

Таблица 6.2

Исходный временной ряд

t	1	2	3	4	5	6	7	8	9	10	11
y_t	86	81	79	72	68	70	65	62	57	55	53

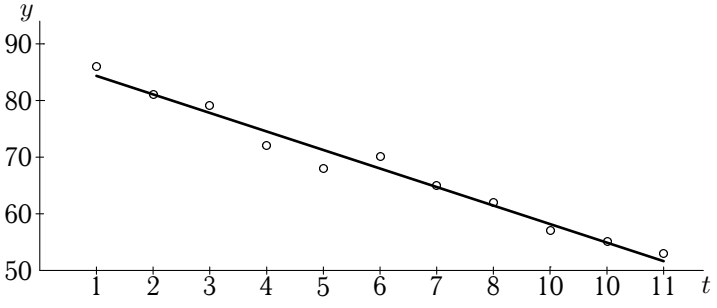


Рис. 6.2

Исходные данные и модель тренда

Таблица 6.3

Промежуточные данные для расчета

t	y	ty	t^2	$\tilde{y}$	ε	ε^2	ε^3	ε^4	$\varepsilon_t-\varepsilon_{t-1}$	$100\frac{\varepsilon}{y}$
1	86	86	1	84,2	1,8	3,2	5,8	10,5	—	2,1
2	81	162	4	81,0	0	—	—	—	-1,8	0
3	79	237	9	78,0	1	1	1	1	1	1,3
4	72	288	16	74,5	-2,5	6,3	-15,6	39,1	-3,5	3,5
5	68	340	25	71,0	-3	9	-27	81,0	-0,5	4,4
6	70	420	36	68,0	2	4	8	16,0	5	2,7
7	65	455	49	65,0	0	—	—	—	-2	0
8	62	486	64	61,5	0,5	0,3	0,2	0,1	0,5	0,8
9	57	513	81	58,0	-1	1	-1	1	-1,5	1,8
10	55	550	100	55,0	0	—	—	—	1	1,8
11	53	583	121	51,7	1,3	1,7	2,2	2,9	1,3	2,4
66	748	4130	506		0,1	26,5	-26,4	151		21

1. По характеру расположения точек на рис. 6.2 принимаем в качестве аппроксимирующей функции полином первой степени, т. е. прямую, задаваемую уравнением:

$$\tilde{y} = a_0 + a_1 t. \tag{6.14}$$

Коэффициенты этого уравнения определим по методу наименьших квадратов из системы нормальных уравнений (4.6):

$$\begin{cases} a_1 \sum\limits_{t=1}^{n} t^2 + a_0 \sum\limits_{t=1}^{n} t = \sum\limits_{t=1}^{n} t y_t, \\ [8pt] a_1 \sum\limits_{t=1}^{n} t + a_0 n = \sum\limits_{t=1}^{n} y_t. \end{cases}$$

Используя данные табл. 6.2 ($n = 11$), приходим к системе двух уравнений с неизвестными коэффициентами a_0 и a_1:

$$\begin{cases} 506a_1 + 66a_0 = 4130, \\ 66a_1 + 11a_0 = 748. \end{cases}$$

Решая эту систему, получим: $a_0 = 87{,}47$; $a_1 = -3{,}25$.

Аппроксимирующая функция принимает вид

$$\tilde{y} = 87{,}47 - 3{,}25t. \tag{6.15}$$

Прямая изображена ни рис. 6.2, а полученные с использованием уравнения прямой точки для заданных значений t приведены в пятом столбце табл. 6.3.

2. Считая функцию (6.15) моделью тренда, проверим ее адекватность, опираясь на четыре требования адекватности, описанные в § 6.3.

2.1. Проверка случайности колебаний уровней остаточной последовательности.

Определим разность уровней исходного ряда (второй столбец табл. 6.3) и значений аппроксимирующей функции (пятый столбец):

$$\varepsilon_t = y_t - \tilde{y}_t.$$

Данные занесены в шестой столбец таблицы. Эти данные позволяют определить две величины, необходимые для дальнейших расчетов. Наибольшую протяженность серии значений ε с одинаковым знаком — это значения ε при $t \in [1;3]$, или $[4;5]$, или $[6;8]$, — $L_{max} = 2$. Число серий — количество интервалов ε с одинаковым знаком $\nu = 5$ — это интервалы t : [1;3], [4;5], [6;8], [9] и [11] (нулевые значения ε в расчет не берутся).

Подставляя полученные данные в зависимости (6.6):

$$L_{max} < [3{,}3(\lg n + 1)];$$

$$\nu > \left[\frac{1}{2}\left(n + 1 - 1{,}96\sqrt{n-1}\right)\right],$$

получим (квадратные скобки обозначают целую часть стоящего в них выражения):

$$2 < [3{,}3(\lg 11 + 1)]; \quad \Rightarrow \quad 2 < 6 \ (!)$$

$$5 > \left[\frac{1}{2}(11 + 1 - 1{,}96\sqrt{11-1})\right] \quad \Rightarrow \quad 5 > 2 \ (!).$$

Оба неравенства выполняются, поэтому гипотеза о случайном характере отклонений уровней временного ряда от тренда не отвергается.

2.2. Проверяем соответствие распределения случайной компоненты нормальному закону распределения.

Предварительно необходимо определить показатель асимметрии α и эксцесса e. По формулам (6.7) находим:

$$\alpha = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^3}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^3}} = \frac{\frac{1}{11}(-26{,}4)}{\sqrt{\left(\frac{1}{11}\cdot 26{,}5\right)^3}} \approx -0{,}645;$$

$$e = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^4}{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^2} - 3 = \frac{\frac{1}{11}\cdot 151}{\left(\frac{1}{11}\cdot 26{,}5\right)^2} - 3 \approx -0{,}61.$$

Используя далее зависимости (6.8), определим средние квадратические отклонения величин α и e:

$$\sigma_\alpha = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} = \sqrt{\frac{6(11-2)}{(11+1)(11+3)}} \approx 0{,}57;$$

$$\sigma_e = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}} = \sqrt{\frac{24\cdot 11(11-2)(11-3)}{(11+1)^2(11+3)(11+5)}} \approx 0{,}77.$$

После этого проверяем выполнение неравенств (6.9):

$$|\alpha| < 1{,}5\sigma_\alpha;\ \Rightarrow\ |-0{,}645| < 1{,}5\cdot 0{,}57;\ \Rightarrow\ 0{,}645 < 0{,}85\ (!)$$

$$\left|e + \frac{6}{n+1}\right| < 1{,}5\sigma_e;\ \Rightarrow\ \left|-0{,}61 + \frac{6}{11+1}\right| < 1{,}5\cdot 0{,}77;\ \Rightarrow$$

$$\Rightarrow\ 0{,}11 < 1{,}16\ (!).$$

Таким образом, гипотеза о соответствии распределения случайной компоненты нормальному закону распределения не отвергается.

2.3. Проверяем равенство нулю математического ожидания случайной компоненты.

Определяя среднее арифметическое величины ε:

$$\overline{\varepsilon} = \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t = \frac{1}{11}\cdot 0{,}1 \approx 0{,}01$$

и ее среднее квадратичное отклонение

$$\sigma_\varepsilon = \frac{1}{n-1}\sum_{t=1}^{n}\varepsilon_t^2 - \overline{\varepsilon}^2 = \frac{1}{11-1}\cdot 26{,}5 - (0{,}01)^2 \approx 2{,}64,$$

находим расчетное значение t-критерия Стьюдента (6.10):

$$t = \frac{\overline{\varepsilon} - 0}{\sigma_\varepsilon}\sqrt{n} = \frac{0{,}01 - 0}{2{,}64}\sqrt{11} = 0{,}0125.$$

По таблице Приложения 3 для уровня значимости $\alpha = 0{,}05$ и числа степеней свободы $n - 1 = 11 - 1 = 10$ определяем: $T_{kp} = 2{,}23$.

Так как $0{,}0125 < 2{,}23$ $(t < T_{kp})$, делаем вывод о том, что гипотеза о равенстве нулю математического ожидания случайной компоненты не отвергается.

2.4. Проверяем независимость значений уровней случайной компоненты.

Используя данные, приведенные в седьмом и предпоследнем столбцах табл. 6.3 (цифры предпоследнего столбца предварительно возводим в квадрат и складываем), найдем расчетные значения d-критерия Дарбина–Уотсона (6.11):

$$d = \frac{\sum\limits_{t=2}^{n}\left(\varepsilon_t - \varepsilon_{t-1}\right)^2}{\sum\limits_{t=2}^{n}\varepsilon_t^2} = \frac{50{,}93}{26{,}5} = 1{,}93.$$

Это значение больше (табл. Приложения 6) величины $d_{\text{В}} = 1{,}36$ — верхнего критического значения d-критерия для $k = 2$ (два параметра: a_0 и a_1) и для $n = 15$ (тем более, будет больше для $n = 11$).

