Министерство образования и науки Российской Федерации

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ
И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

ФАКУЛЬТЕТ ДИСТАНЦИОННОГО ОБУЧЕНИЯ (ФДО)

И. В. Подопригора

МАТЕМАТИЧЕСКИЕ МОДЕЛИ В ЭКОНОМИКЕ

Учебное пособие

Томск
2016

УДК 330.4(075.8)
ББК 65в641я73
П 444

Рецензенты:

Н. А. Ярушкина, канд. экон. наук, доцент кафедры экономики Томского государственного архитектурно-строительного университета;

Г. А. Золотарева, канд. экон. наук, доцент кафедры экономики ТУСУР

Подопригора И. В.

П 444 Математические модели в экономике : учебное пособие / И. В. Подопригора. – Томск : ФДО ТУСУР, 2016. – 161 с.

В учебном пособии «Математические модели в экономике» рассмотрены основные процедуры обработки и анализа данных, дается представление о наиболее распространённых математических методах, используемых для формализации экономико-математических моделей.

Рассматриваются оптимизационные методы математики в экономике, в частности линейное и нелинейное программирование, элементы теории игр и статистических решений, модели управления запасами, теория массового обслуживания, методы сетевого планирования. Приводится описание теоретических основ выборочного метода и методов проверки статистических гипотез. Особое внимание уделено регрессионному анализу, анализу временных рядов. Дается интерпретация результатов экономико-математического моделирования и рекомендации по применению их для обоснования конкретных хозяйственных решений.

Для студентов, бакалавров экономических специальностей.

© Подопригора И. В., 2016
© Оформление.
ФДО, ТУСУР, 2016

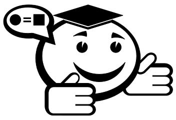
Оглавление

1 Введение 5
1.1 Что такое модель и для чего она нужна? 5
1.2 Как построить модель? 9
2 Оптимизационные методы математики в экономике 12
2.1 Оптимизационные модели 12
2.2 Оптимизационные задачи с линейной зависимостью между переменными 13
2.2.1 Геометрическая интерпретация оптимизационных задач линейного программирования 14
2.2.2 Симплексный метод решения оптимизационных задач линейного программирования 17
2.2.3 Двойственная задача линейного программирования 23
2.2.4 Транспортная задача 25
3 Нелинейное программирование 31
3.1 Специфика нелинейных программ и методы их решения 32
3.2 Теорема Куна – Таккера 35
3.3 Квадратичное программирование. Метод Вулфа – Фрэнка 37
3.4 Дробно-линейное программирование 41
4 Элементы теории игр и статистических решений 44
4.1 Основные понятия теории игр 44
4.2 Матричные игры и линейное программирование 47
4.3 Итеративный метод решения матричных игр 49
4.4 Многошаговые игры. Игры на выживание 50
4.5 Многошаговые игры. Игры погони 52
4.6 Статистические решения. Основные понятия 53
4.7 Выбор критерия принятия решения 55
4.7.1 Критерий Лапласа 55
4.7.2 Критерий Вальда 56
4.7.3 Критерий Гурвица 57
4.7.4 Критерий Сэвиджа 57
5 Ряды динамики 60
5.1 Понятие о рядах динамики 60
5.2 Показатели изменения уровней ряда динамики 61
5.3 Методы выявления основной тенденции (тренда) в рядах динамики 67

5.4 Оценка адекватности тренда и прогнозирование 74
6 Изучение взаимосвязей явлений 77
6.1 Понятие корреляционной зависимости 77
6.2 Методы выявления и оценки корреляционной связи 81
7 Модели управления запасами 96
7.1 Основные понятия теории управления запасами и ее элементы 96
7.2 Классификация моделей управления запасами 101
7.3 Детерминированные модели 105
7.3.1 Модель Уилсона 105
7.3.2 Модель с конечной интенсивностью поступления заказа 108
7.3.3 Модель с учетом неудовлетворенных требований 109
7.3.4 Модель с определением точки заказа 111
8 Теория массового обслуживания 113
8.1 Понятие о задачах теории массового обслуживания 113
8.2 Основы математического аппарата анализа простейших СМО 114
8.3 Основные характеристики СМО 118
8.4 Примеры систем с ограниченной очередью 120
8.5 Дисциплина ожидания и приоритеты 121
8.6 Моделирование систем массового обслуживания и метод Монте-Карло 123
9 Сетевое планирование 126
9.1 Понятие о сетевом графике 126
9.2 Критический путь и другие параметры сетевого графика 130
9.3 Линейная диаграмма проекта (диаграмма Ганта) 135
9.4 Минимизация стоимости проекта при заданной продолжительности 136
9.5 Проблемы применения систем сетевого планирования 144
Заключение 147
Литература 149
Приложение А. Значения интеграла Лапласа 152
Приложение Б. Значения *t*-критерия Стьюдента 153
Приложение В. Значения *F*-критерия Фишера 154
Глоссарий 156

1 Введение

1.1 Что такое модель и для чего она нужна?



В общем виде ***модель*** *можно определить как условный образ (упрощенное изображение) реального объекта (процесса), который создается для более глубокого изучения действительности. Метод исследования, базирующийся на разработке и использовании моделей, называется* ***моделированием****.*

Например, модель самолета продувают в аэродинамической трубе, вместо того чтобы испытывать настоящий самолет, это дешевле. При теоретическом исследовании атомного ядра физики представляют его в виде капли жидкости, имеющей поверхностное натяжение, вязкость и т. п. Необходимость моделирования обусловлена сложностью, а порой и невозможностью прямого изучения реального объекта (процесса). Значительно доступнее создавать и изучать прообразы реальных объектов (процессов), т. е. модели. Можно сказать, что теоретическое знание о чем-либо, как правило, представляет собой совокупность различных моделей. Эти модели отражают существенные свойства реального объекта (процесса), хотя на самом деле действительность значительно содержательнее и богаче [1].

Подобие между моделируемым объектом и моделью может быть физическое, структурное, функциональное, динамическое, вероятностное и геометрическое. При физическом подобии объект и модель имеет одинаковую или сходную физическую природу. Структурное подобие предполагает наличие сходства между структурой объекта и структурой модели. При выполнении объектом и моделью под определенным воздействием сходных функций наблюдается функциональное подобие. При наблюдении за последовательно изменяющимися состояниями объекта и модели отмечается динамическое подобие. Вероятностное подобие отмечается при наличии сходства между процессами вероятностного характера в объекте и модели. Геометрическое подобие имеет место при сходстве пространственных характеристик объекта и модели.

На сегодняшний день общепризнанной единой классификации моделей не существует. Однако из множества моделей можно выделить словесные, графические, физические, экономико-математические и некоторые другие типы моделей.

Словесная, или монографическая, модель представляет собой словесное описание объекта, явления или процесса. Очень часто она выражается в виде определения, правила, теоремы, закона или их совокупности.

Графическая модель создается в виде рисунка, географической карты или чертежа. Например, зависимость между ценой и спросом может быть выражена в виде графика, на оси ординат которого отложен спрос (*D*), а на оси абсцисс – цена (*P*). Кривая нам наглядно иллюстрирует, что с ростом цены спрос падает, и наоборот. Конечно, данную зависимость можно выразить и словесно, но графически она намного нагляднее (рис. 1.1).

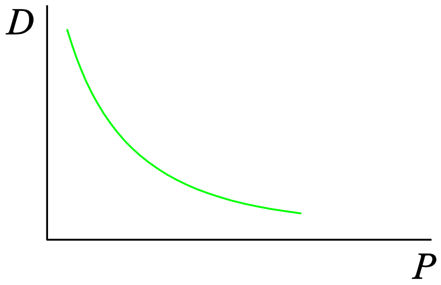


Рис. 1.1 – Графическая модель, отображающая зависимость между спросом и ценой

Физические или вещественные модели создаются для конструирования пока еще не существующих объектов. Создать модель самолета или ракеты для проверки ее аэродинамических свойств значительно проще и экономически целесообразнее, чем изучать эти свойства на реальных объектах.

Экономико-математические модели отражают наиболее существенные свойства реального объекта или процесса с помощью системы уравнений. *Содержанием* любой экономико-математической модели является выраженная в формально-математических соотношениях экономическая сущность условий задачи и поставленной цели. В модели экономическая величина представляется математическим соотношением, но не всегда математическое соотношение является экономическим [2].

Экономико-математическая модель представляет собой концентрированное выражение общих взаимосвязей и закономерностей экономического явления в математической форме (академик В. С. Немчинов).

Единой классификации экономико-математических моделей также не существует, хотя можно выделить наиболее значимые их группы в зависимости от признака классификации.

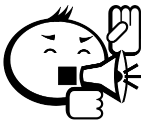
По степени агрегирования объектов моделирования различают модели:

- микроэкономические;
- одно-, двухсекторные (одно-, двухпродуктовые);
- многосекторные (многопродуктовые);
- макроэкономические;
- глобальные.

По учету фактора времени модели подразделяются:

- на статические;
- динамические.

В статических моделях экономическая система описана в статике, применительно к одному определенному моменту времени. Это как бы снимок, срез, фрагмент динамической системы в какой-то момент времени. Динамические модели описывают экономическую систему в развитии.



По цели создания и применения различают модели:

- балансовые;
- эконометрические;
- оптимизационные;
- сетевые;
- систем массового обслуживания;
- имитационные (экспертные).

В *балансовых моделях* отражается требование соответствия наличия ресурсов и их использования.

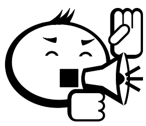
Параметры *эконометрических моделей* оцениваются с помощью методов математической статистики. Наиболее распространены эконометрические модели, представляющие собой системы регрессионных уравнений. В данных уравнениях отражается зависимость эндогенных (зависимых) переменных от экзогенных (независимых) переменных. Данная зависимость в основном выражается через тренд (длительную тенденцию) основных показателей моделируемой экономической системы. Эконометрические модели используются для анализа и прогнозирования конкретных экономических процессов с использованием реальной статистической информации.

Оптимизационные модели позволяют найти из множества возможных (альтернативных) вариантов наилучший вариант производства, распределения или потребления. Ограниченные ресурсы при этом будут использованы наилучшим образом для достижения поставленной цели.

Сетевые модели наиболее широко используются в управлении проектами. Сетевая модель отображает комплекс работ (операций) и событий и их взаимосвязь во времени. Обычно сетевая модель предназначена для выполнения работ в такой последовательности, чтобы сроки выполнения проекта были минимальными. В этом случае ставится задача нахождения критического пути. Однако существуют и такие сетевые модели, которые ориентированы не на критерий времени, а, например, на минимизацию стоимости работ.

Модели систем массового обслуживания создаются для минимизации затрат времени на ожидание в очереди и времени простоев каналов обслуживания.

Имитационная модель наряду с машинными решениями содержит блоки, где решения принимаются человеком (экспертом). Вместо непосредственного участия человека в принятии решений может выступать база знаний. В этом случае ЭВМ, специализированное программное обеспечение, база данных и база знаний образуют экспертную систему. Экспертная система предназначена для решения одной или ряда задач методом имитации действий человека, эксперта в данной области [3].



По учету фактора неопределенности модели подразделяются:

- на детерминированные (с однозначно определенными результатами);
- стохастические (с различными, вероятностными результатами).

По типу математического аппарата различают модели:

- линейного и нелинейного программирования;
- корреляционно-регрессионные;
- матричные;
- сетевые;
- теории игр;

- теории массового обслуживания и т. д.

. . . .

1.2 Как построить модель?

Чтобы воспользоваться математической моделью для конкретной производственно-экономической ситуации, следует применить информационную технологию. Информационная технология позволяет безошибочно выделить из множества реальных производственно-экономических ситуаций именно ту, которая полностью соответствует конкретным обстоятельствам. Эта технология состоит из следующих восьми этапов [4].

Этап 1. Выбор объекта моделирования (например: склад готовой продукции; организация выпуска новой продукции или системы транспортных перевозок и т. п.).

Этап 2. Анализ проблемной ситуации, сложившейся в рассматриваемом объекте моделирования. Например, для нормального функционирования склада готовой продукции необходимо увязать скорость потребления продукции со временем поставки и размерами складских площадей, оборотными средствами, которые всегда оказываются ограниченными.

Этап 3. Тип и число ненаблюдаемых параметров (отыскиваемых значений ЦФ и основных переменных X и j), определение которых позволит выбрать обоснованное управление конкретного экономического объекта.

Этап 4. Тип и число наблюдаемых параметров (задаваемых значений правых частей ограничений $b[i]$, коэффициентов затрат $a[ij]$, граничных условий для отыскиваемых переменных).

Этап 5. Условие адекватности, то есть уверенность в том, что математическая модель экономического объекта полностью (или в главных чертах) характеризует его действительное оптимальное функционирование. Обычно адекватность ставится в зависимость от численного значения критерия оптимальности (или нескольких таких критериев при многокритериальной оптимизации).

Этап 6. Используемый математический аппарат, соответствующий конкретному математическому описанию производственно-экономической ситуации (например, аналитические связи между основными параметрами движения запасов).

Этап 7. Анализ результатов моделирования экономического объекта: оптимальных значений основных переменных и целевой функции. Эти значения составляют основу экономического анализа конкретного объекта, за которым следуют выводы.

Этап 8. Принятие решения. По результатам оптимальных значений и сделанных на этапе 7 выводов принимается решение по управлению экономическим объектом.

Моделирование – циклический процесс. Это означает, что за первым восьмиэтапным циклом может последовать второй, третий и т. д. При этом знания об исследуемом объекте расширяются и уточняются, а исходная модель постепенно совершенствуется. Недостатки, обнаруженные после первого цикла моделирования, обусловленные малым знанием объекта и ошибками в построении модели, можно исправить в последующих циклах. В методологии моделирования, таким образом, заложены большие возможности саморазвития.

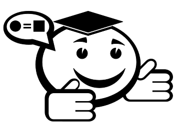
По мере развития и усложнения экономико-математического моделирования его отдельные этапы обособляются в специализированные области исследований, усиливаются различия между теоретико-аналитическими и прикладными моделями, происходит дифференциация моделей по уровням абстракции и идеализации.

Теория математического анализа моделей экономики развилась в особую ветвь современной математики – *математическую экономику*. Модели, изучаемые в рамках математической экономики, теряют непосредственную связь с экономической реальностью; они имеют дело с исключительно идеализированными экономическими объектами и ситуациями. При построении таких моделей главным принципом является не столько приближение к реальности, сколько получение возможно большего числа аналитических результатов посредством математических доказательств. Ценность этих моделей для экономической теории и практики состоит в том, что они служат теоретической базой для моделей прикладного типа.

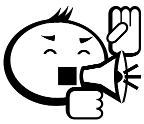
Довольно самостоятельными областями исследований становятся *подготовка и обработка экономической информации и разработка математического обеспечения экономических задач (создание баз данных и банков информации, программ автоматизированного построения моделей и программного сервиса для экономистов-пользователей)*. На этапе практического использования моделей ведущую роль должны играть специалисты в соответствующей области экономического анализа, планирования, управления. Главным участком работы экономистов-математиков остается постановка и формализация экономических задач и синтез процесса экономико-математического моделирования.

Соглашения, принятые в учебном пособии

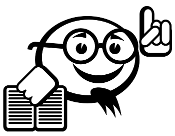
Для улучшения восприятия материала в данном учебном пособии используются пиктограммы и специальное выделение важной информации.



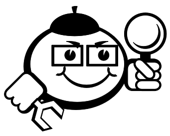
Эта пиктограмма означает определение или новое понятие.



Эта пиктограмма означает «Внимание!». Здесь выделена важная информация, требующая акцента на ней. Автор может поделиться с читателем опытом, чтобы помочь избежать некоторых ошибок.

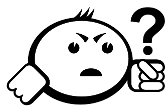


Эта пиктограмма означает теорему.



Пример

Эта пиктограмма означает пример. В данном блоке автор может привести практический пример для пояснения и разбора основных моментов, отраженных в теоретическом материале.



Контрольные вопросы по главе

2 Оптимизационные методы математики в экономике

2.1 Оптимизационные модели

*Экономико-математические задачи, цель которых состоит в нахождении наилучшего (оптимального) с точки зрения некоторого критерия или критериев варианта использования имеющихся ресурсов (труда, капитала и пр.), называются **оптимизационными**.*

Оптимизационные задачи (ОЗ) решаются с помощью оптимизационных моделей (ОМ) методами математического программирования.

Структура оптимизационной модели состоит из целевой функции, области допустимых решений и системы ограничений, определяющих эту область [5].

Область допустимых решений *– это область, в пределах которой осуществляется выбор решений. В экономических задачах она ограничена наличными ресурсами, условиями, которые записываются в виде системы ограничений, состоящей из уравнений и неравенств.*

Если система ограничений несовместима, то область допустимых решений является пустой. Ограничения подразделяются:

а) на линейные (I и II) и нелинейные (III и IV) (рис. 2.1);

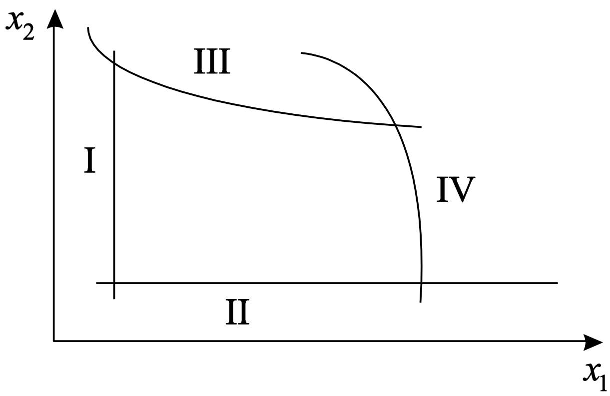


Рис. 2.1 – Линейные и нелинейные ограничения

б) детерминированные (A, B) и стохастические (группы кривых C_i) (рис. 2.2).

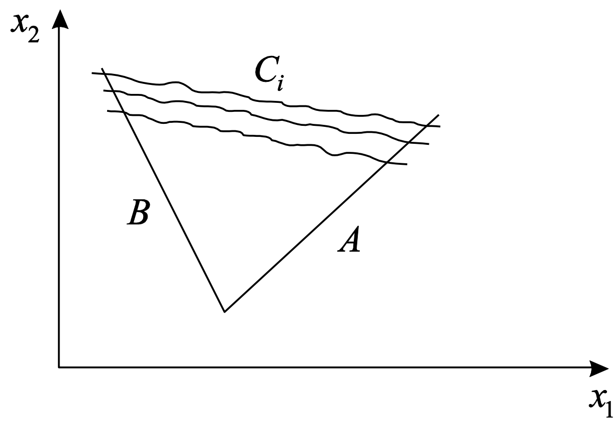


Рис. 2.2 – Детерминированные и стохастические ограничения

Стохастические ограничения являются возможными, вероятностными, случайными.

Оптимизационные задачи решаются методами математического программирования, которые подразделяются:

- на линейное программирование;
- нелинейное программирование;
- динамическое программирование;
- целочисленное программирование;
- выпуклое программирование;
- исследование операций;
- геометрическое программирование и др.

Главная задача математического программирования – это нахождение экстремума функций при ограничениях в форме уравнений и неравенств.

Рассмотрим оптимизационные задачи, решаемые методами линейного программирования.

2.2 Оптимизационные задачи с линейной зависимостью между переменными

Пусть:

b_i – количество ресурса вида i $(i = 1, 2, ..., m)$;

a_{ij} – норма расхода i-го ресурса на единицу j-го вида продукции;

x_j – количество продукции вида j $(j = 1, 2, ..., n)$;

c_j – прибыль (доход) от единицы этой продукции (в задачах на минимум – себестоимость продукции).

Тогда оптимизационная задача линейного программирования (ЛП) в общем виде может быть сформулирована и записана следующим образом:

Найти переменные x_j $(j = 1, 2, \ldots, n)$, при которых целевая функция

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

была бы максимальной (минимальной), не нарушая следующих ограничений:

$$\sum_{j=1}^{n} a_{ij}\ x_j \le b_i, \quad i = 1, 2, \ldots, m_1,$$

$$\sum_{j=1}^{n} a_{ij}\ x_j = b_i, \quad i = m_1 + 1,\ m_1 + 2, \ldots, m_2,$$

$$\sum_{j=1}^{n} a_{ij}\ x_j \ge b_i, \quad i = m_2 + 1,\ m_2 + 2, \ldots, m.$$

Все три случая можно привести к так называемой канонической форме, введя дополнительные переменные:

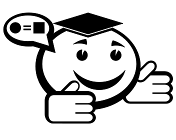
$$\sum_{j=1}^{n+k} a_{ij} x_j = b_i \quad i = 1, 2, \ldots, m,$$

k – количество дополнительных переменных, и условие неотрицательности искомых переменных:

$$x_j \ge 0.$$

В результате решения задачи находится некий план (программа) работы некоторого предприятия. Отсюда и появилось слово «программирование» [6]. Слово «линейное» указывает на линейный характер зависимости как в целевой функции, так и в системе ограничений. Следует еще раз подчеркнуть, что задача обязательно носит экстремальный характер, т. е. состоит в отыскании максимума или минимума (экстремума) целевой функции.

2.2.1 Геометрическая интерпретация оптимизационных задач линейного программирования

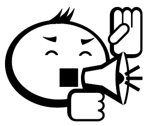


Графический метод *применяется для решения задач с двумя неизвестными* x_j, *j=1,2. Все условия выражаются в виде линейных ограничений, и записывается целевая функция Z.*

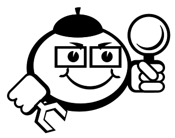
Алгоритм графического метода

Первый шаг при использовании графического метода заключается в геометрическом представлении допустимых решений, т. е. построении области допустимых значений, в которой одновременно удовлетворяются все ограничения модели. Условия неотрицательности переменных $x_1 \geq 0$ и $x_2 \geq 0$ ограничивают область их допустимых значений первым квадрантом. Другие границы пространства решений изображены на плоскости прямыми линиями, построенными по уравнениям, которые получаются при замене знака $\leq$ на знак = в остальных ограничениях. Области, в которых выполняются соответствующие ограничения в виде неравенств, указываются стрелками, направленными в сторону допустимых значений переменных. В каждой точке, принадлежащей внутренней области или границам многоугольника решений, все ограничения выполняются, поэтому решения, соответствующие этим точкам, являются допустимыми.

Пространство решений содержит бесконечное число таких точек, но, несмотря на это, можно найти оптимальное решение, выяснив, в каком направлении возрастает (или убывает) целевая функция. На график наносят ряд параллельных линий, соответствующих уравнению целевой функции при нескольких произвольно выбранных и последовательно возрастающих значениях Z, что позволяет определить наклон целевой функции и направление, в котором происходит ее увеличение (уменьшение).



Чтобы найти оптимальное решение, следует перемещать прямую Z в направлении возрастания (убывания) целевой функции до тех пор, пока она не сместится в область недопустимых решений.



Пример 2.1

Цех выпускает в смену трансформаторы двух видов. Для их изготовления используются железо и проволока. Общий запас железа – 24 кг, проволоки – 18 кг. На один трансформатор первого вида расходуются 3 кг железа и 3 кг проволоки, а на один трансформатор второго вида – 4 кг железа и 2 кг проволоки. За каждый реализованный трансформатор первого вида завод получает прибыль 4 д. е., второго – 2 д. е.

Необходимо составить план выпуска трансформаторов, обеспечивающий заводу максимальную прибыль в смену, если в смену должно выпускаться не более 4 трансформаторов одного вида.

В данной задаче две неизвестные: x_1 – выпуск трансформаторов Т1 за смену, x_2 – выпуск трансформаторов Т2 за смену.

Ограничения накладываются на запасы железа и проволоки, а также на выпуск трансформаторов Т1 за смену:

$$\begin{cases} 3x_1 + 4x_2 \le 24, \\ 3x_1 + 2x_2 \le 18, \\ x_1 \le 4, \\ x_1 \ge 0, \ x_2 \ge 0. \end{cases}$$

Целевая функция имеет вид: $Z = 4x_1 + 2x_2 \to \max$.

Определяем точки, необходимые для построения множества допустимых значений:

1. $3x_1 + 4x_2 = 24$: $x_1 = 0 \Rightarrow x_2 = 24/4 = 6$,
   $x_2 = 0 \Rightarrow x_1 = 24/3 = 8$.
2. $3x_1 + 2x_2 = 18$: $x_1 = 0 \Rightarrow x_2 = 18/2 = 9$,
   $x_2 = 0 \Rightarrow x_1 = 18/3 = 6$.
3. $x_1 = 4 \Rightarrow x_2$ – любое значение (вертикальная прямая).

Для построения целевой функции $4x_1 + 2x_2 = 8$ (значение $Z = 8$ выбрано произвольно):

$$x_1 = 0 \Rightarrow x_2 = 8/2 = 4,$$
$$x_2 = 0 \Rightarrow x_1 = 8/4 = 2.$$

Из рисунка 2.3 видно, что точкой максимума является точка ***B*** (точка пересечения прямых *2* и *3*), координаты которой можно определить как решение системы:

$$\begin{cases} 3x_1 + 2x_2 = 18, \\ x_1 = 4. \end{cases}$$

Подставляя значение $x_1 = 4$ в первое уравнение системы, получаем значение $x_2 = \dfrac{18 - 3 \cdot 4}{2} = 3$. Целевая функция принимает значение $Z = 4 \cdot 3 + 2 \cdot 4 = 20$.

Таким образом, может быть сделан следующий вывод: максимальная прибыль 20 д. е. будет получена при выпуске за смену и реализации 4 трансформаторов Т1 и 3 трансформаторов Т2.

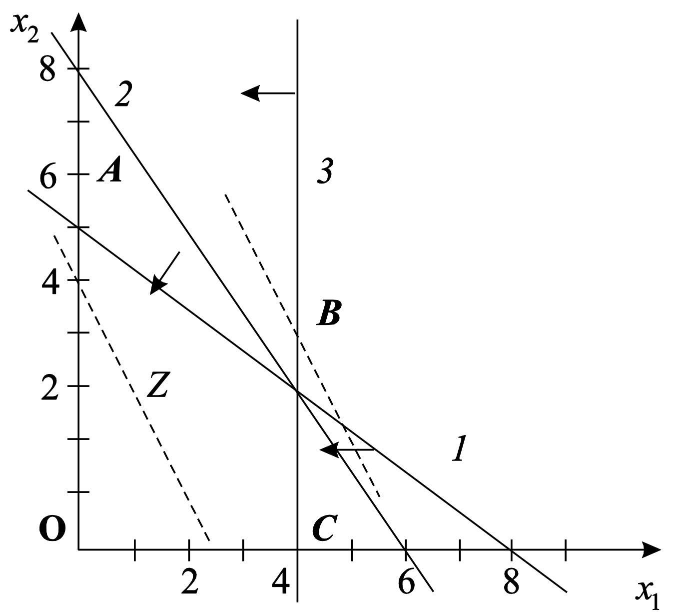
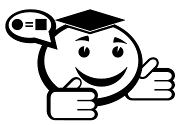


Рис. 2.3 – Графическое решение задачи

. . . .

2.2.2 Симплексный метод решения оптимизационных задач линейного программирования

. . . .



Симплексный метод *– это вычислительная процедура, основанная на принципе последовательного улучшения решений при переходе от одной базисной точки (базисного решения) к другой. При этом значение целевой функции улучшается.*

. . . .

Базисным решением является одно из допустимых решений, находящихся в вершинах области допустимых значений. Проверяя на оптимальность вершину за вершиной симплекса, приходят к искомому оптимуму. На этом принципе основан симплекс-метод [7].

Симплекс – это выпуклый многоугольник в n-мерном пространстве с $n+1$ вершинами, не лежащими в одной гиперплоскости (гиперплоскость делит пространство на два полупространства).

Доказано, что если оптимальное решение существует, то оно обязательно будет найдено через конечное число итераций (шагов), кроме случаев «зацикливания».

Алгоритм симплексного метода [8]

Первый этап. Строится исходная оптимизационная модель. Далее исходная матрица условий преобразуется в приведенную каноническую форму, которая среди всех других канонических форм выделяется тем, что:

а) правые части условий (свободные члены b_i) являются величинами неотрицательными;

б) сами условия являются равенствами;

в) матрица условий содержит полную единичную подматрицу.

Если свободные члены отрицательные, то обе части неравенства умножаются на –1, а знак неравенства меняется на противоположный. Для преобразования неравенств в равенства вводятся дополнительные переменные, которые обычно обозначают объем недоиспользованных ресурсов. В этом их экономический смысл.

Наконец, если после добавления дополнительных переменных матрица условий не содержит полную единичную подматрицу, то вводятся искусственные переменные, которые не имеют никакого экономического смысла. Они вводятся исключительно для того, чтобы получить единичную подматрицу и начать процесс решения задачи при помощи симплексного метода.

В оптимальном решении задачи все искусственные переменные (ИП) должны быть равными нулю. Для этого в целевую функцию задачи вводят искусственные переменные с большими отрицательными коэффициентами ($–M$) при решении задачи на max, и с большими положительными коэффициентами ($+M$), когда задача решается на min. В этом случае даже незначительное ненулевое значение искусственной переменной будет резко уменьшать (увеличивать) значение целевой функции. Обычно M в 1 000 раз должно быть больше, чем значения коэффициентов при основных переменных.

Второй этап. Строится исходная симплекс-таблица и отыскивается некоторое начальное базисное решение. Множество переменных, образующих единичную подматрицу, принимается за начальное базисное решение. Значения этих переменных равны свободным членам. Все остальные внебазисные переменные равны нулю.

Третий этап. Проверка базисного решения на оптимальность осуществляется при помощи специальных оценок коэффициентов целевой функции. Если все оценки коэффициентов целевой функции отрицательны или равны нулю, то имеющееся базисное решение – оптимальное. Если хотя бы одна оценка коэффициента целевой функции больше нуля, то имеющееся базисное решение не является оптимальным и должно быть улучшено.

Четвертый этап. Переход к новому базисному решению. Очевидно, что в оптимальный план должна быть введена такая переменная, которая в наибольшей степени увеличивает целевую функцию. При решении задач на максимум прибыли в оптимальный план вводится продукция, производство которой наиболее выгодно. Это определяется по максимальному положительному значению оценки коэффициента целевой функции.

Столбец симплексной таблицы с этим номером на данной итерации называется *генеральным столбцом*.

Далее, если хотя бы один элемент генерального столбца $\boldsymbol{a_{ij0}}$ строго положителен, то отыскивается генеральная строка (в противном случае задача не имеет оптимального решения).

Для отыскания генеральной строки все свободные члены (ресурсы) делятся на соответствующие элементы генерального столбца (норма расхода ресурса на единицу изделия). Из полученных результатов выбирается наименьший. Соответствующая ему строка на данной итерации называется *генеральной*. Она соответствует ресурсу, который лимитирует производство на данной итерации.

Элемент симплексной таблицы, находящийся на пересечении генеральных столбца и строки, называется *генеральным элементом*.

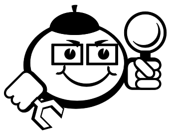
Базисной становится свободная переменная, соответствующая генеральному столбцу.

Затем все элементы генеральной строки (включая свободный член) делятся на генеральный элемент. В результате этой операции генеральный элемент становится равным единице. Далее необходимо, чтобы все другие элементы генерального столбца стали бы равны нулю, т. е. генеральный столбец должен стать единичным. Все строки (кроме генеральной) преобразуются следующим образом. Полученные элементы новой строки умножаются на соответствующий элемент генерального столбца и полученное произведение вычитается из элементов старой строки.

Значения новых базисных переменных получим в соответствующих ячейках столбца свободных членов.

Пятый этап. Полученное базисное решение проверяется на оптимальность (см. третий этап). Если оно оптимально, то вычисления прекращаются.

В противном случае необходимо найти новое базисное решение (четвертый этап) и т. д.



Пример 2.2

Пусть необходимо найти оптимальный план производства двух видов продукции (x_1 и x_2) (табл. 2.1).

Таблица 2.1 – Производство двух видов продукции

Вид продукции	**Норма расхода ресурса на единицу прибыли**		**Прибыль на единицу изделия**
	A	*B*	
1	5	8	7
2	20	4	3
Объем ресурса	20	36	–

1. Построим оптимизационную модель:
   $F(x) = 7x_1 + 3x_2 \to \max;$
   $5x_1 + 2x_2 \le 20$ – ограничение по ресурсу A;
   $8x_1 + 4x_2 \le 36$ – ограничение по ресурсу B.
2. Приведем задачу к канонической форме. Для этого достаточно ввести дополнительные переменные x_3 и x_4. В результате неравенства преобразуются в строгие равенства:
   $5x_1 + 2x_2 + x_3 = 20,$
   $8x_1 + 4x_2 + x_4 = 36.$

Построим исходную симплексную таблицу и найдем начальное базисное решение. Им будут дополнительные переменные, т. к. им соответствует единичная подматрица (табл. 2.2).

$\boldsymbol{x_3} = 20$ и $\boldsymbol{x_4} = 36.$

1-я итерация (табл. 2.3). Находим генеральный столбец и генеральную строку:

$\max(7,3) = 7;$

$$\min\left(\frac{20}{5};\frac{36}{8}\right) = \frac{20}{5}.$$

Генеральный элемент равняется 5.

Таблица 2.2 – Исходная симплекс-таблица

Базисные переменные	**Свободные члены (план)**	x_1	x_2	x_3	x_4
x_3	20	5	2	1	0
x_4	36	8	4	0	1
$F_j - C_j$	–	7	3	0	0

Таблица 2.3 – Первая итерация

Базисные переменные	**Свободные члены (план)**	x_1	x_2	x_3	x_4
x_1	4	1	0,4	0,2	0
x_4	4	0	0,8	–1,6	1
$F_j - C_j$	28	0	0,2	–1,4	0

2-я итерация (табл. 2.4). Найденное базисное решение не является оптимальным, т. к. строка оценок $\left(F_j - C_j\right)$ содержит один положительный элемент. Находим генеральный столбец и генеральную строку:

$$\max(0,\ 0,2,\ -1,4,\ 0) = 0,2;$$

$$\min\left(\frac{4}{0,4}; \frac{4}{0,8}\right) = \frac{4}{0,8}.$$

Таблица 2.4 – Вторая итерация

Базисные переменные	**Свободные члены (план)**	x_1	x_2	x_3	x_4
x_1	2	1	0	1	–0,5
x_2	5	0	1	–2	1,25
$F_j - C_j$	29	0	0	–1	–0,25

Найденное решение оптимально, так как все специальные оценки целевой функции $F_j - C_j$ равны нулю или отрицательны. $F(x) = 29$ $x_1 = 2;$ $x_2 = 5.$

. . . .

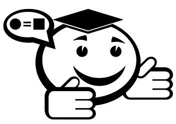
Примечания к симплекс-методу

1. Если в ведущем столбике нет ни одного строго положительного элемента, то задача не имеет оптимального решения, а целевая функция не ограничена снизу (в задаче на минимум) или не ограничена сверху (в задаче на максимум).
2. Несовместимость системы ограничений (в канонической форме) обнаруживается при построении начального допустимого базисного решения (ДБР) (оно не существует).
3. Если в последней (оптимальной) таблице оценка какой-либо небазисной переменной (число в нулевой строке) равна нулю, то задача имеет бесконечное множество оптимальных решений.
4. Симплекс-метод за конечное число итераций либо приводит к оптимальному решению, либо устанавливает неразрешимость задачи (см. п. 1, 2, 3).
5. На каждой итерации симплекс-метод сохраняет допустимость базисного решения, т. е. неотрицательность элементов нулевого столбика – следствие правила выбора ведущей строки.
6. На каждой итерации целевая функция убывает (в задаче на минимум) или возрастаем (в задаче на максимум); это свойство нарушается только в случае зацикливания (см. п. 11, 12).
7. В качестве ведущего столбика можно выбирать любой столбик с положительной оценкой (в задаче на минимум), однако максимальность оценки ведущего столбика ведет к сокращению числа итераций (целевая функция быстро убывает).
8. Слабые переменные со знаком «+» (вводимые для преобразования неравенств вида «≤») можно использовать в качестве базисных переменных, а слабые переменные со знаком «–» (вводимые для преобразования неравенств вида «≥») – нет.
9. Структуру симплекс-таблицы можно упростить, если на каждой итерации исключать из таблицы столбики для базисных переменных. При этом сокращается объем вычислений.
10. При решении симплекс-методом задачи на максимум изменяется только правило выбора ведущей строки (столбик с минимальной отрицательной оценкой) и критерий оптимальности (отсутствие в нулевой строке отрицательных оценок).

11. ДБР, в котором одна или несколько базисных переменных равны нулю, называется вырожденным ДБР. Появление такого ДБР в процессе решения может привести к зацикливанию, т. е. к повторному вхождению переменной в базис (геометрически: возвращение к предыдущей вершине многогранника). Предвестником зацикливания является неоднозначное определение ведущей строки.
12. Для выхода из зацикливания: в критерии определения ведущей строки вместо элементов 0-го столбика применяют элементы 1-го столбика; если и здесь ведущая строка неоднозначна, то применяют элементы 2-го столбика и т. д., пока ведущая строка не будет определена однозначно [9].

2.2.3 Двойственная задача линейного программирования

. . . .



Двойственная задача линейного программирования *может быть сформулирована следующим образом:*

Найти переменные y_i $(i = 1, 2, ..., m)$, *при которых целевая функция была бы минимальной:*

$$Z(y) = \sum_{i=1}^{m} b_i y_i \to \min,$$

не нарушая ограничений:

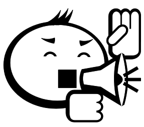
$$\sum_{i=1}^{m} a_{ij} y_i \ge c_j \quad (j = 1, 2, ..., n),$$

$$y_i \ge 0 \qquad (i = 1, 2, ..., m).$$

. . . .

