МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Тверской государственный технический университет»
(ТвГТУ)

А.Н. Балашов, Л.М. Пиджакова, М.А. Шестакова

РЕШЕНИЕ ПРИКЛАДНЫХ ЗАДАЧ АНАЛИТИЧЕСКИМИ И ЧИСЛЕННЫМИ МЕТОДАМИ

Учебное пособие

Тверь 2016

УДК [519.22+519.8+519.63]:004.4(075.8)
ББК [22.18+22.172+22.193]я7

Рецензенты: доктор физико-математических наук профессор кафедры программного обеспечения вычислительной техники ТвГТУ Калабин А.Л.; декан математического факультета, кандидат физико-математических наук, доцент кафедры математических методов современного естествознания ТвГУ Чемарина Ю.В.

Балашов, А.Н. Решение прикладных задач аналитическими и численными методами / А.Н. Балашов, Л.М. Пиджакова, М.А. Шестакова. Тверь: Тверской государственный технический университет, 2016. 160 с.

Изложены математические методы и модели решения широкого класса прикладных задач на основе традиционных методов математического программирования: линейного, целочисленного, динамического и теории игр. Рассмотрены основные разделы математической статистики. Приведены примеры решения задач с использованием современного пакета прикладных программ MS Excel. Изложены основы численных методов, и представлена их реализация в MS Excel.

Предлагаемое издание не только содержит академические знания о методах решения прикладных задач, математической статистики и основам численных методов, но и дает возможность научиться самостоятельно решать эти задачи с использованием современных пакетов прикладных программ.

Предназначено для студентов машиностроительных, энергетических, строительных и гуманитарных специальностей.

ISBN

ОГЛАВЛЕНИЕ

Введение

Моделирование в научных исследованиях применялось еще в глубокой древности и постепенно охватывало все новые области знаний: техническое конструирование, строительство и архитектуру, астрономию, физику, химию, биологию и, наконец, общественные и экономические науки. Признание практически во всех отраслях современной науки принес методу моделирования XX век.

Создание электронно-вычислительных машин, обрабатывающих большой массив информации, позволило решать задачи, требующие огромного объема вычислений, и разработать численные методы решения новых задач. Появились новые математические дисциплины, в частности *математическое программирование* – раздел математики, занимающийся решением задач нахождения экстремумов функции нескольких переменных с ограничениями на область допустимых значений этих переменных. Термин «программирование» появился в связи с тем, что в английском языке слово «programming» означает планирование, составление планов или программ.

Наиболее разработанным разделом математического программирования является линейное программирование. В 1931 г. венгерский математик Б. Эгервари решил задачу линейного программирования, именуемую «проблемой выбора». Примененный им метод решения получил название «венгерский метод». В 1939 г. советский академик Л.В. Канторович, лауреат Нобелевской премии, нашел общий метод решения задач (метод разрешающих множителей), связанных с составлением оптимального плана при организации производственных процессов. В 1949 г. Л.В. Канторович и М.К. Гавурин разработали метод потенциалов, используемый при решении транспортных задач. В 1947 г. американский математик Дж. Данциг предложил основной метод решения задач линейного программирования – симплекс-метод. Термин «линейное программирование» появился в 1951 г. Наряду с разработкой методов решения задач линейного программирования большое внимание уделялось задачам нелинейного, квадратичного, динамического программирования, теории игр и т. д.

Организация производства требует не только наличия современных станков и оборудования, но и разработки новых технологических процессов и современных методов управления производством. Для наилучшего решения поставленных задач разрабатываются математические модели.

Термин «модель» широко используется в различных сферах человеческой деятельности и имеет множество смысловых значений. В качестве модели можно рассматривать совокупность формул, макет реального объекта, рисунок и т. д. Для каждого объекта возможно создание нескольких моделей. *Модель* – это такой материальный или мысленно представляемый объект, который в процессе исследования замещает объект-оригинал так, что его непосредственное изучение дает новые знания об объекте-оригинале. Под моделированием понимается процесс построения, изучения и применения моделей. Необходимость использования метода моделирования определяется тем, что непосредственное исследование тех или иных объектов либо невозможно, либо требует много времени и средств.

Этапы математического моделирования [1]

1. Конструируют модель исходного объекта.

Построение модели предполагает наличие определенных сведений об объекте. Любая модель воспроизводит оригинал в ограниченном виде, отображая лишь некоторые существенные характеристики исходного объекта.

2. Модель выступает как самостоятельный объект исследования (в частности, проводят модельные эксперименты).

3. Осуществляют перенос знаний с модели на оригинал, в результате чего формируются новые знания об исходном объекте.

4. Полученные с помощью модели знания проходят проверку на практике и используются как для улучшения построенной теории реального объекта, так и для его целенаправленного преобразования или управления им.

Моделирование представляет собой циклический процесс, за первым четырехэтапным циклом может последовать второй, третий и т. д. При этом знания об исследуемом объекте расширяются и упорядочиваются, а первоначально построенная модель постепенно совершенствуется.

Глава 1. Математическое программирование

§ 1. Линейное программирование

1.1. Математическая модель общей задачи линейного программирования

Математические методы и модели применяются для отыскания оптимальных решений в управлении. Всякое управление предполагает наличие цели, ради реализации которой осуществляется управление.

Появление достаточно мощной вычислительной техники, способной за короткое время переработать значительные массивы информации, а также развитие математических моделей способствовало преобразованию искусства принятия решения в науку, доступную каждому, кто овладел ее принципами и методологией.

Общая схема составления моделей задач принятия решений

1. Выделяют набор параметров, которые описывают процесс принятия решения. Выбор тех или иных численных значений для параметров набора эквивалентен принятию того или иного решения. Параметры эти принято называть параметрами управления, число их может быть весьма значительным.

2. Выделяют и четко формулируют цель, ради достижения которой принимается то или иное решение. Как правило, целевую установку процесса принятия решения представляют в виде некоторой зависящей от параметров управления функции, значения которой дают оценку качества принятого решения. Эту функцию называют целевой функцией задачи или показателем качества ее решения. Целевая функция дает возможность сравнивать два различных решения: если при одном наборе параметров значение целевой функции меньше, чем при другом (или больше, в зависимости от задачи), то одно решение будет предпочтительнее другого.

3. Поскольку в задачах принятия решений не любой набор параметров управления может быть реализован практически, важно выделить среди всех решений множество возможных решений.

Допустим, что предметом нашего исследования является некоторое предприятие, производящее продукцию одного вида. Назовем эту продукцию конечной. Производство конечной продукции осуществляется с помощью одного из заранее отработанных технологических способов. Каждый технологический способ характеризуется количеством конечной продукции, выпускаемой в единицу времени, и всех расходуемых при этом ресурсов.

Сформулируем условие задачи. Необходимо составить такой план использования различных технологических способов, при котором предприятие выпускало бы максимально возможное количество конечной продукции в пределах имеющихся исходных запасов.

Формализуем условия задачи в общем виде.

Пусть n – число отработанных технологических способов производства; j – j-й технологический способ производства $(j = \overline{1, n})$; m – число исходных ресурсов, необходимых для производства; i – i-й ресурс $(i = \overline{1, m})$; c_j – количество продукции, производимое по j-му технологическому способу в течение единицы времени; a_{ij} – расход i-го исходного ресурса за единицу времени при использовании j-го технологического способа; b_i – запас i-го исходного ресурса, которым располагает предприятие.

Составим математическую модель задачи:

1. Выберем параметры управления (то, что мы можем менять в процессе производства в рамках поставленной задачи). В рамках поставленной задачи мы не можем менять b_i, a_{ij}, c_j, m и n (см. выше). В процессе производства можно менять только время, в течение которого предприятие выпускало продукцию, используя определенный технологический способ, поэтому пусть x_j – время, в течение которого предприятие использовало j-й технологический способ производства. Следовательно, $x_1, x_2, \ldots, x_n$ – искомый набор параметров управления, а составленный план производства эквивалентен выбору конкретных значений параметров x_j.

2. Найдем целевую функцию. По условию задачи необходимо составить такой план производства, при котором предприятие выпускало бы максимально возможное количество конечной продукции. Следовательно, целевая функция должна отражать количество выпускаемой продукции:

$$f(x) = \sum_{j=1}^{n} c_j x_j = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n .$$

Для нахождения оптимального решения необходимо максимизировать целевую функцию $f(x) \to \max$.

3. Поскольку запасы ресурсов ограничены, то на параметры управления накладываются ограничения. Найдем множество допустимых решений. Расход i-го исходного ресурса при работе предприятия по плану $(x_1, x_2, \ldots, x_n)$ можно найти по формуле

$$a_{i1} x_1 + a_{i2} x_2 + \ldots + a_{in} x_n .$$

Расход i-го ресурса не должен превышать имеющихся запасов b_i, т. е. $a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n = \sum_{j=1}^{n} a_{ij}x_j \le b_i$. Необходимо добавить еще ограничения $x_j \ge 0$, так как время использования технологических способов не может быть отрицательным.

В результате мы имеем математическую модель задачи:

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max,$$

$$\begin{cases} \sum_{j=1}^{n} a_{1j}x_j \le b_1,\ \sum_{j=1}^{n} a_{2j}x_j \le b_2,\ ...,\ \sum_{j=1}^{n} a_{mj}x_j \le b_m, \\ x_j \ge 0,\ j = \overline{1, n}. \end{cases}$$

Целевая функция модели и ограничения линейно зависят от параметров управления. Подобные модели принято называть *задачами линейного программирования* (ЗЛП). Зависимости редко бывают строго линейными, однако предположение о линейности дает возможность разрабатывать эффективные методы, использование которых для ряда задач нелинейного программирования является вполне оправданным.

Линейные модели могут содержать в качестве ограничений не только неравенства, но и уравнения.

Общая задача линейного программирования – это определение оптимального значения целевой функции:

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max\ (\min)$$

при ограничениях:

$$\begin{cases} \sum_{j=1}^{n} a_{ij}x_j \le b_i, & i = \overline{1, k}, \\ \sum_{j=1}^{n} a_{ij}x_j \ge b_i, & i = \overline{k+1, l}, \\ \sum_{j=1}^{n} a_{ij}x_j = b_i, & i = \overline{l+1, m}, \\ d_j \le x_j \le D_j, & j = \overline{1, n}, \end{cases}$$

где d_j и D_j – нижняя и верхняя граница переменной x_j соответственно.

Стандартная задача линейного программирования – это определение максимального значения целевой функции:

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max$$

при ограничениях в виде неравенств:

$$\begin{cases} \sum_{j=1}^{n} a_{ij} x_j \le b_i, \quad i = \overline{1, k}, \\ \sum_{j=1}^{n} a_{ij} x_j \ge b_i, \quad i = \overline{k+1, m}, \\ x_j \ge 0, \quad j = \overline{1, n}. \end{cases}$$

Каноническая задача линейного программирования – это определение максимального значения целевой функции:

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max$$

при ограничениях в виде равенств:

$$\begin{cases} \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = \overline{1, m}, \\ x_j \ge 0, \quad j = \overline{1, n}. \end{cases}$$

Совокупность значений переменных $x_1, x_2, ..., x_n$, удовлетворяющих всем линейным ограничениям, называется допустимым (опорным) планом.

Переход от стандартной модели ЗЛП к канонической осуществляется добавлением новых неотрицательных переменных со знаком «+» для неравенств типа $\le$ и со знаком «–» для неравенств типа $\ge$. Само неравенство заменяется на равенство.

Если в ЗЛП на некоторую переменную x_k не накладывается условие неотрицательности, то добавляют две неотрицательные переменные и делают замену переменных $x_k = x_{k_2} - x_{k_1}, x_{k_2} \ge 0, x_{k_1} \ge 0$.

При переходе от задачи на минимум к задаче на максимум изменяется знак целевой функции.

Задача линейного программирования может принимать оптимальное значение в нескольких точках, такие решения называются *альтернативными оптимальными решениями* или альтернативным оптимумом, причем в каждой из этих точек целевая функция имеет одно и то же оптимальное значение.

Если в ЗЛП с двумя параметрами управления нормальный вектор целевой функции параллелен нормальному вектору прямой одного из ограничений, то ЗЛП может иметь множество оптимальных решений, например:

$$f(x) = 2x_1 + 3x_2 \to \max,$$

$$\begin{cases} 4x_1 + 6x_2 \le 15, \\ 7x_1 + 5x_2 \le 20, \\ x_1 \ge 0, x_2 \ge 0. \end{cases}$$

Вектор целевой функции $\overrightarrow{N}(2;3)$ параллелен нормальному вектору $\overrightarrow{N_1}(4;6)$ прямой первого ограничения: $\overrightarrow{N_1} = 2\overrightarrow{N}$. В этом случае возможно множество оптимальных решений.

Если в ЗЛП с тремя параметрами управления нормальный вектор целевой функции параллелен нормальному вектору плоскости одного из ограничений или перпендикулярен к прямой пересечения двух плоскостей ограничений, то ЗЛП может иметь множество оптимальных решений.

Для ЗЛП с большим количеством параметров управления сложно определить, имеется ли множество оптимумов по виду составленной математической модели.

Пример 1.1. Составить математическую модель задачи, привести ее к каноническому виду.

Предприятие при выпуске однотипной продукции может использовать три технологических способа производства (ТСП). Расходы, связанные с каждым технологическим способом, определяются количеством трудозатрат (человеко-часов), электроэнергии (киловатт-часов) и потребляемого материала (кг). Исходные данные задачи:

Расходы ресурсов:	На единицу продукции			Имеется в наличии
	ТСП1	ТСП2	ТСП3	
человеко-часов	10	15	10	320
киловатт-часов	18	16	20	500
материала, кг	70	50	60	2000
Доход с единицы продукции, ден. ед.	50	60	70	max

Определить объем выпускаемой продукции, обеспечивающий наибольший доход.

Решение

1. В процессе производства можно менять количество выпускаемой продукции по определенному ТСП. Поэтому в качестве параметров управления x_1, x_2 и x_3 выберем количество выпускаемой продукции по соответствующему технологическому способу производства ТСП1, ТСП2 и ТСП3.

2. Целевая функция должна максимизировать доход от выпускаемой продукции: $f(x) = 50x_1 + 60x_2 + 70x_3 \to \max$.

3. Ограничивают количество выпускаемой продукции запасы ресурсов. В процессе производства мы не должны превысить количество имеющихся человеко-часов: $10x_1 + 15x_2 + 10x_3 \le 320$. Аналогично для двух других ресурсов: $18x_1 + 16x_2 + 20x_3 \le 500$, $70x_1 + 50x_2 + 60x_3 \le 2000$. Необходимо добавить условие неотрицательности переменных: $x_1 \ge 0$, $x_2 \ge 0$, $x_3 \ge 0$.

В результате имеем математическую модель:

$$f(x) = 50x_1 + 60x_2 + 70x_3 \to \max,$$

$$\begin{cases} 10x_1 + 15x_2 + 10x_3 \le 320, \\ 18x_1 + 16x_2 + 20x_3 \le 500, \\ 70x_1 + 50x_2 + 60x_3 \le 2000, \\ x_j \ge 0,\ j = \overline{1, 3}. \end{cases}$$

Решение с помощью MS Excel [2, Лист 1]:

$f(x)$ = 1790 руб., $x_1 = 0$, $x_2 = 10$, $x_3 = 17$.

Это стандартная ЗЛП. Получим из нее каноническую ЗЛП, добавив в первое, второе и третье неравенства неотрицательные переменные x_4, x_5, x_6:

$$f(x) = 50x_1 + 60x_2 + 80x_3 \to \max,$$

$$\begin{cases} 10x_1 + 15x_2 + 10x_3 + x_4 = 320, \\ 18x_1 + 16x_2 + 20x_3 + x_5 = 500, \\ 70x_1 + 50x_2 + 60x_3 + x_6 = 2000, \\ x_j \ge 0,\ \ j = \overline{1,\ 6}. \end{cases}$$

Пример 1.2. Для изготовления изделий двух видов имеется 400 кг металла. На одно изделие первого вида расходуется 2 кг металла, а на изделие второго вида – 5 кг. Составить план производства, обеспечивающий наибольшую стоимость выпускаемых изделий, если отпускная цена одного изделия первого вида составляет 4 ден. ед., а изделия второго вида – 3 ден. ед., причем изделий каждого вида требуется изготовить не менее 50 и 20 соответственно.

Решение

1. В качестве параметров управления x_1 и x_2 выберем соответственно количество изготовленных изделий первого и второго вида.

2. Целевая функция должна максимизировать прибыль, получаемую от реализации изделий, поэтому $f(x) = 4x_1 + 3x_2 \to \max$.

3. Число выпускаемых изделий в первую очередь ограничено количеством металла (400 кг), которое мы не должны превысить в процессе производства: $2x_1 + 5x_2 \le 400$. Наличие плана также накладывает ограничение. Изделий первого и второго вида необходимо изготовить не менее 50 и 20 соответственно, т. е. $x_1 \ge 50$, $x_2 \ge 20$. Количество изготовленных изделий принадлежит множеству неотрицательных целых чисел. Необходимо добавить условие неотрицательности и целочисленности переменных: $x_{1,\,2} \in N_0$, где N_0 – множество целых неотрицательных чисел, хотя в данной задаче ограничение не-

отрицательности излишне, так как ограничения $x_1 \ge 50,\ \ x_2 \ge 20$ являются более строгими, чем условие неотрицательности переменных.

В итоге имеем математическую модель задачи:

$$f(x) = 4x_1 + 3x_2 \to \max,$$

$$\begin{cases} 2x_1 + 5x_2 \le 400, \\ x_1 \ge 50, \\ x_2 \ge 20, \\ x_{1,2} \in N_0. \end{cases}$$

Решение с помощью MS Excel [2, Лист 2]:

$$f(x) = 660 \text{ ден. ед.}, x_1 = 150, x_2 = 20.$$

1.2. Графический метод решения

Графические методы решения более наглядны, но их применение ограничено числом неизвестных переменных, которых должно быть не более трех. Только в этом случае решение задачи можно наглядно представить на бумаге.

Рассмотрим графическое решение примера 1.2.

Определим множество решения первого линейного неравенства $2x_1 + 5x_2 \le 400$. На координатной плоскости оно отображается полуплоскостью, ограниченной прямой $2x_1 + 5x_2 = 400$. Для построения прямой необходимо найти две точки, принадлежащие прямой. Обычно их находят путем последовательного обнуления каждой из переменных. Если $x_1 = 0$, то $x_2 = 400/5 = 80$. Если $x_2 = 0$, то $x_1 = 400/2 = 200$. Часто именно точки пересечения прямой с координатными прямыми являются оптимальными решениями. В системе координат Ox_1x_2 построим прямую по двум точкам (0; 80) и (200; 0). Построенная прямая делит координатную плоскость на две полуплоскости. Для определения искомой полуплоскости достаточно выбрать одну точку, не лежащую на прямой. Пусть это будет начало координат (0; 0). Подставив координаты выбранной точки в неравенство, получим: $0 \le 400$. Неравенство верное, следовательно, нижняя полуплоскость, которой принадлежит начало координат, является искомой. На рис. 1.1*а* эта полуплоскость заштрихована.

Аналогичная процедура выполняется для второго неравенства. Выберем точку (0; 0) и, подставив во второе неравенство, получим: $0 \ge 50$. Неравенство неверное, следовательно, выбираем правую полуплоскость, которой не принадлежит начало координат (рис. 1.1*б*).

Для третьего неравенства выбираем точку (0; 0) и, подставив в третье неравенство, получим: $0 \ge 20$. Неравенство неверное, следовательно, выбираем верхнюю полуплоскость, которой не принадлежит начало координат (рис. 1.1*в*).

В итоге выбираем общую заштрихованную область, точки которой удовлетворяют всем неравенствам задачи (рис. 1.1*г*). Область допустимых решений, или область *Парето*, изображается в виде замкнутого или незамкнутого многоугольника.

Оптимальное решение ЗЛП находится в одной из вершин многоугольника области Парето.

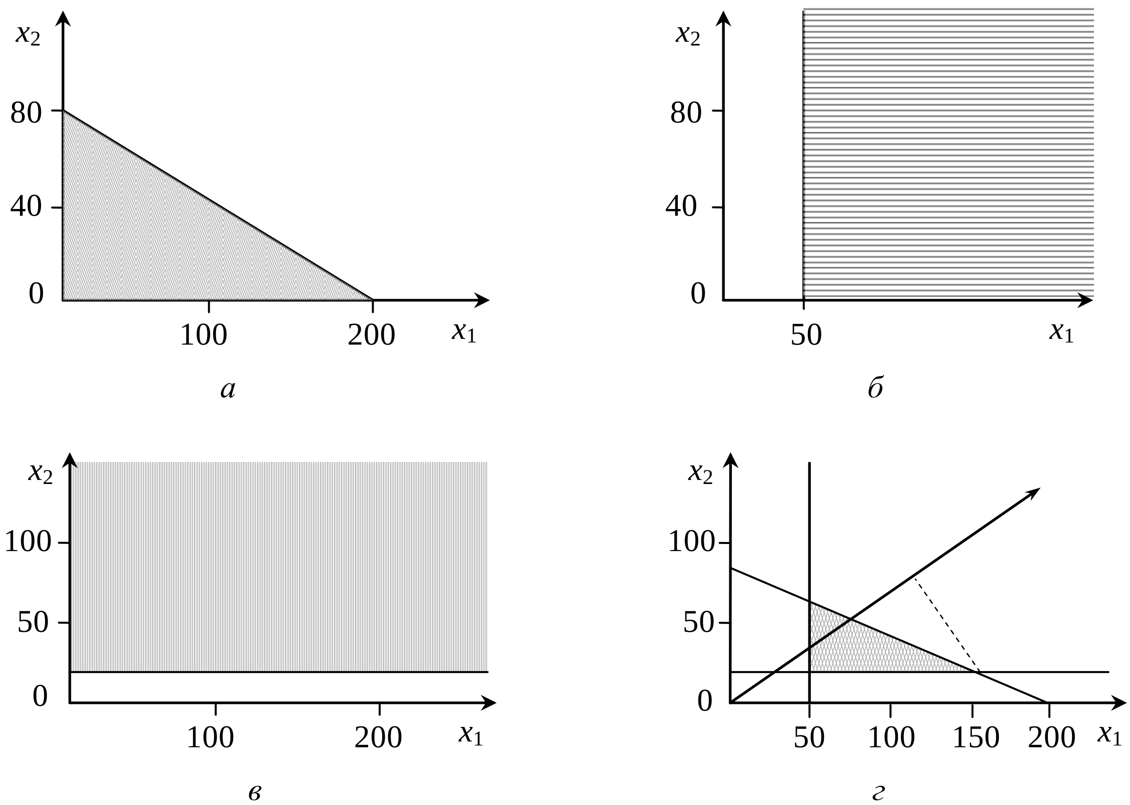


Рис. 1.1

Возможны два пути оптимального решения графическим способом:

1) находят координаты всех вершин многоугольника области Парето и подставляют их в целевую функцию, после чего выбирают оптимальное значение (наибольшее или наименьшее);

2) строят нормальный вектор прямой целевой функции, проецируют область Парето на вектор или его продолжение (см. рис. 1.1*г*). Наиболее или наименее удаленная от начала координат проекция вершины области Парето на вектор или его продолжение будет оптимальным решением.

Для нахождения трех вершин необходимо решить три системы уравнений:

$$\begin{cases} x_1 = 50, \\ x_2 = 20, \end{cases} \quad \begin{cases} x_1 = 50, \\ 2x_1 + 5x_2 = 400, \end{cases} \quad \begin{cases} x_2 = 20, \\ 2x_1 + 5x_2 = 400. \end{cases}$$

В итоге имеем: $A(50;\, 20)$, $B(50;\, 60)$, $C(150;\, 20)$.

Находим значения целевой функции в вершинах многоугольника Парето:

$$F(50, 20) = 260,\ F(50, 60) = 380,\ F(150, 20) = 660.$$

Наибольшая стоимость выпускаемой продукции составит 660 ден. ед. при изготовлении 150 изделий первого и 20 изделий второго вида.

При решении вторым способом находим координаты только одной вершины $C(150;\ 20)$, поскольку проекция этой вершины на продолжение вектора целевой функции $\vec{N}(4; 3)$ будет наибольшей (см. рис. 1.1*г*). Второй способ применяется при использовании одинаковых масштабов на осях координат. В противном случае линия проекции на вектор целевой функции визуально не будет составлять 90°. Если не удается наглядно изобразить область Парето в системе координат с одинаковым масштабом осей координат, то поступают следующим способом. В системе координат с разным масштабом координатных осей рисуют линию уровня целевой функции $f(x) = 0$ или $f(x) = \text{const}$ и перемещают ее параллельно в направлении вектора целевой функции до пересечения с областью Парето.

Пример 1.3. Для изготовления изделий двух видов *A* и *B* используется токарное и фрезерное оборудование. Затраты времени на обработку одного изделия для каждого из типов оборудования, общий фонд рабочего времени каждого из типов используемого оборудования, а также прибыль от реализации одного изделия каждого вида составляют:

Тип оборудования:	Затраты времени на обработку одного изделия вида		Общий фонд рабочего времени оборудования, ч
	A	*B*	
фрезерное, ч	2	5	200
токарное, ч	4	2	160
Прибыль, ден. ед.	10	24	max

Требуется определить, сколько изделий следует изготовить предприятию, чтобы прибыль от их реализации была максимальной.

Решение

1. В качестве параметров управления x_1 и x_2 выберем количество изготовленных изделий первого и второго вида соответственно.

2. Целевая функция должна максимизировать прибыль, получаемую от реализации изделий, поэтому $f(x) = 10x_1 + 24x_2 \to \max$.

3. Количество выпускаемых изделий ограничено общим фондом рабочего времени. В процессе производства мы не должны превысить общий фонд рабочего времени фрезерного оборудования $2x_1 + 5x_2 \le 200$ и

общий фонд рабочего времени токарного оборудования $4x_1 + 2x_2 \le 160$. Необходимо добавить условие неотрицательности и целочисленности переменных: $x_{1,\,2} \in N_0$, где N_0 – множество целых неотрицательных чисел.

Получим математическую модель задачи:

$$f(x) = 10x_1 + 24x_2 \to \max,$$

$$\begin{array}{l} (1) \\ (2) \\ \end{array}\begin{cases} 2x_1 + 5x_2 \le 200, \\ 4x_1 + 2x_2 \le 160, \\ \quad x_{1,\,2} \in N_0. \end{cases}$$

Изобразим область допустимых решений в системе координат Ox_1x_2.

Найдем две точки, лежащие на прямой $2x_1 + 5x_2 = 200$. Пусть $x_1 = 0$, тогда из данного уравнения имеем: $x_2 = 200/5 = 40$. Пусть $x_2 = 0$, тогда из уравнения имеем $x_1 = 200/2 = 100$. В итоге получим координаты двух точек прямой (1): $A(0;\ 40)$, $B(100;\ 0)$. В системе координат Ox_1x_2 построим прямую по точкам A и B (рис. 1.2).

Аналогично для другой прямой $4x_1 + 2x_2 = 160$. Пусть $x_1 = 0$, тогда $x_2 = 160/2 = 80$. Пусть $x_2 = 0$, тогда $x_1 = 160/4 = 40$. В итоге имеем координаты двух точек прямой (2): $C(0;\ 80)$, $D(40;\ 0)$. В системе координат Ox_1x_2 построим прямую по точкам C и D (рис. 1.2).

В результате получили три замкнутые области и одну незамкнутую. Для определения области допустимых решений возьмем контрольную точку с координатами (0; 0), принадлежащую только четырехугольнику *OAED*. Координаты этой точки удовлетворяют всем ограничениям. Следовательно, выбранный четырехугольник *OAED* является искомой областью Парето. Контрольной точкой может быть не только вершина многоугольника, но и любая внутренняя точка полученных многоугольников.

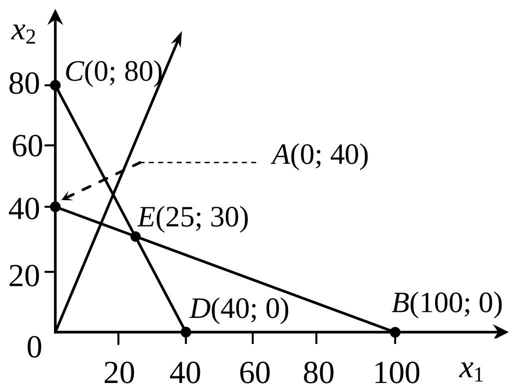


Рис. 1.2

Найдем координаты точки пересечения прямых (1) и (2) из системы уравнений:

$$\begin{cases} 2x_1 + 5x_2 = 200, \\ 4x_1 + 2x_2 = 160. \end{cases}$$

Разделим второе уравнение системы на два и вычтем из второго уравнения первое:

$$\begin{cases} 2x_1 + 5x_2 = 200, \\ -4x_2 = -120. \end{cases}$$

Разделим второе уравнение на (–4):

$$\begin{cases} 2x_1 + 5x_2 = 200, \\ x_2 = 30. \end{cases}$$

Подставим в первое уравнение численное значение $x_2 = 30$. Получим координаты точки *E*(25; 30) пересечения прямых *AB* и *CD*.

Координаты точек пересечения построенных прямых с осями координат найдены при построении прямых. Найдем значения целевой функции в вершинах четырехугольника *O*(0; 0), *A*(0; 40), *E*(25; 30) и *D*(40; 0):

$$F(0, 0) = 0,\ F(0, 40) = 960,\ E(25, 30) = 970,\ F(40, 0) = 400.$$

Наибольшая прибыль составит 970 ден. ед. при изготовлении 25 изделий типа *A* и 30 изделий типа *B*.

Проекции точек *A* и *E* на продолжение вектора целевой функции находятся очень близко друг от друга. При решении вторым способом (см. с. 13) необходимо найти значение целевой функции в обеих точках *A* и *E*.

1.3. Симплексный метод решения

В ЗЛП оптимальное значение целевой функции, если оно существует, достигается в вершинах многогранника, ограничивающего область Парето.

Симплексный метод представляет собой последовательность шагов (итераций) перехода от одной вершины многогранника области Парето к другой с целью получения максимального значения целевой функции. Любые ЗЛП могут быть решены симплексным методом.

Возьмем в качестве примера математическую модель предыдущей задачи без учета ограничений целочисленности переменных:

$$f(x) = 10x_1 + 24x_2 \rightarrow \max,$$

$$\begin{matrix} (1) \\ (2) \\ \\ \end{matrix} \begin{cases} 2x_1 + 5x_2 \le 200, \\ 4x_1 + 2x_2 \le 160, \\ x_1 \ge 0,\ x_2 \ge 0. \end{cases}$$

Добавим в 1-е и 2-е ограничения соответственно неотрицательные переменные x_3 и x_4 и запишем целевую функцию в неявном виде $F(x) = 0$:

$$f(x) - 10x_1 - 24x_2 = 0,$$
$$\begin{cases} 2x_1 + 5x_2 + x_3 = 200, \\ 4x_1 + 2x_2 + x_4 = 160, \\ x_j \geq 0, \ j = \overline{1,4}. \end{cases}$$

Каждый коэффициент целевой функции $F(x)$, если он отрицателен, определяет величину положительного приращения целевой функции $f(x)$, а если положителен – отрицательного приращения целевой функции $f(x)$ при увеличении соответствующей переменной. Если x_1 увеличить на единицу, то целевая функция $f(x)$ увеличится на 10 единиц, а если x_2 увеличить на единицу, то целевая функция $f(x)$ увеличится на 24 единицы. Переменная x_2 сильнее влияет на целевую функцию, чем переменная x_1.

На первом шаге симплексного метода решения канонической ЗЛП необходимо попасть в одну из вершин многогранника области Парето. Выражаем из m ограничений канонической ЗЛП m переменных и разрешаем систему из m уравнений относительно выбранных m переменных. Исключаем эти m переменных из выражения целевой функции. Выбранные m переменных называют базисными или основными, а остальные – небазисными, или неосновными, или свободными.

В нашей задаче в качестве базисных переменных удобно взять переменные x_3 и x_4. Они не входят в целевую функцию и относительно них легко разрешить систему двух уравнений:

$$f(x) - 10x_1 - 24x_2 = 0,$$
$$\begin{cases} x_3 = 200 - 2x_1 - 5x_2, \\ x_4 = 160 - 4x_1 - 2x_2, \\ x_j \geq 0, \ j = \overline{1,4}. \end{cases}$$

Первоначальное допустимое решение получается при нулевых значениях свободных переменных $x_1 = x_2 = 0$, $x_3 = 200$, $x_4 = 160$, $f(x) = 0$.

На втором шаге необходимо перейти в другую вершину многогранника области Парето, в которой целевая функция будет принимать большее значение. Аналитически это означает замену одной базисной переменной на другую, свободную переменную. Выбор новой базисной переменной осуществляют по следующей схеме.

Если в целевой функции имеются свободные переменные, коэффициенты при которых отрицательные, то желательно выбрать из них переменную с наибольшим по модулю значением коэффициента в качестве новой базисной переменной. Если все свободные переменные имеют неотрицательные коэффициенты, то оптимальное решение получено.

Свободная переменная x_2 имеет наибольший по модулю отрицательный коэффициент в целевой функции. Выберем в качестве новой базисной переменной x_2 и разрешим систему ограничений относительно базисных переменных x_2 и x_4, исключим новую базисную переменную x_2 из целевой функции. Эти преобразования удобно выполнять в виде симплекс-таблицы (табл. 1.1), которая составлена на основе канонической ЗЛП с целевой функцией $F(x)$, записанной в неявном виде.

Таблица 1.1

x_1	x_2	x_3	x_4	Свободные члены	Базис
2	5	1	0	200	x_3
4	2	0	1	160	x_4
–10	–24	0	0	0	$F(x)$

Столбцы при переменных x_3 и x_4 являются единичными. Единичный столбец – это столбец, все элементы которого равны нулю, кроме одного, равного единице. В симплекс-таблице количество единичных столбцов должно быть равно количеству базисных переменных. Все свободные члены должны быть неотрицательными. Неотрицательность свободных членов получают при составлении канонической ЗЛП.

Рассмотрим алгоритм ввода новой базисной переменной и исключения старой, реализованный симплекс-таблицей.

1. Выбираем наибольший по модулю отрицательный коэффициент целевой функции – это коэффициент при переменной x_2. Столбец l = 2 симплекс-таблицы, которому принадлежит новая базисная переменная x_2, называется ведущим столбцом.

2. Делим элементы столбца свободных членов на соответствующие элементы ведущего столбца, при этом учитываем только положительные элементы. Выбираем наименьшее частное $\min\left\{\frac{200}{5};\frac{160}{2}\right\}=\min\{40;80\}$. Строка k = 1, которой принадлежит наименьшее частное, называется ведущей строкой. На пересечении ведущего столбца и ведущей строки находится ведущий элемент $a_{12}=5$. Первая строка симплекс-таблицы не нумеруется.

3. Делим элементы ведущей строки на ведущий элемент:

$$a'_{kj}=\frac{a_{kj}}{a_{kl}},$$

где k – номер ведущей строки;

l – номер ведущего столбца;

a_{kl} – ведущий элемент.

Все остальные элементы симплекс-таблицы пересчитываем по формуле

$$a'_{ij} = a_{ij} - \frac{a_{kj}a_{il}}{a_{kl}} \quad \text{или} \quad a'_{ij} = a_{ij} - a'_{kj}a_{il}\,.$$

Элементы ведущего столбца, кроме ведущего элемента, зануляются:

$$a'_{il} = a_{il} - \frac{a_{kl}a_{il}}{a_{kl}} = 0\,.$$

Первая строка (ведущая):

$$a'_{11} = \frac{a_{11}}{a_{12}} = \frac{2}{5} = 0{,}4, \qquad a'_{12} = \frac{a_{12}}{a_{12}} = \frac{5}{5} = 1, \qquad a'_{13} = \frac{a_{13}}{a_{12}} = \frac{1}{5} = 0{,}2,$$

$$a'_{14} = \frac{a_{14}}{a_{12}} = \frac{0}{5} = 0, \qquad a'_{15} = \frac{a_{15}}{a_{12}} = \frac{200}{5} = 40\,.$$

Вторая строка:

$$a'_{21} = a_{21} - a'_{11}a_{22} = 4 - 0{,}4 \cdot 2 = 3{,}2,$$
$$a'_{22} = 0,$$
$$a'_{23} = a_{23} - a'_{13}a_{22} = 0 - 0{,}2 \cdot 2 = -0{,}4,$$
$$a'_{24} = a_{24} - a'_{14}a_{22} = 1 - 0 \cdot 2 = 1,$$
$$a'_{25} = a_{25} - a'_{15}a_{22} = 160 - 40 \cdot 2 = 80\,.$$

Третья строка:

$$a'_{31} = a_{31} - a'_{11}a_{32} = -10 - 0{,}4 \cdot (-24) = -4{,}4,$$
$$a'_{32} = 0,$$
$$a'_{33} = a_{33} - a'_{13}a_{32} = 0 - 0{,}2 \cdot (-24) = 4{,}8,$$
$$a'_{34} = a_{34} - a'_{14}a_{32} = 0,$$
$$a'_{35} = a_{35} - a'_{15}a_{32} = 0 - 40 \cdot (-24) = 480\,.$$

В столбце «Базис» старую базисную переменную ведущей строки x_3 заменяем на новую базисную переменную ведущего столбца x_2 (табл. 1.2):

Таблица 1.2

x_1	x_2	x_3	x_4	Свободные члены	Базис
0,4	1	0,2	0	40	x_2
3,2	0	–0,4	1	80	x_4
–4,4	0	4,8	0	480	$F(x)$

Первый итерационный шаг симплексного метода завершен. Получили новую симплекс-таблицу (см. табл. 1.2) и новый опорный план (0; 40; 0; 80), $f(x)$ = 480.

Первый шаг соответствует переходу от одной вершины $O(0;\ 0)$ многоугольника области Парето к другой $A(0;\ 40)$ (см. рис. 1.2).

Коэффициент при переменной x_1 в целевой функции отрицателен, т. е. оптимальное значение еще не достигнуто. Повторяем шаги 1–3 еще раз.

1. Выбираем наибольший по модулю отрицательный коэффициент целевой функции – это коэффициент при переменной x_1. Столбец $l = 1$ симплекс-таблицы, которому принадлежит новая базисная переменная x_1, является ведущим столбцом.

2. Делим элементы столбца свободных членов на соответствующие элементы ведущего столбца, при этом учитываем только положительные элементы. Выбираем наименьшее частное $\min\left\{\frac{40}{0{,}4}; \frac{80}{3{,}2}\right\} = \min\{100; 25\}$. Строка $k = 2$, которой принадлежит наименьшее частное, является ведущей строкой. На пересечении ведущего столбца и ведущей строки находится ведущий элемент $a_{21} = 3{,}2$.

3. Пересчитываем элементы симплекс-таблицы по формулам, приведенным на первом шаге. Начинаем с ведущей строки.

Вторая строка (ведущая):

$$a'_{21} = \frac{a_{21}}{a_{21}} = 1,\ a'_{22} = \frac{a_{22}}{a_{21}} = \frac{0}{3{,}2} = 0,\ a'_{23} = \frac{a_{23}}{a_{21}} = \frac{-0{,}4}{3{,}2} = -0{,}125,$$

$$a'_{24} = \frac{a_{24}}{a_{21}} = \frac{1}{3{,}2} = 0{,}3125,\ a'_{25} = \frac{a_{25}}{a_{21}} = \frac{80}{3{,}2} = 25.$$

Первая строка:

$$a'_{11} = 0,$$
$$a'_{12} = a_{12} - a'_{22}a_{11} = 1 - 0 \cdot 0{,}4 = 1,$$
$$a'_{13} = a_{13} - a'_{23}a_{11} = 0{,}2 - (-0{,}125) \cdot 0{,}4 = 0{,}25,$$
$$a'_{14} = a_{14} - a'_{24}a_{11} = 0 - 0{,}3125 \cdot 0{,}4 = -0{,}125,$$
$$a'_{15} = a_{15} - a'_{25}a_{11} = 40 - 25 \cdot 0{,}4 = 30.$$

Третья строка:

$$a'_{31} = 0,$$
$$a'_{32} = a_{32} - a'_{22}a_{13} = 0 - 0 \cdot (-4{,}4) = 0,$$
$$a'_{33} = a_{33} - a'_{23}a_{13} = 4{,}8 - (-0{,}125) \cdot (-4{,}4) = 4{,}25,$$
$$a'_{34} = a_{34} - a'_{24}a_{13} = 0 - 0{,}3125 \cdot (-4{,}4) = 1{,}375,$$
$$a'_{35} = a_{35} - a'_{25}a_{13} = 480 - 25 \cdot (-4{,}4) = 590.$$

В столбце «Базис» прежнюю базисную переменную ведущей строки x_4 заменяем на новую базисную переменную ведущего столбца x_1 (табл. 1.3).

Таблица 1.3

x_1	x_2	x_3	x_4	Свободные члены	Базис
0	1	0,25	–0,125	30	x_2
1	0	–0,125	0,3125	25	x_1
0	0	4,25	1,375	590	$F(x)$

Второй шаг симплексного метода завершен. Получили новую симплекс-таблицу (см. табл. 1.3) и новый опорный план (25; 30; 0; 0), $f(x) = 590$. Второй итерационный шаг соответствует переходу от вершины $A(0; 40)$ многоугольника к другой его вершине $B(25; 30)$ (см. рис. 1.2).

Коэффициенты целевой функции неотрицательны, поэтому новый опорный план (25; 30; 0; 0), $f(x) = 590$ является оптимальным.

Рассмотрим решение примера 1.2 симплексным методом.

Запишем математическую модель в каноническом виде без учета целочисленности переменных x_1 и x_2, добавив три неотрицательные переменные x_3, x_4 и x_5:

$$f(x) - 4x_1 - 3x_2 = 0,$$

$$\begin{cases} 2x_1 + 5x_2 + x_3 = 400, \\ x_1 - x_4 = 50, \\ x_2 - x_5 = 20, \\ x_j \geq 0,\ j = \overline{1,5}. \end{cases}$$

Коэффициенты при переменных x_4 и x_5 отрицательны. При составлении симплекс-таблицы они не составят единичных столбцов, поэтому введем две искусственные неотрицательные переменные x_6 и x_7:

$$f(x) - 4x_1 - 3x_2 = 0,$$

$$\begin{cases} 2x_1 + 5x_2 + x_3 = 400, \\ x_1 - x_4 + x_6 = 50, \\ x_2 - x_5 + x_7 = 20, \\ x_j \geq 0,\ j = \overline{1,7}. \end{cases}$$

Искусственные переменные не входят в целевую функцию. Составим симплекс-таблицу (табл. 1.4).

Таблица 1.4

x_1	x_2	x_3	x_4	x_5	x_6	x_7	Свободные члены	Базис
2	5	1	0	0	0	0	400	x_3
1	0	0	–1	0	1	0	50	x_6
0	1	0	0	–1	0	1	20	x_7
–4	–3	0	0	0	–	–	0	$F(x)$

Выводим пошагово из базиса искусственные переменные x_6 и x_7. Геометрически это означает попадание в область Парето.

1. Выбираем вторую строку k = 2 при переменной x_6 в качестве ведущей.

2. Выбираем первый столбец l = 1 при переменной x_1 в качестве ведущего. На пересечении ведущего столбца и ведущей строки находится ведущий элемент $a_{21} = 1$.

3. Пересчитываем элементы симплекс-таблицы по формулам, приведенным в предыдущей задаче. Начинаем с ведущей строки.

Вторая строка (ведущая) остается неизменной:

$$a'_{21} = 1,\ a'_{22} = 0,\ a'_{23} = 0,\ a'_{24} = -1,\ a'_{25} = 0,\ a'_{26} = 1,\ a'_{27} = 0,\ a'_{28} = 50.$$

Первая строка (умножаем ведущую строку на два и результат вычитаем из первой строки):

$$a'_{11} = 0,\ a'_{12} = 5,\ a'_{13} = 1,\ a'_{14} = 2,\ a'_{15} = 0,\ a'_{16} = -2,\ a'_{17} = 0,\ a'_{18} = 300.$$

Третья строка остается неизменной:

$$a'_{31} = 0,\ a'_{32} = 1,\ a'_{33} = 0,\ a'_{34} = 0,\ a'_{35} = -1,\ a'_{36} = 0,\ a'_{37} = 1,\ a'_{38} = 20.$$

Четвертая строка (умножаем ведущую строку на четыре и результат прибавляем к четвертой строке):

$$a'_{41} = 0,\ a'_{42} = -3,\ a'_{43} = 0,\ a'_{44} = -4,\ a'_{45} = 0,\ a'_{48} = 200.$$

В столбце «Базис» прежнюю базисную переменную ведущей строки x_6 заменяем на новую базисную переменную ведущего столбца x_1. Вычеркиваем столбец выведенной искусственной переменной x_6 (табл. 1.5).

Таблица 1.5

x_1	x_2	x_3	x_4	x_5	x_6	x_7	Свободные члены	Базис
0	5	1	2	0		0	300	x_3
1	0	0	–1	0		0	50	x_1
0	1	0	0	–1		1	20	x_7
0	–3	0	–4	0		–	200	$F(x)$

Выводим из базиса искусственную переменную x_7.

1. Выбираем третью строку k = 3 при переменной x_7 в качестве ведущей.

2. Выбираем второй столбец l = 2 при переменной x_2 в качестве ведущего. На пересечении ведущего столбца и ведущей строки находится ведущий элемент $a_{32} = 1$.

3. Пересчитываем элементы симплекс-таблицы (см. табл. 1.5).

Третья строка (ведущая) остается неизменной:

$$a'_{31} = 0,\ a'_{32} = 1,\ a'_{33} = 0,\ a'_{34} = 0,\ a'_{35} = -1,\ a'_{37} = 1,\ a'_{38} = 20.$$

Первая строка (умножаем ведущую строку на пять и результат вычитаем из первой строки):

$$a'_{11}=0,\ a'_{12}=0,\ a'_{13}=1,\ a'_{14}=2,\ a'_{15}=5,\ a'_{17}=-5,\ a'_{18}=200.$$

Вторая строка остается неизменной:

$$a'_{21}=1,\ a'_{22}=0,\ a'_{23}=0,\ a'_{24}=-1,\ a'_{25}=0,\ a'_{27}=0,\ a'_{28}=50.$$

Четвертая строка (умножаем ведущую строку на три и результат прибавляем к четвертой строке):

$$a'_{41}=0,\ a'_{42}=0,\ a'_{43}=0,\ a'_{44}=-4,\ a'_{45}=-3,\ a'_{48}=260.$$

В столбце «Базис» прежнюю базисную переменную ведущей строки x_7 заменяем на новую базисную переменную ведущего столбца x_2. Вычеркиваем столбец выведенной искусственной переменной x_7 (табл. 1.6).

Таблица 1.6

x_1	x_2	x_3	x_4	x_5	Свободные члены	Базис
0	0	1	2	5	200	x_3
1	0	0	–1	0	50	x_1
0	1	0	0	–1	20	x_2
0	0	0	–4	–3	260	$F(x)$

Искусственные базисные переменные x_6 и x_7 выведены из базиса. Геометрически это означает попадание в одну из вершин многоугольника области Парето. Если не удается вывести искусственные переменные из базиса или получена неразрешимая задача, то исходная задача также неразрешима.

Далее используем стандартный алгоритм симплексного метода.

1. Выбираем наибольший по модулю отрицательный коэффициент целевой функции – это коэффициент при переменной x_4. Столбец $l = 4$ симплекс-таблицы, которому принадлежит новая базисная переменная x_4, является ведущим столбцом.

2. Делим элементы столбца свободных членов на соответствующие элементы ведущего столбца, при этом учитываем только положительные элементы. Выбираем наименьшее частное $\min\left\{\frac{200}{2}\right\}=100$. Строка $k = 1$, которой принадлежит наименьшее частное, является ведущей строкой. На пересечении ведущего столбца и ведущей строки находится ведущий элемент $a_{14}=2$.

3. Пересчитываем элементы симплекс-таблицы (см. табл. 1.6). Начинаем с ведущей строки.