Поэтому делаем заключение о том, что гипотеза о независимости уровней случайной компоненты не отвергается.

Так как все четыре условия пунктов **2.1**–**2.4** выполнены, делаем вывод о том, что модель тренда (6.15) адекватна, то есть она правдоподобно описывает тренд рассматриваемого временного ряда, представленного табл. 6.2.

3. Оценим точность модели тренда.

Воспользовавшись зависимостями (6.12) и (6.13), определим в процентах ошибки по среднему квадратичному отклонению:

$$\sigma_\varepsilon = \sqrt{\frac{1}{n-k}\sum_{t=1}^{n}\varepsilon^2} = \sqrt{\frac{1}{11-2}26{,}5} \approx 1{,}7\%$$

и по средней относительной ошибке аппроксимации:

$$\overline{\varepsilon} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{\varepsilon}{y_t}\right| \cdot 100 = \frac{1}{11}\cdot 20{,}9 \approx 1{,}9\%.$$

Полученные значения ошибок меньше 5%, что говорит о хорошей точности модели тренда.

§ 6.6. ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ ТРЕНДОВЫХ МОДЕЛЕЙ

Прогнозирование динамики процессов по трендовым моделям основано на экстраполяции соответствующих кривых, т. е. распространении закономерностей, выявленных в рассматриваемом периоде, за его пределы. Цель прогнозирования — получение представления о будущем на основании сведений о прошлом и настоящем.

На практике используются два основных вида прогноза. Первый осуществляется только с использованием детерминированных моделей тренда, т. е. моделей, являющихся функци-

ями, аппроксимирующими временные ряды процессов, происходящих в реальных условиях. Это так называемая *внутренняя детерминация*. Второй вид прогноза, называемый *внешней детерминацией,* учитывают возможность влияния на развитие процесса дополнительных внешних факторов, которые не учитывались при составлении модели тренда.

Прогноз, осуществляемый по схеме внутренней детерминации, содержит два элемента: точечный и интервальный.

Точечный прогноз определяет единственное значение рассматриваемого показателя. Это значение определяется подстановкой в уравнение тренда конкретных значений $t = n + 1$, $t = n + 2$ и т. д. (n — количество точек заданного временного ряда).

Так как совпадение значений $\tilde{y}_t = \tilde{y}(t = n + l)$ ($l = 1, 2, \ldots$) со значениями, которые примет y в реальных процессах при $t = n + l$, маловероятно, то прогноз должен сопровождаться указанием *доверительных интервалов,* т. е. интервалов изменения величины y, в которые с заданной вероятностью попадут значения прогнозируемой величины. Величина доверительного интервала, установленная путем анализа заданного временного ряда, называется *интервальным прогнозом*.

Хотя временные ряды, строго говоря, нельзя считать совокупностью случайных величин, тем не менее к их анализу применяют методы математической статистики. Так, доверительный интервал считают зависящим от стандартной ошибки в оценке прогнозируемого показателя, от времени упреждения прогноза, от количества уровней в заданном временном ряде и т. д.

Стандартную (среднюю квадратичную) ошибку оценки прогнозируемого показателя рекомендуется определять по формуле:

$$\sigma_{\tilde{y}} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \tilde{y}_t)^2}{n - k}}. \qquad (6.16)$$

Здесь k — количество параметров трендовой модели (число постоянных в аппроксимирующей функции).

Для определения доверительного интервала на величине упреждения l можно использовать выражение:

$$U_y = \tilde{y}_{n+l} \pm \sigma_{\tilde{y}} K, \tag{6.17}$$

где для случая линейной модели тренда

$$K = T_{kp}\sqrt{1 + \frac{1}{n} + \frac{3(n+2l-1)^2}{n(n^2-1)}}, \tag{6.18}$$

T_{kp} — табличное значение критерия Стьюдента для уровня значимости α и для числа степеней свободы, равного $n-2$.

Значения K табулированы. В табл. 6.4 приведены некоторые данные для определения K для ограниченного количества уровней n и протяженности упреждения l.

Таблица 6.4
Значения величины K

	l					
n	1	2	3	4	5	6
7	1,93	2,11	2,30	2,51	2,73	2,97
10	1,69	1,77	1,87	1,96	2,07	2,18
13	1,58	1,63	1,68	1,74	1,80	1,86
15	1,54	1,57	1,61	1,65	1,70	1,75

Формулы подсчета величины K для аппроксимирующих полиномов, имеющих второй и выше порядок, а также для экспоненциальных и других функций имеют более громоздкий по сравнению с (6.18) вид и приведены в специальной литературе по эконометрике. Отметим, что эти формулы существенно отличаются для разных типов трендовых моделей. Но каждая из них отражает тенденцию увеличения доверительного интервала с ростом упреждения l. Следует иметь в виду, что при слишком больших l прогноз становится нереальным.

Важным этапом прогнозирования является его заключительный этап — *верификация прогноза*, т. е. удостоверение

в его подлинности. Верификация осуществляется путем сопоставления расчетных данных по модели с соответствующими фактическими данными развития процесса и заканчивается определением критериев, способов и процедур, позволяющих оценить качество прогноза. Точность прогноза напрямую связана с величиной его интервальной оценки: чем меньше доверительный интервал, тем выше точность прогноза.

Для количественной оценки точности прогноза можно использовать показатель:

$$\tau = \frac{r}{p},$$

где r — число прогнозов, подтвержденных фактическими данными; p — общее число прогнозов.

Приведенная формула оценки прогноза предполагает, что прогнозируемое событие уже произошло. Однако на практике важно иметь возможность убедиться в надежности прогноза еще до реализации процесса в будущем.

Пример. Для временного ряда, представленного в табл. 6.2, дать прогноз на два шага вперед (для $t = 12$ и $t = 13$).

Р е ш е н и е.

Для прогноза используем трендовую модель (6.15):

$$\tilde{y} = 87{,}47 - 3{,}25t.$$

Подставляя в это уравнение заданные значения t, получим два точечных прогноза:

$$\tilde{y}_{12} = 87{,}47 - 3{,}25 \cdot 12 = 48{,}5;$$

$$\tilde{y}_{13} = 87{,}47 - 3{,}25 \cdot 13 = 45{,}3.$$

По формуле (6.16) определим значение среднего квадратичного отклонения оценки прогнозируемого показателя, используя данные седьмого столбца табл. 6.3, т.е. $\sum_{t=1}^{n}(y_t - \tilde{y}_t)^2 =$

$$= \sum_{t=1}^{n} \varepsilon_t^2 = 26{,}5:$$

$$\sigma_{\tilde{y}} = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \tilde{y}_t)^2}{n-k}} = \sqrt{\frac{26{,}5}{12-2}} = 1{,}62.$$

При $l_1 = 1$, $l_2 = 2$ по табл. 6.4 путем линейной интерполяции значений между $n = 10$ и $n = 13$ для $n = 11$ найдем:

$$K_1 = K(l_1) = 1{,}69 - \frac{1{,}69 - 1{,}58}{13 - 10} = 1{,}65;$$

$$K_2 = K(l_2) = 1{,}77 - \frac{1{,}77 - 1{,}63}{13 - 10} = 1{,}72.$$

Теперь определим величины доверительных интервалов для заданных значений t, используя выражение (6.17).

Для $t = 12$ $(l = 1)$:

$$U_y(l = 1) = \tilde{y}_{11+1} \pm 1{,}62 \cdot 1{,}65 = 48{,}5 \pm 2{,}67, \quad \Rightarrow$$

$$U_y^+(l = 1) = 51{,}2; \qquad U_y^-(l = 1) = 45{,}8.$$

Для $t = 13$ $(l = 2)$:

$$U_y(l = 2) = \tilde{y}_{11+2} \pm 1{,}62 \cdot 1{,}72 = 45{,}3 \pm 2{,}79, \quad \Rightarrow$$

$$U_y^+(l = 2) = 48{,}1; \qquad U_y^-(l = 2) = 42{,}5.$$

В приведенных расчетах знаком «+» отмечены верхние границы доверительного интервала, знаком «−» — нижние границы. Так что с принятым при составлении табл. 6.4 уровнем значимости $\alpha = 0{,}2$ (20%) можно утверждать, что с вероятностью 80% величина, например, $U_y(l = 1)$ будет находиться на интервале от 45,8 до 51,2.