Данная задача называется двойственной (симметричной) по отношению к прямой задаче, сформулированной выше. Однако правильным будет и обратное утверждение, т. к. обе задачи равноправны [10]. Переменные двойственной задачи называются *объективно обусловленными оценками*.

. . . .



Прямая и обратная задачи линейного программирования связаны между собой теоремами двойственности.

. . . .

Первая теорема двойственности. Если обе задачи имеют допустимые решения, то они имеют и оптимальное решение, причем значение целевых функций у них будет одинаково:

$$F(x) = Z(y) \text{ или } \max \sum_{j=1}^{n} c_j x_j = \min \sum_{i=1}^{m} b_i y_i.$$

Если же хотя бы одна из задач не имеет допустимого решения, то ни одна из них не имеет оптимального решения.

Вторая теорема двойственности (теорема о дополняющей нежесткости). Для того чтобы векторы $\overline{X} = (x_1, x_2, ..., x_n)$ Ц $\overline{Y} = (y_1, y_2, ..., y_m)$ были оптимальными решениями соответственно прямой и двойственной задачи, необходимо и достаточно, чтобы выполнялись следующие условия:

$$y_i \left(b_i - \sum_{j=1}^{n} a_{ij} x_j \right) = 0 \quad (i = 1, 2, ..., m). \tag{2.1}$$

$$x_j \left(\sum_{i=1}^{m} a_{ij} y_j - c_j \right) = 0 \quad (j = 1, 2, ..., n). \tag{2.2}$$

Следствие 1. Пусть оптимальное значение некоторой переменной двойственной задачи строго положительно:

$$y_i > 0.$$

Тогда из условия (2.1) получим:

$$b_i - \sum_{j=1}^{n} a_{ij} x_j = 0$$

или

$$\sum_{j=1}^{n} a_{ij} x_j = b_i.$$

Экономический смысл данных выражений можно интерпретировать в следующей редакции. Если объективно обусловленная оценка некоторого ресурса больше нуля (строго положительна), то этот ресурс полностью (без остатка) расходуется в процессе выполнения оптимального плана.

Следствие 2. Пусть для оптимального значения некоторой переменной $\boldsymbol{x_i}$ прямой задачи выполняется условие строгого неравенства:

$$\sum_{j=1}^{n} a_{ij} x_j < b_i.$$

Тогда, основываясь на том же первом условии (2.1), можно заключить, что $\boldsymbol{y_i} = 0$.

Экономически это означает, что если в оптимальном плане какой-то ресурс используется не полностью, то его объективно обусловленная оценка обязательно равна нулю.

2.2.4 Транспортная задача

Транспортная модель *используется при разработке плана перевозок одного вида продукции из нескольких пунктов отправления в пункты назначения.*

При построении модели используются:

- величины, характеризующие предложение в каждом исходном пункте и спрос в каждом пункте назначения;
- стоимость перевозки единицы продукции из каждого исходного пункта в каждый пункт назначения.

Поскольку рассматривается только один вид продукции, потребности пункта назначения могут удовлетворяться за счет нескольких исходных пунктов. Цель построения модели состоит в определении количества продукции, которое следует перевезти из каждого исходного пункта в каждый пункт назначения, с тем чтобы общие транспортные расходы были минимальными [5].

Решается транспортная задача с помощью метода потенциалов. При решении транспортных задач прежде всего проверяется условие равенства ресурсов поставщиков потребностям потребителей. Если это условие не выполняется, вводится фиктивный поставщик или потребитель. Фиктивные объемы ресурсов или потребностей при этом включаются в задачу с нулевыми оценками.

Затем заполняется расчетная таблица и составляется первый опорный план, который может быть получен несколькими способами. Более близкий к оптимальному опорный план может быть получен методом минимальных в таблице затрат. При этом способе составление плана начинается с клетки с минимальной оценкой при решении задачи на минимум или с максимальной оценкой при решении на максимум. Если в таблице имеется несколько клеток с одинаковыми «лучшими» оценками, то заполняется прежде клетка, в которую можно записать наибольшую поставку.

После составления первого опорного плана с помощью алгоритма метода потенциалов производится проверка его на оптимальность, если план не оптимальный, то осуществляется его улучшение.

Алгоритм метода потенциалов (решение задачи на минимум)

1. Для всех заполненных клеток рассчитываются потенциалы по формуле:

$$u_i + v_j = c_{ij}, \tag{2.3}$$

   где u_i – потенциалы строк; v_j – потенциалы столбцов; c_{ij} – оценки.

   Для расчета потенциалов одному из них вначале придают любое значение. Обычно $u_i = 0$.

2. Для всех свободных клеток рассчитываются характеристики по формуле:

$$d_{ij} = c_{ij} - (u_i + v_j).$$

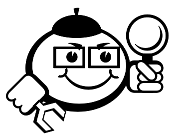
   Если в таблице нет ни одной свободной клетки с отрицательной характеристикой, то план считается оптимальным (при решении задачи на максимум – план оптимальный, если нет положительных характеристик).

3. Среди отрицательных характеристик (при решении на максимум среди положительных) выбирается максимальная по абсолютной величине и для клетки с этой характеристикой строится цепь. Для этого из выбранной свободной клетки проводится по строке или столбцу прямая линия до занятой клетки, затем под углом 90° линия проводится до следующей занятой клетки и так до тех пор, пока цепь не замкнется в исходной клетке.

4. В вершинах цепи, чередуя, проставляются знаки «+» и «–», начиная со свободной клетки. В клетках со знаком «–» выбирается минимальная поставка, которая перераспределяется по цепи: там, где стоит знак «+», она прибавляется, а где «–» – вычитается. Исходная свободная клетка становится занятой, а клетка, в которой выбрана минимальная поставка, свободной.

   Составляется новый план и рассчитывается значение целевой функции.

5. Переход к 1.



Пример 2.3

С пяти асфальтобетонных заводов в течение квартала должен вывозиться асфальт для строительства четырех участков автодорог области. Транспортные издержки при перевозках (д. е.), заказы дорожно-строительных бригад и возможности заводов в совершении машинами рейсов представлены в таблице 2.5.

Таблица 2.5 – Затраты на транспортировку

Участки дорог / **Заводы**	***A***	***B***	***C***	***D***	**Количество рейсов, ед.**
АБЗ-1	5	6	2	2	600
АБЗ-2	9	7	4	6	240
АБЗ-3	7	1	4	5	1360
АБЗ-4	5	2	2	4	1000
АБЗ-5	6	4	3	4	800
Потребность в машинах, ед.	600	800	1 400	1 200	–

Требуется составить такой план перевозки асфальта, при котором транспортные издержки были бы минимальными.

Заполним расчетную таблицу 2.6 и составим первый опорный план методом минимальных затрат, выбирая на каждом шаге наименее затратный маршрут и максимально возможную поставку по этому маршруту. Заполнение таблицы начинается с клетки (3,2) с наименьшими издержками, в которую записывается поставка 800. Затем последовательно заполняются клетки (4,3); (1,4); (5,3); (5,4); (3,4); (3,1); (2,1).

Таблица 2.6 – Итерация 1

Заводы	**Участки дорог**				**Предл.**	u_i
	A	***B***	***C***	***D***		
АБЗ-1	5	6	2	2 600	600	0
АБЗ-2	9 – 240	7	4 +	6	240	5
АБЗ-3	7 + 360	1 800	4	5 – 200	1 360	3
АБЗ-4	5	2	2 1 000	4	1 000	1

Заводы	Участки дорог				Предл.	u_i
	A	*B*	*C*	*D*		
АБЗ-5	6	4	3 – 400	4 + 400	800	2
Спрос	600	800	1 400	1 200	4 000	–
v_j	4	–2	1	2	–	Z = 12 480

Условие, что количество заполненных клеток в таблице должно быть равно $m+n-1=5+4-1=8$, выполняется.

Переходим к анализу первого опорного плана. Рассчитываем значение целевой функции:

$$Z = 600\cdot 2+240\cdot 9+360\cdot 7+800\cdot 1+200\cdot 5+1000\cdot 2+$$
$$+400\cdot 3+400\cdot 4=12\,480 \text{ (д. е.)}.$$

Проверим, является ли план оптимальным, если нет, то улучшаем его.

1. Рассчитываем значения потенциалов по формуле (2.3):
   $u_1=0;\ v_4=2-0=2;\ u_3=5-2=3;\ u_5=4-2=2;\ v_1=7-3=4;$
   $v_2=1-3=-2;\ v_3=3-2=1;\ u_2=9-4=5;\ u_4=2-1=1.$
2. Рассчитываем характеристики для свободных клеток:
   $d_{11}=5-(0+4)=1;\ d_{12}=6-(0+(-2))=8;\ d_{13}=2-(0+1)=1;$
   $d_{22}=7-(5+(-2))=4;\ d_{23}=4-(5+1)=-2;$
   $d_{24}=6-(5+2)=-1;\ d_{33}=4-(3+1)=0;\ d_{41}=5-(1+4)=0;$
   $d_{42}=2-(1+(-2))=3;\ d_{44}=4-(1+2)=1;$
   $d_{51}=6-(2+4)=0;\ d_{52}=4-(2+(-2))=4.$
3. Максимальная по абсолютной величине отрицательная характеристика находится в клетке (2,3), для нее строим цепь.
4. Проставляем по углам цепи, начиная с выбранной клетки, знаки «+», «–». В клетках со знаком «–» минимальная поставка 200. Ее перераспределяем по цепи. Там, где стоит знак «+», прибавляем, а где «–», вычитаем. Заполняем следующую расчетную таблицу 2.7.

Расчеты ведем аналогично. Получены следующие отрицательные характеристики: $d_{11}=-1;\ d_{41}=-2;\ d_{51}=-2.$

Для клетки (4,1) строим цепь. Перераспределяем по цепи поставку 40. Заполняем таблицу для итерации 3.

В расчетной таблице 2.8 нет отрицательных характеристик, следовательно, план оптимальный.

Таблица 2.7 – Итерация 2

Заводы	**Участки дорог**				**Предл.**	u_i
	A	***B***	***C***	***D***		
АБЗ-1	5	6	2	2 600	600	0
АБЗ-2	9 – 40	7	4 + 200	6	240	3
АБЗ-3	7 560	1 800	4	5 200	1 360	1
АБЗ-4	5 +	2	2 – 1 000	4	1 000	1
АБЗ-5	6	4	3 200	4 600	800	2
Спрос	600	800	1 400	1 200	4 000	–
v_j	6	0	1	2	–	Z = 12 080

Таблица 2.8 – Итерация 3

Заводы	**Участки дорог**				**Предл.**	u_i
	A	***B***	***C***	***D***		
АБЗ-1	5	6	2	2 600	600	0
АБЗ-2	9	7	4 240	6	240	3
АБЗ-3	7 560	1 800	4	5 200	1 360	3
АБЗ-4	5 40	2	2 960	4	1 000	1
АБЗ-5	6	4	3 200	4 600	800	2
Спрос	600	800	1 400	1 200	4 000	–
v_j	4	–2	1	2	–	Z = 12 000

Согласно оптимальному плану, необходимо за квартал с АБЗ-1 направить на участок *D* 600 машин асфальта, с АБЗ-2 к участку *C* – 240 машин, с АБЗ-3 к участку *A* – 560 машин и к участку *B* – 800 машин, с АБЗ-4 к участку *A* – 40 машин и к участку *C* – 960 машин, с АБЗ-5 к участку *C* – 200 машин и к участку *C* – 600 машин. При этом минимальные затраты на перевозку составят 12 000 д. е.

Контрольные вопросы по главе 2

1. Что такое оптимизационная модель?
2. Что такое область допустимых решений?
3. Какие существуют ограничения в моделях?
4. Для каких задач применяется графический метод?
5. Опишите алгоритм симплекс-метода.
6. Сформулируйте двойственную задачу линейного программирования.
7. Для чего используется транспортная задача?

3 Нелинейное программирование

В этом разделе мы рассмотрим задачи ***нелинейной оптимизации*** (называемые иначе ***оптимизационными задачами нелинейного программирования***), математические модели которых содержат ***нелинейные*** зависимости от переменных. Источники нелинейности в задачах подобного типа могут относиться, в частности, к одной из двух категорий [11]:

1. Реально существующие и эмпирически наблюдаемые нелинейные соотношения, например непропорциональные зависимости между объемом производства и затратами, между количеством используемого в производстве компонента и некоторыми показателями качества готовой продукции, между затратами сырья и физическими параметрами (давление, температура и т. п.) соответствующего производственного процесса, между выручкой и объемом реализации и т. п.

2. Установленные (постулируемые) руководством правила поведения или задаваемые зависимости, например правила расчета с потребителями энергии или других видов услуг, правила определения страховых уровней запаса продукции, гипотезы о характере вероятностного распределения рассматриваемых в модели случайных величин, различного рода договорные условия взаимодействия между партнерами по бизнесу и др.

В отличие от задачи линейной оптимизации (линейного программирования), не существует одного или нескольких алгоритмов, эффективных для решения любых нелинейных задач. Какой-то алгоритм может быть эффективен при решении задач одного типа и неприемлемым для задач другого типа. В связи с этим разработаны алгоритмы для решения каждого класса (типа) задач. Следует иметь в виду, что даже программы, ориентированные на решение определенного класса задач, не гарантируют правильность решения любых задач этого класса и оптимальность решения следует проверять в каждом конкретном случае.

3.1 Специфика нелинейных программ и методы их решения

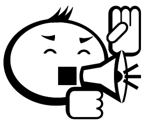
Многообразие методов решения линейных программ имеет в своей основе идею упорядоченного перебора опорных планов (вершин) исходной или сопряженной задачи. Для нелинейных же программ простого метода решения, подобного симплексному, нет по многим причинам [12].

Во-первых, множество планов может оказаться невыпуклым или иметь бесконечное количество «вершин».

Во-вторых, искомые экстремумы могут достигаться как на границе множества планов, так и внутри него.

В-третьих, в нелинейных программах возникает проблема поиска глобального экстремума среди множества локальных.

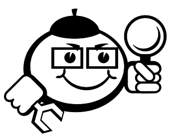
Как мы показали ранее, ни использование аппарата производных, ни прямое табулирование целевой функции над множеством планов не решают проблему в случае более трех переменных. Поэтому каждая нелинейная программа требует индивидуального подхода, учитывающего ее специфику.



Существующие методы нелинейного программирования можно подразделить на следующие основные классы.

1. ***Градиентные методы***, в основе которых лежит свойство градиента функции в точке (вектора частных производных, вычисленного в точке) как указателя направления наибольшего роста функции в окрестности точки.

Так, при отсутствии ограничений одна из простейших разновидностей градиентных методов – *метод наискорейшего спуска* – предлагает выбрать некоторую точку (план) X_k и начальный шаг H_k, вычислить градиент функции в выбранной точке $\operatorname{grad} F(X_k)$ и осуществить переход по градиенту (при максимизации) с выбранным шагом. Если значение функции в новой точке больше предыдущего, новая точка принимается за исходную и повторяется такая же процедура. При попадании в точку с меньшим значением уменьшается шаг (например, вдвое) и переход повторяется от исходной точки. Переходы продолжаются до достаточно малого шага.



Пример 3.1

Например, если нужно решить систему уравнений:

$$\left(X^2+1\right)(Y-1)=3,\ Y=\ln(X+1),$$

то можно заменить эту задачу задачей минимизации функции:

$$F(X,Y)=\left[\left(X^2+1\right)(Y-2)-3\right]^2+\left[Y-\ln(X+1)\right]^2$$

(если система имеет решение, то искомый минимум равен нулю).

Градиент этой функции определяется вектором:

$$\text{grad } F(X,Y)=\Big\{2\cdot\left[\left(X^2+1\right)(Y-2)-3\right]\cdot 2X\cdot(Y-2)-2\cdot\left[Y-\ln(X+1)\right]/(X+1),$$

$$2\cdot\left[\left(X^2+1\right)(Y-2)-3\right]\cdot\left(X^2+1\right)-2\cdot\left[Y-\ln(X+1)\right]\Big\}.$$

Выбираем начальную точку $M_0(2{,}1)$ и шаг $h=1$. Здесь значение функции $F(M_0)\approx 64,$ градиент в этой точке $\text{grad}\, F(M_0)=\left[64{,}066,\ -80{,}197\right]$, нормированный градиент (вектор единичной длины, составленный из компонент, деленных на корень из суммы их квадратов) $\text{grad}_{\text{н}}\ F(M_0)=\left[0{,}62,\ -0{,}78\right]$.

Смещаемся в направлении, обратном градиенту (ищем минимум), с выбранным шагом в точку $M_1=M_0-h\cdot\text{grad}_{\text{н}}\ F(M_0)=\left(2-1\cdot 0{,}62,\ 1+1\cdot 0{,}78\right)=$ $=\left(1{,}38,\ 1{,}78\right)$ и обнаруживаем, что $F(M_1)\approx 14<F(M_0)$.

Аналогичный переход с учетом $\text{grad}_{\text{н}}\ F(M_1)=\left[0{,}19,\ -0{,}98\right]$ приводит в точку $M_2(1{,}19,\ 2{,}76)$, где $F(M_2)\approx 5{,}26,$ $\text{grad}_{\text{н}}\ F(M_2)=\left[-0{,}96,\ -0{,}27\right]$. Переход в очередную точку $M_3(2{,}15,\ 3{,}03)$ дает $F(M_3)\approx 11{,}33>F(M_2)$. Соответственно уменьшаем шаг вдвое $(h=0{,}5)$ и получаем точку $M_4(2{,}15,\ 3{,}03)$, где $F(M_4)\approx 3{,}78,$ $\text{grad}_{\text{н}}\ F(M_2)=\left[0{,}12,\ 0{,}99\right]$. Очередной переход приводит в точку с большим значением функции и приходится еще уменьшать шаг и т. д.

. . . .

Есть и более эффективные переходы по градиенту, связанные с выбором различного шага по разным координатам или с автоматическим определением шага (при каждом переходе решается задача поиска экстремума функции в заданном направлении). Однако гарантии нахождения глобального экстремума нет (при разных начальных точках или шагах можно получить разные решения для многоэкстремальных функций).

Градиентные методы для задач с ограничениями, где при смещениях по градиенту приходится сталкиваться с опасностью «выскочить» за пределы допустимого множества решений, существенно усложняются (модифицированный метод Ньютона, методы возможных направлений Зойтендейка, сопряженных градиентов, проектируемых градиентов Розена и др.) [13].

Существует обширная литература по численному анализу, где значительное внимание уделяется градиентным и другим итерационным методам, но тем не менее решение нелинейных задач оптимизации при наличии ограничений иногда весьма затруднительно.

2. ***Методы Монте-Карло***. Здесь отыскивается n-мерный параллелепипед, включающий в себя множество планов, и затем моделируются N случайных точек с равномерным законом распределения в параллелепипеде (практически во всех программных средах предусмотрено наличие соответствующих датчиков псевдослучайных чисел).

В точках, попавших во множество планов, вычисляются значения функции и запоминается точка текущего экстремума. После этого берется параллелепипед меньших размеров с центром в найденной точке, и в нем вновь моделируются N случайных точек. Процесс такого *стохастического моделирования* заканчивается при малых размерах параллелепипеда. Методы Монте-Карло имеют преимущество над моделированием на детерминированной сетке, так как их точность имеет порядок $1/\sqrt{N}$ и не зависит от размерности задачи. Естественно, этими методами никто не пользуется при ручном счете; они просты для программной реализации и обычно используются при поиске начального приближения для градиентных методов.

3. ***Методы динамического программирования***, сводящие многомерную задачу оптимизации к последовательности задач меньшей размерности.

4. ***Методы выпуклого программирования***, реализующие поиск минимума выпуклой функции или максимума вогнутой на выпуклом множестве планов. Если множество планов – выпуклый многогранник, то эти методы допускают использование симплексного метода [14].

Наиболее эффективно эти и другие методы решения действуют для так называемых *сепарабельных функций*, т. е. функций, представимых в виде суммы функций одной переменной:

$$F(X_1, X_2, ..., X_n) = f_1(X_1) + f_2(X_2) + ... + f_n(X_n).$$

При решении многих задач нелинейного программирования определенный эффект дает *метод множителей Лагранжа*.

Пусть требуется найти экстремумы функции $F(X)$ при условиях $f_i(X) = 0$ $(i = 1...m)$.

Функция $\Phi(X,\lambda)=F(X)+\sum_{i=1}^{m}\lambda_i f_i(X)$ называется *функцией Лагранжа* и коэффициенты λ_i – *множителями Лагранжа.*

Можно доказать, что *необходимым условием экстремума исходной задачи является обращение в нуль всех частных производных функции Лагранжа.*

. . . . Пример 3.2 . . . .

Например, при поиске максимума $F(X_1,X_2)=X_1+X_2$ при единственном условии $X_1^2+X_2^2=1$ строится функция Лагранжа:

$$\Phi(X,\lambda)=X_1+X_2+\lambda\left(X_1^2+X_2^2-1\right).$$

Строим систему уравнений:

$$\frac{\partial\Phi}{\partial X_1}=1+2\lambda X_1=0,\ \frac{\partial\Phi}{\partial X_2}=1+2\lambda X_2=0,\ \frac{\partial\Phi}{\partial\lambda}=X_1^2+X_2^2-1=0,$$

решение которой дает $\lambda=\pm1/\sqrt{2}$, $X_1=X_2=-\lambda$ и экстремальные значения целевой функции $-\sqrt{2}$ и $\sqrt{2}$.

Для определения типа найденного экстремума можно построить матрицу вторых производных $F(X)$, вычисленных в экстремальной точке, и определить знаки главных ее миноров. Если все они положительны, то найден минимум; если они чередуются, начиная с минуса, то найден максимум (*правило Сильвестра*).

. . . .

Сам по себе метод множителей Лагранжа не дает существенного эффекта из-за необходимости решать, как правило, нелинейную систему уравнений; не гарантирует тип отыскиваемого экстремума, кроме глобальных дает и множество локальных экстремумов, но полезен при генерации идей и создании методов нелинейного программирования.

3.2 Теорема Куна – Таккера

Пусть стоит задача минимизации $F(X)$ при условиях:

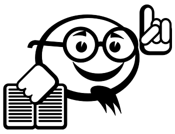
$$f_i(X)\le 0\ \ (i=1...m),$$

$$X\ge 0, \tag{3.1}$$

где X – n-мерный вектор, $F(X)$ и $f_i(X)$ – выпуклые функции, что обеспечивает выпуклость множества планов и единственность искомого минимума (напомним, что мы считаем функцию выпуклой в некоторой точке, если главные миноры матрицы вторых производных положительны) [15].

Введем функцию Лагранжа:

$$\Phi(X,\lambda)=F(X)+\sum_{i=1}^{m}\lambda_i f_i(X).$$



Теорема Куна – Таккера утверждает, что *вектор* $X^*\geq 0$ *является решением поставленной задачи тогда и только тогда, когда существует вектор* $\lambda^*\geq 0$, *такой, что* при всех $X\geq 0,\ \lambda\geq 0$

$$\Phi(X^*,\lambda)\leq\Phi(X^*,\lambda^*)\leq\Phi(X,\lambda^*).$$

Так как функция Лагранжа в точке $\left(X^*,\lambda^*\right)$ принимает минимум по X и максимум по λ, то эта точка называется седловой и эта теорема называется *теоремой о седловой точке* или *теоремой о минимаксе*.

Достаточность условий этой теоремы доказывается сравнительно просто. Доказательство их необходимости предполагает выполнение *условий регулярности,* т. е. существования хотя бы одной допустимой точки X, где $f_i(X)<0$ при всех i.

Если $F(X)$ и $f_i(X)$ дифференцируемы, то условия теоремы эквивалентны «локальным» условиям, утверждающим, что в точке $\left(X^*,\lambda^*\right)$

$$\frac{\partial\Phi}{\partial X}\geq 0,\ \ X^T\frac{\partial\Phi}{\partial X}=0,\ \ X\geq 0, \tag{3.2}$$

$$\frac{\partial\Phi}{\partial\lambda}\leq 0,\ \ \lambda^T\frac{\partial\Phi}{\partial\lambda}=0,\ \ \lambda\geq 0. \tag{3.3}$$

Остановимся на частных случаях задачи.

Если в поставленной задаче отсутствуют требования неотрицательности (3.1), то путем замены X разностью двух неотрицательных векторов можно показать, что условие (3.2) упрощается к виду $\frac{\partial\Phi}{\partial X}=0$ (если нет требования неотрицательности для одной из компонент X, то обратится в нуль производная $\Phi(X)$ по соответствующей переменной).

Если $f_i(X)=0$ при некотором i, то заменой этого равенства системой неравенств $f_i(X)\leq 0,\ \ -f_i(X)\leq 0$ обнаруживаем, что в (3.3) соответствующая производная обращается в нуль и исчезает условие неотрицательности по λ_i; если $f_i(X)=0$ при всех i, то (3.3) упрощается к виду $\frac{\partial \Phi}{\partial \lambda}=0.$

Как частный случай условий Куна – Таккера могут быть построены двойственные задачи линейного программирования.

Так, при минимизации $C^T X$ при условиях $AX \ge B,\ X \ge 0$ функция Лагранжа имеет вид:

$$\Phi(X,\lambda)=C^T X+\lambda^T(AX-B)-$$

и условия Куна – Таккера:

$$\frac{\partial \Phi}{\partial X}=C-A^T\lambda \ge 0,\ \ X^T\frac{\partial \Phi}{\partial X}=X^T\left(C-A^T\lambda\right)=0,\ \ X\ge 0,$$

$$\frac{\partial \Phi}{\partial \lambda}=B-AX\le 0,\ \ \lambda^T\frac{\partial \Phi}{\partial \lambda}=\lambda^T\left(B-AX\right)=0,\ \ \lambda \ge 0.$$

Отсюда получаем с учетом $X^T C=C^T X,\ \lambda^T B=B^T\lambda,\ X^T A^T\lambda=\lambda^T AX$, что в седловой точке достигается минимум по X и максимум по λ, причем

$$C^T X=B^T\lambda,$$

$$AX\ge B,\ A^T\lambda\le C,$$

$$X\ge 0,\ \lambda\ge 0.$$

3.3 Квадратичное программирование. Метод Вулфа – Фрэнка

Рассмотрим задачу минимизации квадратичной функции n переменных:

$$F(X)=C^T X+X^T DX, \tag{3.4}$$

при линейных ограничениях

$$AX\le B, X\ge 0,$$

где A – матрица размерности m на n, $C,\ X$ – n-мерные векторы, B – m-мерный вектор, D – положительно-определенная n-мерная квадратная матрица (матрица называется положительно определенной, если положительны ее главные миноры) [16].

Так, целевая функция

$$F(X)=7X_1-3X_2+22X_1^2-8X_1X_2+X_2^2-10X_1X_3+12X_2X_3+40X_3^2$$

представится в виде (3.4), где

$$C=\begin{Vmatrix}7\\-3\\0\end{Vmatrix},\ \ D=\begin{Vmatrix}22 & -4 & -5\\-4 & 1 & 6\\-5 & 6 & 40\end{Vmatrix}.$$

Положительная определенность матрицы D и линейность ограничений (множество планов выпукло) позволяют использовать теорему Куна – Таккера.

Функция Лагранжа здесь имеет вид:

$$\Phi(X,\lambda) = C^T X + X^T DX + \lambda^T (AX - B) -$$

и условия Куна – Таккера приводятся к форме:

$$\frac{\partial \Phi}{\partial X} = C + 2DX + A^T\lambda \geq 0, \;\; X^T \frac{\partial \Phi}{\partial X} = 0, \;\; X \geq 0, \tag{3.5}$$

$$\frac{\partial \Phi}{\partial \lambda} = AX - B \leq 0, \;\; \lambda^T \frac{\partial \Phi}{\partial \lambda} = 0, \;\; \lambda \geq 0. \tag{3.6}$$

Если ввести векторы ослабляющих переменных, то (3.5)–(3.6) примут вид:

$$C + 2DX + A^T\lambda - V = 0, \;\; X^T V = 0, \;\; X \geq 0, \;\; V \geq 0;$$

$$AX - B + Y = 0, \;\; -\lambda^T Y = 0, \;\; \lambda \geq 0, \;\; Y \geq 0.$$

С учетом неотрицательности переменных можно поставить задачу *минимизации (до нуля)* функции

$$g(X,\lambda,Y,V) = X^T V + \lambda^T Y$$

при условиях:

$$2DX + A^T\lambda - V = -C, \tag{3.7}$$

$$AX + Y = B, \tag{3.8}$$

$$X,\lambda,Y,V \geq 0. \tag{3.9}$$

Если обозначить

$$W^T = \left(X^T, \lambda^T, Y^T, V^T\right),$$

то (3.7)–(3.9) приведутся к виду:

$$RW = S, \;\; W \geq 0,$$

где

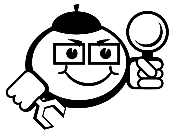
$$R = \begin{Vmatrix} 2D & A^T & 0 & -E \\ A & 0 & E & 0 \end{Vmatrix}, \;\; S = \begin{Vmatrix} -C \\ B \end{Vmatrix}.$$

Метод П. Вулфа и М. Фрэнка сводит решение задачи к форме, допускающей применение симплексной процедуры.

Здесь отыскивается некоторый опорный план W^0. Если $g\left(W^0\right) = 0$, то этот план оптимален. В противном случае отыскивается градиент

$$\operatorname{grad} g\left(W^0\right) = \left(V^T, Y^T, \lambda^T, X^T\right)$$

и его компоненты на один шаг симплексного преобразования (перехода к другому опорному плану) принимаются за коэффициенты «линейной формы». После выбора нового опорного плана выполняются вышеописанные рассуждения.



Пример 3.3

Рассмотрим простой пример.

Минимизировать

$$F(X_1, X_2) = -4X_1 - 6X_2 + X_1^2 + 3X_2^2$$

при условиях:

$$X_1 + 2X_2 \le 4,$$

$$X_1, X_2 \ge 0.$$

Ставим задачу минимизации до нуля функции

$$g(X_1, X_2, \lambda, Y, V_1, V_2) = X_1V_1 + X_2V_2 + \lambda Y$$

при условиях (3.8), где

$$R = \begin{array}{|cc|c|c|cc|} \hline 2 & 0 & 1 & 0 & -1 & 0 \\ 0 & 6 & 2 & 0 & 0 & -1 \\ \hline 1 & 2 & 0 & 1 & 0 & 0 \\ \hline \end{array} \qquad S = \begin{array}{|c|} \hline 4 \\ 6 \\ \hline 4 \\ \hline \end{array}$$

Ищем начальный опорный план методом искусственного базиса (не обращая внимания на коэффициенты, меньшие М) (табл. 3.1):

Таблица 3.1 – Начальный опорный план

C баз	Базис	План W							М	М
			X_1	X_2	λ	Y	V_1	V_2	Z_1	Z_2
М	Z_1	4	2		1	0	–1	0	1	0
М	Z_2	6	0	6	2	0	0	–1	0	1
	Y	4	1	2	0	1	0	0	0	0
Δk		10М	2М	6М	3М	0	–М	–М	0	0

C баз	Базис	План W							М
			X_1	X_2	λ	Y	V_1	V_2	Z_1
М	Z_1	1	2	–3	0	0	–1	1/2	1
	λ	3	0	3	1	0	0	–1/2	0
	Y	4	1	2	0	1	0	0	0
Δk		М	2М	–3М	0	0	–М	М/2	0

Находим начальный опорный план исходной задачи:

$$X_1 = X_2 = V_1 = 0, \;\; V_2 = 2, \;\; Y = 4, \;\; \lambda = 4,$$

для которого $g(W) = 0 \cdot 0 + 0 \cdot 2 + 4 \cdot 4 = 16 > 0$.

Отыскиваем градиент $\operatorname{grad} g(W) = \{0, 2, 4, 4, 0, 0\}$ и берем его компоненты в качестве коэффициентов «линейной формы» (нормали к поверхности, определяемой функцией $g(W)$, в выбранной точке) (табл. 3.2):

Таблица 3.2 – Поиск решения

C баз	Базис	План W	0	2	4	4	0	0
			X_1	X_2	λ	Y	V_1	V_2
0	V_2	2	4	–6	0	0	–2	1
4	λ	4	2	0	1	0	–1	0
4	Y	4	1	2	0	1	0	0
Δk		32	12	6	0	0	–4	0
C баз	Базис	План W	0	0	7/2	3	1/2	0
			X_1	X_2	λ	Y	V_1	V_2
0	X_1	1/2	1	–3/2	0	0	–1/2	1/4
7/2	λ	3	0	3	1	0	0	–1/2
3	Y	7/2	0	7/2	0	1	1/2	–1/4
Δk		21	0	21	0	0	1	–5/2
C баз	Базис	План W	0	0	0	0	2	1
			X_1	X_2	λ	Y	V_1	V_2
0	X_1	**2**	1	0	1/2	0	–1/2	0
0	X_2	**1**	0	1	1/3	0	0	–1/6
0	Y	0	0	0	–7/6	1	1/2	1/3
Δk		**0**						

Получен оптимальный план $X = (2,1)$.

Заметим, что при ограничениях $AX = B$, $X \geq 0$ исчезнет переменная Y, снимутся условия неотрицательности на λ и потребуется минимизировать $X^T V$ при условиях:

$$2DX + A^T\lambda - V = -C,$$
$$AX = B,$$
$$X, V \geq 0.$$

При других отклонениях постановки задачи от выбранного здесь стандарта можно элементарными приемами прийти к нему (вместо максимизации $F(X)$ искать минимум $-F(X)$, условие $AX \geq B$ заменить на $B - AX \leq 0$ и т. п.).

. . . .

3.4 Дробно-линейное программирование

Пусть стоит задача максимизации дробно-линейной функции

$$Q(X)=\frac{C_0+\sum_{j=1}^{n}C_jX_j}{D_0+\sum_{j=1}^{n}D_jX_j}$$

при линейных ограничениях:

$$\sum_{j=1}^{n}A_{ij}X_j=B_i \quad (i=1...m). \tag{3.10}$$

$$X_j\ge 0 \quad (j=1...n). \tag{3.11}$$

Предположим, что знаменатель в (2.1) положителен при всех X, удовлетворяющих (3.10)–(3.11).

Если обозначить

$$D_0+\sum_{j=1}^{n}D_jX_j=\frac{1}{R},\ \ R>0, \tag{3.12}$$

$$Z_j=RX_j,$$

то задача сведется к линейной программе максимизации

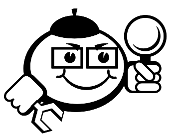
$$Q(Z,R)=C_0R+\sum_{j=1}^{n}C_jZ_j$$

при условиях:

$$D_0R+\sum_{j=1}^{n}D_jZ_j=1,$$

$$-B_iR+\sum_{j=1}^{n}A_{ij}Z_j=0, \quad (i=1...m),$$

$$R>0,\ \ Z_j\ge 0, \quad (j=1...n).$$



Пример 3.4

Рассмотрим в качестве примера задачу максимизации

$$\frac{-3+2X_1+4X_2-5X_3}{5+3X_1-X_2}$$

при условиях:

$$X_1-X_2\ge 0,$$

$$5X_1+3X_2+10X_3\le 15,$$

$$X_1,X_2,X_3\ge 0.$$

С учетом (3.12) получаем задачу максимизации

$$-3R + 2Z_1 + 4Z_2 - 5Z_3$$

при условиях:

$$5R + 3Z_1 - Z_2 = 1,$$

$$Z_1 - Z_2 \geq 0,$$

$$-15R + 5Z_1 + 3Z_2 + 10Z_3 \leq 0,$$

$$R > 0,\ \ Z_1, Z_2, Z_3 \geq 0.$$

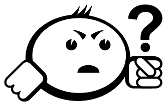
Таблица 3.3 – Поиск решения

C баз	Базис плана	План Z	–3	2	4	–5	0	0	–M	–M
			R	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7
–M	Z_6	1	5	3	–1	0	0	0	1	0
–M	Z_7	0	0	1	–1	0	–1	0	0	1
0	Z_5	0	–15	5	3	10	0	1	0	0
Δk		–M	–5M	–3M	2M	5	M	0	0	0
C баз	Базис плана	План Z	–3	2	4	–5	0	0		–M
			R	Z_1	Z_2	Z_3	Z_4	Z_5		Z_7
–3	R	1/5	1	3/5	–1/5	0	0	0		0
–M	Z_7	0	0	1	–1	0	–1	0		1
0	Z_5	3	0	14	0	10	0	1		0
Δk		–3/5	0	–M	M+	5	M	0		0
C баз	Базис плана	План Z	–3	2	4	–5	0	0		
			R	Z_1	Z_2	Z_3	Z_4	Z_5		
–3	R	1/5	1	0	2/5	0	3/5	0		
2	Z_1	0	0	1	–1	0	–1	0		
0	Z_5	3	0	0	14	10	14	1		
Δk		–3/5	0	0	–7.2	5	–3.8	0		
C баз	Базис плана	План Z	–3	2	4	–5	0	0		
			R	Z_1	Z_2	Z_3	Z_4	Z_5		
–3	R	4/35	1	0	0	–2/7	1/5	–1/35		
2	Z_1	3/14	0	1	0	5/7	0	1/14		
4	Z_2	3/14	0	0	1	5/7	1	1/14		
Δk		33/35	0	0	0	1/7	17/5	13/35		

Из таблицы 3.3 имеем $Z_1 = Z_2 = 3/14,\ \ Z_3 = 0,\ \ R = 4/35$ и соответственно

$$X_{opt} = (15/8, 15/8, 0),\ \ \max Q(X) = 33/35.$$

Заметим, что дробно-линейные программы в приложениях возникают достаточно часто; например, некорректная задача, преследующая взаимно противоположные цели типа максимизации прибыли при минимальных капиталовложениях, сводится к максимизации прибыли на единицу затрат.