Первая строка (делим каждый элемент ведущей строки на ведущий элемент):

$$a'_{11}=0,\ a'_{12}=0,\ a'_{13}=0{,}5,\ a'_{14}=1,\ a'_{15}=2{,}5,\ a'_{16}=100.$$

Вторая строка (первую строку a'_{1j} прибавляем ко второй):

$$a'_{21}=1,\ a'_{22}=0,\ a'_{23}=0{,}5,\ a'_{24}=0,\ a'_{25}=2{,}5,\ a'_{26}=150.$$

Третья строка остается неизменной:

$$a'_{31}=0,\ a'_{32}=1,\ a'_{33}=0,\ a'_{34}=0,\ a'_{35}=-1,\ a'_{36}=20.$$

Четвертая строка (умножаем первую строку a'_{1j} на четыре и результат прибавляем к четвертой строке):

$$a'_{41}=0,\ a'_{42}=0,\ a'_{43}=2,\ a'_{44}=0,\ a'_{45}=7,\ a'_{46}=660.$$

В столбце «Базис» прежнюю базисную переменную ведущей строки x_3 заменяем на новую базисную переменную ведущего столбца x_4 (табл. 1.7).

Таблица 1.7

x_1	x_2	x_3	x_4	x_5	Свободные члены	Базис
0	0	0,5	1	2,5	100	x_4
1	0	0,5	0	2,5	150	x_1
0	1	0	0	–1	20	x_2
0	0	2	0	7	660	$F(x)$

Коэффициенты целевой функции неотрицательны, поэтому новый опорный план (150; 20; 0; 100; 0), $f(x)$ = 660 является оптимальным.

1.4. Теория двойственности

Математический аппарат линейного программирования позволяет не только получить оптимальный план, но и сделать ряд экономически значимых выводов, основанных на свойствах задачи, двойственной к исходной (прямой).

Двойственная задача по отношению к исходной задаче составляется по следующим правилам:

1) в исходной задаче находится максимум целевой функции, а в двойственной – минимум;

2) количество неизвестных двойственной задачи равно количеству ограничений исходной задачи;

3) количество ограничений двойственной задачи равно количеству переменных исходной задачи;

4) ограничению исходной задачи вида $\le$ соответствует неотрицательная переменная двойственной задачи, ограничению исходной задачи вида $\ge$ соответствует неположительная переменная двойственной задачи; ограничению исходной задачи вида $=$ соответствует переменная двойственной задачи, которая может принимать как положительные, так и отрицательные значения;

5) если на переменную исходной задачи наложено (не наложено) условие неотрицательности, то соответствующим ограничением двойственной задачи будет неравенство вида $\ge$ (уравнение);

6) коэффициенты при неизвестных в целевой функции исходной задачи являются свободными членами в ограничениях двойственной задачи;

7) свободные члены ограничений исходной задачи являются коэффициентами при неизвестных в целевой функции двойственной задачи;

8) матрица, составленная из коэффициентов при неизвестных в системе ограничений исходной задачи, в двойственной задаче транспонируется.

Пример 1.4. Составить двойственную задачу для ЗЛП:

$$f(x) = x_1 + 2x_2 + 3x_3 \to \max,$$

$$\begin{cases} x_1 + 5x_2 + 2x_3 = 70, \\ x_1 + x_2 - 4x_3 \le 30, \\ x_1 - 2x_2 + x_3 \ge 40, \\ x_1 \ge 0,\ x_2 \ge 0. \end{cases}$$

Решение

Исходная задача сформулирована на максимум, двойственная задача – на минимум. Исходная задача имеет три ограничения, двойственная задача будет иметь три переменные y_1, y_2 и y_3. Второе ограничение исходной задачи имеет вид $\le$, тогда y_2 будет неотрицательным. Третье ограничение исходной задачи имеет вид $\ge$, тогда y_3 будет неположительным. В исходной задаче $x_1 \ge 0$ и $x_2 \ge 0$, в двойственной задаче первые два ограничения будут записаны в виде неравенств типа $\ge$. Коэффициенты при переменных целевой функции исходной задачи будут свободными членами ограничений двойственной задачи. Матрица коэффициентов при неизвестных в ограничениях исходной задачи транспонируется в двойственной задаче. Коэффициенты при переменных целевой функции двойственной задачи являются свободными членами ограничений исходной задачи.

Получаем двойственную задачу:

$$g(y) = 70y_1 + 30y_2 + 40y_3 \to \min,$$

$$\begin{cases} y_1 + y_2 + y_3 \geq 1, \\ 5y_1 + y_2 - 2y_3 \geq 2, \\ 2y_1 - 4y_2 + y_3 = 3, \\ y_2 \geq 0,\ y_3 \leq 0. \end{cases}$$

Решение исходной задачи с помощью MS Excel [2, Лист 3]:

$$f(x) = 100,\ x_1 = 10,\ x_2 = 0,\ x_3 = 30.$$

Решение двойственной задачи с помощью MS Excel [2, Лист 4]:

$$g(y) = 100,\ y_1 = 2,\ y_2 = 0,\ y_3 = -1.$$

Экономическая интерпретация двойственной задачи заключается в определении оптимальных оценок y_i, называемых неявными или «теневыми» ценами каждого вида ресурса с целью минимизации общей ценности (стоимости) всех ресурсов.

Значения переменных y_i в оптимальном решении двойственной задачи представляют собой оценки влияния свободных членов b_i системы ограничений исходной задачи на величину приращения целевой функции

$$\Delta f(x) = \Delta b_i y_i.$$

Двойственные оценки выступают как инструмент балансировки и сопоставления затрат и результатов, они гарантируют рентабельность оптимального плана.

Самую высокую ценность имеют те ресурсы, которые в наибольшей степени ограничивают выпуск продукции, прибыль предприятия и на увеличение которых предприятие согласно нести значительные расходы. Ресурс, который предприятие не использует полностью в оптимальном плане, получает нулевую оценку.

1.5. Решение задач линейного программирования с помощью электронных таблиц MS Excel

Рассмотрим решение примера 1.3 с помощью MS Excel 2007/10 [2, Лист 5].

Решение задачи линейного программирования с помощью Excel предполагает последовательность действий:

ввод условий задачи;

решение;

анализ оптимального решения.

I. Ввод условий задачи включает:

создание экранной формы для ввода условий задачи;

ввод исходных данных;

ввод зависимостей из математической модели;

назначение целевой функции и переменных;

ввод граничных условий и ограничений.

1. Создаем экранную форму на рабочем листе MS Excel и вводим данные задачи:

	A	B	C	D	E	F
1		Переменные				
2	Имя	x1	x2			
3	Значение			ЦФ	Направление	
4	Коэф. ЦФ	10	24		max	
5						
6		Ограничения		Левая часть	Знак	Правая часть
7	1-ое	2	5		≤	200
8	2-ое	4	2		≤	160

Примечание. ЦФ – целевая функция.

2. Вводим зависимости из математической модели в ячейки D4, D7, D8, используя окно диалога *Мастер функций* (f_x *Вставить функцию*), категория *Математические*, в алфавитном списке *Выберите функцию* выбираем функцию СУММПРОИЗВ. Функция может быть введена непосредственно с клавиатуры.

Содержимое ячеек:

D4=СУММПРОИЗВ($B$3:$C$3;B4:C4);

D7=СУММПРОИЗВ($B$3:$C$3;B7:C7);

D8=СУММПРОИЗВ($B$3:$C$3;B8:C8).

3. Назначение целевой функции и переменных.

Для назначения целевой функции используем команду *Данные/Поиск решения*. Если команда *Поиск решения* недоступна, то ее необходимо установить. Используем команду *Настройка панели быстрого доступа*, выбираем *Другие команды…*, в окне диалога *Параметры Excel* выделяем *Надстройки*, в списке *Надстройки* выделяем *Поиск решения*, в алфавитном списке *Управление* выбираем *Надстройки Excel*, используем кнопку *Перейти*, в диалоговом окне *Надстройки* устанавливаем флажок *Поиск решения*.

Диалоговое окно *Поиск решения* (рис. 1.3):

в поле *Установить целевую ячейку* указываем ячейку целевой функции $D$4. Если перед вызовом инструмента *Поиск решения* выделить ячейку с целевой функцией $D$4, то это поле уже будет заполнено;

в элементе управления *Равной* выбираем *максимальному значению*;

в элементе управления *Изменяя ячейки* указываем диапазон $B$3:$C$3 (рис. 1.3), в котором размещены численные значения переменных.

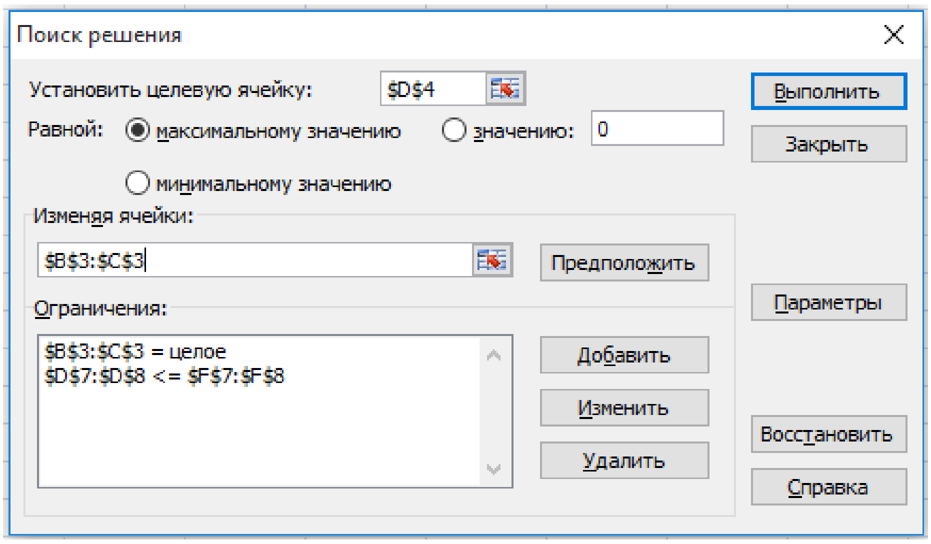


Рис. 1.3

4. Ввод ограничений выполняется с помощью кнопки *Добавить* в списке *Ограничения* (см. рис. 1.3). На экране появится диалоговое окно *Добавление ограничения* (рис. 1.4).

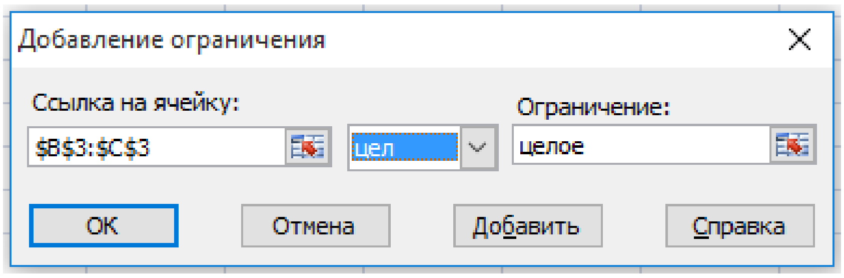


Рис. 1.4

Поле *Ссылка на ячейку* предназначено для указания ссылки на ячейку, где хранятся переменные или формулы, используемые для ввода ограничений. В поле *Ограничение* можно вводить константу, ссылку на ячейки со значениями и формулами. Значения из полей *Ссылка на ячейку* и *Ограничение* сравниваются с помощью операций «>=, <=, =, цел, двоич», которые можно выбрать из списка, расположенного между этими двумя полями.

Кнопка *Добавить* позволяет добавить несколько ограничений, кнопка *OK* добавляет ограничение и закрывает диалоговое окно *Добавление ограничения*.

Вводим условие целочисленности переменных: $B$3:$C$3 *цел* целое, *Добавить* (см. рис. 1.4).

Аналогично вводим ограничения: $D$7:$D$8 ≤ $F$7: F$8, *OK*.

В результате на экране появится диалоговое окно (см. рис. 1.3).

Условие неотрицательности переменных введем во второй части.

На этом ввод условий задачи заканчивается.

II. Решение

Используем кнопку *Параметры* диалогового окна *Поиск решения*.

Команды диалогового окна *Параметры поиска решения*: *Максимальное время*, *Предельное число итераций*, *Относительная погрешность*, *Допустимое отклонение*, используемые по умолчанию, подходят для решения большей части практических задач.

Установим флажок *Линейная модель*, что обеспечит применение симплекс-метода, а также флажок *Неотрицательные значения*, что обеспечит условие неотрицательности переменных, *ОК* (рис. 1.5).

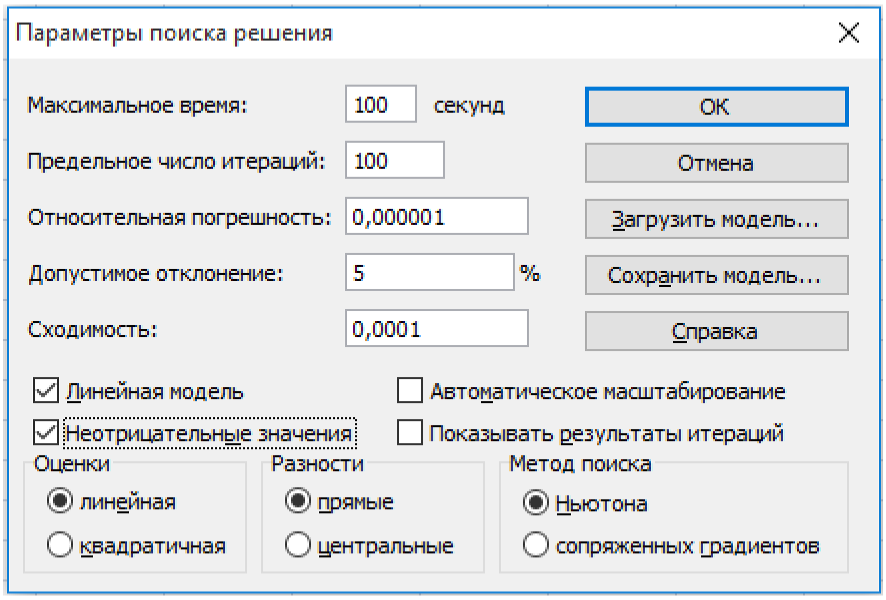


Рис. 1.5

В диалоговом окне *Поиск решения* воспользуемся кнопкой *Выполнить*.

В диалоговом окне *Результаты поиска решения* приводятся результаты решения и даются типы отчета. Выбираем *Сохранить найденное решение*, *ОК*. На рабочем листе отобразится найденный оптимальный план:

	A	B	C	D	E	F
1		Переменные				
2	Имя	x1	x2			
3	Значение	25	30	ЦФ	Направление	
4	Коэф. ЦФ	10	24	970	max	
5						
6		Ограничения		Левая часть	Знак	Правая часть
7	1-ое	2	5	200	≤	200
8	2-ое	4	2	160	≤	160

В результате получили следующий оптимальный план:

$$x = (25; 30), \quad f(x) = 970.$$

III. Анализ оптимального решения

Вызываются отчеты трех типов: по результатам, устойчивости и пределам.

Вызов отчетов выполняется с помощью диалогового окна *Результаты поиска решений* (*Поиск решения*, *Выполнить*). Выбираем *Результаты* в диалоговом окне *Тип отчета, ОК*.

Открываем созданный лист *Отчет по результатам 1*:

Microsoft Excel 12.0 Отчет по результатам
Рабочий лист: [Optimus1.xlsx]Лист5
Отчет создан: 21.12.2015 9:39:11

Целевая ячейка (Максимум)

Ячейка	Имя	Исходное значение	Результат
$D$4	Коэф. в ЦФ ЦФ	970	970

Изменяемые ячейки

Ячейка	Имя	Исходное значение	Результат
$B$3	Значение x1	25	25
$C$3	Значение x2	30	30

Ограничения

Ячейка	Имя	Значение	Формула	Статус	Разница
$D$7	1-ое Левая часть	200	$D$7<=$F$7	связанное	0
$D$9	2-ое Левая часть	160	$D$9<=$F$9	связанное	0
$B$3	Значение x1	25	$B$3=целое	связанное	0
$C$3	Значение x2	30	$C$3=целое	связанное	0
$B$3	Значение x1	25	$B$3=целое	не связан.	25
$C$3	Значение x2	30	$C$3=целое	не связан.	30

Из данного отчета следует, что предприятие может получить максимальную прибыль в размере 970 ден. ед. при изготовлении 25 изделий типа *A* и 30 изделий типа *B*. При этом общий фонд рабочего времени токарного и фрезерного оборудования используется полностью.

Отчет по устойчивости (применяется в случае отсутствия ограничения целочисленности). Отменим условие целочисленности. В диалоговом окне *Поиска решения* в списке *Ограничения* выделяем условие целочисленности и удаляем его. Выбираем кнопку *Выполнить,* в диалоговом окне *Тип отчета* выбираем *Устойчивость*, *ОК*.

Открываем созданный лист *Отчет по устойчивости 1*:

Microsoft Excel 12.0 Отчет по устойчивости
Рабочий лист:[Optimus1.xlsx]Лист5
Отчет создан: 21.12.2015 10:05:25

Изменяемые ячейки

Ячейка	Имя	Результ. значение	Нормир. стоимость	Целевой коэффициент	Допустимое увеличение	Допустимое уменьшение
$B$3	Значение x1	25	0	10	38	0,4
$C$3	Значение x2	30	0	24	1	19

Ограничения

Ячейка	Имя	Результ. значение	Теневая цена	Ограничение Правая часть	Допустимое увеличение	Допустимое уменьшение
$D$7	1-ое Левая часть	200	4,75	200	200	120
$D$8	2-ое Левая часть	160	0,125	160	240	80

Отчет по устойчивости состоит из двух таблиц.

В таблице «Изменяемые ячейки» приводятся значения для переменных:

результат решения исходной задачи;

нормированная стоимость, которая означает следующее. Если продукт входит в оптимальный план, то в данной колонке для этого продукта ставится 0; если же не входит – отрицательное число, показывающее, насколько (по абсолютной величине) нужно увеличить прибыль от производства единицы данного продукта, чтобы он вошел в оптимальный план. Таким образом, нормированная стоимость показывает, насколько изменится значение целевой функции в случае принудительного включения единицы продукции в оптимальное решение;

коэффициенты целевой функции;

предельные значения приращения коэффициентов ΔC_j *целевой функции*, при которых сохраняется первоначальное оптимальное решение.

Информация данной таблицы означает, что для получения максимальной прибыли в размере 970 ден. ед. предприятие должно изготавливать 25 изделий типа *A* и 30 изделий типа *B*, так как нормированная стоимость изделия равна 0.

Интервалы изменения коэффициентов целевой функции, при которых сохраняется первоначальное оптимальное решение, следующие:

$$10-0{,}4 \le C_1 \le 10+38;$$
$$24-19 \le C_2 \le 24+1;$$
$$9{,}6 \le C_1 \le 48;$$
$$5 \le C_2 \le 25.$$

В таблице «Ограничения» приводятся аналогичные значения для ограничений: *теневая цена*, т. е. двойственные оценки y_i, которые показывают, как изменится целевая функция при изменении ресурсов на единицу.

Информация данной таблицы отчета показывает, что общий фонд рабочего времени токарного и фрезерного оборудования используется полностью, оно является дефицитным и для токарного, и для фрезерного оборудования. Двойственная оценка (теневая цена) фонда рабочего времени токарного оборудования равна 4,75, а фрезерного – 0,125. Следовательно, рабочее время токарного оборудования более дефицитно, чем фрезерного.

Интервалы устойчивости двойственных оценок, при которых сохраняется оптимальное решение двойственной задачи:

$$200-120 \le b_1 \le 200+200;$$

$$160-80 \le b_2 \le 160+240.$$

Следовательно, интервалы устойчивости двойственных оценок:

$$80 \le b_1 \le 400;$$

$$80 \le b_2 \le 400.$$

Отчет по пределам неприменим при ограничениях целочисленности. Выбираем диалоговое окно *Поиск решения*, используем кнопку *Выполнить*, выбираем *Пределы* в диалоговом окне *Тип отчета*, *ОК*.

Открываем созданный лист *Отчет по пределам 1*:

Microsoft Excel 12.0 Отчет по пределам
Рабочий лист: [Optimus1.xlsx]Отчет по пределам 1
Отчет создан: 21.12.2015 10:54:03

Ячейка	Целевое имя	Значение
\$D\$4	Коэф. в ЦФ ЦФ	970

Ячейка	Изменяемое имя	Значение	Нижний предел	Целевой результат	Верхний предел	Целевой результат
\$B\$3	Значение x1	25	0	720	25	970
\$C\$3	Значение x2	30	0	250	30	970

В таблице показаны значения целевой функции в случае нулевого значения одного из параметров управления при сохранении оптимальных значений других параметров управления. Если из оптимального плана исключить изделия типа *A*, то прибыль составит $f(0,\ 30) = 24 \cdot 30 =$ $= 720$ ден. ед. Если из оптимального плана исключить изделия типа *B*, прибыль составит $f(25,\ 0) = 10 \cdot 25 = 250$ ден. ед. Аналогично для верхнего предела изменения одного из параметров управления при сохранении оптимальных значений других параметров управления.

1.6. Примеры типовых задач линейного программирования

1.6.1. Планирование производства (распределение ресурсов)

Пример 1.5. При изготовлении трех видов продукции $П_1$, $П_2$ и $П_3$ используют четыре вида ресурсов – P_1, P_2, P_3 и P_4. Расход ресурсов на производство единицы продукции, запасы ресурсов, прибыль, получаемая от единицы продукции, приведены ниже:

Вид ресурса:	Расход ресурсов на производство единицы продукции			Запас ресурса, ед.
	$П_1$	$П_2$	$П_3$	
P_1	10	20	30	500
P_2	12	13	9	600
P_3	2	5	1	100
P_4	7	4	5	200
Прибыль, получаемая от одного изделия, ден. ед.	12	15	10	

Необходимо составить план производства продукции, при котором прибыль от ее реализации будет максимальной.

Решение

1. В качестве параметров управления выберем x_i – количество выпущенной продукции i-го вида.

2. Целевая функция должна максимизировать стоимость выпускаемой продукции $f(x)=\sum_{i=1}^{3} c_i x_i = 12x_1 + 15x_2 + 10x_3 \to \max$.

3. Количество выпускаемой продукции ограничено запасами ресурсов четырех видов, P_1, P_2, P_3 и P_4: $10x_1 + 20x_2 + 30x_3 \le 500$ – для первого ресурса, $12x_1 + 13x_2 + 9x_3 \le 600$ – для второго, $2x_1 + 5x_2 + x_3 \le 100$ – для третьего, $7x_1 + 4x_2 + 5x_3 \le 200$ – для четвертого. Необходимо добавить условие неотрицательности переменных x_i.

Составим математическую модель задачи:

$$f(x) = 12x_1 + 15x_2 + 10x_3 \to \max,$$

$$\begin{cases} 10x_1 + 20x_2 + 30x_3 \le 500, \\ 12x_1 + 13x_2 + 9x_3 \le 600, \\ 2x_1 + 5x_2 + x_3 \le 100, \\ 7x_1 + 4x_2 + 5x_3 \le 200, \\ x_i \ge 0,\ i = \overline{1,3}. \end{cases}$$

Решение с помощью MS Excel [2, Лист 6]:

$f(x)$ = 438,(63) ден. ед., $x_1 = 20,(45)$, $x_2 = 11,(36)$, $x_3 = 2,(27)$.

Ответ: следует выпускать продукцию первого вида в количестве 20,(45), второго вида в количестве 11,(36), третьего вида в количестве 2,(27). Прибыль составит 438,(63) ден. ед.

1.6.2. Выбор оптимальных проектов для финансирования

Пример 1.6. Управляющему банка для выделения кредита было представлено пять проектов. Доступная наличность банка b_j, потребности проектов C_{ij} в каждом квартале и прибыль по ним составляют:

Проект	Квартал				Прибыль, тыс. руб.
	I	II	III	IV	
П1	100	90	60	90	320
П2	120	80	70	60	310
П3	80	70	60	120	280
П4	90	100	50	80	270
П5	130	90	30	70	290
Ресурс банка, тыс. руб.	410	420	460	400	

Какие проекты следует финансировать и какое количество наличности необходимо в течение каждого квартала, чтобы максимизировать прибыль?

Решение

1. В качестве параметров управления выберем $x_i\ (i=\overline{1,5})$, каждый из которых принимает только два значения: 1, если i-й проект получит кредит; 0, если i-й проект не получит кредит.

2. Целевая функция должна максимизировать прибыль, получаемую банком, поэтому $f(x) = 320x_1 + 310x_2 + 280x_3 + 270x_4 + 290x_5 \to \max$.

3. Количество выбранных проектов ограничено только доступной наличностью банка в каждом квартале: $\sum_{i=1}^{5} C_{ij}x_i \le b_j$. Необходимо добавить условие двоичности переменных: $x_i \in \{0;1\}$.

В итоге мы имеем математическую модель задачи:

$$f(x) = 320x_1 + 310x_2 + 280x_3 + 270x_4 + 290x_5 \to \max$$

$$\begin{cases} 100x_1 + 120x_2 + 80x_3 + 90x_4 + 130x_5 \le 410, \\ 90x_1 + 80x_2 + 70x_3 + 100x_4 + 90x_5 \le 420, \\ 60x_1 + 70x_2 + 60x_3 + 50x_4 + 30x_5 \le 460, \\ 90x_1 + 60x_2 + 120x_3 + 80x_4 + 70x_5 \le 400, \\ x_i \in \{0;1\}. \end{cases}$$

Решение с помощью MS Excel [2, Лист 7]:

$$f(x) = 1180000 \text{ руб.},\ x_1 = x_2 = x_3 = x_4 = 1,\ x_5 = 0\ .$$

Ответ: следует финансировать первые четыре проекта. В первом квартале необходимо 390 тыс. руб., во втором – 340 тыс. руб., в третьем – 240 тыс. руб., в четвертом – 350 тыс. руб. Прибыль составит 1180 тыс. руб.

1.6.3. Планирование финансового потока

Пример 1.7. В начале года фирма «Благодетель» решила профинансировать долгосрочный инвестиционный проект, располагая средствами в размере M = 2 млн руб. Необходимые расходы по проекту:

Квартал	I	II	III	IV
Δ_i (расходы, тыс. руб.)	600	520	400	550

Всего за год нужно инвестировать 2,07 млн руб., поэтому было принято решение положить свободные деньги в банк на 3-месячные, 6-месячные и 12-месячные депозитные счета. Процентные ставки по указанным депозитам составляют:

Вид депозита	Годовой доход
1 (3-месячный)	10 %
2 (6-месячный)	12 %
3 (12-месячный)	16 %

Следует составить поквартальный план размещения средств на депозитных счетах, который обеспечил бы необходимое финансирование и максимальный доход от свободных средств к концу года.

Решение

1. В качестве параметров управления выберем x_{ij} – сумму, положенную в банк на i-й вид депозита (i = 1, 2, 3) в начале j-го квартала (j = 1, 2, 3, 4).

2. Целевая функция должна максимизировать доход, полученный от размещения свободных средств. Если в начале первого квартала положить в банк сумму x_{11} на 3-месячный вид депозита, то к началу второго квартала прибыль от вложенной суммы составит $\frac{10\,\%}{100\,\%}\frac{3}{12}x_{11} = 0,025x_{11}$. Если в начале первого квартала положить в банк сумму x_{21} на 6-месячный депозит, то в начале третьего квартала прибыль от вложенной суммы составит $\frac{12\,\%}{100\,\%}\frac{6}{12}x_{21} = 0,06x_{21}$. Если в начале первого квартала положить в банк сумму x_{31} на 12-месячный депозит, то в конце года прибыль от вложенной суммы составит $\frac{16\,\%}{100\,\%}\frac{12}{12}x_{31} = 0,16x_{31}$.

Целевая функция имеет вид:

$$f(x) = 0,025(x_{11} + x_{12} + x_{13} + x_{14}) + 0,06(x_{21} + x_{22} + x_{23}) + 0,16x_{31} \to \max .$$

3. Максимальный доход от свободных средств ограничивают обязательные инвестиции в каждом квартале. В первом квартале можно положить на депозиты всю сумму, за исключением расходов в первом квартале: $x_{11}+x_{21}+x_{31}\le 2\,000-600$. В начале второго квартала получим прибыль $0{,}025x_{11}$ от 3-месячного депозита. В начале третьего квартала получим прибыль $0{,}025(x_{11}+x_{12})+0{,}06x_{21}$ от 3-месячных депозитов первого и второго кварталов и 6-месячного депозита первого квартала. В начале четвертого квартала получим прибыль $0{,}025(x_{11}+x_{12}+x_{13})+0{,}06(x_{21}+x_{22})$ от 3-месячных депозитов первого, второго и третьего кварталов и 6-месячного депозита первого и второго кварталов.

В начале второго квартала в банке может находиться сумма $x_{12}+x_{22}+(x_{21}+x_{31})$, которая не должна превышать $2\,000-600-520+0{,}025x_{11}=880+0{,}025x_{11}$. В начале третьего квартала в банке может находиться сумма $x_{13}+x_{23}+(x_{22}+x_{31})$, которая не должна превышать $480+0{,}025(x_{11}+x_{21})+0{,}06x_{21}$. В начале четвертого квартала в банке может находиться сумма $x_{14}+(x_{23}+x_{31})$, которая не должна превышать $-70+0{,}025(x_{11}+x_{12}+x_{13})+0{,}06(x_{21}+x_{22})$.

Составим математическую модель задачи:

$$f(x)=0{,}025(x_{11}+x_{12}+x_{13}+x_{14})+0{,}06(x_{21}+x_{22}+x_{23})+0{,}16x_{31}\to\max,$$

$$\begin{cases} x_{11}+x_{21}+x_{31}\le 1400,\\ x_{12}+x_{22}+(x_{21}+x_{31})\le 880+0{,}025x_{11},\\ x_{13}+x_{23}+(x_{22}+x_{31})\le 480+0{,}025(x_{11}+x_{12})+0{,}06x_{21},\\ x_{14}+(x_{23}+x_{31})\le -70+0{,}025(x_{11}+x_{12}+x_{13})+0{,}06(x_{21}+x_{22}),\\ x_{ij}\ge 0,\ i=\overline{1,3},\ j=\overline{1,4}. \end{cases}$$

Задача имеет альтернативные оптимумы. Одно из решений с точностью до рубля найдено с помощью MS Excel ([2, Лист 8] при использовании линейной модели и нулевых начальных значениях переменных):

$f(x)$ = 80570 руб., x_{11} = 1013530 руб., x_{21} = 377358 руб.,
x_{22} = 518868 руб., x_{31} = 9112 руб.

Остаток 10570 руб.

Если решать без использования флажка линейной модели в диалоговом окне *Параметры поиска решения*, то получим еще один оптимум:

$f(x)$ = 80570 руб., x_{11} = 576828 руб., x_{13} = 462904 руб.,
x_{21} = 814060 руб., x_{22} = 71249 руб., x_{31} = 9112 руб.

1.6.4. Задача о ранце (рюкзаке)

Имеется n видов предметов, которые турист хочет взять в поход. В рюкзаке он может нести предметы, масса и объем которых соответственно не более M кг и V м^3 . Известны вес m_j, объем v_j каждого предмета и его ценность (эффективность) C_j для туриста. Сколько предметов турист может положить в рюкзак, чтобы суммарная стоимость (эффективность) предметов была максимальной?

Задача о ранце имеет и экономическую интерпретацию. В качестве «предметов» рассматриваются заказы, в качестве полезности – прибыль от выполнения заказа, а в качестве веса – себестоимость заказа.

Составим математическую модель задачи. Введем булевы переменные x_j, которые принимают два значения: ноль, если турист не взял j-й предмет и единицу, если он взял его. Тогда математическая модель задачи будет иметь вид:

$$f(x) = \sum_{j=1}^{n} C_j x_j \to \max,$$

$$\sum_{j=1}^{n} m_j x_j \le M\,,\ \sum_{j=1}^{n} v_j x_j \le V\,,\ x_j \in \{0;1\}.$$

Пример 1.8. Имеются восемь ценных предметов, которые должны быть вывезены в безопасное место в случае пожара. Их характеристики:

Вес, кг	150	200	250	100	400	450	300	400
Габариты, м^3	0,5	0,6	1,1	0,8	0,7	0,4	0,3	1,2
Стоимость, млн руб.	2	3,5	4	3	5	7	4,2	8

Грузоподъемность имеющегося транспортного средства 1,5 т, вместимость 5 м^3. Какие предметы нужно вывезти в первую очередь, чтобы стоимость вывозимых предметов была максимальной?

Решение

1. В качестве параметров управления выберем переменные $x_i\ (i = \overline{1,8})$, каждая из которых принимает только два значения: 1, если i-й предмет вошел в список первоочередных предметов; 0 – в противном случае.

2. Целевая функция должна максимизировать ценность вывозимых предметов:

$$f(x) = \sum_{i=1}^{8} c_i x_i = 2x_1 + 3{,}5x_2 + 4x_3 + 3x_4 + 5x_5 + 7x_6 + 4{,}2x_7 + 8x_8 \to \max .$$

3. Количество вывозимых предметов ограничено грузоподьемностью и вместимостью транспортного средства: $\sum_{i=1}^{8} m_i x_i \le 1500$, $\sum_{i=1}^{8} n_i x_i \le 5$. Необходимо добавить условие двоичности переменных: $x_i \in \{0;1\}$.

Составим математическую модель задачи:

$$f(x) = 2x_1 + 3{,}5x_2 + 4x_3 + 3x_4 + 5x_5 + 7x_6 + 4{,}2x_7 + 8x_8 \to \max,$$

$$\begin{cases} 150x_1 + 200x_2 + 250x_3 + 100x_4 + 400x_5 + 450x_6 + 300x_7 + 400x_8 \le 1500, \\ 0{,}5x_1 + 0{,}6x_2 + 1{,}1x_3 + 0{,}8x_4 + 0{,}7x_5 + 0{,}4x_6 + 0{,}3x_7 + 1{,}2x_8 \le 5, \\ x_i \in \{0;1\},\ i = \overline{1,8}. \end{cases}$$

Решение с помощью MS Excel [2, Лист 9]:

$$f(x) = 25{,}5,\ \ x_2 = x_3 = x_4 = x_6 = x_8 = 1,\ \ x_1 = x_5 = x_7 = 0.$$

Ответ: оптимальная ценность предметов составит 25,5 млн руб., если в первую очередь вывозить второй, третий, четвертый, шестой и восьмой предметы.

1.6.5. Задача составления смесей

Содержательная постановка задачи. Имеется некоторый набор первичных продуктов, из которых можно составлять различные смеси. Обязательное требование к составляемой смеси заключается в том, что количество веществ определенного вида в ней должно быть не ниже заданной нормы, а количество других веществ – не выше заданной нормы. Требуется из имеющихся смесей выбрать оптимальную совокупность с минимальной стоимостью.

К этой группе задач относятся задачи о выборе диеты, составлении кормового рациона в животноводстве, горючих и смазочных смесей в нефтеперерабатывающей промышленности, смесей для получения бетона в строительстве и т. д.

Пример 1.9. Завод выпускает три вида полуфабрикатов в количествах: A_1 – 300 т, A_2 – 200 т и A_3 – 400 т. В результате смешивания этих компонентов получают два вида продукции. Пропорции смешиваемых полуфабрикатов для продукции B_1 – 2:1:2, для B_2 – 1:1:2. Стоимость 1 т продукции B_1 – 500 руб., B_2 – 400 руб. Требуется составить оптимальный план, обеспечивающий максимальную стоимость выпускаемой продукции.

Решение

1. В качестве параметров управления выберем x_i – количество выпущенной продукции i-го вида.

2. Целевая функция должна максимизировать стоимость выпускаемой продукции: $f(x) = 500x_1 + 400x_2 \to \max$.

3. Количество выпускаемой продукции ограничено запасами полуфабрикатов j-го вида: $2x_1 + x_2 \le 300$ – для полуфабриката A_1; $x_1 + x_2 \le 200$ – для полуфабриката A_2; $2x_1 + 2_2 \le 400$ – для полуфабриката A_3. Необходимо добавить условие неотрицательности переменных x_i.

Составим математическую модель задачи:

$$f(x) = 500x_1 + 400x_2 \to \max,$$

$$\begin{cases} 2x_1 + x_2 \le 300, \\ x_1 + x_2 \le 200, \\ 2x_1 + 2x_2 \le 400, \\ x_1, x_2 \ge 0. \end{cases}$$

Решение с помощью MS Excel [2, Лист 10]:

$$f(x) = 90000 \text{ руб.}, \; x_1 = 100, x_2 = 100.$$

Ответ: следует выпускать продукцию первого вида в количестве 100 т, второго вида в количестве 100 т. Максимальная стоимость выпускаемой продукции составит 90 тыс. руб.

1.6.6. Рациональный раскрой промышленных материалов

Применение рационального раскроя промышленных материалов обеспечивает повышение коэффициента использования длинномерных материалов на 2–5 %, а листовых – на 3–7 %.

Особенность рационального раскроя состоит в разработке и использовании таких допустимых раскройных планов, при которых получается необходимый ассортимент заготовок, а отходы по длине, площади или весу сводятся к минимуму. Данная задача особенно актуальна в условиях многономенклатурного производства.

Сформулируем условие задачи в общем виде для одномерного материала.

Необходимо раскроить одномерный материал (прутки стального проката, трубы, брусья, доски и т. д.) длины l. В соответствии с планом или заявками потребителей требуются заготовки i видов длиной Δ_i каждая в соответствующих количествах a_i. Необходимо определить план раскроя, обеспечивающий заявки потребителей при минимальном количестве раскройных материалов (по количественному составу, остаткам по длине или весу).

Для составления математической модели введем обозначения:

l – длина исходного материала;

i – виды заготовок (по длине), которые необходимо получить при раскрое ($i = 1, 2, \ldots, m$);

b_i – необходимое количество заготовок i-го вида по плану;

j – варианты раскроя единицы исходного материала ($j = 1, 2, \ldots, n$);

a_{ij} – количество заготовок i-го вида, получаемое при раскрое единицы исходного материала по j-му варианту;

C_j – отход при раскрое единицы исходного материала по j-му варианту раскроя;

x_j – количество единиц исходного материала, которое будет раскраиваться по j-му варианту.

Целевая функция может минимизировать либо отходы по длине $f(x)=\sum_{j=1}^{n} C_j x_j$, либо количество раскраиваемых единиц материала $f(x)=\sum_{j=1}^{n} x_j$.

В результате имеем математическую модель:

$$f(x)=\sum_{j=1}^{n} C_j x_j \to \min,$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \ (i = 1,2,\ldots,m),\ x_j \in N_0.$$

Если ассортимент заявок потребителей является уникальным (редким), то ограничения записываются в виде равенств. Если же размеры заготовок стандартны, то ограничения можно записывать в виде неравенств.

Обычно количество ограничений гораздо меньше количества раскройных планов ($m < n$), поэтому задача может имеет альтернативные оптимумы, которые дают одно и то же значение целевой функции.

На практике в случае отсутствия полной автоматизации при минимизации затрат необходимо учитывать также количество материала, раскроенного по каждому способу.

Пример 1.10. На складе имеются бревна длиной 600 см. Согласно заявкам потребителей требуются заготовки трех видов в количествах:

Δ_1 = 210 см – 200 шт.;

Δ_2 = 170 см – 300 шт.;

Δ_3 = 140 см – 400 шт.

Необходимо найти план раскроя, удовлетворяющий заявкам потребителей и обеспечивающий:

а) минимальное количество отходов материала;

б) минимальное количество распиленных бревен.

Решение

Определим всевозможные способы распила бревен. Пусть x_j – количество бревен, которые будут распилены по j-му варианту. Для первого варианта раскроя выбираем максимально возможное количество заготовок длиной Δ_1 = 210, т. е. две единицы. Из остатка 600 – 2·210 = 180 выбираем максимально возможное количество заготовок длиной Δ_2 = 170, т. е. одну единицу. Из получившегося остатка 180 – 1·170 = 10 выбираем максимально возможное количество заготовок длиной Δ_3 = 140, т. е. ноль единиц. В итоге получим первый вариант раскроя (2; 1; 0). В результате мы должны перебрать все возможные варианты раскроя (чтобы не допустить потери какого-либо способа раскроя, длины заготовок расположены в порядке убывания):

Способ распила, j	Число заготовок i-го вида, получаемых при раскрое одного бревна			Остаток, C_j
	$\Delta_1 = 210$	$\Delta_2 = 170$	$\Delta_3 = 140$	
1	2	1	0	10
2	2	0	1	40
3	1	2	0	50
4	1	1	1	80
5	1	0	2	110
6	0	3	0	90
7	0	2	1	120
8	0	1	3	10
9	0	0	4	40

В результате имеем математическую модель для первой части задачи:

$$f(x) = \sum_{j=1}^{n} C_j x_j,$$

$$f(x) = 10x_1 + 40x_2 + 50x_3 + 80x_4 + 110x_5 + 90x_6 + 120x_7 + 10x_8 + 40x_9 \to \min,$$

$$\begin{cases} 2x_1 + 2x_2 + x_3 + x_4 + x_5 = 200, \\ x_1 + 2x_3 + x_4 + 3x_6 + 2x_7 + x_8 = 300, \\ x_2 + x_4 + 2x_5 + x_7 + 3x_8 + 4x_9 = 400, \\ x_j \in N_0. \end{cases}$$

Оценим возможность получения альтернативных оптимумов. Составим расширенную матрицу системы ограничений и приведем ее к ступенчатому виду:

$$\begin{pmatrix} 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 200 \\ 1 & 0 & 2 & 1 & 0 & 3 & 2 & 1 & 0 & 300 \\ 0 & 1 & 0 & 1 & 2 & 0 & 1 & 3 & 4 & 400 \end{pmatrix}.$$

Умножим вторую строку матрицы на два и вычтем из нее первую строку:

$$\begin{pmatrix} 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 200 \\ 0 & -2 & 3 & 1 & -1 & 6 & 4 & 2 & 0 & 400 \\ 0 & 1 & 0 & 1 & 2 & 0 & 1 & 3 & 4 & 400 \end{pmatrix}.$$

Умножим третью строку на два и прибавим к ней вторую строку:

$$\begin{pmatrix} 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 200 \\ 0 & -2 & 3 & 1 & -1 & 6 & 4 & 2 & 0 & 400 \\ 0 & 0 & 3 & 3 & 3 & 6 & 6 & 8 & 8 & 1200 \end{pmatrix}.$$

Ранг матрицы равен трем, количество базисных переменных равно трем, количество свободных переменных равно 9 – 3 = 6, поэтому задача может иметь альтернативные оптимумы. Решение с помощью MS Excel (файл Optimus1.xlsx Лист 11 при использовании нулевых начальных значениях переменных):

$f(x)$ = 4600 см, $x_1 = 99$, $x_3 = x_4 = x_9 = 1$, $x_6 = 21$, $x_7 = 2$, $x_8 = 131$, $x_2 = x_5 = 0$.

Всего необходимо распилить 256 бревен.

Решение с помощью MS Excel ([2, Лист 11] при использовании линейной модели и нулевых начальных значениях переменных):

$f(x)$ = 4600 см, $x_1 = 86$, $x_2 = 4$, $x_6 = 14$, $x_8 = 132$, $x_4 = x_5 = x_7 = x_9 = 0$.

Всего необходимо распилить 256 бревен.

Решение с помощью MS Excel для стандартных заявок (ограничения записаны в виде неравенств $\geq$):

$f(x)$ = 3000 см, $x_1 = 129$, $x_8 = 171$, $x_2 = x_3 = x_4 = x_5 = x_6 = x_7 = x_9 = 0$.

Всего необходимо распилить 300 бревен, остается 300 заготовок длиной 10 см. Любой набор параметров

$$x_1 = 129 + z,\ x_8 = 171 - z,$$

где $z \in [-29; 37]$, $z \in \mathbf{Z}$, также является оптимальным решением.

Целевая функция второй части задачи имеет вид

$$f(x) = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 \rightarrow \min.$$

Ограничения остаются неизменными. Задача может иметь множество оптимумов. Решение с помощью MS Excel дает [2, Лист 12]:

$f(x) = 256$, $x_1 = 90$, $x_3 = 18$, $x_5 = 2$, $x_6 = 14$, $x_8 = 132$, $x_2 = x_4 = x_7 = x_9 = 0$.

Другая разновидность задачи о раскрое заключается в том, что из имеющегося ограниченного количества материала необходимо изготовить максимальное количество заготовок в количествах, пропорциональных определенным числам.

Пример 1.11. Для изготовления оконных рам необходимы рейки длиной 40 и 120 см в соотношении 8:8. Фирма имеет в наличии 400 реек длиной 250 см. Определить план распила, обеспечивающий максимальное число комплектов.

Решение

Определим все возможные способы распила реек:

Способ распила, j	Число заготовок i-го вида, получаемых при раскрое одной рейки	
	$\Delta_1 = 120$	$\Delta_2 = 40$
1	2	0
2	1	3
3	0	6

Пусть x_j – количество реек, которые будут распилены по j-му варианту, x – количество комплектов реек. Учитывая, что все рейки должны быть распилены, а число реек каждого размера должно соответствовать условию комплектности, составим математическую модель:

$$f(x) = x \to \max,$$
$$\begin{cases} x_1 + x_2 + x_3 = 400, \\ 2x_1 + x_2 = 8x, \\ 3x_2 + 6x_3 = 8x, \\ x, x_j \in N_0. \end{cases}$$

Решение с помощью MS Excel дает [2, Лист 13]:

$$f(x) = 75,\ x_1 = 240,\ x_2 = 120,\ x_3 = 40,\ x = 75.$$

Рассмотрим аналитическое решение задачи. Составим расширенную матрицу системы ограничений и приведем ее к ступенчатому виду:

$$\begin{pmatrix} 1 & 1 & 1 & 400 \\ 2 & 1 & 0 & 8x \\ 0 & 3 & 6 & 8x \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 400 \\ 0 & -1 & -2 & 8x-800 \\ 0 & 3 & 6 & 8x \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 400 \\ 0 & -1 & -2 & 8x-800 \\ 0 & 0 & 0 & 32x-2400 \end{pmatrix}.$$

Умножим первую строку первой матрицы на два и результат вычтем из второй строки. Умножим вторую строку второй матрицы на три и результат прибавим к третьей строке.

Из последней строки имеем: $32x - 2400 = 0$, $x = 75$. Из второй строки получим: $x_2 + 2x_3 = 800 - 8x$, $x_2 = 200 - 2x_3$. Из первого уравнения имеем: $x_1 + x_2 + x_3 = 400$, $x_1 = 400 - x_2 - x_3$, $x_1 = 200 + x_3$.

Оптимальным планом является любой набор:

$$x_1 = 200 + x_3,\ x_2 = 200 - 2x_3,\ x_{1,2,3} \in N_0.$$

§ 2. Транспортная задача

2.1. Основные понятия, построение опорного плана перевозок

Решение *транспортной задачи* (ТЗ) состоит из трех этапов:

составление математической модели;

нахождение опорного (первоначального) плана;

нахождение оптимального плана.

Второй этап желателен, но необязателен. Сформулируем условие ТЗ в общем виде.

Имеется несколько пунктов производства или сосредоточения какого-либо однородного продукта (например, нефти или картофеля), а также некоторое количество пунктов, испытывающих потребность в этом продукте (пунктов потребления). Запасы в каждом пункте известны. Транспортные издержки, связанные с произвольной перевозкой, заданы. Необходимо составить такой план перевозок продукта из пунктов производства в пункты потребления, при котором весь продукт будет перевезен в пункты потребления, каждый пункт потребления будет полностью удовлетворен и суммарные транспортные затраты окажутся минимальными.

Переведем условие задачи на математический язык. Введем обозначения. Пусть A_i – пункты производства (их число m), B_j – пункты потребления (их число n), a_i – количество продукта в пункте A_i, b_j – количество продукта в пункте B_j. Примем допущение, что стоимость перевозки продукции линейно зависит от количества перевозимой продукции. Пусть $c_{ij}(x)$ – стоимость перевозки x единиц продукта из A_i в B_j. С учетом наших допущений имеем: $c_{ij}(x) = c_{ij}x$, где c_{ij} – стоимость перевозки одной единицы продукта из A_i в B_j.

В качестве параметров управления выберем x_{ij} – количество продукта, перевозимого из A_i в B_j. Тогда суммарные транспортные издержки (целевую функцию задачи) можно определить как

$$f(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} .$$

Составим ограничения. Необходимо вывести весь продукт из каждого пункта производства $\sum_{j=1}^{n} x_{ij} = a_i$, ($i = 1, 2, \ldots, m$) и полностью удовлетворить потребность пунктов потребления $\sum_{i=1}^{m} x_{ij} = b_j$, ($j = 1, 2, \ldots, n$). Поскольку перевозки идут из пунктов производства в пункты потребления, то добавим ограничения $x_{ij} \geq 0$ для любых i и j ($\forall i, j$).