§ 6.7. АДАПТИВНЫЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ

В реальных процессах в условиях изменения внешних воздействий, особенно на последних стадиях протекания процессов, эти изменения могут оказывать существенное влияние на дальнейшее поведение показателей рядов. Неравноценность уровней временного ряда при прогнозировании поведения рассматриваемых процессов учитывается в адаптивных моделях (от английского слова «adapt» — приспосабливать).

Адаптивные модели прогнозирования — это такие модели, которые учитывают влияние изменения внешних условий на протекание процесса.

В адаптивных моделях при оценке показателей процессов (уровней временных рядов) им присваиваются различные веса, величина которых зависит от того, насколько сильным признается влияние внешних условий на текущие уровни ряда. В адаптивных моделях используются две схемы: скользящего среднего и авторегрессии.

В *схеме скользящего среднего* оценкой текущих уровней ряда являются взвешенные средние всех предшествующих уровней, причем их веса увеличиваются по мере приближения к последнему уровню. Такие модели отражают тенденции изменения внешних условий, но не влияют на воздействие колебаний, наблюдаемых в некоторых процессах.

В *авторегрессионной схеме* оценкой текущего уровня служит взвешенная сумма не всех, а только ближайших предшествующих уровней. Отметим, что авторегрессионные модели приспособлены к описанию временных рядов, не имеющих тенденций в развитии, потому не представляют большого интересы для экономических показателей и здесь не рассматриваются.

Схема построения адаптивных моделей следующая. По нескольким первым уровням ряда оцениваются значения параметров модели и строится текущая модель процесса. По этой модели делается прогноз на один шаг вперед. Отклонение прогноза от фактического уровня ряда принимается за ошибку прогнозирования, которая позволяет скорректировать модель.

По скорректированной модели рассчитывается прогноз на следующий шаг и т. д. Модель постоянно насыщает себя новой информацией (обучается) и к концу обучения отражает тенденцию развития процесса.

Одной из самых распространенных моделей, основанной на схеме скользящего среднего, является модель Брауна.

Модель Брауна прогноза уровней рядов предполагает использовать для аппроксимации процессов функцию, представляющую собой полином до второго порядка:

$$y_{t+l} = A_0 + A_1 l + A_2 l^2, \tag{6.19}$$

где l — количество шагов, на которые строится прогноз.

В зависимости от количества слагаемых, удерживаемых в модели (6.19), различают модели Брауна *нулевого порядка* (удерживается только слагаемое A_0), *первого порядка* (удерживается только слагаемые с параметрами A_0 и A_1) — это линейная модель и *второго порядка* — параболическая модель, записываемая в виде (6.19), т. е. с удержанием всех трех слагаемых. Последняя модель учитывает не только скорость изменения уровней (как вторая модель), но и их ускорение.

Порядок выбранной модели обычно определяют по характеру поведения во времени уровней исходного ряда.

Остановимся на линейной модели Брауна и рассмотрим этапы ее построения.

1. По первым n_1 (чаще всего принимают $n_1 = 5$) точкам заданного временного ряда оцениваются значения параметров A_0 и A_1 модели с помощью метода наименьших квадратов для линейной аппроксимации:

$$\tilde{y}(t) = A_0 + A_1 t \qquad (t = \overline{1, n_1}).$$

Отметим, что процедура определения коэффициентов этого уравнения методом наименьших квадратов осуществляется в модели Брауна, как правило, только для исходного шага, т. е. для оценки расчетной переменной при $t = 1$. Считая A_0 и $A_1 = B_0$ исходными (нулевыми) значениями параметров модели (при $t = 1$), получим выражение для определения расчетного значения показателя процесса $\tilde{y}_1 = \tilde{y}(1)$ для первой

точки:

$$\tilde{y}_1 = A_0 + B_0.$$

Параметры аппроксимации временного ряда для первых n_1 точек определяются методом наименьших квадратов по формулам (4.9):

$$B_0 = \frac{\sum\limits_{t=1}^{n_1}(t-\overline{t})(y_t-\overline{y})}{\sum\limits_{t=1}^{n_1}(t-\overline{t})^2};$$

$$A_0 = \overline{y} - B_0\overline{t}, \tag{6.20}$$

в которых $\overline{t} = \dfrac{1}{n_1}\sum\limits_{t=1}^{n_1} t$, $\overline{y} = \dfrac{1}{n_1}\sum\limits_{t=1}^{n_1} y_t$ — средние арифметические величин t и y_t на рассматриваемом интервале $t \in [1, n_1]$ изменения временного интервала.

2. Расчетное значение $\tilde{y}_1$ показателя временного ряда сравнивается с его заданным значением y_1 в первой рассматриваемой точке (при $t = 1$) и находится величина расхождения этих величин:

$$e_1 = y_1 - \tilde{y}_1.$$

3. Определяются параметры линейной модели аппроксимации путем корректировки их значений, найденных в исходном приближении (6.20), с использованием величины расхождения:

$$B_1 = B_0 + \alpha^2 e_1;$$

$$A_1 = A_0 + B_1.$$

В этих формулах $\alpha \in [0; 1]$ — параметр сглаживания, величина которого вбирается в зависимости от степени влияния на рассматриваемый процесс более поздних наблюдений. Чем ближе α к единице, тем больше это влияние. При $\alpha = 0$ влияние истории наблюдений на рассматриваемый процесс отсутствует и параметры аппроксимации принимаются равными исходным (A_0 и B_0) во всех точках ряда.

4. По найденным параметрам A_1 и B_1 находят расчетное значение показателя ряда во второй точке:

$$\tilde{y}_2 = A_1 + B_1.$$

После этого повторяются вычисления, следуя пунктам **2–4** для $t = 2$ и т. д.

Для некоторого значения t с рассчитанными $\tilde{y}_t$ и заданными y_t значениями уровня, найденными A_t и B_t, рекуррентные формулы метода Брауна выглядят следующим образом.

2.

$$e_t = y_t - \tilde{y}_t.$$

3.

$$B_{t+1} = B_t + \alpha^2 e_t;$$

$$A_{t+1} = A_t + B_{t+1}.$$

4.

$$\tilde{y}_{t+2} = A_{t+1} + B_{t+1}.$$

После этого повторяются вычисления, следуя пунктам **2–4** для $t+2$ и т. д. до $t = T$, т. е. до достижения последней точки.

Прогноз на l точек вперед после $t = T$ строится по линейной модели аппроксимации, в которой значения параметров найдены для последней точки заданного ряда:

$$\tilde{y}_{T+l} = A_{(T)} + B_{(T)} l.$$

Интервальный прогноз делается, как для линейной модели кривой роста, т. е. для определения доверительного интервала используются формулы (6.17)–(6.18).

Пример. Используя линейную модель Брауна, построить прогноз процесса, заданного временным рядом табл. 6.2.

Р е ш е н и е задачи осуществим, следуя пунктам **1–4** и занося промежуточные данные в табл. 6.5.

1. Условимся все расчеты проводить с точностью до удержания первого знака после запятой.

Начальные оценки параметров линейной модели Брауна получим по первым пяти точкам ($t \in [1;5]$) заданного временного ряда, который представлен первыми двумя столбцами табл. 6.5.

Таблица 6.5
Реализация линейной модели Брауна

t	y_t	A_t	B_t	$\tilde{y}_t$	e_t
		90,7	-4,5		
1	86	86,2	-4,5	86,2	-0,2
2	81	81,6	-4,6	81,7	-0,7
3	79	77,3	-4,3	77,0	2,0
4	72	72,8	-4,5	73,0	-1,0
5	68	68,3	-4,5	68,3	-0,3
6	70	64,8	-3,5	63,8	6,2
7	65	61,9	-2,9	61,3	3,7
8	62	59,6	-2,3	58,0	4,0
9	57	57,3	-2,3	57,3	-0,3
10	55	55,2	-2,3	55,0	0,0
11	53	54,7	-2,6	55,0	-2,0
12				52,1	
13				49,5	

Предварительно определим средние арифметические значения t и y_t по первым пяти точкам заданного ряда:

$$\overline{t} = \frac{1}{5}(1+2+3+4+5) = 3;$$

$$\overline{y} = \frac{1}{5}(86+81+79+72+68) = 77{,}2.$$

Используя соотношения (6.20), найдем:

$$B_0 = \frac{\sum\limits_{t=1}^{5}(t-\overline{t})(y_t-\overline{y})}{\sum\limits_{t=1}^{5}(t-\overline{t})^2} =$$

$$= \frac{(1-3)(86-77{,}2)+(2-3)(81-77{,}2)+\ldots+(5-3)(68{,}3-77{,}2)}{(1-3)^2+(2-3)^2+\ldots+(5-3)^2} \approx$$

$$\approx -4{,}5;$$

$$A_0 = \overline{y} - B_0\overline{t} = 77{,}2 - (-4{,}5)\cdot 3 = 90{,}7.$$

Полученные начальные значения параметров модели позволяют определить расчетное значение показателя ряда $\tilde{y}_1$, т. е. показателя, соответствующего одному шагу вперед ($l = 1$) по сравнению с фиктивной нулевой точкой:

$$\tilde{y}_1 = A_0 + B_0 = 90{,}7 + (-4{,}5) = 86{,}2.$$

Это значение внесено в пятый столбец табл. 6.5 в строку, соответствующую $t = 1$.