Контрольные вопросы по главе 3

1. Почему в экономико-математических задачах может возникнуть нелинейность?
2. Какие существуют методы решения задач нелинейной оптимизации (нелинейного программирования)?
3. Что утверждает теорема Куна – Таккера?
4. В чем суть квадратичного программирования?
5. В чем суть дробно-линейного программирования?

4 Элементы теории игр и статистических решений

4.1 Основные понятия теории игр

***Теория игр** занимается изучением так называемых конфликтных ситуаций, где сталкиваются интересы индивидов, партий, государств и т. п.*

Как утверждал Г. Лейбниц, «...и игры заслуживают изучения; и если какой-нибудь проницательный математик посвятит себя их изучению, то получит много важных результатов, ибо нигде человек не показывает столько изобретательности, как в игре» [17, С. 15].

Нет математической теории, которая могла бы дать алгоритм любой реальной игры, но существуют ситуации, подобные игровым и допускающие математический анализ.

Остановимся на классификации игр.

Интересы участников игры (игроков) могут оказаться несовпадающими и даже противоположными. В этом случае игра называется *антагонистической*.

В игре могут участвовать два или более игроков. Случай игры с одним участником (пасьянс, управление физическим объектом и т. д.) в сущности является игрой двух лиц, где вторым участником выступает природа (судьба, рок, провидение).

Игроки могут в игре выступать каждый за себя или объединяться в группы. В последнем случае игра называется *коалиционной*.

Игры, в которых игроки осведомлены о состоянии своем и партнеров, а также о прошлом поведении участников игры, относятся к категории игр *с полной информацией* (типичные примеры – шахматы, крестики-нолики и т. п.). Большинство же игр протекает в условиях неполной информации, где сведения о состоянии партнеров исчерпываются лишь вероятностными характеристиками (домино, карточные игры, игры против природы).

Антагонистическую игру, где выигрыш одного коллектива равен проигрышу другого, называют *игрой с нулевой суммой*.

Система правил, однозначно определяющая выбор хода игрока в зависимости от сложившейся ситуации, называется *стратегией*.

Каждая фиксированная стратегия игрока, где любой ситуации сопоставлен конкретный выбор, называется *чистой*. В реальности чаще используются *смешанные стратегии*, где чистые стратегии смешиваются с некоторыми частотами [18].

Простейшими являются игры двух лиц с нулевой суммой.

Пусть в такой игре игрок 1 имеет m выборов и игрок 2 – n выборов. Если игрок 1 делает свой i-й выбор, а игрок 2 – свой j-й выбор, то выигрыш игрока 1 (проигрыш игрока 2) равен R_{ij}. Такая игра называется *матричной* и матрица $R = \left[R_{ij} / i = 1...m, j = 1...n \right]$ называется *матрицей выигрышей* (платежной матрицей).

При ведении игры игрок должен ориентироваться на оптимальную политику партнера и наказывать его за отступления от таковой.

Проведем рассуждения за игрока 1. Если Я воспользуюсь i-м выбором, мой противник для минимизации моего выигрыша сделает тот из своих выборов, который даст $\min R_{ij}$. Соответственно, Я должен использовать тот выбор, который гарантирует мне выигрыш, не меньший

$$V_1 = \max_{i=1...m} \min_{j=1...n} R_{ij}.$$

Противник, рассуждая аналогично, приходит к выводу о гарантированном проигрыше, не превышающем

$$V_2 = \max_{i=1...m} \min_{j=1...n} R_{ij}.$$

*Если в матрице выигрышей существует элемент $R_{kl} = V_1 = V_2$, то говорят о наличии оптимальной политики «в пространстве чистых стратегий» и оптимальными выборами для игроков соответственно являются выборы k и l. Пару (k, l) называют **седловой точкой**.*

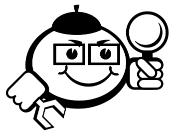
Пример 4.1

Пусть игра определяется матрицей:

$$R=\begin{Vmatrix}2&3&4&5\\3&4&5&6\\4&3&4&6\\6&6&8&10\end{Vmatrix},\ V_1=\max\begin{Vmatrix}2\\4\\3\\6\end{Vmatrix}=6,\ V_2=\min\left[6,6,8,10\right]=6.$$

Седловые точки – (4, 1) и (4, 2). Цена игры = 6; оптимальный выбор для игрока 1 – четвертый, для игрока 2 равнозначны первый и второй (*под ценой игры понимают гарантированный выигрыш-проигрыш при оптимальной политике обоих игроков*).

. . . .



Пример 4.2

Пусть игра определяется матрицей:

$$R=\begin{Vmatrix}5&4&3&7&6\\6&2&4&3&5\\2&7&3&5&4\\7&3&4&4&4\end{Vmatrix},\ V_1=\max\begin{Vmatrix}3\\2\\2\\3\end{Vmatrix}=3,\ V_2=\min\left[7,7,4,7,6\right]=4.$$

Здесь равенство $V_1=V_2$ не выполняется; оптимальной чистой стратегии для игроков нет.

. . . .

При анализе игр часто прибегают к попыткам обнаружить доминирование между строками и столбцами. Так, в примере 4.1 элементы четвертой строки больше элементов других строк: использование выбора 4 выгоднее других выборов при любой политике противника. Противник видит, что в такой ситуации использовать выборы 3 и 4 неразумно.

Использование доминирования таким образом позволяет уменьшить размеры изучаемой матрицы исключением «невыгодных» строк и столбцов.

При отсутствии седловой точки среди чистых стратегий приходится искать таковую среди смешанных [19].

Если игрок 1 прибегает к своему выбору *i* с вероятностью P_i, а игрок 2 – к своему *j*-му выбору с вероятностью Q_j, то ожидаемый выигрыш игрока 1 (проигрыш игрока 2) равен $\sum_{i=1}^{m} R_{ij}P_iQ_j = P^T RQ.$

Основная теорема теории игр (*теорема Джона фон Неймана*) утверждает, что любая матричная игра с нулевой суммой всегда имеет седловую точку, т. е. существуют векторы P и Q, такие, что

$$\max_P \min_Q P^T RQ = \min_Q \max_P P^T RQ = V,$$

(V – цена игры).

4.2 Матричные игры и линейное программирование

Очевидно, что если игрок 1 отступит от оптимальной политики, а игрок 2 будет действовать оптимально, то выигрыш игрока 1 будет меньше цены игры, и если игрок 2 отступит от оптимальной политики при сохранении оптимального поведения игроком 1, то его проигрыш превысит цену игры:

$$P^T RQ_{opt} \le V = P_{opt}^T RQ_{opt} \le P_{opt}^T RQ.$$

Рассуждения игрока 1: мне хотелось бы максимизировать цену игры, т. е. мой гарантированный выигрыш, и я должен подобрать систему значений P_i так, чтобы при любом выборе противника мой ожидаемый выигрыш был больше цены игры [20].

Рассуждения игрока 2: мне хочется уменьшить мой гарантированный проигрыш, т. е. цену игры, и мне надо подобрать значения Q_j так, чтобы при любом выборе противника мой проигрыш был меньше цены игры.

Отсюда возникают две задачи:

максимизировать V	минимизировать V
при условиях	при условиях
$\sum_{i=1}^{m} R_{ij} P_i \ge V, \quad j = 1...n$	$\sum_{j=1}^{n} R_{ij} Q_j \le V, \quad i = 1...m$
$\sum_{i=1}^{m} P_i = 1$	$\sum_{j=1}^{n} Q_j = 1$
$P_i \ge 0, \quad i = 1...m$	$Q_j \ge 0, \quad j = 1...n$

Легко показать, что эти задачи образуют пару двойственных задач линейного программирования.

Таким образом решение матричной игры сводится к решению пары двойственных линейных программ.

Обратим внимание на то, что при увеличении элементов матрицы R на любую константу C цена игры увеличится на C и это изменение не окажет влияния на искомые вероятности выборов. Таким образом, можно добиться, например, положительности элементов матрицы и, следовательно, цены игры. Поэтому можно допустить, что цена игры V положительна.

В предположении $V > 0$ проведем замену переменных:

$$X_i = P_i / V, \quad Y_j = Q_j / V.$$

Отсюда видно, что

$$V = 1 \Big/ \sum X_i = 1 \Big/ \sum Y_j .$$

Соответственно, поставленные задачи можно преобразовать к задачам с меньшим числом переменных:

минимизировать

$$\sum_{i=1}^{m} X_i$$

при условиях

$$\sum_{i=1}^{m} R_{ij} X_i \ge 1, \quad j = 1...n$$

$$X_i \ge 0, \quad i = 1...m$$

максимизировать

$$\sum_{j=1}^{n} Y_j$$

при условиях

$$\sum_{j=1}^{n} R_{ij} Y_j \le 1, \quad i = 1...m$$

$$Y_j \ge 0, \quad j = 1...n$$

Например, для игры с матрицей

$$\begin{Vmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ 2 & 3 & 0 \end{Vmatrix}$$

возникают задачи:

максимизировать

$$Y_1 + Y_2 + Y_3$$

при

$$Y_1 + 2Y_2 + 3Y_3 \le 1$$

$$4Y_1 + Y_3 \le 1$$

$$2Y_1 + 3Y_2 \le 1$$

$$Y_1, Y_2, Y_3 \ge 0$$

минимизировать

$$X_1 + X_2 + X_3$$

при

$$X_1 + 4X_2 + 2X_3 \ge 1$$

$$2X_1 + 3X_3 \ge 1$$

$$3X_1 + X_2 \ge 1$$

$$X_1, X_2, X_3 \ge 0$$

Решение этих задач симплексным методом дает оптимальные значения:

$$X = \{11/37, 4/37, 5/37\}, \quad Y = \{8/37, 7/37, 5/37\}$$

и экстремумы целевых функций, равные 20/37.

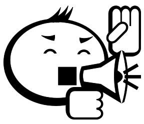
Отсюда

$$V = 37/20,$$

$$P = \{11/20, 4/20, 5/20\},$$

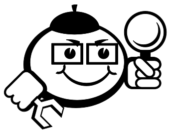
$$Q = \{8/20, 7/20, 5/20\}.$$

4.3 Итеративный метод решения матричных игр



Как мы показали выше, игры могут решаться методами линейного программирования. Здесь мы рассмотрим *итеративный метод Брауна – Робинсон*, обычно используемый при решении игр большой размерности.

Используется многократная реализация игры на основе знания предыстории с последовательным совершенствованием стратегий [21].



Пример 4.3

Для примера возьмем задачу, которую мы только что решили.

Пусть игрок 1 сделал выбор 1 с ожидаемыми выигрышами 1, 2, 3. Противник, стремясь минимизировать свой проигрыш, прибегнет к выбору 1 с ожиданием проигрыша 1, 4, 2. Игрок 1 в стремлении максимизировать свой выигрыш прибегнет к выбору 2, что даст ему надежду на суммарный выигрыш (1 + 4, 2 + 0, 3 + 1). Но тогда его противник найдет среди этих значений меньшее и прибегнет к выбору 2 с ожидаемым суммарным проигрышем (1 + 2, 4 + 0, 2 + 3) и т. д. (табл. 4.1):

Таблица 4.1 – Расчет с использованием итеративного метода

Шаг	**Выбор** ***i***	**Суммарный выигрыш**			**Выбор** ***j***	**Суммарный проигрыш**		
1	1	1	2	3	1	1	4	2
2	2	5	2	4	2	3	4	5
3	3	7	5	4	3	6	5	5
4	1	8	7	7	2	8	5	8
5	1	9	9	10	1	9	9	10
6	3	11	12	10	3	12	10	10

Шаг	Выбор i	Суммарный выигрыш			Выбор j	Суммарный проигрыш		
7	1	12	14	13	1	13	14	12
8	2	16	14	13	3	16	15	12
9	1	17	16	16	2	18	15	15
10	1	18	18	19	1	19	19	17

Этот процесс реализуется достаточно большое число раз с последующим поиском частоты использования выборов и усреднением значений выигрышей-проигрышей.

В результате 10 выборов для 1-го игрока частоты составили 0,6, 0,2, 0,2; для игрока 2 – 0,4, 0,3, 0,3; оценка цены игры в диапазоне от 1,7 до 1,9.

. . . .

4.4 Многошаговые игры. Игры на выживание

Предыдущее рассмотрение игр проводилось в предположении, что реализация игры может осуществляться любое число раз. Например, для игры «орел или решка», где в случае совпадения предъявляемых сторон монеты выигрывает игрок 1 и при несовпадении – игрок 2, оптимальная политика игроков состоит в равновероятностном выборе «орла» и «решки» и цена игры равна 0.

Однако в реальной игре с ограниченными ресурсами политика игроков зависит от результата предыдущих действий и от длительности игры.

Соответственно для матричной игры

$$F_k(A,B) = \max_P \min_Q \sum_{i=1}^{m}\sum_{j=1}^{n} P_i\, Q_j F_{k-1}(A+R_{ij}, B-R_{ij}) =$$

$$= \min_Q \max_P \sum_{i=1}^{m}\sum_{j=1}^{n} P_i\, Q_j F_{k-1}(A+R_{ij}, B-R_{ij}),$$

где $F_k(A,B)$ – ожидаемый выигрыш игрока 1 в k последовательных реализациях при начальных ресурсах A и B и использовании оптимальной политики.

Пусть общий начальный ресурс игроков $A+B=C$ и игра продолжается до разорения одного из игроков. Обозначим через $F(A)$ ожидаемую вероятность выживания (шансы не разориться) игрока 1 при его начальном ресурсе A и оптимальной политике обоих игроков.

Тогда

$$F(A) = \max_P \min_Q \sum_{i=1}^{m}\sum_{j=1}^{n} P_i\, Q_j F(A+R_{ij}) =$$

$$= \min_Q \max_P \sum_{i=1}^{m}\sum_{j=1}^{n} P_i \, Q_j F(A + R_{i\,j}),$$

$$F(A \le 0) = 0,$$

$$F(A \ge C) = 1.$$

Если игра не обладает седловой точкой в пространстве чистых стратегий, то оптимальные значения вероятностей использования выборов соответствуют внутренним точкам множества планов $(0 < P < 1, 0 < Q < 1)$ и напрашивается мысль прибегнуть к аппарату производных.

Пример 4.4

Рассмотрим игру на выживание с матрицей $\begin{Vmatrix} 2 & -1 \\ -2 & 1 \end{Vmatrix}$ при полном капитале игроков $C = 4$.

Здесь в силу целочисленности данных берем целочисленные значения A от 0 до 4. Если обозначить вероятности соответствующих выборов игроков через $P, 1-P, Q, 1-Q,$ то $F(A \le 0) = 0, F(A \ge 4) = 1,$

$$F(1) = \max_P \min_Q \left[PQ\,F(3) + P(1-Q)\,F(0) + (1-P)Q\,F(-1) + (1-P)(1-Q)\,F(2) \right] =$$

$$= \max_P \min_Q \left[PQ\,F(3) + (1-P)(1-Q)\,F(2) \right],$$

$$F(2) = \max_P \min_Q \left[PQ\,F(4) + P(1-Q)\,F(1) + (1-P)Q\,F(0) + (1-P)(1-Q)\,F(3) \right] =$$

$$= \max_P \min_Q \left[PQ + P(1-Q)\,F(1) + (1-P)(1-Q)\,F(3) \right],$$

$$F(3) = \max_P \min_Q \left[PQ\,F(5) + P(1-Q)\,F(2) + (1-P)Q\,F(1) + (1-P)(1-Q)\,F(4) \right] =$$

$$= \max_P \min_Q \left[PQ + P(1-Q)\,F(2) + (1-P)Q\,F(1) + (1-P)(1-Q) \right].$$

Отыскиваем частные производные и строим системы уравнений для поиска оптимальных значений $P(A), Q(A)$:

$$1.\ \frac{\partial}{\partial P} F(1) = Q\,F(3) - (1-Q)\,F(2) = 0,$$

$$\frac{\partial}{\partial Q} F(1) = P\,F(3) - (1-P)\,F(2) = 0.$$

$$2.\ \frac{\partial}{\partial P} F(2) = Q + (1-Q)\,F(1) - (1-Q)\,F(3) = 0,$$

$$\frac{\partial}{\partial Q} F(2) = P - P\,F(1) - (1-P)\,F(3) = 0.$$

$$3.\ \frac{\partial}{\partial P}F(3)=Q+(1-Q)\ F(2)-Q\ F(1)-(1-Q)=0,$$

$$\frac{\partial}{\partial Q}F(3)=P-P\ F(2)+(1-P)\ F(1)-(1-P)=0.$$

Решение приведенных систем дает

$$P(1)=Q(1)=\frac{F(2)}{F(2)+F(3)},$$

$$P(2)=\frac{F(3)}{1-F(1)+F(3)},\quad Q(2)=\frac{F(3)-F(1)}{1-F(1)+F(3)},$$

$$P(3)=\frac{1-F(1)}{2-F(1)-F(2)},\quad Q(3)=\frac{1-F(2)}{2-F(1)-F(2)}.$$

Подставляя полученные выражения в исходные выражения функций, имеем:

$$F(1)=\frac{F(2)\cdot F(3)}{F(2)+F(3)},\quad F(2)=\frac{F(3)}{1-F(1)+F(3)},\quad F(3)=\frac{1-F(1)\cdot F(2)}{2-F(1)-F(2)}.$$

Решая полученную нелинейную систему, имеем оценки:

$$F(1)=0,3,\quad F(2)=0,5,\quad F(3)=0,7$$

и

$$P(1)=0,41,\ P(2)=0,5,\ P(3)=0,59,\ Q(1)=0,41,\ Q(2)=0,3,\ Q(3)=0,41.$$

. . . .

4.5 Многошаговые игры. Игры погони

Простейшим примером таких игр может служить задача для двух игроков, расположившихся на прямой на расстоянии d. На каждом шаге игры игроки могут одновременно смещаться влево или вправо при полной информации о позиции друг друга После очередного шага игрок 2 уплачивает игроку 1 величину $G(S)$, где S – расстояние между ними. С вероятностью $A(d)$ игра может быть продолжена и с вероятностью $1-A(d)$ окончена [22].

Если обозначить через P_1, P_1, Q_1, Q_2 вероятности смещения игроков в ту или иную сторону, то одна из возможных формулировок задачи имеет вид:

$$F(d)=G(d)+A(d)\max_P\min_Q\left[P_1Q_1F(d)+P_1Q_2F(d+2)+\right.$$

$$\left.+P_2Q_1F(d-2)+P_2Q_2F(d)\right]=$$

$$=G(d)+A(d)\min_Q\max_P\left[P_1Q_1F(d)+\ldots\right].$$

Существенно больший интерес может представить игра погони на плоскости или в пространстве, где устанавливается принципиальная возможность поимки одного игрока другим или отыскивается траектория, минимизирующая время поимки. Эти игры относятся к так называемым непрерывным многошаговым играм, решение которых сводится к дискретным моделям [18].

4.6 Статистические решения. Основные понятия

Теория статистических решений может быть истолкована как теория поиска оптимального недетерминированного поведения в условиях неопределенности. Современная концепция статистического решения выдвинута А. Вальдом и считает поведение оптимальным, если оно минимизирует риск в последовательных экспериментах, т. е. математическое ожидание убытков статистического эксперимента. В такой постановке любая задача статистических решений может рассматриваться как игра двух лиц, в которой одним из игроков является природа.

Выбор наилучших решений в условиях неполной информации является одним из основных занятий людей [7].

Собираясь в туристический поход, мы укладываем вещи в рюкзак с учетом неизвестной погоды и преследуем цель получить максимум удовольствий, не превращаясь в рекордсмена по переноске тяжестей.

Проектируя гидротехнические сооружения, мы стремимся сделать их надежными, несмотря на непредсказуемые землетрясения, паводки и т. п.

Создавая систему профилактических и аварийных ремонтов, мы преследуем какую-то цель, не зная в точности времени возникновения аварий.

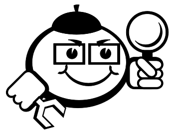
Если процесс определяется повторяющимися ситуациями, то его усредненные характеристики испытывают тенденцию к стабилизации и появляется возможность либо замены случайного процесса детерминированным, либо использования каких-то методов исследования стационарных случайных процессов (в частности, *методов теории массового обслуживания*).

Однако большинство процессов характеризуется «дурной неопределенностью» и невозможно найти законы распределения и другие вероятностные характеристики. В таких ситуациях приходится прибегнуть к экспертным оценкам.

Возникает и проблема выбора критерия оптимальности, поскольку решение, оптимальное для каких-то условий, бывает неприемлемым в других и приходится искать некоторый компромисс.

Пусть задан некоторый вектор $S=\left(S_1,S_2,...,S_n\right)$, описывающий n состояний внешней среды, и вектор $X=\left(X_1,X_2,...,X_m\right)$, описывающий m допустимых решений. Требуется найти вектор $X^*=\left(0,0,...,0,X_i,0,...,0\right)$, который обеспечивает оптимум некоторой функции полезности $W(X,S)$ по некоторому критерию K.

Информация об указанной функции представляют матрицей размерности $m\times n$ с элементами $W_{ij}=F(X_i,S_j)$, где F – *решающее правило*.



Пример 4.5

Рассмотрим типичный пример формирования такой матрицы.

Планируется выпуск новой продукции, для чего необходимо закупить станки. Система оптовой торговли может поставить не более 50 станков; комплект поставки – 10 станков. Минимальный объем поставок – 20 станков. Соответственно, вектор решений об объеме поставок $X=(20,30,40,50)$.

Ежегодный доход от продукции, снимаемой с одного станка, составляет 21,9 тыс. руб. Оптовая цена одного станка 4,775 тыс. руб., эксплуатационные расходы – 3,6 тыс. руб. Затраты на подготовку производства составляют 25,5 тыс. руб. и не зависят от числа станков и объема выпуска.

Пусть спрос пропорционален количеству продукции, снимаемой с S работающих станков, и для простоты ограничимся вектором состояний спроса $S=(0,10,20,30,40,50)$.

Если решающее правило сформулировать как «доход – издержки», то можно рассчитать элементы матрицы полезности (табл. 4.2):

$$W_{ij}=(21{,}9-3{,}6)\times\min\left(X_i,S_j\right)-4{,}775X_i-25{,}5.$$

Таблица 4.2 – Матрица полезности

	$S_1=0$	$S_1=10$	$S_1=20$	$S_1=30$	$S_1=40$	$S_1=50$
$X_1=20$	–121	62	245	245	245	245
$X_2=30$	–168,75	14,25	197,25	380,25	380,25	380,25
$X_3=40$	–216,5	–33,5	149,5	332,5	515,5	515,5
$X_4=50$	–264,25	–81,25	101,75	284,75	467,75	650,75

Например, $W_{11}=-(4,775\times20+25,5)=-121$,

$W_{12}=(21,9-3,6)\times10-(4,775\times20+25,5)=62$,

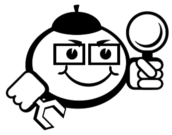
$W_{13}=(21,9-3,6)\times20-(4,775\times20+25,5)=245$,

$W_{14}=W_{15}=245$ (спрос останется неудовлетворенным).

4.7 Выбор критерия принятия решения

Предположим, что в нашем распоряжении имеются статистические данные, позволяющие оценить вероятность того или иного спроса, и этот опыт может быть использован для оценки будущего. При известных вероятностях P_j для спроса S_j можно найти математическое ожидание $W(X,S,P)$ и определить вектор X^*, дающий

$$W=\max_{i=1\ldots m}\sum_{j=1}^{n}W_{ij}P_j.$$



Пример 4.6

Если для вышеприведенного примера задать вектор $P=(0,01,\ 0,09,\ 0,2,\ 0,3,\ 0,3,\ 0,1)$, то математические ожидания прибыли при разных выборах:

$$W_1=-121\times0,01+62\times0,09+245\times0,2+245\times0,3+245\times0,3+245\times0,1=224,87,$$

$$W_2=305,22,\ W_3=330,675,\ W_4=301,12-$$

и выбор максимального значения обнаруживает оптимальность варианта 40 станков с ожидаемой прибылью 330,675 тыс. руб.

4.7.1 Критерий Лапласа

В основе этого критерия лежит *принцип недостаточного основания* [23].

Если нет достаточных оснований считать, что вероятности того или иного спроса имеют неравномерное распределение, то они принимаются одинаковыми и задача сводится к поиску варианта, дающего

$$W=\max_{i=1\ldots m}\frac{1}{n}\sum_{j=1}^{n}W_{ij}.$$

Пример 4.7

Для нашего примера

$$W_1 = (-121 + 62 + 245 + 245 + 245 + 245)/6 = 153{,}5,$$

$$W_2 = 197{,}25,\ W_3 = 210{,}5,\ W_4 = 193{,}5$$

и выбор максимального значения обнаруживает оптимальность выбора варианта 40 станков с ожидаемой прибылью 210,5 тыс. руб.

4.7.2 Критерий Вальда

Критерий Вальда [24] обеспечивает выбор осторожной, пессимистической стратегии в той или иной деятельности, и его суждения близки к тем суждениям, которые мы использовали в теории игр для поиска седловой точки в пространстве чистых стратегий: для каждого решения X_i выбирается самая худшая ситуация (наименьшее из W_{ij}) и среди них отыскивается гарантированный максимальный эффект:

$$W = \max_{i=1\ldots m} \min_{j=1\ldots n} W_{ij}.$$

Пример 4.8

В нашем примере $W = \max(-121,\ -168{,}75,\ -216{,}5,\ -264{,}25) = -121,$ т. е. по этому критерию следует закупить 20 станков и максимальный возможный убыток не превысит 121 тыс. руб. (если бы мы включили и вариант отказа от покупки станков вообще, то этот критерий рекомендовал бы нам воздержаться от какой-либо деятельности, но «кто не рискует, тот не пьет шампанского»).

Можно принять и критерий выбора оптимистической стратегии:

$$W = \min_{i=1\ldots m} \max_{j=1\ldots n} W_{ij},$$

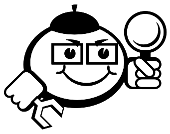
где оценивается гарантированный выигрыш при самых благоприятных условиях. Для нашего примера $W = \min(245,\ 380{,}25,\ 515{,}5,\ 650{,}75) = 245.$

4.7.3 Критерий Гурвица

Ориентация на самый худший исход является своеобразной перестраховкой [25]. Однако опрометчиво выбирать политику, которая излишне оптимистична. Критерий Гурвица предлагает некоторый компромисс:

$$W = \max_{i=1\ldots m}\left[\alpha \max_{j=1\ldots n} W_{ij} + (1-\alpha)\min_{j=1\ldots n} W_{ij}\right],$$

где параметр α принимает значение от 0 до 1 и выступает как коэффициент оптимизма.



Пример 4.9

Так, в нашем примере при различных α значения W определяются таблицей 4.3:

Таблица 4.3 – Расчет по критерию Гурвица

	$\alpha = 0{,}1$	$\alpha = 0{,}2$	$\alpha = 0{,}5$	$\alpha = 0{,}8$	$\alpha = 0{,}9$
$X_1 = 20$	–84,4	–47,0	62	171	206,4
$X_2 = 30$	–113,85	–58,95	105,75	270,45	325,35
$X_3 = 40$	–140,3	–70,1	149,5	369,1	442,3
$X_4 = 50$	–172,75	–81,25	193,25	467,75	559,25

При $\alpha = 0{,}5$ (равновероятных шансах на успех и неудачу) следует закупить 50 станков и ожидать прибыль порядка 193,25 тыс. руб.

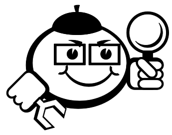
При вероятности успеха 0,2 не следует закупать более 20 станков с надеждой, что убытки не превысят 47 тыс. руб.

4.7.4 Критерий Сэвиджа

Суть этого критерия заключается в нахождении минимального риска. При выборе решения по этому критерию сначала матрице функции полезности (эффективности) сопоставляется *матрица сожалений*:

$$D_{ij} = W_{ij} - \max_i (W_{ij}),$$

элементы которой отражают убытки от ошибочного действия, т. е. выгоду, упущенную в результате принятия i-го решения в j-м состоянии. Затем по матрице D выбирается решение по пессимистическому критерию Вальда, дающее наименьшее значение максимального сожаления [22].



Пример 4.10

Для нашего примера отыскиваем матрицу D, вычитая (–121) из первого столбца матрицы полезности, 62 из второго и т. д. (табл. 4.4).

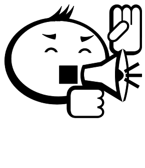
Таблица 4.4 – Расчет по критерию Сэвиджа

	$S_1=0$	$S_1=10$	$S_1=20$	$S_1=30$	$S_1=40$	$S_1=50$
$X_1=20$	0	0	0	–135,25	–270,5	–405,75
$X_2=30$	–47,75	–47,75	–47,75	0	–135,25	–270,5
$X_3=40$	–95,5	–95,5	–95,5	–47,75	0	–135,25
$X_4=50$	–143,25	–143,25	–143,25	–95,5	–47,75	0

Наибольшее значение среди минимальных элементов строк здесь равно $\max\left[-405{,}75,\ -270{,}5,\ -135{,}25,\ -143{,}25\right]=-135{,}25$, и, покупая 40 станков, мы уверены, что в худшем случае убытки не превысят 135,25 тыс. руб.

Таким образом, различные критерии приводят к различным выводам:

1) по критерию Лапласа приобретать 40 станков,
2) по критерию Вальда – 20 станков,
3) по критерию Гурвица – 20 при пессимистическом настроении и 50 в состоянии полного оптимизма,
4) по критерию Сэвиджа – 40 станков.



Возможность выбора критерия дает свободу лицам, принимающим экономические решения, при условии, что они располагают достаточными средствами для постановки подобной задачи. Всякий критерий должен согласовываться с намерениями решающего задачу и соответствовать его характеру, знаниям и убеждениям.

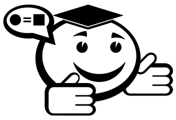
Контрольные вопросы по главе 4

1. Что изучает теория игр?
2. Что такое игры с природой?
3. Что называют седловой точкой?
4. Что представляет собой матрица выигрышей?
5. Что такое цена игры?
6. Опишите алгоритм итеративного метода.
7. Перечислите критерии принятия решения.

5 Ряды динамики

5.1 Понятие о рядах динамики

Одной из важнейших экономических задач является изучение изменений анализируемых показателей во времени, то есть их динамика. Эта задача решается при помощи анализа рядов динамики (временных рядов).



Ряд динамики *– это числовые значения определенного экономического показателя в последовательные моменты или периоды времени (т. е. расположенные в хронологическом порядке).*

Числовые значения того или иного показателя, составляющего ряд динамики, называют *уровнями* ряда (будем обозначать их y). Первый член ряда y_1 называют начальным (базисным) уровнем, а последний y_n – конечным. Моменты или периоды времени, к которым относятся уровни, обозначают через t.

Ряды динамики, как правило, представляют в виде таблицы (см. табл. 5.1) или графически (см. рис. 5.1), причем по оси абсцисс строится шкала времени t, а по оси ординат – шкала уровней ряда y.

Таблица 5.1 – Внешнеторговый оборот России за 2000–2006 гг.

Год	2000	2001	2002	2003	2004	2005	2006
Млрд долл. США	149,9	155,6	168,3	212,0	280,6	368,9	468,4

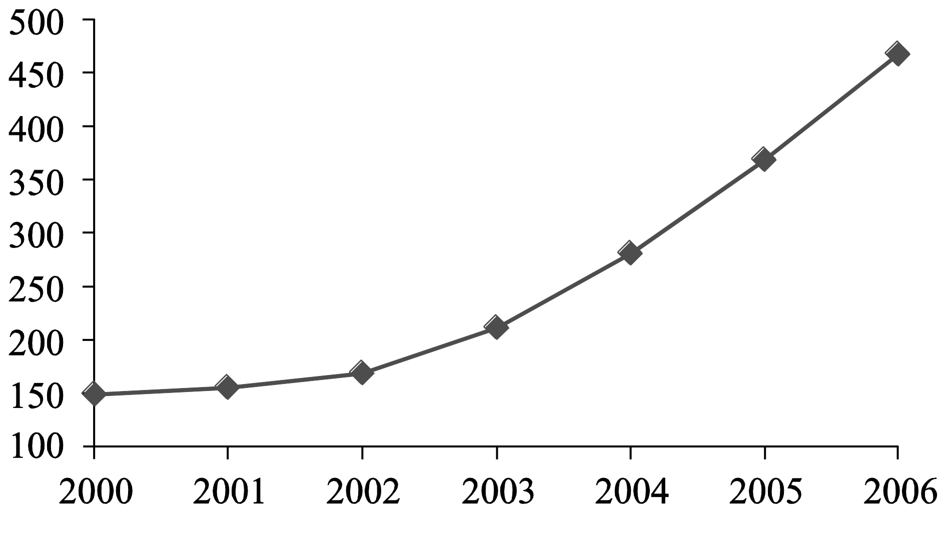


Рис. 5.1 – Внешнеторговый оборот России за 2000–2006 гг.

Данные таблицы 5.1 и рисунка 5.1 наглядно иллюстрируют ежегодный рост внешнеторгового оборота (ВО) в России за 2000–2006 гг.

5.2 Показатели изменения уровней ряда динамики

Цепные и базисные показатели динамики

Анализ рядов динамики начинается с определения того, как именно изменяются уровни ряда (увеличиваются, уменьшаются или остаются неизменными) в абсолютном и относительном выражении. Чтобы проследить за направлением и размером изменений уровней во времени, для рядов динамики рассчитывают *показатели изменения уровней ряда динамики*:

- абсолютное изменение (абсолютный прирост);
- относительное изменение (темп роста или индекс динамики);
- темп изменения (темп прироста) [26].

Все эти показатели могут определяться *базисным* способом, когда уровень данного периода сравнивается с первым (базисным) периодом, либо *цепным* способом, когда сравниваются два уровня соседних периодов.

Абсолютное изменение *(абсолютный прирост) уровней рассчитывается как разность между двумя уровнями ряда:*
для базисного способа сравнения по формуле:

$$\Delta y_i^{\text{Б}} = y_i - y_1;$$

или для цепного по формуле:

$$\Delta y_i^{\text{Ц}} = y_i - y_{i-1}.$$

Оно показывает, на сколько (в единицах показателей ряда) уровень одного (i-того) периода отличается от уровня предшествующего или первоначального периода соответственно.

Темп роста *(индекс динамики) – относительное изменение уровней ряда, которое рассчитывается как отношение (деление) двух уровней:*

для базисного способа сравнения по формуле

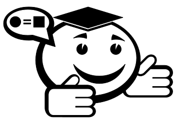
$$T_{p_i}^{Б} = y_i / y_1;$$

или для цепного по формуле:

$$T_{p_i}^{Ц} = y_i / y_{i-1}.$$

Относительное изменение показывает, во сколько раз уровень данного периода больше уровня какого-либо предшествующего периода (при $i_i > 1$) или какую его часть составляет (при $i_i < 1$). Относительное изменение может выражаться в виде *коэффициентов*, то есть простого кратного отношения (если база сравнения принимается за единицу), и в *процентах* (если база сравнения принимается за 100 единиц) путем домножения относительного изменения на 100%.

Темп прироста уровней – относительный показатель, показывающий, на сколько процентов данный уровень больше (или меньше) другого, принимаемого за базу сравнения. Он рассчитывается путем вычитания из относительного изменения 100%, соответственно для цепных и базисных показателей, то есть по формулам:



Темп прироста базисный:

$$T_{np_i}^{Б} = T_{p_i}^{Б} - 100\%.$$

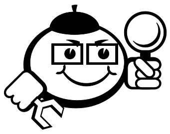
Темп прироста цепной:

$$T_{np_i}^{Ц} = T_{p_i}^{Ц} - 100\%.$$

Абсолютное значение одного процента прироста *представляет собой отношение абсолютного прироста к темпу прироста и определяется по формуле:*

$$\text{А} = 0{,}01 y_{i-1}.$$

Абсолютное значение одного процента прироста определяется только для цепных приростов. Среднее абсолютное значение одного процента прироста не рассчитывается.

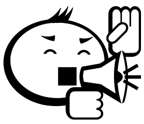


Пример 5.1

В таблице 5.2 рассчитаны цепные и базисные показатели динамики.

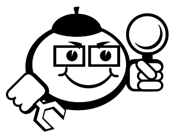
Таблица 5.2 – Анализ динамики ВО России

Год	y	$\Delta y_i^{\text{Б}}$	$\Delta y_i^{\text{Ц}}$	$T_{p_i}^{\text{Б}}$	$T_{p_i}^{\text{Ц}}$	$T_{пр_i}^{\text{Ц}}$, %	$T_{пр_i}^{\text{Б}}$, %	А
2000	149,9	–	–	–	–	–	–	–
2001	155,6	5,7	5,7	1,038	1,038	3,8	3,8	1,499
2002	168,3	18,4	12,7	1,123	1,082	12,3	8,2	1,556
2003	212,0	62,1	43,7	1,414	1,260	41,4	26,0	1,683
2004	280,6	130,7	68,6	1,872	1,324	87,2	32,4	2,120
2005	368,9	219,0	88,3	2,461	1,315	146,1	31,5	2,806
2006	468,4	318,5	99,5	3,125	1,270	212,5	27,0	3,689
Итого	1 803,7	–	318,5	–	3,125	–	–	–



Между базисными и цепными абсолютными изменениями существует взаимосвязь: сумма цепных абсолютных изменений равна последнему базисному изменению, то есть

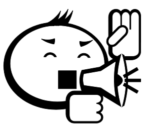
$$\sum_{i=1}^{n} \Delta y_i^{\text{Ц}} = \Delta y_n^{\text{Б}}.$$



Пример 5.2

В нашем примере про ВО подтверждается правильность расчета абсолютных изменений по формуле: $\sum_{i=1}^{n} \Delta y_i^{\text{Ц}} = 318{,}5$ рассчитана в итоговой строке 4-го столбца, а $\Delta y_n^{\text{Б}} = 318{,}5$ – в предпоследней строке 3-го столбца.