В данной задаче производство полностью удовлетворяет потребление (выполняется условие баланса): $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$.

В результате имеем математическую модель ТЗ:

$$f(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min,$$

$$\begin{cases} \sum_{j=1}^{n} x_{ij} = a_i, \; i = \overline{1, m}, \\ \sum_{i=1}^{m} x_{ij} = b_j, \; j = \overline{1, n}, \\ x_{ij} \geq 0, \forall i, j. \end{cases}$$

Для нахождения опорного плана обычно используют методы:
«северо-западного угла»;
метод добротностей;
метод наименьших стоимостей или минимального элемента;
метод двойного предпочтения.

Транспортные задачи, в которых производство полностью удовлетворяет потреблению, получили название *закрытых* или *замкнутых* ТЗ с выполнением *условия баланса*.

Пример 1.12. Имеются два пункта производства товара $a_1 = 10$, $a_2 = 20$ и два пункта потребления: $b_1 = 5$, $b_2 = 25$. Стоимость перевозок единицы груза из пунктов производства в пункты потребления задана $c_{11} = 1$, $c_{12} = 2$, $c_{21} = 3$, $c_{22} = 1$. Необходимо составить математическую модель плана перевозок продукта из пунктов производства в пункты потребления, при котором весь продукт будет перевезен в пункты потребления, каждый пункт потребления полностью удовлетворен и суммарные транспортные затраты окажутся минимальными.

Решение

Поскольку условие баланса выполняется (10 + 20 = 5 + 25), то ТЗ является замкнутой и все ограничения записываются в виде равенств.

Пусть x_{ij} – количество продукта, перевозимого из i-го пункта производства в j-й пункт потребления. Тогда

$$F(x) = x_{11} + 2x_{12} + 3x_{21} + x_{22} \to \min,$$

$$\begin{cases} x_{11} + x_{12} = 10, \\ x_{21} + x_{22} = 20, \\ x_{11} + x_{21} = 5, \\ x_{12} + x_{22} = 25, \\ x_{ij} \geq 0, \forall i, j. \end{cases}$$

Первое ограничение показывает, что из первого пункта производства необходимо вывезти товар в первый и второй пункты потребления $x_{11} + x_{12}$ в количестве $a_1 = 10$. Аналогично второе ограничение.

Третье ограничение показывает, что в первый пункт потребления необходимо привезти товар из первого и второго пунктов производств $x_{11} + x_{21}$ в количестве $b_1 = 5$. Аналогично четвертое ограничение.

Составим первоначальный план перевозок методом наименьших стоимостей с двойным предпочтением. Исходные данные ТЗ представим в виде транспортной таблицы:

a_i/b_j	$b_1 = 5$	$b_2 = 25$
$a_1 = 10$	1^{**}	2
$a_2 = 20$	3	1^{**}

Отмечаем клетки с наименьшими стоимостями перевозок по каждой строке и каждому столбцу. Клетки, имеющие две отметки, заполняем в первую очередь. В клетке a_1b_1 записываем число 5 и вычеркиваем полностью заполненный столбец b_1. В клетке a_2b_2 записываем число 20 и вычеркиваем полностью заполненную строку a_2:

a_i/b_j	$b_1 = 5$	$b_2 = 25$
$a_1 = 10$	**5**	
$a_2 = 20$		**20**

Заполняем клетки с одной отметкой с учетом заявок производителей и потребителей, затем заполняем клетки без отметок. Осталась одна клетка a_1b_2. В нее записываем число 5:

a_i/b_j	$b_1 = 5$	$b_2 = 25$
$a_1 = 10$	**5**	**5**
$a_2 = 20$		**20**

Решение с помощью MS Excel дает [3, Лист 1]:

$$f(x) = 35,\ x_{11} = 5,\ x_{12} = 5,\ x_{22} = 20.$$

В нашем случае опорный план совпал с оптимальным решением.

Рассмотрим ТЗ, в которой не выполняется условие баланса. Такую ТЗ задачу называют *открытой*.

Пример 1.13. Имеются два пункта производства товара, $a_1 = 10$, $a_2 = 25$, и два пункта потребления, $b_1 = 10$, $b_2 = 30$. Стоимость перевозок товара задана $c_{11} = 4$, $c_{12} = 2$, $c_{21} = 3$, $c_{22} = 1$. За недопоставку одной единицы товара первый пункт потребления терпит убыток, равный 1, а второй – 0,5. Необходимо составить математическую модель плана перевозок продукта из пунктов производства в пункты потребления, при котором весь продукт был бы перевезен в пункты потребления, а суммарные транспортные затраты и убытки оказались минимальными.

Решение

Введем обозначения. Пусть x_{ij} – количество продукта, перевозимого из i-го пункта производства в j-й пункт потребления, $z_1 = b_1 - x_{11} - x_{21}$, $z_2 = b_2 - x_{12} - x_{22}$ – количество недопоставленного товара соответственно в первый и второй пункты потребления.

Составим математическую модель:

$$F(x) = 4x_{11} + 2x_{12} + 3x_{21} + x_{22} + z_1 + 0{,}5z_2 \to \min,$$

$$\begin{cases} x_{11} + x_{12} = 10, \\ x_{21} + x_{22} = 25, \\ x_{11} + x_{21} \le 10, \\ x_{12} + x_{22} \le 30, \\ x_{ij} \ge 0, \forall i, j. \end{cases}$$

Прежде чем решать открытую ТЗ, ее необходимо сбалансировать (при решении задачи в MS Excel это можно не делать).

Вводим дополнительный *фиктивный* пункт производства a_{i+1} (в случае дефицита товара) или фиктивный пункт потребления b_{j+1} (в случае профицита товара). Мощность дополнительно введенного фиктивного пункта равна количеству недопоставленного или избыточного товара.

Для нашего случая введем фиктивный пункт производства $a_3 = (30 + 10) - (25 + 10) = 5$. Стоимость перевозок из фиктивного пункта производства равна нулю, поэтому целевая функция не меняется, а система ограничений принимает вид:

$$\begin{cases} x_{11} + x_{12} = 10, \\ x_{21} + x_{22} = 25, \\ x_{31} + x_{32} = 5, \\ x_{11} + x_{21} + x_{31} = 10, \\ x_{12} + x_{22} + x_{32} = 30, \\ x_{ij} \geq 0, \forall i, j. \end{cases}$$

Решение с помощью MS Excel дает [3, Лист 2]:

$$f(x) = 60,\ x_{12} = 10,\ x_{21} = 5,\ x_{22} = 20.$$

Пример 1.14. Завод имеет три цеха *A*, *B*, *C* и четыре склада №№ 1, 2, 3, 4. Цех *A* производит 30 тыс. изделий, цех *B* – 40 тыс. шт., цех *C* – 20 тыс. шт. за плановое время. Пропускная способность складов за то же время: склад № 1 – 20 тыс. шт., склад № 2 – 30 тыс. шт., склад № 3 – 30 тыс. шт., склад № 4 – 10 тыс. шт. Стоимость перевозок одной тысячи изделий в склады №№ 1, 2, 3, 4 соответственно составляет: из цеха *A* – 2, 3, 3, 4 ден. ед.; из цеха *B* – 3, 2, 5, 1 ден. ед.; из цеха *C* – 4, 3, 2, 6 ден. ед.

Составить план перевозок, при котором расходы на перевозку всех изделий минимальны.

Решение

Условие баланса выполняется $\sum_{i=1}^{3} a_i = \sum_{j=1}^{4} b_j = 90$, поэтому ТЗ является замкнутой и все ограничения записываем в виде равенств. Пусть x_{ij} (тыс. шт.) – количество изделий, перевозимых из i-го цеха в j-й склад. Тогда математическая модель задачи будет иметь вид:

$$F(x) = 2x_{11} + 3x_{12} + 3x_{13} + 4x_{14} + 3x_{21} + 2x_{22} + 5x_{23} + x_{24} + \\ + 4x_{31} + 3x_{32} + 2x_{33} + 6x_{34} \to \min;$$

цеха: $\begin{cases} x_{11} + x_{12} + x_{13} + x_{14} = 30, \\ x_{21} + x_{22} + x_{23} + x_{24} = 40, \\ x_{31} + x_{32} + x_{33} + x_{34} = 20, \end{cases}$ склады: $\begin{cases} x_{11} + x_{21} + x_{31} = 20, \\ x_{12} + x_{22} + x_{32} = 30, \\ x_{13} + x_{23} + x_{33} = 30, \\ x_{14} + x_{24} + x_{34} = 10, \\ x_{ij} \geq 0. \end{cases}$

Составим первоначальный план перевозок методом наименьших стоимостей с двойным предпочтением. Исходные данные ТЗ представим в виде транспортной таблицы. Отметим клетки с наименьшими стоимостями перевозок по каждой строке и каждому столбцу:

a_i/b_j	$b_1 = 20$	$b_2 = 30$	$b_3 = 30$	$b_4 = 10$
$a_1 = 30$	2**	3	3	4
$a_2 = 40$	3	2*	5	1**
$a_3 = 20$	4	3	2**	6

Клетки, имеющие две отметки, заполняем в первую очередь с учетом заявок производителей и потребителей: x_{11} = 20 (вычеркиваем первый полностью заполненный столбец), x_{24} = 10 (вычеркиваем четвертый столбец), x_{33} = 20 (вычеркиваем третью строку). Затем заполняем клетки с одной отметкой с учетом заявок производителей и потребителей: x_{22} = 30 (вычеркиваем второй столбец и вторую строку). Остается клетка $a_1b_3 = x_{13} = 10$:

a_i/b_j	$b_1 = 20$	$b_2 = 30$	$b_3 = 30$	$b_4 = 10$
$a_1 = 30$	2** **20**	3	3 **10**	4
$a_2 = 40$	3	2* **30**	5	1** **10**
$a_3 = 20$	4	3	2** **20**	6

Получили начальный план перевозок: x_{11} = 20, x_{13} = 10, x_{22} = 30, x_{24} = 10, x_{33} = 20, $f(x)$ = 180 ден. ед. Нахождение оптимальнго плана перевозок описано в п. 2.3 данной главы.

2.2. Создание оптимального плана перевозок

Для получения оптимального плана перевозок разработано несколько методов:

распределительный;
потенциалов;
дельта-метод;
аппроксимация Фогеля.

Рассмотрим метод потенциалов, который широко применяется на практике и используется при программировании. На первом этапе составляют начальный план перевозок. На втором этапе строят систему потенциалов и проверяют начальный план на оптимальность. Каждому i-му поставщику устанавливается потенциал U_i, который можно интерпретировать как цену продукта в пункте поставщика, а каждому j-му потребителю устанавливается потенциал V_j, который характеризует цену продукта в пункте потребителя. Если начальный план является неоптимальным, то переходят к третьему этапу, суть которого заключается в корректировке плана прикрепления потребителей и поставщиков. Затем снова переходят ко второму этапу и т. д.

Пример 1.15. Фирма должна отправить шкафы с трех складов в три магазина. На складах имеется соответственно 15, 25 и 10 шкафов, а для трех магазинов требуется соответственно 20, 12, и 18 шкафов. Стоимость перевозки C_{ij} одного шкафа со складов a_i в магазины b_j составляет:

Склады	Магазины		
	b_1	b_2	b_3
a_1	6	3	4
a_2	3	7	2
a_3	2	8	6

Как следует спланировать перевозку, чтобы ее стоимость была минимальной?

Решение

Условие баланса выполняется: $\sum_{i=1}^{3} a_i = \sum_{j=1}^{3} b_j = 50$, поэтому ТЗ является замкнутой и все ограничения записываем в виде равенств.

На первом этапе получаем первоначальный план перевозок методом наименьших стоимостей с двойным предпочтением:

a_i/b_j	$b_1 = 20$	$b_2 = 12$	$b_3 = 18$	U_i
$a_1 = 15$	6 **3**	3** **12**	4	$U_1 = 0$
$a_2 = 25$	3 **7**	7	2** **18**	$U_2 = 3$
$a_3 = 10$	2** **10**	8	6	$U_3 = 4$
V_j	$V_1 = 6$	$V_2 = 3$	$V_3 = 5$	

Начальный план: $x_{11} = 3$, $x_{12} = 12$, $x_{21} = 7$, $x_{23} = 18$, $x_{31} = 10$, $f(x) = 131$.

На втором этапе построим систему потенциалов U_i и V_j для начального опорного плана.

Для занятых клеток составляем систему линейных уравнений вида $V_j = U_i + C_{ij}$:

$$V_1 = U_1 + 6,\ V_1 = U_2 + 3,\ V_1 = U_3 + 2,\ V_2 = U_1 + 3,\ V_3 = U_2 + 2.$$

Полагаем $U_1 = 0$, тогда $V_1 = 6$, $U_2 = 3$, $U_3 = 4$, $V_2 = 3$, $V_3 = 5$. Найденные потенциалы указаны в вышеприведенной таблице.

Для того чтобы опорный план был оптимальным, необходимо и достаточно, чтобы система потенциалов U_i и V_j удовлетворяла условию:

$$V_j = U_i + C_{ij} \text{ для } x_{ij} > 0 \text{ (для занятой клетки);}$$

$$V_j \le U_i + C_{ij} \text{ для } x_{ij} = 0 \text{ (для незанятой клетки).}$$

Для незанятых клеток проверим условие оптимальности:

$$V_2 = 3 < U_2 + C_{22} = 3 + 7 = 10,\ V_2 = 3 < U_3 + C_{32} = 4 + 8 = 12,$$

$$V_3 = 5 > U_1 + C_{13} = 0 + 4 = 4,\ V_3 = 5 < U_3 + C_{33} = 4 + 6 = 10.$$

Третье неравенство, или клетка a_1b_3, не удовлетворяет условию оптимальности, поэтому переходим к третьему этапу.

Для незанятых клеток, не удовлетворяющих условию оптимальности, находим наибольшую величину $\alpha_{ij} = V_j - U_i - C_{ij}$:

$$\alpha_{13} = V_3 - U_1 - C_{13} = 5 - 0 - 4 = 1.$$

Клетка a_1b_3 является положительной вершиной цикла. Строим замкнутый контур, начальная вершина которого находится в выбранной клетке, а остальные вершины контура – в занятых клетках. Линии контура могут быть горизонтальными либо вертикальными отрезками. Число отрезков и вершин будет четным. Первая вершина контура свободной клетки имеет знак «+», а в остальных вершинах контура расставляются поочередно знаки «–» и «+»:

a_i/b_j	$b_1 = 20$	$b_2 = 12$	$b_3 = 18$	U_i
$a_1 = 15$	6 – **3**	3 **12**	4 +	$U_1 = 0$
$a_2 = 25$	3 + **7**	7	2 – **18**	$U_2 = 3$
$a_3 = 10$	2 **10**	8	6	$U_3 = 4$
V_j	$V_1 = 6$	$V_2 = 3$	$V_3 = 5$	

Выбираем наименьшее из величин в вершинах контура с отрицательным знаком $\min\{3, 18\} = 3$. Уменьшаем на три единицы величину каждой вершины со знаком «–» и увеличиваем на три единицы величину каждой вершины со знаком «+». Получим:

a_i/b_j	$b_1 = 20$	$b_2 = 12$	$b_3 = 18$	U_i
$a_1 = 15$	6	3 **12**	4 **3**	$U_1 = 0$
$a_2 = 25$	3 **10**	7	2 **15**	$U_2 = 2$
$a_3 = 10$	2 **10**	8	6	$U_3 = 3$
V_j	$V_1 = 5$	$V_2 = 3$	$V_3 = 4$	

Новый опорный план:

$$x_{12} = 12,\ x_{13} = 3,\ x_{21} = 10,\ x_{23} = 15,\ x_{31} = 10,\ f(x) = 128.$$

Найдем потенциалы и проверим новый опорный план на оптимальность.

Для занятых клеток составляем систему линейных уравнений вида $V_j = U_i + C_{ij}$:

$$V_1 = U_2 + 3,\ V_1 = U_3 + 2,\ V_2 = U_1 + 3,\ V_3 = U_1 + 4,\ V_3 = U_2 + 2.$$

Полагаем $U_1 = 0$, тогда $V_2 = 3$, $V_3 = 4$ $U_2 = 2$, $V_1 = 5$, $U_3 = 3$. Найденные потенциалы указаны в вышеприведенной таблице.

Для незанятых клеток проверим условие оптимальности:

$$V_1 = 5 < U_1 + C_{11} = 0 + 6 = 6,\ V_2 = 3 < U_2 + C_{22} = 2 + 7 = 9,$$

$$V_2 = 3 < U_3 + C_{32} = 3 + 8 = 11,\ V_3 = 4 < U_3 + C_{33} = 3 + 6 = 9.$$

Полученный опорный план удовлетворяет всем условиям оптимальности, поэтому новый опорный план $x_{12} = 12$, $x_{13} = 3$, $x_{21} = 10$, $x_{23} = 15$, $x_{31} = 10$, $f(x) = 128$ является оптимальным.

2.3. Решение транспортной задачи с помощью MS Excel

Решим пример 1.14 с применением стандартных средств табличного процессора MS Excel. Создадим экранную форму на листе MS Excel и введем данные задачи.

В ячейках A2:A4 указаны производственные мощности цехов 1–3, в ячейках B1:E1 – производственные мощности складов 1–4, в ячейках B2:E4 – тарифы по перевозке тысячи единиц изделий, перевозимых из i-го цеха в j-й склад.

Ячейки G2:J4 соответствуют численным значениям переменных x_{ij}, ячейки K2:K4 – левым частям ограничений для цехов, ячейки G5:J5 – левым частям ограничений для складов, ячейка F2 является целевой функцией $f(x)$.

В ячейки G2:J4 введем значения переменных начального плана: $x_{11} = 20$ (ячейка G2), $x_{13} = 10$ (ячейка I2), $x_{22} = 30$ (ячейка H3), $x_{24} = 10$ (ячейка J3), $x_{33} = 20$ (ячейка I4).

Получим:

	A	B	C	D	E	F	G	H	I	J	K
1		**20**	**30**	**30**	**10**	$f(x)$	x_{ij}				
2	**30**	2	3	3	4		20		10		
3	**40**	3	2	5	1			30		10	
4	**20**	4	3	2	6				20		
5											

Вводим зависимости из математической модели в ячейки K2:K4, G5:J5, F2, используя окно диалога *Мастер функций* (*f_x Вставить функцию*), категория *Математические*, в алфавитном списке *Выберите функцию* выбираем функцию СУММ для ячеек K2:K4, G5:J5 и СУММПРОИЗВ для целевой ячейки F2. Функция может быть введена непосредственно с клавиатуры.

Содержимое ячеек:

K2=СУММ(G2:J2), K3=СУММ(G3:J3), K4=СУММ(G4:J4),
G5=СУММ(G2:G4), H5=СУММ(H2:H4), I5=СУММ(I2:I4),
J5=СУММ(J2:J4), F2=СУММПРОИЗВ(B2:E4;G2:J4).

В результате получим:

	A	B	C	D	E	F	G	H	I	J	K
1		**20**	**30**	**30**	**10**	$f(x)$	x_{ij}				
2	**30**	2	3	3	4	180	20		10		**30**
3	**40**	3	2	5	1			30		10	**40**
4	**20**	4	3	2	6				20		**20**
5							**20**	**30**	**30**	**10**	

Для назначения целевой функции используем команду *Данные/Поиск решения*. В диалоговом окне *Поиск решения* (рис. 1.6):

в поле *Установить целевую ячейку* указываем ячейку целевой функции $F$2. Если перед вызовом инструмента *Поиск решения* выделить ячейку с целевой функцией $F$2, то это поле уже будет заполнено;

в элементе управления *Равной* выбираем *максимальному значению*;

в элементе управления *Изменяя ячейки* указываем диапазон $G$2:$J$4, в котором размещены численные значения переменных.

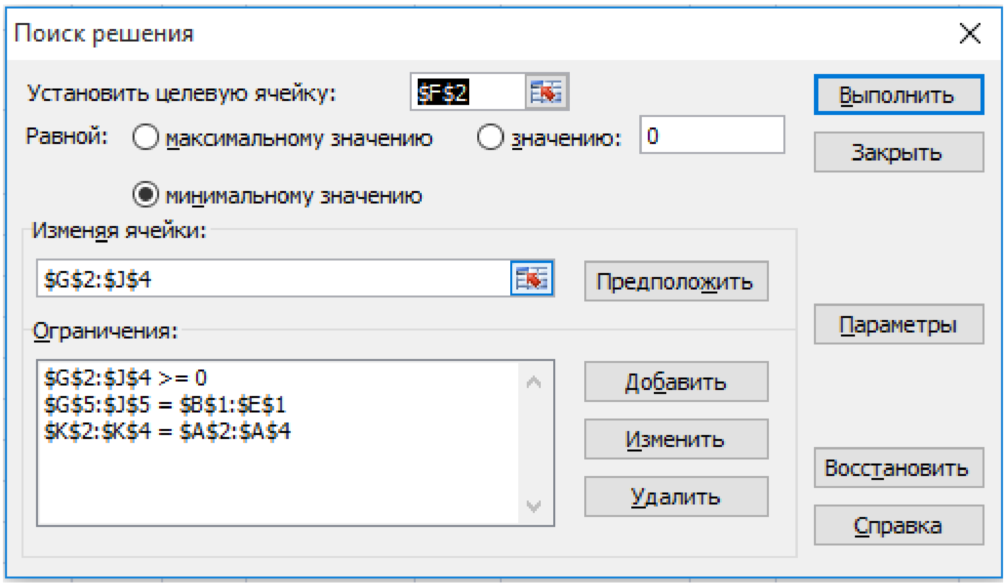


Рис. 1.6

Ввод ограничений выполняется с помощью кнопки *Добавить* в списке *Ограничения* (см. рис. 1.6). На экране появится диалоговое окно (рис. 1.7).

Вводим поочередно условие неотрицательности переменных: $G$2:$J$4 ≥ 0, *Добавить*.

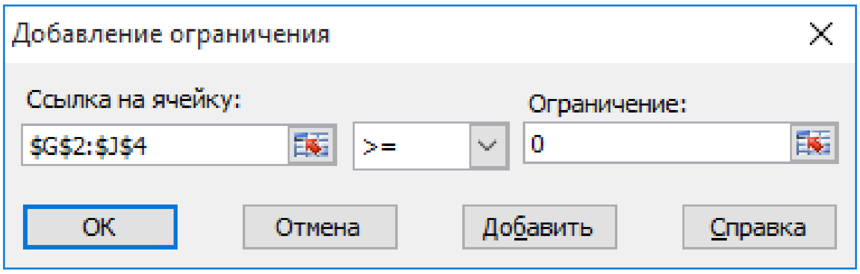


Рис. 1.7

Вводим ограничения поставщиков продукции (объем производимых изделий каждым цехом): $K$2:$K$4 = $A$2:$A$4 (рис. 1.8), *Добавить*.

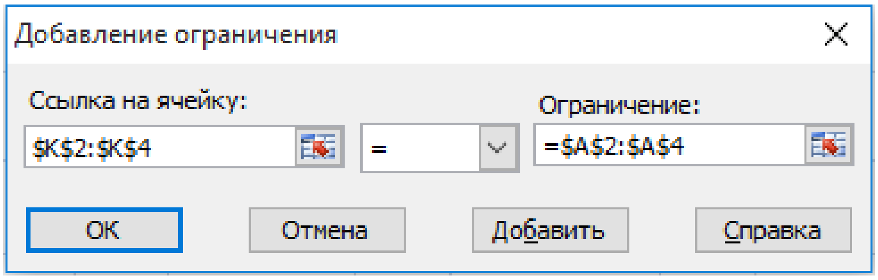


Рис. 1.8

Аналогично добавляем ограничения потребителей продукции (объем пропускной способности изделий каждого склада): $G$5:$J$5 = $B$1:$E$1 (рис. 1.9), *OK*.

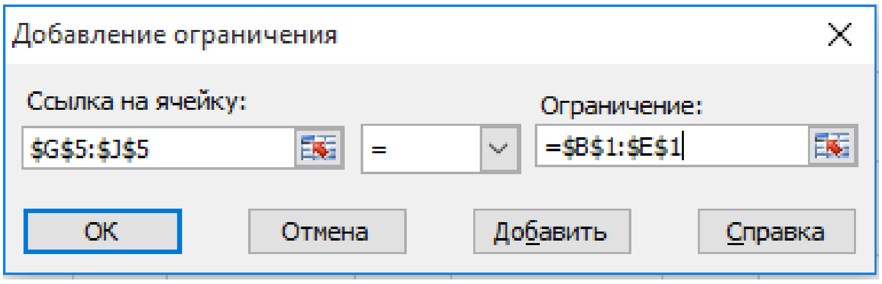


Рис. 1.9

На этом ввод условий задачи заканчивается.

Для решения используем кнопку *Параметры* диалогового окна *Поиск решения*. Устанавливаем флажок *Линейная модель*, что обеспечит применение симплекс-метода, и флажок *Неотрицательные значения*, что обеспечит условие неотрицательности переменных (рис. 1.10).

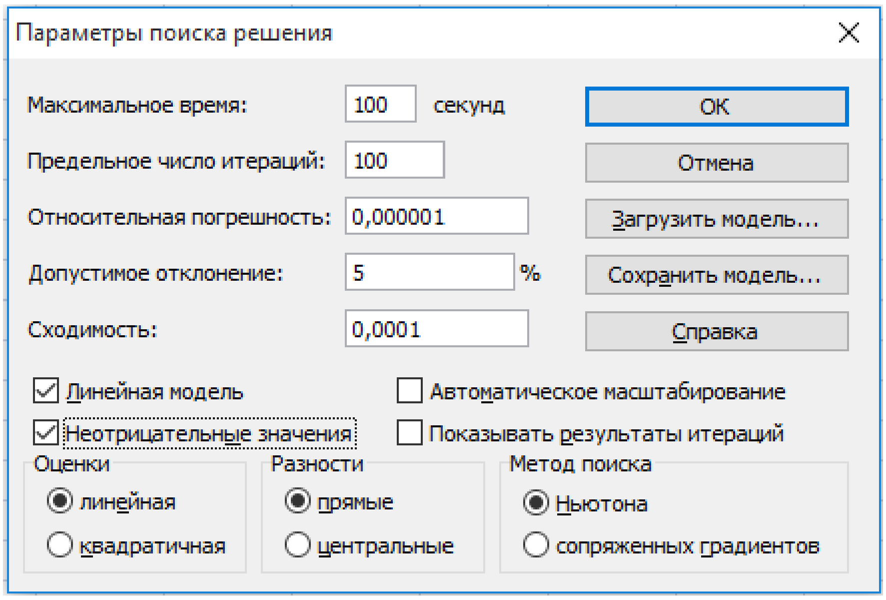


Рис. 1.10

В диалоговом окне *Поиск решения* выбираем *Выполнить*.

В диалоговом окне *Результаты поиска решения* приводятся результаты решения и даются типы отчета. Выбираем *Сохранить найденное решение*, *OK*. Решение с помощью MS Excel [3, Лист 3]:

$$x_{11} = 20,\ x_{13} = 10,\ x_{22} = 30,\ x_{24} = 10,\ x_{33} = 20,\ f(x) = 180.$$

2.4. Приложение транспортной задачи

Транспортная задача применяется при решении экономических задач, которые по своему характеру не имеют ничего общего с транспортировкой груза, поэтому величины C_{ij} могут иметь различный смысл: означать стоимость, расстояние, время, производительность и т. д. Рассмотрим постановку и математические модели некоторых задач.

2.4.1. Оптимальное закрепление за станками операций по обработке деталей

На предприятии имеется m видов станков, максимальное время работы которых соответственно равно a_i ($i = 1, 2,\ldots, m$) ч. Каждый из станков может выполнять n видов операций. Суммарное время выполнения каждой операции равно b_j ($j = 1, 2, \ldots, n$) ч. Известна производительность C_{ij} i-го вида станка при выполнении j-й операции. Требуется определить, сколько времени и на какой операции нужно использовать каждый из станков, чтобы обработать максимальное количество деталей.

Пример 1.16. На предприятии имеются два различных станка, каждый из которых может выполнять три вида элементарных операций по обработке детали, причем операции могут производиться в любом порядке. Максимальное время работы станков равно 380 и 400 ч соответственно; каждая операция должна выполняться соответственно в течение 336, 224 и 220 ч.

Требуется определить, сколько времени и на какой операции нужно использовать каждый из станков, чтобы обработать максимальное количество деталей. Производительность станков C_{ij} (i = 1, 2; j = 1, 2, 3) составляет:

Станки	Операции		
	4	3	6
	3	4	3

Решение

Обозначим через x_{ij} время, в течение которого i-й станок должен выполнять j-ю операцию. Тогда количество деталей, обработанных на i-м станке, равно $\sum_{j=1}^{3} C_{ij} x_{ij}$. Количество деталей, обработанных на обоих станках, можно выразить целевой функцией $f(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij} x_{ij}$. Поскольку максимально возможное время работы i-го станка ограничено (a_1 = 380, a_2 = 400), получаем ограничения $\sum_{j=1}^{3} x_{ij} \le a_i$. С другой стороны, время, отведенное на j-ю операцию, равно b_j ч, поэтому $\sum_{i=1}^{m} x_{ij} = b_j$.

Получаем математическую модель, записанную в общем виде:

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij} x_{ij} \to \max;$$

$$\sum_{i=1}^{n} x_{ij} \le a_j, \sum_{j=1}^{m} x_{ij} = b_i, x_{ij} \ge 0.$$

С учетом численных значений имеем:

$$f(x) = 4x_{11} + 3x_{12} + 6x_{13} + 3x_{21} + 4x_{22} + 3x_{23} \to \max,$$

$$\begin{cases} x_{11} + x_{12} + x_{13} \le 380, \\ x_{21} + x_{22} + x_{23} \le 400, \end{cases} \quad \begin{cases} x_{11} + x_{21} = 336, \\ x_{12} + x_{22} = 224, \\ x_{13} + x_{23} = 220, \end{cases} \quad x_{ij} \ge 0.$$

Поскольку суммарное время работы станков равно суммарному времени обработки деталей (380 + 400 = 336 + 224 + 220 = 780), то первую группу ограничений можно записать в виде равенств.

Решение с помощью MS Excel [3, Лист 4]:

$f(x)$ = 3384, $x_{11} = 160, x_{13} = 220, x_{21} = 176, x_{22} = 224.$

2.4.2. Оптимальные назначения или проблема выбора

Пусть имеются m лиц, которые могут выполнять различные m работ. Известна производительность C_{ij} i-го лица при выполнении j-й работы. Необходимо определить, кого и на какую работу следует назначить, чтобы добиться максимальной суммарной производительности при условии, что каждое лицо может быть назначено только на одну работу.

Для составления математической модели обозначим через x_{ij} назначение i-го лица на j-ю работу. Тогда, так как количество лиц равно количеству работ и каждое из них может быть назначено только на одну работу, x_{ij} принимает только два значения: 1, если i-е лицо назначается для выполнения j-й работы; 0, если i-е лицо не назначается для выполнения j-й работы, поэтому $\sum_{i=1}^{m} x_{ij} = 1$ и $\sum_{j=1}^{m} x_{ij} = 1$. При назначении i-го лица на j-ю работу производительность равна $C_{ij}x_{ij}$.

Таким образом, приходим к задаче линейного программирования:

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{m} C_{ij} x_{ij} \to \max;$$

$$\sum_{i=1}^{m} x_{ij} = 1, \sum_{j=1}^{m} x_{ij} = 1, x_{ij} \ge 0.$$

Если количество лиц больше количества работ, то вводят либо фиктивную работу с нулевой производительностью, либо вторую часть ограничений записывают в виде неравенств.

Пример 1.17. На предприятии пять станков разных видов, каждый из которых может выполнять пять различных операций по обработке деталей. Приведена производительность каждого станка при выполнении каждой операции:

	Операции				
Станки	5	3	4	6	7
	6	2	6	4	5
	4	3	5	6	6
	3	4	3	4	3
	5	6	3	2	5

Необходимо определить, какую операцию и за каким станком следует закрепить, чтобы суммарная производительность была максимальной при условии, что за каждым станком закреплена только одна операция.

Задача имеет несколько оптимумов. Решение с помощью MS Excel ([3, Лист 5] при использовании линейной модели):

$$f(x) = 28,\ x_{15} = x_{23} = x_{34} = x_{41} = x_{52} = 1,$$
$$f(x) = 28,\ x_{15} = x_{21} = x_{33} = x_{44} = x_{52} = 1,$$
$$f(x) = 28,\ x_{15} = x_{21} = x_{34} = x_{43} = x_{52} = 1.$$

2.4.3. Задача об использовании мощностей или задача о загрузке оборудования

Предприятию задан план производства продукции по времени и номенклатуре: требуется за время T выпустить n_1, n_2, …, n_k единиц продукции P_1, P_2, …, P_k. Продукция производится на станках S_1, S_2, …, S_m. Для каждого станка известны производительность a_{ij} (число единиц продукции P_j, которое можно произвести на станке S_i) и затраты b_{ij} на изготовление продукции P_j на станках S_i в единицу времени.

Необходимо составить такой план работы станков, чтобы затраты на производство всей продукции были минимальными.

Пример 1.18. Предприятию необходимо выпустить за месяц 1200 изделий типа A и 1 000 изделий типа B. Предприятие имеет два станка. Первый станок за час производит шесть изделий типа A или пять изделий типа B, а второй – соответственно четыре или три. При изготовлении изделий A типа на первом станке затраты составляют 10 ден. ед. в час, а на втором 8 ден. ед. в час. При изготовлении изделий B типа на первом станке затраты составляют 5 ден. ед. в час, а на втором – 4 ден. ед. в час. Необходимо составить такой план работы станков, чтобы затраты на производство всей продукции были минимальными. Каждый станок может работать в две смены (каждая по 8 ч) в течение 20 дней.

Решение

Обозначим через x_{ij} количество изделий j-го типа, изготовленных на i-м станке. Если на первом станке будет произведено x_{11} изделий типа A, то суммарное время, израсходованное первым станком на изготовление изделий типа A, составит $x_{11}/6$ ч. Аналогично для изделия типа B – $x_{12}/5$ ч, а для другого станка – $x_{21}/4$ ч и $x_{22}/3$ ч. Затраты на производство всей продукции составят:

$$f(x) = \frac{10}{6}x_{11} + x_{12} + 2x_{21} + \frac{4}{3}x_{22}.$$

Каждый станок может работать в две смены (каждая по 8 ч) в течение 20 дней, т. е. $2 \cdot 8 \cdot 20 = 320$ ч. В результате получаем два ограничения:

$$\frac{x_{11}}{6} + \frac{x_{12}}{5} \le 320, \ \frac{x_{21}}{4} + \frac{x_{22}}{3} \le 320.$$

Предприятию необходимо выпустить за месяц 1200 изделий типа A и 1000 изделий типа B: $x_{11} + x_{21} = 1200$, $x_{12} + x_{22} = 1000$.

В итоге получаем математическую модель:

$$f(x) = \frac{10}{6}x_{11} + x_{12} + 2x_{21} + \frac{4}{3}x_{22} \to \min,$$

$$1)\ \begin{cases} \dfrac{x_{11}}{6}+\dfrac{x_{12}}{5}\le 320, \\ \dfrac{x_{21}}{4}+\dfrac{x_{22}}{3}\le 320, \end{cases} \qquad 2)\ \begin{cases} x_{11}+x_{21}=1200, \\ x_{12}+x_{22}=1000, \\ x_{ij}\in N_0. \end{cases}$$

Решение с помощью MS Excel [3, Лист 6]:

$f(x)$ = 3113 ден. ед., x_{11} = 1200, x_{12} = 600, x_{22} = 400.

§ 3. Целочисленное программирование

Целочисленность и дискретность в задачах линейного программирования возникают из-за физической неделимости объектов. При массовом производстве (значительных объемах) применение сложных методов решения целочисленных ЗЛП нецелесообразно. В этих случаях применяют методы ЗЛП с последующим округлением полученных нецелочисленных значений переменных до их целого значения. Одним из критериев массовости производства ЗЛП является относительное изменение оптимального значения целевой функции при изменении параметра управления на единицу. Исследователь сам выбирает количественное значение этого параметра. Это может быть 0,1 или 0,01 %.

При производстве небольшого (относительно небольшого) количества дорогостоящих изделий должны применяться методы целочисленного программирования. Задачу целочисленного программирования решают одним из методов линейного программирования, например симплексным. Если найденное оптимальное решение не является целочисленным, то дополнительно вводят одно или несколько ограничений и продолжают поиск оптимального целочисленного решения специальными методами:

Балаша;

Фора – Мальгранжа;

Гомори;

ветвей и границ;

графический и т. д.

Рассмотрим на примере применение графического метода решения, который используется для задач с двумя или тремя параметрами управления.

Пример 1.19. Намечается выпуск двух видов костюмов – мужских и женских. Для каждого костюма определены нормы затрат ресурсов и прибыль от их реализации. Исходные данные:

Вид ресурса	Норма затрат ресурса на изготовление костюма		Запас ресурса
	Мужской	Женский	
Шерсть, м	3,5	1	350
Лавсан, м	0,5	2	240
Человеко-недель	1	1	138
Прибыль, ден. ед.	20	15	

Требуется определить наилучшую программу выпуска костюмов, обеспечивающую максимальную прибыль.

Решение

1. В качестве параметров управления x_1 и x_2 выберем соответственно количество изготовленных костюмов мужского и женского вида.

2. Целевая функция должна максимизировать прибыль, получаемую от реализации костюмов, поэтому $f(x) = 20x_1 + 15x_2 \to \max$.

3. Количество выпускаемых костюмов ограничено запасами ресурсов. В процессе производства мы не должны превысить имеющиеся запасы ресурсов. Необходимо добавить условие неотрицательности и целочисленности переменных $x_{1,2} \in N_0$.

В итоге имеем математическую модель задачи:

$$f(x) = 20x_1 + 15x_2 \to \max,$$

$$\begin{cases} 3{,}5x_1 + x_2 \le 350, \\ 0{,}5x_1 + 2x_2 \le 240, \\ x_1 + x_2 \le 138, \\ x_{1,2} \in N_0. \end{cases}$$

Графический метод решения

Изобразим область допустимых решений в системе координат Ox_1x_2.

Найдем две точки, лежащие на прямой $3{,}5x_1 + x_2 = 350$. Пусть $x_1 = 0$, тогда из данного уравнения имеем $x_2 = 350$. Пусть $x_2 = 0$, тогда из уравнения имеем $x_1 = 350/3{,}5 = 100$. Получены координаты двух точек прямой AB: $A(0;\ 350)$ и $B(100;\ 0)$.

Найдем две точки, лежащие на прямой $0{,}5x_1 + 2x_2 = 240$. Пусть $x_1 = 0$, тогда $x_2 = 240/2 = 120$. Пусть $x_2 = 0$, тогда $x_1 = 240/0{,}5 = 480$. В итоге имеем координаты двух точек прямой CD: $C(0;\ 120)$ и $D(480;\ 0)$.

Найдем две точки, лежащие на прямой $x_1 + x_2 = 138$. Пусть $x_1 = 0$, тогда $x_2 = 138$. Пусть $x_2 = 0$, тогда $x_1 = 138$. В итоге имеем координаты двух точек прямой EF: $E(0;\ 138)$ и $F(138;\ 0)$.

Получили шесть замкнутых и одну незамкнутую область. Для определения области допустимых решений возьмем контрольную точку с координатами (0; 0), принадлежащую только пятиугольнику $OCKLB$ (рис. 1.11). Координаты этой точки удовлетворяют всем ограничениям. Следовательно, выбранный пятиугольник $OCKLB$ является искомой областью Парето.

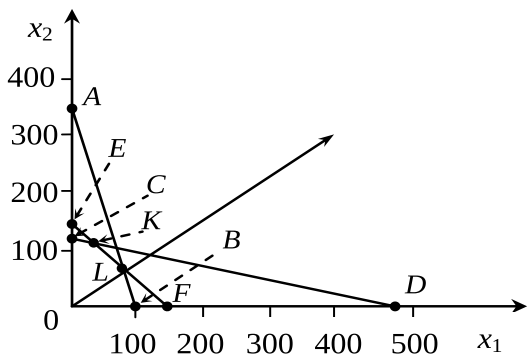


Рис. 1.11

Найдем координаты точки пересечения прямых *AB* и *EF* из системы уравнений

$$\begin{cases} 3{,}5x_1 + x_2 = 350, \\ x_1 + x_2 = 138. \end{cases}$$

Вычтем из первого уравнение второе. Разделим полученное первое уравнение на 2,5 и подставим во второе уравнение численное значение $x_1 = 84{,}8$. Получим координаты точки $L(84{,}8;\ 53{,}2)$ пересечения прямых *AB* и *EF:*

$$\begin{cases} 3{,}5x_1 + x_2 = 350, \\ x_1 + x_2 = 138, \end{cases} \Rightarrow \begin{cases} 2{,}5x_1 = 212, \\ x_1 + x_2 = 138, \end{cases} \Rightarrow \begin{cases} x_1 = 84{,}8, \\ x_2 = 53{,}2. \end{cases}$$

Проекция вершины $L(84{,}8;\ 53{,}2)$ пятиугольника *OCKLB* на продолжение вектора целевой функции $\vec{N}(20;15)$ будет наибольшей среди проекций остальных его вершин. Найдем оптимальное значение целевой функции в точке $L(84{,}8;\ 53{,}2)$:

$$F(84{,}8,\ 53{,}2) = 2494.$$

Оптимальное решение является нецелочисленным.

Найдем оптимальное целочисленное решение графическим методом. На рис. 1.12 сплошными жирными линиями изображены прямые *AB* и *EF*, точка их пересечения $L(84{,}8;\ 53{,}2)$; пунктирными линиями изображены две линии уровня целевой функции. Одна линия уровня проходит через оптимальную точку *L* и имеет уравнение $20x_1 + 15x_2 = 2494$. Вторая пунктирная линия получена параллельным переносом первой линии уровня до пересечения с первой точкой области Парето, которая имеет целые значения координат – (84; 54). Полученная точка с координатами (84; 54) является оптимальным решением задачи на целочисленном множестве решений.

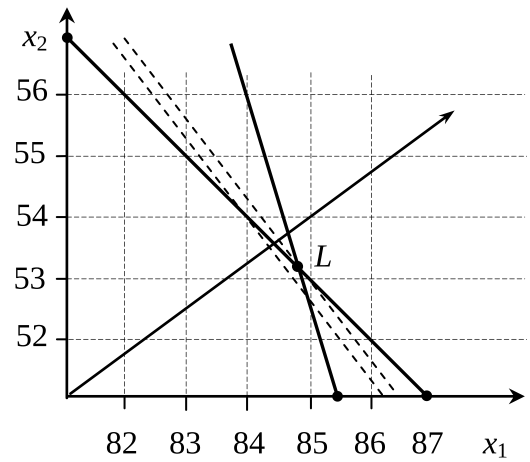


Рис. 1.12

Решение методом ветвей и границ

На первом шаге находим оптимальное решение без учета требования целочисленности переменных

$$f(x) = 2494,\ x_1 = 84{,}8,\ x_2 = 53{,}2,\ x_{1,2} \in R\,.$$

На втором шаге осуществляем ветвление по переменной с нецелым значением в оптимальном плане с наибольшим коэффициентом целевой функции $f(x) = 20x_1 + 15x_2$ или имеющей дробную часть, наиболее близкое к 0,5. Решаем две оптимизационные задачи. В первую добавляем ограничение $x_1 \le 84$. Во вторую добавляем ограничение $x_1 \ge 85$.

Решение первой задачи с помощью MS Excel [3, Лист 7]:

$$f(x) = 2494, x_1 = 84, x_2 = 54, x_{1,2} \in R.$$

Решение второй задачи с помощью MS Excel [3, Лист 8]:

$$f(x) = 2487{,}5, x_1 = 85, x_2 = 52{,}5, x_{1,2} \in R.$$

В первом случае получили целочисленное решение, во втором – нецелочисленное. Оптимальное целочисленное решение первой задачи больше оптимального решения второй задачи. Дальнейшее ветвление второй задачи по другой переменной не имеет смысла. Задача решена.

Ответ: $f(x) = 2494, x_1 = 84{,}8, x_2 = 53{,}2, x_{1,2} \in R$;

$$f(x) = 2490, x_1 = 84, \quad x_2 = 54, x_{1,2} \in N_0.$$

§ 4. Многокритериальная оптимизация

Эффективность функционирования системы определяется, как правило, несколькими критериями. В оптимизационных задачах математической формой эффективности является целевая функция. Некоторые критерии могут противоречить друг другу, другие – индифферентны, третьи действуют в одном направлении.

Пусть имеется k критериев, которые можно записать в виде целевой функции $f_i(x), i = \overline{1, k}$. Изменением знака целевой функции можно свести задачу на минимум к задаче на максимум, тогда задача многокритериальной оптимизации запишется как

$$F(x) = \{f_1(x), f_2(x), ..., f_k(x)\} \to \max$$

при ограничениях:

$$\begin{cases} g_j(x) \le b_j, \quad j = \overline{1, l}, \\ g_j(x) = b_j, \quad j = \overline{l+1, m}, \\ x \ge 0. \end{cases}$$

Если найденные оптимальные значения x_i^* для каждого критерия $f_i(x)$ не совпадают, то общее решение многокритериальной задачи может быть только компромиссным. Областью компромиссов называется подмножество допустимого множества решений x, обладающего тем свойством, что все принадлежащие ему решения не могут быть улучшены одновременно по всем локальным критериям $f_i(x)$.

Найденный допустимый план x_0 является оптимальным по Парето, если не существует другого плана x^*, для которого $f_i(x^*) \ge f_i(x_0), \ i = \overline{1, k}$, и хотя бы для одного критерия выполняется строгое неравенство.

Используют несколько методов решения задач многокритериальной оптимизации:

одного наиболее важного критерия, остальные критерии играют роль дополнительных ограничений;

последовательных уступок;

свертывания критериев, основанный на сведении всех критериев к одному введением весовых коэффициентов.

Пример 1.20. При изготовлении трех видов изделий *A*, *B* и *C* используют четыре вида ресурсов – P_1, P_2, P_3 и P_4. Известны запасы ресурсов и расход ресурсов на производство одного изделия:

Вид ресурса	Расход ресурсов на производство одного изделия			Запас ресурса, ед.
	A	*B*	*C*	
P_1	1	2	3	900
P_2	2	3	1	800
P_3	2	1	1	500
P_4	3	1	2	200

Все изделия обрабатываются на станках трех типов. Указаны норма времени на обработку одного изделия и фонд времени работы станков:

Вид станков	Норма времени на обработку одного изделия			Фонд времени, ч
	A	*B*	*C*	
Токарные	3	2	4	500
Фрезерные	4	5	3	600
Шлифовальные	2	3	1	300

Оптовая цена и себестоимость одного изделия составляют:

	Тип изделия		
	A	*B*	*C*
Оптовая цена, ден. ед.	12	13	8
Себестоимость, ден. ед.	3	5	2

Количество каждого типа изделия должен быть не менее 20. Критерии эффективности предприятия располагаются по степени их значимости:

1) прибыль;
2) валовой объем выпуска продукции в стоимостном выражении;
3) себестоимость продукции;
4) уровень загрузки оборудования.

Задачу необходимо решить двумя методами:

последовательных уступок, если уступку по каждому критерию полагать равной 10 % от его оптимального значения;

свертывания критериев, при этом вектор весовых коэффициентов выбирается равным (0,4; 0,3; 0,2; 0,1).

Решение

1. В качестве параметров управления выберем x_i – количество выпущенных изделий i-го вида.

2. Первая целевая функция должна максимизировать прибыль предприятия $f_1(x) = 9x_1 + 8x_2 + 6x_3 \to \max$. Прибыль вычисляется как разница между оптовой ценой и себестоимостью.

Вторая целевая функция максимизирует валовой объем в стоимостном выражении $f_2(x) = 12x_1 + 13x_2 + 8x_3 \to \max$.

Третья целевая функция минимизирует себестоимость $f_3(x) = 3x_1 + 5x_2 + 2x_3 \to \min$.

Четвертая целевая функция минимизирует уровень загрузки станков $f_4(x) = 9x_1 + 10x_2 + 8x_3 \to \min$.