2. Сравниваем значение $\tilde{y}_1$ с заданной величиной уровня y_1 для первой точки и находим величину расхождения:

$$e_1 = y_1 - \tilde{y}_1 = 86 - 86{,}2 = -0{,}2.$$

3. Задаваясь величиной параметра сглаживания $\alpha = 0{,}4$, определим модифицированные значения параметров модели для $t = 1$:

$$B_1 = B_0 + \alpha^2 e_1 = -4{,}5 + (0{,}4)^2(-0{,}2) \approx -4{,}5;$$

$$A_1 = A_0 + B_1 = 90{,}7 + (-4{,}5) = 86{,}2.$$

Эти значения занесены в третий и четвертый столбцы табл. 6.5.

4. Находим расчетное значение параметра модели для второй точки:

$$\tilde{y}_2 = A_1 + B_1 = 86{,}2 + (-4{,}5) = 81{,}7.$$

Возвращаемся в пункту **2** и повторяем вычисления для $t = 2$. После этого повторяем неоднократно вычисления по пунктам **2**–**4** до $t{=}T{=}11$. Результаты вычислений занесены в табл. 6.5.

Для последующих точек ряда (для прогноза) получаем модель линейной аппроксимации в виде:

$$\tilde{y}_{11} = A_{11} + B_{11}l = 54{,}7 - 2{,}6l.$$

Эта модель позволяет прогнозировать поведение показателя $\tilde{y}_{T+l}$ на заданное число шагов (l) вперед по отношению к последней точке $t{=}T{=}11$.

Сделаем такой прогноз для $l = 1$:

$$\tilde{y}_{12} = A_{11} + B_{11} \cdot 1 = 54{,}7 - 2{,}6 \cdot 1 = 52{,}1$$

и $l = 2$:

$$\tilde{y}_{13} = A_{11} + B_{11} \cdot 2 = 54{,}7 - 2{,}6 \cdot 2 = 49{,}5.$$

Результаты занесены в последние две строки табл. 6.5.

Отметим в заключение, что интервальные оценки по линейной модели Брауна рекомендуется проводить по формуле:

$$U_l = \tilde{y}_{T+l} \pm \sigma_{\tilde{y}} t_\alpha \sqrt{1 + \frac{1}{T} + \frac{(T + l - \overline{t})^2}{\sum\limits_{t=T-n_1}^{T} (t - \overline{t})^2}}.$$

Величины, входящие в это выражение, имеют то же значение, что и для формул (6.17)–(6.18).

§ 6.8. РЕГРЕССИОННЫЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ

Обсудим задачи прогнозирования на линейных моделях регрессии. Качество таких моделей может быть оценено, как и трендовых моделей, по адекватности и точности. Адекватность регрессионных моделей может быть установлена путем анализа остаточной последовательности, т. е. разности между совокупностью значений аппроксимирующей функции и соответствующими значениями задаваемых показателей ряда.

Остаточная последовательность проверяется на выполнение свойств случайной компоненты:

— близости нулю математического ожидания;

— случайного характера отклонений;

— отсутствия автокорреляции;

— нормальности закона распределения.

Перечисленные проверки совпадают в общем с проверками, рассмотренными в § 6.3, и выполняются с использованием описанных в упомянутом разделе методов и с применением тех же статистических критериев.

О качестве регрессионных моделей можно судить и по значениям коэффициентов корреляции. Формулы для расчета этих коэффициентов для однофакторных моделей и для моделей множественной корреляции, а также примеры их определения приведены в гл. 4. Чем ближе коэффициенты корреляции (некоторые из них по абсолютной величине) к единице, тем теснее связь между изучаемым признаком и выбранными факторами и тем с большей уверенностью можно говорить об адекватности модели.

Точность регрессионных моделей можно оценить по таким же статистическим критериям, какие используются для трендовых моделей, в том числе по средней относительной ошибке аппроксимации (6.13).

Проверка *значимости модели* регрессии проводится с использованием F-критерия Фишера. Расчетные значения критерия находятся как отношение дисперсии показателей исходного ряда и несмещенной оценки дисперсии остаточной последовательности регрессионной модели. Если расчетное значение этого критерия больше табличного значения F-критерия, найденного при заданном уровне значимости, то модель признается значимой.

После проверки регрессионной модели на точность, адекватность и значимость и подтверждения возможности применения к исследованию процесса можно использовать ее для прогнозирования.

Для осуществления прогнозирования на l шагов вперед необходимо знать значения всех входящих в модель факторов. Прогнозные значения могут быть получены экстраполяцией с использованием средних приростов факторных признаков (могут быть получены, например, путем экспертных оценок). Прогнозные значения факторов подставляют в регрессионную модель и получают точечные оценки интересующего показателя.

При интервальной оценке прогноза (определении доверительных интервалов) необходимо учитывать возможность

влияния на процесс ошибок первого и второго рода. Ошибки первого рода обусловлены рассеянием наблюдений относительно линии (плоскости) регрессии (в случае линейной аппроксимации) и их можно учесть, например, величиной средней квадратичной ошибки аппроксимации по отношению к заданному показателю.

Ошибки второго рода обусловлены тем, что однозначно определенные в модели коэффициенты регрессии на самом деле являются случайными величинами. Эти ошибки учитываются введением поправочных коэффициентов в формулы для расчета доверительных интервалов. Формулы для определения доверительных интервалов при прогнозировании с использованием регрессионных моделей мало отличаются от (6.17)–(6.18).

§ 6.9. ОБОБЩЕНИЕ СТАТИСТИЧЕСКИХ МОДЕЛЕЙ

Важным с прикладной точки зрения элементом исследования статистических закономерностей является синтез результатов, полученных математическими методами, с привлечением теоретического анализа.

Предметом анализа служит конкретная *статистическая модель,* которая выступает как средство анализа и прогнозирования процессов.

Статистические модели можно классифицировать по ряду признаков.

По аналитической форме модели делят на:

— линейные;

— нелинейные.

К нелинейным относят степенные (с наибольшим показателем степени больше единицы), показательные (экспоненциальные) и т. д. Примером нелинейной модели, получившей достаточно широкое распространение, является *модель Брандона*:

$$y = \overline{y} f_1(x_1) f_2(x_2) \dots f_k(x_k),$$

где y — изучаемый показатель, который называют *результативным признаком,* черта над y указывает на то, что

это среднее арифметическое (математическое ожидание); x_1, $x_2, \ldots, x_k$ — величины, влияющие на y и называемые *факторными признаками*.

Существует разница между показателями, входящими в статистические модели. Эти показатели могут быть как переменными, так и постоянными и подразделяются на:

— *эндогенные (или внутренние, выходные);*

— *экзогенные (или внешние, входные).*

По направлению и сложности связей модели подразделяются на:

— регрессионные модели;

— взаимозависимые системы;

— рекурсивные системы.

Регрессионные модели основываются на уравнениях прямых регрессии, а также на уравнениях множественной регрессии, представляющих собой уравнения гиперплоскостей, наилучшим образом описывающих зависимости между одной эндогенной переменной y и k экзогенными переменными x_1, x_2, ...:

$$y = a_0 + a_1x_1 + a_2x_2 + \ldots + a_kx_k.$$

Такие модели для одно- и двухмерных пространств рассмотрены в гл. 4.

Взаимозависимые системы описывают системы, содержащие определенные количества (не по одному) эндогенных (m) и экзогенных (k) переменных:

$$\begin{cases} y_1 = a_{10} + a_{11}x_1 + a_{12}x_2 + \ldots \\ \qquad\qquad \ldots + a_{1k}x_k + b_{12}y_2 + b_{13}y_3 + \ldots + b_{1m}y_m; \\ y_2 = a_{20} + a_{21}x_1 + a_{22}x_2 + \ldots \\ \qquad\qquad \ldots + a_{2k}x_k + b_{21}y_1 + b_{23}y_3 + \ldots + b_{2m}y_m; \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ y_m = a_{m0} + a_{m1}x_1 + a_{m2}x_2 + \ldots \\ \qquad \ldots + a_{mk}x_k + b_{m1}y_1 + b_{m2}y_2 + \ldots + b_{m,m-1}y_{m-1}. \end{cases}$$

Эта система уравнений требует специальных методов решения, которые не отличаются простотой. Наиболее распространенным подходом к получению решения таких систем

является сведение системы уравнений к так называемому *рекурсивному* виду.