В столбце 5 рассчитаны базисные относительные изменения по формуле, а в столбце 6 – цепные относительные изменения по формуле.



Между базисными и цепными относительными изменениями существует взаимосвязь: произведение цепных относительных изменений равно последнему базисному изменению, то есть

$$\prod_{i=1}^{n} i_i^{\text{Ц}} = i_n^{\text{Б}}.$$

Пример 5.3

В нашем примере про ВО подтверждается правильность расчета относительных изменений:

$$\prod_{i=1}^{n} i_i^{\text{Ц}} = 1{,}038 \cdot 1{,}082 \cdot 1{,}260 \cdot 1{,}324 \cdot 1{,}315 \cdot 1{,}270 = 3{,}125$$

рассчитано по данным 6-го столбца, а $i_n^{\text{Б}} = 3{,}125$ – в предпоследней строке 5-го столбца.

Средние показатели ряда динамики

Каждый ряд динамики можно рассматривать как некую совокупность n меняющихся во времени показателей, которые можно обобщить в виде средних величин. Такие обобщенные (средние) показатели особенно необходимы при сравнении динамики изменений того или иного показателя внешнеэкономической деятельности в разные периоды, в разных странах и т. д. [17].

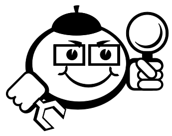
Обобщенной характеристикой ряда динамики служит прежде всего средний уровень ряда $\bar{y}$. Для разных видов рядов динамики он рассчитывается неодинаково.

Ряды динамики бывают равномерные (с равными интервалами времени между уровнями), для которых средний уровень определяется по простой формуле средней величины, и неравномерные (с неравными интервалами), для которых используются формулы средних взвешенных (по интервалам времени) величин.

В интервальном ряду динамики (в котором время задано в виде промежутков времени, к которым относятся уровни) $\bar{y}$ определяется по формуле средней арифметической, а в моментном ряду (в котором время задано в виде конкретных моментов времени или дат, к которым относятся уровни) – по формуле средней хронологической. В таблице 5.3 приводятся виды рядов динамики и соответствующие формулы для расчета их среднего уровня $\bar{y}$.

Таблица 5.3 – Виды средних величин,
применяемых при расчете среднего уровня

Вид ряда динамики	**Название средней величины**	**Формула средней величины**
Равномерный интервальный	Арифметическая простая	$\bar{y} = \dfrac{\sum y}{n}$
Равномерный моментный	Хронологическая простая	$\bar{y} = \dfrac{\dfrac{y_1}{2} + y_2 + y_3 + ... + y_{n-1} + \dfrac{y_n}{2}}{n-1} = \dfrac{\dfrac{y_1 + y_n}{2} + \sum\limits_{i=2}^{n-1} y_i}{n-1}$
Неравномерный интервальный	Арифметическая взвешенная	$\bar{y} = \dfrac{\sum\limits_{i=1}^{n} y_i t_i}{\sum\limits_{i=1}^{n} t_i}$
Неравномерный моментный	Хронологическая взвешенная	$\bar{y} = \dfrac{\sum\limits_{i=1}^{n-1} (y_i + y_{i+1}) t_i}{2\sum\limits_{i=1}^{n-1} t_i}$

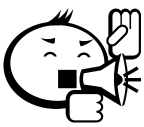


Пример 5.4

В нашем примере про ВО России за 2000–2006 гг. имеем равномерный интервальный ряд динамики, поэтому его средний уровень определяем:

$$\bar{y} = 1803{,}7/7 = 257{,}671,$$

то есть ВО России в 2000–2006 гг. составлял ежегодно в среднем 257,671 млрд долл. США.



Кроме среднего уровня ряда рассчитываются и другие средние показатели:

- средний абсолютный прирост;
- средний темп роста;
- средний темп прироста.

Каждый из этих показателей может рассчитываться базисным и цепным способом.

Средний абсолютный прирост *– это частное от деления последнего базисного абсолютного изменения на количество изменений уровней или частное от деления суммы всех цепных абсолютных изменений на количество изменений:*

$$\Delta\bar{y} = \frac{\Delta y_n^{\text{Б}}}{n-1} = \frac{\sum \Delta y^{\text{Ц}}}{n-1}.$$

По знаку средних абсолютных изменений также судят о характере изменения явления в среднем: рост, спад или стабильность. Очевидно, что числители формулы и равны между собой по формуле, значит, среднее абсолютное изменение не зависит от способа расчета (базисный или цепной), так как результат получится одинаковый.

Пример 5.5

В нашей задаче:

$$\Delta\bar{y} = 318{,}5/6 = 53{,}083,$$

то есть ежегодно в среднем ВО растет на 53,083 млрд долл.

Наряду со средним абсолютным изменением рассчитывается и среднее относительное.

Средний темп роста *показывает, во сколько раз в среднем изменялось значение ряда динамики в течение рассматриваемого периода и определяется по формуле:*

$$\bar{T}_p = \sqrt[n-1]{T_{p_i}^{\text{Б}}} = \sqrt[n-1]{y_n / y_1}.$$

Пример 5.6

В нашем примере про ВО:

$$\bar{T}_p = \sqrt[6]{3{,}125} = 1{,}209,$$

то есть ежегодно в среднем в 2000–2006 гг. ВО России растет в 1,209 раза.

Вычитанием 100% из среднего относительного изменения образуется соответствующий ***средний темп прироста****,*

$$\bar{T}_{пр} = \bar{T}_{р} - 100\%.$$

По его знаку также можно судить о характере изменения изучаемого явления, отраженного данным рядом динамики.

Пример 5.7

В нашем примере:

$$\bar{T}_{пр} = 1,209 - 1 = 0,209,$$

то есть ежегодно в среднем в 2000–2006 гг. ВО России растет на 20,9%.

5.3 Методы выявления основной тенденции (тренда) в рядах динамики

Одна из основных задач изучения рядов динамики – это выявление тренда.

Тренд *– это основная тенденция (закономерность) в изменении уровней ряда.*

Закономерность в изменении уровней ряда в одних случаях проявляется наглядно, в других – может маскироваться колебаниями случайного или неслучайного характера. Поэтому, чтобы сделать правильные выводы о закономерностях развития того или иного показателя, надо суметь отделить тренд от колебаний, вызванных случайными кратковременными причинами. На основании выделенного тренда можно экстраполировать (прогнозировать) развитие явления в будущем. С этой целью (устранить колебания, вызванные случайными причинами) ряды динамики подвергают *обработке*.

Существуют несколько методов обработки рядов динамики, помогающих выявить основную тенденцию изменения уровней ряда, а именно: метод укрупнения интервалов, метод скользящей средней и аналитическое выравнивание. Во всех методах вместо фактических уровней при обработке ряда рассчитываются иные (расчетные) уровни, в которых тем или иным способом взаимопогашается действие случайных факторов и тем самым уменьшается колеблемость уровней. Последние в результате становятся как бы «выравненными», «сглаженными» по отношению к исходным фактическим данным. Такие методы обработки рядов динамики называются *сглаживанием* или *выравниванием* рядов динамики.

Простейший метод сглаживания уровней ряда – *укрупнение интервалов* для определения итогового значения или средней величины исследуемого показателя.

Этот метод особенно эффективен, если первоначальные уровни ряда относятся к коротким промежуткам времени.

Пример 5.8

Например, если имеются данные о ежесуточном производстве мороженого на предприятии за месяц, то, естественно, в таком ряду возможны значительные колебания уровней, так как чем меньше период, за который приводятся данные, тем больше влияние случайных факторов. Чтобы устранить это влияние, рекомендуется укрупнить интервалы времени, например до 5 или 10 дней, и для этих укрупненных интервалов рассчитать общий или среднесуточный объем производства (соответственно по пятидневкам или декадам).

В ряду с укрупненными интервалами времени закономерность изменения уровней будет более наглядной.

По своей сути метод *скользящей средней* похож на метод укрупнения интервалов, но в данном случае фактические уровни заменяются средними уровнями, рассчитанными для последовательно подвижных (скользящих) укрупненных интервалов, охватывающих m уровней ряда.

Каждое звено скользящей средней – это уровень за соответствующий период, который относится к середине выбранного периода, если число уровней скольжения нечетное. Например, для ряда динамики, представленного уровнями y_1, y_2, y_3, y_4, y_5, скольжение 3-уровневой средней будет иметь вид:

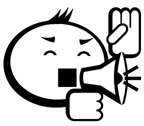
$$y'_2 = \frac{y_1 + y_2 + y_3}{3};\ \ y'_3 = \frac{y_2 + y_3 + y_4}{3};\ \ y'_4 = \frac{y_3 + y_4 + y_5}{3}.$$

Нахождение скользящей средней по четному числу уровней несколько сложнее, так как средняя может быть отнесена к середине между двумя периодами, находящимися в середине интервала сглаживания.

Чтобы ликвидировать такой сдвиг, применяют способ *центрирования*. Центрирование заключается в нахождении средней из двух смежных скользящих средних для отнесения полученного уровня к определенному периоду.

$$y'_{2-3} = \frac{y_1 + y_2 + y_3 + y_4}{4};\ \ y'_{3-4} = \frac{y_2 + y_3 + y_4 + y_5}{4}.$$

Центрированием получаем $y'_3 = \dfrac{y_{2-3} + y_{3-4}}{2}.$



Наиболее совершенным методом обработки рядов динамики в целях устранения случайных колебаний и выявления тренда является *выравнивание уровней ряда по аналитическим формулам* (или *аналитическое выравнивание*).

Суть аналитического выравнивания заключается в замене эмпирических (фактических, исходных) уровней y_i теоретическими $\hat{y}_t$, которые рассчитаны по определенному уравнению, принятому за математическую модель тренда, где теоретические уровни рассматриваются как функция времени: $\hat{y}_t = f(t)$.

При этом каждый фактический уровень y_i рассматривается обычно как сумма двух составляющих:

$$y_i = f(t) + \varepsilon_t,$$

где $f(t) = \hat{y}_t$ – систематическая составляющая, отражающая тренд и выраженная определенным уравнением; ε_t – случайная величина, вызывающая колебания уровней вокруг тренда.

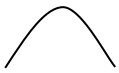
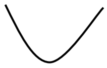
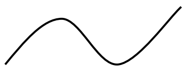
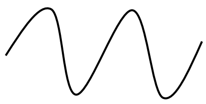
Задача аналитического выравнивания сводится к следующему:

1) определение на основе фактических данных формы (вида) гипотетической функции $\hat{y}_t = f(t)$, способной наиболее адекватно отразить тенденцию развития исследуемого показателя;

2) нахождение по эмпирическим данным параметров указанной функции (уравнения);
3) расчет по найденному уравнению теоретических (выравненных) уровней.

В аналитическом выравнивании наиболее часто используются простейшие функции, представленные в таблице 5.4, где обозначено $\hat{y}_t$ – теоретические (выравненные) уровни (читается как «игрек, выравненный по *t*»); *t* – условное обозначение времени (1, 2, 3,…); $a_0, a_1, a_2, ...$ – параметры аналитической функции; *k* – число гармоник (при выравнивании по ряду Фурье).

Таблица 5.4 – Виды математических функций, используемые при выравнивании

Название функции	**Вид функции**	**Формула**
Прямая линия		$\hat{y}_t = a_0 + a_1 t$
Парабола 2-го порядка	 или 	$\hat{y}_t = a_0 + a_1 t + a_2 t^2$
Парабола 3-го порядка		$\hat{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$
Гипербола		$\hat{y}_t = a_0 + \frac{a_1}{t}$
Показательная		$\hat{y}_t = a_0 a_1^t$
Степенная		$\hat{y}_t = a_0 t^{a_1}$
Ряд Фурье		$\hat{y}_t = a_0 + \sum_{k=1}^{m} \left(a_k \cos kt + b_k \sin kt \right)$

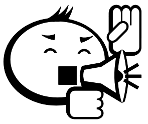
Выбор той или иной функции для выравнивания ряда динамики осуществляется на основании графического изображения эмпирических данных. Если по тем или иным причинам уровни эмпирического ряда трудно описать одной функцией, следует разбить анализируемый период на отдельные части и затем выравнять каждую часть по соответствующей кривой [7].

Нередко один и тот же ряд можно выравнять по разным аналитическим функциям и получить довольно близкие результаты. В нашем примере про ВО России можно произвести выравнивание и по прямой линии, и по параболе.

Чтобы решить вопрос о том, использование какой кривой дает лучший результат, обычно сопоставляют суммы квадратов отклонений эмпирических уровней от теоретических (*остатки*), рассчитанным по разным функциям, то есть:

$$\sum\left(\hat{y}_t - y\right)^2.$$

Та функция, при которой эта сумма минимальна, считается наиболее адекватной, приемлемой. Однако сравнивать непосредственно суммы квадратов отклонений можно в том случае, если сравниваемые уравнения имеют одинаковое число параметров. Если же число параметров k разное, то каждую сумму квадратов делят на разность $(n-k)$, выступающую в роли числа степеней свободы, и сравнивают уже квадраты отклонений уровней, рассчитанные на одну степень свободы (т. е. остаточные дисперсии на одну степень свободы).



Параметры искомых уравнений $(a_0, a_1, a_2, \ldots)$ при аналитическом выравнивании могут быть определены по-разному, но наиболее распространенным методом является *метод наименьших квадратов* (МНК). При этом методе учитываются все эмпирические уровни и должна обеспечиваться минимальная сумма квадратов отклонений эмпирических значений уровней y от теоретических уровней $\hat{y}_t$:

$$\sum\left(\hat{y}_t - y\right)^2 \to \min.$$

В частности, при выравнивании по прямой параметры a_0 и a_1 отыскиваются по МНК следующим образом. В вместо $\hat{y}_t$ записываем его конкретное выражение $a_0 + a_1 t$. Тогда

$$S = \sum\left(a_0 + a_1 t - y\right)^2 \to \min.$$

Дальнейшее решение сводится к задаче на экстремум, т. е. к определению того, при каком значении a_0 и a_1 функция двух переменных S может достигнуть минимума. Как известно, для этого надо найти частные производные S по a_0 и a_1, приравнять их к нулю и после элементарных преобразований решить систему двух уравнений с двумя неизвестными.

В соответствии с вышеизложенным найдем частные производные:

$$\begin{cases} \dfrac{\partial S}{\partial a_0} = 2\sum\left(a_0 + a_1 t - y\right) = 0 \\ \dfrac{\partial S}{\partial a_1} = 2\sum\left(a_0 + a_1 t - y\right)t = 0. \end{cases}$$

Сократив каждое уравнение на 2, раскрыв скобки и перенеся члены с y в правую сторону, а остальные – оставив в левой, получим систему нормальных уравнений:

$$\begin{cases} na_0 + a_1\sum t = \sum y \\ a_0\sum t + a_1\sum t^2 = \sum yt, \end{cases}$$

где n – количество уровней ряда;

t – порядковый номер в условном обозначении периода или момента времени;

y – уровни эмпирического ряда.

Эта система и, соответственно, расчет параметров a_0 и a_1 упрощаются, если отсчет времени ведется от середины ряда[1]. Например, при *нечетном* числе уровней (как в нашем примере про ВО России – 7 уровней) серединная точка времени (год, месяц) принимается за нуль, тогда предшествующие периоды обозначаются соответственно –1, –2, –3 и т. д., а следующие за средним (центральным) – соответственно 1, 2, 3 и т. д. При *четном* числе уровней два серединных момента (периода) времени обозначают –1 и +1, а все последующие и предыдущие соответственно через два интервала: ±3, ±5, ±7 и т. д.

При таком порядке отсчета времени (от середины ряда) $\sum t = 0,$ поэтому система уравнений для нахождения коэффициентов a_0 и a_1 упрощается до следующих двух уравнений, каждое из которых решается самостоятельно:

$$\begin{cases} na_0 = \sum y \Rightarrow a_0 = \dfrac{\sum y}{n} \\ a_1\sum t^2 = \sum yt \Rightarrow a_1 = \dfrac{\sum yt}{\sum t^2}. \end{cases}$$

Как видим, при такой нумерации периодов параметр a_0 представляет собой средний уровень равномерного интервального ряда.

[1]При расчете параметров уравнения тренда на ЭВМ необходимость вести отсчет от середины ряда динамики отпадает. Например, для получения уравнения тренда в Microsoft Office Excel необходимо построить его график с помощью «Мастера диаграмм», после чего вызвать контекстное меню, нажав на правую кнопку мыши на построенном графике, и выбрать пункт «Добавить линию тренда», в появившемся окне выбрать подходящую математическую функцию и установить галочку «показывать уравнение на диаграмме».

Пример 5.9

Определим параметры уравнения прямой для нашего примера про ВО России, для чего исходные данные и все расчеты необходимых сумм представим в таблице 5.5.

Таблица 5.5 – Вспомогательные расчеты для линейного тренда

Год	y	t	t^2	yt	$\hat{y}_t$	$(\hat{y}_t - y)^2$	$(\hat{y}_t - \bar{y})^2$	$(y - \bar{y})^2$
2000	149,9	–3	9	–449,7	97,557	2 739,775	25 636,584	11 614,681
2001	155,6	–2	4	–311,2	150,929	21,822	11 394,038	10 418,577
2002	168,3	–1	1	–168,3	204,300	1 296,000	2 848,509	7 987,252
2003	212	0	0	0	257,671	2 085,879	0,000	2 085,879
2004	280,6	1	1	280,6	311,043	926,768	2 848,509	525,719
2005	368,9	2	4	737,8	364,414	20,122	11 394,038	12 371,795
2006	468,4	3	9	1 405,2	417,786	2 561,806	25 636,584	44 406,531
Итого	1 803,7	0	28	1 494,4	1 803,700	9 652,171	79 758,263	89 410,434

Из таблицы 5.5 получаем, что:

$$a_0 = 1803{,}7/7 = 257{,}671,$$

$$a_1 = 1494{,}4/28 = 53{,}371.$$

Отсюда искомое уравнение тренда:

$$\hat{y}_t = 257{,}671 + 53{,}371t.$$

В 6-м столбце таблицы 5.5 приведены теоретические (трендовые) уровни, рассчитанные по этому уравнению, а в итоге 7-го столбца – остатки. Для иллюстрации построим график эмпирических и трендовых уровней (рис. 5.2).

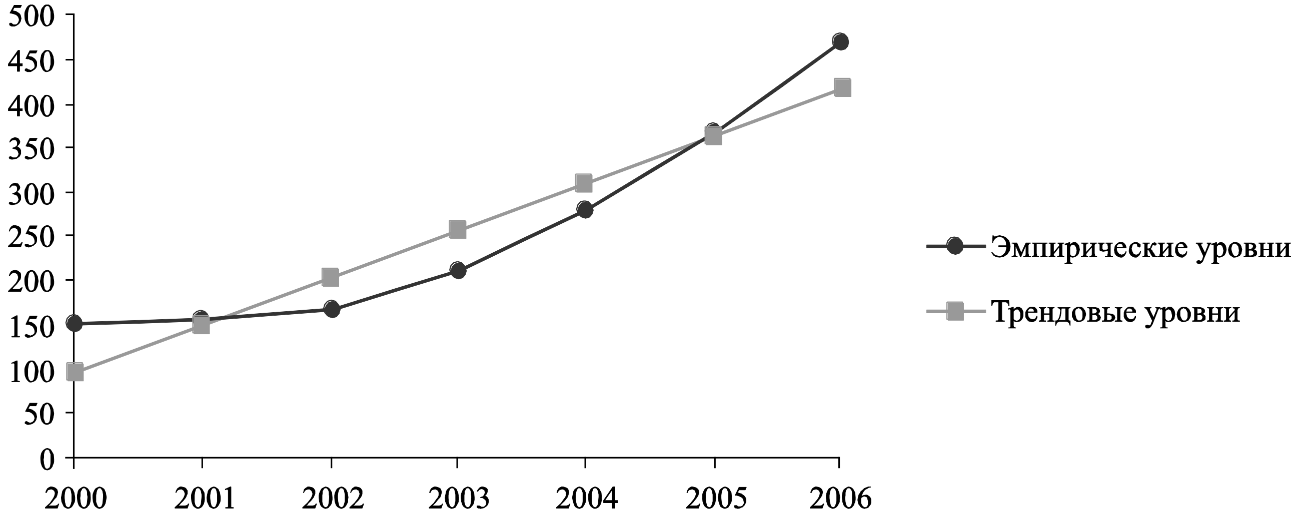
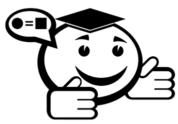


Рис. 5.2 – Эмпирические и трендовые уровни ряда динамики ВО России

. . . .

5.4 Оценка адекватности тренда и прогнозирование

. . . .



Для найденного уравнения тренда необходимо провести оценку его ***надежности (адекватности)****, что осуществляется обычно с помощью критерия Фишера, его расчетное значение F_p сравнивается с теоретическим (табличным) значением F_T (прил. В). При этом расчетный критерий Фишера определяется по формуле:*

$$F_p = \frac{(n-k)\sum\left(\hat{y}_t - \bar{y}\right)^2}{(k-1)\sum\left(\hat{y}_t - y\right)^2},$$

где k – число параметров (членов) выбранного уравнения тренда.

. . . .

Для проверки правильности расчета сумм для F_p можно использовать следующее равенство:

$$\sum(y-\bar{y})^2 = \sum\left(\hat{y}_t - y\right)^2 + \sum\left(\hat{y}_t - \bar{y}\right)^2.$$

Сравнение расчетного и теоретического значений критерия Фишера ведется при заданном уровне значимости (вероятности сделать неверный прогноз) с учетом степеней свободы: $\nu_1 = k-1$ и $\nu_2 = n-k$. При условии $F_p > F_T$ считается, что выбранная математическая модель ряда динамики адекватно отражает обнаруженный в нем тренд [27].

Пример 5.10

В нашем примере про ВО это равенство соблюдается (необходимые суммы рассчитаны в трех последних столбцах таблицы 5.5):

$$89\ 410{,}434 = 9\ 652{,}171 + 79\ 758{,}263.$$

Проверим тренд на адекватность:

$$F_p = 79\ 758{,}263 \cdot 5 / (9\ 652{,}171 \cdot 1) = 41{,}32 > F_T,$$

значит, модель адекватна и ее можно использовать для прогнозирования $(F_T = 6{,}61$ находим по прил. В в 1-м столбце $[\nu_1 = k - 1 = 2 - 1 = 1]$ и 5-й строке $[\nu_2 = n - k = 5])$.

Как уже было отмечено ранее, в нашем примере можно произвести выравнивание не только по прямой линии, но и по параболе, чего делать не будем, так как уже найденный линейный тренд адекватно описывает тенденцию.

При составлении прогнозов уровней социально-экономических явлений обычно оперируют не точечной, а интервальной оценкой, рассчитывая так называемые *доверительные интервалы прогноза*. Границы интервалов определяются по формуле:

$$\hat{y}_t \pm t_\alpha \sigma_{\hat{y}}, \tag{5.1}$$

где $\hat{y}_t$ – точечный прогноз, рассчитанный по модели тренда; t_α – *коэффициент доверия по распределению Стьюдента* при уровне значимости α и числе степеней свободы $\nu = n - 1$ (прил. Б)[1]; $\sigma_{\hat{y}}$ – *ошибка аппроксимации*, определяемая по формуле:

$$\sigma_{\hat{y}} = \sqrt{\frac{\sum (\hat{y}_t - y)^2}{n - k}}.$$

Пример 5.11

Спрогнозируем ВО России на 2007 и 2008 гг. с вероятностью 0,95 (значимостью 0,05), для чего найдем ошибку аппроксимации по формуле:

$$\sigma_{\hat{y}} = \sqrt{9652{,}171 / (7 - 2)} = 43{,}937$$

---

[1]Используется при малом количестве уровней ($n < 30$), в противном случае ($n > 30$) вместо t_α используют коэффициент доверия t нормального закона распределения (прил. А).

и найдем коэффициент доверия по распределению Стьюдента по прил. Б:

$$t_\alpha = 2{,}4469 \text{ при } \nu = 7 - 1 = 6.$$

Прогноз на 2007 и 2008 гг. с вероятностью 0,95:

$$Y_{2007} = (257{,}671 + 53{,}371 \cdot 4) \pm 2{,}4469 \cdot 43{,}937$$

или

$$363{,}6 < Y_{2007} < 578{,}7 \text{ (млрд долл.);}$$

$$Y_{2008} = (257{,}671 + 53{,}371 \cdot 5) \pm 2{,}4469 \cdot 43{,}937$$

или

$$417{,}0 < Y_{2008} < 632{,}0 \text{ (млрд долл.).}$$

Как видно из полученных прогнозов, доверительный интервал достаточно широк (из-за достаточно большой величины ошибки аппроксимации). Более точный прогноз можно получить при выравнивании по параболе 2-го порядка.

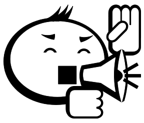
Контрольные вопросы по главе 5

1. Что такое ряд динамики? Что он характеризует?
2. Какие виды рядов динамики Вы знаете?
3. Назовите цепные и базисные показателя рядов динамики. Как они рассчитываются и что показывают?
4. Назовите средние показатели рядов динамики. Как они рассчитываются и что показывают?
5. Какие Вы знаете методы выравнивания рядов динамики?
6. Что такое тренд и как определить его параметры?
7. Как оценить адекватность тренда?

6 Изучение взаимосвязей явлений

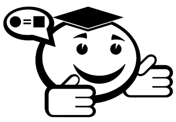
6.1 Понятие корреляционной зависимости

Один из наиболее общих законов объективного мира – закон всеобщей связи и зависимости между явлениями. Естественно, что, исследуя явления в самых различных областях, мы неизбежно сталкиваемся с зависимостями как между количественными, так и между качественными показателями, признаками. Наша задача – обнаружить (выявить) такие зависимости и дать им количественную характеристику.



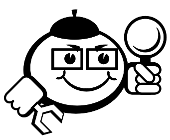
Среди взаимосвязанных признаков (показателей) одни могут рассматриваться как определенные факторы, влияющие на изменение других (*факторные*), а вторые (*результативные*) – как следствие, результат влияния первых.

Существуют два вида связи между отдельными признаками: функциональная и стохастическая (статистическая), частным случаем которой является корреляционная.



Связь между двумя переменными x и y называется ***функциональной****, если определенному значению переменной x строго соответствует одно или несколько значений другой переменной y, и с изменением значения x значение y меняется строго определенно.*

Такие связи обычно встречаются в точных науках.



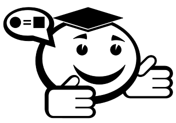
Пример 6.1

Например, известно, что площадь квадрата равна квадрату его стороны $\left(S=a^2\right)$. Это соотношение характерно для каждого единичного случая (квадрата), это так называемая *жестко детерминированная* связь. Такие связи можно встретить и в области экономических явлений.

Например, при простой сдельной оплате труда связь между оплатой труда y и количеством изготовленных изделий x при фиксированной расценке за одну деталь, например 5 руб., легко выразить формулой $y = 5x$.

Для изучения функциональных связей применяется *индексный метод*.

Существуют и иного рода связи, где взаимно действуют многие факторы, комбинация которых приводит к вариации значений результативного признака (показателя) при одинаковом значении факторного признака. Например, при изучении зависимости величины таможенных платежей, поступающих в федеральный бюджет, от количества товаров, перемещаемых через таможенную границу государства, (или от стоимостного товарооборота) последние будут рассматриваться как факторный признак, а величина таможенных платежей – как результативный. Между ними нет жестко детерминированной связи, т. е. при одном и том же количестве перемещенных через таможенную границу товаров (или стоимости товарооборота) величина таможенных платежей, перечисленных разными таможнями будет различной, так как кроме количества товаров, перемещаемых через таможенную границу государства, (или стоимость товарооборота) на величину таможенных платежей влияет много других факторов (различная номенклатура товаров, для которых применяются различные таможенные пошлины, сборы и льготы; различные таможенные режимы перемещения товаров через таможенную границу и др.), комбинация которых вызывает вариацию величины таможенных платежей.

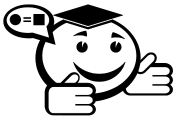


Там, где взаимодействует множество факторов, в том числе и случайных, выявить зависимости, рассматривая единичный случай, невозможно. Такие связи можно обнаружить только при массовом наблюдении как статистические закономерности[1]. *Выявленная таким образом связь именуется* ***стохастической***[2].

[1]Проявление стохастических связей подвержено *действию закона больших чисел*: лишь в достаточно большом числе единиц индивидуальные особенности сгладятся, случайности взаимопогасятся и зависимость, если она имеет существенную силу, проявится достаточно отчетливо.

[2]Термин «стохастический» происходит от гр. *stochos* – мишень. Стреляя в мишень, даже хороший стрелок редко попадает в ее центр, выстрелы ложатся в некоторой близости от него. Другими словами стохастическая связь означает приблизительный характер значений признака.

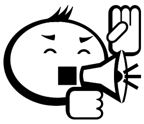
Корреляционная связь[1] – понятие более узкое, чем стохастическая связь, это ее частный случай.



***Корреляционная связь** – это связь, проявляющаяся при большом числе наблюдений в виде определенной зависимости между средним значением результативного признака и признаками-факторами.*

Другими словами, корреляционную связь условно можно рассматривать как своего рода функциональную связь средней величины одного признака (результативного) со значением другого (или других) [25]. При этом, если рассматривается связь средней величины результативного показателя y с одним признаком-фактором x, корреляция называется *парной*, а если факторных признаков 2 и более $(x_1, x_2, ..., x_m)$ – *множественной*[2].

По характеру изменений x и y в парной корреляции различается *прямая* и *обратная* связь. При прямой связи значения обоих признаков изменяются в одном направлении, т. е. с увеличением (уменьшением) значений x увеличиваются (уменьшаются) и значения y. При обратной связи значения факторного и результативного признаков изменяются в разных направлениях.



Изучение корреляционных связей сводится в основном к решению следующих задач[3]:

1) выявление наличия (отсутствия) корреляционной связи между изучаемыми признаками;
2) измерение тесноты связи между двумя (и более) признаками с помощью специальных коэффициентов (эта часть исследования именуется корреляционным анализом);
3) определение уравнения регрессии – математической модели, в которой среднее значение результативного признака y рассматривается как функция одной или несколь-

---

[1]Термин «корреляция» ввел в статистику английский биолог и статистик Ф. Гальтон в конце XIX в., под ним понималась «как бы связь», т. е. связь в форме, отличающейся от функциональной. Еще ранее этот термин применил француз Ж. Кювье в палеонтологии, где под законом корреляции частей животных он понимал возможность восстановить по найденным в раскопках частям облик всего животного.

[2]Множественная корреляция изучается в курсе эконометрики на основе применения компьютерных программ (напр., специальная надстройка к Excel, SPSS и др.), в курсе статистики изучается только парная корреляция.

[3]Теория статистики : учебник для вузов / под ред. Р. А. Шмойловой. – 4-е изд., доп., перераб. – М. : Финансы и статистика, 2007.

ких переменных – факторных признаков (эта часть исследования именуется регрессионным анализом).

Общий термин *корреляционно-регрессионный анализ* подразумевает всестороннее исследование корреляционных связей (т. е. решение всех трех задач).

Корреляционно-регрессионный анализ находит широкое применение в экономике. Рассмотрим его практическое применение на примере данных таможенной статистики внешней торговли России в 2006 гг. (табл. 6.1).

Пример 6.2

Таблица 6.1 – Величина внешнеторгового оборота и таможенных платежей

Месяц	**Оборот, млрд долл.**	**Платеж, млрд руб.**
Январь	27,068	172,17
Февраль	29,889	200,90
Март	34,444	231,83
Апрель	33,158	232,10
Май	37,755	233,40
Июнь	37,554	236,99
Июль	37,299	246,53
Август	40,370	253,62
Сентябрь	37,909	256,43
Октябрь	38,348	261,89
Ноябрь	39,137	259,36
Декабрь	46,298	278,87

В качестве факторного признака x примем стоимостной внешнеторговый товарооборот в млрд долл. США, а в качестве результативного признака y – величину таможенных платежей в федеральный бюджет в млрд руб.

6.2 Методы выявления и оценки корреляционной связи

Для выявления наличия и характера корреляционной связи между *двумя* признаками используется ряд методов.

1. *Рассмотрение параллельных данных* (значений x и y в каждой из n единиц). Единицы наблюдения необходимо расположить по возрастанию значений факторного признака x как в таблице слева и затем сравнить с ним (визуально) поведение результативного признака y.

В нашей задаче в 6 случаях по мере увеличения значений x увеличиваются и значения y, а в 5 случаях этого не происходит, поэтому затруднительно говорить о прямой связи между x и y.

2. *Графический метод* – это графическое изображение корреляционной зависимости. Для этого, имея n взаимосвязанных пар значений x и y и пользуясь прямоугольной системой координат, каждую такую пару изображают в виде точки на плоскости с координатами x и y. Совокупность полученных точек представляет собой *корреляционное поле* (рис. 6.1).

Пример 6.3

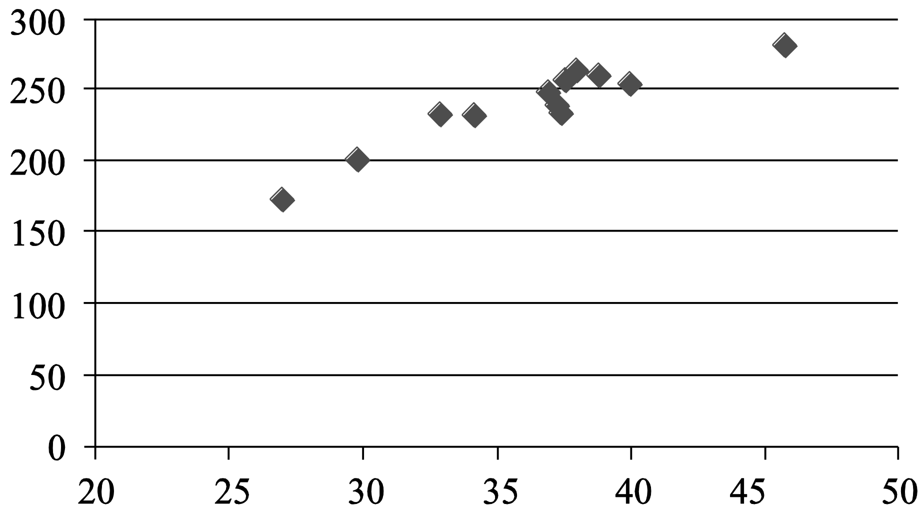


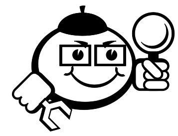
Рис. 6.1 – Корреляционное поле

Визуально анализируя график, можно предположить характер зависимости между признаками x и y. В нашей задаче можно выдвинуть гипотезу о наличии прямой зависимости между величиной стоимостного внешнеторгового товарооборота и величиной таможенных платежей в федеральный бюджет.

3. *Коэффициент корреляции знаков (Фехнера)* – простейший показатель тесноты связи, основанный на сравнении поведения отклонений индивидуальных значений каждого признака $(x$ и $y)$ от своей средней величины. При этом во внимание принимаются не величины отклонений $(x_i - \bar{x})$ и $(y_i - \bar{y})$, а их знаки («+» или «–»). Определив знаки отклонений от средней величины в каждом ряду, рассматривают все пары знаков и подсчитывают число их совпадений (C) и несовпадений (H). Тогда коэффициент Фехнера рассчитывается как отношение разности чисел пар совпадений и несовпадений знаков к их сумме, т. е. к общему числу наблюдаемых единиц:

$$K_{\Phi} = \frac{\sum C - \sum H}{\sum C + \sum H}.$$

Очевидно, что если знаки всех отклонений по каждому признаку совпадут, то $K_{\Phi} = 1$, что характеризует наличие прямой связи. Если все знаки не совпадут, то $K_{\Phi} = -1$ (обратная связь). Если же $\sum C = H$, то $K_{\Phi} = 0$. Итак, как и любой показатель тесноты связи, коэффициент Фехнера может принимать значения от 0 до ± 1. Однако, если $K_{\Phi} = 1$, то это ни в коей мере нельзя воспринимать как свидетельство функциональной зависимости между x и y.



Пример 6.4

Средние значения факторного и результативного признаков определяем по формуле средней арифметической простой:

$$\bar{x} = \frac{\sum x}{n} = \frac{439,229}{12} = 36,602;$$

$$\bar{y} = \frac{\sum y}{n} = \frac{2864,09}{12} = 238,674.$$

В двух последних столбцах таблицы 6.2 приведены знаки отклонений каждого x и y от своей средней величины. Число совпадений знаков – 10, а несовпадений – 2, тогда определяем коэффициент корреляции знаков (Фехнера):

$$K_{\Phi} = \frac{10-2}{10+2} = \frac{8}{12} = 0,667.$$

Таблица 6.2 – Вспомогательная таблица для расчета коэффициента Фехнера

№ п/п	x	y	$x-\bar{x}$	$y-\bar{y}$
1	27,068	172,17	–	–
2	29,889	200,90	–	–

№ п/п	x	y	$x-\bar{x}$	$y-\bar{y}$
3	33,158	232,10	–	–
4	34,444	231,83	–	–
5	37,299	246,53	+	+
6	37,554	236,99	+	–
7	37,755	233,40	+	–
8	37,909	256,43	+	+
9	38,348	261,89	+	+
10	39,137	259,36	+	+
11	40,370	253,62	+	+
12	46,298	278,87	+	+
Итого	439,229	2 864,09		

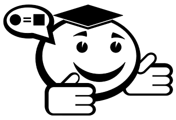
. . . .