3. Ограничения по расходу ресурсов:

$$\begin{cases} x_1 + 2x_2 + 3x_3 \le 900, \\ 2x_1 + 3x_2 + x_3 \le 800, \\ 2x_1 + x_2 + x_3 \le 500, \\ 3x_1 + x_2 + 2x_3 \le 200. \end{cases}$$

Ограничения по фонду времени работы станков:

$$\begin{cases} 3x_1 + 2x_2 + 4x_3 \le 500, \\ 4x_1 + 5x_2 + 3x_3 \le 600, \\ 2x_1 + 3x_2 + x_3 \le 300. \end{cases}$$

Необходимо добавить условие целочисленности и ограниченности переменных:

$$x_i \ge 20,\ x_i \in N,\ i = \overline{1,3}.$$

Находим оптимальное решение *методом последовательных уступок*.

1. Находим оптимальное решение однокритериальной задачи на максимум с целевой функцией $f_1(x) = 9x_1 + 8x_2 + 6x_3 \to \max$.

Решение с помощью MS Excel [3, Лист 9]:

$$f_1(x) = 980,\ f_2(x) = 1484,\ f_3(x) = 504,\ f_4(x) = 1196,\ x_1 = 20,\ x_2 = 76,\ x_3 = 32.$$

Определяем величину уступки по первому критерию $\Delta_1 = 980 \cdot 0{,}1 = 98$. Целевая функция $f_1(x)$ не должна быть меньше $980 - 98 = 882$.

2. Находим оптимальное решение однокритериальной задачи на максимум с целевой функцией $f_2(x)$ и дополнительным ограничением $f_1(x) \ge 882$.

Решение с помощью MS Excel [3, Лист 10]:

$$f_1(x) = 980,\ f_2(x) = 1484,\ f_3(x) = 504,\ f_4(x) = 1196,\ x_1 = 20,\ x_2 = 76,\ x_3 = 32.$$

Определяем величину уступки по второму критерию: $\Delta_2 = 1484 \cdot 0{,}1 = 148{,}4$. Целевая функция $f_2(x)$ не должна быть меньше $1484 - 148{,}4 = 1335{,}6$.

3. Находим оптимальное решение однокритериальной задачи на минимум с целевой функцией $f_3(x)$ и дополнительными ограничениями: $f_1(x) \geq 882,\ f_2(x) \geq 1335{,}6$.

Решение с помощью MS Excel [3, Лист 11]:

$$f_1(x) = 897,\ f_2(x) = 1336,\ f_3(x) = 439,\ f_4(x) = 1057,\ x_1 = 33,\ x_2 = 60,\ x_3 = 20.$$

Определяем величину уступки по третьему критерию $\Delta_3 = 439 \cdot 0{,}1 = 43{,}9$. Целевая функция $f_3(x)$ не должна быть больше $439 + 43{,}9 = 482{,}9$.

4. Находим оптимальное решение однокритериальной задачи на минимум с целевой функцией $f_4(x)$ и дополнительными ограничениями $f_1(x) \geq 882,\ f_2(x) \geq 1335{,}6,\ f_3(x) \leq 482{,}9.$

Решение с помощью MS Excel [3, Лист 12]:

$$f_1(x) = 897,\ f_2(x) = 1336,\ f_3(x) = 439,\ f_4(x) = 1057,\ x_1 = 33,\ x_2 = 60,\ x_3 = 20.$$

При решении задачи многокритериальной оптимизации методом последовательных уступок имеем: прибыль – 897, валовой объем – 1336, себестоимость – 439, загрузка оборудования – 1057. Предприятие должно выпускать изделий типа *A* в количестве 33, типа *B* – 60, типа *C* – 20.

Находим оптимальное решение *методом свертывания критериев*. Необходимо провести нормировку критериев по формуле

$$f_i^N(x) = \frac{f_i(x)}{f_i^*(x)},$$

где $f_i^*(x)$ – оптимальное решение задачи с одним критерием $f_i(x)$.

Оптимальные значения каждого критерия находим с помощью MS Excel [3, Лист 9]:

$$f_1^*(x) = 980,\ f_2^*(x) = 1484,\ f_3^*(x) = 200,\ f_4^*(x) = 540.$$

Составляем функцию свертки с учетом заданных весовых коэффициентов (0,4; 0,3; 0,2; 0,1):

$$F(x) = 0{,}4 f_1^N(x) + 0{,}3 f_2^N(x) - 0{,}2 f_3^N(x) - 0{,}1 f_4^N(x).$$

Решение с помощью MS Excel [3, Лист 13]:

$$f_1(x) = 700,\ f_2(x) = 980,\ f_3(x) = 280,\ f_4(x) = 782,\ x_1 = 46,\ x_2 = 20,\ x_3 = 21.$$

Если прибыль сделать приоритетной, например, взять весовые коэффициенты (0,6; 0,2; 0,1; 0,1), то решение с помощью MS Excel [3, Лист 14]:

$$f_1(x) = 965,\ f_2(x) = 1457,\ f_3(x) = 492,\ f_4(x) = 1151,\ x_1 = 29,\ x_2 = 73,\ x_3 = 20.$$

§ 5. Динамическое программирование

Метод динамического программирования является одним из главных методов оптимизации управления, специально приспособленным к «многошаговым» операциям.

Методы динамического программирования используются при решении задач распределения ресурсов между предприятиями, замене или ремонте оборудования, прокладке коммуникаций и т. д.

В основе решения задач динамического программирования лежит *принцип оптимальности*: каково бы ни было состояние системы S перед очередным шагом, нужно выбирать управление на этом шаге так, чтобы выигрыш на данном шаге плюс оптимальный выигрыш на всех последующих шагах был максимальным. Математически этот принцип записывается в виде *рекуррентного соотношения Беллмана*:

$$f_i(S) = \max\{\varphi_i(S, x_i) + f_{i+1}(S_i)\},$$

где $\varphi_i(S, x_i)$ – функция выигрыша (прибыли) при использовании управления x_i на i-м шаге; x_i – все допустимые управления при условии, что система находится в состоянии S; S_i – следующее состояние, в которое переходит система под воздействуем управления x_i; $f_i(S)$ – эффект за оставшиеся i шагов.

Вычисления в задачах динамического программировании выполняют обычно от конца к началу. Таким способом находят условно-оптимальные решения на каждом шаге. Затем выполняют вычисления от первого шага к последнему, учитывая полученные условно-оптимальные решения.

Рассмотрим задачу распределения ресурса между предприятиями.

Пример 1.21. Имеется три предприятия, между которыми необходимо распределить пять единиц ресурсов с максимальным доходом от их размещения. Приведены функции дохода φ_i от вложения денежных средств в каждое предприятие:

x	$\varphi_1(x)$	$\varphi_2(x)$	$\varphi_3(x)$
1	10	5	2
2	14	9	6
3	16	13	12
4	17	16	14
5	17	18	15

Решение

Рассмотрим пошаговый алгоритм решения. Количество шагов будет равно количеству предприятий. Начнем с последнего предприятия.

На первом шаге распределяем денежные средства только третьему предприятию.

$$f_3(1) = \varphi_3(1) = 2,\ f_3(2) = \varphi_3(2) = 6,\ f_3(3) = \varphi_3(3) = 12,\ f_3(4) = \varphi_3(4) = 14,$$

$$f_3(5) = \varphi_3(5) = 15.$$

На втором шаге распределяем денежные средства между вторым и третьим предприятиями:

$$f_2(1) = \max\{\underline{\varphi_2(1)};\ f_3(1)\} = \max\{5;\ 2\} = 5,$$

$$f_2(2) = \max\{\underline{\varphi_2(2)};\ \varphi_2(1) + f_3(1);\ f_3(2)\} = \max\{9;\ 5+2;\ 6\} = 9;$$

$$f_2(3) = \max\{\underline{\varphi_2(3)};\ \varphi_2(2) + f_3(1);\ \varphi_2(1) + f_3(2);\ f_3(3)\} = \max\{13;\ 9+2;\ 5+6;\ 12\} = 13;$$

$$f_2(4) = \max\{\varphi_2(4);\ \varphi_2(3) + f_3(1);\ \varphi_2(2) + f_3(2);\ \underline{\varphi_2(1) + f_3(3)};\ f_3(3)\} = \max\{16;\ 13+2;\ 9+6;\ 5+12;\ 14\} = 17;$$

$$f_2(5) = \max\{\varphi_2(5);\ \varphi_2(4) + f_3(1);\ \varphi_2(3) + f_3(2);\ \underline{\varphi_2(2) + f_3(3)};\ \varphi_2(1) + f_3(4);\ f_3(5)\} = \max\{18;\ 16+2;\ 13+6;\ 9+12;\ 5+14;\ 15\} = 21.$$

На третьем шаге распределяем денежные средства между первым и вторым предприятием:

$$f_1(1) = \max\{\underline{\varphi_1(1)};\ f_2(1)\} = \max\{10;\ 5\} = 10;$$

$$f_1(2) = \max\{\underline{\varphi_1(2)};\ \varphi_1(1) + f_2(1);\ f_2(2)\} = \max\{14;\ 10+5;\ 9\} = 14;$$

$$f_1(3) = \max\{\varphi_1(3);\ \underline{\varphi_1(2) + f_2(1)};\ \underline{\varphi_1(1) + f_2(2)};\ f_2(3)\} = \max\{16;\ 14+5;\ 10+9;\ 13\} = 19;$$

$$f_1(4) = \max\{\varphi_1(4);\ \varphi_1(3) + f_2(1);\ \underline{\varphi_1(2) + f_2(2)};\ \underline{\varphi_1(1) + f_2(3)};\ f_2(4)\} = \max\{17;\ 16+5;\ 14+9;\ 10+13;\ 17\} = 23;$$

$$f_1(5) = \max\{\varphi_1(5);\ \varphi_1(4) + f_2(1);\ \varphi_1(3) + f_2(2);\ \underline{\varphi_1(2) + f_2(3)};\ \underline{\varphi_1(1) + f_2(4)};\ f_2(5)\} = \max\{17;\ 17+5;\ 16+9;\ 14+13;\ 10+17;\ 21\} = 27.$$

Выполним обратные вычисления и уточним распределение денежных средств. Максимальный доход от вложенных средств может быть реализован в двух вариантах: $\varphi_1(2) + f_2(3) = \varphi_1(2) + \varphi_2(3)$ и $\varphi_1(1) + f_2(4) = \varphi_1(1) + \varphi_2(1) + f_3(3) = \varphi_1(1) + \varphi_2(1) + \varphi_3(3)$. По первому плану первому предприятию необходимо выделить две единицы, а второму – три. По второму плану первому предприятию необходимо выделить одну единицу, второму – одну, третьему – три.

§ 6. Теория игр

6.1. Основные понятия теории игр

В предыдущих разделах рассматривались задачи, в которых решение принималось одним лицом. На практике часто приходится решать задачи, в которых сталкиваются интересы двух и более конкурирующих сторон. Такие задачи называются конфликтными. Математическая дисциплина, изучающая конфликтные ситуации, называется *теорией игр*.

В экономике конфликтные ситуации встречаются довольно часто, например взаимоотношения между покупателем и продавцом, поставщиком и потребителем, банком и клиентом.

Игра, в которой участвуют два игрока, называется *парной*. Если же число игроков больше двух, то игру называют *множественной*. Парная игра, в которой выигрыш одного игрока равен проигрышу другого, называется *игрой с нулевой суммой* или *антагонистической*.

Совокупность правил, которые в зависимости от ситуации определяют выбор действий игрока, называется *стратегией игры*. Принятие игроком того или иного решения называется *ходом*. Ходы могут быть личные (сознательные) и случайные.

В зависимости от числа возможных стратегий игры делятся на *конечные*, если каждый игрок имеет конечное число стратегий, и *бесконечные*, если хотя бы один из игроков имеет бесконечное число стратегий.

Стратегия, приносящая игроку максимальный выигрыш, называется *оптимальной*.

Рассмотрим парную конечную игру. Предположим, что имеются две конкурирующие фирмы (игроки) A и B, выпускающие однотипные товары. Пусть фирма A разработала три стратегии, которые обозначим через A_1, A_2, A_3; фирма B – четыре стратегии B_1, B_2, B_3 и B_4. В результате выбора игроками любой пары стратегий A_i и B_j определяется исход игры. Если фирма A от реализации своей стратегии получит доход (выигрыш положительный), равный a_{ij}, то фирма B потерпит убыток, и наоборот. Матрица (a_{ij}), элементами которой являются выигрыши, соответствующие стратегиям A_i и B_j, называется *платежной* матрицей или матрицей игры. Для нашего примера платежная матрица имеет вид:

Стратегии	B_1	B_2	B_3	B_4
A_1	a_{11}	a_{12}	a_{13}	a_{14}
A_2	a_{21}	a_{22}	a_{23}	a_{24}
A_3	a_{31}	a_{32}	a_{33}	a_{34}

Строки этой таблицы соответствуют стратегиям фирмы A, а столбцы – стратегиям фирмы B.

Пусть α_i – минимальное значение в i-й строке (минимально возможный выигрыш при принятии i-й стратегии игроком A), $\alpha_i = \min\limits_{j=1,2,3,4} a_{ij}$. Среди всех α_i выберем наибольшее, $\alpha = \max\limits_{i=1,2,3} \alpha_i$, – это максимальный гарантированный выигрыш (максимин) независимо от выбора стратегии игрока B. Значение α называют *нижней ценой игры*.

Игрок B заинтересован в том, чтобы уменьшить выигрыш игрока A. Пусть β_j – это максимальное значение в j-м столбце (максимально возможный выигрыш игрока A при принятии j-й стратегии игроком B),

$\beta_j = \max\limits_{i=1,2,3} a_{ij}$. Среди всех β_j выберем наименьшее, $\beta = \min\limits_{j=1,2,3,4} \beta_j$, – минимальный гарантированный проигрыш (минимакс) игрока B. Значение β называют *верхней ценой игры*.

Если нижняя и верхняя цена игры совпадают ($\alpha = \beta = \nu$), то их общее значение называют *ценой игры*, а саму игру называют игрой с *седловой* точкой. Решение игры будет обладать устойчивостью. Минимаксные стратегии, соответствующие цене игры, называются оптимальными стратегиями. При принятии оптимальных стратегий игрок A получает максимальный гарантированный выигрыш ν, а игрок B – минимальный гарантированный проигрыш ν. Если, например, игрок A придерживается своей оптимальной стратегии, то игроку B невыгодно отклоняться от своей оптимальной стратегии.

Пример 1.22. Определить, существует ли седловая точка матрицы игры с нулевой суммой

$$\begin{pmatrix} 4 & 3 & 6 \\ 5 & 2 & 2 \end{pmatrix}.$$

Решение

Находим минимальные значения в каждой строке и максимальные значения в каждом столбце матрицы:

$$\alpha_1 = \min_{j=1,2,3} a_{1j} = \min(4, 3, 6) = 3,\ \alpha_2 = \min_{j=1,2,3} a_{2j} = \min(5, 2, 2) = 2,$$

$$\beta_1 = \max_{i=1,2} a_{i1} = \max(4, 5) = 5,\ \beta_2 = \max_{i=1,2} a_{i2} = \max(3, 2) = 3,\ \beta_3 = \max(6, 2) = 6.$$

Находим максимальное значение α_i и минимальное значение β_j:

$$\alpha = \max_{i=1,2} a_i = \max(3, 2) = 3,\ \beta = \min_{j=1,2,3} \beta_j = \min(5, 3, 6) = 3.$$

Результаты вычислений:

<table>
<tr><td></td><td>B_1</td><td>B_2</td><td>B_3</td><td>$\alpha_i = \min\limits_{j=1,2,3} a_{ij}$</td><td>$\alpha = \max\limits_{i=1,2} \alpha_i$</td></tr>
<tr><td>A_1</td><td>4</td><td>3</td><td>6</td><td>3</td><td rowspan="2">3</td></tr>
<tr><td>A_2</td><td>5</td><td>2</td><td>2</td><td>2</td></tr>
<tr><td>$\beta_j = \max\limits_{i=1,2} a_{ij}$</td><td>5</td><td>3</td><td>6</td><td colspan="2" rowspan="2">$\alpha = 3$
$\beta = 3$</td></tr>
<tr><td>$\beta = \min\limits_{j=1,2,3} \beta_j$</td><td colspan="3">3</td></tr>
</table>

Равенство $\alpha = \beta = 3$ означает, что цена игры $\nu = 3$ и игра имеет седловую точку (стратегии A_1 и B_2). Игрок A должен придерживаться первой стратегии, а игрок B – второй.

Если матрица игры содержит одинаковые строки (столбцы), то из них оставляют только одну строку (столбец), а другую строку (столбец) вычеркивают. В этом случае говорят о *дублировании* стратегий.

Если одна строка поэлементно не меньше другой, то первая строка называется *доминирующей* по отношению ко второй. Игрок A не использует стратегию второй строки.

Если один столбец поэлементно не меньше другого столбца, то второй столбец называется *доминирующим* по отношению к первому столбцу. Игрок B не использует стратегию первого столбца.

Стратегия, над которой доминирует другая стратегия, отбрасывается. Отбрасываем «меньшую» строку и «больший» столбец. На цену игры отбрасывание дублирующих и доминантных стратегий не влияет, а размер матрицы игры уменьшается.

Если конечная игра имеет седловую точку, то после упрощения получится платежная матрица 1×1.

В примере 1.22 второй столбец доминирует над первым (3 < 4, 2 < 5), поэтому отбрасываем первый столбец. Получим матрицу

$$\begin{pmatrix} 3 & 6 \\ 2 & 2 \end{pmatrix}.$$

Первая строка доминирует над второй (3 > 2, 6 > 2), поэтому отбрасываем вторую строку. Получим матрицу $\begin{pmatrix} 3 & 6 \end{pmatrix}$.

Первый столбец доминирует над вторым (3 < 6), поэтому отбрасываем второй столбец. Получим матрицу (3).

Пример 1.23. Игрок A записывает одно из двух чисел: 1 или 2, игрок B – одно из трех чисел: 3, 4 или 5. Если оба числа одинаковой четности, то игрок A выигрывает и выигрыш равен сумме этих чисел; если четности выбранных игроками чисел не совпадают, то игрок B выигрывает и выигрыш равен сумме этих чисел. Составить платежную матрицу игры, определить верхнюю и нижнюю цену игры.

Решение

Если A и B пишут соответственно 1 и 3, то 1 + 3 = 4 – это выигрыш игрока A. Если A и B пишут соответственно 1 и 4, то 1 + 4 = 5 – проигрыш игрока A и т. д. В итоге получаем платежную матрицу

$$\begin{pmatrix} 4 & -5 & 6 \\ -5 & 6 & -7 \end{pmatrix}.$$

Найдем верхнюю и нижнюю цены игры:

	B_1	B_2	B_3	$\alpha_i = \min\limits_{j=1,2,3} a_{ij}$	$\alpha = \max\limits_{i=1,2} \alpha_i$
A_1	4	–5	6	–5	–5
A_2	–5	6	–7	–7	
$\beta_j = \max\limits_{i=1,2} a_{ij}$	4	6	6		$\alpha = -5$
$\beta = \min\limits_{j=1,2,3} \beta_j$	4			$\beta = 4$	

Нижняя цена игры $\alpha = -5$, а верхняя $\beta = 4$. Задача не имеет седловой точки. Цена игры $\alpha \le \nu \le \beta$, $-5 \le \nu \le 4$.

Если конечная матричная игра не имеет седловой точки, то используют *смешанные* стратегии. Для игрока *A* каждая стратегия имеет вероятность (частоту) p_i применения i-й стратегии. Для игрока *B* каждая стратегия имеет вероятность (частоту) q_j применения j-й стратегии. Тогда средний выигрыш игрока *A* будет равен $\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} p_i q_j$.

Чистая стратегия, входящая в оптимальную смешанную стратегию, называется *активной*. Если один из игроков придерживается своей оптимальной стратегии, то выигрыш остается неизменным и равным цене игры ν, если второй игрок не выходит за пределы своих активных стратегий. Фиксированная стратегия (использование только одной стратегии) одного из игроков в игре без седловой точки не обладает свойством устойчивости.

Смешанные стратегии представляют собой модель гибкой тактики. При использовании этой тактики противник не знает заранее, с какой ситуацией ему предстоит столкнуться. Поскольку игроки выбирают свои чистые стратегии случайно и независимо друг от друга, игра имеет случайный характер и случайной становится величина выигрыша.

Использование в игре оптимальных смешанных стратегий обеспечивает игроку *A* не меньший выигрыш, чем при использовании им любой другой стратегии.

Теорема фон Неймана. Каждая конечная матричная игра имеет по крайней мере одно оптимальное решение, возможно, среди смешанных стратегий.

6.2. Графический способ решения игр с нулевой суммой

6.2.1. Игра 2×2

Рассмотрим платежную матрицу $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Пусть с вероятностью p игрок A использует первую стратегию, тогда с вероятностью $1 - p$ он использует вторую стратегию. Средний выигрыш игрока A при применении игроком *B* первой стратегии $w_1(p) = a_{11}p + a_{21}(1-p)$. Средний выигрыш игрока A при применении игроком B второй стратегии $w_2(p) = a_{12}p + a_{22}(1-p)$. Оптимальное значение смешанной стратегии игрока *A* получим из уравнения $w_1(p) = w_2(p)$ (пересечение двух прямых):

$$a_{11}p + a_{21} - a_{21}p = a_{12}p + a_{22} - a_{22}p,\ (a_{11} + a_{22} - a_{12} - a_{21})p = a_{22} - a_{21},$$

$$p = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}},\ 1 - p = \frac{a_{11} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \qquad (1)$$

для случая $a_{11} + a_{22} - a_{12} - a_{21} \ne 0$.

Средний выигрыш игрока *B* при применении игроком *A* первой стратегии $w_1(q) = a_{11}q + a_{12}(1-q)$. Средний выигрыш игрока *B* при применении игроком *A* второй стратегии $w_2(q) = a_{21}q + a_{22}(1-q)$. Оптимальное значение смешанной стратегии игрока *B* получим из уравнения $w_1(q) = w_2(q)$ (пересечение двух прямых):

$$a_{11}q + a_{12} - a_{12}q = a_{21}q + a_{22} - a_{22}q,\ (a_{11} + a_{22} - a_{12} - a_{21})q = a_{22} - a_{12},$$

$$q = \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}},\ 1 - q = \frac{a_{11} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} \qquad (2)$$

для случая $a_{11} + a_{22} - a_{12} - a_{21} \neq 0$.

Если хотя бы одна из величин *p*, 1 – *p*, *q*, 1 – *q* окажется отрицательной, то это значит, что имеется решение в чистых стратегиях.

Пример 1.24. Найдите решение матричной игры $\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix}$.

Решение

Найдем оптимальную смешанную стратегию игрока *A* по формуле (1):

$$p = \frac{4-1}{5+4-2-1} = \frac{1}{2},\ 1 - p = \frac{5-2}{5+4-2-1} = \frac{1}{2}.$$

Каждую из стратегий игроку *A* нужно применять с частотой 1/2.

Найдем оптимальную смешанную стратегию игрока *B* по формуле (2):

$$q = \frac{4-2}{6} = \frac{1}{3},\ 1 - q = \frac{5-1}{6} = \frac{2}{3}.$$

Первую стратегию игрок *B* должен применять с частотой 1/3, а вторую – с частотой 2/3.

6.2.2. Игра 2×n

При игре 2×*n* игрок *B* имеет *n* стратегий. При нахождении оптимального значения смешанной стратегии игрока *A* графическим способом необходимо построить *n* прямых на промежутке (0; 1), а также нижнюю огибающую. Нижняя огибающая – это такая ломаная, состоящая из отрезков прямых, выше которой лежит вся картинка. Вершина нижней огибающей с наибольшей ординатой является оптимальным значением смешанной стратегии игрока *A*.

Пример 1.25. Генерал *A*, командующий тремя полками, преодолевает горный хребет. Горный хребет имеет два перевала, которые охраняет генерал *B*, командующий двумя полками. При встрече *n* полков генерала *A* с *m* полками генерала *B* на одном из перевалов прорывается *n* – *m* полков генерала *A*, если *n* > *m*, и ни одного, если *n* ≤ *m*. Генералу *A* необходимо сохранить как можно больше полков. Генерал *B* должен пропустить как можно меньше полков. Каждый полк преодолевает и охраняет перевал в полном составе. Каждый потерянный полк генерал A оценивает как «–1».

Решение

Генерал A имеет четыре стратегии игры:

A_1 – послать на первый перевал все три полка;

A_2 – послать на первый перевал два полка, а на второй перевал – один полк;

A_3 – послать на первый перевал один полк, а на второй перевал – два полка;

A_4 – послать на третий перевал все три полка.

Стратегии генерала B:

B_1 – охранять первый перевал двумя полками;

B_2 – охранять первый перевал одним полком, а другой перевал – другим полком;

B_3 – охранять только второй перевал двумя полками.

Составим платежную матрицу игры генерала A:

$A \backslash B$	$B_1(2, 0)$	$B_2(1, 1)$	$B_3(0, 2)$
$A_1(3, 0)$	–2	–1	0
$A_2(2, 1)$	–2	–2	–1
$A_3(1, 2)$	–1	–2	–2
$A_4(0, 3)$	0	–1	–2

Если к каждому элементу платежной матрицы добавить три единицы, то получим платежную матрицу количества полков генерала A, прошедших через горный хребет, причем это не повлияет на выбор оптимальных стратегий генералов, а цена игры увеличится на три.

Стратегия A_1 доминирует над A_2, а стратегия A_4 – над A_3. Вычеркиванием стратегии A_2 и A_3 игра сводится к игре 2×3:

$A \backslash B$	$B_1(2, 0)$	$B_2(1, 1)$	$B_3(0, 2)$
$A_1(3, 0)$	–2	–1	0
$A_4(0, 3)$	0	–1	–2

Пусть с вероятностью p игрок A использует стратегию A_1, тогда с вероятностью 1 – p он использует стратегию A_4. Средний выигрыш игрока A при применении игроком B первой стратегии $w_1(p) = -2p$. Средний выигрыш игрока A при применении игроком B второй стратегии

$w_2(p) = -p - 1(1-p) = -1$. Средний выигрыш игрока *А* при применении игроком *В* третьей стратегии $w_3(p) = -2(1-p) = -2 + 2p$.

Изобразим их графики (рис. 1.13).

Ломаная *АВС* – нижняя огибающая. Точка *В* – это точка с наибольшей ординатой нижней огибающей. Точка *В* является точкой пересечения трех прямых, w_1, w_2 и w_3. Найдем абсциссу точки *В*:

$$w_1 = w_2\ ,\ -2p = -1,\ p = \frac{1}{2},\ 1 - p = \frac{1}{2}.$$

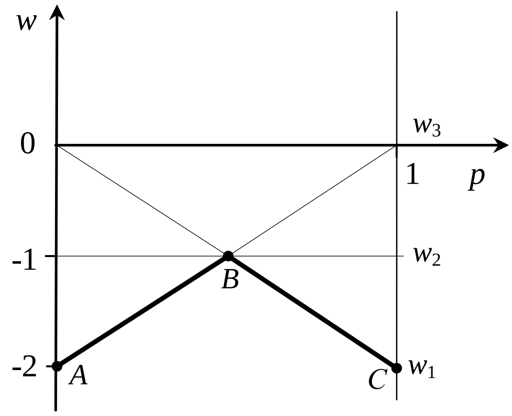


Рис. 1.13

Ответ: с вероятностью $p = 1/2$ генерал *А* должен использовать стратегию A_1 и с вероятностью $1 - p = 1/2$ – стратегию A_4. Стратегии A_2 и A_3 генерал *А* не должен использовать. Данное решение будет оптимальным только в случае многократного повторения игры. Для одношаговой задачи нет оптимальной чистой стратегии. Игрок *А* должен выбрать стратегию A_1 либо стратегию A_4 случайным образом, например подбрасывая монету. Результат одношаговой игры непредсказуем и зависит от хода игрока *В*.

6.2.1. Игра m×2

При игре $m\times2$ игрок *А* имеет m стратегий. При нахождении оптимального значения смешанной стратегии игрока *В* графическим способом необходимо построить m прямых на промежутке (0; 1), а также построить верхнюю огибающую. Верхняя огибающая – это такая ломаная, состоящая из отрезков прямых, ниже которой лежит вся картинка. Вершина с наименьшей ординатой верхней огибающей является оптимальным значением смешанной стратегии игрока *В*.

Пример 1.26. Найдите решение матричной игры $\begin{pmatrix} 1 & 5 \\ 4 & 3 \\ 7 & 2 \end{pmatrix}$.

Решение

Пусть с вероятностью q игрок *В* использует первую стратегию, тогда с вероятностью 1 – q он использует вторую стратегию. Средний выигрыш игрока *В* при применении игроком *А* первой стратегии $w_1(q) = q + 5(1-q) = 5 - 4q$. Средний выигрыш игрока *В* при применении игроком *А* второй стратегии $w_2(q) = 4q + 3(1-q) = 3 + q$. Средний выигрыш игрока *В* при применении игроком *А* третьей стратегии $w_3(q) = 7q + 2(1-q) = 2 + 5q$.

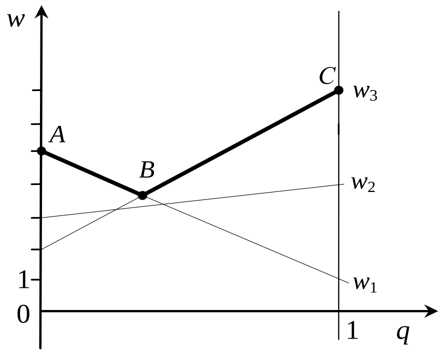


Рис. 1.14

Изобразим их графики (рис. 1.14).

Ломаная *ABC* – верхняя огибающая. Точка *B* – это точка с наименьшей ординатой верхней огибающей. Точка *B* является точкой пересечения прямых w_1 и w_3. Найдем абсциссу точки *B*:

$$w_1(q) = w_3(q),\ 5 - 4q = 2 + 5q,\ q = \frac{1}{3},$$

$$1 - q = \frac{2}{3}.$$

Цена игры $v = w_1\left(\frac{1}{3}\right) = w_2\left(\frac{1}{3}\right) = \frac{11}{3} = 3\frac{2}{3}$.

Найдем решение для игрока *A*. Точка *B* является точкой пересечения прямых w_1 и w_3, поэтому игрок *A* не должен пользоваться второй стратегией в оптимальном решении.

Пусть с вероятностью *p* игрок *A* использует стратегию A_1, тогда с вероятностью 1 – *p* он использует стратегию A_3. Средний выигрыш игрока *A* при применении игроком *B* первой стратегии составит $w_1(p) = p + 7(1 - p) = 7 - 6p$. Средний выигрыш игрока *A* при применении игроком *B* третьей стратегии $w_3(p) = 5p + 2(1 - p) = 2 + 3p$. Найдем *p*:

$$w_1(p) = w_3(p),\ 7 - 6p = 2 + 3p,\ p = \frac{5}{9},\ 1 - p = \frac{4}{9}.$$

Зная цену игры, можно было иначе найти *p*:

$$w_1(p) = v,\ 7 - 6p = \frac{11}{3},\ 6p = \frac{10}{3},\ p = \frac{5}{9}.$$

6.3. Матричная игра и метод линейного программирования

Рассмотрим решение матричной игры с нулевой суммой с помощью линейного программирования. Любая игра двух лиц с нулевой суммой эквивалентна некоторой ЗЛП. Обозначим через p_i вероятность выбора игроком *A* стратегии A_i и через q_j – вероятность выбора игроком *B* стратегии B_j. Если игрок *A* будет строго применять оптимальную стратегию, то игрок *B* не сможет изменить ход игры без дополнительного ущерба для себя. Оптимальная стратегия игрока *A* гарантирует ему выигрыш v. При любой чистой стратегии B_j игрока *B* игрок *A* получает средний выигрыш $v_j = \sum_{i=1}^{m} a_{ij} p_i$. Задача об определении оптимальной смешанной стратегии для игрока *A* сводится к тому, чтобы все средние выигрыши были не меньше цены игры v.

В результате получаем задачу линейного программирования в виде:

$$\nu = \sum_{j=1}^{n} \nu_j \to \max,$$

$$\begin{cases} \sum\limits_{i=1}^{m} a_{ij} p_i \ge \nu, \ \ j = \overline{1,n}, \\ \sum\limits_{i=1}^{m} p_i = 1, \qquad p_i \ge 0 \ . \end{cases}$$

Разделим каждое ограничение на число $\nu > 0$ и введем новые переменные: $x_i = \dfrac{p_i}{\nu}$. Тогда последнее ограничение системы примет вид $x_1 + x_2 + ... + x_m = \dfrac{1}{\nu}$. Максимизация цены игры ν равносильна минимизации величины $\dfrac{1}{\nu}$. Задача линейного программирования примет вид:

$$f(x) = \sum_{i=1}^{m} x_i \to \min,$$

$$\begin{cases} \sum\limits_{i=1}^{m} a_{ij} x_i \ge 1, \ \ j = \overline{1,n}, \\ \qquad x_i \ge 0. \end{cases}$$

После нахождения оптимального решения $f^*(x)$, цену игры определяют из соотношения $\nu = \dfrac{1}{f^*(x)}$.

Для игрока *B* математическая модель примет вид:

$$f(x) = \sum_{j=1}^{n} y_j \to \max,$$

$$\begin{cases} \sum\limits_{j=1}^{n} a_{ij} y_j \le 1, \ \ i = \overline{1,m}, \\ \qquad y_j \ge 0. \end{cases}$$

Рассмотрим решение примера 1.23 о записи цифр. Поскольку нижняя цена игры $\alpha = -5$, что меньше нуля, прибавим к каждому элементу платежной матрицы число большее $\alpha = -5$ по абсолютному значению, например число $\gamma = 7$. Получим платежную матрицу

$$\begin{pmatrix} 11 & 2 & 13 \\ 2 & 13 & 0 \end{pmatrix}.$$

Составим математическую модель модифицированной задачи с учетом новой платежной матрицы:

$$f(x) = x_1 + x_2 \to \min,$$

$$\begin{cases} 11x_1 + 2x_2 \geq 1, \\ 2x_1 + 13x_2 \geq 1, \\ \quad 13x_1 \geq 1, \\ \quad x_1, x_2 \geq 0. \end{cases}$$

Решение с помощью MS Excel [3, Лист 15]:

$f^*(x)$ = 0,143884892, x_1 = 0,079136691, x_2 = 0,064748201.

Цена модифицированной игры $v^* = \dfrac{1}{f^*(x)} = 6{,}95$. Вероятности для оптимальной смешанной стратегии игрока A:

$$p_1 = v^* x_1 = 0{,}55, \; p_2 = v^* x_2 = 0{,}45.$$

Цена исходной задачи $v = v^* - \gamma = 6{,}95 - 7 = -0{,}05$.

Итак, с вероятностью $p_1 = 0{,}55$ игрок A должен писать цифру 1 и с вероятностью $p_2 = 0{,}45$ – цифру 2. При этом в среднем он будет проигрывать 0,05. Причем свои оптимальные стратегии игрок A должен осуществлять в случайном порядке. Данная игра невыгодна игроку A. Если же игрок B не владеет азами теории игр, то у игрока A появляется шанс.

Если, например, игра повторяется 20 раз, то игрок A должен написать $20 \cdot 0{,}55 = 11$ раз единицу и $20 \cdot 0{,}45 = 9$ раз двойку, причем это сделать в произвольном порядке. В результате проигрыш (наихудший вариант) составит $20 \cdot 0{,}05 = 1$.

Например, если игрок B всегда выбирает третью стратегию, то выигрыш игрока A составит $11 \cdot 6 + 9 \cdot (-7) = 3$.

Самостоятельно найдите оптимальное решение для игрока B.

Пример 1.27. Две фирмы выделяют денежные средства (тыс. руб.) на строительство двух объектов. Каждая из фирм имеет по две стратегии финансирования. Прибыль фирмы A в зависимости от объема финансирования выражена матрицей

$$\begin{pmatrix} 40 & 30 \\ 20 & 50 \end{pmatrix}.$$

Требуется найти оптимальные стратегии фирм. Предполагается, что убыток предприятия B равен прибыли фирмы A.

Решение

Обозначим чистые стратегии фирм A и B через A_1, A_2 и B_1, B_2 соответственно. Тогда чистая стратегия A_1 – это выделение a_1 тыс. руб. фирмой A на строительство первого объекта; чистая стратегия A_2 – это выделение a_2 тыс. руб. фирмой A на строительство второго объекта. Найдем верхнюю и нижнюю цену игры:

	B_1	B_2	$\alpha_i = \min\limits_{j=1,2} a_{ij}$	$\alpha = \max\limits_{i=1,2} \alpha_i$
A_1	40	30	30	30
A_2	20	50	20	
$\beta_j = \max\limits_{i=1,2} a_{ij}$	40	50	$\beta = 40$	$\alpha = 30$
$\beta = \min\limits_{j=1,2} \beta_j$	40			

Нижняя цена игры $\alpha = 30$, а верхняя $\beta = 40$. Задача не имеет седловой точки, поэтому решение игры определяется в смешанных стратегиях. Цена игры заключена между α и β, т. е. $30 \le \nu \le 40$. Составим математическую модель для игрока A:

$$f(x) = x_1 + x_2 \to \min,$$

$$\begin{cases} 40x_1 + 20x_2 \ge 1, \\ 30x_1 + 50x_2 \ge 1, \\ x_{1,\,2} \ge 0. \end{cases}$$

Решение с помощью MS Excel [3, Лист 16]:

$$f^*(x) = 0{,}028571429,\ x_1 = 0{,}021428571,\ x_2 = 0{,}007142857.$$

Цена $\nu = \dfrac{1}{f(x)} = 35$. Вероятности для оптимальных смешанных стратегий игрока A: $p_1 = \nu \cdot x_1 = 0{,}75$, $p_2 = \nu \cdot x_2 = 0{,}25$ при гарантированном получении фирмой A, независимо от стратегий фирмы B, прибыли не менее 35 тыс. руб. Убыток фирмы B при этом составит не более 35 тыс. руб.

Таким образом, фирма A из выделяемых средств должна направить на строительство первого объекта 75 %, а второго – 25 %.

6.4. Биматричные игры

Антагонистические игры встречаются редко. Гораздо чаще интересы игроков A и B не противоположны. В этом случае у каждого игрока имеется своя платежная матрица игры, поэтому игра называется *биматричной*. Антагонистическая игра является частным случаем биматричной.

6.4.1. Равновесие в доминирующих стратегиях обоих игроков

Рассмотрим равновесие доминирующих стратегий на примере.

Две оптовые базы реализуют одинаковый товар. Каждая из них может продавать товар по прежней цене или повысить отпускную цену товара. Платежная матрица игры имеет вид:

	Новая цена	Прежняя цена
Новая цена	(9, 10)	(2, 13)
Прежняя цена	(12, 3)	(7, 8)

Если игрок *B* выберет первую стратегию, то игроку *A* выгодно выбрать вторую стратегию. Отметим * лучший результат в первом столбце для игрока *A*. Если игрок *B* выберет вторую стратегию, то игроку *A* тоже выгодно выбрать вторую стратегию. Отметим * лучший результат во втором столбце для игрока *A*. Аналогично для игрока *B* выбираем лучший результат для каждой строки платежной матрицы игрока *B*:

	Новая цена	Прежняя цена
Новая цена	(9, 10)	(2, 13*)
Прежняя цена	(12*, 3)	(7*, 8*)

У игрока *A* вторая стратегия доминирует над первой (9 < 12, 2 < 7). У игрока *B* – аналогичная ситуация (10 < 13, 3 < 8). Пересечение доминирующей строки игрока *A* и доминирующего столбца игрока *B* представляет собой точку равновесия в доминирующих стратегиях. Точка равновесия в таблице отмечена двойным знаком **. Независимо от действий другого игрока каждый из них имеет гарантированную прибыль: игрок *A* – 7, игрок *B* – 8.

Равновесие по Нэшу. Пара чистых стратегий i_0 и j_0 в биматричной игре игроков *A* и *B* называется равновесием по Нэшу в чистых стратегиях, если при выборе игроком *B* стратегии j_0 для игрока *A* наиболее выгодной является стратегия i_0, и наоборот.

Существуют и другие способы определения оптимального решения игры, например, *равновесие по Парето*: всякое изменение, которое никому не приносит убытков, а некоторым людям приносит пользу (по их собственной оценке), является улучшением. Таким образом, признается право на все изменения, которые никому не приносят дополнительного вреда. Чтобы определить оптимум по Парето, необходимо перебрать все исходы игры.

В данном примере оптимальным решением по Парето является ситуация (A_1, B_1), (A_1, B_2) и (A_2, B_1). Для игр с ненулевой суммой оптимальное решение по Парето может не совпадать с равновесием по Нэшу.

6.4.2. Равновесие с доминирующей стратегией одного игрока

Рассмотрим равновесие с доминирующей стратегией одного игрока на примере.

Два предприятия могут производить однотипный товар разного качества и реализовывать его в разных районах города. Платежная матрица игры имеет вид:

	Высокое качество	Низкое качество
Высокое качество	(8, 9)	(9, 13)
Низкое качество	(13, 12)	(10, 11)

Если игрок *B* выберет первую стратегию, то игроку *A* выгодно выбрать вторую стратегию. Отметим звездочкой лучший результат в первом столбце для игрока *A*. Если игрок *B* выберет вторую стратегию, то игроку *A* также выгодно выбрать вторую стратегию. Отметим звездочкой лучший результат во втором столбце для игрока *A*. Аналогично для игрока *B* выбираем лучший результат для каждой строки платежной матрицы игрока *B*:

	Высокое качество	Низкое качество
Высокое качество	(8, 9)	(9, 13*)
Низкое качество	(13*, 12*)	(10*, 11)

У игрока *A* вторая стратегия доминирует над первой (8 < 13, 9 < 10). У игрока *B* нет доминирующей стратегии (звездочки расположены в разных столбцах). Игрок *A* независимо от действий игрока *B* выбирает вторую доминирующую стратегию. В этом случае игрок *B* выбирает первую стратегию (12 > 11). Точка равновесия в таблице отмечена двойным знаком **. Получаем *равновесие по Нэшу*: если один игрок не меняет своей стратегии, то и другому менять стратегию невыгодно. Оптимальным решением по Парето является ситуация (A_1, B_2) и (A_2, B_1).

6.4.3. Равновесие без доминирующих стратегий

Рассмотрим на примере равновесие без доминирующих стратегий биматричной игры, платежная матрица которой имеет вид:

	Высокое качество	Низкое качество
Высокое качество	(12, 12)	(15, 10)
Низкое качество	(10, 15)	(16, 16)

Если игрок *B* выберет первую стратегию, то игроку *A* тоже выгодно выбрать первую стратегию. Отметим звездочкой лучший результат в первом столбце для игрока *A*. Если игрок *B* выберет вторую стратегию, то игроку *A* также выгодно выбрать вторую стратегию. Отметим звездочкой лучший результат во втором столбце для игрока *A*. Аналогично для игрока *B* выбираем лучший результат для каждой строки платежной матрицы игрока *B*:

	Высокое качество	Низкое качество
Высокое качество	(12*, 12*)	(15, 10)
Низкое качество	(10, 15)	(16*, 16*)

У игрока *A* нет доминирующей стратегии (12 > 10, 15 < 16). У игрока *B* тоже нет доминирующей стратегии (12 > 10, 15 < 16). Получаем две точки неоднозначного равновесия по Нэшу, отмеченные в таблице двойным знаком **. Попав в одну из них, ни одному из игроков не выгодно выходить из нее. Равновесие по Парето совпадает с равновесием по Нэшу.

6.4.4. Равновесие по Нэшу в смешанных стратегиях

Рассмотрим равновесие по Нэшу в смешанных стратегиях биматричной игры, платежная матрица которой имеет вид:

(20, 5)	(30, 10)
(10, 20)	(40, 15)

Если игрок *B* выберет первую стратегию, то игроку *A* также выгодно выбрать первую стратегию. Отметим звездочкой лучший результат в первом столбце для игрока *A*. Если игрок *B* выберет вторую стратегию, то игроку *A* тоже выгодно выбрать вторую стратегию. Отметим звездочкой лучший результат во втором столбце для игрока *A*. Аналогично для игрока *B* выбираем лучший результат для каждой строки платежной матрицы игрока *B*:

(20*, 5)	(30, 10*)
(10, 25*)	(40*, 15)

В чистых стратегиях равновесия по Нэшу нет, оптимальным решением по Парето в чистых стратегиях является ситуация (A_2, B_1) и (A_2, B_2).

Пара смешанных стратегий игроков *A* и *B* $x_0 = (p_0, 1 - p_0)$, $y_0 = (q_0, 1 - q_0)$ называется *равновесием по Нэшу в смешанных стратегиях*, если ни одному из игроков невыгодно отклоняться от этих стратегий в одиночку, т. е. когда

$\bar{f}_1(x, y_0) \le \bar{f}_1(x_0, y_0)$, $\bar{f}_2(x_0, y) \le \bar{f}_2(x_0, y_0)$ для любых $x \ne x_0$, $y \ne y_0$,

где $\bar{f}_1$ – средний выигрыш игрока *A*;

$\bar{f}_2$ – средний выигрыш игрока *B*.

Аналитическое решение биматричной игры 2×2 в смешанных стратегиях найдем по формулам, аналогичным для игры 2×2 с нулевой суммой:

$$p = \frac{b_{22} - b_{21}}{b_{11} + b_{22} - b_{12} - b_{21}},\ p = \frac{15 - 25}{5 + 15 - 10 - 25} = \frac{-10}{-15} = \frac{2}{3},\ 1 - p = \frac{1}{3},$$

$$q = \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}},\ q = \frac{40 - 30}{20 + 40 - 30 - 10} = \frac{10}{20} = \frac{1}{2},\ 1 - q = \frac{1}{2}.$$

Если хотя бы одна из величин *p*, 1 – *p*, *q*, 1 – *q* окажется отрицательной, то это значит, что имеется решение в чистых стратегиях.

В оптимальном решении для игрока *A* присутствуют только компоненты платежной матрицы игрока *B* и наоборот.

Средний выигрыш игрока *A* при равновесии по Нэшу составляет

$$\bar{f}_1(x_0, y_0) = (x_0)(a_{ij})(y_0)^T = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 20 & 30 \\ 10 & 40 \end{pmatrix} \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix} = \begin{pmatrix} \frac{50}{3} & \frac{100}{3} \end{pmatrix} \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix} = 25.$$

Средний выигрыш игрока *B* при равновесии по Нэшу

$$\bar{f}_2(x_0, y_0) = (x_0)(b_{ij})(y_0)^T = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 5 & 10 \\ 25 & 15 \end{pmatrix} \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix} = \begin{pmatrix} \frac{35}{3} & \frac{35}{3} \end{pmatrix} \begin{pmatrix} 0{,}5 \\ 0{,}5 \end{pmatrix} = \frac{35}{3}.$$

Средний выигрыш игрока *A* при использовании оптимальной стратегии больше наихудшего результата в чистых стратегиях (25 > 10) и меньше наилучшего результата (25 < 40). Аналогично для игрока *B*.

6.5. Игры с природой (статистические игры)

В играх с природой вторым игроком является природа, которая действует случайным образом. Мы не имеем представления о возможных вероятностях природы, а можем только перечислить ее состояния и определить ожидаемые доходы. Природа может случайным образом улучшить выигрыш первого игрока или ухудшить его, поэтому существует несколько критериев выбора стратегий. Выбор как критериев, так и параметров является субъективным. Целесообразно исследовать ситуацию с точки зрения каждого критерия. Если большинство критериев рекомендует одни и те же стратегии, то их и необходимо использовать. Если рекомендации противоречат друг другу, то целесообразно принять решение с учетом достоинств и недостатков каждой стратегии.

1. Критерий Лапласа (L-критерий) применяют, когда выбор стратегии осуществляется многократно, состояния природы неизвестны и считаются равновероятными.

Находят средние значения по каждой строке платежной матрицы, выбирают из них наибольшее, и соответствующая ему стратегия принимается в качестве наилучшей.

2. Критерий Байеса – Лапласа (BL-критерий) можно применять, когда выбор стратегии осуществляется многократно, состояния природы известны и не зависят от времени. При малом количестве реализаций выбора допускается некоторый риск.