Рекурсивные системы представляют собой следующее. Сначала выбирают одну эндогенную переменную y_1 такую, которая зависит только от экзогенных переменных. Затем выбирается следующая эндогенная переменная y_2, зависящая от экзогенных переменных и y_1. Третья эндогенная переменная будет зависеть от экзогенных переменных и в общем уже двух экзогенных переменных y_1 и y_2 и т. д. до тех пор, пока не будут найдены все y_i $(i = \overline{1,m})$. Таким образом, каждая последующая эндогенная переменная будет зависеть от всех (или части) предыдущих и от экзогенных переменных.

Параметры первого уравнения рекурсивной системы находят обычно методом наименьших квадратов. Их подставляют вместе с найденным значением y_1 во второе уравнение, снова применяют метод наименьших квадратов и т. д.

Процесс построения статстических моделей включает в себя этапы:

— определение цели исследования, построение системы показателей и логический отбор наиболее существенных факторов, влияющих на процесс;

— выбор формы связи между параметрами модели и упомянутыми факторами;

— сбор исходных данных и анализ информации;

— построение модели;

— проверка адекватности модели изучаемому процессу;

— использование математической модели для расчета реальных процессов (анализа и прогнозирования).

При реализации перечисленных этапов важно из большого количества показателей отобрать те, которые в наибольшей мере влияют на рассматриваемый процесс, а также выявить факторы, более всего влияющие на главные показатели.

Существенного сокращения числа факторов, влияющих на процесс, можно добиться, используя пошаговую процедуру оценивания этих факторов с привлечением статистических методов.

Различают две разновидности пошаговых методов: исключения и включения.

Метод включения заключается в последовательном включении в регрессионную модель все новых и новых переменных до тех пор, пока модель не будет адекватно описывать исследуемое явление. Вопрос о том, включать или не включать очередной показатель или фактор в модель определяется значением частного коэффициента корреляции. Чем больше этот коэффициент, тем большее влияние на процесс оказывает рассматриваемый показатель. Показатели с максимальным частным коэффициентом корреляции первыми включаются в модель.

Метод исключения, в противоположность методу включения, предполагает первоначальное включение в модель максимально возможного количества показателей или факторов с последующим исключением из нее тех, которые мало влияют на рассматриваемый процесс.

§ 6.10. РЕЗЮМЕ

Прогнозирование развития статистических процессов можно осуществлять, опираясь на тренды — выделенные из временного ряда составляющие, ответственные за тенденцию развития процесса.

Для удобства анализа тренда подбирается соответствующая регрессионная модель. Наиболее часто используются линейные регрессионные модели, как наиболее простые и с помощью которых можно осуществить кусочно-линейную аппроксимацию нелинейных функций. В силу того, что многие процессы описываются функциями, близкими к полиномам различных степеней, экспоненциальным и логарифмическим функциям, основанные на них регрессионные модели также широко используются в анализе.

После выбора регрессионной модели осуществляется ее проверка на адекватность по отношению к тренду и точность. Соответствующие методы проверки описаны в § 6.3–§ 6.4.

Прогноз развития процессов может быть осуществлен на основе трендовых, адаптивных или регрессионных моделей. В любом из указанных методов достоверность прогноза определяется доверительным интервалом, построенным на

основе поведения исследуемого показателя, представленного случайным образом изменяемыми величинами.

§ 6.11. ВОПРОСЫ

1. В чем суть прогнозирования процессов на основе метода экстраполяции?
2. Перечислите основные типы кривых роста для построения трендовых моделей и опишите их свойства.
3. Перечислите методы предварительного выбора типа кривых роста. Как находятся параметры этих кривых?
4. Как оценивается адекватность трендовых моделей? Какие статистические критерии при этом используются?
5. Как оценивается точность моделей прогнозирования процессов? Какие статистические критерии при этом используются?
6. Перечислите этапы прогнозирования динамики процессов на основе одномерных временных рядов.
7. Что такое точечный и интервальный прогноз показателей временных рядов?
8. Как оценить точность прогноза? Какие факторы влияют на величину доверительного интервала в интервальном прогнозе?
9. Перечислите основные адаптивные модели. Поясните их суть.
10. Восстановите последовательность построения адаптивной модели Брауна.

§ 6.12. ЗАДАНИЕ НА РАСЧЕТНУЮ РАБОТУ, ЧАСТЬ 8

Для заданного временного ряда

t	1	2	3	4	5	6	7	8	9	10
y_t	45	43	50	48	55	56	61	59	68	64

выполнить действия, сформулированные в пунктах **1–5.**

1. Выбрать наилучшую кривую роста методами:

а) конечных разностей;

б) характеристик прироста.

2. Построить линейную модель, определив ее параметры методом наименьших квадратов.

3. Построить адаптивную модель Брауна с параметрами сглаживания $\alpha = 0{,}4$ и $\alpha = 0{,}8$.

4. Оценить адекватность и точность моделей, построенных в пункте **3.** Выбрать лучшую из моделей.

5. На основании лучшей модели, выбранной в пункте **4,** сделать точечный и интервальный прогноз на один и на два шага вперед при параметре Стьюдента $T_{кр} = 1{,}09$.

Результаты расчетов отразить на графиках.

ПРИЛОЖЕНИЯ

Приложение 1

Значения функции $f(t) = \dfrac{1}{\sqrt{2\pi}} e^{-t^2/2} \big(f(-t) = f(t)\big)$

	сотые доли t									
t	0	1	2	3	4	5	6	7	8	9
0,0	0,389	0,399	0,399	0,399	0,399	0,398	0,398	0,398	0,398	0,397
0,1	0,397	0,397	0,396	0,396	0,395	0,395	0,394	0,393	0,393	0,392
0,2	0,391	0,390	0,389	0,389	0,388	0,387	0,386	0,385	0,384	0,383
0,3	0,381	0,380	0,379	0,378	0,377	0,375	0,374	0,373	0,371	0,370
0,4	0,368	0,367	0,365	0,364	0,362	0,361	0,359	0,357	0,356	0,354
0,5	0,352	0,350	0,348	0,347	0,345	0,343	0,341	0,339	0,337	0,335
0,6	0,333	0,331	0,329	0,327	0,325	0,323	0,321	0,319	0,317	0,314
0,7	0,312	0,310	0,308	0,307	0,304	0,301	0,299	0,297	0,294	0,292
0,8	0,290	0,287	0,285	0,283	0,280	0,278	0,276	0,273	0,271	0,269
0,9	0,266	0,264	0,261	0,259	0,257	0,254	0,252	0,249	0,247	0,244
1,0	0,242	0,240	0,237	0,235	0,232	0,230	0,228	0,225	0,223	0,220
1,1	0,218	0,216	0,213	0,211	0,208	0,206	0,204	0,201	0,199	0,197
1,2	0,194	0,192	0,190	0,187	0,185	0,183	0,180	0,178	0,176	0,174
1,3	0,171	0,169	0,167	0,165	0,163	0,160	0,158	0,156	0,154	0,152
1,4	0,150	0,148	0,146	0,144	0,142	0,139	0,137	0,135	0,133	0,132
1,5	0,130	0,128	0,126	0,124	0,122	0,120	0,118	0,116	0,115	0,113
1,6	0,111	0,109	0,107	0,106	0,104	0,102	0,101	0,099	0,097	0,096
1,7	0,094	0,093	0,091	0,089	0,088	0,086	0,085	0,083	0,082	0,080
1,8	0,079	0,078	0,076	0,075	0,073	0,072	0,071	0,069	0,068	0,067
1,9	0,066	0,064	0,063	0,062	0,061	0,060	0,058	0,057	0,056	0,055
2,0	0,054	0,053	0,052	0,051	0,050	0,049	0,048	0,047	0,046	0,045
2,1	0,044	0,043	0,042	0,041	0,040	0,040	0,039	0,038	0,037	0,036
2,2	0,036	0,035	0,034	0,033	0,033	0,032	0,031	0,030	0,030	0,029
2,3	0,028	0,028	0,027	0,026	0,026	0,025	0,025	0,024	0,024	0,023
2,4	0,022	0,022	0,021	0,021	0,020	0,020	0,020	0,019	0,019	0,018
2,5	0,018	0,017	0,017	0,016	0,016	0,015	0,015	0,015	0,014	0,014
2,6	0,014	0,013	0,013	0,013	0,012	0,012	0,012	0,011	0,011	0,011
2,7	0,010	0,010	0,010	0,010	0,009	0,009	0,009	0,009	0,008	0,008
2,8	0,008	0,008	0,008	0,007	0,007	0,007	0,007	0,007	0,006	0,006
2,9	0,006	0,006	0,006	0,006	0,005	0,005	0,005	0,005	0,005	0,005
3,0	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,003