Обычно такое значение показателя тесноты связи характеризует заметную прямую зависимость между x и y, однако следует иметь в виду, что поскольку K_{Φ} зависит только от знаков и не учитывает величину самих отклонений x и y от их средних величин, то он практически характеризует не столько тесноту связи, сколько ее наличие и направление.

4. *Линейный коэффициент корреляции* – самый распространенный измеритель линейной зависимости между двумя количественными признаками x и y. Он основан на предположении, что при *полной независимости признаков* x и y отклонения значений факторного признака от средней $(x-\bar{x})$ носят случайный характер и должны случайно сочетаться с различными отклонениями $(y-\bar{y})$. При наличии значительного перевеса совпадений или несовпадений таких отклонений делается предположение о наличии связи между x и y [15].

В отличие от K_{Φ} в линейном коэффициенте корреляции учитываются не только знаки отклонений от средних величин, но и значения самих отклонений, выраженные для сопоставимости в единицах среднего квадратического отклонения t:

$$t_x = \frac{x-\bar{x}}{\sigma_x} \text{ и } t_y = \frac{y-\bar{y}}{\sigma_y}.$$



. . . .

Линейный коэффициент корреляции r *представляет собой среднюю величину из произведений нормированных отклонений для* x и y:

$$r = \frac{\sum\left(\frac{x-\bar{x}}{\sigma_x}\right)\left(\frac{y-\bar{y}}{\sigma_y}\right)}{n} = \frac{\sum t_x t_y}{n},$$

или

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{n\sigma_x\sigma_y}.$$

. . . .

Числитель формулы, деленный, представляющий собой среднее произведение отклонений значений двух признаков от их средних значений, называется *коэффициентом ковариации* – это мера совместной вариации факторного x и результативного y признаков:

$$\operatorname{cov}(x,y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{n} = \overline{(x-\bar{y})(y-\bar{y})}.$$

Недостатком коэффициента ковариации является то, что он не нормирован, в отличие от линейного коэффициента корреляции. Очевидно, что линейный коэффициент корреляции представляет собой частное от деления ковариации между x и y на произведение их средних квадратических отклонений:

$$r = \frac{\operatorname{cov}(x,y)}{\sigma_x\sigma_y}.$$

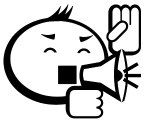
Путем несложных математических преобразований можно получить и другие модификации формулы линейного коэффициента корреляции, например:

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x\sigma_y},$$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}},$$

$$r = \frac{n\sum xy - \sum x\sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}},$$

$$r = \frac{\sum xy - \sum x \frac{\sum y}{n}}{\sqrt{\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]\left[\sum y^2 - \frac{\left(\sum y\right)^2}{n}\right]}}.$$



Линейный коэффициент корреляции может принимать значения от –1 до +1, причем знак определяется в ходе решения.

Если $r > 0,$ то зависимость между x и y прямая.

Если $r < 0,$ то зависимость между x и y обратная.

Если $r = 0,$ то линейная зависимость между x и y отсутствует.

При $r = 1$ существует функциональная зависимость между x и y.

Следовательно, всякое промежуточное значение r от 0 до 1 характеризует степень приближения корреляционной связи между x и y к функциональной. Существует эмпирическое правило (шкала Чэддока) для оценки тесноты связи, представленное в таблице 6.3.

Таблица 6.3 – Шкала Чэддока

$\lvert r \rvert$	**Теснота связи**
Менее 0,1	Отсутствует линейная связь
0,1 ÷ 0,3	Слабая
0,3 ÷ 0,5	Умеренная
0,5 ÷ 0,7	Заметная
Более 0,7	Сильная (тесная)

Таким образом, коэффициент корреляции при линейной зависимости служит как мерой тесноты связи, так и показателем, характеризующим степень приближения корреляционной зависимости между x и y к линейной. Поэтому близость значения r к 0 в одних случаях может означать отсутствие связи между x и $y,$ а в других свидетельствовать о том, что зависимость не линейная.

В нашей задаче для расчета r построим вспомогательную таблицу 6.4.

Пример 6.5

Таблица 6.4 – Вспомогательные расчеты линейного коэффициента корреляции

№ п/п	x	y	$(x-\bar{x})^2$	$(y-\bar{y})^2$	t_x	t_y	$t_x \cdot t_y$	$(x-\bar{x})\times$ $\times(y-\bar{y})$	xy
1	27,068	172,17	90,905	4 422,804	–1,993	–2,408	4,799	634,078	4 660,298
2	29,889	200,90	45,070	1 426,888	–1,403	–1,368	1,919	253,594	6 004,700
3	33,158	232,10	11,864	43,220	–0,720	–0,238	0,171	22,644	7 695,972
4	34,444	231,83	4,659	46,843	–0,451	–0,248	0,112	14,773	7 985,153
5	37,299	246,53	0,485	61,714	0,146	0,284	0,041	5,472	9 195,322
6	37,554	236,99	0,906	2,836	0,199	–0,061	–0,012	–1,603	8 899,922
7	37,755	233,40	1,328	27,817	0,241	–0,191	–0,046	–6,079	8 812,017
8	37,909	256,43	1,707	315,270	0,273	0,643	0,176	23,199	9 721,005
9	38,348	261,89	3,047	538,975	0,365	0,841	0,307	40,525	10 042,958
10	39,137	259,36	6,424	427,904	0,530	0,749	0,397	52,430	10 150,572
11	40,37	253,62	14,195	223,378	0,788	0,541	0,426	56,310	10 238,639
12	46,298	278,87	94,004	1 615,705	2,027	1,455	2,950	389,722	12 911,123
Итого	439,229	2 864,09	274,594	9 153,353	–	–	11,241	1 485,066	106 317,681

$$\sigma_x = \sqrt{274{,}594/12} = 4{,}784;$$

$$\sigma_y = \sqrt{9\ 153{,}353/12} = 27{,}618.$$

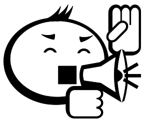
Тогда линейный коэффициент корреляции:

$$r = 11{,}241/12 = 0{,}937.$$

$$r = 1\ 485{,}066/(12 \cdot 4{,}784 \cdot 27{,}618) = 0{,}937.$$

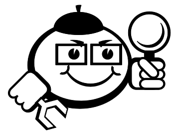
$$r = (106\ 317{,}681/12 - 36{,}602 \cdot 238{,}674)/(4{,}784 \cdot 27{,}618) = 0{,}937.$$

Найденное значение свидетельствует о том, что связь между величиной стоимостного внешнеторгового товарооборота и величиной таможенных платежей в федеральный бюджет очень близка к функциональной (сильная по шкале Чэддока).



Проверка коэффициента корреляции на значимость (существенность).

Интерпретируя значение коэффициента корреляции, следует иметь в виду, что он рассчитан для ограниченного числа наблюдений и подвержен случайным колебаниям, как и сами значения x и y, на основе которых он рассчитан. Другими словами, как любой выборочный показатель, он содержит случайную ошибку и не всегда однозначно отражает действительно реальную связь между изучаемыми показателями. Для того чтобы оценить существенность (значимость) самого r и, соответственно, реальность измеряемой связи между x и y, необходимо рассчитать среднюю квадратическую ошибку коэффициента корреляции σ_r. Оценка существенности (значимости) r основана на сопоставлении значения r с его средней квадратической ошибкой: $\frac{|r|}{\sigma_r}$.



Пример 6.6

Существуют некоторые особенности расчета σ_r в зависимости от числа наблюдений (объема выборки) – *n*.

1. Если число наблюдений достаточно велико $(n > 30)$, то σ_r рассчитывается по формуле:

$$\sigma_r = \frac{1 - r^2}{\sqrt{n}}.$$

Обычно, если $\frac{|r|}{\sigma_r} > 3$, то r считается значимым (существенным), а связь – реальной. Задавшись определенной вероятностью, можно определить *доверительные пределы (границы)* $r = (r \pm t\sigma_r)$, где t – коэффициент доверия, рассчитываемый по интегралу Лапласа (см. прил. А).

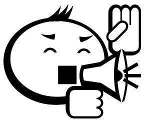
2. Если число наблюдений небольшое $(n < 30)$, то σ_r рассчитывается по формуле:

$$\sigma_r = \frac{\sqrt{1 - r^2}}{\sqrt{n - 2}},$$

а значимость r проверяется на основе t-критерия Стьюдента, для чего определяется расчетное значение критерия по формуле и сопоставляется с $t_{\text{табл}}$.

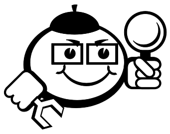
$$t_{\text{расч}} = \frac{|r|}{\sigma_r} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

Табличное значение $t_{\text{табл}}$ находится по таблице распределения t-критерия Стьюдента (см. прил. Б) при уровне значимости $\alpha = 1 - \beta$ и числе степеней свободы $\nu = n - 2$.



Если $t_{\text{расч}} > t_{\text{табл}}$, то r считается значимым, а связь между x и y – реальной.

В противном случае $\left(t_{\text{расч}} < t_{\text{табл}}\right)$ считается, что связь между x и y отсутствует, и значение r, отличное от нуля, получено случайно.



Пример 6.7

В нашей задаче число наблюдений небольшое, значит, оценивать существенность (значимость) линейного коэффициента корреляции будем по следующей формуле:

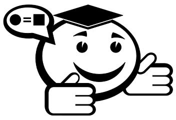
$$\sigma_r = \frac{\sqrt{1-0{,}937^2}}{\sqrt{12-2}} = 0{,}349/3{,}162 = 0{,}110;$$

$$t_{\text{расч}} = \frac{|r|}{\sigma_r} = 0{,}937/0{,}110 = 8{,}482.$$

Из прил. Б видно, что при числе степеней свободы $\nu = 12 - 2 = 10$ (в 10-й строке) и вероятности $\beta = 95\%$ (уровень значимости $\alpha = 1 - \beta = 0{,}05$) $t_{\text{табл}} = 2{,}2281$, а при вероятности 99% $(\alpha = 0{,}01)$ $t_{\text{табл}} = 3{,}169$, значит, $t_{\text{расч}} > t_{\text{табл}}$, что дает возможность считать линейный коэффициент корреляции $r = 0{,}937$ значимым.

5. *Подбор уравнения регрессии*[1].

[1]Термин «регрессия» ввел в статистику Ф. Гальтон, который, изучив большое число семей, установил, что в группе семей с высокорослыми отцами сыновья в среднем ниже ростом, чем их отцы, а в группе семей с низкорослыми отцами сыновья в среднем выше отцов, т. е. отклонение роста от среднего в следующем поколении уменьшается – регрессирует.



. . . .

Уравнение регрессии *представляет собой математическое описание изменения взаимно коррелируемых величин по эмпирическим (фактическим) данным.*

. . . .

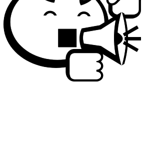
Уравнение регрессии должно определить, каким будет среднее значение результативного признака y при том или ином значении факторного признака x, если остальные факторы, влияющие на y и не связанные с x, не учитывать, т. е. абстрагироваться от них. Другими словами, уравнение регрессии можно рассматривать как вероятностную гипотетическую функциональную связь величины результативного признака y со значениями факторного признака x.

Уравнение регрессии можно также назвать *теоретической линией регрессии.* Рассчитанные по уравнению регрессии значения результативного признака называются *теоретическими*. Они обычно обозначаются $\hat{y}_x$ или $\bar{y}_x$ (читается: «игрек, выравненный по икс») и рассматриваются как функция от x, т. е. $\hat{y}_x = f(x)$.

Найти в каждом конкретном случае тип функции, с помощью которой можно наиболее адекватно отразить ту или иную зависимость между признаками x и y, – одна из основных задач регрессионного анализа. Выбор теоретической линии регрессии часто обусловлен формой эмпирической линии регрессии; теоретическая линия как бы сглаживает изломы эмпирической линии регрессии. Кроме того, необходимо учитывать природу изучаемых показателей и специфику их взаимосвязей [28].

Для аналитической связи между x и y могут использоваться виды уравнений, приведенные в таблице 6.4 (при условии замены t на x). Обычно зависимость, выражаемую уравнением прямой, называют *линейной* (или *прямолинейной*), а все остальные – *криволинейными зависимостями*.

Выбрав тип функции, по эмпирическим данным определяют параметры уравнения. При этом отыскиваемые параметры должны быть такими, при которых рассчитанные по уравнению теоретические значения результативного признака $\hat{y}_x$ были бы максимально близки к эмпирическим данным.



. . . .

Существует несколько методов нахождения параметров уравнения регрессии. Наиболее часто используется *метод наименьших квадратов* (МНК). Его суть заключается в следующем требовании:

искомые теоретические значения результативного признака $\hat{y}_x$ должны быть такими, при которых бы обеспечивалась минимальная сумма квадратов их отклонений от эмпирических значений, т. е.

$$S = \sum\left(y - \hat{y}_x\right)^2 \to \min.$$

Поставив данное условие, легко определить, при каких значениях a_0, a_1 и т. д. для каждой аналитической кривой эта сумма квадратов отклонений будет минимальной. Данный метод уже использовался нами в п. 4.3., поэтому, для нахождения параметров теоретической линии регрессии, заменив параметр t на x, получим:

$$\begin{cases} na_0 + a_1\sum x = \sum y \\ a_0\sum x + a_1\sum x^2 = \sum xy. \end{cases}$$

Выразив из первого уравнения системы a_0, получим:

$$a_0 = \frac{\sum y}{n} - a_1\frac{\sum x}{n} = \bar{y} - a_1\bar{x}.$$

Подставив во второе уравнение системы, затем разделив обе его части на n, получим:

$$\left(\bar{y} - a_1\bar{x}\right)\frac{\sum x}{n} + a_1\frac{\sum x^2}{n} = \frac{\sum xy}{n}.$$

Применяя 3 раза формулу средней арифметической, получим:

$$\left(\bar{y} - a_1\bar{x}\right)\bar{x} + a_1\overline{x^2} = \overline{xy}.$$

Раскрыв скобки и перенеся члены без a_1 в правую часть уравнения, выразим a_1:

$$a_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x^2}.$$

Параметр a_1 в уравнении линейной регрессии называется *коэффициентом регрессии*, который показывает, на сколько изменяется значение результативного признака y при изменении факторного признака x на единицу.

Пример 6.8

Исходные данные и расчеты для нашего примера представим в таблице 6.5.

Таблица 6.5 – Вспомогательные расчеты для нахождения уравнения регрессии

№ п/п	x	y	x^2	xy	$\hat{y}_x$	$(y-\hat{y}_x)^2$	$(\hat{y}_x-\bar{y})^2$
1	27,068	172,17	732,677	4 660,298	187,124	223,612	2 657,453
2	29,889	200,90	893,352	6 004,700	202,377	2,181	1 317,497
3	33,158	232,10	1 099,453	7 695,972	220,052	145,147	346,774
4	34,444	231,83	1 186,389	7 985,153	227,006	23,274	136,153
5	37,299	246,53	1 391,215	9 195,322	242,443	16,706	14,202
6	37,554	236,99	1 410,303	8 899,922	243,821	46,669	26,495
7	37,755	233,40	1 425,440	8 812,017	244,908	132,441	38,864
8	37,909	256,43	1 437,092	9 721,005	245,741	114,256	49,940
9	38,348	261,89	1 470,569	10 042,958	248,115	189,761	89,122
10	39,137	259,36	1 531,705	10 150,572	252,381	48,710	187,871
11	40,370	253,62	1 629,737	10 238,639	259,048	29,459	415,076
12	46,298	278,87	2 143,505	12 911,123	291,100	149,580	2 748,498
Итого	439,229	2 864,09	16 351,437	106 317,681	2 864,115	1 121,795	8 027,945

$$a_1=\frac{106\ 317,681/12-36,602\cdot 238,674}{4,784^2}=5,407.$$

$$a_0=238,674-5,407\cdot 36,602=40,767.$$

Отсюда получаем уравнение регрессии:

$$\hat{y}_x=40,72+5,4082x,$$

подставляя в которое вместо x эмпирические значения факторного признака (2-й столбец таблицы 6.5), получаем выравненные по прямой линии теоретические значения результативного признака $\hat{y}_x$ (6-й столбец таблицы 6.5). Для иллюстрации различий между эмпирическими и теоретическими линиями регрессии построим график (рис. 6.2).

Из рисунка видно, что небольшие различия между эмпирическими данными и теоретической линиями регрессии существуют, поэтому необходимо *оценить существенность* коэффициента регрессии и уравнения связи, для чего определяют среднюю ошибку параметров уравнения регрессии и сравнивают их с этой ошибкой.

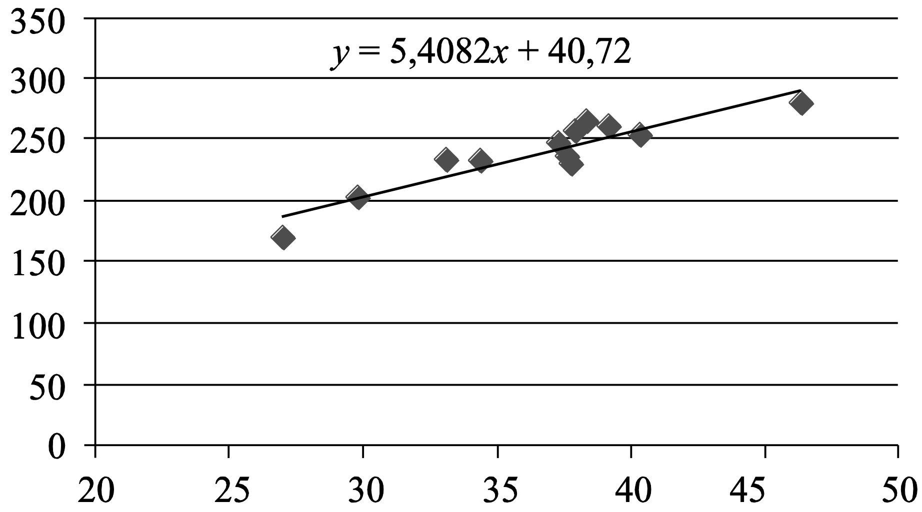


Рис. 6.2 – График эмпирических данных и теоретической линий регрессии

. . . .

Расчет ошибок параметров уравнения регрессии основан на использовании остаточной дисперсии, характеризующей расхождение (отклонение) между эмпирическими и теоретическими значениями результативного признака. Для линейного уравнения регрессии $\left(\hat{y}_x = a_0 + a_1 x\right)$ средние ошибки параметров a_1 и a_2 определяются соответственно:

$$\mu_{a_0} = \frac{\sigma_{\text{ост}}}{\sqrt{n-2}},$$

$$\mu_{a_0} = \frac{\sigma_{\text{ост}}}{\sigma_x \sqrt{n-2}},$$

$$\sigma_{\text{ост}} = \sqrt{\frac{\sum\left(y - \hat{y}_x\right)^2}{n}}.$$

Значимость параметров проверяется путем сопоставления его значения со средней ошибкой. Обозначим это соотношение как t:

$$t_{a_i} = \frac{a_i}{\mu_{a_i}}.$$

При большом числе наблюдений $(n > 30)$ параметр a_i считается значимым, если $t_{a_i} > 3.$

Если выборка малая $(n < 30),$ то значимость параметра a_i проверяется путем сравнения с табличным значения t-критерия Стьюдента при числе степеней свободы $\nu = n - 2$ и заданном уровне значимости α (прил. Б). Если рассчитанное значение t больше табличного, то параметр считается значимым.

Пример 6.9

В нашем примере:

$$\sigma_{\text{ост}} = \sqrt{\frac{1121{,}795}{12}} = 9{,}669.$$

Находим среднюю ошибку параметра a_0:

$$\mu_{a_0} = \frac{9{,}669}{\sqrt{12-2}} = 3{,}06.$$

Теперь находим среднюю ошибку параметра a_1:

$$\mu_{a_0} = \frac{9{,}669}{4{,}784\sqrt{12-2}} = 0{,}639.$$

Теперь для параметра a_0:

$$t_{a_0} = \frac{40{,}767}{3{,}06} = 13{,}3.$$

И по той же формуле для параметра a_1:

$$t_{a_1} = \frac{5{,}407}{0{,}639} = 8{,}46.$$

Так как выборка малая, то задавшись стандартной значимостью $\alpha = 0{,}05$ находим в 10-й строке прил. Б табличное значение $t_a = 2{,}23$, которое значительно меньше полученных значений 13,3 и 8,46, что свидетельствует о значимости обоих параметров уравнения регрессии.

Наряду с проверкой значимости отдельных параметров осуществляется *проверка значимости уравнения регрессии* в целом или, что то же самое, проверка адекватности модели с помощью критерия Фишера по прил. В.

Пример 6.10

Данный метод уже использовался нами для проверки адекватности уравнения тренда, поэтому в нашем примере получим:

$$F_p = \frac{(12-2)8027{,}945}{(2-1)1121{,}795} = 71{,}56.$$

Сравнивая расчетное значение критерия Фишера $F_p = 71{,}56$ с табличным $F_m = 4{,}96$, определяемое по прил. В при числе степеней свободы $\nu_1 = k-1 = 2-1 = 1$ и $\nu_2 = n-k = 12-2 = 10$ (т. е. 1-й столбец и 10-я строка) и стандартном уровне значимости $\alpha = 0,05$, можно сделать вывод, что уравнение регрессии значимо.

6. ***Коэффициент эластичности*** *показывает, на сколько процентов изменяется в среднем результативный признак y при изменении факторного признака x на 1%. Он рассчитывается на основе уравнения регрессии:*

$$Э = \frac{\partial \hat{y}_x}{\partial x} \frac{x}{\hat{y}_x},$$

где $\frac{\partial \hat{y}_x}{\partial x}$ – первая производная уравнения регрессии y по x.

Коэффициент эластичности – величина переменная, т. е. изменяется с изменением значений фактора x. Так, для линейной зависимости $\hat{y}_x = a_0 + a_1 x$:

$$Э = a_1 \frac{x}{a_0 + a_1 x}.$$

Пример 6.11

Применительно к рассмотренному уравнению регрессии, выражающему зависимость величины таможенных платежей в федеральный бюджет от величины стоимостного внешнеторгового оборота $(\hat{y}_x = 40,767 + 5,407x)$, коэффициент эластичности:

$$Э = \frac{5,407x}{40,767 + 5,407x}.$$

Подставляя в данное выражение разные значения x, получаем и разные значения Э. Так, например, при $x = 40$ коэффициент эластичности

$$Э = \frac{5,407 \cdot 40}{40,767 + 5,407 \cdot 40} = 0,84,$$

а при $x = 50$ соответственно

$$Э = \frac{5,407 \cdot 50}{40,767 + 5,407 \cdot 50} = 0,87 \text{ и т. д.}$$

Это значит, что при увеличении внешнеторгового товарооборота x с 40 до 40,4 млрд долл. (т. е. на 1%) величина таможенных платежей возрастет в среднем на 0,84% прежнего уровня; при увеличении x с 50 до 50,5 млрд долл. (т. е. на 1%) y возрастет на 0,87% и т. д.

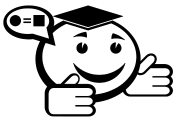
Контрольные вопросы по главе 6

1. Что такое функциональная зависимость?
2. Что такое стохастическая зависимость?
3. Что такое корреляционная связь?
4. Как определяется линейный коэффициент корреляции?
5. Что такое уравнение регрессии?
6. В чем суть метода наименьших квадратов?
7. Как оценить адекватность параметров уравнения регрессии?

7 Модели управления запасами

7.1 Основные понятия теории управления запасами и ее элементы

Одним из важнейших этапов планирования работы любой производственной единицы – цеха, предприятия или объединения предприятий – является определение рационального уровня запасов сырья, полуфабрикатов, инструментов. Основными причинами создания производственных запасов служат необходимость обеспечения бесперебойного снабжения производственного процесса, периодичность производства различных видов продукции поставщиками, осуществление транспортировки большинства видов продукции от поставщика к потребителю партиями, а также несовпадение ритма производства с ритмом потребления.



Предметом теории управления запасами *является отыскание такой организации поставок или производства, при которых суммарные затраты на функционирование системы были бы минимальными.*

Под *организацией поставок* понимается определение объемов поставок и периодичность заказов, а при планировании производства нескольких видов продукции на одном и том же оборудовании – определение размера партии и периодичности запуска продукции в производство. Существуют четыре основных вида затрат, которые могут оказать влияние на выбор решения по управлению запасами:

- затраты на приобретение запасов,
- затраты на организацию заказа,
- издержки хранения запасов,
- потери от дефицита [26].

Запасы делятся:

1) на текущие (обеспечивают ритм производства на определенном интервале времени);
2) страховые (на случай срыва ритма поставок).

Затраты, которые не зависят от принимаемых решений, не учитываются при анализе. Так, *затраты на приобретение продукции* целесообразно учитывать, только если цена единицы продукции зависит от величины партии, что обычно выражается в виде оптовых скидок.

К затратам на организацию заказа, учитываемым в анализе функционирования систем управления запасами, относят постоянные расходы по размещению заказов: расходы на разъезды и командировки, почтово-телеграфные расходы, транспортные расходы, не зависящие от размера партии.

В общем случае в стоимости поставки, кроме постоянных, входят затраты, пропорциональные объему партии и количеству заказываемых номенклатур. Однако принимаемые решения никак не влияют на величину затрат, пропорциональных размеру партии. Затраты, пропорциональные количеству номенклатур в заказе, учитываются только в многономенклатурных (многопродуктовых) моделях. Эти затраты представляют собой стоимостное выражение трудозатрат, связанных с поиском и обработкой информации по отдельным продуктам, упаковкой у поставщика, а также приемом и размещением на складе потребителя.

Если складскую систему снабжает предприятие-поставщик, то при условии серийного выпуска продукции стоимость переналадки оборудования перед выпуском очередной партии тоже попадает в эту категорию затрат. Иногда сюда относят также издержки вследствие более низкой производительности труда и более высокого процента брака в начале производственного периода. Затраты, связанные с началом выпуска очередной партии, называют *затратами на подготовительно-заключительные операции*.

К *издержкам хранения запасов*, учитываемым в моделях управления запасами, относятся лишь издержки, зависящие от величины запасов. К ним относятся издержки физического присутствия материальных ценностей на складе (естественная убыль, плата за производственные фонды) и потери от иммобилизации средств в запасах. Если рассматривать средства, вложенные в запасы как банковскую ссуду, то издержки задаются процентной ставкой [27].

Потери от дефицита на промышленных предприятиях исчисляются как суммарные потери прибыли в расчете на одну денежную единицу стоимости дефицитных материалов. Прибыль предприятия при дефиците может снизиться за счет простоя производственных мощностей и рабочих, переналадки производственного процесса, замены дефицитных материалов другими, более дорогими, выпуск продукции в сверхурочное время после ликвидации причины простоя, штраф за нарушение сроков поставки.

Элементы задачи управления запасами:

- система снабжения;
- спрос на предметы снабжения;
- возможность восполнения запасов;
- функции затрат;
- ограничения;
- стратегия управления запасами.

Из параметров управления запасами принято выделять:

1) управляемые параметры
   - объем и номенклатура необходимого сырья (ресурсов);
   - момент (время) выдачи заказа на пополнение ресурса;
2) неуправляемые параметры
   - затраты на организацию снабжения;
   - ограничение на запасы поставщика;
   - выбор системы снабжения (централизованная, децентрализованная).

Многообразие реальных ситуаций вызвало необходимость разработки разнообразных моделей управления запасами. Основным фактором, влияющим на тип модели, является характер спроса или потребности в материальных ресурсах. Спрос может быть детерминированным или вероятностным. В свою очередь детерминированный спрос может быть статическим, неизменным во времени, или динамическим, изменяющимся во времени. Вероятностный спрос может быть стационарным, с неизменной во времени плотностью вероятности, и нестационарным с изменяющейся во времени плотностью вероятности.

Другим важным фактором, учитывающимся при построении модели, является срок выполнения заказа, т. е. интервал времени между моментом размещения заказа и его поставкой. Если этот фактор учитывается, то модель называется моделью с запаздыванием поставок.

В модели может быть учтена интенсивность поставок. При пополнении запасов из внешнего источника обычно доставляется вся партия одновременно. Пополнение запаса с некоторой интенсивностью чаще осуществляется самим предприятием, когда продукция одного цеха используется другим.

Число видов продукции учитывается в модели при условии наличия взаимосвязи между ними. Связь может возникать до поставки и после нее. Взаимодействие до поставки проявляется в снабжении из одного источника (заказ на несколько партий различных видов продукции подается одновременно), в требовании комплектности, в ограниченной мощности оборудования. Взаимодействие после поставки имеет место, когда несколько видов продукции хранится в одном складском помещении или ограничена величина оборотных средств, вложенных в запасы.

В работе системы может допускаться дефицит или, наоборот, выдвигаться требование бездефицитной работы.

Системы управления запасами классифицируют по многим признакам:

- вид запасов (сырье, полуфабрикаты, готовая продукция, инструменты, запчасти);
- место хранения (производитель, потребитель, снабженческая база или другие элементы товаропроводящей сети);
- структура системы (изолированный склад, последовательная система складов, иерархическая система, с ремонтными возможностями или без них);
- свойства запасов (одно- или многономенклатурные запасы, их взаимозаменяемость, ограниченность срока годности, порча при хранении);
- статистические характеристики процессов спроса и поставок (стационарность, коррелированность спроса, управляемость, случайность поставок);
- цели системы (стоимостные вероятностные критерии, многокритериальность);
- ограничения (на объем и номенклатуру запасов, размеры партий, надежность и экономические характеристики процесса снабжения);
- информационные характеристики (периодичность сброса данных, наблюдаемость спроса, полнота знаний о коэффициентах потерь).

Стратегия управления запасами

Оптимальное управление запасами – выбор таких объемов и моментов поставок, когда суммарные издержки на функционирование системы снабжения будут минимальными [29].

Простейшие стратегии:

1) периодические (со временем контроля T);

2) по критическим уровням $(H, h, y_i$ – текущий уровень запаса $q)$.

1. ***Стратегия постоянного уровня***.

В данном случае через каждый интервал контроля T запас пополняется до верхнего уровня (рис. 7.1).

$q_1 \neq q_2 \neq q_3 \neq \text{const}$;

$q^*_{\text{опт}} = H - y_{\text{тек}}$;

$y_{1,\,2}$ – текущие уровни.

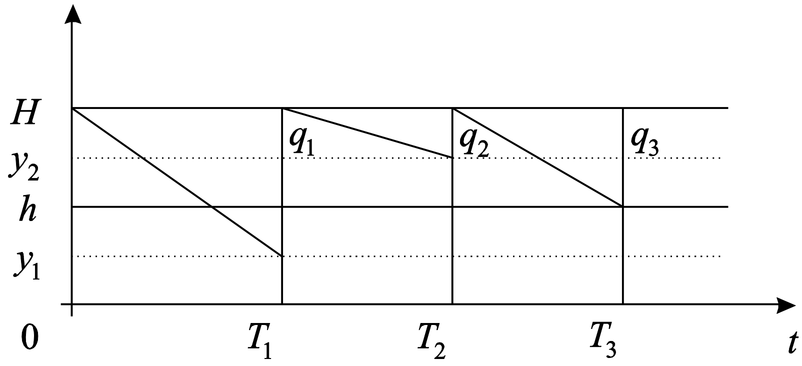


Рис. 7.1 – Стратегия постоянного уровня

2. ***Стратегия фиксированного объема поставок***.

Здесь $Q^* = \text{const}$;

$q_1 = q_2 = q_3 = \text{const}$ (рис. 7.2).

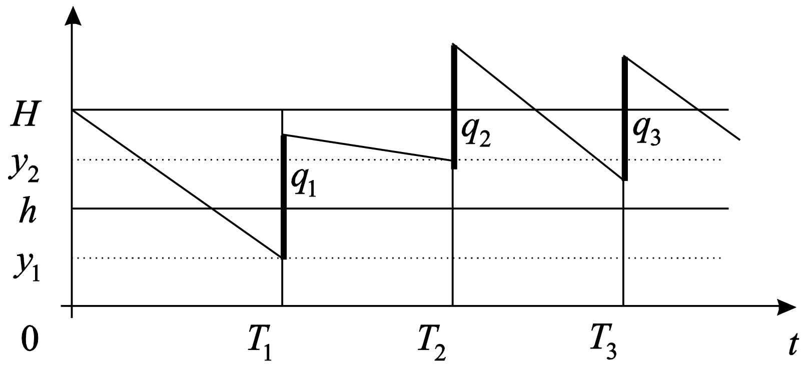


Рис. 7.2 – Стратегия фиксированного объема поставок

3. ***Стратегия с контролем за текущим уровнем***.

а) если $y < h,$ то: $-y < h \Rightarrow q^* = \text{const}$;

$-y \geq h \Rightarrow q^* = 0$ (не заказываем сырье);

б) если $y < h,$ то $-y < h \Rightarrow q^* = H - y_{\text{тек}}$;

$$-y \geq h \Rightarrow q^* = 0.$$

7.2 Классификация моделей управления запасами

Многообразие реальных ситуаций вызвало необходимость в рассмотрении огромного числа вариантов задачи управления запасами, которые систематизированы лишь частично. Использование богатейшего материала, накопленного теорией управления запасами, немыслимо без его упорядочения в рамках единой классификации. Попытки такой классификации и введения унифицированных обозначений предпринимались неоднократно, но оказались малопродуктивными [30].

В некоторых источниках классификация моделей проводится по 45 элементам. Укрупненно она различает модели:

- по числу номенклатур;
- числу складов;
- характеру восполнения;
- характеру спроса;
- способу рассмотрения динамики;
- целевой функции;
- стратегии восполнения;
- способу контроля уровня запаса;
- учету недостач;
- задержки поставок.

Детальная классификация моделей управления запасами имеет реальную полезность лишь при создании для последних компьютерной базы знаний. Тогда можно построить диалоговую систему, опрашивающую пользователя и последовательно формирующую код нужной модели или ближайшей к ней. По коду модели можно войти в базу знаний, найти библиографический источник, посмотреть подробности, лучше уяснить допущения и метод получения результата, при необходимости и способности – модифицировать модель. Затем можно ввести релевантную задаче количественную информацию.

Классификация моделей управления запасами может исходить из основных элементов данной модели, которые упоминались выше. Под системой снабжения понимается совокупность источников заявок и складов, между которыми в ходе операций снабжения осуществляются перевозки хранимого имущества. В состав системы могут входить звенья, обеспечивающие восстановление отказавших у пользователей устройств. В простейшем случае система сводится к единственному складу. Функция затрат составляется и минимизируется для системы в целом.

Обобщенная модель управления запасами выглядит довольно простой. Чем же тогда объясняется столь большое разнообразие моделей этого класса и методов решения соответствующих задач, базирующихся на различном математическом аппарате: от простых схем дифференциального и интегрального исчисления до сложных алгоритмов динамического и других видов математического программирования? Ответ на этот вопрос определяется характером спроса, который может быть детерминированным (достоверно известным) или вероятностным (задаваемым плотностью вероятности). *Детерминированный спрос* может быть *статическим*, в том смысле, что интенсивность потребления остается неизменной во времени, или *динамическим*, когда спрос известен достоверно, но изменяется в зависимости от времени. *Вероятностный спрос* может быть *стационарным*, когда функция плотности вероятности спроса неизменна во времени, и *нестационарным*, когда функция плотности вероятности спроса изменяется во времени.

Типичными примерами нестационарных ситуаций являются торговля сезонными и модными товарами, а также периоды пикового (предпраздничного) спроса [31].

В случае дискретного спроса каждое отдельное требование дополнительно характеризуется своим объемом (числом заказанных единиц). Нестационарный спрос в очередной период может быть зависимым или независимым в смысле связи с предысторией процесса. Детали с зависимым спросом нельзя планировать независимо, поскольку спрос на них будет иметь сильную положительную корреляцию. В случае взаимозаменяемых изделий имеется отрицательная корреляция.

В реальных условиях случай детерминированного статистического спроса встречается редко. Такой случай можно рассматривать как простейший. Так, например, хотя спрос на такие продукты массового потребления, как хлеб, может меняться от одного дня к другому, эти изменения могут быть столь незначительными, что предположение статичности спроса несущественно искажает действительность.

Наиболее точно характер спроса может быть, возможно, описан посредством *вероятностных нестационарных* распределений. Однако с математической точки зрения модель значительно усложняется, особенно при увеличении рассматриваемого периода времени. При переходе от детерминированного статического спроса к вероятностному стационарному спросу происходит возрастание математической сложности модели управления запасами. По существу, классификацию можно считать представлением различных *уровней абстракции* описания спроса.

На первом уровне предполагается, что распределение вероятности спроса стационарно во времени. Это означает, что для описания спроса в течение всех исследуемых периодов времени используется одна и та же функция распределения вероятностей. При таком предположении влияние сезонных колебаний спроса в модели не учитывается.

На втором уровне абстракции учитывается изменение спроса от одного периода к другому. Однако при этом функции распределения не меняются, а потребности в каждом периоде описываются средней величиной спроса. Это упрощение означает, что элемент риска в управлении запасами не учитывается. Однако оно позволяет исследовать сезонные колебания спроса, которые вследствие аналитических и вычислительных трудностей нельзя учесть в вероятностной модели. Другими словами, здесь возникает определенный компромисс: можно использовать, с одной стороны, стационарные распределения вероятностей, а с другой – переменную, но известную функцию спроса при допущении «определённости» [3].

На третьем уровне упрощения исключаются как элементы риска, так и изменения спроса. Тем самым спрос в течение любого периода предполагается равным среднему значению известного (по предположению) спроса по всем рассматриваемым периодам. В результате этого упрощения спрос можно оценить его *постоянной* интенсивностью.