Находят математическое ожидание по каждой строке платежной матрицы, выбирают наименьшее значение, и соответствующая ему стратегия принимается в качестве наилучшей.

3. Критерий Вальда (MM-критерий) применяют, когда выбор стратегии осуществляется однократно, вероятности состояний природы неизвестны.

Находят наименьшие значения по каждой строке платежной матрицы, выбирают наибольшее из них, и соответствующая ему стратегия принимается в качестве наилучшей. Этот критерий, совпадающий с нижней ценой игры, является пессимистическим. Его применяют при принятии разовых ответственных решений. Считается, что природа будет действовать наиболее неблагоприятным для человека образом.

4. Критерий максимума (оптимизма) применяют, когда выбор стратегии осуществляется один раз, вероятности состояний природы неизвестны.

Находят наибольшие значения по каждой строке платежной матрицы, выбирается наибольшее, и соответствующая ему стратегия принимается в качестве наилучшей. Этот критерий применяют при авантюристических, очень рискованных решениях. Считается, что природа будет действовать наиболее благоприятным для человека образом.

5. Критерий Сэвиджа (S-критерий) применяют, когда выбор стратегии осуществляется однократно, вероятности состояний природы неизвестны.

Выбор оптимальной стратегии определяется по формуле

$$S_S = \min_i(\max_j(\max_i a_{ij} - a_{ij})).$$

Находят наибольший элемент $\max_i a_{ij}$ для каждого столбца платежной матрицы. Из каждого элемента $\max_i a_{ij}$ каждого столбца вычитают соответствующие элементы столбца. Получают матрицу рисков:

$$r_{ij} = \max_i a_{ij} - a_{ij}.$$

Смысл каждого элемента матрицы рисков состоит в определении недополученной предприятием прибыли, если для каждого текущего состояния природы будет выбрана неоптимальная стратегия.

В каждой строке таблицы рисков выбирается наибольшее значение риска, и из них выбирается наименьшее значение, и соответствующая ему стратегия принимается в качестве наилучшей.

Критерий Сэвиджа является пессимистическим. Худшим считается не минимальный выигрыш, а максимальная потеря выигрыша по сравнению с тем, чего можно было бы добиться в данных условиях.

6. Критерий Гурвица (HW-критерий) можно применять, когда выбор стратегии осуществляется малое число раз, вероятности состояний природы неизвестны, допускается некоторый риск.

Сам игрок определяет вероятность своего «везения» по формуле

$$S_{HW} = \max_i(\alpha \min_j a_{ij} + (1-\alpha)\max_j a_{ij}).$$

Находят наименьший и наибольший элемент для каждой строки платежной матрицы. Выбирают численное значение параметра α, обычно при начальных исследованиях принимают параметр $\alpha = 0{,}5$. Для каждой строки вычисляют значения

$$S_i = \alpha \min_j a_{ij} + (1-\alpha) \max_j a_{ij}.$$

Из полученных S_i выбирают максимальное значение, и соответствующая ему стратегия принимается в качестве наилучшей.

При α = 1 критерий Гурвица превращается в критерий Вальда, при α = 0 – в критерий максимума. Параметр α выбирается из субъективных соображений: чем больше возможные потери, тем значение параметра α ближе к единице.

Пример 1.28. Предприятие планирует запустить новые цеха по производству новой продукции. Для выпуска новой продукции и полного оснащения цеха можно приобрести 20 станков типа *A*, либо типа *B*, либо типа *C*. Станок *A* имеет наименьшую стоимость и наименьшую производительность, станок *C* имеет наибольшую стоимость и наибольшую производительность, станок *B* занимает промежуточное положение по стоимости и производительности между станками *A* и *C*. Прибыль предприятия зависит от экономической ситуации в течение года, которая может быть благоприятной, стабильной или кризисной. Необходимо определить план закупок станков, обеспечивающий максимальную прибыль. Приведена зависимость прибыли в денежных единицах от состояния экономики в расчете на один станок:

Тип станка	Экономическая ситуация		
	Благоприятная	Стабильная	Кризисная
A	110	100	80
B	130	110	70
C	160	90	10

Решение

1. Критерий Лапласа

Находим среднее значение для каждой строки:

$$\bar{x}_A = \frac{110+100+80}{3} = \frac{290}{3},\ \bar{x}_B = \frac{130+110+70}{3} = \frac{310}{3},\ \bar{x}_C = \frac{260}{3}.$$

Согласно критерию Лапласа следует закупать станки типа *B*.

2. Критерий Байеса – Лапласа

Допустим, имеется экономический прогноз на этот год: вероятность стабильной ситуации составит 0,6, благоприятной – 0,3, кризисной – 0,1.

Находим математические ожидания для каждой строки:

$$m_A = 110\cdot 0{,}3 + 100\cdot 0{,}6 + 80\cdot 0{,}1 = 33 + 60 + 8 = 101,\ m_B = 112,\ m_C = 103.$$

Согласно критерию Байеса – Лапласа следует закупать станки типа *B*.

3. Критерий Вальда

Находим наименьшее значение для каждой строки и из них выбираем наибольшее:

$$\max(80, 70, 10) = 80.$$

Согласно критерию Вальда следует закупать станки типа *A*.

4. Критерий максимума

Находим наибольшие значения по каждой строке и выбираем наибольшее из них: $\max(110, 130, 160) = 160$.

Согласно критерию максимума следует закупать станки типа *C*.

5. Критерий Сэвиджа

Находим наибольшее значение $\max\limits_i a_{ij}$ для каждого столбца и от каждого элемента $\max\limits_i a_{ij}$ каждого столбца вычитаем соответствующие элементы столбца. Получаем матрицу рисков:

$$\begin{pmatrix} 50 & 10 & 0 \\ 30 & 0 & 10 \\ 0 & 20 & 70 \end{pmatrix}.$$

В каждой строке таблицы рисков выбирается наибольшее значение риска, и из них выбирается наименьшее значение: $\min(50, 30, 70) = 30$.

Согласно критерию Сэвиджа следует закупать станки типа *B*.

6. Критерий Гурвица

Выбираем значение параметра α = 0,5. Для каждой строки вычислим значения $S_i = \alpha \min\limits_j a_{ij} + (1-\alpha)\max\limits_j a_{ij}$:

$$S_A = 0{,}5 \cdot 80 + 0{,}5 \cdot 110 = 95,\ S_B = 0{,}5 \cdot 70 + 0{,}5 \cdot 130 = 100,\ S_C = 85.$$

Выбираем максимальное значение:

$$\max(S_A, S_B, S_C) = \max(95, 100, 85) = 100.$$

Рассмотрим еще один параметр α = 0,6, сместимся в сторону пессимизма:

$$S_A = 0{,}6 \cdot 80 + 0{,}4 \cdot 110 = 92,\ S_B = 0{,}6 \cdot 70 + 0{,}4 \cdot 130 = 94,\ S_C = 70.$$

Выбираем максимальное значение:

$$\max(S_A, S_B, S_C) = \max(92, 94, 70) = 94.$$

Согласно критерию Гурвица при α = 0,5 и α = 0,6 следует закупать станки типа *B*.

Результаты применения критериев представим в табличной форме:

Критерий	Рекомендуемая стратегия		
	A	*B*	*C*
Лапласа		+	
Байеса – Лапласа		+	
Вальда	+		
Максимума			+
Сэвиджа		+	
Гурвица		+	

В заданных условиях непредсказуемости состояния экономики целесообразно закупать станки типа *B*.

Глава 2. Элементы математической статистики

В современном обществе статистическими методами пользуются все – от политиков, желающих предсказать исход выборов, до предпринимателей, стремящихся оптимизировать прибыль при тех или иных вложениях капитала.

Основная цель статистики – получить знания объективным способом на основе наблюдений и анализа реальности. Методы математической статистики помогают обоснованно выбрать из возможных моделей ту, которая наилучшим образом соответствует имеющимся статистическим данным, характеризующим исследуемую систему. Основные задачи математической статистики – оценка неизвестных параметров генеральной совокупности и проверка статистических гипотез. В математической статистике обосновывается, что выводы, полученные путем анализа данных выборки наблюдений, можно распространить на всю исследуемую совокупность, а методы теории вероятностей позволяют оценить надежность этих выводов. Развитию статистической науки способствовало применение экономико-математических методов и использование компьютерной техники в анализе социально-экономических явлений.

§ 1. Статистические оценки параметров распределения

1.1. Основные понятия

Генеральной совокупностью назовем всю подлежащую изучению совокупность объектов, относительно некоторого признака A, характеризующего эти объекты.

Выборочной совокупностью (*выборкой*) назовем ту часть объектов генеральной совокупности, которую отобрали для непосредственного изучения признака A.

Если удалось установить вид распределения, к которому относится признак A, то возникает задача нахождения значений параметров, характеризующих данное распределение. Например, если известно, что изучаемый признак имеет распределение Пуассона, то необходимо оценить параметр a, которым это распределение определяется.

Параметр генеральной совокупности (*параметр*) – показатель, вычисленный для всей генеральной совокупности (например, среднее квадратичное отклонение случайной величины).

Параметр выборки (*выборочный параметр, статистика*) – некоторый показатель, вычисленный на основе данных выборки.

Параметр – фиксированное, но неизвестное число. Обычно в распоряжении исследователя имеются лишь данные выборки, например значения количественного признака $x_1, x_2, ..., x_n$, полученные в результате n наблюдений. Через эти данные и выражают оцениваемый параметр. По результатам выборки, какого бы объема она ни была, нельзя определить точное значение неизвестного параметра θ, но можно найти его приближенное значение θ^*. В связи с этим возникает необходимость в оценке приближенного значения параметра. Под статистической оценкой параметров генеральной совокупности понимают методы, позволяющие делать научно обоснованные выводы о значениях числовых параметров генеральной совокупности по случайной выборке из нее.

Построение и анализ выборочного распределения является основным математическим способом исследования реальной случайной величины.

Теоретическую основу применимости выборочного метода составляет *закон больших чисел,* согласно которому при неограниченном увеличении объема выборки случайные выборочные характеристики сколь угодно приближаются (сходятся по вероятности) к определенным параметрам генеральной совокупности.

Пусть $X_1, X_2, ..., X_n$ – выборка, полученная в результате наблюдения из некоторой генеральной совокупности X. По выборкам, как отмечалось выше, можно получить только приближенные значения неизвестного параметра θ, которые служат его оценкой. Оценки, как правило, меняются от одной выборки к другой.

Оценкой θ^* (*статистической оценкой, оценочной функцией*) неизвестного параметра θ называют произвольную функцию $\theta^* = \theta^*(x_1, x_2, ..., x_n)$, зависящую от выборки $x_1, x_2, ..., x_n$. Оценка θ^*, как функция от случайной выборки, является случайной величиной. Для одного и того же параметра можно построить различные оценочные функции. Основная задача статистического оценивания неизвестных параметров θ по выборке состоит в построении такой функции от имеющихся данных статистических наблюдений, которая давала бы наиболее точные приближенные значения реальных, но неизвестных исследователю значений этих параметров. Например, оценкой математического ожидания служит функция

$$\theta^* = \frac{x_1 + x_2 + ... + x_n}{n}.$$

Пример 2.1. Пусть произведено n испытаний в схеме Бернулли с неизвестной вероятностью успеха θ. X – число успехов в одном испытании. В результате наблюдений получена выборка $x_1, x_2, ..., x_n$, где x_i – число успехов в i-м испытании. Ряд распределения наблюдаемой величины X имеет вид

x	0	1
P	$1-\theta$	θ

$k = x_1 + x_2 + ... + x_n$ – суммарное число успехов в n испытаниях Бернулли. В качестве оценки θ^* рассмотрим функцию

$$\theta^* = \frac{k}{n}.$$

Параметр выборки k распределен по биноминальному закону с параметром θ. Оценка θ^* имеет следующий ряд распределения:

θ^*	0	$1/n$	$\cdots$	m/n	$\cdots$	1
P	$(1-\theta)^n$	$n\theta(1-\theta)^{n-1}$	$\cdots$	$C_n^m \theta^m (1-\theta)^{n-m}$	$\cdots$	θ^n

Статистические оценки дают достоверные представления об оцениваемых параметрах, если последние обладают определенными свойствами. Укажем эти свойства.

1. Оценка θ^* неизвестного параметра θ называется *несмещенной*, если при любом объеме выборки n ее математическое ожидание равно оцениваемому параметру θ, т. е. $M(\theta^*) = \theta$. В противном случае оценка называется *смещенной*.

Разность $M(\theta^*) - \theta$ называют *смещением* (*систематической ошибкой* оценивания). Для несмещенных оценок систематическая ошибка равна нулю.

Чтобы не сделать систематических ошибок в сторону увеличения или уменьшения, необходимо для приближенного неизвестного параметра брать несмещенные оценки.

Несмещенная оценка не всегда дает «хорошее» приближение параметра. Получаемые по различным выборкам оценки могут быть сильно рассеяны вокруг своего среднего значения, т. е. дисперсия оценки может быть велика. В этом случае оценка, полученная по выборке, будет сильно отличаться от оцениваемого параметра.

2. Оценка θ^* неизвестного параметра θ называется *состоятельной*, если с ростом объема выборки она сходится по вероятности к оцениваемому параметру θ, т. е. для любого положительного ε выполняется соотношение $\lim\limits_{n\to\infty} P\{|\theta^* - \theta| > \varepsilon\} = 0$. Выполнение условия состоятельности гарантирует при достаточно больших n отсутствие грубых ошибок в оценке θ. Поэтому только состоятельные оценки имеют практический смысл.

Замечание 1. Несмещенная оценка будет состоятельной, если ее дисперсия стремится к нулю при $n \to \infty$.

Замечание 2. При ограниченном n оценки, обладающие одновременно свойствами состоятельности и несмещенности, могут иметь разные дисперсии.

3. Оценка θ^* неизвестного параметра θ называется *эффективной* в некотором классе точечных оценок, если ее дисперсия наименьшая среди дисперсий других оценок из этого класса. Эффективная оценка имеет наименьший разброс по сравнению с другими несмещенными и состоятельными оценками относительно истинного значения оцениваемого параметра, поэтому эффективность является решающим свойством, определяющим качество оценки, и она, вообще говоря, не предполагает обязательного наличия свойства несмещенности.

Замечание 3. Для упрощения вычислений целесообразно использовать незначительно смещенные оценки или оценки, обладающие большей дисперсией по сравнению с эффективными оценками, так как достичь выполнения перечисленных условий удается не всегда.

Оценка θ^* в примере 2.1 является несмещенной, состоятельной и эффективной оценкой неизвестной вероятности успеха θ.

Статистические оценки параметров подразделяют по способу их представления на *точечные* (числом) и *интервальные* (интервалом).

1.2. Точечные оценки

Точечной называют статистическую оценку $\theta^* = \theta^*(X_1, X_2, ..., X_n)$ параметра θ, определяемую *одним* числом. Точечные оценки зависят от объема выборки и обычно используются в выборках большого объема. Оценки параметров совокупности, полученные по разным выборкам, как правило, отличаются друг от друга. *Ошибкой выборки* (оценивания) называют абсолютную разность $|\theta^* - \theta|$.

Выборочная средняя $\bar{x}_в = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i$ есть несмещенная и состоятельная точечная оценка математического ожидания $M(X)$. Выборочная дисперсия $D_в = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x}_в)^2 \cdot n_i$ – это точечная смещенная оценка генеральной дисперсии $D(X)$. Несмещенной оценкой генеральной дисперсии $D(X)$ является «исправленная» выборочная дисперсия $S_в^2 = \frac{n}{n-1} \cdot D_в$. «Исправленная» выборочная дисперсия является состоятельной оценкой дисперсии $D(X)$. При достаточно больших значениях n выборочная и «исправленная» дисперсии различаются мало.

Если известно математическое ожидание $M(X)$ случайной величины X, то выборочная дисперсия $D_в = \frac{1}{n}\sum_{i=1}^{k}(x_i - M(X))^2 \cdot n_i$ – несмещенная, состоятельная и эффективная оценка генеральной дисперсии $D(X)$.

Относительная частота $\frac{n_i}{n}$ – несмещенная и состоятельная оценка вероятности $P(X = x_i)$.

Пример 2.2. Необходимо найти несмещенную оценку дисперсии по данным, приведенным в табл. 2.1.

Таблица 2.1

Федеральный округ	Средний размер заработной платы, x_i, тыс. руб.	Численность занятых, n_i, млн чел.
Центральный	35,67	20,33
Северо-Западный	32,29	7,15
Южный	22,38	6,41
Северо-Кавказский	19,30	3,95
Приволжский	22,41	14,74
Уральский	34,03	6,08
Сибирский	26,48	9,05
Дальневосточный	37,26	3,19

Решение

Найдем средний размер заработной платы по стране:

$$\overline{x}_в = \frac{\sum_{i=1}^{8} x_i \cdot n_i}{\sum_{i=1}^{8} n_i} = (35{,}67 \cdot 20{,}33 + 32{,}29 \cdot 7{,}15 + 22{,}38 \cdot 6{,}41 + 19{,}30 \cdot 3{,}95 +$$

$$+ 22{,}41 \cdot 14{,}74 + 34{,}03 \cdot 6{,}08 + 26{,}48 \cdot 9{,}05 + 37{,}26 \cdot 3{,}19) : (20{,}33 + 7{,}15 + 6{,}41 +$$

$$+ 3{,}95 + 14{,}74 + 6{,}08 + 9{,}05 + 3{,}19) = \frac{2071{,}4646}{70{,}9} \approx 29{,}21670804.$$

Для вычисления выборочной дисперсии воспользуемся формулой

$$D_в = \overline{x^2} - (\overline{x}_в)^2,$$

где $\overline{x^2}$ – выборочная средняя квадратов вариантов выборки.

$$\sum_{i=1}^{8} x_i^2 \cdot n_i = (35{,}67^2 \cdot 20{,}33 + 32{,}29^2 \cdot 7{,}15 + 22{,}38^2 \cdot 6{,}41 + 19{,}30^2 \cdot 3{,}95 +$$

$$+ 22{,}41^2 \cdot 14{,}74 + 34{,}03^2 \cdot 6{,}08 + 26{,}48^2 \cdot 9{,}05 + 37{,}26^2 \cdot 3{,}19) = 63221{,}545186,$$

$$\overline{x^2} = \frac{\sum_{i=1}^{8} x_i^{\,2} \cdot n_i}{\sum_{i=1}^{8} n_i} = \frac{63221{,}545186}{70{,}9} \approx 891{,}7002142,$$

$$D_в = 891{,}7002142 - (29{,}21670804)^2 \approx 38{,}08418554$$

Находим несмещенную оценку дисперсии:

$$S_в^2 = \frac{n}{n-1} \cdot D_в = \frac{70{,}9}{69{,}9} \cdot 38{,}08418554 \approx 38{,}62902367.$$

При большом объеме выборки выборочное среднее, выборочную и исправленную дисперсию удобно вычислять, используя вспомогательную табл. 2.2.

Таблица 2.2

x_i	n_i	$x_i \cdot n_i$	$x_i - \overline{x}_в$	$(x_i - \overline{x}_в)^2$	$(x_i - \overline{x}_в)^2 \cdot n_i$
x_1 ...	n_1 ...	$x_1 \cdot n_1$ ...	$x_1 - \overline{x}_в$ ...	$(x_1 - \overline{x}_в)^2$ ...	$(x_1 - \overline{x}_в)^2 \cdot n_1$ ...
x_k	n_k	$x_k \cdot n_k$	$x_k - \overline{x}_в$	$(x_k - \overline{x}_в)^2$	$(x_k - \overline{x}_в)^2 \cdot n_k$
	$n = \sum_{i=1}^{k} n_i$	$\sum_{i=1}^{k} x_i \cdot n_i$			$\sum_{i=1}^{k} (x_i - \overline{x}_в)^2 \cdot n_i$

Замечание. Для проведения статистической обработки информации используют также табличный процессор Microsoft Excel, включающий в себя программную надстройку «Пакет анализа» и библиотеку из 78 статистических функций, технология работы с которыми описывается, например, в [4]. В повседневной деятельности такого набора инструментов бывает вполне достаточно для проведения довольно полного и качественного статистического анализа информации. Если же пользователя не удовлетворяет Excel, необходимо обратиться к мощным специализированным пакетам статистического анализа, в частности к пакету STATISTICA фирмы StatSoft.

Рассмотрим методы нахождения точечных оценок.

1.2.1. Метод моментов

При заданном виде закона распределения случайной величины X неизвестные параметры этого распределения можно оценить, т. е. выразить как функцию вариант выборки, на основе *метода моментов* (К. Пирсон, 1894 г.).

Метод моментов состоит в приравнивании определенного числа выборочных моментов к соответствующим теоретическим моментам (т. е. вычисленным с использованием функции $f(x,\theta)$) исследуемой случайной величины, причем последние являются функциями от неизвестных параметров $\theta^{(1)}$, …, $\theta^{(k)}$. Число рассматриваемых моментов совпадает с числом подлежащих оценке параметров. Для оценки k параметров составляем k уравнений, решая которые относительно параметров, получаем искомые оценки. Например, для оценки двух параметров закона распределения записываем два уравнения:

$$\nu_1 = M_1,\ \mu_2 = m_2,$$

где ν_1 – начальный момент первого порядка закона распределения случайной величины X;

M_1 – эмпирический момент первого порядка;

μ_2 – центральный момент второго порядка закона распределения случайной величины X;

m_2 – центральный эмпирический момент второго порядка.

Учитывая, что $\nu_1 = M(X)$, $\mu_2 = D(X)$ [5, гл.VIII, § 10], $M_1 = \overline{x}_в$, $m_2 = D_в$ [5, гл. VII, § 2], имеем:

$$\begin{cases} M(X) = \overline{x}_в, \\ D(X) = D_в. \end{cases}$$

Пример 2.3. На предприятии изготавливаются электронные лампы. Время работы лампы является случайной величиной, для характеристики которой принят показательный закон распределения $f(x) = \lambda e^{-\lambda x}$, $(x \ge 0)$. Результаты испытания 100 электронных ламп на длительность работы представлены в табл. 2.3.

Таблица 2.3

Время работы лампы	(0; 400)	(400; 800)	(800; 1200)	(1200; 1600)	(1600; 2000)
Число ламп	31	26	19	15	9

Найти оценку параметра λ.

Решение

Показательный закон распределения зависит от одного параметра λ, поэтому для его оценки записываем одно уравнение $\nu_1 = M_1$. Учитывая, что $\nu_1 = M(X)$ [5, гл. VIII, § 10], $M_1 = \overline{x}_в$ [5, гл. VII, § 2], получаем $M(X) = \overline{x}_в$.

Используя данные интервального ряда, подсчитаем выборочную среднюю по формуле

$$\overline{x}_в = \frac{1}{\sum_{i=1}^{k} n_i} \cdot \sum_{i=1}^{k} x_i' \cdot n_i ,$$

где x_i' – середина соответствующего интервала.

$\sum_{i=1}^{5} n_i = 31 + 26 + 19 + 15 + 9 = 100$, поэтому

$$\overline{x}_в = \frac{1}{100} \cdot \sum_{i=1}^{5} x_i' \cdot n_i = \frac{200 \cdot 31 + 600 \cdot 26 + 1000 \cdot 19 + 1400 \cdot 15 + 1800 \cdot 9}{100} = 780.$$

Вычислим математическое ожидание:

$$M(X) = \int_0^{+\infty} x\, f(x) dx = \int_0^{+\infty} x\lambda e^{-\lambda x} dx = \begin{vmatrix} u = x & dv = \lambda e^{-\lambda x} dx \\ du = dx & v = -e^{-\lambda x} \end{vmatrix} = -xe^{-\lambda x}\Big|_0^{+\infty} +$$

$$+ \int_0^{+\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x}\Big|_0^{+\infty} = \frac{1}{\lambda},$$

тогда по методу моментов $\frac{1}{\lambda} = \overline{x}_в$.

Последнее равенство является приближенным, так как его правая часть является случайной величиной, и оно задает не точное значение параметра λ, а его оценку λ^*:

$$\frac{1}{\lambda^*} = \overline{x}_в .$$

Следовательно, $\frac{1}{\lambda^*} = 780$, поэтому $\lambda^* = \frac{1}{780}$.

1.2.2. Метод наибольшего правдоподобия

Данный метод, *разработанный* английским статистиком Р. Фишером в 1912 г., опирается на использование условий экстремума функций одной или нескольких случайных величин. В качестве такой функции рассматривают *функцию правдоподобия.*

В случае непрерывной случайной величины функция правдоподобия имеет вид

$$L = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \ldots \cdot f(x_k, \theta),$$

где $f(x_i, \theta)$ – функция плотности вероятности в точках выборки x_i. В качестве оценки неизвестного параметра θ, найденной по методу наибольшего правдоподобия, выбирается такая функция $\theta^* = \theta^*(x_1, x_2, \ldots, x_n)$, которая максимизирует функцию правдоподобия.

Для этого рассматриваем систему уравнений $\dfrac{\partial L}{\partial \theta_i} = 0 \quad (i = 1,\ 2,\ \dots,\ k)$ по числу оцениваемых параметров. Среди решений выбираем такое, которое обращает функцию правдоподобия в максимум.

Замечание 1. Функция $\ln L$ является возрастающей, поэтому функции $\ln L$ и L достигают максимума при одних и тех же значениях. Определять максимум функции $\ln L$ (*логарифмическая функция правдоподобия*) удобнее.

Замечание 2. Метод наибольшего правдоподобия дает состоятельные оценки, распределенные асимптотически нормально и имеющие наименьшую дисперсию по сравнению с другими асимптотически нормальными оценками. Недостаток данного метода по сравнению с методом моментов – приходится решать систему сложных уравнений.

Точечные оценки – приближенные значения неизвестных параметров даже в тех случаях, когда эти оценки являются несмещенными. Если точечные оценки определяются, когда объем выборки невелик, то расхождение между оценкой и истинным значением параметра может быть значительным, что приводит к грубым ошибкам. При оценивании неизвестного параметра θ возникают вопросы: какая ошибка нами допущена, каков должен быть объем выборки n, чтобы обеспечить заданную точность, как при заданном объеме выборки определить точность оценивания. Возникает необходимость в построении доверительных интервалов, дающих представление о точности и надежности полученной оценки.

1.3. Интервальные оценки

Статистическая оценка параметров закона распределения случайной величины X, характеризующаяся двумя числовыми значениями – концами интервала, называется *интервальной.*

Интервал, в который с заданной вероятностью (надежностью), попадает оцениваемый параметр, называется *доверительным*. На практике обычно используют два типа доверительных интервалов: *симметричные* и *односторонние*.

Вероятность γ, с которой выполняется неравенство $|\theta^* - \theta| < \varepsilon$, называется *доверительной вероятностью* (*надежностью*) оценки θ для заданного ε. Доверительная вероятность связана не с величиной параметра θ, а лишь с границами интервала, которые меняются при изменении выборки. Доверительную вероятность γ чаще всего задают заранее в зависимости от конкретных условий и накладывают требование быть близкой к единице. Общепринятые значения – 0,95; 0,99; 0,999.

Вероятность $\alpha = 1-\gamma$ называется *уровнем значимости* (*вероятностью ошибок*).

Доверительными границами (*критическими значениями*) называют границы доверительного интервала. Доверительные границы зависят от закона распределения параметра θ^*. Доверительный интервал, определяемый выборкой – случайной величиной, носит случайный характер и относительно длины, и по расположению относительно θ^*. Доверительные границы тоже случайная величина, поэтому принято говорить, что доверительный интервал покрывает параметр θ с доверительной вероятностью γ.

Методы построения доверительных интервалов различны. Впервые были разработаны американским статистиком Ю. Нейманом, использовавшим идеи Р. Фишера.

Точностью оценки ε называют число, равное половине длины доверительного интервала.

Точность оценки ε, доверительная вероятность γ и объем выборки n связаны между собой.

Алгоритм построения доверительного интервала:

1) производим выборку объема n случайной величины X из генеральной совокупности с известным распределением $f(x;\theta)$;

2) находим точечную оценку θ^* неизвестного параметра θ по данным выборки;

3) задаем доверительную вероятность γ;

4) определяем границы доверительного интервала $(\theta^* - \varepsilon;\, \theta^* + \varepsilon)$. Для этого возьмем произвольное значение θ и, используя плотность вероятности, найдем функцию распределения из условия $P\left\{|\theta-\theta^*| < \varepsilon\right\} = \int\limits_{\theta^*-\varepsilon}^{\theta^*+\varepsilon} f(x,\theta)dx = \gamma$. Границы интервала определим из решения уравнений: $P(X(\theta) < \theta^* - \varepsilon) = \dfrac{1-\gamma}{2}$, $P(X(\theta) > \theta^* + \varepsilon) = \dfrac{1-\gamma}{2}$.

Полученный интервал с доверительной вероятностью γ покрывает неизвестный параметр θ и является его интервальной оценкой.

Замечание. При малом объеме выборки построение доверительных интервалов трудоемко, так как оно сводится к перебору значений неизвестного параметра. Существуют специальные таблицы ([6], табл. 5.2), позволяющие по числу успехов и неудач определить границы доверительного интервала доверительной вероятности γ.

Наиболее часто на практике встречается нормальное распределение. В этом случае нужно оценить параметры, полностью определяющие нормальное распределение, – математическое ожидание и дисперсию. Укажем интервальные оценки для параметров нормального распределения.

Произведем выборку случайной величины X из генеральной совокупности, подчиненной закону нормального распределения $N(a,\sigma)$. Пусть оценка $\tilde{\theta}$ является несмещенной оценкой параметра θ, тогда $M(\theta^*)=\theta=a$.

1.3.1. Оценка математического ожидания при известном σ

Доверительный интервал для оценки математического ожидания случайной величины X с заданной доверительной вероятностью γ в случае нормального закона распределения имеет вид

$$\left(\bar{x}_в - z\frac{\sigma}{\sqrt{n}}\,;\ \bar{x}_в + z\frac{\sigma}{\sqrt{n}}\right), \qquad (1)$$

где $z=\Phi^{-1}\left(\dfrac{\gamma}{2}\right)$, Φ^{-1} – функция, обратная функции Лапласа $\Phi(x)$, ее значения можно находить, пользуясь таблицей значений функции $\Phi(x)$, (см. прил. 2);

σ – известное среднее квадратичное отклонение или его оценка;

$\bar{x}_в$ – выборочное среднее;

n – объем выборки.

Здесь $z\dfrac{\sigma}{\sqrt{n}}$ – точность оценки (предельная погрешность).

Этот интервал покрывает математическое ожидание с заданной доверительной вероятностью γ.

Минимальный объем выборки, который обеспечит заданную точность ε при заданной доверительной вероятности находим из неравенства $n\ge\dfrac{z^2\sigma^2}{\varepsilon^2}$.

1.3.2. Оценка математического ожидания при неизвестном σ

Доверительный интервал для оценки математического ожидания случайной величины X, распределенной по нормальному закону, с заданной доверительной вероятностью γ имеет вид

$$\left(\bar{x}_в - t_\gamma\frac{S_в}{\sqrt{n}}\,;\ \bar{x}_в + t_\gamma\frac{S_в}{\sqrt{n}}\right), \qquad (2)$$

где $S_в$ – исправленное выборочное среднее квадратичное отклонение;

$\bar{x}_в$ – выборочное среднее;

n – объем выборки;

t_γ находим по таблице значений t_γ распределения Стьюдента в зависимости от числа степеней свободы $k=n-1$ и надежности γ (см. прил. 3).

Этот интервал накрывает математическое ожидание с заданной доверительной вероятностью γ.

Замечание 1. Распределение Стьюдента, связанное с нормальным распределением, часто используется в математической статистике. Оно широко применяется при обработке статистических данных, так как зависит только от одного параметра – степени свободы k.

Замечание 2. При возрастании объема выборки n (практически при $n \geq 30$), распределение Стьюдента стремится к нормальному распределению. В этом случае можно использовать нормальное распределение. При малом объеме выборки n $(n < 30)$ замена распределения Стьюдента нормальным распределением приводит к грубым ошибкам – неоправданному сужению доверительного интервала.

Замечание 3. Если измерения физической величины производят в равных условиях и на одном приборе, то результаты измерений – независимые случайные величины, имеющие одно и то же математическое ожидание, дисперсию и нормальное распределение. Поэтому при нахождении доверительных интервалов можно использовать записанные выше интервалы.

1.3.3. Оценка среднего квадратичного отклонения

Доверительный интервал для оценки среднего квадратичного отклонения случайной величины X с заданной доверительной вероятностью γ в случае нормального закона распределения случайной величины имеет вид

$$\left(S_в(1-q);\, S_в(1+q)\right),\ q<1, \tag{3}$$

где $S_в$ – исправленное выборочное среднее квадратичное отклонение;

q – параметр, который находится по таблице (см. прил. 4) на основе известного значения объема выборки n и заданной доверительной вероятности γ.

Если $q > 1$, то интервал имеет вид

$$\left(0;\, S_в(1+q)\right). \tag{4}$$

Доверительные интервалы (3)–(4) используют для оценки точности измерений, которая характеризуется средним квадратическим отклонением случайных ошибок измерений.

1.3.4. Оценка вероятности

Пусть проведено n испытаний, в каждом из которых событие A может произойти с вероятностью p. Если при этом в m случаях событие A произошло, тогда доверительный интервал для неизвестной вероятности p имеет вид

$$\left(\tilde{p} - z\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}};\ \tilde{p} + z\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n}}\right), \tag{5}$$

где $\widetilde{p} = \dfrac{m}{n}$ – относительная частота (состоятельная и несмещенная оценка вероятности p);

$z = \Phi^{-1}(\gamma/2)$, Φ^{-1} – функция, обратная функции Лапласа $\Phi(x)$, ее значения можно находить, пользуясь таблицей значений функции $\Phi(x)$, (см. прил. 2).

Замечание. Число испытаний должно удовлетворять неравенству $n \geq 36$.

Пример 2.4. Было проверено качество 16 болтов, средняя длина которых составляет 120 мм. Ошибка при их изготовлении подчинена нормальному закону распределения со средним квадратическим отклонением $\sigma = 4$ (мм). Определите доверительный интервал с надежностью 0,95 для оценки математического ожидания.

Решение

Используя соотношение $\Phi(z) = \dfrac{\gamma}{2}$ и надежность $\gamma = 0{,}95$, находим значение функции Лапласа: $\Phi(z) = \dfrac{0{,}95}{2} = 0{,}475$. По таблице значений функции $\Phi(x)$ (см. прил. 2), определяем аргумент функции Лапласа $z = 1{,}96$. Составляем доверительный интервал (1): $\left(120 - 1{,}96 \cdot \dfrac{4}{\sqrt{16}};\ 120 + 1{,}96 \cdot \dfrac{4}{\sqrt{16}}\right)$, или $(118{,}04;\ 121{,}96)$.

Пример 2.5. Глубина озера измеряется прибором, систематическая ошибка которого равна 0, а случайные ошибки подчинены нормальному закону распределения со средним квадратическим от-клонением $\sigma = 30$ см. Сколько нужно провести независимых измерений, чтобы определить глубину озера с ошибками не более 10 см при доверительной вероятности 0, 99.

Решение

По условию задачи случайная величина, распределена по нормальному закону с $\sigma = 30$, $\varepsilon = 10$, $\gamma = 0{,}99$. По таблице значений функции $\Phi(x)$ (см. прил. 2) определяем аргумент функции Лапласа $z = 2{,}58$. Находим значение n, удовлетворяющее неравенству $n \geq \dfrac{z^2\sigma^2}{\varepsilon^2}$.

Следовательно, $n \geq \dfrac{2{,}58^2 \cdot 30^2}{10^2}$, $n \geq (2{,}58 \cdot 3)^2$ или $n \geq 60$, т. е. для выполнения условий задачи нужно провести не менее 60 независимых измерений.

Пример 2.6. Случайная величина X распределена нормально. По выборке объема $n = 50$ найдено несмещенное значение выборочного среднего квадратичного отклонения $s_в = 4$. Определите доверительный интервал для оценки среднего квадратичного отклонения случайной величины с заданной доверительной вероятностью $0,999$.

Решение

По заданным значениям доверительной вероятности $\gamma = 0,999$ и объема выборки $n = 50$ из табл. прил. 4 определяем значение $q = 0,43$. Используя (3), составляем доверительный интервал: $\left(4 \cdot (1 - 0,43); 4 \cdot (1 + 0,43)\right)$, или $(2,28;\ 5,72)$.

Пример 2.7. С целью проверки усвоения темы был проведен тест из 15 вопросов. Для испытаний отобрано 50 чел. Результаты проверки представлены в табл. 2.4.

Таблица 2.4

Число правильных ответов	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Число студентов	–	2	–	1	–	3	6	8	7	7	5	3	2	4	2

Тест считается пройденным, если число правильных ответов более 8. Оцените с доверительной вероятностью 0,9 вероятность прохождения студентом теста.

Решение

По условию задачи $n = 50$, $m = 30$, тогда $\tilde{p} = \dfrac{30}{50} = \dfrac{3}{5} = 0,6$, $1 - \tilde{p} = 1 - 0,6 = 0,4$. Используя соотношение $\Phi(z) = \dfrac{\gamma}{2}$ и доверительную вероятность $\gamma = 0,9$, находим значение функции Лапласа: $\Phi(z) = \dfrac{0,9}{2} = 0,45$. По таблице значений функции $\Phi(x)$ (см. прил. 2) определяем аргумент функции Лапласа $z = 1,65$. Составляем доверительный интервал (5):

$$\left(0,6 - 1,65 \cdot \sqrt{\frac{0,6 \cdot 0,4}{50}};\ 0,6 + 1,65 \cdot \sqrt{\frac{0,6 \cdot 0,4}{50}}\right), \text{ или } (0,49;\ 0,71).$$

§ 2. Статистическая проверка статистических гипотез

В реальности нам часто приходится принимать решения. Одним из компонентов принятия решения можно считать проверку статистических гипотез. Она дает, например, информацию об эффективности (преимуществах) той или иной технологии, новой методики обучения, значимости математической модели и неразрывно связана со статистическим оцениванием. Нахождение точечных или интервальных оценок, как правило – предварительная стадия статистических исследований. Затем обычно формулируется предположение относительно вида функции распределения (*непараметрическая статистическая гипотеза*), либо предположения относительно значений параметров функции распределения (дисперсия или математическое ожидание) известного вида (*параметрическая статистическая гипотеза*). Проблемы, возникающие при исследовании случайной величины, часто сводятся к оценке истинности одной или нескольких выдвигаемых гипотез на основе результатов анализа накопленной информации об изучаемой величине. Следовательно, дальнейшая задача экспериментатора состоит в проверке выдвинутых гипотез на основе выборки.

Любое предположение о виде и свойствах наблюдаемых в эксперименте случайных величин называется *статистической гипотезой*. Например, в схеме Бернулли одной из гипотез будет предположение «вероятность успеха равна 0,25», в случае нормального распределения – «теоретическая функция распределения нормальна с дисперсией, не превосходящей квадрата среднего значения».

В ходе решения задач часто выдвигаются следующие гипотезы:

об общем виде закона распределения исследуемой случайной величины (проверка таких гипотез дает возможность установить закон распределения с точностью до параметров, которые характеризуют неизвестный экспериментатору закон распределения);

об однородности двух или нескольких выборок (такие гипотезы позволяют сделать вывод о равенстве или различии законов распределения случайной величины, характеризующих изучаемое свойство);

о числовых значениях характеристик исследуемого явления или процесса (используя эти гипотезы, сравнивают значения числовых параметров с заданными значениями);

об общем виде зависимости, существующей между компонентами исследуемого многомерного признака (данные гипотезы позволяют определить характер зависимости между свойствами исследуемого признака);

о независимости и стационарности ряда наблюдений.

Главная задача статистических гипотез – выбор правильного решения из двух альтернативных.

Основную (выдвинутую) гипотезу называют *нулевой* гипотезой (H_0). Обычно нулевые гипотезы учитывают, что различие между сравниваемыми величинами отсутствуют, а наблюдаемые отклонения объясняются лишь случайными колебаниями выборки.

Альтернативной (H_a) называется гипотеза, конкурирующая с нулевой гипотезой в том смысле, что если нулевая гипотеза отвергается, то принимается альтернативная. Выбор альтернативной гипотезы зависит от поставленной задачи. Например, H_0: $\sigma = \sigma_0$, тогда H_a: $\sigma \neq \sigma_0$ или $\sigma < \sigma_0$, $\sigma > \sigma_0$.

Различают *простую* и *сложную* гипотезы. Простой называется гипотеза, полностью определяющая функцию распределения $F(x)$. «Вероятность успеха в схеме Бернулли равна 0,25» – простая гипотеза. Сложной называется гипотеза, не являющаяся простой. «Теоретическая функция распределения не является нормальной» – сложная гипотеза.

Непосредственно определить истинность гипотезы нельзя, поэтому проверка статистических гипотез осуществляется статистическими методами на основе данных выборки и заключается в установлении согласования данных наблюдения с выдвинутой гипотезой. Процесс обоснованного сравнения сформулированной гипотезы с полученными в результате выборки данными называют *статистической проверкой гипотез*. При проверке гипотез обязательно учитывают все условия задачи. Если данные наблюдения противоречат высказанной гипотезе, то говорят, что результат сопоставления высказанной гипотезы с выборочным данными *отрицателен*. В этом случае гипотезу следует *отклонить*. Если данные наблюдения не противоречат высказанной гипотезе, то говорят, что результат сопоставления высказанной гипотезы с выборочным данными *неотрицателен*. В этом случае гипотезу можно *принять* в качестве одного из возможных решений.

Проверка статистической гипотезы не дает логического доказательства ее верности или неверности и в чем-то похожа на обоснование математических утверждений. Чтобы опровергнуть утверждение, достаточно привести один контрпример, однако для доказательства справедливости недостаточно любого числа примеров. Примеры показывают, что на данный момент нет фактов, опровергающих утверждение. Так и в статистике говорят: нулевая гипотеза не отвергается, так как полученные данные ей не противоречат. Такими свойствами могут обладать и другие гипотезы. Кроме того, гипотеза может быть отвергнута на основании других выборочных данных или по другим причинам, поэтому принятие нулевой гипотезы нельзя расценивать как точно установленный, содержащийся в ней факт, а только как достаточно правдоподобное, не противоречащее эксперименту утверждение.

Для проверки нулевой гипотезы используют специальным образом подобранную случайную величину, точное или приближенное распределение которой известно, и называют ее *статистическим критерием (критерием)*. В зависимости от вида распределения случайную величину обозначают по-разному: в случае нормального распределения – U или Z, в случае распределения Фишера – Снедекора $-F$, в случае распределения Стьюдента – T, по закону хи-квадрат (χ^2). В общем случае ее обозначают K. Например, если проверяют гипотезу о равенстве дисперсий двух нормальных генеральных совокупностей, то в качестве критерия K принимают отношение исправленных выборочных дисперсий:

$$K = \frac{S_1^2}{S_2^2}.$$

В различных опытах дисперсии будут принимать различные, наперед неизвестные значения. Следовательно, K – случайная величина, распределенная по закону Фишера – Снедекора.

Под *статистическим критерием* также понимают правило, показывающее, когда статистическую гипотезу на основании опытных значений $x_1, x_2, \ldots, x_n$ величины ξ следует принять, а когда – отвергнуть.

Критерии, применяемые в гипотезах, не требующих знаний о виде функции распределения, называются *непараметрическими.*

Критерии, использующиеся для проверки гипотез о параметрах распределения генеральной совокупности, называются *параметрическими* (*критериями значимости*).

Критерии, относящиеся к предположениям о виде функции распределения, называются *критериями согласия.*

Проверка статистических гипотез статистическими методами заключается в исследовании генеральной совокупности по выборке n независимых случайных величин $x_1, x_2, \ldots, x_n$. Параметрические гипотезы проверяют, используя критерии значимости; для проверки непараметрических используют критерии согласия. При проверке гипотезы выборочные данные могут либо согласовываться с основной гипотезой, либо ей противоречить, а значит, всегда есть риск принять неверное решение.

Причины расхождения результатов выборки и теоретических характеристик различны. Это и неудачный способ группировки данных, и недостаточный объем выборки, и неправильный выбор распределения, и т. п. При проверке гипотезы сначала по данным выборки определяют частные значения входящих в критерий величин, а затем – сам критерий. Значения критерия, вычисленные по выборке, называются наблюдаемым значением – $K_{набл}$. Для каждого критерия (распределения) имеются соответствующие таблицы, по которым, используя данные выборки, находят $K_{крит}$. Для большинства критериев при $K_{набл} \geq K_{крит}$ нулевую гипотезу отвергают, при $K_{набл} < K_{крит}$ нет оснований, чтобы отвергнуть нулевую гипотезу (в этом случае считают, что данные наблюдений согласуются с нулевой гипотезой).

Множество значений критерия, при которых гипотеза H_0 отклоняется и принимается гипотеза H_a, называется *критической областью*. Границы критической области называют *критическими точками* ($K_{крит}$).

Критическая область, представляющая собой промежуток $(k_{кр}^{п};\ \infty)$, называется *правосторонней (*$k_{кр}^{п}$ определяется из условия $P(K > k_{кр}^{п}) = \alpha$).

Критическая область, представляющая собой промежуток $(-\infty;\ k_{кр}^{л})$, называется *левосторонней (*$k_{кр}^{л}$ определяется из условия $P(K < k_{кр}^{л}) = \alpha$).

Критическая область, представляющая собой два промежутка $(-\infty;\ k_{кр}^{л})$ и $(k_{кр}^{п};\ \infty)$, называется *двусторонней* $\left(k_{кр}^{л}\right.$ и $k_{кр}^{п}$ определяются из условий: $\left.P(K < k_{кр}^{л}) = \dfrac{\alpha}{2}\right.$ и $\left.P(K > k_{кр}^{п}) = \dfrac{\alpha}{2}\right)$.

Основной принцип проверки статистических гипотез: гипотеза, попадающая в критическую область, отвергается, а альтернативная гипотеза принимается как одна из возможных.

В реальной жизни мы обнаруживаем массу различных ошибок, не является исключением и проверка статистических гипотез, которые принимаются на основе выборочных данных. Ошибки могли возникнуть, например, в процессе сбора и обработки данных. С точки зрения статистической проверки гипотез существует только два вида ошибок, называемых *ошибкой I рода и ошибкой II рода*.

Ошибкой I рода называется ошибка отклонения верной нулевой гипотезы (H_0), которая на самом деле верна.

Ошибкой II рода называется ошибка принятия ложной гипотезы (H_0).

Вероятность $\alpha = P_{H_0}(H_a)$ совершения ошибки I рода называется *уровнем значимости* статистического критерия (*размером критерия*). Выбор величины уровня значимости α зависит от размера потерь в случае ошибочного решения. В большинстве практических задач используют стандартные значения уровня значимости: α = 0,1; 0,05; 0,025; 0,01; 0,005; 0,001. Наиболее распространенной является величина уровня значимости α = 0,05, которая означает, что высказанная гипотеза будет ошибочно отклонена в среднем в пяти случаях из ста.

Если альтернативная гипотеза H_a единственная, то можно вычислить вероятность *ошибки II рода* – $\beta = P_{H_a}(H_0)$. Вероятность несовершения ошибки II рода, т. е. вероятность отклонения неверной гипотезы, называется *мощностью статистического критерия*.

Пример 2.8. Крупная партия товара может содержать дефектные изделия. Поставщик полагает, что их доля составляет 3 %, покупатель – 20 %. Достигнута следующая договоренность: сделка состоится, если при проверке 10 случайно отобранных изделий будет обнаружено не более одного дефектного. Сформулируйте нулевую и альтернативную гипотезы с точки зрения покупателя. Определите критическую область и область принятия нулевой гипотезы. Сформулируйте, в чем состоят ошибки I и II рода. Найдите их вероятности.

Решение

С точки зрения покупателя нулевой гипотезой H_0 будем считать гипотезу о 20 % дефектных изделий, а альтернативной гипотезой H_a версию поставщика о 3 % дефектных изделий.

Поскольку отбирается 10 изделий, а затем фиксируется число дефектных, то множеством всевозможных результатов испытаний будет: $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, так как дефектных изделий может оказаться 0, 1, …, 10. По условиям сделки гипотеза покупателя отвергается, если $K \le 1$, следовательно, критическая область – $A_{крит} = \{0, 1\}$, а область принятия гипотезы – $\{2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

Ошибка *I* рода: сделка состоится, даже если в ней 20 % дефектных изделий.