Приложение 2

Значения функции $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \exp\left[-t^2/2\right] dt$

$\big(\Phi(-x) = 1 - \Phi(x)\big)$

	сотые доли t									
t	0	1	2	3	4	5	6	7	8	9
0,0	0,500	0,504	0,508	0,512	0,516	0,520	0,524	0,528	0,532	0,536
0,1	0,540	0,544	0,548	0,551	0,555	0,559	0,563	0,567	0,571	0,575
0,2	0,579	0,583	0,587	0,591	0,595	0,598	0,602	0,606	0,610	0,614
0,3	0,618	0,621	0,625	0,629	0,633	0,637	0,640	0,644	0,648	0,653
0,4	0,655	0,659	0,663	0,666	0,670	0,673	0,678	0,681	0,684	0,688
0,5	0,691	0,695	0,698	0,702	0,705	0,709	0,712	0,715	0,719	0,722
0,6	0,726	0,729	0,732	0,735	0,739	0,742	0,745	0,748	0,752	0,755
0,7	0,758	0,761	0,764	0,767	0,770	0,773	0,776	0,779	0,782	0,785
0,8	0,788	0,791	0,794	0,797	0,799	0,802	0,805	0,808	0,810	0,813
0,9	0,816	0,818	0,821	0,824	0,826	0,829	0,831	0,834	0,836	0,839
1,0	0,841	0,844	0,846	0,848	0,851	0,853	0,855	0,857	0,860	0,862
1,1	0,864	0,866	0,868	0,871	0,873	0,875	0,877	0,879	0,881	0,883
1,2	0,885	0,887	0,889	0,890	0,892	0,894	0,896	0,898	0,899	0,901
1,3	0,903	0,905	0,906	0,908	0,910	0,911	0,913	0,914	0,916	0,918
1,4	0,919	0,921	0,922	0,923	0,925	0,926	0,928	0,929	0,930	0,931
1,5	0,933	0,934	0,936	0,937	0,938	0,939	0,940	0,942	0,943	0,944
1,6	0,945	0,946	0,947	0,948	0,949	0,950	0,951	0,952	0,953	0,954
1,7	0,955	0,956	0,957	0,958	0,959	0,960	0,961	0,961	0,962	0,963
1,8	0,964	0,965	0,965	0,966	0,967	0,968	0,968	0,969	0,970	0,970
1,9	0,971	0,972	0,972	0,973	0,974	0,974	0,975	0,975	0,976	0,976
2,0	0,977	0,978	0,978	0,979	0,979	0,980	0,980	0,981	0,981	0,981
2,1	0,982	0,982	0,983	0,983	0,984	0,984	0,984	0,985	0,985	0,986
2,2	0,986	0,986	0,987	0,987	0,987	0,988	0,988	0,988	0,988	0,989
2,3	0,989	0,989	0,990	0,990	0,990	0,990	0,991	0,991	0,991	0,991
2,4	0,992	0,992	0,992	0,992	0,992	0,993	0,993	0,993	0,993	0,993
2,5	0,994	0,994	0,994	0,994	0,994	0,994	0,995	0,995	0,995	0,995
2,6	0,995	0,995	0,995	0,996	0,996	0,996	0,996	0,996	0,996	0,996
2,7	0,996	0,996	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997
2,8	0,997	0,997	0,997	0,997	0,998	0,998	0,998	0,998	0,998	0,998
2,9	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998
3,0	0,998	0,998	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999

Приложение 3

t-распределения Стьюдента

(T_{kp} в процентах в зависимости от числа k степеней свободы и вероятности α)

k / α	1	2	3	4	5	7	10	15	20	30
0,10	6,31	2,92	2,35	2,13	2,02	1,90	1,81	1,75	1,73	1,65
0,05	12,7	4,30	3,18	2,78	2,57	2,37	2,23	2,13	2,09	1,96
0,01	63,7	9,93	5,84	4,60	4,03	3,50	3,20	2,95	2,85	2,75

Приложение 4

Распределения χ^2

(K_{kp} в процентах в зависимости от числа степеней свободы q и вероятности α)

q / α	1	2	3	4	5	6	7	8	9	10
0,010	6,6	9,2	11,3	13,3	15,1	16,8	18,5	20,1	21,7	23,2
0,025	5,0	7,4	9,4	11,1	12,8	14,4	16,0	17,5	19,0	20,5
0,050	3,8	6,0	7,8	9,5	11,1	12,6	14,1	15,5	16,9	18,3

Приложение 5

F-распределение (Фишера) для $\alpha = 0{,}05$
(число степеней свободы: m_1 — для большей дисперсии, m_2 — для меньшей дисперсии)

m_2 \ m_1	1	2	3	4	5	6	8	12	24	∞
2	18,5	19,0	19,2	19,3	19,3	19,4	19,4	19,4	19,5	19,5
3	10,1	9,6	9,3	9,1	9,0	8,9	8,8	8,7	8,6	8,5
4	7,7	6,9	6,6	6,4	6,3	6,2	6,0	5,9	5,8	5,6
5	6,6	5,8	5,4	5,2	5,1	5,0	4,8	4,7	4,5	4,4
7	5,6	4,7	4,4	4,1	4,0	3,9	3,7	3,6	3,4	3,2
10	5,0	4,1	3,7	3,5	3,3	3,2	3,1	2,9	2,7	2,5
15	4,5	3,7	3,3	3,1	2,9	2,8	2,6	2,5	2,3	2,1
20	4,4	3,5	3,1	2,9	2,7	2,6	2,5	2,3	2,1	1,8
∞	3,8	3,0	2,6	2,4	2,2	2,1	1,9	1,8	1,5	1,0

Приложение 6

Критерии Ирвина

n	2	3	10	20	30	50	100
λ_t	2,8	2,3	1,5	1,3	1,2	1,1	1,0

Приложение 7

Критерии Дарбина–Уотсона

	$k=1$		$k=2$		$k=3$	
n	$d_{\text{Н}}$	$d_{\text{В}}$	$d_{\text{Н}}$	$d_{\text{В}}$	$d_{\text{Н}}$	$d_{\text{В}}$
15	1,08	1,36	0,95	1,54	0,82	1,75
20	1,20	1,41	1,10	1,55	1,00	1,68
30	1,35	1,49	1,28	1,57	1,21	1,65

Приложение 8

Нормально распределенные случайные числа
$(\boldsymbol{m} = 3,\ \boldsymbol{\sigma}^2 = 1)$

3,464	3,137	5,455	2,677	2,932	3,296	2,712	4,298	3,060	0,744
2,469	2,806	3,543	1,442	3,187	1,810	4,486	2,646	2,366	3,697
3,926	4,375	3,785	2,037	4,022	2,528	4,279	6,521	3,571	1,149
3,194	4,192	4,394	2,445	3,046	3,321	5,945	4,974	2,742	3,412
3,906	2,487	2,475	3,595	3,881	2,066	4,579	3,161	4,179	1,945
3,007	3,769	3,971	3,712	4,090	2,369	1,499	2,512	2,838	2,864
4,033	3,303	3,448	3,748	2,310	3,756	1,382	2,655	2,689	0,949
2,543	2,782	4,372	3,225	3,378	3,769	3,181	2,264	3,960	1,470
2,518	4,678	2,943	1,771	2,514	3,856	2,509	1,017	1,624	2,850
4,356	2,439	2,744	2,788	3,219	3,779	1,990	3,598	2,082	4,598
4,065	3,415	2,831	3,313	2,995	2,101	3,012	2,275	3,147	2,879
4,096	3,181	4,393	1,837	2,089	4,231	2,801	2,754	4,239	0,426

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

А

Автокорреляция 211
Анализ
    регрессионный 140
Аппроксимация 130
    квадратичная 145
    линейная 141
Асимметрия 249

В

Величина
    определяющая 33
Верификация
    прогноза 259
Волна
    сезонная 217

Д

Детерминация
    внешняя 258
    внутренняя 258
Дисперсия
    межгрупповая 162

Е

Единица
    измерения 14

З

Зависимости
    статистические 129
Зависимость
    корреляционная 129

И

Идентификация 186
Индекс
    корреляции
        множественной 176
Интервал
    доверительный 258
Интерполирование 130
    линейное 131
    параболой 132
Испытания
    статистические 128