Хотя характер спроса является одним из основных факторов при построении модели управления запасами, имеются другие факторы, влияющие на выбор типа модели. К их числу относятся:

1. *Задержка поставок*. Может увеличиваться в период низкого спроса, когда поставщик накапливает заказы перед запуском производства. Тот же эффект может наблюдаться и при очень высоком спросе, создающем очередь заявок. В некоторых моделях с задержкой, кроме обычной, вводится экспертная поставка, которая, как правило, принимается мгновенной. Возможность такой поставки исключает отрицательные уровни запаса. Может существовать различие в объеме поставок: поставка равна требуемому количеству или поставка равна случайной величине, в общем случае, зависимой от величины заказа. Если слу-

чайность является следствием плохой организации снабжения, необходимо организованными мерами добиваться своевременного и полного выполнения заказов.

2. *Пополнение запасов.* Всегда происходит с некоторой случайной задержкой относительно момента выдачи требования. Однако роль и длина этой задержки сильно зависят от конкретных условий, что позволяет в ряде случаев упростить задачу. Степень возможного упрощения определяет один из следующих вариантов:

- мгновенная поставка;
- задержка поставок на фиксированный срок (в частности, кратный длине периода);
- случайная задержка с известным распределением длительности.

3. *Период времени.* Определяет интервал, в течение которого осуществляется регулирование уровня запаса. В зависимости от отрезка времени, на котором можно надежно прогнозировать, рассматриваемый период принимается конечным или бесконечным.

4. *Число пунктов накопления запаса*. В систему управления запасами может входить несколько пунктов хранения запаса. В некоторых случаях эти пункты организованы таким образом, что один выступает в качестве поставщика для другого. Эта схема иногда реализуется на различных уровнях, так что пункт-потребитель одного уровня может стать пунктом-поставщиком на другом. В таком случае принято говорить о системе управления запасами с разветвленной структурой.

5. *Число видов продукции*. В системе управления запасами может фигурировать более одного вида продукции. Этот фактор учитывается при условии наличия некоторой зависимости между различными видами продукции. Так, для различных изделий может использоваться одно и то же складское помещение или же их производство может осуществляться при ограничениях на общие производственные фонды [31].

Функция затрат образует показатель эффективности принятой стратегии и учитывает следующие издержки:

- расходы на хранение;
- транспортные расходы и затраты, связанные с заказом каждой новой партии;
- затраты на штрафы.

Иногда в минимизируемую функцию включаются доходы, полученные от продажи остатков запаса в конце каждого периода. В некоторых случаях ставится задача максимизации доходов.

Ограничения в задачах управления запасами могут быть различного характера. Известны следующие виды ограничений:

- по максимальному объему (весу, стоимости) запасов;
- по средней стоимости;
- по числу поставок в заданном интервале времени;
- по максимальному объему (весу, стоимости) поставки или кратности этого объема некоторой минимальной величине (целое число стандартных «упаковок» – вагонов, бочек, коробок);
- по доле требований, удовлетворяемых из наличного запаса (без дополнительных задержек).

Необходимо отметить, что область применения теории управления запасами отнюдь не ограничивается складскими операциями. Под запасами можно подразумевать: наличие товара; рабочую силу, планируемую для выполнения конкретного задания; объем информации в базе данных; численность персонала данной квалификации и т. д. Таким образом, при переосмысливании элементов модели методами теории управления запасами может быть решен широкий круг задач оптимального планирования.

7.3 Детерминированные модели

В этом параграфе обсуждаются пять моделей. Большинство из них однопродуктовые, и только в одной из них учитывается влияние нескольких «конкурирующих» видов продукции. Основное различие между моделями определяется допущением о характере спроса (статический или динамический). Важным фактором с точки зрения формулировки и решения задачи является также вид функции затрат. Используются различные методы решения. Эти примеры наглядно показывают, что при решении задач управления запасами следует применять различные методы оптимизации [12].

7.3.1 Модель Уилсона

Рассмотрение моделей управления запасами начнем с простейшего случая.

Модель Уилсона, в определенном смысле классическая, основана на выборе такого фиксированного размера заказываемой партии, который минимизирует расходы на оформление заказа, доставку и хранение товара.

Экономическая партия товара вычисляется при следующих упрощениях реальной ситуации:

- уровень запасов убывает с постоянной интенсивностью, и в тот момент, когда все запасы товара исчерпаны, подается заказ на поставку новой партии;
- выполнение заказа осуществляется мгновенно, т. е. время доставки равно нулю и уровень запасов восстанавливается до значения, равного q;
- накладные расходы, связанные с размещением заказа и поставкой товара, не зависят от объема партии и равны постоянной величине;
- ежедневная стоимость хранения единицы товара равна постоянной величине.

Данная политика, проводимая складом, характерна для тех случаев, когда интенсивность потребления запасов близка к постоянной величине, а поставки производятся регулярно.

Простейшая модель оптимальной партии поставки строится при следующих предложениях: спрос ν в единицу времени является постоянным; заказанная партия доставляется одновременно; дефицит недопустим; затраты K на организацию поставки постоянны и не зависят от величины q партии; издержки содержания единицы продукции в течение единицы времени составляют s. На рисунке 7.3 показана динамика изменения уровня I запасов.

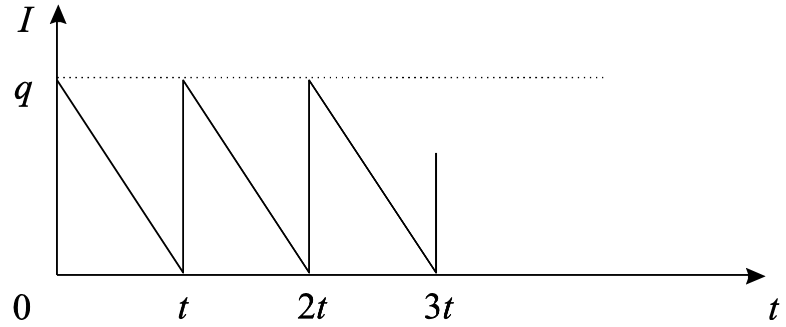


Рис. 7.3 – Динамика изменения уровня запасов

Уровень снижается равномерно от q до 0, после чего подается заказ на доставку новой партии величиной q. Заказ выполняется мгновенно и уровень запаса восстанавливается до величины q.

Интервал времени длиной t между поставками называется циклом.

Издержки в течение цикла $L_ц$ состоят из стоимости доставки заказа K и затрат на содержание запаса ν которые пропорциональны средней величине запаса $I = q/2$ и длине цикла $t = q/\nu$,

$$L_ц = K + s\frac{q}{2}\frac{q}{\nu}.$$

Разделив это выражение на длину цикла $(t = q/\nu)$, получим издержки в единицу времени (L):

$$L = K\frac{\nu}{q} + s\frac{q}{2}.$$

Оптимальный размер партии определяется из уравнения:

$$\frac{dL}{dq} = -\frac{K\nu}{q2} + \frac{s}{2} = 0 \text{ (необходимый признак экстремума).}$$

Отсюда находим оптимальный размер $\boldsymbol{q^*}$ партии:

$$q^* = \sqrt{\frac{2K\nu}{s}}.$$

Так как $d^2L/dq^2 > 0$ (достаточный признак экстремума), то для всех $q > 0$ это выражение является минимумом функции затрат. Это уравнение известно под многими названиями. Его называют *формулой наиболее экономной величины заказа, формулой Уилсона, формулой квадратного корня.*

Чтобы найти оптимальные параметры работы системы, поставляем значение q^* в соответствующие выражение. Получаем, что оптимальная стратегия предусматривает заказ q^* через каждые

$$t^* = \frac{q^*}{\nu} = \sqrt{\frac{2K}{s\nu}}$$

единиц времени. Наименьшие суммарные затраты работы системы в единицу времени:

$$L^* = \sqrt{2Ks\nu} = sq^*.$$

Пример 7.1

Жидкие продукты нескольких видов разливаются в пакеты на одной линии упаковки. Затраты на подготовительно-заключительные операции составляют 700 ден. ед., потребность в продуктах составляет 140 000 л в месяц, стоимость хранения 1 л в течение месяца – 4 ден. ед. Определить оптимальные параметры системы. Сравнить минимальные затраты с затратами при действующей системе разлива одного продукта в течение трех дней.

Решение.

Оптимальные параметры модели Уилсона:

$$q^* = \sqrt{\frac{2 \cdot 700 \cdot 140\ 000}{4}} = 7\ 000 \text{ (л)},$$

$$t^* = \sqrt{\frac{2 \cdot 700}{4 \cdot 140\ 000}} = 0{,}05 \text{ (месяца)} = 1{,}5 \text{ (дня)},$$

$$L^* = \sqrt{2 \cdot 700 \cdot 4 \cdot 140\ 000} = 28\ 000 \text{ (ден. ед.)}.$$

При действующей системе $t_{\text{д}} = 3$ (дня) $= 0{,}1$ (месяца), $q_{\text{д}} = t_{\text{д}} \nu = 14\ 000$ (л).

Величина затрат при действующей системе равна:

$$L^* = \frac{700 \cdot 140\ 000}{14\ 000} + \frac{4 \cdot 14\ 000}{2} = 35\ 000 \text{ (ден. ед.)}.$$

7.3.2 Модель с конечной интенсивностью поступления заказа

Пусть заказанная партия поступает с интенсивностью u единиц в единицу времени. Очевидно, система может работать без дефицита, если интенсивность поставок u превосходит интенсивность потребления ν. Таким образом рассматривается система типа заводского склада, куда продукция, произведенная одним цехом, поступает с определенной интенсивностью и используется в производстве другого цеха [32]. Изменение уровня запаса для рассматриваемого случая изображено на рисунке 7.4. В течение времени t_1 запас одновременно и поступает, и расходуется, это время накопления запаса. В течение t_2 запас только расходуется. Длина цикла $t = t_1 + t_2$. Учитывая, что максимальный наличный запас $I_{\text{м}} = q(1 - \nu/u)$, издержки системы в единицу времени составят:

$$L = \frac{K\nu}{q} + \frac{sq}{2}\left(1 - \frac{\nu}{u}\right).$$

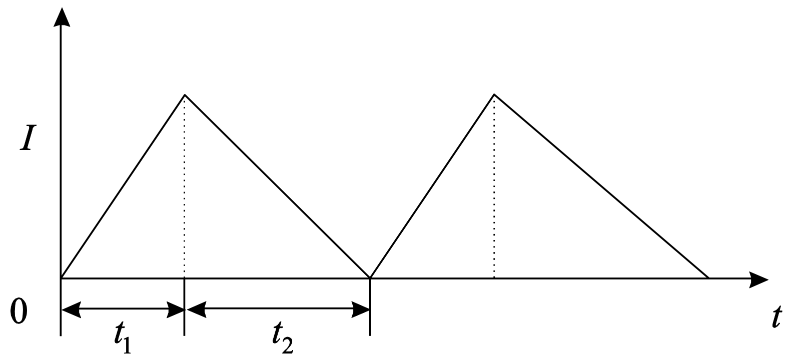


Рис. 7.4 – Модель с конечной интенсивностью поступления заказа

Оптимальные параметры работы системы определяются обычным образом. Величины оптимальной партии:

$$q^* = \sqrt{\frac{2K\nu}{s}}\frac{1}{\sqrt{1-\nu/u}},$$

оптимальный период возобновления заказа:

$$t^* = \sqrt{\frac{2K}{s\nu}}\frac{1}{\sqrt{1-\nu/u}},$$

и его составляющие:

$$t_1^* = \frac{q^*}{u}, t_2^* = \sqrt{\frac{2K}{s\nu}}\sqrt{1-\nu/u},$$

минимальные издержки в единицу времени:

$$L^* = \sqrt{2Ks\nu}\sqrt{1-\nu/u}.$$

В случае, когда интенсивность поставки значительно больше интенсивности потребления $\nu/u \to 0,$ расчетные параметры становятся параметрами обычной системы Уилсона.

7.3.3 Модель с учетом неудовлетворенных требований

В некоторых случаях, когда потери из-за дефицита сравнимы с издержками хранения, дефицит допускается. Пусть требования, поступающие в момент отсутствия запаса, берутся на учет. Обозначим через y максимальную величину задолженного спроса (рис. 7.5). Максимальная величина наличного запаса $Y = q - y$ расходуется за время t_1 (время существования наличного запаса), а затем поступающие требования ставятся на учет в течение времени t_2 (время дефицита). При поступлении очередной партии в первую очередь удовлетворяется задолженный спрос, а затем пополняется запас [33]. Убытки, связанные с дефицитом единицы запаса в единицу времени, составляют d. Затраты на хранение продукции пропорциональны средней величине запаса $(q-y)/2$ и времени его существования $(q-y)/\nu$; аналогично убытки от дефицита пропорциональны средней величине дефицита $y/2$ и времени его существования y/ν. Средние издержки работы системы в течение цикла, включающие затраты на размещение заказа, содержание запаса и потери от дефицита:

$$L_{\text{ц}} = K + s\frac{q-y}{2}\frac{q-y}{\nu} + d\frac{y}{2}\frac{y}{\nu}.$$

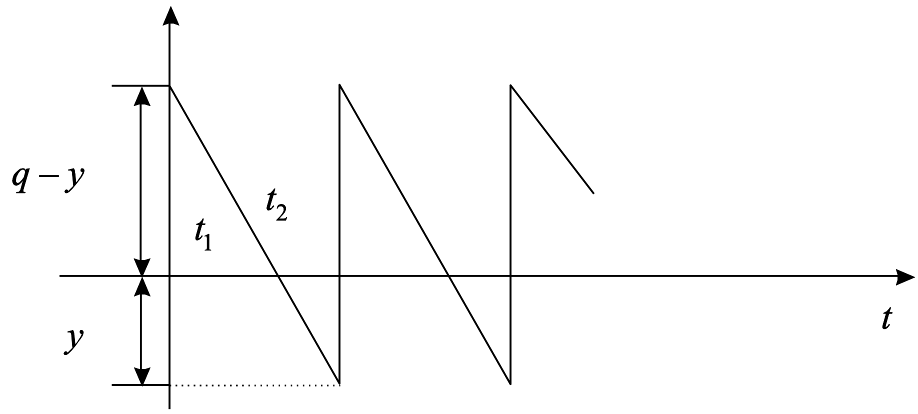


Рис. 7.5 – Модель с учетом неудовлетворенных требований

Разделим издержки цикла на его величину $t = q/\nu$ и получим издержки работы системы в единицу времени:

$$L = \frac{K\nu}{q} + s\frac{(q-y)2}{2q} + d\frac{y2}{2q}.$$

Откуда обычным способом находим:

$$q^* = \sqrt{\frac{2K\nu}{s}}\sqrt{1+\frac{s}{d}},$$

$$y^* = \frac{s}{d}\sqrt{\frac{2K\nu}{s}}\frac{1}{\sqrt{1+s/d}},$$

$$L^* = \sqrt{2Ks\nu}\frac{1}{\sqrt{1+s/d}}.$$

Подставив значения q^* и y^* в соответствующие выражения, найдем другие оптимальные параметры системы:

$$Y^* = q^* - y^* = \sqrt{\frac{2K\nu}{s}}\frac{1}{\sqrt{1+s/d}},$$

$$t_1^* = \frac{Y^*}{\nu} = \sqrt{\frac{2K}{s\nu}}\frac{1}{\sqrt{1+s/d}},$$

$$t_2^* = \frac{y^*}{\nu} = \frac{s}{d}\sqrt{\frac{2K}{s\nu}}\frac{1}{\sqrt{1+s/d}},$$

$$t^* = t_1^* + t_2^* = \frac{q^*}{\nu} = \sqrt{\frac{2K}{s\nu}}\sqrt{1+s/d}.$$

В более сложных моделях управления запасами сохраняется общий подход: строится функция затрат на приобретение запаса, строится функция потерь при хранении запаса и при его нехватке, находится уравнение, при котором минимизируются затраты и потери.

7.3.4 Модель с определением точки заказа

В реальных ситуациях следует учитывать время выполнения заказа Q. Для обеспечения бесперебойного снабжения заказ должен подаваться в момент, когда уровень запаса достаточен для удовлетворения потребности на время выполнения заказа. Этот уровень называется точкой возобновления заказа и обозначается j. Для систем, в которых дефицит не допускается, заказ должен размещаться в момент, когда величина наличного запаса равна:

$$j = Q\nu - \left[\frac{Q}{t^*}\right]q^*,$$

где [.] – целая часть числа (.).

Для обеспечения бездефицитной работы необходим минимальный начальный запас I_0, величина которого $I_0 = Q\nu$. Пусть I – фактический начальный запас. Для непрерывной работы необходимо, чтобы $I \geq Q\nu$. Время потребления начального запаса равно I/ν. Чтобы заказанная партия была доставлена не позже полного расхода начального запаса, ее нужно разместить в момент $T_0 = I/\nu - Q$. В общем случае заказы нужно размещать в моменты:

$$T = \left(\frac{I}{\nu} - Q\right) + kt^*,\ k = 0,1,2,...$$

В системе с конечной интенсивностью поступления заказа при определении оптимальной точки рассматриваются два случая:

если $Q - \left[\frac{Q}{t^*}\right]t^* > t_2^*$, то $t = Q(\nu - u) + \left(\left[\frac{Q}{t^*}\right] + 1\right)\left(\frac{u}{\nu} - 1\right)q^*$;

если $Q - \left[\frac{Q}{t^*}\right]t^* < t_2^*$, то $t = Q\nu - \left[\frac{Q}{t^*}\right]q^*$.

Для системы с учетом неудовлетворенных требований точка заказа определяется по формуле

$$j = Q\nu - \left[\frac{Q}{t^*}\right] q^* - y^*$$

и может быть отрицательной величиной. Это означает, что заявки на пополнение запаса должны посылаться, когда величина дефицита составляет $[j]$.

Контрольные вопросы по главе 7

1. Что изучает теория управления запасами?
2. Как классифицируются системы управления запасами?
3. Опишите модель Уилсона.
4. Из чего складываются затраты на закупку товара?
5. Как определить оптимальный объем закупки?
6. Как определить период между поставками?

8 Теория массового обслуживания

8.1 Понятие о задачах теории массового обслуживания

Очереди возникают практически во всех системах массового обслуживания (СМО) и *теория массового обслуживания (теория очередей)* занимается оценкой функционирования системы при заданных параметрах и поиском параметров, оптимальных по некоторым критериям.

Эта теория представляет особый раздел теории случайных процессов и использует, в основном, аппарат теории вероятностей. Первые публикации в этой области относятся к 20-м гг. XX в. и принадлежат датчанину А. Эрлангу, занимавшемуся исследованиями функционирования телефонных станций – типичных СМО, где случайны моменты вызова, факт занятости абонента или всех каналов, продолжительность разговора. В дальнейшем теория очередей нашла развитие в работах К. Пальма, Ф. Поллачека, А. Я. Хинчина, Б. В. Гнеденко, А. Кофмана, Р. Крюона, Т. Саати и других советских и зарубежных математиков [34].

В качестве основных элементов СМО следует выделить входной поток заявок, очередь на обслуживание, систему (механизм) обслуживания и выходящий поток заявок. В роли заявок (требований, вызовов) могут выступать покупатели в магазине, телефонные вызовы, поезда при подходе к железнодорожному узлу, вагоны под разгрузкой, автомашины на станции техобслуживания, самолеты в ожидании разрешения на взлет и т. д. Роль обслуживающих приборов (каналов, линий) играют продавцы или кассиры в магазине, таможенники, пожарные машины, взлетно-посадочные полосы, экзаменаторы, ремонтные бригады.

В зависимости от характеристик этих элементов СМО классифицируются следующим образом.

1. *По характеру поступления заявок*. Если интенсивность входного потока (количество заявок в единицу времени) постоянна или является заданной функцией от времени, поток называют *регулярным*. Если па-

раметры потока независимы от конкретного момента времени, поток называют *стационарным*.

2. *По количеству одновременно поступающих заявок*. Поток с вероятностью одновременного появления двух и более заявок, равной нулю, называется *ординарным*.
3. *По связи между заявками.* Если вероятность появления очередной заявки не зависит от количества предшествующих заявок, имеем дело с потоком *без последействия*.
4. *По однородности заявок* выделяют *однородные* и *неоднородные потоки*.
5. *По ограниченности потока заявок* различают *замкнутые* и *разомкнутые* системы (система с ограниченной клиентурой называется замкнутой). Так, универсальный магазин является разомкнутой системой, тогда как оптовый магазин с постоянными клиентами – замкнутая система.
6. *По поведению в очереди* системы делятся на системы *с отказами* (заявка покидает систему, если нет мест в очереди), *с ограниченным ожиданием* и *с ожиданием без ограничения времени*.
7. *По дисциплине выбора на обслуживание*. Здесь можно выделить системы с обслуживанием в порядке поступления, в случайном порядке, в порядке, обратном поступлению (последний пришел – первым обслужен), или с учетом приоритетов.
8. *По числу каналов обслуживания* системы разделяют на одно- и многоканальные.
9. *По времени обслуживания* выделяют системы *с детерминированным* и *случайным* временем.
10. *По количеству этапов обслуживания* различают *однофазные* и *многофазные* системы.

8.2 Основы математического аппарата анализа простейших СМО

Рассмотрим стационарный поток однородных заявок без последействия. Пусть $P_k(\tau)$ – вероятность появления k заявок в интервале времени τ. Эта вероятность зависит только от τ и не зависит от начала отсчета времени, от поступления заявок в предыдущих временных интервалах. Пусть к тому же поток является ординарным, т. е. $P_k(dt)$ при $k>1$ бесконечно мала в сравнении с малым интервалом dt. Если обозначить через λ число заявок в единицу времени (интенсивность потока), то можно показать, что для такого *простейшего* потока

$$P_k(t)=\frac{(\lambda t)^k}{k!}e^{-\lambda t}.$$

Эта формула определяет *распределение Пуассона*. Для пуассоновского потока можно обнаружить, что промежутки времени T между поступлениями заявок распределены по экспоненциальному (показательному) закону:

$$P(T<t)=1-e^{-\lambda t}$$

(вероятность, что промежуток времени не превышает t).

Естественно, что входной поток может описываться не только пуассоновским, но и другими распределениями (Эрланга, гиперэкспоненциальным и т. п.).

Аналогичная ситуация имеет место и для выходного потока. Чаще всего используется показательный закон распределения времени обслуживания:

$$P(t)=1-e^{-\mu t},$$

где $\mu=1/t_{\text{обс}}$ – интенсивность обслуживания (среднее число обслуживаний в единицу времени),

$t_{\text{обс}}$ – среднее время обслуживания одной заявки.

Пусть S – множество состояний системы и $P(l,t+\tau/i,t)$ – вероятность того, что система, находившаяся в момент t в состоянии i, в момент $t+\tau$ окажется в состоянии l. Для марковской системы (она привлекает нас отсутствием последействия) можно записать *уравнения Чепмена – Колмогорова*:

$$P(l,t+\tau/i,t)=\sum_{j\in S}P\left(j,t+\tau^*/i,t\right)P\left(l,t+\tau/j,t+\tau^*\right).$$

Если под состояниями понимать число заявок, то эти уравнения можно записать в виде:

$$P_k(t+\tau)=\sum_{i+j=k}P_i(t)\,P_j(\tau).$$

Рассмотрим *случай разомкнутой системы* с простейшим входным потоком интенсивности λ и одним каналом обслуживания с интенсивностью μ.

Возьмем интервал времени $[t,t+dt]$. В силу разомкнутости системы множество состояний системы

$$S=\{S_0,S_1,S_2,\ldots,S_k,S_{k+1},\ldots\},$$

где S_k – состояние, когда в системе находится k заявок.

Оценим вероятность перехода между состояниями с учетом того, что вероятность появления заявки в этом интервале времени равна $\lambda\,dt$ и вероятность завершения обслуживания предшествующей заявки равна $\mu\,dt$.

Очевидно, что вероятность перехода $S_0 \to S_1$ равна $\lambda\,dt$ и вероятность перехода $S_1 \to S_0$ равна $1-\lambda\,dt$. Если в системе присутствовали $k>0$ заявок (состояние S_k), то для перехода в состояние S_{k-1} необходимо, чтобы заявка была обслужена и не поступило новой заявки; отсюда вероятность перехода $S_k \to S_{k-1}$ равна $\mu\,dt(1-\lambda\,dt) \cong \mu\,dt$. Для перехода из состояния S_k в состояние S_{k+1} необходимо, чтобы поступила новая заявка, но ни одна из ранее поступивших не была обслужена: вероятность перехода $S_k \to S_{k+1}$ равна $\lambda\,dt(1-\mu\,dt) \cong$ $\cong \lambda\,dt$. Вероятность для системы остаться в том же состоянии составит $1-(\lambda+\mu)dt$.

Тогда имеем

$$P_0(t+dt) = (1-\lambda\,dt)P_0(t) + \mu\,dt\;P_1(t),$$

$$P_k(t+dt) = \lambda\,dt\;P_{k-1}(t) + (1-\lambda\,dt - \mu\,dt)P_k(t) + \mu\,dt\;P_{k+1}(t),\quad k>0.$$

При $dt \to 0$ получаем дифференциальные уравнения состояний СМО:

$$\frac{d}{dt}P_0(t) = -\lambda\,P_0(t) + \mu\,P_1(t),$$

$$\frac{d}{dt}P_k(t) = \lambda\,P_{k-1}(t) - (\lambda+\mu)P_k(t) + \mu\,P_{k+1}(t),\quad k>0.$$

Это решение при заданных начальных условиях достаточно сложно (здесь можно использовать преобразование Лапласа или численное решение задачи Коши для системы обыкновенных дифференциальных уравнений большого порядка, но едва ли это посильно непрофессионалу) [34].

Ограничимся рассмотрением *установившегося* режима, признаком которого является существование предела

$$\lim_{t\to\infty} P_k(t) = P_k,\quad k = 0,1,2,\ldots$$

В этом случае мы получим бесконечную систему линейных алгебраических уравнений с трехдиагональной матрицей коэффициентов:

$$\lambda\,P_0 = \mu\,P_1,$$

$$(\lambda+\mu)P_k = \lambda\,P_{k-1} + \mu\,P_{k+1},\quad k = 1,2,\ldots$$

Обозначив

$$\rho = \lambda/\mu,$$

имеем

$$P_1 = \rho\, P_0, P_2 = \rho^2 P_0, P_3 = \rho^3 P_0, P_4 = \rho^4 P_0, \ldots, P_k = \rho^k P_0, \ldots,$$

откуда с учетом

$$P_0 + P_1 + P_2 + P_3 + \ldots + P_k + \ldots = 1$$

получаем при $\rho < 1$

$$P_0 \left(1 + \rho + \rho^2 + \rho^3 + \ldots + \rho^k + \ldots\right) = P_0 \frac{1}{1-\rho} = 1.$$

Тогда

$$P_0 = 1 - \rho$$

и

$$P_k = \rho^k P_0 \text{ для } k = 1, 2, \ldots$$

Обратите внимание на требование $\rho < 1$. Если это требование нарушено, то ни о каком установившемся режиме не может быть речи: очередь растет неограниченно (средняя продолжительность обслуживания больше среднего интервала времени между заявками).

Теперь обратимся к аналогичной *замкнутой системе* с числом заявок, не превышающим n. Здесь система уравнений приведется к конечной системе:

$$\frac{d}{dt} P_0(t) = -\lambda\, P_0(t) + \mu\, P_1(t),$$
$$\frac{d}{dt} P_k(t) = \lambda\, P_{k-1}(t) - (\lambda + \mu) P_k(t) + \mu\, P_{k+1}(t), \quad k = 1, 2, \ldots, n-1,$$
$$\frac{d}{dt} P_n(t) = -\mu\, P_n(t) + \lambda\, P_{n-1}(t),$$

которая для установившегося режима дает конечную систему линейных алгебраических уравнений

$$\lambda\, P_0 = \mu\, P_1,$$
$$(\lambda + \mu) P_k = \lambda\, P_{k-1} + \mu\, P_{k+1}, \quad k = 1, 2, \ldots, n-1,$$
$$\lambda\, P_{n-1} = \mu\, P_n.$$

Решение этой системы

$$P_1 = \rho\, P_0, P_2 = \rho^2 P_0/2, P_3 = \rho^3 P_0/(2\times 3), P_4 = \rho^4 P_0/(2\times 3\times 4), \ldots, P_n = \rho^n P_0/n!$$

дает

$$P_0 = 1 \Big/ \sum_{k=0}^{n} \frac{\rho^k}{k!}$$

и

$$P_k = P_0 \rho^k / k! \text{ для } k = 1,2,...,n.$$

Полученные выше решения можно обобщить на случай многоканальных систем с ограниченным ожиданием. Так, если СМО имеет N однотипных каналов обслуживания (интенсивность обслуживания равна $N\mu$), m мест в очереди и к тому же число n возможных заявок превышает $N+m$ (в противном случае нет проблем), то возникает система:

$$\frac{d}{dt}P_0(t) = -\lambda P_0(t) + \mu P_1(t),$$

$$\frac{d}{dt}P_k(t) = \lambda P_{k-1}(t) - (\lambda + k\mu) P_k(t) + (k+1)\mu P_{k+1}(t), \quad k = 1,...,N-1,$$

$$\frac{d}{dt}P_k(t) = \lambda P_{k-1}(t) - (\lambda + N\mu) P_k(t) + N\mu P_{k+1}(t), \quad k = N,...,N+m-1,$$

$$\frac{d}{dt}P_{N+m}(t) = -N\mu P_{N+m}(t) + \lambda P_{N+m-1}(t),$$

из которой получаем для установившегося режима:

$$\lambda P_0 = \mu P_1,$$

$$(\lambda + k\mu) P_k = \lambda P_{k-1} + (k+1)\mu P_{k+1}, \quad k = 1,2,...,N-1,$$

$$(\lambda + N\mu) P_k = \lambda P_{k-1} + N\mu P_{k+1}, \quad k = N, N+1,...,N+m-1,$$

$$\lambda P_{N+m-1} = N\mu P_{N+m}.$$

Решение этой системы дает:

$$P_k = P_0 \rho^k / k! \text{ для } k = 1,2,...,N-1,$$

$$P_k = P_0 \rho^k / \left(N^{k-N} N!\right) \text{ для } k = N, N+1,...,N+m,$$

$$P_0 = \left[\sum_{k=0}^{N} \frac{\rho^k}{k!} + \frac{\rho^{N+1}}{N \cdot N!} \cdot \frac{(\rho/N)^m - 1}{(\rho/N) - 1}\right]^{-1}.$$

Умение найти значения P_k дает возможность отыскать и ряд основных характеристик СМО.

8.3 Основные характеристики СМО

Значение P_0 определяет вероятность того, что все каналы обслуживания свободны (находятся в состоянии простоя).

Значение P_k определяет вероятность того, что в системе (в очереди и на обслуживании) находятся k заявок. Если k не превышает числа каналов N, то все заявки находятся на обслуживании и очередь отсутствует; в противном случае все каналы заняты и $k-N$ заявок находится в очереди [16].

Вероятность $P_{\text{отк}}$ отказа в обслуживании определяется ситуацией занятости всех N каналов и всех m мест в очереди и равна P_{N+m}.

Среднее число занятых каналов $N_{\text{зан}}$ определяется математическим ожиданием дискретной случайной величины:

$$N_{\text{зан}} = \sum_{k=1}^{N} k \cdot P_k + \sum_{k=N+1}^{N+m} N \cdot P_k = \rho \cdot \left[1 - \frac{\rho^{N+m}}{N!\cdot N^m} P_0\right]$$

(мы опускаем здесь достаточно простые преобразования).

Среднее число свободных каналов:

$$N_{\text{своб}} = N - N_{\text{зан}}.$$

Коэффициент простоя каналов:

$$K_{\text{прост}} = N_{\text{своб}} / N.$$

Коэффициент занятости каналов:

$$K_{\text{занят}} = N_{\text{зан}} / N.$$

Относительная пропускная способность (доля обслуженных заявок в общем числе поступавших в систему) определяется величиной

$$q = 1 - P_{\text{отк}}.$$

Абсолютная пропускная способность (среднее число заявок, обслуживаемых в единицу времени) определяется величиной

$$A = \lambda \cdot q. \qquad (8.1)$$

Средняя длина очереди:

$$L_{\text{очер}} = \sum_{k=N+1}^{N+m} (k-N) P_k = \frac{\rho^{N+1}}{N!\cdot N} \cdot \frac{1-(\rho/N)^m (m+1-m\rho/N)}{(1-\rho/N)^2} P_0.$$

Среднее число заявок, находящихся в системе, складывается из средних значений занятости каналов и длины очереди:

$$L = N_{\text{зан}} + L_{\text{очер}}.$$

Среднее время пребывания заявки в очереди равно

$$T_{\text{очер}} = L_{\text{очер}} / \lambda.$$

Общее время пребывания заявки в очереди будет складываться из $T_{\text{очер}}$ и среднего времени обслуживания:

$$T_{\text{сист}} = T_{\text{очер}} + q/\mu.$$

Полученные характеристики дают возможность анализа замкнутых и разомкнутых систем с отказами $(m=0)$, с очередью или с ожиданием $(m \to \infty)$ при простейшем входном потоке и однотипных параллельных каналах обслуживания с показательным законом длительности обслуживания (в частности, с фиксированной длительностью) [16].

8.4 Примеры систем с ограниченной очередью

Пример 8.1

Пусть на аэродром самолеты прибывают с интенсивностью 27 самолетов в час, время приземления составляет 2 минуты, допускается нахождение над аэродромом не более $m=10$ самолетов. Нужно определить число N посадочных полос, гарантирующее вероятность отказа, не превышающую 0,05, и среднее время ожидания, не превышающее 5 минут.

Здесь $\lambda=27, \mu=30, \rho=\lambda/\mu=0,9.$

Отыскиваем вероятность простоя диспетчеров службы посадки:

$$P_0=\left[\sum_{k=0}^{N}\frac{0,9^k}{k!}+\frac{0,9^{N+1}}{N\cdot N!}\cdot\frac{(0,9/N)^{10}-1}{(0,9/N)-1}\right]^{-1}.$$

Вероятность отказа в посадке равна

$$P_{\text{отк}}=P_0\,0,9^{N+10}\big/\left(N^{10}\cdot N!\right).$$

Среднее время ожидания в воздухе равно

$$T_{\text{очер}}=L_{\text{очер}}/\lambda,$$

где

$$L_{\text{очер}}=\sum_{k=N+1}^{N+m}(k-N)\cdot P_k=\frac{\rho^{N+1}}{N!\cdot N}\cdot\frac{1-(\rho/N)^m(m+1-m\rho/N)}{(1-\rho/N)^2}\cdot P_0=$$

$$=\frac{0,9^{N+1}}{N!\cdot N}\cdot\frac{1-(0,9/N)^{10}(11-10\cdot 0,9/N)}{(1-0,9/N)^2}\cdot P_0.$$

Выполняя арифметические действия при $N=1,$ обнаруживаем, что

$$P_0\cong 0,14;\ P_{\text{отк}}\cong 0,04;\ L_{\text{очер}}\cong 0,045;\ T_{\text{очер}}\cong 0,9 \text{ мин}$$

и что одной посадочной полосы при указанных условиях вполне достаточно.

Пример 8.2

Пусть имеются станки, которые могут выходить из строя с частотой в среднем 2 раза за смену. Продолжительность ремонта одним оператором составляет около трех часов (оператор одновременно может ремонтировать лишь один станок и не переходит к другому, не отремонтировав предыдущий). Хотелось бы определить число операторов, при котором потери от простоя станков и оплаты лишнего числа операторов были бы минимальны.

Такую замкнутую систему можно представить системой с N каналами (операторами) и очередью с m местами ожидания (совпадает с числом станков). Если известны потери C_n от простоя станка в течение часа и оплата C_p часа работы оператора, то при семичасовой смене задача сводится к нахождению значения N, которое минимизировало бы значение

$$C_n \times T_{\text{очер}} + C_p \times 7 \times N,$$

где $T_{\text{очер}}$ определяется при $\lambda = 2/7$, $\mu = 1/3$, $\rho = \lambda/\mu = 6/7$.

. . . .

Можно привести множество подобных задач для определения числа кассиров в универмаге, наилучшего с позиций минимума потерянных покупателей; для определения числа бригад грузчиков на железнодорожной станции, минимизирующего штраф за простой вагонов; для определения числа полос движения на проектируемой автомагистрали и т. п.

8.5 Дисциплина ожидания и приоритеты

Выше мы рассматривали простейший поток однотипных заявок с дисциплиной выборки на обслуживание в порядке поступления [8].

Можно показать, что и в случае случайного выбора на обслуживание полученные выше оценки не претерпят изменения, но их дисперсия (разброс относительно ожидаемой величины) возрастет. Очевидно, что среднее время сидения в очереди не изменится от того, что кто-то пройдет без очереди, но для отдельных клиентов время ожидания увеличится. Так, отношение дисперсий времени ожидания в неупорядоченной и упорядоченной очереди имеет порядок $(2+\rho)/(2-\rho)$, где $\rho = \lambda/\mu$ (мы обычно предпочитаем систему с жесткой дисциплиной обслуживания из-за предсказуемости ее поведения и всякое «возмущение» в ее работе отрицательно действует на нашу психику).