Ошибка *II* рода: сделка не состоится, даже если в ней 3 % дефектных изделий.

Найдем вероятности этих ошибок. Если нулевая гипотеза верна, то вероятность того, что одно случайно выбранное изделие окажется дефектным, составляет 0,2. Ошибка I рода произойдет, если из 10 изделий дефектным будет не более чем одно (0 или 1).

В случае истинности нулевой гипотезы H_0 число дефектных изделий K является биноминальной случайной величиной $B_i(10; 0,2)$, поэтому $P(K=0) = 0,8^{10} \approx 0,107$, $P(K=1) = 10 \cdot 0,8^9 \cdot 0,2 \approx 0,268$.

Вычислим вероятность α ошибки I рода:

$$\alpha = P(K \le 1 | H_0) = P(K = 0 | H_0) + P(K = 1 | H_0) = 0,107 + 0,268 = 0,375.$$

Если верна альтернативная гипотеза H_a, то вероятность выбрать дефектное изделие составляет $0,03$. Ошибка II рода произойдет, если из 10 изделий дефектными окажутся два или более.

В случае истинности альтернативной гипотезы число дефектных изделий K также является биноминальной случайной величиной, но с другим параметром – $B_i(10; 0,03)$, поэтому $P(K=0) = 0,97^{10} \approx 0,737$, $P(K=1) = 10 \cdot 0,97^9 \cdot 0,03 \approx 0,228$. Следовательно, для вероятности β ошибки II рода имеем:

$$\beta = P(K > 1 | H_a) = 1 - P(K \le 1 | H_a) = 1 - P(K = 0 | H_a) - P(K = 1 | H_a) = 0,035.$$

Замечание. Из сравнения вероятностей α и β следует, что оговоренная процедура проверки выгодна скорее поставщику.

Задачи, сводящиеся к оценке истинности нулевой гипотезы H_0 по отношению к конкурирующей гипотезе H_a, можно решить, используя различные статистические критерии. Несмотря на разнообразие как самих гипотез, так и применяемых статистических критериев, существует некий алгоритм:

1. На основе выборочных данных $x_1, x_2, \ldots, x_n$ с учетом конкретных условий задачи выдвигают нулевую гипотезу (H_0) и конкурирующую с ней альтернативную гипотезу H_a.

2. Выбирают *уровень значимости α.*

3. Выбирают (если он не задан) объем выборки n и число степеней свободы r.

4. На основе выборочных данных выбирают критерий (статистику) – некоторую функцию K от результатов наблюдений и условий рассматриваемой статистической задачи, подчиненную некоторому закону распределения. Статистика должна быть несмещенной, состоятельной и эффективной. Критерий строится на основании *принципа отношения правдоподобия.*

5. По таблицам в зависимости от объема выборки и уровня значимости α критических точек ($K_{крит}$) определяют критическую область W. В этом случае возможно совершение ошибки I рода.

6. Определяют на основе выборочных данных $x_1, x_2, \ldots, x_n$ численную величину критерия $K_{набл}$ по формулам, учитывающим характер проверяемой гипотезы.

7. Вырабатывают решение: если $K_{набл}$ попадает в критическую область, то нулевая гипотеза отклоняется и принимается альтернативная.

Проанализировать выбранное решение можно, используя табл. 2.5.

Таблица 2.5

Выбранное решение	Реальная ситуация	
	H_0 – истинная H_a – ложная	H_0 – ложная H_a – истинная
Выбрали H_0 и отвергли H_a	$P_{H_0}(H_0) = 1 - \alpha$ Правильное решение	$P_{H_a}(H_0) = \beta$ Ошибка II рода
Выбрали H_a и отвергли H_0	$P_{H_0}(H_a) = \alpha$ Ошибка I рода	$P_{H_a}(H_a) = 1 - \beta$ Правильное решение

Для большей уверенности перед принятием гипотезы, желательно подвергнуть ее проверке с помощью других критериев или повторить эксперимент, увеличив объем выборки.

Пример 2.9. При правильной технологии хранения в течение 8 месяцев порча груш составляет 10 %. Для проверки соответствия технологии хранения установленным требованиям в хранилище была произведена случайная выборка 500 груш, 60 из которых оказались испорченными. Согласуются ли данные выборки с утверждением «доля испорченных груш в хранилище (генеральной совокупности) соответствует установленному нормативу»? Уровень значимости при проверке гипотезы принять $\alpha = 0,05$.

Решение

В данной задаче нужно проверить гипотезу о том, что доля признака p равна определенной величине δ. Нулевая гипотеза H_0: $p = \delta = 0,1$, альтернативная гипотеза $H_a: p > 0,1$. С целью проверки нулевой гипотезы используем критерий $K = \frac{m}{n}$ – выборочную долю признака в выборке. Данный критерий имеет биноминальное распределение, но при большом n и не очень маленьком δ ($n > 50$, $n\delta \geq 10$) это распределение приближается к нормальному с центром в точке δ и средним квадратическим отклонением $\sigma\left(\frac{m}{n}\right) = \sqrt{\frac{\delta(1-\delta)}{n}}$. Следовательно, величина $z = \dfrac{\frac{m}{n} - \delta}{\sqrt{\frac{\delta(1-\delta)}{n}}}$ распределена по стандартному нормальному закону. Критическую точку определяем из условия

$$P\left\{\frac{\frac{m}{n} - \delta}{\sqrt{\frac{\delta(1-\delta)}{n}}} \leq z\right\} = 0,5 + \Phi(z) = 1 - \alpha .$$

Отсюда $\Phi(z) = 1 - \alpha - 0,5 = 0,45$.

По таблице функций Лапласа (см. прил. 1) находим $z = 1,65$ и определяем критическую точку

$$K_{крит} = \delta + z\sqrt{\frac{\delta(1-\delta)}{n}} = 0,1 + 1,65 \cdot \sqrt{\frac{0,1(1-0,1)}{500}} = 0,1 + 0,022 = 0,122.$$

Выборочная доля $K_{набл} = \frac{m}{n} = \frac{60}{500} = 0,12$. Имеем: $K_{набл} < K_{крит}$. Последнее означает, что нулевая гипотеза H_0 не отвергается. Доля испорченных груш в выборочной совокупности (12 %) превысила норматив, однако такое отклонение находится в допустимых пределах и может объясняться случайностью отбора.

Установление теоретического закона распределения случайной величины, характеризующей изучаемый признак по опытному (*эмпирическому*) распределению, – основная задача математической статистики. Для ее решения необходимо установить вид и параметры закона распределения. Как бы хорошо ни подобрали теоретический закон распределения, расхождения между эмпирическим и теоретическим распределениями неизбежны. Случайно ли расхождение? Можно ли объяснить его малым числом наблюдений, способом группировки или другими причинами? Возможно, расхождение не случайно и объясняется ложной гипотезой о виде распределения. Ответить на эти вопросы и провести проверку согласия эмпирической функции распределения с предположением относительно теоретической функции распределения $F(x)$ позволяют *критерии согласия,* в которых гипотеза либо полностью, либо с точностью до небольшого числа параметров определяет закон распределения.

Замечание. Критерии согласия не доказывают справедливость гипотезы, а лишь устанавливают на принятом уровне ее согласие или несогласие с данными наблюдений.

Рассмотрим наиболее простые и употребляемые критерии согласия.

Критерий χ^2 (χ*-квадрат или критерий Пирсона*)

Критерий Пирсона наиболее часто используется при решении экономических задач средствами математической статистики. Этот критерий используют для проверки гипотезы «эмпирическая частота m_i мало отличается от соответствующей теоретической частоты nP_i».

Модель. Пусть выдвинута простая непараметрическая нулевая гипотеза H_0 о том, что генеральная совокупность X имеет функцию распределения $F(x)$. Из генеральной совокупности извлечена выборка объема n $(n \geq 50)$. Диапазон значений вариантов разобьем на k частичных интервалов равной длины. Получим интервалы наблюдаемых значений случайной величины X: $(x_0;\ x_1]$, $(x_1;\ x_2]$, $(x_2;\ x_3]\ldots(x_{k-1};\ x_k]$. Пусть m_i – число выборочных значений, попавших в интервал $(x_{i-1};\ x_i]$, тогда $\sum_{i=1}^{k} m_i = n$. Числа $p_i^* = \dfrac{m_i}{n}$ – относительные частоты попадания в i-й интервал (статистическая вероятность попадания в интервал).

Затем воспользуемся алгоритмом применения критерия.

1. Используя теоретическую функцию распределения $F(x)$, вычислим вероятности P_i попадания случайной величины X в частичный интервал $(x_{i-1};\ x_i]$:

$$P_i = P\left(x_{i-1} < x < x_i\right) = F\left(x_i\right) - F\left(x_{i-1}\right) = \int\limits_{x_{i-1}}^{x_i} f(x)\,dx.$$

Если величины $p_i^* = \frac{m_i}{n}$ и P_i (i = 1, 2, ..., k) мало различаются, то будем считать, что выдвинутая гипотеза не противоречит опытным данным. Это обстоятельство лежит в основе χ^2.

2. Умножая полученные вероятности P_i на объем выборки n, получим теоретические частоты nP_i частичных интервалов $(x_{i-1};\ x_i]$, т. е. частоты, которые следует ожидать, если нулевая гипотеза справедлива.

3. Вычисляем критерий χ^2 по формуле $\chi^2 = \sum_{i=1}^{k} \frac{(m_i - nP_i)^2}{nP_i} = \chi^2_{набл}$. Это значение будет решающим при проверке гипотезы. Критерий χ^2 показывает, насколько экспериментальные значения m_i расходятся с теоретически наиболее вероятными значениями nP_i. χ^2 принимает только неотрицательные значения, отсюда и обозначение «χ-квадрат».

4. По уровню значимости α ($\alpha = 0,05$или $\alpha = 0,1$) и числу степеней свободы r из таблицы распределения χ^2 с r степенями свободы и α (см. прил. 6) находим $\chi^2_{крит}$ такое, что выполняется равенство

$$P\left(\chi^2 \geq \chi^2_{крит}\right) = \alpha.$$

Здесь $r = k - 1 - m$, где k – число частичных интервалов; m – число неизвестных параметров предположительного распределения, оцениваемых по выборке.

При больших значениях n критерий χ^2 практически не зависит от гипотетического (предположительного) распределения *F(x)* и имеет распределение χ^2 с r степенями свободы. Например, для нормального распределения $r = k - 3$, для показательного распределения $r = k - 2$.

5. Сравниваем $\chi^2_{набл}$ с $\chi^2_{крит}$.

Если значение χ^2, вычисляемое по выборке, окажется больше или равно $\chi^2_{крит}$, то гипотеза отвергается, в противном случае гипотеза принимается (не противоречит наблюдениям). По этому правилу верная гипотеза отвергается с вероятностью α.

Замечание 1. В случае верно выбранного гипотетического распределения $F(x)$, график выбранной функции плотности должен мало отличаться от гистограммы.

Замечание 2. Асимптотическое распределение этого критерия нашел К. Пирсон. Он показал, что таковым является распределение χ^2 с r степенями свободы: $\chi^2(r)$.

С помощью критерия согласия χ^2 можно проверить гипотезу о согласии данных выборки с любым теоретическим законом распределения для любой случайной величины, как непрерывной, так и дискретной.

Пример 2.10. Предприниматель владеет тремя магазинами, расположенными недалеко друг от друга. Он решил выяснить, посещают ли покупатели все три магазина одинаково или имеется различие.

Для проверки собрали информацию о количестве покупателей в течение недели. В первом магазине это число составило 160 чел., во втором – 225, в третьем – 215.

Решение

Нулевая гипотеза H_0: вероятности посещения p_i $(i=1,\ 2,\ 3)$ покупателями магазинов равны: $p_1 = p_2 = p_3 = \frac{1}{3}$.

В результате испытания получили: $m_1 = 160$, $m_2 = 225$, $m_3 = 215$, $n = 160 + 225 + 215 = 600$, $np_i = \frac{1}{3} \cdot 600 = 200$.

Вычислим критерий $\chi^2_{набл.}{}^2 = \frac{(160-200)^2}{200} + \frac{(225-200)^2}{200}\frac{(215-200)^2}{200} =$ $= 12,25$. $\chi^2_{крит}$ найдем по таблице критических значений при $r = 3 - 1 = 2$. Даже при уровне значимости $\alpha = 0,01$ имеем $\chi^2_{крит} = 9,21$. Значит, $\chi^2_{набл} > \chi^2_{крит}$. Разницу в посещаемости магазинов в течение недели нельзя объяснить случайными колебаниями, это связано с более или менее выгодным расположением магазинов либо с разной квалификацией персонала и т. п.

Пример 2.11. В случайном порядке отобрано 60 студентов. Их оценки по высшей математике следующие: 4, 3, 5, 3, 4, 4, 3, 3, 4, 2, 3, 3, 3, 5, 3, 2, 3, 4, 5, 5, 4, 5, 4, 4, 3, 4, 2, 4, 3, 4, 4, 4, 3, 3, 4, 4, 3, 2, 4, 4, 2, 3, 3, 2, 5, 3, 2, 5, 3, 4, 3, 2, 3, 3, 5, 3, 4, 3, 3, 3. Проверьте при уровне значимости 0,05 гипотезу о нормальном законе распределения студентов по уровню знаний по высшей математике.

Решение

Нулевая гипотеза H_0: распределение студентов по уровню знаний по высшей математике (по оценкам) подчинено нормальному закону распределения. Альтернативная гипотеза H_a: распределение оценок существенно отличается от нормального распределения.

Чтобы получить первое представление о распределении студентов по уровню знаний по высшей математике (по оценкам), построим вариационный ряд. Изучаемый признак (оценка) принимает ограниченное число целых значений, поэтому удобно построить дискретный ряд и подсчитать число студентов, имеющих одинаковую оценку. Сгруппированные данные представим в виде табл. 2.6.

Таблица 2.6

Оценка (x_i)	2	3	4	5
Число студентов (m_i)	8	25	19	8

Из таблицы получим: $n=\sum_{i=1}^{4} m_i = 8+25+19+8=60$. Найдем по имеющимся данным параметры распределения: выборочную среднюю (средний балл) $\bar{x}_в$ и среднее квадратическое отклонение выборки s:

$$\bar{x}_в = \frac{2\cdot 8+3\cdot 25+4\cdot 19+5\cdot 8}{60} = \frac{207}{60} = 3,45;$$

$$\overline{x_в^2} = \frac{2^2\cdot 8+3^2\cdot 25+4^2\cdot 19+5^2\cdot 8}{60} = \frac{761}{60} \approx 12,68;$$

$$D_в = \overline{x_в^2} - (\bar{x}_в)^2 = 12,68-3,45^2 \approx 0,78,\; s=\sqrt{D_в} = \sqrt{0,78} \approx 0,88.$$

Для проверки нулевой гипотезы вначале рассчитаем теоретические частоты нормального распределения с параметрами a и σ. В качестве оценок a и σ возьмем оценки наибольшего правдоподобия, положив $a=\bar{x}_в=3,45$ и $\sigma \approx s \approx 0,88$, где $\bar{x}_в$ и s – выборочное среднее и выборочное среднее квадратичное отклонение соответственно. Расчеты запишем в виде:

№	x_i	m_i	$x_i - a$	$z_i = \frac{x_i - a}{\sigma}$	$\varphi(z_i)$	$P_i = \Delta z_i \cdot \varphi(z_i)$	P_i исправленное	$n \cdot P_i$
1	2	8	−1,45	−1,65	0,1023	0,117	0,123	7,4
2	3	25	−0,45	−0,51	0,3503	0,400	0,400	24,0
1	4	19	0,55	0,625	0,3281	0,374	0,374	22,4
4	5	8	1,55	1,76	0,0848	0,097	0,103	6,2
Σ		60				0,988	1,000	60,0

Замечание. z_i – нормированное значение отклонений x_i от математического ожидания a; $\varphi(z_i)$ – значения локальной функции Лапласа: $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ (значения $\varphi(z_i)$ взяты из прил. 1); $\Delta z_i = z_2 - z_1 = z_3 - z_2 = z_4 - z_3 = -0,51+1,65=1,14$; $n\cdot P_i$ – теоретические частоты.

Сумма вероятностей P_i: $\sum_{i=1}^{4} P_i = 0,988$ отлична от единицы. Чтобы сумма вероятностей равнялась единице, к крайним значениям z_i, т. е. к z_1 и z_4, прибавим по $\frac{1-0,988}{2} = 0,006$. В последнем столбце таблицы задействованы «исправленные» вероятности P_i.

Если нулевая гипотеза H_0 справедлива, то частоты выборки m_i должны незначительно отличаться от теоретических частот $n \cdot P_i$. Для оценки расхождения эмпирических и теоретических частот используют критерий $\chi^2 = \sum_{i=1}^{4} \frac{(m_i - nP_i)^2}{nP_i} = \chi^2_{набл}$.

Расчеты запишем в виде:

№	m_i	$n \cdot P_i$	$m_i - nP$	$(m_i - nP_i)^2$	$\frac{(m_i - nP_i)^2}{nP_i}$
1	8	7,4	0,6	0,36	0,05
2	25	24,0	1,0	1,00	0,04
1	19	22,4	–3,4	11,56	0,52
4	8	6,2	1,8	3,24	0,52
$\sum$	60	60,0			1,13

Наблюдаемое значение критерия $\chi^2_{набл} = 1,13$. Случайная величина X принимает всего четыре значения ($k = 4$), неизвестных параметров – два ($m = 2$), поэтому число степеней свободы $r = 4 - 3 = 1$.

Из табл. прил. 6 при уровне значимости $0,05$ и числе степеней свободы $r = 1$ находим: $\chi^2_{крит} = 3,84$. $\chi^2_{набл} = 1,13 < \chi^2_{крит} = 3,84$.

Вывод: $\chi^2_{набл}$ попадает в допустимую область критерия. Нет оснований для отказа от гипотезы H_0, т. е. с вероятностью $0,95$можно утверждать, что распределение студентов по уровню знаний по математике (по оценкам) согласуется с нормальным распределением.

Критерий согласия λ Колмогорова

Применяется для проверки простых гипотез о законах распределения только непрерывных случайных величин. Его отличие от критерия согласия χ^2 Пирсона состоит в том, что при применении критерия χ^2 сравнивались эмпирические и теоретические постоянные распределения, а при применении критерия λ сравниваются эмпирическая и теоретическая функции распределения $\overline{F}(x)$ и $F(x)$. Кроме того, при применении критерия Колмогорова считаются известными теоретические значения параметров предполагаемой (теоретической) функции распределения, а в критерии согласия χ^2 они могут определяться по данным выборки.

Эти ограничения несколько сужают область практического применения критерия Колмогорова. При использовании данного критерия неизвестные параметры оцениваются по данным выборки большого объема.

Модель. Пусть выдвинута простая непараметрическая нулевая гипотеза H_0 о том, что исследуемая случайная величина X имеет непрерывную функцию распределения $F(x)$. Пусть $x_1, x_2, \ldots, x_n$ – выборка объема n $(n \geq 50)$. Требуется проверить гипотезу H_0.

Алгоритм проверки критерия согласия Колмогорова

1. Располагаем результаты наблюдения в возрастающем порядке или представляем их в виде интервального статистического ряда.

2. Находим эмпирическую функцию распределения $\overline{F}(x) = \frac{n_x}{x}$.

3. Вычисляем значения теоретической функции распределения $F(x)$, соответствующие значениям случайной величины X.

4. Для каждого значения X (интервала) находим $\left|\overline{F}(x) - F(x)\right|$.

5. Вычисляем наблюдаемое значение выборочной статистики λ Колмогорова $\lambda_{набл} = \sqrt{n} \cdot \max\limits_{x}\left|\overline{F}(x) - F(x)\right|$.

6. Задаем уровень значимости и сравниваем $\lambda_{набл}$ с $\lambda_{крит}$, используя значения $\lambda_{крит}$:

Уровень значимости α	0,20	0,10	0,05	0,02	0,01	0,001
$\lambda_{крит}$	1,073	1,224	1,358	1,520	1,627	10

Заключение о правдоподобии гипотезы выносится так же, как и при использовании критерия Пирсона: если $\lambda_{набл} \geq \lambda_{крит}$, то гипотеза H_0 отвергается, в противном случае нет оснований для отклонения нулевой гипотезы и гипотеза признается правдоподобной.

Замечание. А.Н. Колмогоров нашел асимптотическое распределение рассматриваемой статистики. Из его результата следует, что решение t_q уравнения $P\left\{\left[\lambda_{набл} \geq \lambda_{крит}\right]\middle| H_0\right\} \leq \alpha$ можно найти по приближенной формуле $\lambda_{крит} = \lambda_\alpha$, где λ_α – решение уравнения $P(\lambda) = \alpha$ и функция $P(\lambda)$ определена формулой $P(\lambda) = 1 - \sum\limits_{k=-\infty}^{\infty}(-1)^k e^{-2k^2\lambda^2}$. С помощью этой формулы можно вычислять значения функции $P(\lambda)$, а также находить решения уравнения $P(\lambda) = \alpha$ относительно λ.

Если проведены две случайные выборки из одной и той же генеральной совокупности объема n_1 и n_2 $(n_{1,2} \geq 50)$, то, используя λ-критерий Колмогорова – Смирнова, можно проверить нулевую гипотезу H_0. Проверка основывается на вычислении выборочной статистики λ

$\lambda_{набл} = \sqrt{n} \cdot \max\limits_{x} \left| \overline{F}_1(x) - \overline{F}_2(x) \right|$, где $n = \dfrac{n_1 \cdot n_2}{n_1 + n_2}$, $\overline{F}_1(x)$ и $\overline{F}_2(x)$ – эмпирические функции распределения, построенные по данным первой и второй выборки. При $\lambda_{набл} \geq \lambda_{крит}$ гипотеза H_0 отвергается, при $\lambda_{набл} < \lambda_{крит}$ нет оснований для отклонения нулевой гипотезы H_0.

Критерий знаков

На практике часто приходится иметь дело со случайными величинами, закон распределения которых неизвестен. Критерий знаков – один из простейших непараметрических критериев, с помощью которого проверяется сложная непараметрическая нулевая гипотеза о том, что две выборки извлечены из одной и той же генеральной совокупности. Преимущество этого критерия – в отсутствии ограничений относительно вида функции распределения, кроме ее непрерывности.

Критерий знаков обычно применяется как критерий «спаренных» наблюдений. Обычно сравнивают результаты двух выборок одинакового объема. Исследуют знаки разностей спаренных результатов обоих выборок и находят r – число тех знаков, которых меньше. Если $r_{набл} > r_{крит}$ (см. табл. прил. 6), то считается, что нет оснований для отклонения нулевой гипотезы о том, что две выборки извлечены из одной и той же генеральной совокупности. Если $r_{набл} \leq r_{крит}$, то нулевая гипотеза отклоняется, т. е. считают, что две выборки извлечены из генеральных совокупностей с различными функциями распределения.

Пример 2.12. Пусть известны результаты двух выборок: 3, 2, 2, 3, 3, 3, 5, 3, 3, 3, 3, 3, 3, 3, 4, 3, 3, 5, 3, 3, 4, 4 и 5, 3, 3, 3, 3, 3, 2, 3, 3, 3, 4, 4, 4, 3, 4, 3, 3, 3, 4, 4, 4, 3. Проверить выполнение нулевой гипотезы о том, что две выборки извлечены из генеральных совокупностей одной функции распределения.

Решение

Сведем данные в табл. 2.7 и определим знаки разностей спаренных результатов обеих выборок.

Таблица 2.7

1-я выборка	3	2	2	3	3	3	5	3	3	3	3	3	3	3	4	3	3	5	3	3	4	4
2-я выборка	5	3	3	3	3	3	2	3	3	3	4	4	4	3	4	3	3	3	4	4	4	3
Знаки разностей	–	–	–				+				–	–	–					+	–	–		+

Из таблицы видно, что $n = 11$. Меньше знаков «+». Их число $r_{набл} = 3$. По табл. прил. 5 сравниваем $r_{крит}$ для $n = 11$и $r_{набл}$. Получаем $r_{набл} > r_{крит}$, поэтому оснований для отклонения нулевой гипотезы нет.

§ 3. Элементы теории корреляции

Одна из важных задач статистики – исследование зависимости между двумя или несколькими переменными. Две переменные могут быть либо независимыми, либо связанными функциональной или статистической зависимостью.

Функциональная зависимость между переменными x и y со множествами их значений X и Y – это правило, которое каждому элементу x из X ставит в соответствие определенный элемент y из Y.

Можно привести примеры, когда взаимосвязь двух переменных является очевидной, но тем не менее установить между ними функциональную зависимость (например, зависимость урожая от количества осадков) не представляется возможным. Такие зависимости называют статистическими (корреляционными). *Корреляционная связь* – согласованное изменение двух признаков, показывающее, что изменчивость одного признака находится в зависимости от изменчивости другого. Корреляционные связи – вероятностные изменения, поэтому их можно изучать только методами математической статистики по выборкам большого объема.

Замечание. Термин «корреляция» был введен английским естествоиспытателем Френсисом Гальтоном в 1886 г.

Статистическая (корреляционная) зависимость между случайными величинами X и Y – это правило, которое каждому значению x ставит в соответствие условный закон распределения y. Изменение x влечет в общем случае изменение закона распределения y. Корреляционная зависимость возникает тогда, когда один из признаков Y зависит не только от заданного второго X, но и от ряда случайных факторов. Корреляция бывает линейной и нелинейной, положительной и отрицательной.

Для каждого значения x_i можно найти одно вполне определенное условное среднее значение $\overline{y}(x_i)$ – среднее арифметическое наблюдавшихся значений Y, соответствующих $X = x$. Аналогично можно для каждого значения y_i найти одно вполне определенное условное среднее значение $\overline{x}(y_i)$.

При обработке экспериментальных данных наиболее часто требуется:

1) установить направление, т. е. понять, с увеличением переменной x переменная y в среднем увеличивается (положительное направление) или имеет тенденцию к уменьшению (отрицательное направление);

2) по данным $(x_1; y_1)$, $(x_2; y_2), \ldots,$ $(x_n; y_n)$, полученным в результате n независимых испытаний, определить форму корреляционной связи, т. е. вид функции $\overline{y}_x = f(x)$ или $\overline{x}_y = \varphi(y)$;

3) оценить силу корреляционной связи, т. е. дать оценку степени рассеяния значений y около условного среднего $\overline{y}_x$ или x около $\overline{x}_y$.

В общем случае $\overline{y}_x$ – некоторая функция от x. Если $\overline{y}_x$ – некоторое выборочное среднее, то уравнение $\overline{y}_x = f(x)$ называют *выборочным уравнением регрессии* y на x. Функцию $f(x)$ называют *выборочной регрессией*, а ее график – *выборочной линией регрессии* y на x. Аналогично определяется уравнение регрессии x на y.

Если график функции регрессии $\overline{y}_x = f(x)$ или $\overline{x}_y = \varphi(y)$ изображается прямой линией, то корреляцию называют *линейной,* а если кривой линией – *криволинейной*.

Например, функции регрессии y на x, используемые при количественной оценке связей между переменными, могут иметь вид: $\overline{y}_x = ax + b$ – линейная; $\overline{y}_x = ax^2 + bx + c$ – квадратичная; $\overline{y}_x = b + \frac{a}{x}$ ($\overline{y}_x = \frac{1}{ax+b}$) – гиперболическая; $\overline{y}_x = bx^a$ – степенная; $\overline{y}_x = be^{ax}$ – экспоненциальная; $\overline{y}_x = ba^x$ – показательная; $\overline{y}_x = a\ln x + b$ – полулогарифмическая и т. д.

Для определения вида функции регрессии строят точки $(x; y_x)$ (*корреляционное поле*) и по их расположению делают заключение о примерном виде функции регрессии. Не существует общего правила для выбора подходящего вида функции. Можно лишь догадываться о виде уравнения регрессии. Однако существуют способы, с помощью которых можно проверить, является ли догадка удачной.

Экономико-математические модели прогноза строятся в виде уравнений регрессии, в которых в качестве зависимой переменной величины (функции) выступает экономический показатель, в качестве независимых переменных – формирующие его факторы. Процесс нахождения теоретической линии регрессии заключается в выборе и обосновании типа кривой и расчете параметров ее уравнения. Регрессионный анализ является методом статистической обработки наблюдений, в результате которой появляется возможность составить уравнение регрессии и получить количественную оценку влияния факторных признаков на результативный признак.

Классический подход к оцениванию параметров линейной регрессии основан на методе наименьших квадратов (МНК), который служит для оценки неизвестных величин по результатам измерений, содержащим случайные погрешности. МНК позволяет получить такие оценки параметров a и b, при которых сумма квадратов отклонений фактических значений результативного признака y от расчетных (теоретических) $\overline{y}_x$ минимальна:

$$S = \sum_i \left(y_i - \overline{y}_{x_i}\right)^2 \to \min.$$

Иными словами, из всего множества линий линия регрессии по графику выбирается так, чтобы сумма квадратов расстояний по вертикали между точками и этой линией была минимальной:

$$S = \sum_i \left(y_i - \overline{y}_{x_i}\right)^2 = \sum_i \left(y_i - ax_i - b\right)^2 = \sum_i \varepsilon_i^2 \to \min.$$

Среди многих приложений МНК наиболее важным является нахождение наилучшего уравнения (функциональной зависимости) определенного вида для представления опытных данных. Процесс выражения опытных данных функциональной зависимостью с помощью МНК состоит из двух этапов: сначала выбирают вид искомой формулы, а затем для данной формулы подбирают параметры.

Довольно часто в качестве эмпирической формулы (полученной на основании опытных данных) можно принять линейную зависимость. Если выборочное уравнение регрессии y на x имеет линейную зависимость $\overline{y}_x = ax + b$, то параметры a и b находят из системы линейных уравнений:

$$\begin{cases} \left(\sum\limits_{i=1}^{n} x_i^2\right) \cdot a + \left(\sum\limits_{i=1}^{n} x_i\right) \cdot b = \left(\sum\limits_{i=1}^{n} x_i\, y_i\right), \\ \left(\sum\limits_{i=1}^{n} x_i\right) \cdot a + n \cdot b = \left(\sum\limits_{i=1}^{n} y_i\right), \end{cases}$$

где n – количество измерений или данных.

Параметр a называется коэффициентом регрессии. Его величина показывает среднее изменение признака y на a единиц при изменении фактора x на одну единицу. Коэффициент b равен значению y при $x = 0$, т. е. как бы характеризует начальное состояние рассматриваемой зависимости – уровень отсчета. Например, если уравнение регрессии имеет вид $\overline{y}_x = 2x + 5$, где x – количество вносимых удобрений в центнерах на 1 ед. площади, а $\overline{y}_x$ – полученная средняя урожайность культуры в центнерах, то смысл коэффициентов $a = 2$ и $b = 5$ таков: при увеличении количества вносимых удобрений 1 ц/га урожайность (теоретически, в среднем) повышается на 2 ц/га. Если удобрений не вносить вообще, планируемая урожайность составит 5 ц/га.

Знак при коэффициенте регрессии a показывает направление связи: при $a > 0$ связь прямая, при $a < 0$ – обратная.

Уравнение регрессии может быть найдено также в виде $\overline{x}_y = cx + d$.

Возможность четкой экономической интерпретации коэффициента регрессии сделала линейное уравнение регрессии достаточно распространенным в экономических исследованиях.

Параметр b может не иметь экономического содержания. Интерпретировать можно лишь знак при параметре b. Если $b > 0$, то относительное изменение результата y происходит медленнее, чем изменение фактора x. Иными словами, вариация результата y меньше вариации фактора – коэффициент вариации по фактору x выше коэффициента вариации для результата y: $V_x > V_y$.

Если уравнение регрессии y на x имеет вид $\overline{y}_x = ax^2 + bx + c$, то параметры a, b и c находят из системы линейных уравнений:

$$\begin{cases} \left(\sum\limits_{i=1}^{n} x_i^4\right)\cdot a + \left(\sum\limits_{i=1}^{n} x_i^3\right)\cdot b + \left(\sum\limits_{i=1}^{n} x_i^2\right)\cdot c = \left(\sum\limits_{i=1}^{n} x_i^2\, y_i\right), \\ \left(\sum\limits_{i=1}^{n} x_i^3\right)\cdot a + \left(\sum\limits_{i=1}^{n} x_i^2\right)\cdot b + \left(\sum\limits_{i=1}^{n} x_i\right)\cdot c = \left(\sum\limits_{i=1}^{n} x_i\, y_i\right), \\ \left(\sum\limits_{i=1}^{n} x_i^2\right)\cdot a + \left(\sum\limits_{i=1}^{n} x_i\right)\cdot b + n\cdot c = \left(\sum\limits_{i=1}^{n} y_i\right) \end{cases}$$

и т. д.

Зная уравнение регрессии, по точкам легко построить его наглядное представление – линию регрессии. С помощью полученного уравнения регрессии можно определить выравненное значение показателя признака y_i, подставив фактические значения x_i в уравнение регрессии. Если в найденную функцию подставить задаваемые значения фактора X в прогнозируемом периоде, то получим планируемую величину показателя.

Пример 2.13. Найдите формулу вида $y = ax + b$ МНК по данным опыта:

x	1	2	3	4	6	8
y	4	3	3	1	2	0

Решение

Для определения коэффициентов a и b в линейной функции $y = ax + b$ предварительно вычислим $\sum\limits_{i=1}^{6} y_i x_i$ =4 + 6 + 9 + 4 + 12 + 0 = 35, $\sum\limits_{i=1}^{6} x_i^2$ = 1 + 4 + 9 + 16 + 36 + 64 = 130, $\sum\limits_{i=1}^{6} y_i$ = 13, $\sum\limits_{i=1}^{6} x_i$ = 24, n = 6.

Система для определения параметров примет вид:

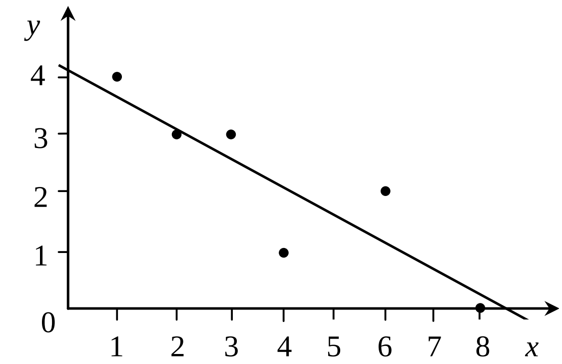
$$\begin{cases} 130a + 24b = 35, \\ 24a + 6b = 13, \end{cases}$$

откуда $a = -\frac{1}{2}$; $b = \frac{25}{6} = 4\frac{1}{6}$.

Искомая прямая есть

$$y = -\frac{1}{2}x + 4\frac{1}{6}.$$

Полученная линейная функция и экспериментальные данные показаны на рисунке.



Решение задачи нахождения линейной регрессии с помощью MS Excel

С помощью встроенной статистической функции *ЛИНЕЙН* определим параметры линейной регрессии $\overline{y}_x = ax + b$. Порядок вычисления:

1) вводим исходные данные;

2) выделяем область пустых ячеек 5×2 (5 строк, 2 столбца) с целью вывода результатов регрессионной статистики или область 1×2 – для получения только оценок коэффициентов регрессии;

3) активизируем *Мастер функций* любым из способов:

а) Вставка/Функция в главном меню;

б) кнопка *Вставка функции* на панели инструментов *Стандартная*;

4) в окне *Категория* выбираем *Статистические*, в окне *Функция – ЛИНЕЙН* и щелкаем по кнопке *OK*;

5) заполняем аргументы функции:

Известные значения y – диапазон, содержащий данные результативного признака;

Известные значения x – диапазон, содержащий данные факторов независимого признака;

Константа – логическое значение, указывающее на наличие или отсутствие свободного члена в уравнении; при *Константе* = 1 свободный член рассчитывается обычным способом; при *Константе* = 0 свободный член равен 0;

Статистика – логическое значение, указывающее на возможность вывода дополнительной информации по регрессионному анализу. При *Статистике* = 1 дополнительная информация выводится, при *Статистике* = 0 выводятся только оценки параметров уравнения;

щелкаем по кнопке *OK*. В левой верхней ячейке выделенной области появится первый элемент итоговой таблицы. Чтобы раскрыть всю таблицу, выделим таблицу 5×2 и нажмем на клавишу <F2>, а затем используем комбинацию клавиш <CTRL> +<SHIFT>+<ENTER>.

Дополнительная регрессионная статистика будет выводиться в следующем порядке:

Значение коэффициента a	Значение коэффициента b
Среднеквадратическое отклонение a	Среднеквадратическое отклонение b
Коэффициент детерминации R^2	Среднеквадратическое отклонение y
F-статистика	Число степеней свободы
Регрессионная сумма квадратов	Остаточная сумма квадратов

Параметры показательной кривой $y = ba^x$ в MS Excel вычисляются с помощью встроенной статистической функции *ЛГРФПРИБЛ*. Порядок вычисления аналогичен применению функции *ЛИНЕЙН*.

При обработке информации на компьютере вид уравнения регрессии обычно выбирают экспериментальным методом, т. е. путем сравнения величины остаточной дисперсии $D_{ост}$, рассчитанной при разных моделях. В практических исследованиях имеет место некоторое рассеяние точек относительно линии регрессии. Оно обусловлено влиянием прочих, не учитываемых в уравнении регрессии факторов. Имеют место отклонения фактических данных от теоретических $y - \overline{y}_x$. Величина этих отклонений и лежит в основе расчета остаточной дисперсии: $D_{ост} = \frac{1}{n}\sum(y - \overline{y}_x)^2$. Чем меньше величина остаточной дисперсии, тем лучше уравнение регрессии подходит к исходным данным. Если остаточная дисперсия оказывается примерно одинаковой для нескольких функций, то на практике предпочтение отдается более простым видам функций, так как они в большей степени поддаются интерпретации и требуют меньшего объема наблюдений. Результаты многих исследований подтверждают, что число наблюдений должно в 6–7 раз превышать число рассчитываемых параметров при переменной x. Это означает, что искать линейную регрессию, имея менее 6 наблюдений, бессмысленно. Для других видов функциональных зависимостей каждый параметр при x должен рассчитываться хотя бы по 6 наблюдениям. Для параболы второй степени $y = ax^2 + bx + c$ необходимо не менее 12 данных, а для параболы третьей степени $y = ax^3 + bx^2 + cx + d$ – не менее 18 наблюдений.

Уравнение регрессии всегда дополняется показателем «тесноты» связи между наблюдавшимися значениями x_i и y_i признаков X и Y.

При использовании линейной регрессии в качестве такого показателя выступает линейный коэффициент корреляции r_{xy}. Имеются разные модификации формулы линейного коэффициента корреляции, например:

$$r_{xy} = a\frac{\sigma_x}{\sigma_y} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{yx} - \overline{y} \cdot \overline{x}}{\sigma_x \sigma_y},$$

где $\overline{xy}=\frac{\sum_{i=1}^{n} x_i\ y_i}{n}$; $\sigma_x=\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}-(\overline{x})^2}$; $\sigma_y=\sqrt{\frac{\sum_{i=1}^{n} y_i^2}{n}-(\overline{y})^2}$; $\overline{y}=\frac{\sum_{i=1}^{n} y_i}{n}$; $\overline{x}=\frac{\sum_{i=1}^{n} x_i}{n}$.

Коэффициенты a, c и r_{xy} связаны соотношением: $a\cdot c=r_{xy}$.

Линейный коэффициент корреляции находится в границах $-1\le r_{xy}\le 1$. Если $r_{xy}>0$, то зависимость между X и Y такова, что возрастание одной из величин приводит к увеличению другой. При $r_{xy}<0$ увеличение одной приводит к уменьшению другой. Чем ближе значение r_{xy} к ± 1, тем сильнее (теснее) X и Y связаны линейной функциональной зависимостью. Если $r_{xy}=\pm 1$, то линии регрессии $\overline{y}_x$ и $\overline{x}_y$ сливаются в одну, что соответствует строгой линейной зависимости между X и Y, в этом случае все наблюдаемые значения располагаются на общей прямой. При $r_{xy}=-1$ имеет место отрицательная линейная зависимость, при $r_{xy}=1$ – положительная. Если $r_{xy}=0$, то признаки X и Y некоррелированы. В этом случае линии регрессии $\overline{y}_x$ и $\overline{x}_y$ параллельны координатным осям. Величина линейного коэффициента корреляции оценивает тесноту связи рассматриваемых признаков в ее линейной форме. Если допускается отклонение этой зависимости от линейной, то близость абсолютной величины линейного коэффициента корреляции к нулю еще не означает, что X и Y независимы. Следовательно, $r_{xy}=0$ не означает независимости исследуемой пары признаков. В то же время, если X и Y независимы, то $r_{xy}=0$. При $r_{xy}=0$ необходимо дополнительное статистическое исследование. Таким образом, при отклонении парной статистической зависимости от линейной коэффициент корреляции теряет свой смысл как характеристика степени тесноты связи. В этом случае следует воспользоваться другим измерителем связи – корреляционным отношением.

Допустим, что выборочный коэффициент корреляции (найденный по выборке) оказался отличным от нуля. Так как выборка отобрана случайно, то отсюда еще нельзя заключить, что коэффициент корреляции генеральной совокупности также отличен от нуля. В этом случае проверяют гипотезу о значимости (существенности) выборочного коэффициента корреляции. Другая формулировка нулевой гипотезы: коэффициент корреляции генеральной совокупности $r_{г.с}$ равен нулю. Для оценки коэффициента корреляции $r_{г.с}$ нормально распределенной генеральной совокупности можно воспользоваться формулой

$$r_{xy} - 3 \cdot \frac{1 - r_{xy}^2}{\sqrt{n}} \le r_{г.с} \le r_{xy} + 3 \cdot \frac{1 + r_{xy}^2}{\sqrt{n}} \text{ (при } n \ge 50).$$

Для оценки качества подбора уравнения регрессии рассматривается квадрат линейного коэффициента корреляции r_{xy}^2, называемый *коэффициентом детерминации*. Коэффициент детерминации характеризует долю дисперсии результативного признака Y, объясняемую регрессией, в общей дисперсии результативного признака:

$$r_{xy}^2 = \frac{\sigma_{y_{объясн}}^2}{\sigma_{y_{общ}}^2}.$$

Соответственно величина $1 - r^2$ характеризует долю дисперсии Y, вызванную влиянием остальных не учтенных в модели факторов.

Величина коэффициента детерминации является одним из критериев оценки качества линейной модели. Чем больше доля объясненной вариации, тем меньше роль прочих факторов, т. е. линейная модель хорошо аппроксимирует исходные данные, и ею можно воспользоваться для прогноза значений результативного признака. Например, если коэффициент корреляции $r_{xy} = 0,76$, то коэффициент детерминации $r_{xy}^2 = 0,76^2 \approx 0,58$. Как следует из полученного результата, 58 % рассеяния зависимой переменной Y объясняются линейной регрессией y на x, а необъясненные 42 % рассеяния могут быть вызваны либо случайными ошибками, либо тем, что линейная регрессионная модель плохо согласуется с опытными данными.

На практике при заданном уровне значимости α обычно проверяется гипотеза о том, что линейная регрессионная модель *не* согласуется с опытными данными, т. е. гипотеза, отвергающая наличие линейной связи между X и Y.

Линейный коэффициент корреляции как измеритель тесноты линейной связи признаков логически связан не только с коэффициентом регрессии a, но и с коэффициентом эластичности. При линейной связи признаков X и Y средний коэффициент эластичности в целом по совокупности определяется как

$$Э_{xy} = a \cdot \frac{\bar{x}}{\bar{y}_x} = \frac{x}{ax + b}.$$

Коэффициенты регрессии и эластичности являются показателями силы связи. Коэффициент регрессии является абсолютной мерой, так как имеет единицы измерения признаков Y и X, а коэффициент эластичности – относительным показателем силы связи, выраженной в процентах.

Степень близости полученной экономико-математической модели к фактическим данным можно определить по корреляционному отношению и ошибке уравнения регрессии (среднеквадратическому отклонению уравнения регрессии).

§ 4. Однофакторный дисперсионный анализ

Дисперсионный анализ, основанный на сравнении дисперсий, является одним из методов статистической обработки наблюдений и служит для оценки влияния на наблюдаемую величину различных *факторных* признаков, т. е. признаков, от которых зависит наблюдаемая величина. Первоначально дисперсионный анализ был предложен Р. Фишером (1925 г.) для обработки результатов агрономических опытов по выявлению условий, при которых испытываемый сорт сельскохозяйственной культуры дает максимальный урожай.

На практике дисперсионный анализ применяют, чтобы установить, оказывает ли существенное влияние некоторый качественный фактор F (*однофакторный анализ*), имеющий p уровней $F_1, F_2, \ldots, F_p$, на изучаемую величину X. Например, если требуется выяснить, какой вид удобрений наиболее эффективен для получения наибольшего урожая, то фактор F – удобрение, а его уровни – виды удобрений.

Суть дисперсионного анализа заключается в расчленении общей дисперсии признака на компоненты согласно влиянию конкретных факторов, в сравнении *факторной дисперсии*, порождаемой воздействием фактора, и *остаточной дисперсии*, обусловленной случайными причинами, проверке гипотез о значимости их влияния. Для этого при дисперсионном анализе совокупность разбивается на группы, различающиеся по уровню факторов. Если различие между факторной и остаточной дисперсиями значимо, то фактор оказывает существенное влияние на X. В этом случае средние наблюдаемых значений на каждом уровне (групповые средние) также значимо различаются, что определяется в соответствии с законом распределения вероятностей случайных ошибок распределения.

Объектом дисперсионного анализа являются *средние квадраты* (несмещенные оценки дисперсий), получающиеся делением сумм квадратов отклонений на соответствующее число степеней свободы. Число степеней свободы определяется как общее число наблюдений минус число связывающих их уравнений. Дисперсионный анализ, позволяющий лишь обнаружить наличие систематических расхождений, непригоден для их численной оценки с последующим исключением из результатов наблюдений. Эта цель может быть достигнута только при многократных измерениях.

Дисперсионный анализ применяется также для установления однородности нескольких совокупностей. В более сложных случаях исследуют воздействие нескольких факторов на нескольких постоянных или случайных уровнях и выясняют влияние отдельных уровней и их комбинаций (*многофакторный* анализ).

Пусть на количественный нормально распределенный признак X воздействует фактор F, который имеет p уровней. Пусть число наблюдений (испытаний) на каждом уровне одинаково и равно q. Пусть наблюдалось $n = p \cdot q$ значений x_{ij} признака X, где i – номер испытания ($i = 1, 2, ..., q$), j – номер уровня фактора ($j = 1, 2, ..., p$). Результаты наблюдений сводят в таблицу.

Рассматривается нулевая гипотеза H_0 о равенстве групповых средних. Другими словами, гипотеза H_0 утверждает равенство факторной и остаточной дисперсий. Альтернативная гипотеза H_a отрицает утверждение гипотезы H_0 (некоторые средние могут быть различными). По критерию Фишера для проверки нулевой гипотезы H_0 достаточно проверить нулевую гипотезу о равенстве факторной и остаточной дисперсий.

Схема метода дисперсионного анализа – это схема проверки нулевой гипотезы H_0:

1. По данным испытаний вычисляют групповые средние $\bar{x}_{гр.j}$, $j = 1, 2, ..., p$, например $\bar{x}_{гр.1} = \dfrac{x_{11} + x_{21} + ... + x_{q1}}{q}$.

2. По данным испытаний вычисляют общую среднюю:

$$\bar{x}_0 = \frac{1}{pq} \cdot \sum_{i=1}^{q} \sum_{j=1}^{p} x_{ij} =$$

$$= \frac{1}{pq} \cdot (x_{11} + x_{12} + ... + x_{1p} + x_{21} + x_{22} + ... + x_{2p} + ... + x_{q1} + x_{q2} + ... + x_{qp}).$$

3. Находят общую сумму квадратов отклонений наблюдаемых значений x_{ij} от общей средней $\bar{x}_0$:

$$S_{общ} = \sum_{i=1}^{q} \sum_{j=1}^{p} (x_{ij} - \bar{x}_0)^2 .$$

Можно заметить: $S_{общ} = \overline{\sigma^2} \cdot n$.

4. Находят факторную сумму квадратов отклонений групповых средних от общей средней. Факторная сумма характеризует рассеяние «между группами»:

$$S_{факт} = q \cdot \sum_{j=1}^{p} (\bar{x}_{гр.j} - \bar{x}_0)^2 .$$

Справедливо соотношение: $S_{факт} = n \cdot \delta^2$, где δ^2 – межгрупповая дисперсия.