К

Ковариация
    выборочная 151
Колебания
    сезонные 217
Комплекс
    степенной 16
Компонента
    сезонная 197
    случайная 197
    циклическая 197
Коррелограмма 212
Корреляционная
    зависимость 140
    таблица 136
Корреляционное
    отношение 154

Корреляция
множественная 163
Коэффициент
автокорреляции 212
детерминации 155, 167
дисконтирования 208
корреляции 152, 155
частный 176
корреляции совокупный 165
прироста 216
регрессии 151
роста 216
Кривая
Гомперца 242
логистическая 243
полиномиальная 239
регрессии 140
экспоненциальная 241
Критерий
Дарбина–Уотсона 250
Ирвина 199
Стьюдента 156
Фишера 155, 168, 177

Л

Лаг
временной 212

М

Медиана 248
Месштаб 24
Метод
Ирвина 199
Фостера–Стюарта 203
Четверикова 219
включения 273
двухшаговый 187
исключения 273
итераций 71
конечных разностей 243
наименьших квадратов 140, 187
последовательных приближений 71
скользящей средней 206
характеристик прироста 244
экспоненциального сглаживания 208
Моделирование
аналогичное 20
стохастическихкое 128
Модель 9
Брандона 270
Брауна 263
авторегрессионная 262
адаптивная 262
аналогичная 11
двухфакторная 176
детерминированная 12
логическая 11
математическая 11
материальная 11
регрессионная 271
статистическая 270
стохастическая 12
трендовая 198
условная 11
физическая 30

О

Оригинал 9
Отклонение
среднее квадратичное 199
Ошибка
аппроксимации 155
второго рода 198
первого рода 198
случайная 156
средняя квадратическая 258

П
Переменные
лаговые 186
экзогенные 186
эндогенные 186
Плоскость
регрессии 163
Погрешность
абсолютная 70
истинная 70
относительная 70
Подобие 23
Признак
результативный 174
факторный 271
Прирост
средний абсолютный 216
Прогноз
интервальный 258
Прогнозирование 217, 258

Р
Размер
объекта 15
Размерность 17
независимая 33
Разность
конечная 110
Ряд
временной 196
динамический 196
интервальный 196
моментный 196
тренд-сезонный 217

С
Сглаживание
ряда 219
Серия 248
Система уравнений
взаимосвязанных 185
независимых 185
нормальная 142, 146, 164
рекурсивных 185
Системы
взаимозависимые 271
рекурсивные 272
Сплайн-функция 133
Сходственные
переменные 19
постоянные 19
уравнения 20
функции 19

Т
Таблица
корреляционная 134
Темп
роста
средний 216
Точность
вычисления 72
Тренд 197

У
Уравнение
критериальное 35
линейное
разностное 111
определяющее 35
прямой регрессии 149
разностное 110
регрессии 140
Уравнения
масштабные 28
Уровень
аномальный 198
ряда средний 210
Ускорение роста 240

Ф
Фильтрация ряда 217
Форма
  приведенная 186
  структурная 185
Формула
  Брауна 209
  прямоугольников 103
  рекуррентная 71, 176
  трапеций 104
Функция
  корреляционная 212
  унимодальная 77

Э
Экстраполирование 130, 219
Эксцесс 249

ЛИТЕРАТУРА

1. *Айвазян С. А., Мхитарян В. С.* Прикладная статистика и основы эконометрики. — М.: ЮНИТИ, 1998, 1024 с.

   Учебник охватывает всю проблематику вероятностно-статистического моделирования.
2. *Бартини Р.* Соотношение между физическими величинами. — Докл. АН СССР, 163, № 4, 1965.

   Обсуждаются проблемы размерностей применительно к пространственно-временным многомерным процессам во вселенной, гравитации, авиации, оптике и т. д. Подробно обсуждается подход Брауна к описанию всех физических явлений двумя показателями размерности: длиной и временем.
3. *Блехман И. И., Мышкис А. Д., Пановко Н. Г.* Прикладная математика. Предмет, логика, особенности подходов. — М.: УРСС, 2006. — 376 с.

   Общие подходы к проблеме математического моделирования. Примеры из механики.
4. *Бронштейн И. Н. Семендяев К. А.* Справочник по математике для инженеров и учащихся втузов. — М.: Наука, 1980, — 978 с.

   Широко известный и многократно переиздаваемый справочник по математике. Охватывает все используемые на практике разделы математики.
5. Введение в математическое моделирование. Учебное пособие. Под ред. П. В. Трусова – М.: Логос, 2004. — ISBN 5-94010-272-7.

   Рассмотрены основные понятия и подходы к математическому моделированию и технологии построения математических моделей. Приведены примеры.
6. *Горлач Б. А.* Математика (учебный комплекс). — М.: ЮНИТИ, 2006, — 913 с. — ISBN 978-5-8114-1428-4.

Учебное пособие содержит теоретический материал, разработку семинарских занятий, задания для самостоятельной работы, вопросы для тестирования и типовые контрольные работы. Материал пособия включает разделы: линейная алгебра, математический анализ, теория вероятностей и математическая статистика.

7. *Гусева Е. Н.* Экономико-математическое моделирование. — М.: Флинта, МПСИ, 2008. — 216 с. — ISBN 978-5-89349-976-6.

   Изложены основные методы и приемы проектирования информационных экономических моделей. Приведен комплекс лабораторных работ.

8. *Елисеева И. И.* Эконометрика. — М.: «Финансы и статистика», 2003. — 344 с.

   Учебник, написанный на базе многолетнего преподавания курса эконометрики в Санкт-Петербургском государственном университете экономики и финансов. Много материала последних глав заимствовано из этого учебника.

9. *Замков О. О., Толстопятенко А. В., Черемных Ю. Н.* Математические методы в экономике. — М.: Изд. МГУ, «ДИС», 1997,— 366 с.

   Учебник, рассчитанный на студентов университетов, охватывает, в основном, вопросы применения математического анализа и линейной алгебры к решению задач экономики.

10. *Карасев А. И., Кремер Н. Ш., Савельева Т. И.* Математические методы и модели в планировании. — М.: Экономика, 1987. — 420 с.

    Традиционные для математического программирования вопросы изложены доходчиво и доступно. Приведены примеры решения многих прикладных задач экономики.

11. Математические методы принятия решений в экономике /Под редакцией В. А.Колемаева. — М.: ЗАО «Финстатинформ», 1999, — 386 с.

    На достаточно высоком уровне изложены основные вопросы теории исследования операций. Теоретические положения сопровождаются строгими доказательствами их справедливости. Изложение сопровождается примерами решения задач. Приведены вопросы и условия задач для самостоятельного решения. Рекомендуется для углубленного изучения материала.

12. *Краснощеков П. С., Петров А. А.* Принципы построения моделей — М.: ФАЗИС; ВЦ РАН, 2000, — 412 с. (Математическое моделирование; вып 1) – ISBN 5-7036-0061-8.

Приводятся и обсуждаются модели, описывающие физические явления, технические системы и коллективное поведение людей.

13. *Лебедев А. Н.* Моделирование в научно-технических исследованиях. — М.: Радио и связь, 1989. — 224 с.

Часть материала первых двух глав заимствовано из этой книги.

14. *Мышкис А. Д.* Элементы теории математических моделей. — М.: КомКнига, 2007. — 192 с. ISBN 978-5-484-00953-4.

Обсуждаются вопросы построения уравнений, описывающих поведение реальных процессов и явлений, их уточнений и упрощений. Приводятся примеры построения моделей из физики и механики.

15. *Петров А. А., Поспелов И. Г., Шананин А. А.* Опыт математического моделирования экономики. — М.: Энергоатомиздат, 1996. — 544 с. — ISBN 5-7036-0061-8.

Рассмотрены вопросы построения математических моделей экономики, в частности, модель перестройки и реформы российской экономики.

16. *Самарский А. А., Михайлов А. П.* Математическое моделирование. Идеи. Методы. Примеры. — М.: Физматлит, 2001. — 320 с. — ISBN 5-9221-0120-X.

Описаны общие подходы к построению математических моделей изучаемых объектов. Приведены примеры построения математических моделей механики, физики, биологии, экономики, социологии на основе законов природы, вариационных принципов, метода аналогий.

17. *Советов Б. Я., Яковлев С. А.* Моделирование систем. — М.: Высш. шк., 2001. — 343 с. — ISBN 5-06-003860-2.

Изложены фундаментальные основы теории моделирования, введены понятия компьютерной имитации, приведены примеры имитации явлений природы и общества, статистических систем, рассмотрена интеллектуальная система моделирования.

18. Справочник по математике для экономистов /Под ред. В. И. Ермакова. — М.: Высшая школа, 1997, — 384 с.