Существует множество систем, в которых присутствует $N > 1$ входных потоков с различной интенсивностью $\lambda_i \left(i = 1,\dots,N \right)$, время обслуживания заявок которых распределено по показательному закону с параметрами μ_i. Здесь при условии пуассоновости входных потоков можно считать, что суммарный поток будет пуассоновским с интенсивностью $\Lambda = \sum \lambda_i$; функция распределения времени обслуживания заявок суммарного потока в одноканальной системе:

$$S(t) = \sum_{i=1}^{N} \frac{\lambda_i}{\Lambda}\left(1 - e^{-\mu_i t}\right);$$

среднее время ожидания определяется формулой Поллачека – Хинчина:

$$W = \frac{\Lambda/2 \int_0^{\infty} t^2 dS(t)}{1 - R_N}, \quad R_N = \Lambda \int_0^{\infty} t \, dS(t),$$

которая для данного случая дает $R_N = \sum \left[\lambda_i / \mu_i\right] = \sum \rho_i$ и в случае стационарности режима $\left(R_N < 1\right)$:

$$W = \frac{\sum_{i=1}^{N} \rho_i/\mu_i}{1 - R_N}.$$

Определенный интерес представляют системы, где каждому входному потоку сопоставлено целое число k – показатель приоритета потока (наивысший приоритет определяется $k = 1$) [21].

Если обслуживание заявки не прерывается ни при каких условиях и выбор на обслуживание происходит с учетом приоритета (при одинаковом приоритете выбирается первый пришедший в систему), то такая система называется *системой с относительными приоритетами*.

Можно показать, что поскольку время ожидания заявки с приоритетом k складывается из времени завершения обработки требования, вошедшего в канал, времени обслуживания ранее поступивших требований приоритета от 1 до $k-1$ и ранее поступивших требований с приоритетом k, то его среднее значение равно

$$W_k = \frac{\sum_{i=1}^{N} \rho_i/\mu_i}{\left(1 - R_k\right)\left(1 - R_{k-1}\right)}, \quad R_k = \sum_{i=1}^{k} \rho_i \quad \left(k = 1 \ldots N\right).$$

На этой основе можно определить среднюю длину очереди заявок k-го приоритета $L_k = \lambda_k W_k$ и среднее число таких заявок в системе $L_k + \rho_k$. Показано, что введение приоритетов улучшает функционирование системы, если более высокое преимущество присваивается заявкам с меньшей длительностью обслуживания. Если учитывать стоимостные характеристики, то более высокое преимущество предоставляется заявкам с большим значением $C_k \times \mu_k$, где C_k – средняя стоимость ожидания.

Существуют системы с абсолютными приоритетами, где появление заявки более высокого уровня прерывает обслуживание текущей заявки, которая вернется в очередь и потом снова поступит на обслуживание с места прерывания (или с начала). Здесь среднее время ожидания заявки с приоритетом k:

$$W_k = \frac{\sum_{i=1}^{N} \rho_i / \mu_i}{(1-R_k)(1-R_{k-1})} + \frac{R_{k-1}}{\mu_k (1-R_{k-1})}, \quad R_k = \sum_{i=1}^{k} \rho_i \quad (k = 1 \ldots N).$$

В [15] рекомендуется для минимизации затрат на пребывание заявок в очереди в системах с относительными и абсолютными приоритетами, равных

$$\Phi = \sum_{k=1}^{N} \alpha_k \lambda_k W_k,$$

где α_k – издержки на ожидание заявки k-го приоритета в единицу времени, более высокий приоритет давать заявкам с наибольшим значением $\alpha_k \mu_k$.

Исключительно сложно установить разумные приоритеты в случае многофазных систем, где заявка проходит обслуживание в нескольких последовательных подсистемах [20]. Здесь относительно простые выводы удается сделать лишь для случая двух подсистем, и для получения выводов для более сложных систем приходится прибегать к моделированию.

8.6 Моделирование систем массового обслуживания и метод Монте-Карло

До сих пор мы рассматривали системы, для которых удавалось описать результаты обследования в аналитическом виде. Однако многие реальные системы (многофазные, с оригинальными приоритетами, с признаками нестационарности, с непуассоновскими входными потоками, с непоказательным распределением длительности обслуживания и т. д.) не поддаются такому решению.

Суть математического моделирования системы заключается в следующем.

Время функционирования системы разделяется на достаточно большое количество подынтервалов (единиц времени, в течение которых не может возникнуть более одной заявки или завершиться выполнение более одной заявки). Для каждого такого подынтервала последовательно моделируется факт появления новой заявки (да/нет), проверяется наличие свободного канала (закончено ли обслуживание какой-то заявки) и загрузка его заявкой из очереди, проверяется наличие мест в очереди с последующим выводом (принять в очередь/отказать в обслуживании) и т. д. При этом фиксируется число отказов, время ожидания заявок в очереди и в системе вообще, число заявок в очереди в каждый момент и другие значения, которые позволяют найти вероятность отказа, распределение времени ожидания и среднее время, вероятность простоя каналов и т. п. Для надежности выводов такое разовое моделирование повторяется достаточно много раз.

Очевидно, что ни о каком ручном моделировании не может быть речи (объем работы здесь слишком велик для нормального индивида). Здесь приходится использовать компьютер со встроенным или программным датчиком псевдослучайных чисел с равномерным законом распределения в интервале от 0 до 1. Псевдослучайные числа получаются по какому-то алгоритму, но в совокупности подчиняются всем законам проверки на случайность (мы не останавливаемся на методах их получения, так как есть отличные программные датчики во всех системах программирования).

В процессе моделирования возникает необходимость генерации случайных чисел с законом распределения, отличным от вышеуказанного.

Пусть R – случайные числа с равномерным законом распределения в [0,1] и X – создаваемые случайные числа с плотностью распределения $p(X)$. Между ними можно установить соотношение:

$$R = \int_{-\infty}^{x} p(X)dX.$$

Так, для показательного распределения $p(X) = \lambda \exp(-\lambda x)$ для $X > 0$ легко установить, что $X = -\ln(1-R)/\lambda$. Для равномерного распределения в $[a,b]$ с очевидностью $X = a + (b-a)R$. Получение дискретных случайных чисел сводится к поиску наименьшего значения X, при котором

$$R \le e^{-\lambda} \sum_{k=0}^{x} \lambda^k / k!.$$

Если взятие интеграла и представление X через R составит затруднение, можно воспользоваться методом Неймана. Здесь при неограниченности области значений X усекаем ее до некоторого интервала $[a,b]$; например, для нормального распределения концы интервала берем отстоящими от среднего на 3–4 стандартных отклонения. Затем генерируется пара случайных чисел R_1 и R_2; если $R_2 \le (b-a) \times p\left(a+(b-a)R_1\right) / p^*$, где $p^* = \max\left[(b-a) \cdot p(x)\right]$, то берем

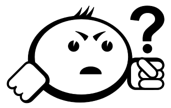
$X = a + (b - a)R_2$ и в противном случае берем следующую пару случайных чисел.

Таким путем мы можем моделировать интервалы времени между заявками входного потока, продолжительность обслуживания заявки, вероятность выхода канала из строя и т. п.

Вопрос о числе N отдельных реализаций системы решается на основе закона больших чисел и вывода о том, что погрешность оценок имеет порядок

$$1/\sqrt{N}.$$

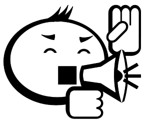
Описанный подход к поиску характеристик сложной системы называют методом статистических испытаний (методом Монте-Карло), который обычно используют там, где другие методы не работают (моделирование сложных систем, вычисление интегралов кратности 10 и выше, поиск экстремумов функций с очень большим числом переменных и др.).



Контрольные вопросы по главе 8

1. Что изучает теория массового обслуживания?
2. Какие существуют виды заявок?
3. Что такое интенсивность потока?
4. Как определяется распределение Пуассона?
5. Перечислите характеристики СМО.
6. Что представляет собой ограниченная очередь?
7. Что такое отказ СМО?

9 Сетевое планирование



В последние десятилетия исключительно усложнились управление и организация различных технических и научных разработок. Большие масштабы разработок усложнили координацию работ исполнителей и оценку хода выполнения работ. Одним из важнейших факторов стало максимальное сокращение сроков разработок с целью недопущения морального старения разрабатываемой системы.

Формулируя закон необходимого разнообразия, У. Р. Эшби указывал, что для обеспечения процесса управления управляющая система должна обладать по крайней мере такой же сложностью, как и управляемая.

Соответственно возникла необходимость создания системы, обеспечивающей возможность оценки текущего состояния и предсказания последующего хода разработки. Результатом исследований в этом направлении явилось создание систем, базирующихся на так называемых сетевых графиках.

Первые методы такого рода получили название CPM (метод критического пути, впервые апробированный при управлении строительными работами) и PERT (метод оценки и обзора программ). Система PERT впервые была применена для управления разработкой ракеты «Полярис», позволив сократить срок разработки, по мнению специалистов, на 2–3 года.

В дальнейшем методы сетевого планирования применялись при проектировании инженерных сооружений, постановке театральных спектаклей, организации переподготовки специалистов и так далее [32].

Было время, когда сетевое планирование было «модным» среди специалистов; сейчас им пользуются там, где оно действительно может оказаться полезным.

9.1 Понятие о сетевом графике

В терминологии теории графов сетевым графиком называют конечный ориентированный граф без контуров, в котором имеются единственная вершина с отсутствующими прообразами и единственная вершина, не имеющая образов.

Иначе ***сетевым графиком*** *можно назвать ориентированную транспортную сеть с одним входом и одним выходом, в которой нет путей с повторяющимися вершинами.*

Дуги указанного графа понимаются как некоторые работы. Вершины графа называются событиями.

Пример 9.1

Информация о сетевом графике некоторого проекта может задаваться в виде рисунка или различных списков. Если информация дана списком (табл. 9.1), то ее можно представить в графическом виде (рис. 9.1); если ввести обозначения или нумерацию для вершин (событий), то можно использовать и графическое представление (рис. 9.2), где числа на дугах определяют продолжительность работ.

Таблица 9.1 – Пример списка работ

Работа	**Последующие работы**	**Продолжительность**	**Работа**	**Последующие работы**	**Продолжительность**
1	2,4	3	6	5,9	4
2	8	4	7	6	3
3	5,9	5	8	–	4
4	6	3	9	8	2
5	–	3	–	–	–

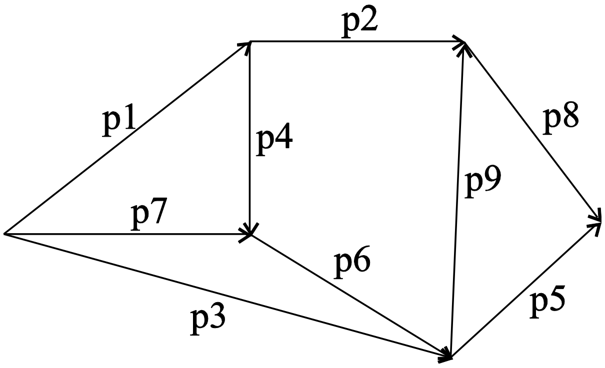


Рис. 9.1 – Графический вид

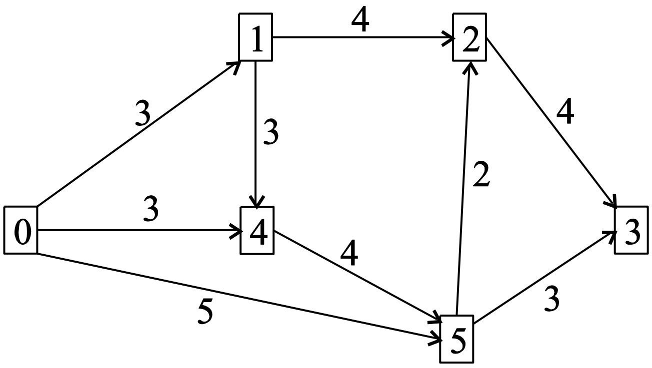
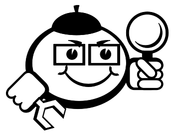


Рис. 9.2 – Графическое представление работ

. . . .

При наличии нумерации вершин информацию о графике можно представить также перечнем работ с указанием начального и конечного событий и продолжительностей работ. Можно ту же самую информацию представить и в виде матрицы связей между событиями (элементы матрицы равны продолжительности соответствующих работ или не указаны при отсутствии таких работ).



. . . . Пример 9.2 . . . .

Матричное задание особенно удобно и наглядно при обработке сетевого графика на ЭВМ (табл. 9.2).

Таблица 9.2 – Матричный вид сетевой задачи

Работа	Продолжительность	Работа	Продолжительность
0–1	3	4–5	4
1–2	4	0–4	3
0–5	5	2–3	4
1–4	3	5–2	2
5–3	3		

Событие	0	1	2	3	4	5
0		3			3	5
1			4		3	
2				4		
3						
4						4
5			2	3		

. . . .

При рисовании сетевого графика часто удобно использовать *фиктивные работы* – работы с нулевой продолжительностью, изображаемые пунктиром и служащие для указания порядка следования работ. Например, при наличии двух выходов их можно связать фиктивной работой.

Могут использоваться и *фиктивные события*. Так, если обнаружится, что пара событий связана более чем одной работой, то можно воспользоваться фиктивными событиями и работами (рис. 9.3).

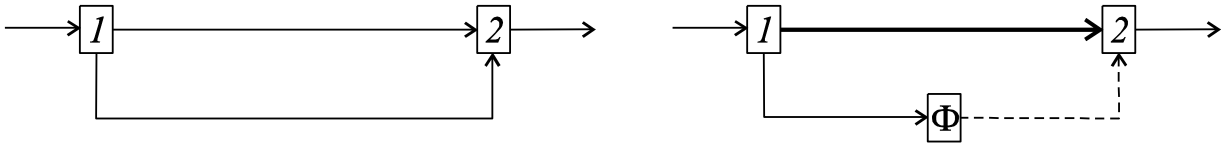


Рис. 9.3 – Фиктивные события и работы

В большинстве реальных проектов точное значение продолжительности работ T_{ij} неизвестно, но на основании экспертизы могут быть предложены верхняя (пессимистическая) оценка B_{ij}, определяющая максимальную оценку продолжительности с учетом всех возможных срывов, и нижняя (оптимистическая) оценка A_{ij}, определяющая продолжительность работы в идеальных условиях. При наличии некоторого опыта может существовать и наиболее вероятная оценка M_{ij}. Соответственно, планируемая продолжительность работы определяется по одной из формул:

$$T_{ij} = \left(3A_{ij} + 2B_{ij}\right)/5,$$

$$T_{ij} = \left(A_{ij} + 4M_{ij} + B_{ij}\right)/6.$$

Заметьте, что требование отсутствия контуров неслучайно, т. к. наличие таковых дает возможность возврата к повторению ранее выполненных работ.

При обработке сетевых графиков существенное влияние на объем вычислительных работ оказывает порядок просмотра (нумерация) событий [34].

Начальному событию (входу) сопоставляется ранг 0. Ранг 1 получают события, в которые приводят работы, начинающиеся только в событиях ранга 0. Ранг 2 получают события, в которые приводят работы, начинающиеся только в событиях ранга 0 и 1 и т. д.

Последующая нумерация ведется в соответствии с рангами по возрастанию номеров (при одинаковом ранге порядок произволен; нумерация не обязательно сплошная).

Для приведенного выше примера ранг 0 получает событие 0, ранг 1 – событие 1, ранг 2 – событие 4, ранг 3 – событие 5, ранг 4 – событие 2 и ранг 5 – событие 3.

9.2 Критический путь и другие параметры сетевого графика

Если продолжительность работы принять за длину соответствующей дуги сетевого графика, то ***критическим*** *можно назвать путь максимальной длины от входа до выхода графика. Длина этого пути определяет критическое время выполнения проекта, т. е. минимальное время, в пределах которого коллектив исполнителей в состоянии выполнить весь комплекс работ сетевого графика.*

Пусть для определенности начальное событие имеет номер 0 и конечное – номер N. Обозначим через L_j длину пути наибольшей протяженности от события 0 до события j.

По известному принципу оптимальности Р. Беллмана

$$L_j = \max_{(i\,j)}\left[L_i + T_{ij}\right],\ \ j > 0;\ \ L_0 = 0.$$

Величина L_j соответствует наиболее раннему возможному времени T_j^0 наступления j-го события, т. е. самому раннему сроку завершения всех работ, предшествующих этому событию. Значение L_N соответствует критическому времени выполнения проекта $T_{\text{крит}}$.

Обозначим через M_i длину пути наибольшей протяженности от события i до события N. Тогда по тому же принципу

$$M_i = \max_{(i\,k)}\left[T_{ik} + M_k\right],\ \ i < N;\ \ L_N = 0.$$

Величина $T_i^1 = T_{\text{крит}} - M_i$ соответствует наиболее позднему допустимому времени наступления i-го события, т. е. самому позднему сроку начала всех работ, последующих за этим событием. Совершенно очевидно, что для событий на критическом пути самое раннее и самое позднее времена их наступления будут совпадать.

Пример 9.3

Рассмотрим сетевой график (рис. 9.4, табл. 9.3).

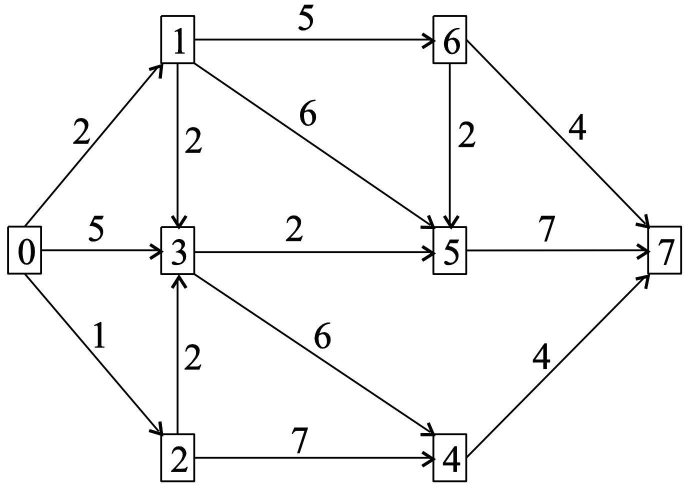


Рис. 9.4 – Исходные данные

Таблица 9.3 – Исходные данные

Работа	**Продолжи-тельность**	**Работа**	**Продолжи-тельность**
0–1	2	2–4	7
0–2	1	3–4	6
0–3	5	3–5	2
1–3	2	4–7	4
1–5	6	5–6	2
1–6	5	5–7	7
2–3	2	6–7	4

Далее рассчитываем значения L_j в порядке роста номеров:

$$L_j = 0;\ L_1 = 2\ \ (i = 0);\ L_2 = 1\ \ (i = 0);$$

$$L_3 = \max\left[L_0 + 5, L_1 + 2, L_2 + 2\right] = 5\ \ (i = 0)\ \text{ и т. д.}$$

Затем рассчитываем значения M_i в порядке убывания номеров (табл. 9.4):

$$M_7 = 0;\ M_6 = 4\ \ (k = 7);\ M_5 = \max\left[2 + M_6, 7 + M_7\right] = 7\ \ (k = 7);$$

$$M_4 = 4\ \ (k = 7);\ M_3 = \max\left[6 + M_4, 2 + M_5\right] = 10\ \ (k = 4)\ \text{ и т. д.}$$

В итоге имеем информацию о наиболее ранних и наиболее поздних моментах наступления событий и индексы предшествующих и последующих событий в самых длинных путях, проходящих через данное событие.

Таблица 9.4 – Расчетные данные

Событие j	L_j T_j^0	i **пред.**	M_j	k **посл.**	T_j^1
0	0	–	15	1,3	0
1	2	0	13	5	2
2	1	0	12	3	3
3	5	0	10	4	5
4	11	3	4	7	11
5	8	1	7	7	8
6	10	5	4	7	11
7	15	4,5	0	–	15

По информации из колонок 3 или 5 можно выявить критические пути с длиной 15: [0 – 1 – 5 – 7] и [0 – 3 – 4 – 7].

Очевидно, что работы, не лежащие на критических путях, обладают *резервами* времени – их выполнение при некоторых условиях может быть задержано на какое-то время. Существуют 4 вида резервов:

1. *Полный резерв* $R_{ij} = T_j^1 - T_i^0 - T_{ij}$.

2. *Свободный резерв* $R_{ij} = T_j^0 - T_i^0 - T_{ij}$.

3. *Независимый резерв* $R_{ij} = \max\left[T_j^0 - T_i^1 - T_{ij}, 0\right]$.

4. *Частный резерв* $R_{ij} = T_j^1 - T_i^1 - T_{ij}$.

Так, полный резерв работы можно понимать как время, на которое можно замедлить выполнение работы, если предшествующие работы завершатся к самому раннему возможному сроку, но комплекс последующих работ будет выполняться в кратчайший возможный срок. Независимый резерв предполагает завершение предшествующих работ к самому позднему, но начало последующих в самый ранний срок.

Результаты обработки приведенного сетевого графика можно представить следующей таблицей (табл. 9.5).

Полученные данные позволяют выделить так называемые *подкритические работы*, т. е. работы, лежащие на путях, отличающихся по длине от критического не более чем на заданную величину. Основной характеристикой здесь может служить *полный резерв*.

Таблица 9.5 – Результаты обработки сетевого графика

Работа	Продолжительность	Раннее время		Позднее время		Резервы			
		начала	конца	начала	конца	полн.	своб.	незав.	част.
0–1	2	0	2	0	2	0	0	0	0
0–2	1	0	1	2	3	2	0	0	2
0–3	5	0	5		5	0	0	0	0
1–3	2	2	4	3	5	1	1	1	1
1–5	6	2	8	2	8	0	0	0	0
1–6	5	2	7	6	11	4	3	3	4
2–3	2	1	3	3	5	2	2	0	0
2–4	7	1	8	4	11	3	3	1	1
3–4	6	5	11	5	11	0	0	0	0
3–5	2	5	7	6	8	1	1	1	1
4–7	4	11	15	11	15	0	0	0	0
5–6	2	8	10	9	11	1	0	0	1
5–7	7	8	15	8	15	0	0	0	0
6–7	4	10	14	11	15	1	1	0	0

Для нашего графика на уровне критичности 1 подкритическими будут работы 1–3, 3–5, 5–6, 6–7. Чтобы убедиться в этом, возьмем работу 5–6 и найдем путь максимальной длины, проходящий через нее, по индексам предшествующих и последующих событий. Так, событию 5 в пути максимальной длины предшествует событие 1, а ему – событие 0. Событию 6 в таком пути последует событие 7. Длина пути 0–1–5–6–7 равна 14 и задержка на 1 при выполнении работ 5–6 или 6–7 сделает его критическим.

Однако полный резерв не полностью характеризует уровень критичности работ.

Возьмем для примера два графика (рис. 9.5).

В первом графике все некритические работы имеют одинаковый полный резерв, равный 8. Однако напряженность работ пути 0–1–4 составляет 24 единицы времени на интервале 32, тогда как напряженность работ пути 0–2–3 составляет 2 единицы на 10. Нет сомнения, что работы второго пути можно выполнять с большей «прохладцей», чем первого (отдыхаем 8 дней из 10 и 8 дней из 32 соответственно).

Во втором графике работы 0–2 и 2–3 имеют резерв 3, а работы 0–1 и 1–4 – резерв 8, но напряженность у них одинакова.

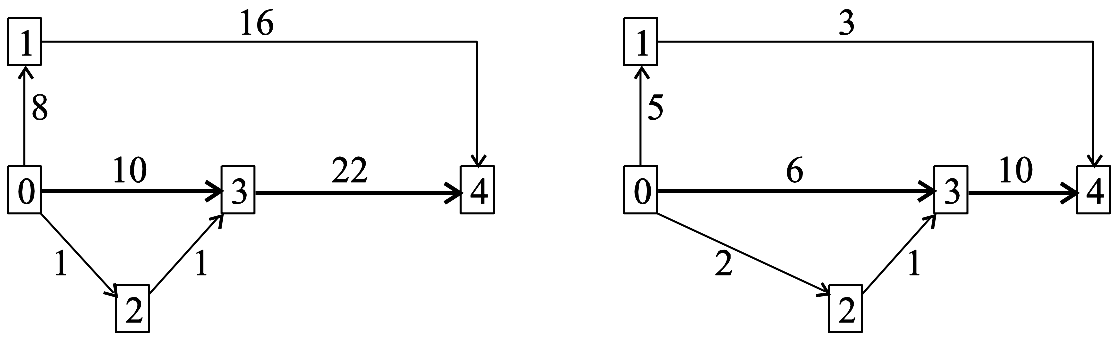


Рис. 9.5 – Пример различного уровня критичности работ

Поэтому всякая работа характеризуется *коэффициентом напряженности*. Для отыскания этого коэффициента отыскивается путь максимальной длины, проходящий через данную работу: при этом используются индексы предшествующих и последующих событий, которые мы находили при поиске T^0 и T^1. На этом пути ищутся ближайшие «слева» и «справа» события, принадлежащие критическому пути (путям), и определяется отношение длины пути между этими событиями, проходящего через данную работу, к длине соответствующего отрезка критического пути.

Так, для рассмотренного выше сетевого графика выберем некритическую работу 1–3. Обнаруживаем, что через нее проходит путь (максимальной длины) 0–1–3–4–7. Ближайшими соседями на критических путях 0–1–5–7 и 0–3–4–7 будут события 1,7 и 0,3 соответственно. Отсюда находим коэффициент напряженности:

$$K_{13} = \max\left|\frac{T_{13}+T_{34}+T_{47}}{T_{15}+T_{57}}, \frac{T_{01}+T_{13}}{T_{03}}\right| = \max\left|\frac{12}{13}, \frac{4}{5}\right| = \frac{12}{13}.$$

Аналогично получаем:

$$K_{35} = \max\left|\frac{T_{03}+T_{35}}{T_{01}+T_{15}}, \frac{T_{35}+T_{57}}{T_{34}+T_{47}}\right| = \max\left|\frac{7}{8}, \frac{9}{10}\right| = \frac{9}{10},$$

$$K_{16} = \max\left|\frac{T_{16}+T_{67}}{T_{15}+T_{57}}, \frac{T_{01}+T_{16}+T_{67}}{T_{03}+T_{34}+T_{47}}\right| = \max\left|\frac{9}{13}, \frac{11}{15}\right| = \frac{11}{15},$$

$$K_{56} = K_{67} = \max\left|\frac{T_{56}+T_{67}}{T_{57}}, \frac{T_{01}+T_{15}+T_{56}+T_{67}}{T_{03}+T_{34}+T_{47}}\right| = \max\left|\frac{6}{7}, \frac{14}{15}\right| = \frac{14}{15},$$

$$K_{02} = K_{23} = \max\left|\frac{T_{02}+T_{23}+T_{34}+T_{47}}{T_{01}+T_{15}+T_{57}}, \frac{T_{02}+T_{23}}{T_{03}}\right| = \max\left|\frac{13}{15}, \frac{3}{5}\right| = \frac{13}{15},$$

$$K_{24} = \max\left|\frac{T_{02}+T_{24}+T_{47}}{T_{01}+T_{15}+T_{57}}, \frac{T_{02}+T_{24}}{T_{03}+T_{34}}\right| = \max\left|\frac{12}{15}, \frac{8}{11}\right| = \frac{12}{15}.$$

9.3 Линейная диаграмма проекта (диаграмма Ганта)

Для небольших проектов с целью большей наглядности выполнения работ во времени после составления пронумерованного сетевого графика можно построить *линейную диаграмму* (рис. 9.6).

При ее построении отрезок $i-j$ откладываем так, чтобы его начало лежало на одной вертикали с самым правым концом отрезков-работ, заканчивающихся в вершине i, что соответствует T_i^0. Самый правый конец всех отрезков соответствует $T_{\text{крит}}$.

Пример 9.4

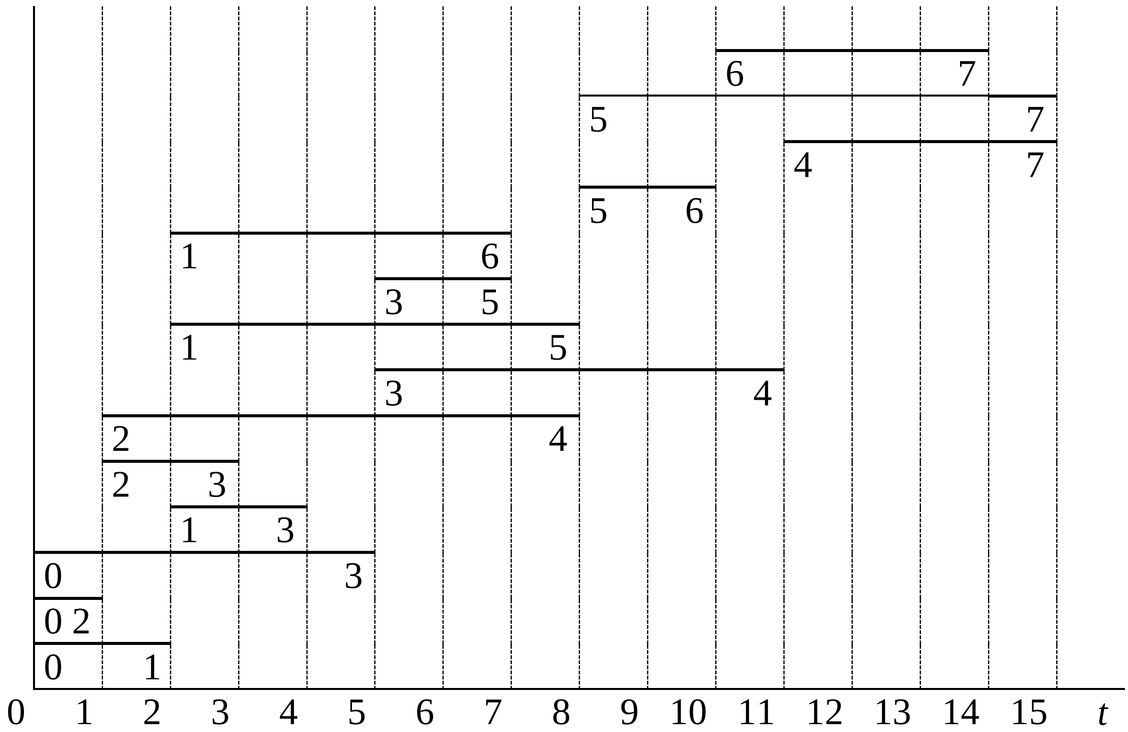


Рис. 9.6 – Диаграмма Ганта

Если сдвинуть отрезки вправо так, чтобы отрезки $i-j$ заканчивались самым левым концом отрезков с начальным индексом j, то эти самые левые концы будут соответствовать T_j^1. Величина такого сдвига определяет полный резерв работы.

Если сдвигать отрезок $i-j$ вправо без сдвига отрезков с начальным индексом *j*, то величина такого сдвига определит свободный резерв.

Существенным достоинством линейной диаграммы является возможность оценить загрузку исполнителей во времени. Так, для t в интервале от 2 до 3 выполняется наибольшее количество работ (шесть). Если работы выполняются взаимозаменимыми исполнителями, то, сдвинув работу 1–6 или 2–4, мы обнаружим возможность обойтись пятью исполнителями (непосредственно из сетевого графика это трудно увидеть).

9.4 Минимизация стоимости проекта при заданной продолжительности

Выполнение всякой работы связано с затратами. В ряде случаев ускоренное выполнение работы связано с увеличением затрат (авралы, срочные поставки и т. п.) и затраты являются обратной функцией от времени выполнения, которая выбирается в виде:

1) линейный вариант:

$$C_{ij} = -A_{ij} \cdot T_{ij} + B_{ij};$$

2) выпуклый вариант:

$$C_{ij} = -A_{ij} / T_{ij}.$$

В таких условиях может быть поставлена задача поиска *оптимального по стоимости безрезервного плана*, т. е. плана, в котором стоимость снижется удлинением выполнения работ до предельного допустимого времени.

Пусть проект требуется выполнить за время T, не меньшее $T_{\text{крит}}$.

Если обозначить через D_{ij} минимальное необходимое время выполнения работы $i-j$, а через T_j – момент наступления j-го события, то продолжительность работы $i-j$ равна $T_j - T_i$ и возникает задача:

минимизировать функцию

$$Z = \sum_{(ij)} \left[-A_{ij} \left(T_j - T_i \right) + B_{ij} \right]$$

при условиях:

$$T_j - T_i \ge D_{ij} \text{ при всех } (ij);$$

$$T_0 = 0, \quad T_{\text{вых}} = T.$$

Пример 9.5

Пусть для рассмотренного выше сетевого графика заданы параметры стоимости и продолжительности работ (табл. 9.6) и предельное время $T = 25$.

Таблица 9.6 – Сетевой график

$i-j$	A_{ij}	B_{ij}	D_{ij}	$i-j$	A_{ij}	B_{ij}	D_{ij}	$i-j$	A_{ij}	B_{ij}	D_{ij}
0–1	10	100	2	2–4	4	140	7	5–6	5	50	2
0–2	3	40	1	3–4	2	90	6	4–7	9	180	4
0–3	5	150	5	1–5	1	80	6	5–7	5	140	7
1–3	2	70	2	3–5	5	120	2	6–7	9	200	4
2–3	6	170	2	1–6	2	60	5	–	–	–	–

Минимизируемую функцию Z можно преобразовать к виду (обычное приведение подобных):

$$Z = \sum_{(ij)} B_{ij} - \sum_{k} A_k T_k,$$

где

$$A_k = \sum_{i} A_{ik} - \sum_{j} A_{kj},$$

и задачу минимизации заменить задачей максимизации суммы $\sum_{k} A_k T_k$.

Для нашего примера

$$A_0 = -18;\ A_1 = 5;\ A_2 = -7;\ A_3 = 6;\ A_4 = -3;\ A_5 = -4;\ A_6 = -2;\ A_7 = 23.$$

И возникает задача максимизации функции

$$-18T_0 + 5T_1 - 7T_2 + 6T_3 - 3T_4 - 4T_5 - 2T_6 + 23T_7$$

при условиях:

$$T_1 - T_0 \ge 2 T_3 - T_1 \ge 2 T_4 - T_3 \ge 6 T_6 - T_1 \ge 5 T_7 - T_5 \ge 7$$

$$T_2 - T_0 \ge 1 T_3 - T_2 \ge 2 T_4 - T_1 \ge 6 T_6 - T_5 \ge 2 T_7 - T_6 \ge 4$$

$$T_3 - T_0 \ge 5 T_4 - T_2 \ge 7 T_5 - T_3 \ge 2 T_7 - T_4 \ge 2$$

$$T_0 = 0,\quad T_7 \le 25.$$

Решение этой задачи симплексным методом дает оптимальные времена наступления событий $T_1 = 11;\ T_2 = 1;\ T_3 = 15;\ T_4 = 21;\ T_5 = 17;\ T_6 = 19$ и значение $Z = 1\,146.$

Легко видеть, что использование симплексного метода требует значительных затрат энергии (даже в случае модифицированного симплекс-метода). Поэтому рассмотрим решение задачи по *алгоритму И. А. Радчик*, являющемуся модификацией *венгерского метода Форда – Фалкерсона*.

Поставленная задача сводится к максимизации

$$F(T) = \sum_k A_k T_k,$$

при условиях:

$$T_i - T_j \le -D_{ij} \text{ при всех } ij;$$

$$T_n - T_0 \le T, \;\; T_0 = 0.$$

(индекс n соответствует выходу графика).

Сопряженная (двойственная) задача состоит в максимизации

$$G(X) = -\sum_{(ij)} D_{ij} X_{ij} + T \times X_{n0},$$

при ограничениях:

$$\sum_j X_{ij} - \sum_k X_{ki} = A_i \text{ при всех } i,$$

$$X_{ij} \ge 0 \text{ при всех } ij, \;\; X_{n0} > 0.$$

Величину X_{ij} можно трактовать как количество вещества, протекающего по дуге $i - j$ в единицу времени, и значения A_i – как разницу между притоком и оттоком вещества в вершине i.

Соответственно возникает задача о потоке в сети, определенной сетевым графиком и дополненной дугой от выхода до входа.

На предварительном этапе отыскивается какое-нибудь допустимое решение исходной задачи. Например, оценки T_j берутся равными минимальным временам наступления событий (значение T_n берется равным T).

Затем строится матрица значений $R_{ij} = T_j - T_i - D_{ij}$ (значение $R_{n0} = 0$).

На каждом очередном шаге отыскивается путь от какого-то источника $(A_i > 0)$ до какого-то стока $(A_i < 0)$.

Для этого отмечаем строки с $A_i > 0$ символом * (эти же метки переносим на столбцы). В строке *, например i_0, берем клетки с $R_{i0,j} = 0$ и отмечаем ранее не отмеченные столбцы индексами $(i_0, b_1 = A_{i0})$. Затем по значениям $X_{i0,j} < 0$ отмечаем ранее не отмеченные столбцы индексами $\left(i_0, b_1 = \min\left(A_{i0}, \left|X_{i0,j}\right|\right)\right)$. Метки столбцов переносим на строки.

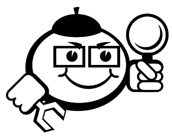
Затем выбираем отмеченную индексом (i_0, b_1) строку i и столбцы с значениями $R_{ij} = 0$ и значениями $X_{ij} < 0$ метим индексами $(i, b_2 = b_1)$ или $\left(i, b_2 = \min\left(b_1, \left|X_{ij}\right|\right)\right)$. Метки столбцов переносим на строки и т. д., пока не будет отмечен некоторый сток или не обнаружится невозможность дальнейшего отмечания.

В первом случае, если некий j_0-й сток помечен индексом (k, b_s), отыскиваем величину $V = \min\left(b_s, A_{j0}\right)$ и обратным ходом по первому из индексов отыскиваем путь, который привел к этому стоку. Значения X_{ij} на этом пути увеличиваем на V и симметричные значения уменьшаем на V. Уменьшаем на V значение A_{i0} и увеличиваем на V значение A_{j0}.