5. Находят остаточную сумму квадратов отклонений наблюдаемых значений группы от своей групповой средней. Она характеризует рассеяние «внутри групп»: $S_{ост} = S_{общ} - S_{факт}$.

Легко заметить: $S_{ост} = \tilde{\sigma}^2 \cdot n$, где $\tilde{\sigma}^2$ – средняя арифметическая групповых дисперсий.

$S_{общ}$ характеризует влияние фактора и случайных причин, $S_{факт}$ – воздействие фактора F; $S_{ост}$ отражает влияние случайных причин.

6. Находят общую, факторную и остаточную дисперсии:

$$S^2_{общ} = \frac{S_{общ}}{p \cdot q - 1}, \quad S^2_{факт} = \frac{S_{факт}}{p-1}, \quad S^2_{ост} = \frac{S_{ост}}{p \cdot (q-1)},$$

где p – число уровней фактора;

q – число наблюдений на каждом уровне.

7. Находят наблюдаемое значение фактора $F_{набл}$ – отношение большей из найденных дисперсий к меньшей:

$$F_{набл} = \frac{S^2_{факт}}{S^2_{ост}}, \text{ если } S^2_{факт} > S^2_{ост} \text{ и } F_{набл} = \frac{S^2_{ост}}{S^2_{факт}}, \text{ если } S^2_{ост} > S^2_{факт}.$$

8. Число степеней свободы для $S^2_{факт}$ равно $k_1 = p - 1$, число степеней свободы для $S^2_{ост}$ – $k_2 = p \cdot q - 1$. Учитывая число степеней свободы и уровень значимости α, на котором рассматривается гипотеза H_0, по таблице критических точек распределения F Фишера – Снедекора (см. прил. 7) находят критическую точку: $F_{крит}(\alpha; k_1; k_2)$.

9. Вычисленное $F_{набл}$ сравнивается с $F_{крит}$. Если $F_{набл} > F_{крит}$, то различие средних значимое и нулевая гипотеза H_0 о равенстве групповых средних *отвергается.* Если $F_{набл} < F_{крит}$ – различие групповых средних незначимое и нет оснований для отказа от гипотезы H_0.

Замечание 1. Если $S^2_{ост} > S^2_{факт}$, то гипотеза H_0 справедлива.

Замечание 2. Критерий Фишера является равномерно наиболее мощным несмещенным критерием для проверки нулевой гипотезы H_0.

Замечание 3. Оценка доли влияния факторов и их сочетаний на изучаемый (результативный) признак производится с помощью корреляционных отношений $\eta_1^2 = \frac{S^2_{факт}}{S^2_{общ}}$ и $\eta_2^2 = \frac{S^2_{факт}}{S^2_{общ}}$. Чем ближе значение корреляционного отношения η_1^2 к единице, тем больше влияние соответствующего признака (фактора).

Замечание 4. Если наблюдаемое значение x_{ij} увеличивается (уменьшается) в одно и то же число раз k, то дисперсия увеличивается (уменьшается) в k^2 раз, но отношение дисперсий ($F_{набл}/\eta_i$) не изменится.

Замечание 5. При увеличении (уменьшении) наблюдаемых значений x_{ij} на одно и то же число c дисперсия не меняется.

Замечание 6. Параметры x_{ij} могут принимать как дискретные, так и непрерывные значения.

Пример 2.14. Указано число студентов первого курса, не посещавших лекции по высшей математике в течение февраля. Проверьте выполнение нулевой гипотезы H_0 о равенстве групповых средних:

N	Группы				
	1	2	3	4	Σ
1	1	2	4	2	9
2	2	2	4	2	10
3	3	3	8	5	20
4	1	3	8	3	15
5	3	3	5	2	13
6	2	2	7	4	15
Σ	12	15	36	18	81

Решение

По условию $p=4$, $q=6$, поэтому $pq=24$, $k_1=p-1=4-1=3$, $k_2=p\cdot q-1=24-1=23$. Используя вышеприведенные соотношения, вычислим групповые и общую средние: $\bar{x}_{гр.1}=\frac{12}{6}=2$, $\bar{x}_{гр.2}=\frac{15}{6}=2{,}5$, $\bar{x}_{гр.3}=\frac{36}{6}=6$, $\bar{x}_{гр.4}=\frac{18}{6}=3$, $\bar{x}_0=\frac{81}{24}=3{,}375$. Найдем теперь общую, факторную и остаточную суммы: $S_{общ}\approx 89{,}63$, $S_{факт}\approx 58{,}13$, $S_{ост}=S_{общ}-S_{факт}=89{,}63-58{,}13=31{,}5$. Определим факторную и остаточную дисперсии:

$$S^2_{факт}=\frac{58{,}13}{3}\approx 19{,}38,\quad S^2_{ост}=\frac{31{,}5}{20}\approx 1{,}58.$$

Для проверки нулевой гипотезы воспользуемся критерием Фишера – Снедекора.

Используя полученные величины, вычислим наблюдаемое значение критерия Фишера: $F_{набл}=\frac{19{,}38}{1{,}57}\approx 12{,}34$.

По таблице критических точек (см. прил. 7) $F_{крит}(0{,}05;\ 3;\ 23)=3{,}028$.

Ответ: $F_{набл}>F_{крит}$, поэтому различие средних значимое и нулевая гипотеза H_0 о равенстве групповых средних *отвергается.*

Глава 3. Элементы численных методов

Численные методы являются одним из мощных математических средств решения различных задач прикладного характера. При этом ни одна серьезная разработка в любой отрасли науки и производства не обходится без трудоемких математических расчетов.

Microsoft Excel – средство для работы с электронными таблицами, намного превосходящее по своим возможностям существующие редакторы таблиц. Первая версия данного продукта была разработана фирмой Microsoft в 1985 г. Microsoft Excel – простое и удобное средство, позволяющее проанализировать данные и, при необходимости, проинформировать о результате заинтересованную аудиторию. MS Excel разработан фирмой Microsoft и на сегодняшний день является самым популярным табличным редактором в мире. Помимо стандартных возможностей, он выводит на поверхность центральные функции электронных таблиц и делает их более доступными для всех пользователей. Для облегчения работы пользователя упрощены основные функции, создание формул, форматирование, печать и построение графиков.

§ 1. Аппроксимация функций

Аппроксимировать – значит приближенно заменять.

Допустим, известны значения некоторой функции в заданных точках. Требуется найти промежуточные значения данной функции. (Это так называемая задача о восстановлении функции.) Кроме того, при проведении расчетов сложные функции удобно заменять алгебраическими многочленами или другими элементарными функциями, которые достаточно просто вычисляются (задача о приближении функции).

Постановка задачи интерполяции

На интервале $[a;b]$ заданы точки x_i, $i=\overline{1,N}$, $a \le x_i \le b$, и значения неизвестной функции в этих точках f_i, $i=\overline{1,N}$. Требуется найти функцию $F(x)$, принимающую в точках x_i те же значения f_i. Точки x_i называются узлами интерполяции, а условия $F(x_i)=f_i$ – условиями интерполяции. При этом $F(x)$ ищем только на отрезке $[a;b]$. Если необходимо найти функцию вне отрезка, то это задача экстраполяции. Мы будем рассматривать только интерполяционные задачи.

Задача имеет много решений, так как через заданные точки (x_i, f_i), $i=\overline{1,N}$ можно провести бесконечно много кривых, причем каждая из них будет графиком функции, для которой выполнены все условия интерполяции. Для практики важен случай аппроксимации функции многочленами, т. е. $F(x) = a_0 + a_1x + a_2x^2 + ... + a_mx^m$.

Все методы интерполяции можно разделить на локальные и глобальные. В случае локальной интерполяции на каждом интервале $[x_{i-1};\ x_i]$ строится отдельный полином. В случае глобальной интерполяции отыскивается единый полином на всем интервале $[a; b]$. При этом искомый полином называется интерполяционным полиномом.

Простейшим и часто используемым видом локальной интерполяции является линейная (или кусочно-линейная) интерполяция. Она состоит в том, что заданные точки (x_i, f_i) $(i = \overline{1, N})$ соединяются прямолинейными отрезками и функция $F(x)$ приближается к ломаной с вершинами в данных точках.

Уравнения каждого отрезка ломаной в общем случае разные. Поскольку имеется N интервалов $(x_{i-1};\ x_i)$, то для каждого из них в качестве уравнения интерполяционного многочлена используется уравнение прямой, проходящей через две точки. В частности, для i-го интервала можно написать уравнение прямой, проходящей через точки $(x_{i-1}; f_{i-1})$ и $(x_i; f_i)$, в виде

$$\frac{y - f_{i-1}}{f_i - f_{i-1}} = \frac{x - x_{i-1}}{x_i - x_{i-1}},$$

отсюда

$$y = a_ix + b_i, \text{ при } x_{i-1} \le x \le x_i,$$

где

$$a_i = \frac{f_i - f_{i-1}}{x_i - x_{i-1}},\ b_i = f_{i-1} - a_ix_{i-1}.$$

Следовательно, при использовании линейной интерполяции сначала нужно определить интервал, в который попадает значение аргумента x, а затем подставить его в вышеуказанную формулу и найти приближенное значение функции в этой точке.

Рассмотрим теперь случай квадратичной интерполяции. В качестве интерполяционной функции на отрезке $[x_{i-1};\ x_{i+1}]$ принимается квадратный трехчлен. Такую интерполяцию называют также параболической.

Уравнение квадратного трехчлена

$$y = a_ix^2 + b_ix + c_i \text{ при } x_{i-1} \le x \le x_{i+1}$$

содержит три неизвестных коэффициента, a_i, b_i, c_i, для определения которых необходимы три уравнения. Ими служат условия прохождения параболы через три точки – $(x_{i-1};\ f_{i-1})$, $(x_i;\ f_i)$ и $(x_{i+1};\ f_{i+1})$. Эти условия можно записать в виде:

$$\begin{cases} a_i x_{i-1}^2 + b_i x_{i-1} + c_i = f_{i-1}, \\ a_i x_i^2 + b_i x_i + c_i = f_i, \\ a_i x_{i+1}^2 + b_i x_{i+1} + c_i = f_{i+1}. \end{cases}$$

При интерполировании функций мы использовали условие равенства значений интерполяционного полинома и данной функции в узлах интерполяции. Если же исходные данные получены в результате опытных измерений, то требование точного совпадения не нужно из-за погрешности данных. В таких случаях можно требовать лишь приближенного выполнения условий интерполяции $|F(x_i) - f_i| < \varepsilon$. Это условие означает, что интерполирующая функция $F(x)$ проходит не точно через заданные точки, а в некоторой их окрестности, например так, как показано на рис. 3.1.

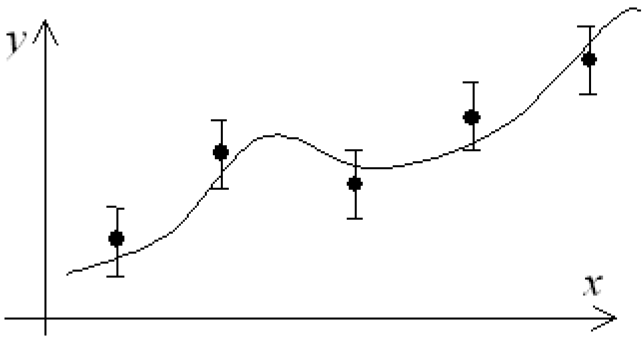


Рис. 3.1

Для практики важен случай аппроксимации функции многочленами, т. е. $F(x) = a_0 + a_1 x + a_2 x^2 + ... + a_m x^m$.

После того как выбран вид эмпирической зависимости, степень близости к эмпирическим данным определяют, используя **минимум суммы квадратов отклонений вычисленных и экспериментальных данных.**

Метод наименьших квадратов – это математическая процедура составления линейного уравнения $y = ax + b$, максимально соответствующего набору упорядоченных пар, путем нахождения значений для коэффициентов a и b. Цель МНК состоит в минимизации общей квадратичной ошибки между значениями f и y.

Коэффициенты a и b в случае линейной интерполяции определяются из системы уравнений:

$$\begin{cases} Nb + a\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} f_i, \\ b\sum_{i=1}^{N} x_i + a\sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} x_i f_i. \end{cases}$$

Линейная и квадратичная интерполяция методом наименьших квадратов в MS Excel

Пример 3.1. Найдите эмпирические формулы для функции $F(x)$, заданной таблицей, с помощью линейной и квадратичной интерполяции. Найдите значение функции при $x = 50$.

x	0	10	11	16	21	27	32	37	43	48
$F(x)$	8,4	6,2	5,6	5,1	4,2	3,4	3,1	2,5	2,1	1,9

Решение

Аналитический способ нахождения линейной интерполяции подробно описан на примере 2.13.

Построим таблицу в Excel 2007/10 с двумя столбцами: x и y, а затем построим точечный график (рис. 3.2). Используем команду *Вставка /Диаграммы*. В списке *Точечная* выбираем *Точечная с маркерами*.

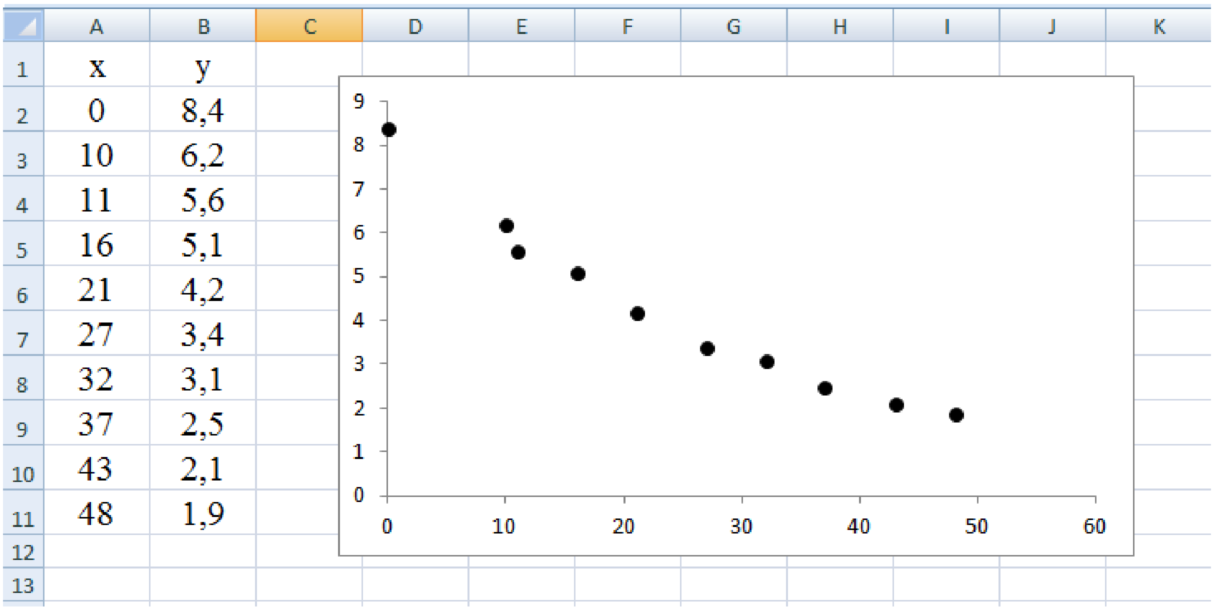


Рис. 3.2

Линейная аппроксимация

Курсором выделяем точечные маркеры и нажатием правой кнопкой мыши выводим меню, в котором выбираем команду *Добавить линию тренда…*

В диалоговом окне *Формат линии тренда* выбираем линию тренда *Линейная*, устанавливаем флажок *показывать уравнение на диаграмме* (рис. 3.3).

Услуга *линии тренда* реализует МНК для поиска коэффициентов эмпирической функции и построения ее графика.

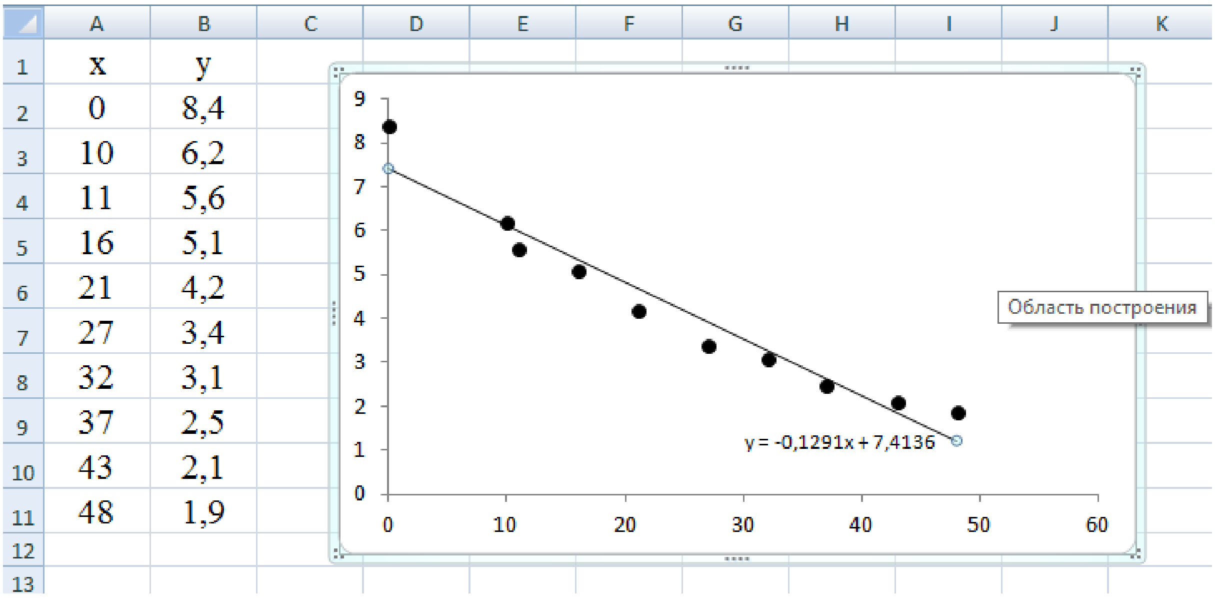


Рис. 3.3

Квадратичная аппроксимация

Курсором выделяем точечные маркеры и нажатием правой кнопкой мыши выводим меню, в котором выбираем команду *Добавить линию тренда…*

В диалоговом окне *Формат линии тренда* выбираем линию тренда *Полиномиальная*, устанавливаем флажок *показывать уравнение на диаграмме*, *OK* (рис. 3.4).

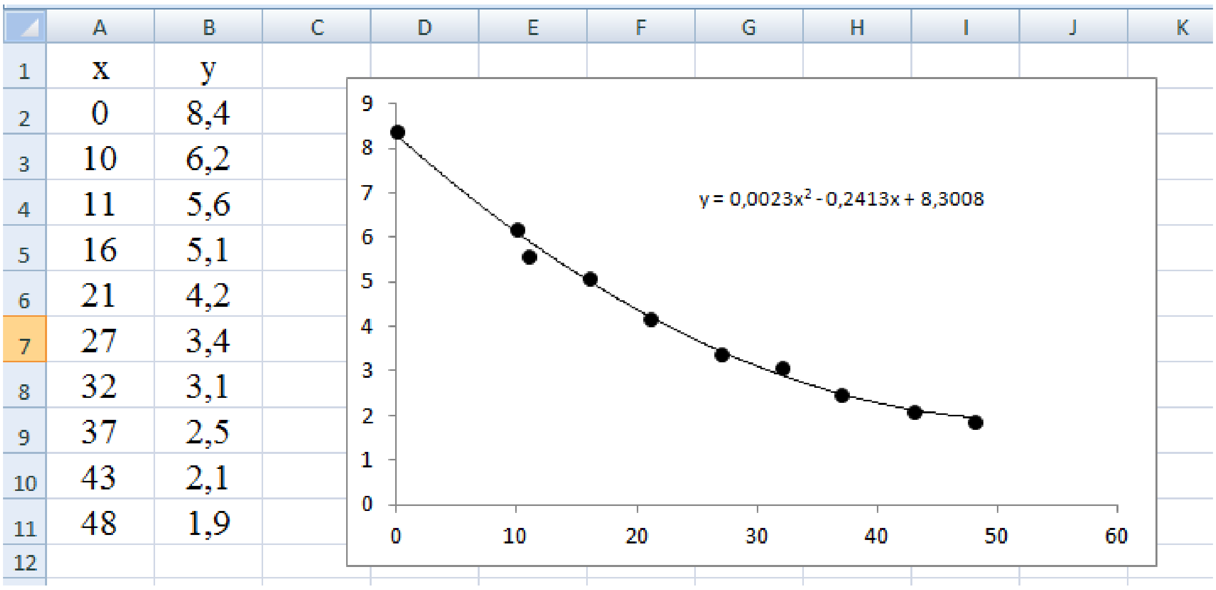


Рис. 3.4

В MS Excel имеется функция для расчета значения линейно интерполированной функции по МНК. Эта функция называется *ТЕНДЕНЦИЯ*.

Синтаксис у нее следующий:
ТЕНДЕНЦИЯ (известные значения *Y*; известные значения *X*; новые значения *X*; конст),
где известные значения *Y* – массив зависимых переменных;
известные значения *X* – массив независимых переменных;
новые значения *X* – новые значения *X*, для которого функция *ТЕНДЕНЦИЯ* возвращает ожидаемое значение зависимых переменных, *конст* – необязательный.

Для нахождения ожидаемого значения y при значении аргумента $x = 50$ используем окно диалога *Мастер функций* (f_x *Вставить функцию*), *категория Статистические*, в алфавитном списке *Выберите функцию* выбираем функцию *ТЕНДЕНЦИЯ*. В результате имеем: $y(50) = 0{,}957293...$

§ 2. Дифференцирование и интегрирование

2.1. Численное дифференцирование

Для решения многих инженерных задач часто требуется вычислить производную функции, заданной таблично.

Первый способ заключается в том, что заданный набор точек аппроксимируется стандартной функцией Excel, т. е. подбирается функция, которая лучше всего ложится на заданные точки (см. п. 1.1 гл. 3).

Второй способ – численное дифференцирование, представляющее собой совокупность методов вычисления значения производной дискретно заданной функции.

Напомним, что производной функции $f(x)$ в точке x называется предел отношения приращения Δf функции в точке x к приращению Δx аргумента, когда последнее стремится к нулю:

$$\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = f'(x) = y'.$$

Для расчета производной будем брать очень малые значения приращения аргумента, т. е. Δx.

Для того чтобы найти приближенное значение производной в нужных точках, можно заменить значение самой производной в точке x_0 на отношение $\dfrac{\Delta y}{\Delta x} = \dfrac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$.

Получим:

$$f'(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Для вычисления этой производной в каждой точке следует производить вычисления с использованием двух соседних точек: с координатой x_0 и с координатой $x_0 + \Delta x$. Вычисленная таким образом производная называется **разностной производной вправо с шагом Δx.**

Аналогично, взяв уже другие две соседние точки, $x_0 - \Delta x$ и x_0, получим формулу для вычисления **разностной производной влево с шагом $-\Delta x$**:

$$f'(x_0) \approx \frac{f(x_0) - f(x_0 - \Delta x)}{\Delta x}.$$

Из предыдущих формул получим формулу, которая позволяет вычислять **центральную разностною производную** с шагом $2\Delta x$ и чаще других используется для численного дифференцирования:

$$f'(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x}.$$

Численное дифференцирование в Excel

Рассмотрим пример с известной функцией $y = x^3$.

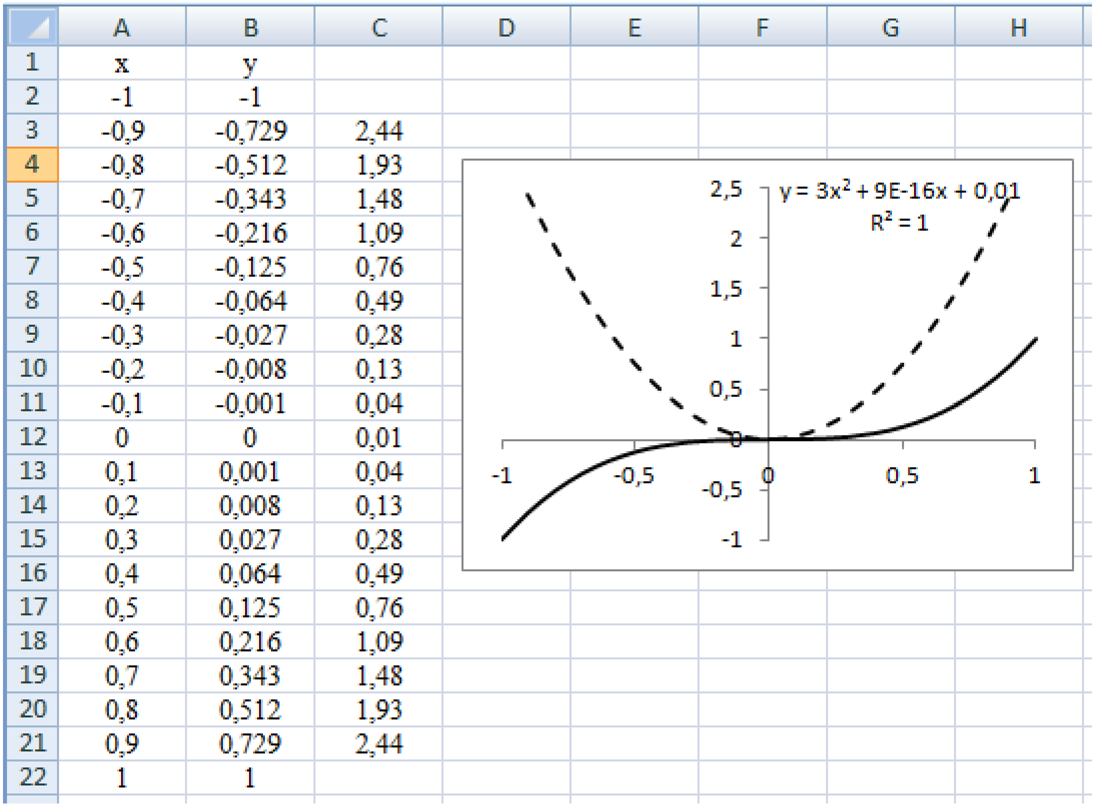
Построим таблицу в Excel с двумя столбцами, x и y, на отрезке [–1; 1] с шагом 0,1. В ячейки A2:A21 последовательно вводим числа от –1 до 1 с шагом 0,1. В ячейку B2 вводим формулу «=A1^3» и копируем ее в ячейки B3:B21 (рис. 3.5). Производная функции $y = x^3$ это $y = 3x^2$, график которой, т. е. параболу, мы и должны получить с использованием формул.

Вычислим значения центральной разностной производной в точках *x*. Для этого в ячейку C2 вводим формулу «=(B2-B4)/(A2-A4)». Затем копируем эту формулу во все нижние ячейки B2:B21.

Для построения графика функции выделяем ячейки A2:B21, используем команду *Вставка/Диаграммы*, в списке *Точечная* выбираем *Точечная с гладкими кривыми* (сплошная линия на рисунке), после чего строим график с использованием уже имеющихся значений x и полученных значений центральной разностной производной (пунктирная линия).

Для определения аналитической формы зависимости y от x в диалоговом окне *Формат линии тренда* выбираем линию тренда *Полиномиальная*, устанавливаем флажки *показывать уравнение на диаграмме* и *поместить на диаграмму величину достоверности аппроксимации* (*R*^2).

С учетом погрешности имеем: $y = 3x^2$.



	A	B	C
1	x	y	
2	-1	-1	
3	-0,9	-0,729	2,44
4	-0,8	-0,512	1,93
5	-0,7	-0,343	1,48
6	-0,6	-0,216	1,09
7	-0,5	-0,125	0,76
8	-0,4	-0,064	0,49
9	-0,3	-0,027	0,28
10	-0,2	-0,008	0,13
11	-0,1	-0,001	0,04
12	0	0	0,01
13	0,1	0,001	0,04
14	0,2	0,008	0,13
15	0,3	0,027	0,28
16	0,4	0,064	0,49
17	0,5	0,125	0,76
18	0,6	0,216	1,09
19	0,7	0,343	1,48
20	0,8	0,512	1,93
21	0,9	0,729	2,44
22	1	1	

Рис. 3.5

2.2. Численное интегрирование

Задача численного интегрирования состоит в замене исходной подынтегральной функции $f(x)$, для которой трудно или невозможно записать первообразную, некоторой аппроксимирующей функцией $\varphi(x)$. Такой функцией обычно является полином $\varphi(x)=\sum_{i=1}^{n} c_i\varphi_i(x)$, т. е.

$$I=\int_a^b f(x)dx=\int_a^b \varphi(x)dx+R,$$

где $R=\int_a^b r(x)dx$ – априорная погрешность метода на интервале интегрирования;

$r(x)$ – априорная погрешность метода на отдельном шаге интегрирования.

Методы вычисления однократных интегралов называются квадратурными (для кратных интегралов – кубатурными).

Основные методы интегрирования

1. Методы Ньютона – Котеса. Здесь $\varphi(x)$ – полином различных степеней. Сюда относятся метод прямоугольников, трапеций, Симпсона.

2. Методы статистических испытаний (методы Монте-Карло). Здесь узлы сетки для квадратурного или кубатурного интегрирования выбираются с помощью датчика случайных чисел, ответ носит вероятностный характер. В основном применяются для вычисления кратных интегралов.

Метод Симпсона

Подынтегральная функция $f(x)$ заменяется интерполяционным полиномом второй степени $P(x)$ – параболой, проходящей через три узла, например, как показано на рис. 3.6 ((1) – функция, (2) – полином).

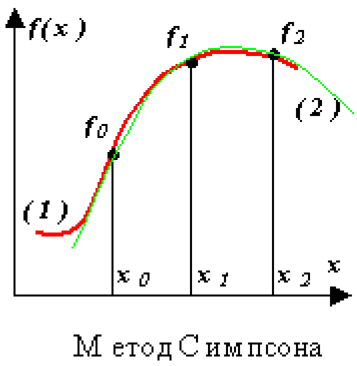


Рис. 3.6

Рассмотрим два шага интегрирования ($h = const = x_{i+1} - x_i$), т. е. три узла x_0, x_1, x_2, через которые проведем параболу, воспользовавшись уравнением Ньютона:

$$P(x) = f_0 + \frac{x - x_0}{h}(f_1 - f_0) + \frac{(x - x_0)(x - x_1)}{2h^2}(f_0 - 2f_1 + f_2).$$

Пусть $z = x - x_0$, тогда

$$P(z) = f_0 + \frac{z}{h}(f_1 - f_0) + \frac{z(z-h)}{2h^2}(f_0 - 2f_1 + f_2) =$$

$$= f_0 + \frac{z}{2h}(-3f_0 + 4f_1 - f_2) + \frac{z^2}{2h^2}(f_0 - 2f_1 + f_2).$$

Теперь, воспользовавшись полученным соотношением, сосчитаем интеграл по данному интервалу:

$$\int_{x_0}^{x_2} P(x)dx = \int_0^{2h} P(z)dz = 2hf_0 + \frac{(2h)^2}{4h}(-3f_0 + 4f_1 - f_2) + \frac{(2h)^3}{6h^2}(f_0 - 2f_1 + f_2) =$$

$$= \frac{h}{3}(6f_0 - 9f_0 + 12f_1 - 3f_2 + 4f_0 - 8f_1 + 4f_2).$$

В итоге получим:

$$\int_{x_0}^{x_2} f(x)dx == \frac{h}{3}\left(f_0 + 4f_1 + f_2\right) + r.$$

Для равномерной сетки и четного числа шагов формула Симпсона принимает вид

$$\int_a^b f(x)dx = \frac{h}{3}\left(f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + ... + 2f_{n-2} + 4f_{n-1} + f_n\right) + R.$$

Здесь $r = -\frac{h^5}{90} f^{(4)}(x_i)$, а $R = -\frac{h^4}{180}\int_a^b f^{(4)}(x)dx$ в предположении непрерывности четвертой производной подынтегральной функции.

Методы Монте-Карло

1. Одномерная случайная величина – статистический вариант метода прямоугольников. В качестве текущего узла x_i берется случайное число, равномерно распределенное на интервале интегрирования $[a,b]$. Проведя N вычислений, определим значение интеграла:

$$\int_a^b f(x)dx = \frac{b-a}{N}\sum_{i=1}^{N} f(x_i) + R.$$

Для R можно взять хотя бы $\frac{1}{\sqrt{N}}$ (рис. 3.7).

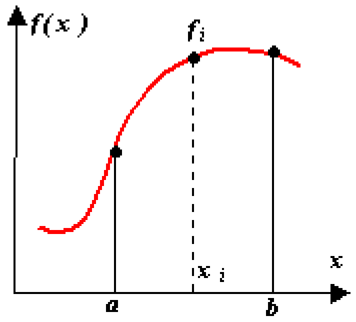


Рис. 3.7

2. Двумерная случайная величина – оценка площадей. Две равномерно распределенные случайные величины, x_i и y_i, можно рассматривать как координаты точки в двумерном пространстве. За приближенное значение интеграла принимается отношение количества точек S, попавших под кривую $y = f(x)$, к общему числу испытаний N, т. е.

$$\int_a^b f(x)dx \approx \frac{S}{N}.$$

И первый, и второй случаи легко обобщаются на кратные интегралы.

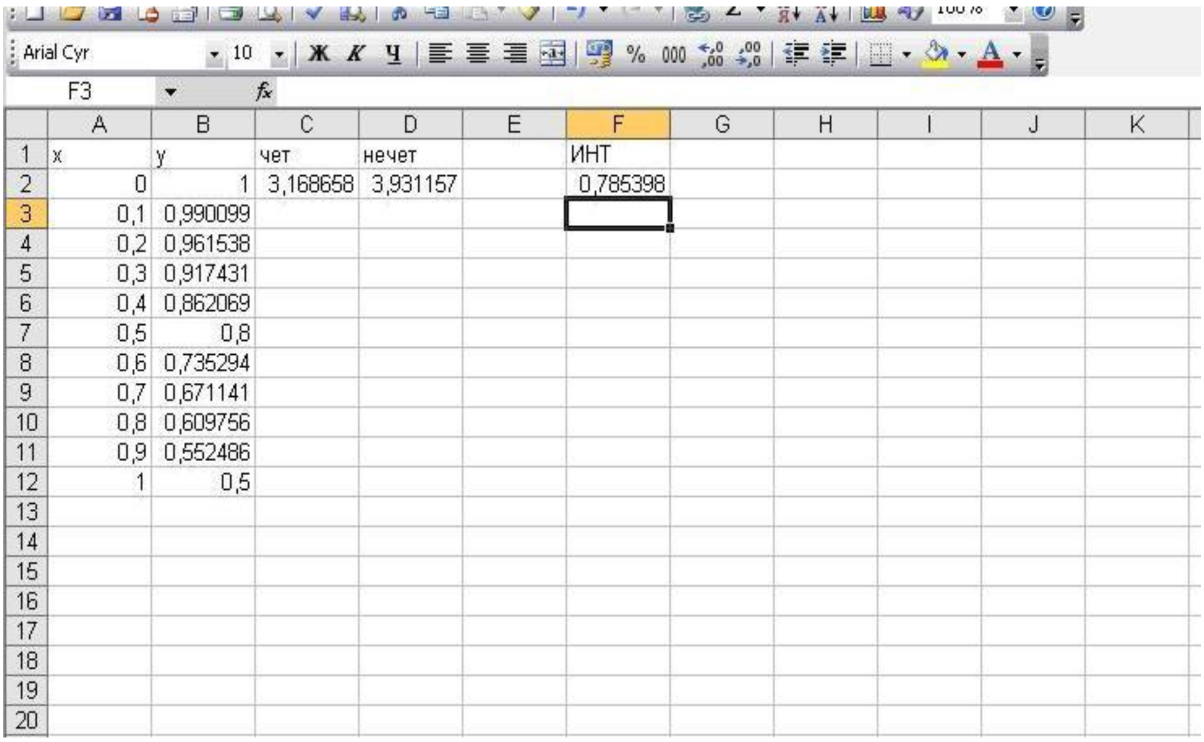
Численное интегрирование в Excel

Пример 3.2. Необходимо приближенно вычислить интеграл $\int_0^1 \frac{dx}{1+x^2} = I$ методом Симпсона.

Решение

Промежуток интегрирования разобьем на 10 частей ($n = 10$), $h = 0{,}1$. В таблицу внесем абсциссы точек деления x_i и соответствующие им ординаты $y_i = \frac{1}{1+x_i^2}$. В ячейку B2 вводим формулу «=1/(1+A2^2)» и копируем ее в ячейки B3:B12 (рис. 3.8).

В ячейке C2 суммируем значения y_i, стоящие на четных местах, кроме ячеек B2 и B12: «=B4+B6+B8+B10». В ячейке D2 суммируем значения y_i, стоящие на нечетных местах: «=B3+B5+B7+B9+B11». В ячейке F2 вводим формулу Симпсона: «=1/30*(B2+2*C2+4*D2+B12)». В ячейке F2 получим приближенное значение интеграла.



	A	B	C	D	E	F
1	x	y	чет	нечет		ИНТ
2	0	1	3,168658	3,931157		0,785398
3	0,1	0,990099				
4	0,2	0,961538				
5	0,3	0,917431				
6	0,4	0,862069				
7	0,5	0,8				
8	0,6	0,735294				
9	0,7	0,671141				
10	0,8	0,609756				
11	0,9	0,552486				
12	1	0,5				

Рис. 3.8

Точное значение интеграла

$$\int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4} \approx 0{,}7853.$$

Пример 3.3. Требуется приближенно вычислить интеграл $\int\limits_0^1 \frac{dx}{1+x^2} = I$ методом Монте-Карло.

Решение

Промежуток интегрирования разобьем на 20 частей (N = 20), $\frac{b-a}{N} = 0{,}05$. В таблицу внесем абсциссы точек деления x_i и соответствующие им ординаты $y_i = \frac{1}{1+x_i^2}$, а также вычислим в ячейке D2 сумму значений функции «=СУММ(B2:B22)» (рис. 3.9).

Используя формулу $\int\limits_a^b f(x)dx \approx \frac{b-a}{N}\sum\limits_{i=1}^{N} f(x_i)$, вычислим приближенное значение интеграла в ячейке F2: «=(A22–A2)/20*D2» (рис. 3.9).

F2 f_x =(A22-A2)/20*D2

	A	B	C	D	E	F
1	x	y				
2	0	1		16,45588		0,822794
3	0,05	0,997506				
4	0,1	0,990099				
5	0,15	0,977995				
6	0,2	0,961538				
7	0,25	0,941176				
8	0,3	0,917431				
9	0,35	0,890869				
10	0,4	0,862069				
11	0,45	0,831601				
12	0,5	0,8				
13	0,55	0,767754				
14	0,6	0,735294				
15	0,65	0,702988				
16	0,7	0,671141				
17	0,75	0,64				
18	0,8	0,609756				
19	0,85	0,580552				
20	0,9	0,552486				
21	0,95	0,525624				
22	1	0,5				

Рис. 3.9

Точное значение интеграла $\int\limits_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4} \approx 0{,}82279.$

Для получения более точного результата следует взять большее количество точек разбиения.

§ 3. Нелинейные уравнения

Как правило, нелинейное уравнения общего вида $f(x)=0$ невозможно решить аналитически. Для практических задач достаточно найти приближенное значение x, в определенном смысле близкое к точному решению уравнения.

В большинстве случаев поиск приближенного решения включает два этапа. На первом этапе отделяют корни, т. е. определяют такие отрезки, внутри которых находится строго один корень. На втором этапе уточняют корень на одном из таких отрезков, т. е. находят его значение с требуемой точностью.

Достигнутая точность может оцениваться либо «по функции» (в найденной точке x функция достаточно близка к 0), либо «по аргументу» (найден достаточно маленький отрезок $[a, b]$, внутри которого находится корень).

Отделение корней может производиться сочетанием графического и аналитического исследования функции. Такое исследование опирается на теорему Вейерштрасса. В соответствии с этой теоремой, для непрерывной на отрезке $[a,\ b]$ функции $f(x)$ и любого числа y, отвечающего условию $f(a)\leq y\leq f(b)$, на этом отрезке существует точка x, в которой функция равна y. Следовательно, для непрерывной функции достаточно найти отрезок, на концах которого функция имеет разные знаки, и тогда на этом отрезке есть корень уравнения $f(x)=0$.

Для ряда методов уточнения желательно, чтобы найденный на первом этапе отрезок содержал только один корень уравнения. Это условие выполняется, если функция на отрезке монотонна. Монотонность можно проверить либо по графику функции, либо по знаку производной.

Реализация отделения корней в MS Excel

Пример 3.4. Необходимо найти с точностью до целых все корни нелинейного уравнения $y(x)=x^3-10x+7=0$.

Решение

В Excel создается таблица, содержащая аргументы и значения функции, и по ней строится точечная диаграмма. На рис. 3.10 приведен снимок решения.

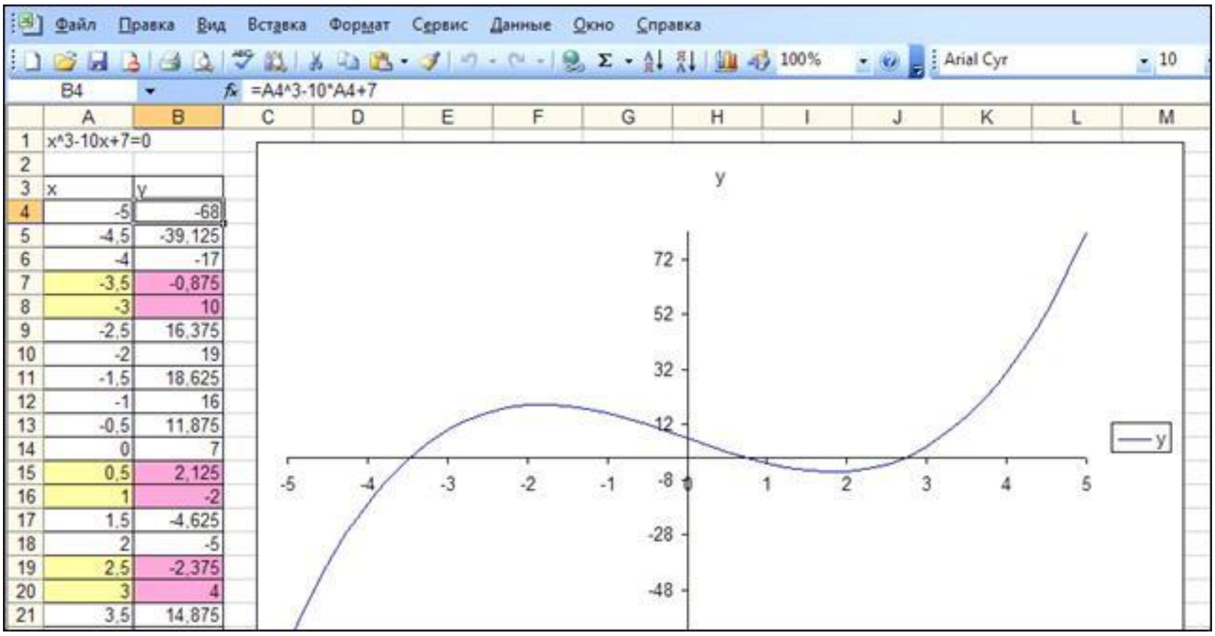


Рис. 3.10

На графике видно, что уравнение имеет три корня, принадлежащие отрезкам $[-4; -3]$, $[0;1]$ и $[2;3]$. Эти отрезки можно выявить также, наблюдая за сменой знаков функции в таблице (см. рис. 3.10). По построенному графику можно сделать вывод, что на указанных отрезках функция $f(x)$ монотонна и, следовательно, на каждом из них содержится только по одному корню.

Найти корень уравнения на выделенном отрезке можно, используя опции «Подбор параметра» и «Поиск решения».

Уточнение корней

Во всех методах уточнения корней необходимо задать начальное приближение, которое затем будет уточняться. Если уравнение имеет несколько корней, в зависимости от выбранного начального приближения будет найден один из них. При неудачно выбранном начальном приближении решение может быть не найдено. Если в результате первого этапа расчетов уже выделен отрезок, содержащий единственный корень уравнения, в качестве начального приближения можно взять любую точку этого отрезка.

Уточнение корней стандартными средствами Excel

В Excel для уточнения значений корней можно использовать опции «Подбор параметра» и «Поиск решения». Примеры оформления решения приведены на рис. 3.11 и 3.12.

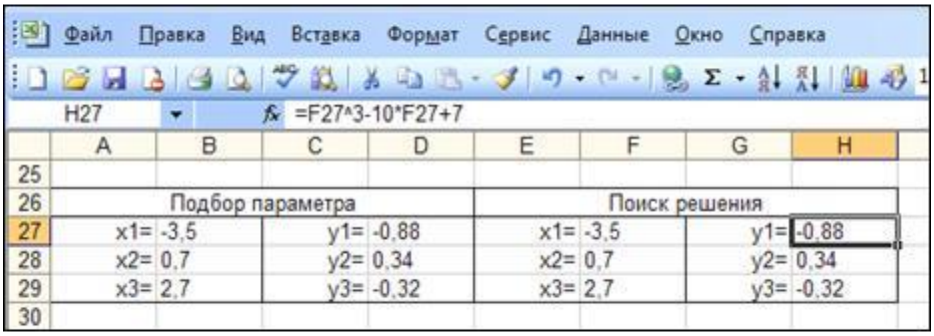


Рис. 3.11

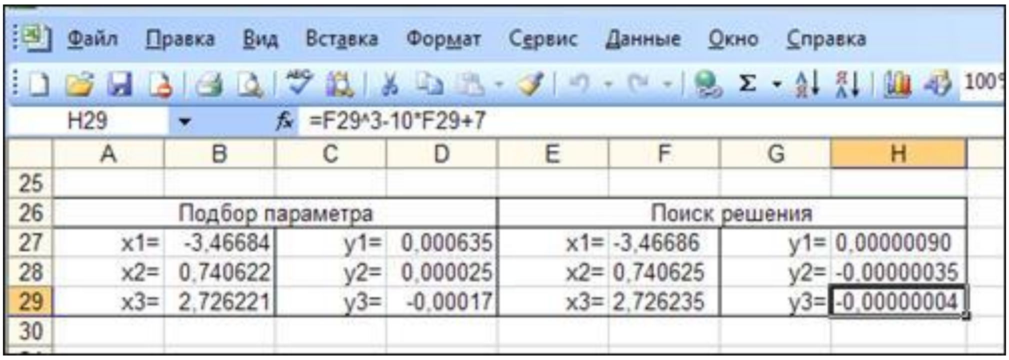


Рис. 3.12

Очевидно, что каждый стандартный инструмент находит решение уравнения с определенной точностью. Эта точность зависит от метода, используемого в пакете и, в некоторой степени, от настроек пакета. Управлять точностью результата здесь достаточно сложно, а часто и невозможно.

Ниже рассмотрены несколько наиболее распространенных методов. Отметим очевидный момент: при прочих равных условиях наиболее эффективным будет метод уточнения корней, в котором результат с той же погрешностью найден с меньшим числом вычислений функции $f(x)$ (при этом достигается и максимальная точность при одинаковом числе вычислений функции).

Метод деления отрезка пополам

На каждом шаге отрезок делят на две равные части. Затем сравнивают знаки функции на концах каждой из двух половинок (например, по знаку произведения значений функций на концах), определяют ту из них, в которой содержится решение (знаки функции на концах должны быть разные), и сужают отрезок, перенося в найденную точку его границу (a или b). Условием окончания служит малость отрезка, где содержится корень («точность по x»), либо близость к 0 значения функции в середине отрезка («точность по y»). Решением уравнения считают середину отрезка, найденного на последнем шаге.

Реализация метода деления отрезка пополам в MS Excel

Пример 3.5. Требуется построить таблицу для уточнения корня уравнения $y(x)=x^3-10x+7=0$ на отрезке $[-4;-3]$ методом деления отрезка пополам. Определим, сколько шагов надо сделать методом деления отрезка пополам и какая при этом достигается точность по x для достижения точности по y, равной 0,1; 0,01; 0,001.