Основные разделы математики, применяемые в современной экономической науке, представлены соответствующими формулами и описанием особенностей их применения, примерами применения, а порой и выводами. Справочник полезен для постоянного использования в практической работе по использованию математики для решения экономических задач.

19. *Фомин Г. П.* Математические методы и модели в коммерческой деятельности. — М.: «Финансы и статистика», 2002, — 544 с.

Обсуждаются основные положения исследования операций и описываются алгоритмы решения задач математического программирования. Приводится решения многих экономических задач. Предлагаются вопросы и задачи для самостоятельного решенияю. Книга полезна для приобретения навыков решения конкретных экономических задач.

20. Экономико-математические методы и прикладные модели /Под редакцией В. В. Федосеева. — М.: Высшая школа, 1999, — 392 с.

С использованием современного подхода рассмотрены характерные математические модели экономических процессов, в том числе базирующиеся на использовании элементов линейной алгебры и математического программирования.

Полезен для углубления знаний по математическому моделированию в экономике.

ОГЛАВЛЕНИЕ

Предисловие 3

Введение 5

Глава 1. Моделирование объектов 9
§ 1.1. Основные понятия и определения 9
§ 1.2. Классификация моделей 10
§ 1.3. Условное моделирование 13
§ 1.4. Размеры и числовые значения 14
§ 1.5. Основные и производные единицы измерения . . . 15
§ 1.6. Размерные и безразмерные величины 17
§ 1.7. Аналогия. Аналогичное моделирование 18
§ 1.8. Резюме 21
§ 1.9. Вопросы 22

Глава 2. Математическое и физическое моделирование 23
§ 2.1. Подобие 23
§ 2.2. Подобие степенных комплексов 24
§ 2.3. Подобное моделирование 29
§ 2.4. Физическое моделирование. Модельный эксперимент 30
§ 2.5. Определяющие зависимости. π-теорема 33
§ 2.6. Резюме 37
§ 2.7. Вопросы 37
§ 2.8. Упражнения к разделу 2.1 38
§ 2.9. Упражнения к разделам 2.2, 2.3 44
§ 2.10. Упражнения к разделам 2.4, 2.5 53
§ 2.11. Задачи для самостоятельного решения 66

Глава 3. Численные методы 69
§ 3.1. Приближения и ошибки 70
§ 3.2. Итерационные методы 71
§ 3.3. Паутинная модель 74
§ 3.4. Выбор и исключение интервалов 77
§ 3.5. Метод золотого сечения 80
§ 3.6. Метод хорд 85
§ 3.7. Метод касательных 87

§ 3.8. Метод хорд и касательных 89
§ 3.9. Симплексный метод поиска экстремума 93
§ 3.10. Метод градиентного спуска 98
§ 3.11. Приближенное вычисление интегралов. Формулы прямоугольников и трапеций 102
§ 3.12. Формула Симпсона приближенного вычисления определенного интеграла 104
§ 3.13. Пример численного определения площади 106
§ 3.14. Разностные уравнения 110
§ 3.15. Численные методы решения задачи Коши 114
§ 3.16. Резюме 117
§ 3.17. Вопросы 119
§ 3.18. Расчетная работа, части 1–5 125

Глава 4. Теория корреляции 128
§ 4.1. Статистическая и корреляционная зависимости . . 129
§ 4.2. Интерполирование функций 130
§ 4.3. Пример корреляционной зависимости 133
§ 4.4. Корреляционная таблица 136
§ 4.5. Метод наименьших квадратов 140
§ 4.6. Линейная аппроксимация 141
§ 4.7. Квадратичная аппроксимация 145
§ 4.8. Уравнения линейной регрессии 148
§ 4.9. Коэффициент корреляции и корреляционные отношения 152
§ 4.10. Оценка качества и анализ уравнений регрессии . . 154
§ 4.11. Построение корреляционных соотношений и уравнений регрессии (пример) 157
§ 4.12. Плоскости регрессии 162
§ 4.13. Примеры решения задач парной регрессии 166
§ 4.14. Множественная регрессия и корреляция 174
§ 4.15. Реализация моделей множественной регрессии . . 177
§ 4.16. Системы взаимосвязанных уравнений 185
§ 4.17. Пример решения систем уравнений 187
§ 4.18. Резюме 191
§ 4.19. Вопросы 192
§ 4.20. Задание на расчетную работу, часть 6 194

Глава 5. Динамика случайных процессов 196
§ 5.1. Временные ряды 196
§ 5.2. Предварительный анализ временных рядов 198
5.2.1. Метод Ирвина 199
5.2.2. Метод проверки разностей средних уровней 201
5.2.3. Метод Фостера–Стюарта 203
§ 5.3. Сглаживание временных рядов 206

5.3.1. Метод скользящей средней 206
5.3.2. Метод экспоненциального сглаживания . . 208
§ 5.4. Характеристики временного ряда 210
§ 5.5. Расчет показателей динамики временного ряда . . 215
§ 5.6. Тренд-сезонные процессы и их анализ 217
§ 5.7. Метод фильтрации Четверикова 219
§ 5.8. Пример фильтрации ряда 222
§ 5.9. Резюме 235
§ 5.10. Вопросы 236
§ 5.11. Задание на расчетную работу, часть 7 237

Глава 6. Прогнозирование 239
§ 6.1. Трендовые модели одномерных процессов 239
§ 6.2. Выбор аппроксимирующей кривой 243
6.2.1. Метод конечных разностей 243
6.2.2. Метод характеристик прироста 244
§ 6.3. Адекватность трендовых моделей 247
6.3.1. Проверка случайности характера уровней последовательности 248
6.3.2. Проверка соответствия распределения случайной компоненты нормальному закону 249
6.3.3. Проверка равенства нулю математического ожидания случайной компоненты 249
6.3.4. Проверка независимости значений уровней случайной компоненты 250
§ 6.4. Точность трендовых моделей 251
§ 6.5. Пример исследования трендовой модели на адекватность и точность 252
§ 6.6. Прогнозирование на основе трендовых моделей . . 257
§ 6.7. Адаптивные модели прогнозирования 262
§ 6.8. Регрессионные модели прогнозирования 268
§ 6.9. Обобщение статистических моделей 270
§ 6.10. Резюме 273
§ 6.11. Вопросы 274
§ 6.12. Задание на расчетную работу, часть 8 274

Приложения 276

Предметный указатель 281

Литература 285

Борис Алексеевич ГОРЛАЧ,
Валентин Гаврилович ШАХОВ

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПОСТРОЕНИЕ МОДЕЛЕЙ И ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ

Учебное пособие

Зав. редакцией физико-математической
литературы *Н. Р. Крамор*
Верстка *А. Г. Сандомирская*
Выпускающий *Н. А. Крылова*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lanbook.ru; www.lanbook.com
196105, Санкт-Петербург, пр. Ю. Гагарина, д. 1, лит. А.
Тел./факс: (812) 336-25-09, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

ГДЕ КУПИТЬ

ДЛЯ ОРГАНИЗАЦИЙ:

Для того, чтобы заказать необходимые Вам книги, достаточно обратиться в любую из торговых компаний Издательского Дома «ЛАНЬ»:

по России и зарубежью
«ЛАНЬ-ТРЕЙД». 192029, Санкт-Петербург, ул. Крупской, 13
тел.: (812) 412-85-78, 412-14-45, 412-85-82; тел./факс: (812) 412-54-93
e-mail: trade@lanbook.ru; ICQ: 446-869-967
www.lanpbl.spb.ru/price.htm

в Москве и в Московской области
«ЛАНЬ-ПРЕСС». 109263, Москва, 7-я ул. Текстильщиков, д. 6/19
тел.: (499) 178-65-85; e-mail: lanpress@lanbook.ru

в Краснодаре и в Краснодарском крае
«ЛАНЬ-ЮГ». 350901, Краснодар, ул. Жлобы, д. 1/1
тел.: (861) 274-10-35; e-mail: lankrd98@mail.ru

ДЛЯ РОЗНИЧНЫХ ПОКУПАТЕЛЕЙ:

интернет-магазин
Издательство «Лань»: http://www.lanbook.com

магазин электронных книг
Global F5: http://globalf5.com/

Подписано в печать 25.02.16.
Бумага офсетная. Гарнитура Школьная. Формат 84×108 $^{1}/_{32}$.
Печать офсетная. Усл. п. л. 15,33. Тираж 200 экз.

Заказ № 046-16.

Отпечатано в полном соответствии
с качеством предоставленного оригинал-макета
в ПАО «Т8 Издательские Технологии».
109316, г. Москва, Волгоградский пр., д. 42, к. 5.