Во втором случае отыскиваем величину H, равную минимальному из значений R_{ij}, лежащих на пересечении отмеченных строк и неотмеченных столбцов.

Если вход и выход помечены, вычитаем H из всех R_{ij}, находящихся в не помеченных столбцах, и прибавляем к находящимся в не помеченных строках. Уменьшаем на H и значения T_j для непомеченных столбцов. Если вход и выход не помечены, то вычитание H проводим для помеченных строк и помеченных столбцов, причем значения T_j для помеченных столбцов увеличиваем на H.

Алгоритм завершает работу, когда все A_j станут равными нулю, т. е. все источники опустошены и все стоки насыщены.



Пример 9.6

Строим начальную таблицу значений R_{ij} на основе ранее найденных оценок T_j^0 (значения X_{ij} равны 0) (табл. 9.7).

Отмечаем «источники» 1, 3 и 7 символом *. В строке 1 обнаруживается $R_{15} = 0$ и 5-й столбец метим индексом (1,5) (из пункта 1 поток емкости $b_1 = 5$). Этот столбец соответствует стоку с потребностью 4. Таким образом найден путь 1–5 с величиной потока $V = \min(5, |-4|)$.

Увеличиваем на V значение A_5 и X_{15}, уменьшая A_5 и X_{51} (значения X_i записываем в этой же таблице под дробной чертой (табл. 9.8)).

Таблица 9.7 – Расчетная таблица

$i \backslash j$	0 (—)	1 (*)	2 (—)	3 (*)	4 (—)	5 (1,5)	6 (—)	7 (*)	A_i	
0		0	0	0					–18	
1				1		0	3		5	*
2				2	3				–7	
3					0	1			6	*
4								10	–3	
5							0	10	–4	1,5
6								11	–2	
7	0								23	*
T_j	0	2	1	5	11	8	10	25		

Таблица 9.8 – Расчетная таблица

$i \backslash j$	0 (—)	1 (*)	2 (—)	3 (*)	4 (—)	5 (1,1)	6 (5,1)	7 (*)	A_i	
0		0	0	0					–18	
1				1		0/4	3		1	*
2				2	3				–7	
3					0	1			6	*
4								10	–3	
5		/–4					0	10	0	1,1
6								11	–2	5,1
7	0								23	*
T_j	0	2	1	5	11	8	10	25		

В полученной матрице обнаруживается путь 1–5–6 интенсивности 1 и $V = \min(1, |-2|) = 1$. Увеличиваем X_{15}, X_{56}, A_6 и уменьшаем X_{51}, X_{65}, A_1 (табл. 9.9).

Таблица 9.9 – Расчетная таблица

	7,23			*	3,6			*		
$i \backslash j$	0	1	2	3	4	5	6	7	A_i	
0		0	0	0					–18	7,23
1				1		0/5	3		0	
2				2	3				–7	
3					0	1			6	*
4								10	–3	3,6
5		/–5					0/1	10	0	
6						/–1		11	–1	
7	0								23	*
T_j	0	2	1	5	11	8	10	25		

В полученной матрице обнаруживается путь 3–4 интенсивности 3 и путь 7–0 интенсивности 18.

Корректируем X_{34}, X_{43}, X_{07}, X_{70} и A_4, A_3, A_0, A_7 (табл. 9.10).

Таблица 9.10 – Расчетная таблица

	7,5		0,5	*	3,3			*		
$i \backslash j$	0	1	2	3	4	5	6	7	A_i	
0		0	0	0				/–18	0	7,5
1				1		0/5	3		0	
2				2	3				–7	0,5
3					0/3	1			3	*
4				/–3				10	0	3,3
5		/–5					0/1	10	0	
6						/–1		11	–1	
7	0/18								5	*
T_j	0	2	1	5	11	8	10	25		

В полученной матрице обнаруживается путь 7–0–2 интенсивности 5 и корректируем X_{70}, X_{02}, X_{07}, X_{20} и A_7, A_2 (табл. 9.11).

Таблица 9.11 – Расчетная таблица

$i \backslash j$	0 (—)	1 (—)	2 (—)	3 (*)	4 (3,3)	5	6	7	A_i	
0		0	0/5	0				/–23	0	
1				1		0/5	3		0	
2	/–5			2	3				–2	
3					0/3	1			3	*
4				/–3				10	0	3,3
5		/–5					0/1	10	0	
6						/–1		11	–1	
7	0/23								0	
T_j	0	2	1	5	11	8	10	25		

В полученной матрице обнаруживается невозможность достижения стоков 2 и 6.

Отыскиваем в строках 3, 4 в непомеченных столбцах минимальный из R_{ij}: $H = \min(R_{35}, R_{47}) = 1$. Вычитаем H из строк 3, 4 и добавляем к столбцам 3,4 (в том числе и к значениям T_3, T_4 (табл. 9.12)).

Таблица 9.12 – Расчетная таблица

$i \backslash j$	0	1	2	3 (*)	4 (3,3)	5 (3,3)	6 (5,3)	7	A_i	
0		0	0/5	0				/–23	0	
1				1		0/5	3		0	
2	/–5			3	4				–2	
3					0/3	0			3	*
4				/–3				9	0	3,3
5		/–5					0/1	10	0	3,3
6						/–1		11	–1	5,3
7	0/23								0	
T_j	0	2	1	6	12	8	10	25		

В полученной матрице обнаруживается путь 3–5–6 интенсивности 1 и корректируем X_{35}, X_{56}, X_{53}, X_{65}, A_3, A_6 (табл. 9.13).

В отличие от предыдущих таблиц здесь используем отмечание не только по $R_{ij}=0$, но и по $X_{ij}<0$ (столбец 1). Добраться до стока 2 невозможно. Потому отыскиваем $H=\min R_{ij}$ при $i=1,3,4,5,6$ и $j=0,2,7$ $\left(H=\min[9,10,11]=9\right)$ и вычитаем 9 из отмеченных строк с добавлением к отмеченным столбцам (табл. 9.14).

Таблица 9.13 – Расчетная таблица

		5,2		*	3,2	3,2	5,2			
i \ j	0	1	2	3	4	5	6	7	A_i	
0		0	0/5	1				/–23	0	
1				2		0/5	3		0	5,2
2	/–5			3	4				–2	
3					0/3	0/1			2	*
4				/–3				9	0	3,2
5		/–5		/–1			0/2	10	0	3,2
6						/–2		11	0	5,2
7	0/23								0	
T_j	0	2	1	6	12	8	10	25		

Таблица 9.14 – Расчетная таблица

	7,2	5,2	0,2	*	3,2	3,2	5,2	4,2		
i \ j	0	1	2	3	4	5	6	7	A_i	
0		9	0/5	10				/–23	0	7,2
1				2		0/5	3		0	5,2
2	/–5			12	13				–2	0,2
3					0/3	0/1			2	*
4				/–3				0	0	3,2
5		/–5		/–1			0/2	1	0	3,2
6						/–2		2	0	5,2
7	0/23								0	4,2
T_j	0	11	1	15	21	17	19	25		

Здесь обнаруживается путь 3–4–7–0–2 с интенсивностью 2 и после корректуры значения A_3 и A_2 обращаются в нуль, что служит признаком конца решения задачи.

Оптимальный безрезервный план определяется временами наступления событий:

$$T_0 = 0,\ T_1 = 11,\ T_2 = 1,\ T_3 = 15,\ T_4 = 21,\ T_5 = 17,\ T_6 = 19,\ T_7 = 25.$$

. . . .

В рассмотренной задаче предполагалась ограниченность продолжительности работ снизу. При наличии ограничения сверху задача может быть поставлена в виде:
минимизировать функцию

$$Z = \sum_{(ij)} \left[-A_{ij} t_{ij} + B_{ij} \right]$$

при условиях:

$$T_j - T_i \ge t_{ij} \text{ при всех } ij,\ D_{ij} \le t_{ij} \le W_{ij} \text{ при всех } ij,\ T_0 = 0,\ T_{\text{вых}} = T.$$

Решение задачи дает *оптимальный по стоимости резервный план*. Если считать T переменной величиной, то поставленная задача становится задачей параметрического линейного программирования.

Определенный интерес представляет и задача минимизации времени выполнения проекта при заданной его стоимости.

Минимизировать T_n при условиях:

$$T_j - T_i \ge t_{ij} \text{ при всех } ij,$$

$$D_{ij} \le t_{ij} \le W_{ij} \text{ при всех } ij,$$

$$T_0 = 0,$$

$$\sum_{(ij)} \left[-A_{ij} t_{ij} + B_{ij} \right] \le C.$$

9.5 Проблемы применения систем сетевого планирования

Выше мы ориентировались на сетевой график некоторого самостоятельного проекта, на выполнение которого направлены усилия коллектива исполнителей. В реальности один и тот же коллектив выполняет работы «одновременно» по нескольким проектам и даже при идеальной отработке графиков каждого из них нет уверенности в выполнении всех проектов, т. к. выполнение многих тем приходится на один и тот же промежуток времени. При выполнении таких «многотемных» разработок наряду с оценкой работ по времени приходится учитывать трудоемкость работ (количество человеко-часов или количество специалистов в этой области), мощность подразделений исполнителей, возможность выполнения работы подразделениями.

Поэтому сначала для каждой темы разрабатывают сетевой график и проводят оценки не только по времени, но и по исполнителям. Работы из всех тем сортируют по подразделениям и накладывают на календарь.

Оценивают возможности подразделения и, если все работы выполнить в данный период невозможно, часть из них переносят на более поздние сроки с соответствующими отметками в исходных графиках.

Сетевые графики могут иметь стохастическую структуру по оценке времени выполнения работ. Здесь по заданным пессимистической и оптимистической оценкам отыскивают математическое ожидание и дисперсию продолжительности работ:

$$T_{ij} = \frac{3A_{ij} + 2B_{ij}}{5}; \qquad D_{ij} = \left| \frac{B_{ij} - A_{ij}}{5} \right|^2.$$

После традиционной обработки сетевого графика отыскивается дисперсия длины критического пути как сумма дисперсий составляющих его работ. Оценка вероятности той или иной продолжительности выполнения проекта (или отдельной работы) проводится на основе функции нормального распределения.

Так, при полученной длине критического пути в 38 дней и дисперсии 6,31, т. е. стандартном отклонении 2,51, вероятность того, что фактическое время выполнения проекта лежит в интервале от 35 до 42 дней, определится как разность значений функции нормального распределения при аргументах $(35-38)/2{,}51$ и $(42-38)/2{,}51$ и составит 0,822.

Контрольные вопросы по главе 9

1. В каких областях применяется сетевое планирование?
2. Что такое сетевой график?
3. Что представляет собой фиктивная работа?
4. Что такое ранг события?
5. Что такое критический путь?
6. Что такое резерв?
7. Что такое напряженность?

8. Как выглядит диаграмма Ганта?
9. Как минимизировать стоимость проекта?

Заключение

Характерной особенностью научно-технического прогресса в развитых странах является возрастание роли экономической науки. Экономика выдвигается на первый план именно потому, что она в решающей степени определяет эффективность и приоритетность направлений научно-технического прогресса, раскрывает широкие пути реализации экономически выгодных достижений.

Применение математики в экономической науке дало толчок развитию как самой экономической науки, так и прикладной математики, в части методов экономико-математической модели. Пословица говорит: «Семь раз отмерь – один раз отрежь». Расчёты по моделям противостоят волевым решениям, поскольку позволяют заранее оценить последствия каждого решения, отбросить недопустимые варианты и рекомендовать наиболее удачные.

На всех уровнях управления, во всех отраслях используются методы экономико-математического моделирования. Выделим условно следующие направления их практического применения, по которым получен уже большой экономический эффект.

Первое направление – прогнозирование и перспективное планирование. Прогнозируются темпы и пропорции развития экономики, на их основе определяются темпы и факторы роста национального дохода, его распределение на потребление и накопление и т. д. Важным моментом является использование экономико-математических методов не только при составлении планов, но и в деле оперативного руководства по их реализации.

Второе направление – разработка моделей, которые используются как инструмент согласования и оптимизации плановых решений, в частности это межотраслевые и межрегиональные балансы производства и распределения продукции. По экономическому содержанию и характеру информации выделяют балансы стоимостные и натурально-продуктовые, каждый из которых может быть отчетным и плановым.

Третье направление – использование экономико-математических моделей на отраслевом уровне (выполнение расчетов оптимальных планов отрасли, анализ с помощью производственных функций, прогнозирование основных производственных пропорций развития отрасли). Для решения задачи размещения и специализации предприятия, оптимального прикрепления к поставщикам или потребителям и т. д. используются модели оптимизаций двух типов: в одних для заданного объёма производства продукции требуется найти вариант реализации плана с наименьшими затратами, в других требуется определить масштабы производства и структуру продукции с целью получения максимального эффекта. В продолжение расчетов осуществляется переход от статистических моделей к динамическим и от моделирования отдельных отраслей к оптимизации многоотраслевых комплексов. Если раньше были попытки создать единую модель отрасли, то теперь наиболее перспективным считается использование комплексов моделей, взаимоувязанных как по вертикали, так и по горизонтали.

Четвертое направление – экономико-математическое моделирование текущего и оперативного планирования промышленных, строительных, транспортных и других объединений, предприятий и фирм. Область практического применения моделей включает также подразделения сельского хозяйства, торговли, связи, здравоохранения, охрану природы и т. д. В машиностроении используется большое количество разнообразных моделей, наиболее «отлаженными» из которых являются оптимизационные, позволяющие определить производственные программы и наиболее рациональные варианты использования ресурсов, распределить производственную программу во времени и эффективно организовать работу внутризаводского транспорта, существенно улучшить загрузку оборудования и разумно организовать контроль продукции и др.

Пятое направление – территориальное моделирование, начало которому положила разработка отчетных межотраслевых балансов некоторых регионов в конце 1950-х гг.

В качестве шестого направления можно выделить экономико-математическое моделирование материально-технического обеспечения, включающее оптимизацию транспортно-экономических связей и уровня запасов.

К седьмому направлению относятся модели функциональных блоков экономической системы: движение населения, подготовка кадров, формирование денежных доходов и спроса на потребительские блага и др.

Особенно большую роль приобретают экономико-математические методы по мере внедрения информационных технологий во всех областях жизнедеятельности.

Литература

1. Орлова И. В. Экономико-математические методы и модели: компьютерное моделирование / И. В. Орлова, В. А. Половников. – М. : Вузовский учебник ; ИНФРА-М, 2011. – 389 с.
2. Горбунова Р. И. Экономико-математические методы и модели / Р. И. Горбунова, С. И. Макаров, М. В. Мищенко. – М. : КНОРУС, 2009. – 540 с.
3. Замков О. О. Математические методы в экономике / О. О. Замков, А. В. Толстопятенко, Ю. Н. Черемных. – М. : Дело и Сервис, 2009. – 384 с.
4. Просветов Г. И. Математические методы в экономике / Г. И. Просветов. – М. : РДЛ, 2005. – 158 с.
5. Федосеев В. В. Экономико-математические методы и прикладные модели : учеб. пособие для вузов / В. В. Федосеев, А. Н. Гармаш, И. В. Орлова и др. ; под ред. В. В. Федосеева. – 2-е изд. – М. : ЮНИТИ, 2005. – 304 с.
6. Волгина О. А., Голодная Н. Ю., Одияко Н. Н., Шуман Г. И. Математическое моделирование экономических процессов и систем : учеб. пособие / О. А. Волгина, Н. Ю. Голодная, Н. Н. Одияко, Г. И. Шуман. – 2-е изд. – М. : КНОРУС, 2012. – 200 с.
7. Бережная Е. В. Математические методы моделирования экономических систем / Е. В. Бережная, В. И. Бережной. – М. : Финансы и статистика, 2006. – 432 с.
8. Просветов Г. И. Математические методы и модели в экономике : учеб.-практ. пособие / Г. И. Просветов. – М. : Альфа-Пресс, 2008. – 344 с.
9. Кундышева Е. С. Математическое моделирование в экономике : учеб. пособие / Е. С. Кундышева ; под науч. ред. проф. Б. А. Суслакова. – М. : Издательско- торговая корпорация «Дашков и К», 2004. – 352 с.
10. Минюк Е. А. Математические методы и модели в экономике : учеб. пособие / Е. А. Минюк, Е. А. Ровба, К. К. Кузьмич. – Минск : ТетраСистемс, 2002. – 432 с.

11. Шелобаев С. И. Математические методы и модели в экономике, финансах, бизнесе : учеб. пособие для вузов / С. И. Шелобаев. – 2-е изд. – М. : ЮНИТИ-ДАНА, 2005. – 287 с.
12. Вентцель Е. С. Исследование операций: Задачи, принципы, методология / Е. С. Вентцель. – М. : Высшая школа, 2001. – 208 с.
13. Имитационное моделирование экономических процессов : учеб. пособие / под ред. А. А. Емельянова. – 2-е изд. – М. : Финансы и статистика, 2006. – 416 с.
14. Канторович Л. В. Экономический расчёт наилучшего использования ресурсов / Л. В. Канторович. – М. : Изд-во АН СССР, 1960. – 346 с.
15. Колемаев В. А. Математическая экономика : учебник для вузов / В. А. Колемаев. – М. : ЮНИТИ, 1998. – 240 с.
16. Моделирование экономических процессов : учебник для студентов вузов, обучающихся по специальностям экономики и управления (060000) / под ред. М. В. Грачёвой, Л. Н. Фадеевой, Ю. И. Черемных. – М. : ЮНИТИ-ДАНА, 2005. – 351 с.
17. Светлов Н. М. Альбом наглядных пособий по курсу «Моделирование микро- и макроэкономических процессов» / Н. М. Светлов – М. : ФГОУ ВПО РГАУ–МСХА им. К. А. Тимирязева, 2006. – 205 с.
18. Численные методы / Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков. – 4-е изд. – М. : БИНОМ. Лаборатория знаний, 2006. – 636 с.
19. Лебедев В. В. Математическое моделирование социально-экономических процессов / В. В. Лебедев. – М. : Изограф, 1997. – 222 с.
20. Ширяев А. Н. Основы стохастической финансовой математики / А. Н. Ширяев. – М. : ФАЗИС, 2004. – Т. 1–2.
21. Маркин Ю. П. Математические методы и модели в экономике : учеб. пособие для вузов / Ю. П. Маркин. – М., 2007.
22. Подопригора И. В. Общая теория статистики : учеб. пособие / И. В. Подопригора. – Томск : ФДО ТУСУР, 2015. – 110 с.
23. Росс С. И. Математическое моделирование и исследование национальной экономики : учеб. пособие / С. И. Росс. – СПб. : СПб ГУ ИТМО, 2006. – 61 с.
24. Стариков А. В. Экономико-математическое и компьютерное моделирование : учеб. пособие / А. В. Стариков. – Воронеж : ГОУ ВПО «ВГЛТА», 2008. – 132 с.

25. Ломкова Е. Н. Экономико-математические модели управления производством (теоретические аспекты) : учеб. пособие / Е. Н. Ломкова, А. А. Эпов. – Волгоград : ВолгГТУ, 2005. – 67 с.
26. Колемаев В. А. Экономико-математическое моделирование. Моделирование макроэкономических процессов и систем : учебник / В. А. Колемаев. – М. : ЮНИТИ-ДАНА, 2005. – 295 с.
27. Ильченко А. Н. Экономико-математические методы / А. Н. Ильченко. – М. : Финансы и статистика, 2006. – 287 с.
28. Коробов П. Н. Математическое программирование и моделирование экономических процессов / П. Н. Коробов. – М. : ДНК, 2006. – 375 с.
29. Просветов Г. И. Математические модели в экономике / Г. И. Просветов. – СПб.: РДЛ, 2006. – 151 с.
30. Пелих А. С. Экономико-математические методы и модели в управлении производством / А. С. Пелих, Л. Л. Терехов, Л. А. Терехова. – Ростов н/Д : Феникс, 2005. – 246 с.
31. Федосеев В. В. Экономико-математические методы и прикладные модели / В. В. Федосеев. – М. : ЮНИТИ, 2005. – 391 с.
32. Фомин Г. П. Математические методы и модели в коммерческой деятельности / Г. П. Фомин. – М. : Финансы и статистика, 2005. – 615 с.
33. Шелобаев С. И. Экономико-математические методы и модели / С. И. Шелобаев. – М. : ЮНИТИ, 2005. – 286 с.
34. Конюховская Л. В. Математические методы исследования операций в экономике : учеб. пособие / Л. В. Конюховская. – СПб. : Питер, 2000. – 208 с.

Приложение А. Значения интеграла Лапласа

$$p(t) = \frac{1}{\sqrt{2\pi}} \int_{-t}^{+t} e^{-\frac{t^2}{2}} dt.$$

t	Сотые доли									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0,00	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,10	0,0797	0,0876	0,0955	0,1034	0,1113	0,1192	0,1271	0,1350	0,1428	0,1507
0,20	0,1585	0,1663	0,1741	0,1819	0,1897	0,1974	0,2051	0,2128	0,2205	0,2282
0,30	0,2358	0,2434	0,2510	0,2586	0,2661	0,2737	0,2812	0,2886	0,2961	0,3035
0,40	0,3108	0,3182	0,3255	0,3328	0,3401	0,3473	0,3545	0,3616	0,3688	0,3759
0,50	0,3829	0,3899	0,3969	0,4039	0,4108	0,4177	0,4245	0,4313	0,4381	0,4448
0,60	0,4515	0,4581	0,4647	0,4713	0,4778	0,4843	0,4907	0,4971	0,5035	0,5098
0,70	0,5161	0,5223	0,5285	0,5346	0,5407	0,5467	0,5527	0,5587	0,5646	0,5705
0,80	0,5763	0,5821	0,5878	0,5935	0,5991	0,6047	0,6102	0,6157	0,6211	0,6265
0,90	0,6319	0,6372	0,6424	0,6476	0,6528	0,6579	0,6629	0,6680	0,6729	0,6778
1,00	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,10	0,7287	0,7330	0,7373	0,7415	0,7457	0,7499	0,7540	0,7580	0,7620	0,7660
1,20	0,7699	0,7737	0,7775	0,7813	0,7850	0,7887	0,7923	0,7959	0,7995	0,8029
1,30	0,8064	0,8098	0,8132	0,8165	0,8198	0,8230	0,8262	0,8293	0,8324	0,8355
1,40	0,8385	0,8415	0,8444	0,8473	0,8501	0,8529	0,8557	0,8584	0,8611	0,8638
1,50	0,8664	0,8690	0,8715	0,8740	0,8764	0,8789	0,8812	0,8836	0,8859	0,8882
1,60	0,8904	0,8926	0,8948	0,8969	0,8990	0,9011	0,9031	0,9051	0,9070	0,9090
1,70	0,9109	0,9127	0,9146	0,9164	0,9181	0,9199	0,9216	0,9233	0,9249	0,9265
1,80	0,9281	0,9297	0,9312	0,9328	0,9342	0,9357	0,9371	0,9385	0,9399	0,9412
1,90	0,9426	0,9439	0,9451	0,9464	0,9476	0,9488	0,9500	0,9512	0,9523	0,9534
2,00	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9615	0,9625	0,9634
2,10	0,9643	0,9651	0,9660	0,9668	0,9676	0,9684	0,9692	0,9700	0,9707	0,9715
2,20	0,9722	0,9729	0,9736	0,9743	0,9749	0,9756	0,9762	0,9768	0,9774	0,9780
2,30	0,9786	0,9791	0,9797	0,9802	0,9807	0,9812	0,9817	0,9822	0,9827	0,9832
2,40	0,9836	0,9840	0,9845	0,9849	0,9853	0,9857	0,9861	0,9865	0,9869	0,9872
2,50	0,9876	0,9879	0,9883	0,9886	0,9889	0,9892	0,9895	0,9898	0,9901	0,9904
2,60	0,9907	0,9909	0,9912	0,9915	0,9917	0,9920	0,9922	0,9924	0,9926	0,9929
2,70	0,9931	0,9933	0,9935	0,9937	0,9939	0,9940	0,9942	0,9944	0,9946	0,9947
2,80	0,9949	0,9950	0,9952	0,9953	0,9955	0,9956	0,9958	0,9959	0,9960	0,9961
2,90	0,9963	0,9964	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972
3,00	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9978	0,9979	0,9979	0,9980
3,10	0,9981	0,9981	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986
3,20	0,9986	0,9987	0,9987	0,9988	0,9988	0,9988	0,9989	0,9989	0,9990	0,9990
3,30	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,40	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995	0,9995
3,50	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9997

Приложение Б. Значения *t*-критерия Стьюдента

При уровне значимости α: 0,10; 0,05; 0,01.

Число степеней свободы ν	**α**			**Число степеней свободы** ν	**α**		
	0,1	0,05	0,01		0,1	0,05	0,01
1	6,314	12,706	63,66	18	1,734	2,101	2,878
2	2,92	4,3027	9,925	19	1,729	2,093	2,861
3	2,353	3,1825	5,841	20	1,725	2,086	2,845
4	2,132	2,7764	4,604	21	1,721	2,08	2,831
5	2,015	2,5706	4,032	22	1,717	2,074	2,819
6	1,943	2,4469	3,707	23	1,714	2,069	2,807
7	1,895	2,3646	3,5	24	1,711	2,064	2,797
8	1,86	2,306	3,355	25	1,708	2,06	2,787
9	1,833	2,2622	3,25	26	1,706	2,056	2,779
10	1,813	2,2281	3,169	27	1,703	2,052	2,771
11	1,796	2,201	3,106	28	1,701	2,048	2,763
12	1,782	2,1788	3,055	29	1,699	2,045	2,756
13	1,771	2,1604	3,012	30	1,697	2,042	2,75
14	1,761	2,1448	2,977	40	1,684	2,021	2,705
15	1,753	2,1315	2,947	60	1,671	2	2,66
16	1,746	2,1199,	2,921	120	1,658	1,98	2,617
17	1,74	2,1098	2,898	∞	1,645	1,96	2,576

Приложение В. Значения *F*-критерия Фишера

При уровне значимости $\alpha = 0{,}05$.

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
1	161,5	200	215,7	224,6	230,2	234	238,9	243,9	249	254,3
2	18,5	19	19,16	19,25	19,3	19,33	19,37	19,41	19,45	19,5
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,9	2,71
10	4,96	4,1	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,2	3,09	2,95	2,79	2,61	2,4
12	4,75	3,88	3,49	3,26	3,11	3	2,85	2,69	2,5	2,3
13	4,67	3,8	3,41	3,18	3,02	2,92	2,77	2,6	2,42	2,21
14	4,6	3,74	3,34	3,11	2,96	2,85	2,7	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,9	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,2	2,96	2,81	2,7	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,9	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,1	2,87	2,71	2,6	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,3	3,44	3,05	2,82	2,66	2,55	2,4	2,23	2,03	1,78
23	4,28	3,42	3,03	2,8	2,64	2,53	2,38	2,2	2	1,76
24	4,26	3,4	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,6	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,3	2,13	1,93	1,67
28	4,2	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,7	2,54	2,43	2,28	2,1	1,9	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2	1,79	1,52
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,4	2,29	2,13	1,95	1,72	1,44

ν_2 \ ν_1	1	2	3	4	5	6	8	12	24	∞
60	4	3,15	2,76	2,52	2,37	2,25	2,1	1,92	1,7	1,39
70	3,98	3,13	2,74	2,5	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,1	2,71	2,47	2,32	2,2	2,04	1,86	1,64	1,28
100	3,94	3,09	2,7	2,46	2,3	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,6	1,21
150	3,9	3,06	2,66	2,43	2,27	2,16	2	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,8	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79.	1,55	1,1
400	3,86	3,02	2,63	2,4	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3	2,61	2,38	2,22	2,1	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,6	2,37	2,21	2,09	1,94	1,75	1,52	

Глоссарий

Адекватность модели – требование к модели, состоящее в ее способности воспроизводить свойства, состояние и поведение исследуемого объекта с достаточной для поставленных целей точностью и в достаточно широком диапазоне изменения ее состояния и состояния ее среды. Достаточным условием адекватности модели является ее гомоморфизм исследуемому объекту.

Аналитические модели – математические модели, разрабатываемые для исследования структуры моделируемой системы. В экономико-математическом моделировании, как правило, имеют целью выявление резервов повышения эффективности функционирования моделируемой системы либо факторов, влияющих на исследуемые показатели хозяйственной деятельности, а также формы и степени их влияния.

Апостериорное решение – в стохастических двухэтапных моделях – вектор оптимальных значений переменных, характеризующих плановые задания, выполняемые после поступления информации о наступлении определённого случайного события, влияющего на хозяйственные результаты. Для каждого варианта (исхода) случайных событий предусматривается отдельное апостериорное решение.

Априорное решение – в стохастических двухэтапных моделях – вектор оптимальных значений переменных, характеризующих плановые задания, требующие выполнения до поступления информации о случайных событиях, влияющих на ожидаемые хозяйственные результаты.

Баланс – в экономико-математическом моделировании – уравнение или неравенство, устанавливающее соответствие между источниками ресурса и направлениями его использования. Уравнения используются, если имеющиеся источники ресурса должны быть использованы без остатка либо если наличие остатка влияет на экономический эффект (требует затрат на утилизацию либо может быть с выгодой продан на рынке). В противном случае балансы записываются в форме неравенств, согласно которым совокупное использование ресурса не превышает размера имеющихся его источников.

Векторное программирование – (1) раздел математики, исследующий методы решения задач векторного программирования; (2) формализм, используемый для представления знаний о структуре моделируемых объектов в форме задачи векторного программирования.

Взаимность (в математическом программировании) – свойство задач выпуклого (в том числе линейного) программирования, состоящее в инвариантности оптимального решения и инвариантности с точностью до масштаба множителей Лагранжа к замене целевой функции любым эффективным ограничением при дополнительном условии, что значение прежней целевой функции останется равным оптимальному.

Вычислительный эксперимент – метод исследования явления, процесса или машины, для которых разработана компьютерная модель.

Гомоморфизм – одностороннее отношение подобия структур двух систем. Система называется гомоморфной другой системе, если можно указать отношение, отображающее любой допустимый вектор ее переменных на вектор некоторых выбранных переменных другой системы, компоненты которого являются компонентами некоторого допустимого вектора ее состояния. Обязательное требование к модели – ее гомоморфизм моделируемому объекту.

Двойственная оценка ограничения – величина, характеризующая прирост значения целевой функции задачи математического программирования при малом изменении величины свободного члена данного ограничения; частная производная оптимального значения целевой функции, рассматриваемого в качестве функции свободных членов ограничений задачи математического программирования, по величине свободного члена данного ограничения. Измеряется в единицах измерения целевой функции в расчёте на единицу измерения ограничения.

Дескриптивные модели – модели, целью которых является формализованное представление знания о структуре моделируемого объекта.

Динамическая модель – математическая модель, описывающая развитие процесса во времени для определения или улучшения их характеристик; для рационализации способов их построения; для управления и прогнозирования.

Задача векторного программирования – задача отыскания оптимума по Парето заданной вектор-функции на заданном множестве допустимых значений переменных.

Имитационная модель – математическая модель, воспроизводящая поведение исследуемого объекта и применяемая для постановки компьютерных экспериментов, выявляющих особенности функционирования объекта при различных внешних условиях и управляющих воздействиях.

Исторические модели – экономические модели, с помощью которых можно анализировать изменения и ситуации в реальной жизни.

Компьютерная имитация – реализованная на ЭВМ математическая модель, оперирующая системой динамических уравнений.

Компьютерная модель – математическая модель, оперирующая нечисленными алгоритмами и реализованная на ЭВМ.

Линейное программирование – (1) раздел математического программирования, исследующий задачи отыскания экстремума линейной функции на множестве допустимых значений переменных, заданном системой линейных уравнений и (или) неравенств; (2) формализм, используемый для представления знаний о структуре моделируемых объектов в форме задачи отыскания экстремума линейной функции на множестве допустимых значений переменных, заданном системой линейных уравнений и (или) неравенств.

Макроэкономическая модель – экономико-математическая модель, в которой не выделяются переменные, описывающие отдельных хозяйствующих субъектов (предприятия, отрасли), составляющих моделируемую хозяйственную систему, и которая отражает только связи, присущие этой системе как целому.

Математическая модель – совокупность математических зависимостей, гомоморфная исследуемой системе и используемая для суждения о ее свойствах и поведении.

Математическое моделирование – метод исследования реальных объектов при помощи постановки экспериментов на их математических моделях.

Математическое программирование – (1) раздел математики, исследующий методы решения задач отыскания экстремума на заданном множестве допустимых значений переменных; (2) формализм, используемый для представления знаний о структуре моделируемых объектов в форме задачи отыскания экстремума на заданном множестве допустимых значений переменных.

Матричная модель – экономико-математическая модель, предназначенная для планирования и анализа производства и распределения продукции на разных уровнях материального производства.

Межотраслевой баланс – экономико-математическая балансовая модель в виде системы линейных уравнений, характеризующих связи между выпуском продукции в одной отрасли (в стоимостном измерении) и затратами продукции всех участвующих отраслей, необходимыми для обеспечения этого выпуска.

Микроэкономическая модель – экономико-математическая модель, в которой присутствуют переменные, характеризующие различных хозяйствующих субъектов (предприятия, отрасли), составляющих моделируемую хозяйственную систему, и математическое описание связей между этими субъектами.

Моделирование – исследование объектов познания на их моделях. Моделирование предполагает построение и изучение моделей реально существующих предметов, явлений и конструируемых объектов.

Модель – упрощённое подобие реального объекта, используемое для его исследования.

Модель оптимизации – математическая модель, исходящая из того, что некоторые переменные в моделируемых процессах или ситуациях максимизируются или минимизируются.

Модель управления запасами – экономико-математическая модель, позволяющая рассчитать рациональную структуру использования ресурсов.

Неограниченность целевой функции – ситуация, при которой множество допустимых значений переменных задачи математического программирования содержит значения, доставляющие сколь угодно большое значение целевой функции. Если имеет место неограниченность целевой функции, оптимального решения задачи не существует.

Несовместность системы ограничений – ситуация, при которой множество допустимых значений переменных задачи математического программирования пусто вследствие наличия взаимоисключающих уравнений или неравенств, определяющих это множество. Вследствие отсутствия допустимых значений при несовместности системы ограничений оптимального решения задачи не существует.

Объективно обусловленная оценка ресурса (продукции) – величина прироста экономического эффекта, обусловленного малым изменением доступного объёма ресурса или величины планового задания по выпуску продукции. При использовании экономико-математического моделирования численно равна двойственной оценке соответствующего ограничения. Измеряется в единицах измерения экономического эффекта в расчёте на единицу ресурса (продукции).

Оптимальный план – план, доставляющий максимум целевой функции, отражающей выбранный критерий эффективности функционирования объекта планирования при соблюдении требований, заданных в форме системы уравнений и неравенств. Оптимальный план не обязательно является наилучшим планом, подлежащим утверждению и последующему выполнению, поскольку учитывает только те условия хозяйственной деятельности, которые удалось описать в математической форме. Во многих случаях процесс планирования требует использования информации, содержащейся во множестве разнообразных оптимальных планов.

Оптимизационная модель – математическая модель, имеющая форму задачи математического программирования.

Синтетические модели – математические модели, разрабатываемые для проектирования новых, отличающихся от известных, систем с заданными свойствами. К числу синтетических экономико-математических моделей относятся, например, модели машинно-тракторного парка, модели формирования инвестиционных программ и др.

Система математических моделей – совокупность логически, информационно и алгоритмически связанных математических моделей, отражающих существенные закономерности функционирования экономического объекта (экономические, организационные, технологические, финансовые и др.) в реальных условиях среды (П. П. Пастернак).

Системное моделирование – процесс имитации свойств, состояния и поведения во внешней среде систем со сложной или очень сложной структурой в целях управления ими, осуществляемый при помощи системы математических моделей (П. П. Пастернак).

Системный анализ – метод научного познания, нацеленный на установление структуры исследуемой системы. Метод системного анализа является необходимой предпосылкой метода математического моделирования.

Стохастические двухэтапные модели – экономико-математические модели, содержащие переменные, описывающие план, реализуемый до поступления информации о случайных условиях (априорное решение) и варианты планов, зависящих от поступившей информации о случайных условиях (апостериорное решение).

Теоретическая модель – математическая модель, описывающая структуру исследуемого объекта в общем виде, без спецификации конкретных числовых значений параметров.

Устойчивость оптимального плана – свойство математических моделей, имеющих форму задачи линейного программирования, состоящее в неизменности двойственных оценок ограничений при изменениях свободных членов ограничений в определённых пределах и в неизменности значений переменных при изменениях параметров целевой функции в определённых пределах.

Форма представления систем – класс символьных представлений знаний о системе, выделяемый по признаку применимости для решения определённого круга исследовательских или прикладных задач. Например, форма кибернетической системы ориентирована на исследование информационных процессов, опосредующих управление данной системой.

Формализм – (1) знаковая система, используемая для представления знаний; (2) совокупность языковых (изобразительных) и процедурных (вычислительных) средств представления знаний.

Целевая функция – математическое выражение, отражающее выбранный критерий эффективности функционирования исследуемой системы в ее математической модели.

Эконометрические модели – экономико-математические модели, целью которых является установление значений параметров исследуемой экономической системы, не поддающихся непосредственному наблюдению. Как правило, представляют собой эмпирическую спецификацию теоретической модели исследуемой системы, содержащей требуемый параметр, которую оценивают на основе имеющихся эмпирических данных с помощью того или иного статистического метода (например, метода наименьших квадратов, метода оболочки данных, метода максимальной энтропии и т. п.).

Экономико-математическое моделирование – концентрированное выражение наиболее существенных взаимосвязей и закономерностей поведения управляемой системы в математической форме.

Эмпирическая модель – математическая модель, содержащая числовые параметры, значения которых обоснованы данными опыта или наблюдения.