Решение

Для решения можно использовать табличный процессор Excel, позволяющий автоматически продолжать строки. На первом шаге заносим в таблицу значения левого и правого концов выбранного начального отрезка и вычисляем значение середины отрезка $c=\dfrac{a+b}{2}$, а затем вводим формулу для вычисления функции в точке a $f(a)$ и копируем ее для вычисления $f(c)$ и $f(b)$. В последнем столбце вычисляем выражение $\dfrac{b-a}{2}$, характеризующего степень точности вычислений. Все набранные формулы можно скопировать во вторую строку таблицы (рис. 3.13).

На втором шаге автоматизируем процесс поиска той половины отрезка, где содержится корень, используя логическую функцию ЕСЛИ (Меню: Вставка→Функция→Логические). Для нового левого края отрезка проверяем истинность условия $f(a)\cdot f(c)>0$. Если оно верно, то в качестве нового значения левого конца отрезка берем число c (так как это условие показывает, что корня на отрезке $[a;c]$ нет), иначе оставляем значение a. Аналогично, для нового правого края отрезка проверяем истинность условия $f(b)\cdot f(c)>0$. Если оно верно, то в качестве нового значения правого конца отрезка берем число c (так как это условие показывает, что корня на отрезке $[c;b]$ нет), иначе оставляем значение b.

Копируя вторую строку таблицы, можно создать необходимое число последующих строк.

Итерационный процесс завершается, когда очередное значение в последнем столбце становится меньшим, чем заданный показатель точности. При этом значение середины отрезка в последнем вичислении принимается в качестве приближенного значения искомого корня нелинейного уравнения.

Файл Правка Вид Вставка Формат Сервис Данные Окно Справка

B28 fx =ЕСЛИ(E27*F27>0;C27;B27)

	A	B	C	D	E	F	G	H
26	№	a	c	b	f(a)	f(c)	f(b)	ε
27	1	-4	-3,5	-3	-17	-0,875	10	0,5
28	2	-3,5	-3,25	-3	-0,875	5,171875	10	0,25
29	3	-3,5	-3,375	-3,25	-0,875	2,306641	5,171875	0,125
30	4	-3,5	-3,4375	-3,375	-0,875	0,756104	2,306641	0,0625
31	5	-3,5	-3,46875	-3,4375	-0,875	-0,049286	0,756104	0,03125
32	6	-3,46875	-3,45313	-3,4375	-0,04929	0,355938	0,756104	0,015625
33	7	-3,46875	-3,46094	-3,45313	-0,04929	0,15396	0,355938	0,007813
34	8	-3,46875	-3,46484	-3,46094	-0,04929	0,052496	0,15396	0,003906
35	9	-3,46875	-3,4668	-3,46484	-0,04929	0,001644	0,052496	0,001953
36	10	-3,46875	-3,46777	-3,4668	-0,04929	-0,023811	0,001644	0,000977
37	11	-3,46777	-3,46729	-3,4668	-0,02381	-0,011081	0,001644	0,000488
38	12	-3,46729	-3,46704	-3,4668	-0,01108	-0,004717	0,001644	0,000244
39	13	-3,46704	-3,46692	-3,4668	-0,00472	-0,001536	0,001644	0,000122
40	14	-3,46692	-3,46686	-3,4668	-0,00154	0,0000541	0,0016445	0,0000610

Рис. 3.13

Метод хорд

Нелинейная функция $f(x)$ на отделенном интервале $[a,b]$ заменяется линейной – уравнением хорды, т. е. прямой, соединяющей граничные точки графика на отрезке. Условие применимости метода – монотонность функции на начальном отрезке, обеспечивающая единственность корня на этом отрезке. Расчет по методу хорд аналогичен расчету методом деления отрезка пополам, но теперь на каждом шаге новая точка x внутри отрезка $[a;b]$ рассчитывается по любой из формул:

$$x = a - \frac{b-a}{f(b)-f(a)} \cdot f(a) = b - \frac{b-a}{f(b)-f(a)} \cdot f(b) = \frac{f(b)\cdot a - f(a)\cdot b}{f(b)-f(a)}.$$

Метод Ньютона (касательных)

Идея, на которой основан метод, аналогична той, что реализована в методе хорд, только на каждом шаге кривая $f(x)$ заменяется касательной к ней, проведенной в предыдущей найденной точке. В качестве начальной точки в зависимости от свойств функции берется или левая граница отрезка, содержащего корень $x_0 = a$ (если $f(a)\cdot f''(x) > 0$), или правая его граница $x_0 = b$ (если $f(b)\cdot f''(x) > 0$). Расчет нового приближения на следующем шаге $i+1$ производится по формуле

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}.$$

Алгоритм применим для монотонных функций, сохраняющих выпуклость или вогнутость в промежутке между начальным приближением и корнем уравнения (т. е. должен сохраняться знак первой и второй производных функции $f(x)$). В расчетах нет необходимости отслеживать две границы отрезка, поэтому достаточно на каждом шаге вычислять значения x, $f(x)$ и $f'(x)$. При этом легко оценить «точность по y», по значению левой части уравнения на очередном шаге. Для оценки «точности по x» нужно отслеживать разницу приближений на предыдущем и последующих шагах, которая связана с разницей между найденным приближением и точным значением корня.

Главным достоинством метода касательных является квадратичная скорость сходимости, что во многих случаях может привести к сокращению числа вычислений функции.

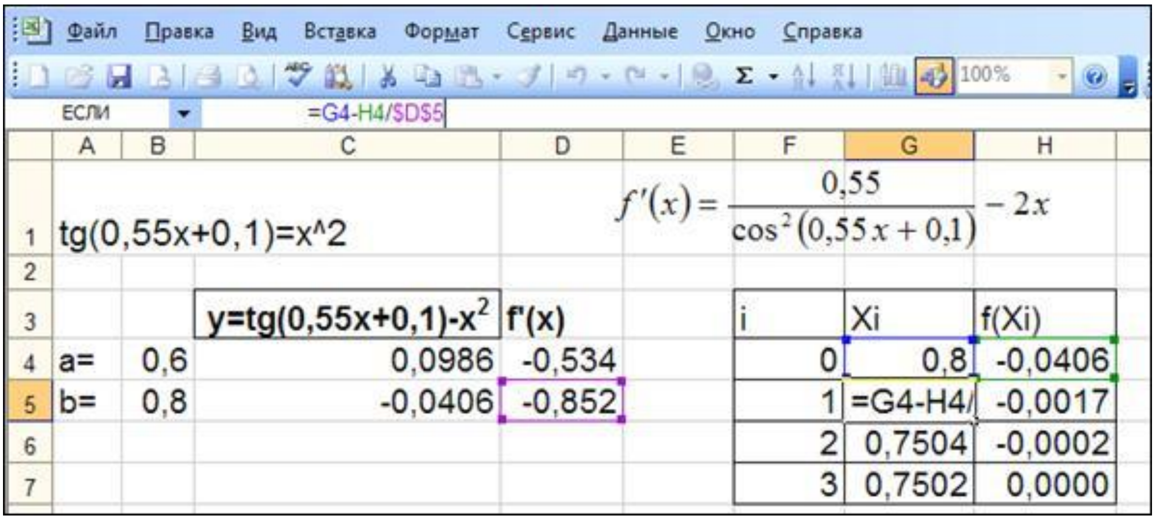
Пример 3.6. Определить корень уравнения $tg\left(0,55x+0,1\right)-x^2=0$ на отрезке $[0,6; 0,8]$ методом касательных до точности 0,001.

Решение

Поскольку $f(0,6)>0,\ f(0,8)<0,\ f''(x)<0$, то за начальное приближение примем $x_0=0,8$.

Для вычислений в Excel составляется таблица, показанная на рис. 3.14. Получим: x = 0,7502. Заданная точность обеспечивается выполнением условия $\left|f(x_i)\right|\le 0,001$.

Наиболее трудоемким элементом расчетов по методу Ньютона является вычисление производной на каждом шаге.



	A	B	C	D	E	F	G	H
1	tg(0,55x+0,1)=x^2				$f'(x)=\dfrac{0,55}{\cos^2(0,55x+0,1)}-2x$			
2								
3			$y=tg(0,55x+0,1)-x^2$	f'(x)		i	Xi	f(Xi)
4	a=	0,6	0,0986	-0,534		0	0,8	-0,0406
5	b=	0,8	-0,0406	-0,852		1	=G4-H4/	-0,0017
6						2	0,7504	-0,0002
7						3	0,7502	0,0000

Рис. 3.14

Решение систем нелинейных уравнений

Систему n нелинейных уравнений с n неизвестными $x_1, x_2, \ldots, x_n$ записывают в виде:

$$\begin{cases} F_1\left(x_1, x_2, ..., x_n\right) = 0, \\ F_2\left(x_1, x_2, ..., x_n\right) = 0, \\ \qquad ... \\ F_n\left(x_1, x_2, ..., x_n\right) = 0. \end{cases}$$

где $F_1,\ F_2, \ldots, F_n$ – функции независимых переменных, среди которых есть нелинейные.

Как и в случае систем линейных уравнений, решением системы является такой вектор X, который при подстановке обращает одновременно все уравнения системы в тождества:

$$X = \left(x_1;\, x_2; ...;\, x_n\right).$$

Система уравнений может не иметь решений, иметь единственное решение, конечное или бесконечное количество решений. Вопрос о количестве решений должен решаться для каждой конкретной задачи отдельно.

Численные методы решения системы уравнений носят итерационный характер и требуют задания начального приближения X_0.

Решение систем нелинейных уравнений методом Ньютона

Рассмотрим этот метод на примере системы двух нелинейных уравнений с двумя неизвестными:

$$F\left(X\right) = 0 \Leftrightarrow \begin{cases} f(x, y) = 0, \\ u(x, y) = 0. \end{cases}$$

Начальные значения x_0 и y_0 определяются графически. Для нахождения каждого последующего приближения $\left(x_{i+1}, y_{i+1}\right)$ используют вектор значений функций и матрицу значений их первых производных, рассчитанные в предыдущей точке $\left(x_i, y_i\right)$:

$$F\left(x^1, y^i\right) = \begin{pmatrix} f(x^1, y^1) \\ u(x^1, y^1) \end{pmatrix},$$

$$F^/\left(x^1, y^i\right) = \begin{pmatrix} f_x^/(x^1, y^i) & f_y^/(x^1, y^1) \\ u_x^/(x^i, y^i) & u_y^/(x^i, y^i) \end{pmatrix}.$$

Для расчета новых приближений на шаге $i+1$ используется матричная формула

$$\begin{pmatrix} x^{i+1} \\ y^{i+1} \end{pmatrix} = \begin{pmatrix} x^i \\ y^i \end{pmatrix} - \left(F^/(x^i, y^i)\right)^{-1} \cdot F(x^i, y^i).$$

Следует обратить внимание, что в последней формуле используется вычисление матрицы, обратной к матрице первых производных.

Расчет останавливают при выполнении одного (а иногда и обоих) из двух условий. Первое из них заключается в том, что на очередном шаге максимальное по модулю из изменений аргументов x и y становится меньше заданной погрешности по аргументам. Второе условие: на очередном шаге максимальное по модулю значение левых частей уравнений должно отличаться от нуля меньше, чем заданная погрешность по функциям.

При правильном использовании матричных операций эти формулы достаточно просто записываются в Excel. Правда, здесь придется заранее получить формулы для вычисления производных.

Решение систем нелинейных уравнений методами итераций

Для реализации этих методов исходную систему уравнений необходимо путем алгебраических преобразований явно выразить каждую переменную через остальные. Для случая двух уравнений с двумя неизвестными новая система будет иметь вид

$$\begin{cases} x = \varphi(x, y), \\ y = \psi(x, y). \end{cases}$$

Для решения такой системы задаются начальным приближением x_0 и y_0. Уточненные решения получают по шагам, подставляя в правые части уравнений значения, найденные на предыдущем шаге. В методе простых итераций для уточнения решения используют формулы:

$$\begin{cases} x^{i+1} = \varphi(x^i, y^i), \\ y^{i+1} = \psi(x^i, y^i). \end{cases}$$

Если одно из решений системы и начальные значения x_0 и y_0 лежат в области D, задаваемой неравенствами: $a \le x \le b$, $c \le y \le d$, то расчет по методу простых итераций сходится при выполнении в области D соотношений:

$$\left|\frac{\partial\varphi}{\partial x}\right| + \left|\frac{\partial\psi}{\partial x}\right| < 1, \ \left|\frac{\partial\varphi}{\partial y}\right| + \left|\frac{\partial\psi}{\partial y}\right| < 1.$$

§ 4. Методы оптимизации

Задача поиска экстремума некоторой функции при наличии ограничений или при отсутствии таковых решается элементарно лишь в случае одной переменной. Так, при поиске максимума $F(x)$ на интервале $[a; b]$ можно взять производную, приравнять ее нулю, решить уравнение

$F'(x) = 0$ (выше мы рассмотрели многообразие методов решения уравнений) и вычислить $F(x)$ в полученных «критических» точках, лежащих внутри интервала, и на концах интервала с последующим выбором максимального значения.

Если этот путь решения не подходит, то для «достаточно гладких» функций можно предложить эвристический прием. Сначала интервал разбивается на N частей (10, 100 и т. п.) длиной $h = \dfrac{b-a}{N}$ и отыскивается точка x полученной сетки значений аргумента с наибольшим значением $F(x)$. Затем берется окрестность найденной точки $(x \pm h)$, разбивается на N частей и т. д. до тех пор, пока длина очередного интервала поиска не станет меньше заданной точности. При этом существует опасность «проскочить» через глобальный экстремум за счет большого начального шага.

Если $F(x)$ заведомо унимодальна (например, имеет единственный максимум), то можно начать «прогулку» от точки a с некоторым начальным шагом до встречи с перевалом (значение функции не возросло, а наоборот уменьшилось), после чего отступить на шаг и продолжить с уменьшенным шагом.

Если эти пути нас не устраивают из-за значительного объема вычислений, то можно воспользоваться идеей градиентного спуска. Выбираем некоторую точку x_0 и находим в ней значения $F(x_0)$ и $F'(x_0)$. Смещаемся от этой точки влево или вправо, в зависимости от знака производной, на выбранный шаг h. Если значение функции в новой точке меньше значения в предыдущей, то уменьшаем шаг вдвое (втрое, в 10 раз), и возвращаемся в исходную точку с последующим переходом. В противном случае продолжаем процесс переходов либо до выхода на концы исходного интервала, либо до получения шага в пределах точности. Отметим, что здесь подстерегает опасность выйти на точку локального максимума (все зависит от начальной точки и начального шага).

MS Excel обладает мощным встроенным средством для нахождения экстремальных значений функции одной или нескольких переменных. Для одноэкстремальных функций можно найти безусловный глобальный экстремум, для многоэкстремальных – условный локальный экстремум.

Рассмотрим различные примеры поиска экстремальных значений функции.

Пример 3.7. Найти минимум функции $y = \dfrac{x^2}{x^2 + 16}$.

Решение

Для поиска безусловного экстремума функции сформируем лист электронной таблицы. Функцию запишем в ячейку A2: «=A1^2/(A1^2+16)», где вместо переменной *x* следует указать адрес ячейки A1, которая содержит начальное приближение экстремума равное нулю.

Алгоритм для поиска минимума:

выполнить команду *Данные/Поиск решения;*

заполнить диалоговое окно *Поиск решения*;

щелкнуть левой клавишей мыши в поле *Установить целевую ячейку* и щелкнуть на ячейке с формулой (в нашем случае это ячейка A2, абсолютный адрес которой $A$2 появится в поле);

выбрать поле *Минимальному значению*;

в поле *Изменяя ячейки* ввести адреса ячеек, значения которых будут варьироваться в процессе поиска решения (в нашем случае это ячейка A1, абсолютный адрес которой $A$1).

После выполнения лист электронной таблицы будет выглядеть так, как показано на рис. 3.15.

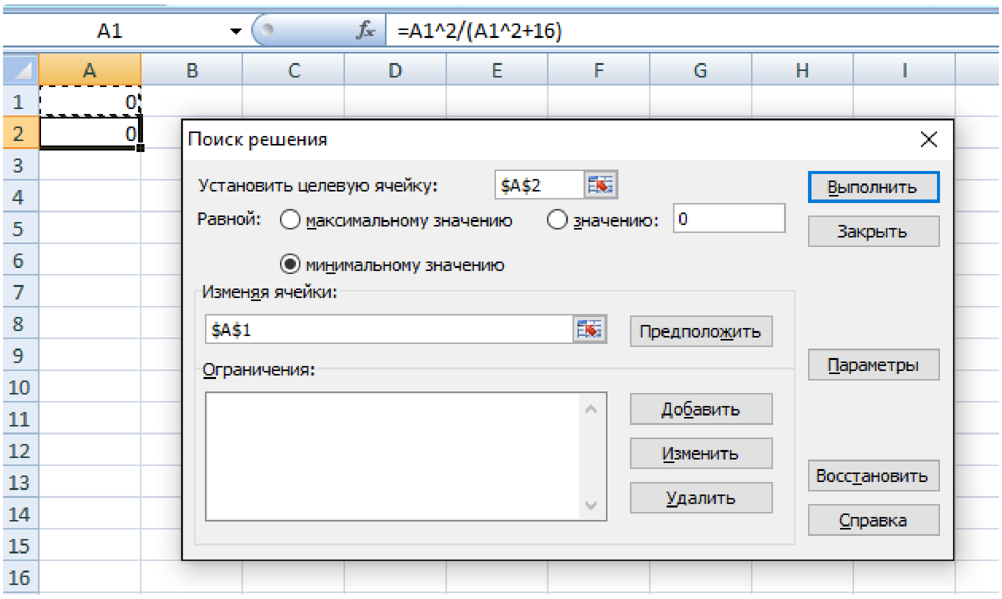


Рис. 3.15

После щелчка на кнопке *Выполнить* получим решение поставленной задачи.

В клетке A1 находится значение переменной *x*, равное нулю, при котором функция достигает минимального значения, равного нулю.

Отметим, что для многоэкстремальных функций определить, какой из локальных экстремумов будет найден, невозможно без построения графика функции на интересующем нас интервале. Численные методы нахождения экстремума ориентированы на поиск ближайшего решения к точке начального приближения и требуют унимодальности функции.

Условный экстремум

Для функции одной переменной поиск экстремума возможен как на всей числовой оси, так и на некотором интервале. Поиск на интервале уже можно считать поиском условного экстремума функции, так как появляются ограничения на изменение значений аргумента.

В диалогом окне *Поиск решения* есть поле *Ограничения* и соответствующие ему команды: *Добавить*, *Заменить*, *Удалить*.

Рассмотрим предыдущую задачу, добавив условие поиска минимального значения на интервале $[1;5]$. Тогда диалоговое окно *Поиск решения* следует видоизменить, добавив ограничения. Щелкнув левой клавишей мыши в поле *Ограничения* и затем на кнопке *Добавить*, откроем диалоговое окно *Добавление ограничения*, которое следует заполнить так как показано на рис. 3.16.

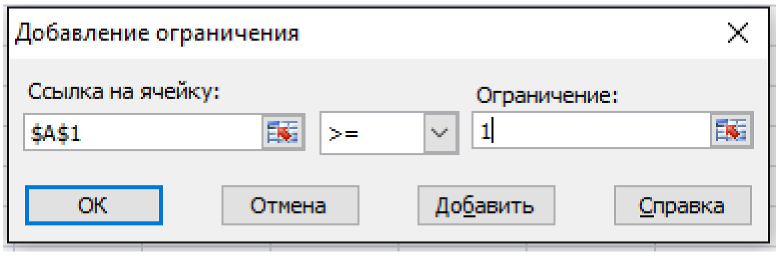


Рис. 3.16

После добавления последнего ограничения диалоговое окно *Поиск решения* будет содержать математическую постановку задачи экстремума и выглядеть следующим образом (рис. 3.17).

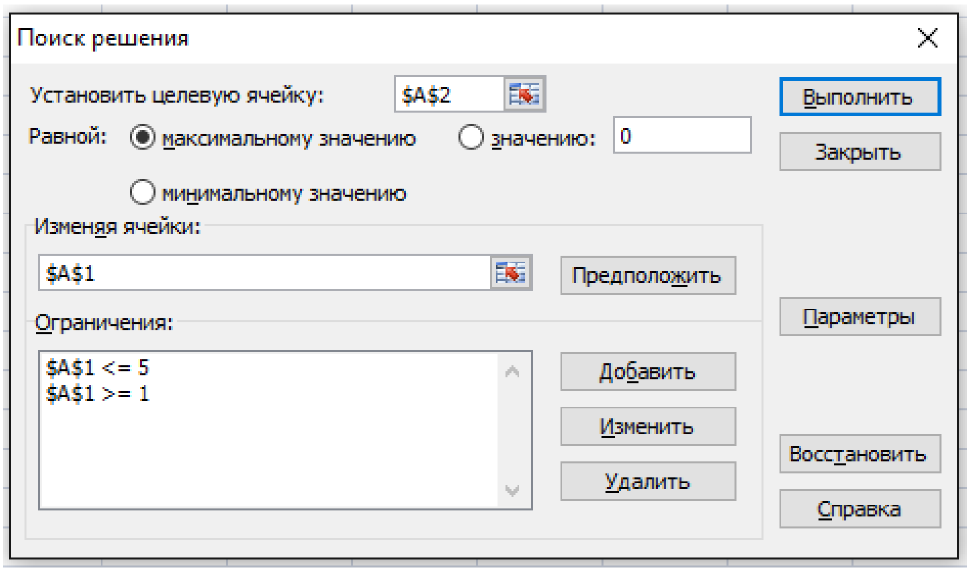


Рис. 3.17

После щелчка на кнопке *Выполнить* получим: $y = 0,058824$ при $x = 1$, что отличается от решения, полученного в предыдущем случае. Здесь в качестве минимального значения выступает наименьшее значение функции на интервале $[1; 5]$, совпадающее с левой границей интервала.

§ 5. Обыкновенные дифференциальные уравнения

Многие задачи науки и техники сводятся к решению обыкновенных дифференциальных уравнений (ОДУ), которые содержат одну или несколько производных от искомой функции. В общем виде ОДУ можно записать следующим образом:

$$F\left(x, y, y', y'', ..., y^{(n)}\right) = 0,$$

где x – независимая переменная;

$y^{(i)}$ – i-я производная от искомой функции;

n – порядок уравнения.

Общее решение ОДУ n-го порядка содержит n произвольных постоянных $c_1, c_2, ..., c_n$, т. е. общее решение имеет вид $y = \varphi\left(x, c_1, c_2, ..., c_n\right)$.

Для выделения единственного решения необходимо задать n дополнительных условий. В зависимости от способа задания дополнительных условий существуют два различных типа задач: задача Коши и краевая. Если дополнительные условия задаются в одной точке, то такая задача называется задачей Коши. Дополнительные условия в задаче Коши называются начальными условиями. Если же дополнительные условия задаются в более чем одной точке, т. е. при различных значениях независимой переменной, то такая задача называется краевой. Сами дополнительные условия называются краевыми или граничными.

Решить такие задачи аналитически удается лишь для некоторых специальных типов уравнений.

Численные методы решения задачи Коши для ОДУ первого порядка

Постановка задачи

Найдите решение ОДУ первого порядка $\dfrac{dy}{dx} = f\left(x, y\right)$ на отрезке $\left[x_0, x_n\right]$ при условии $y\left(x_0\right) = y_0$.

При нахождении приближенного решения будем считать, что вычисления проводятся с расчетным шагом $h = \dfrac{x_n - x_0}{n}$, расчетными узлами служат точки $x_i = x_0 + ih$, $(i = \overline{1, n})$ промежутка $\left[x_0, x_n\right]$.

Целью является построение таблицы:

x_i	x_0	x_1	…	x_n
y_i	y_0	y_1	…	y_n

т. е. находятся приближенные значения функции y в узлах сетки.

Интегрируя уравнение на отрезке $[x_i, x_{i+1}]$, получим:

$$y_{i+1} - y_i = \int\limits_{x_i}^{x_{i+1}} f(x, y)dx.$$

Возможным (но не единственным) путем получения численного решения является замена в нем интеграла какой-либо квадратурной формулой численного интегрирования. Если воспользоваться простейшей формулой левых прямоугольников первого порядка

$$\int\limits_{x_i}^{x_{i+1}} f(x, y)dx = hf(x_i, y_i),$$

то получим **явную формулу Эйлера**:

$$y_{i+1} = y_i + hf(x_i, y_i),\ i = \overline{1, n-1}.$$

Таким образом, порядок расчетов следующий:
зная x_0, y_0 и $f(x_0, y_0)$, находим $y_1 = y_0 + hf(x_0, y_0)$, затем $y_2 = y_1 + hf(x_1, y_1)$ и т. д.

Геометрическая интерпретация метода Эйлера (рис. 3.18)

Пользуясь тем, что в точке x_0 известно решение $y(x_0) = y_0$ и значение его производной $y'(x_0) = f(x_0, y_0)$, можно записать уравнение касательной к графику искомой функции $y = y(x)$ в точке (x_0, y_0):

$$y = y_0 + f(x_0, y_0)(x - x_0).$$

При достаточно малом шаге h ордината $y_1 = y_0 + hf(x_0, y_0)$ этой касательной, полученная подстановкой в правую часть значения $x_1 = x_0 + h$, должна мало отличаться от ординаты $y(x_1)$ решения $y(x)$ задачи Коши. Следовательно, точка (x_1, y_1) пересечения касательной с прямой $x = x_1$ может быть приближенно принята за новую начальную точку. Через эту точку снова проведем прямую $y = y_1 + f(x_1, y_1)(x - x_1)$, которая приближенно отражает поведение касательной к $y = y(x)$ в точке $(x_1, y(x_1))$. Подставив сюда $x_2 = x_1 + h$ (т. е. пересечение с прямой $x = x_2$), получим приближенное значение $y(x)$ в точке x_2: $y_2 = y_1 + hf(x_1, y_1)$ и т. д. В итоге для i-й точки получим формулу Эйлера.

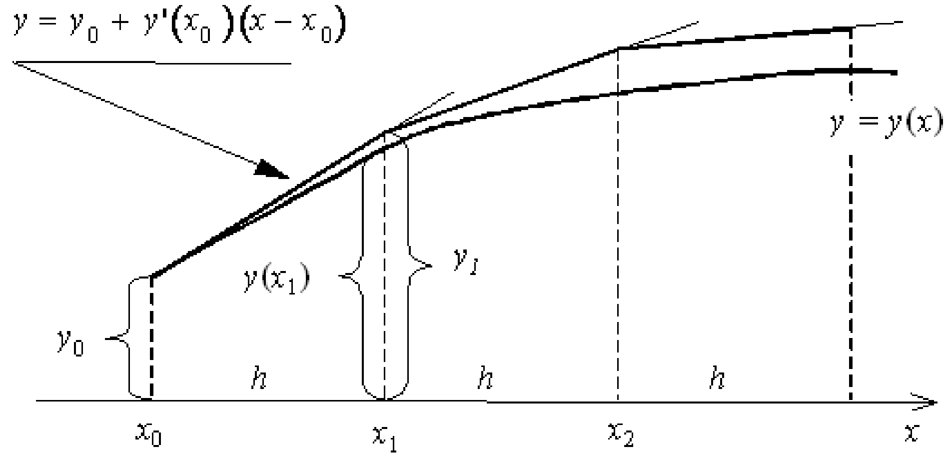


Рис. 3.18

Явный метод Эйлера имеет первый порядок точности или аппроксимации. Если использовать формулу правых прямоугольников $\int\limits_{x_i}^{x_{i+1}} f(x,y)dx = hf(x_{i+1}, y_{i+1})$, то придем к методу

$$y_{i+1} = y_i + hf(x_{i+1}, y_{i+1}),\ i = \overline{1, n-1}.$$

Этот метод называют **неявным методом Эйлера**, поскольку для вычисления неизвестного значения $y_{i+1} \approx y(x_{i+1})$ по известному значению $y_i \approx y(x_i)$ требуется решать уравнение, в общем случае нелинейное.

Неявный метод Эйлера имеет первый порядок точности или аппроксимации.

Модифицированный метод Эйлера

Вычисление y_{i+1} состоит из двух этапов:

$$\tilde{y}_{i+1} = y_i + hf(x_i, y_i),$$

$$y_{i+1} = y_i + \frac{h}{2}\left[f(x_i, y_i) + f(x_{i+1}, \tilde{y}_{i+1})\right].$$

Данная схема называется еще методом предиктор – корректор (предсказывающее – исправляющее). На первом этапе приближенное значение предсказывается с невысокой точностью (h), а на втором это предсказание исправляется, так что результирующее значение имеет второй порядок точности.

Методы Рунге – Кутта

Идея построения явных методов Рунге – Кутты p-го порядка заключается в получении приближений к значениям $y(x_{i+1})$ по формуле

$$y_{i+1} = y_i + h \cdot \varphi(x_i, y_i, h),$$

где

$$\varphi(x_i, y_i, h) = \sum_{n=1}^{q} c_n k_n^i(h),$$
$$k_1^i(h) = f(x_i, y_i),$$
$$k_2^i(h) = f(x_i + \alpha_2 h, y_i + \beta_{21} k_1^i(h)),$$
$$k_3^i(h) = f(x_i + \alpha_3 h, y_i + \beta_{31} k_1^i(h) + \beta_{32} k_2^i(h)),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$k_q^i(h) = f(x_i + \alpha_q h, y_i + \beta_{q1} k_1^i(h) + \ldots + \beta_{q,q-1} k_{q-1}^i(h)).$$

Здесь α_n, β_{nj}, p_n, $0 < j < n \le q$ – некоторые фиксированные числа (параметры).

При построения методов Рунге–Кутты параметры функции $\varphi(x_i, y_i, h)$ (α_n, β_{nj} , p_n) подбирают таким образом, чтобы получить нужный порядок аппроксимации.

Схема Рунге – Кутта четвертого порядка точности:

$$k_1 = f(x_i, y_i),$$
$$k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_1}{2}\right),$$
$$k_3 = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_2}{2}\right),$$
$$k_4 = f\left(x_i + h, y_i + hk_3\right),$$
$$y_{i+1} = y_i + \frac{h}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right),\ i = 0,1,2,\ldots$$

Решение задачи Коши в MS Excel (рис. 3.19)

Требуется решить задачу Коши $\frac{dy}{dx} = 2\left(x^2 + y\right)$, $y(0) = 1$ на отрезке $\left[0,1\right]$ с шагом $h = 0{,}1$.

Рассмотрим три метода: явный метод Эйлера, модифицированный метод Эйлера, метод Рунге – Кутта.

Точное решение: $y(x) = 1{,}5e^{2x} - x^2 - x - 0{,}5$.

1. Расчетные формулы по явному методу Эйлера для данного примера:

$$y_0 = 1,$$
$$y_{i+1} = y_i + h \cdot 2 \cdot \left(x_i^2 + y_i\right),$$
$$x_i = i \cdot h,\ i = \overline{1,10}.$$

2. Расчетные формулы модифицированного метода Эйлера:

$$y_0 = 1,$$
$$\tilde{y}_{i+1} = y_i + h \cdot 2 \cdot \left(x_i^2 + y_i\right),$$
$$y_{i+1} = y_i + \frac{h}{2} \cdot \left[2 \cdot \left(x_i^2 + y_i\right) + 2 \cdot \left(x_{i+1}^2 + \tilde{y}_{i+1}\right)\right],$$
$$x_i = i \cdot h,\ i = \overline{1,10}.$$

3. Расчетные формулы метода Рунге – Кутта:

$$y_0 = 1,$$
$$k_1 = 2 \cdot \left(x_i^2 + y_i\right),$$
$$k_2 = 2 \cdot \left(\left(x_i + \frac{h}{2}\right)^2 + y_i + \frac{hk_1}{2}\right),$$
$$k_3 = 2 \cdot \left(\left(x_i + \frac{h}{2}\right)^2 + y_i + \frac{hk_2}{2}\right),$$
$$k_4 = 2 \cdot \left((x_i + h)^2 + y_i + hk_3\right),$$
$$y_{i+1} = y_i + \frac{h}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right),$$
$$x_i = i \cdot h,\ i = \overline{1,10}.$$

	A	B	C	D	E	F	G	H	I	J	K
1	x	Эйлер	$\tilde{y}$	Модиф.Эйлер	k1	k2	k3	k4	Рунге-Кутта	Точное	
2	0	1	1	1					1	1	
3	0,1	1,2	1,2	1,221	2	2,205	2,2255	2,4651	1,22210167	1,222104	
4	0,2	1,442	1,442	1,4923	2,464203	2,735624	2,762766	3,076756	1,49773064	1,497737	
5	0,3	1,7384	1,7384	1,82837	3,075461	3,428007	3,463262	3,868114	1,84316587	1,843178	
6	0,4	2,10408	2,10408	2,246615	3,866332	4,317965	4,363128	4,878957	2,27829046	2,278311	
7	0,5	2,556896	2,556896	2,7679661	4,876581	5,449239	5,506505	6,157882	2,82738964	2,827423	
8	0,6	3,118275	3,118275	3,41759023	6,154779	6,875257	6,947305	7,76424	3,52012537	3,520175	
9	0,7	3,81393	3,81393	4,225742277	7,760251	8,661276	8,751378	9,770526	4,3927268	4,3928	
10	0,8	4,674716	4,674716	5,228788134	9,765454	10,887	10,99915	12,26528	5,48944418	5,489549	
11	0,9	5,73766	5,73766	6,470432901	12,25889	13,64978	13,78887	15,35666	6,86432478	6,864471	
12	1	7,047191	7,047191	8,003195337	15,34865	17,06851	17,2405	19,17675	8,58338196	8,583584	

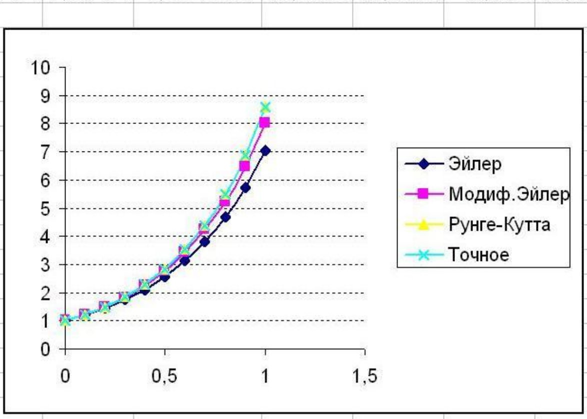


Рис. 3.19

Приложения

Приложение 1

Таблица значений функции $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$

	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	074	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Приложение 2

Таблица значений функции Лапласа $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{z^2}{2}}dz$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,0000	0,40	0,1554	0,80	0,2881	1,20	0,3849	1,60	0,4452	2,00	0,4772	2,80	0,4974
0,01	0,0040	0,41	0,1591	0,81	0,2910	1,21	0,3869	1,61	0,4463	2,02	0,4783	2,82	0,4976
0,02	0,0080	0,42	0,1628	0,82	0,2939	1,22	0,3883	1,62	0,4474	2,04	0,4793	2,84	0,4977
0,03	0,0120	0,43	0,1664	0,83	0,2967	1,23	0,3907	1,63	0,4484	2,06	0,4803	2,86	0,4979
0,04	0,0160	0,44	0,1700	0,84	0,2995	1,24	0,3925	1,64	0,4495	2,08	0,4812	2,88	0,4980
0,05	0,0199	0,45	0,1736	0,85	0,3023	1,25	0,3944	1,65	0,4505	2,10	0,4821	2,90	0,4981
0,06	0,0239	0,46	0,1772	0,86	0,3051	1,26	0,3962	1,66	0,4515	2,12	0,4830	2,92	0,4982
0,07	0,0279	0,47	0,1808	0,87	0,3078	1,27	0,3980	1,67	0,4525	2,14	0,4838	2,94	0,4984
0,08	0,0319	0,48	0,1844	0,88	0,3106	1,28	0,3997	1,68	0,4535	2,16	0,4846	2,96	0,4985
0,09	0,0359	0,49	0,1879	0,89	0,3133	1,29	0,4015	1,69	0,4545	2,18	0,4854	2,98	0,4986
0,10	0,0398	0,50	0,1915	0,90	0,3159	1,30	0,4032	1,70	0,4554	2,20	0,4861	3,00	0,49865
0,11	0,0438	0,51	0,1950	0,91	0,3186	1,31	0,4049	1,71	0,4564	2,22	0,4868	3,20	0,49931
0,12	0,0478	0,52	0,1985	0,92	0,3212	1,32	0,4066	1,72	0,4573	2,24	0,4875	3,40	0,49966
0,13	0,0517	0,53	0,2019	0,93	0,3238	1,33	0,4082	1,73	0,4582	2,26	0,4881	3,60	0,49984
0,14	0,0557	0,54	0,2054	0,94	0,3264	1,34	0,4099	1,74	0,4591	2,28	0,4887	3,80	0,49993
0,15	0,0596	0,55	0,2088	0,95	0,3289	1,35	0,4115	1,75	0,4599	2,30	0,4893	4,00	0,49997
0,16	0,0636	0,56	0,2123	0,96	0,3315	1,36	0,4131	1,76	0,4608	2,32	0,4898	4,50	0,49999
0,17	0,0675	0,57	0,2157	0,97	0,3340	1,37	0,4147	1,77	0,4616	2,34	0,4904	5,00	0,49999
0,18	0,0714	0,58	0,2190	0,98	0,3365	1,38	0,4162	1,78	0,4625	2,36	0,4909	если $x > 0,5$,	
0,19	0,0753	0,59	0,2224	0,99	0,3389	1,39	0,4177	1,79	0,4633	2,38	0,4913	то $\Phi(x) \cong 0,5$	
0,20	0,0793	0,60	0,2257	1,00	0,3413	1,40	0,4192	1,80	0,4641	2,40	0,4918		
0,21	0,0832	0,61	0,2291	1,01	0,3438	1,41	0,4207	1,81	0,4649	2,42	0,4922		
0,22	0,0871	0,62	0,2324	1,02	0,3461	1,42	0,4222	1,82	0,4656	2,44	0,4927		
0,23	0,0910	0,63	0,2357	1,03	0,3485	1,43	0,4236	1,83	0,4664	2,46	0,4931		
0,24	0,0948	0,64	0,2389	1,04	0,3508	1,44	0,4251	1,84	0,4671	2,48	0,4934		
0,25	0,0987	0,65	0,2422	1,05	0,3531	1,45	0,4265	1,85	0,4678	2,50	0,4938		
0,26	0,1026	0,66	0,2454	1,06	0,3554	1,46	0,4279	1,86	0,4686	2,52	0,4941		
0,27	0,1064	0,67	0,2486	1,07	0,3577	1,47	0,4292	1,87	0,4693	2,54	0,4945		
0,28	0,1103	0,68	0,2517	1,08	0,3599	1,48	0,4306	1,88	0,4699	2,56	0,4948		
0,29	0,1141	0,69	0,2549	1,09	0,3621	1,49	0,4319	1,89	0,4706	2,58	0,4951		
0,30	0,1179	0,70	0,2580	1,10	0,3643	1,50	0,4332	1,90	0,4713	2,60	0,4953		
0,31	0,1217	0,71	0,2611	1,11	0,3665	1,51	0,4345	1,91	0,4719	2,62	0,4956		
0,32	0,1255	0,72	0,2642	1,12	0,3686	1,52	0,4357	1,92	0,4726	2,64	0,4959		
0,33	0,1293	0,73	0,2673	1,13	0,3708	1,53	0,4370	1,93	0,4732	2,66	0,4961		
0,34	0,1331	0,74	0,2703	1,14	0,3729	1,54	0,4382	1,94	0,4738	2,68	0,4963		
0,35	0,1368	0,75	0,2734	1,15	0,3749	1,55	0,4394	1,95	0,4744	2,70	0,4965		
0,36	0,1406	0,76	0,2764	1,16	0,3770	1,56	0,4406	1,96	0,4750	2,72	0,4967		
0,37	0,1443	0,77	0,2794	1,17	0,3790	1,57	0,4418	1,97	0,4756	2,74	0,4969		
0,38	0,1480	0,78	0,2823	1,18	0,3810	1,58	0,4429	1,98	0,4761	2,76	0,4971		
0,39	0,1517	0,79	0,2852	1,19	0,3830	1,59	0,4441	1,99	0,4767	2,78	0,4973		

Приложение 3

Таблица значений t_γ распределения Стьюдента
в зависимости от числа степеней свободы k и надежности γ

k \ α	0,8	0,9	0,95	0,98	0,99	0,999
1	3,078	6,314	12,706	31,821	63,657	636,619
2	1,886	2,920	4,303	6,965	9,925	31,598
3	1,638	2,353	3,182	4,541	5,841	12,941
4	1,533	2,132	2,776	3,747	4,604	8,610
5	1,476	2,015	2,571	3,365	4,032	6,859
6	1,440	1,943	2,447	3,143	3,707	5,959
7	1,415	1,895	2,365	2,998	3,499	5,405
8	1,397	1,860	2,306	2,896	3,355	5,041
9	1,363	1,833	2,262	2,821	3,250	4,781
10	1,372	1,812	2,228	2,764	3,169	4,587
11	1,363	1,796	2,201	2,718	3,106	4,487
12	1,356	1,782	2,179	2,681	3,055	4,318
13	1,350	1,771	2,160	2,650	3,012	4,221
14	1,345	1,761	2,145	2,624	2,977	4,140
15	1,341	1,753	2,131	2,602	2,947	4,073
16	1,337	1,746	2,120	2,583	2,921	4,015
17	1,333	1,740	2,110	2,567	2,898	3,965
18	1,330	1,734	2,103	2,552	2,878	3,922
19	1,328	1,729	2,093	2,539	2,861	3,883
20	1,325	1,725	2,086	2,528	2,845	3,850
21	1,323	1,721	2,080	2,518	2,831	3,819
22	1,321	1,717	2,074	2,508	2,819	3,792
23	1,319	1,714	2,069	2,500	2,807	3,767
24	1,318	1,711	2,064	2,492	2,797	3,745
25	1,316	1,708	2,060	2,485	2,787	3,725
30	1,310	1,697	2,042	2,457	2,750	3,646
40	1,303	1,684	2,021	2,423	2,704	3,551
60	1,296	1,671	2,000	2,390	2,660	3,460
120	1,289	1,658	1,980	2,358	2,617	3,373
∞	1,282	1,645	1,960	2,326	2,576	3,291

Приложение 4

Значения чисел q в зависимости от объема выборки n и надежности γ для определения доверительного интервала среднего квадратичного отклонения σ

n	γ				n	γ		
	0,95	0,99	0,999			0,95	0,99	0,999
7	0,92	1,62	2,98		25	0,32	0,49	0,73
8	0,80	1,38	2,42		30	0,28	0,43	0,63
9	0,71	1,20	2,06		35	0,26	0,38	0,56
10	0,65	1,08	1,80		40	0,24	0,35	0,50
11	0,59	0,98	1,60		45	0,22	0,32	0,46
12	0,55	0,90	1,45		50	0,21	0,30	0,43
13	0,52	0,83	1,33		60	0,188	0,269	0,38
14	0,48	0,78	1,23		70	0,174	0,245	0,34
15	0,46	0,73	1,15		80	0,161	0,226	0,31
16	0,44	0,70	1,07		90	0,151	0,211	0,29
17	0,42	0,66	1,01		100	0,143	0,198	0,27
18	0,40	0,63	0,96		150	0,115	0,160	0,211
19	0,39	0,60	0,92		200	0,099	0,136	0,185
20	0,37	0,58	0,88		250	0,089	0,120	0,162

Приложение 5

Критические значения $r_{крит}$ критерия знаков в зависимости от n и уровня значимости α

n	α				n	α		
	0,01	0,05	0,10			0,01	0,05	0,10
6		0	0		13	1	2	3
7		0	0		14	1	2	3
8	0	0	1		15	2	3	3
9	0	1	1		20	3	5	5
10	0	1	1		25	5	7	7
11	0	1	2		30	7	9	10
12	1	2	2		35	9	11	12

Приложение 6

Значения χ_{α}^{2} в зависимости от k – числа степеней свободы

k \ α	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,455	1,074	1,642	2,71	3,84	5,41	6,64	10,83
2	1,386	2,41	3,22	4,60	5,99	7,82	9,21	13,82
3	2,37	3,66	4,64	6,25	7,82	9,84	11,34	16,27
4	3,36	4,88	5,99	7,78	9,49	11,67	13,28	18,46
5	4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,5
6	5,35	7,23	8,56	10,64	12,59	15,03	16,81	22,5
7	6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,3
8	7,34	9,52	11,03	13,36	15,51	18,17	20,1	26,1
9	8,34	10,66	12,24	14,68	16,92	19,68	21,7	27,9
10	9,34	11,78	13,44	15,99	18,31	21,2	23,2	29,6
11	10,34	12,90	14,63	17,28	19,68	22,6	24,7	31,3
12	11,34	14,01	15,81	18,55	21,0	24,1	26,2	32,9
13	12,34	15,12	16,98	19,81	22,4	25,5	27,7	34,6
14	13,34	16,22	17,15	21,1	23,7	26,9	29,1	36,1
15	14,34	17,32	19,31	22,3	25,0	28,3	30,6	37,7
16	15,34	18,42	20,5	23,5	26,3	29,6	32,0	39,3
17	16,34	19,51	21,6	24,8	27,6	31,0	33,4	40,8
18	17,34	20,6	22,8	26,0	28,9	32,3	34,8	42,3
19	18,34	21,7	23,9	27,2	30,1	33,7	36,2	43,8
20	19,34	22,8	25,0	28,4	31,4	35,0	37,6	45,3
21	20,3	23,9	26,2	29,6	32,7	36,3	38,9	46,8
22	21,3	24,9	27,3	30,8	33,9	37,7	40,3	48,3
23	22,3	26,0	28,4	32,0	35,2	39,0	41,6	49,7
24	23,3	27,1	29,6	33,2	36,4	40,3	43,0	51,2
25	24,3	28,2	30,7	34,4	37,7	41,7	44,3	52,6
26	25,3	29,2	31,8	35,6	38,9	42,9	45,6	54,1
27	26,3	30,3	32,9	36,7	40,1	44,1	47,0	55,5
28	27,3	31,4	34,0	37,9	41,3	45,4	48,3	56,9
29	28,3	32,5	35,1	39,1	42,6	46,7	49,6	58,3
30	29,3	33,5	36,2	40,3	43,8	48,0	50,9	59,7

Приложение 7

Критические точки распределения F Фишера – Снедекора

Уровень значимости $\alpha = 0{,}05$											
k_1 / k_2	1	2	3	4	5	6	7	8	9	10	12
1	161	200	216	225	230	234	237	239	241	242	244
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,41
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,74
4	7,71	6,94	6,59	6,39	6,26	6,16	6,00	6,04	6,00	5,96	5,91
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,68
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,57
8	5,32	4,45	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,28
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,13	3,07
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,97	2,91
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,86	2,79
12	4,75	3,88	3,49	3,26	3,11	3,00	2,92	2,85	2,80	2,76	2,69
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72	2,67	2,60
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65	2,60	2,53
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59	2,55	2,48
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50	2,45	2,38
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16

Библиографический список

1. Федосеева, В.В. Экономико-математические методы и прикладные модели / В.В. Федосеева [и др.]. М.:ЮНИТИ, 2000. 391 с.

2. URL: http://elearning.tver.ru/course/view.php?id=257, пароль: Optimus, файл: Optimus1.xlsx.

3. URL: http://elearning.tver.ru/course/view.php?id=257, пароль: Optimus, файл: Optimus2.xlsx.

4. Макарова, Н.В. Статистика в Excel.: учеб. пособие / Н.В. Макарова, В.Я. Трофимец. М.: Финансы и статистика, 2002. 368 с.

5. Гмурман, В.Е. Теория вероятностей и математическая статистика: учеб. пособие для вузов / В.Е. Гмурман. 11-е изд. М.: Издательство Юрайт, 2010. 479 с.

6. Большев, Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. М.: Наука, 1983.

Балашов Александр Николаевич
Пиджакова Любовь Михайловна
Шестакова Маргарита Аркадьевна

Решение прикладных задач аналитическими и численными методами

Редактор М.Б. Юдина
Корректор Я.А. Петрова
Технический редактор Ю.Ф. Воробьева

Подписано в печать
Формат 60х84/16 Бумага писчая
Физ. печ. л. Усл. печ. л. Уч.-изд. л.
Тираж 100 экз. Заказ №

Редакционно-издательский центр
Тверского государственного технического университета
170026, г. Тверь, наб. А. Никитина, 22