**Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Российская академия народного хозяйства и государственной службы
при Президенте Российской Федерации»**

СИБИРСКИЙ ИНСТИТУТ УПРАВЛЕНИЯ

**А. Л. ОСИПОВ
Е. А. РАПОЦЕВИЧ**

Основы математического моделирования социально-экономических процессов

УЧЕБНОЕ ПОСОБИЕ

для студентов всех форм обучения
по направлению 081100.62 — Государственное и муниципальное управление

**НОВОСИБИРСК
2014**

ББК 65в6я73
О 74-1

Издается в соответствии с планом учебно-методической работы

Рецензенты:

А. А. Добрынин — кандидат физико-математических наук,
доцент кафедры прикладных информационных технологий НГУЭУ;

В. Н. Крупчатников — доктор физико-математических наук,
профессор кафедры информатики и математики
Сибирского института управления — филиала РАНХиГС

Осипов, А. Л.
О 74-1 **Основы математического моделирования социально-экономических процессов** : учеб. пособие / А. Л. Осипов, Е. А. Рапоцевич ; РАНХиГС, Сиб. ин-т упр. — Новосибирск : Изд-во СибАГС, 2014. — 154 с.

ISBN 978-5-8036-0596-6

Учебное пособие по дисциплине федерального компонента «Основы математического моделирования социально-экономических процессов» включает четыре части. Три части посвящены детерминированным методам (линейное программирование, методы прогнозирования, многокритериальные задачи и метод анализа иерархий), стохастическим методам (регрессионный анализ и анализ зависимостей в слабых шкалах) и игровым методам (матричные игры и игры с природой), которые используются для принятия управленческих решений. В четвертой части представлены приложения в социологии.

Предназначено для студентов всех форм обучения по направлению 081100.62 — Государственное и муниципальное управление.

ББК 65в6я73

ISBN 978-5-8036-0596-6

Оглавление

Предисловие

Учебное пособие по дисциплине «Основы математического моделирования социально-экономических процессов» предназначено для студентов всех форм обучения по направлению 081100.62 — Государственное и муниципальное управление. Дисциплина является федеральным компонентом образовательного цикла. В научном сообществе еще нет устоявшегося понимания содержания этой дисциплины, и многие авторы указывают в своих пособиях такие модели, которые применяются в экономике, но которые не слишком интересны для управленцев. Авторы данного пособия считают, что в обучении по направлению «Государственное и муниципальное управление» необходимо шире применять известные подходы, математические модели и методы, которые управленец может использовать при планировании своей деятельности. Исходя из этого в пособие включены три класса методов (детерминированные, стохастические, игровые), выбор которых основывается на условиях ситуации принятия решения.

При выборе оптимального решения, которое основывается на разных критериях, используются теоретико-игровые методы и критерии принятия решения в условиях неопределенности.

В учебном пособии приводятся примеры решения основных типов задач, типовых задач с решениями, которые могут послужить основой для проведения практических занятий, а также задачи, предназначенные для самостоятельного решения. Детальность решения типовых задач и наличие дополнительных задач позволяют студенту самостоятельно изучить материал, что делает пособие пригодным для организации дистанционной технологии обучения.

Следует отметить, что большинство учебников по дисциплине либо вообще не содержат прикладных примеров, либо эти примеры не рассматриваются подробно, а для эффективного использования на практике полученных знаний студент должен уметь самостоятельно строить математические модели для конкретных прикладных задач. Поэтому в пособии приведены различные примеры, позволяющие студенту глубже понять смысл изучаемой дисциплины.

ВВЕДЕНИЕ

С незапамятных времен человечество с помощью метода проб и ошибок, интуиции и опыта, накапливаемого в каждой конкретной ситуации, создавало искусство выработки наилучших решений в самых разных областях своей деятельности.

Успехи использования математических методов и стиля мышления в естественных науках привели к мысли о необходимости включения в сферу математического влияния и проблемы принятия решений с целью превращения древнего искусства в современную науку. Испытанный метод проб и ошибок в наши дни часто теряет свою универсальность: слишком катастрофическими могут оказаться ошибки и слишком мало времени отпущено для проб. Становится все более ясным, что сегодня меньше чем когда-либо ранее допустимы произвольные, чисто волевые решения. На первый план выходит не задача создания все новых и новых образцов техники, а проблема организации и управления, причем управления не только (и не столько) машинами, но и людьми, сложными человеко-машинными системами.

Математический аппарат, столь давно используемый в сфере точных и опытных наук, только сравнительно недавно стал применяться в гуманитарных науках. Предмет гуманитарных наук неизмеримо сложнее тех, которыми занимаются точные науки. Многие гуманитарные явления гораздо труднее поддаются формализации, если вообще поддаются. Для каждого из этих явлений гораздо шире спектр причин, от которых оно зависит, и в их числе — психология живых людей и коллективов, людские пристрастия и антагонизмы.

И все же, кроме традиционных областей приложения, математика начинает заниматься такими вопросами, как конфликтные ситуации, иерархические отношения в коллективах, согласие, авторитет, общественное мнение. При этом строятся и анализируются математические модели, применяются математические методы. Математика не только проникает в ранее чуждые для нее области, но и трансформируется, становится менее «формальной», меняет свои методологические черты, гибко приближаясь к гуманитарным наукам. Ее методы в области гуманитарных и смежных с ними наук могут служить мощным вспомогательным средством, позволяющим исследователю более глубоко проникнуть в существо явления, проследить его закономерности, обнаружить скрытые связи, малодоступные наблюдению простым, невооруженным глазом.

Математические методы можно рассматривать как достаточно эффективное средство структурированного, более компактного и обозримого представления имеющейся информации. Это особенно ясно в тех случаях, когда информация задается в виде числовых массивов, в графической форме и др. Анализ результатов математической обработки данных часто позволяет высказать некоторые рекомендации относительно тех или иных способов действия.

Кроме того, существует целый ряд типичных управленческих ситуаций, допускающих известную формализацию, где именно математические подходы и соображения обоснованно становятся решающими. Этому способствуют заметные продвижения в разработке математических подходов к решению количественных проблем управления, заметные теоретические продвижения в теории игр и теории полезности (Дж. фон Нейман) и линейном программировании (Дж. Данциг, Л. В. Канторович), а также создание новых мощных средств вычислений. Многие задачи управления удалось достаточно хорошо формализовать, и сейчас они уже весьма широко и довольно успешно решаются стандартными математическими методами.

В первой половине XX в. начали разрабатывать (и довольно успешно) элементы научного подхода к поиску решений задач управления, а модели, хорошо показавшие себя при проведении естественно-научных и инженерно-технических изысканий, стали приспосабливать к решению управленческих задач.

Под моделированием понимается исследование явлений, процессов и объектов путем построения и изучения их моделей. Модель — это такой материальный или мысленно представляемый объект, который в процессе познания замещает объект-оригинал, сохраняя некоторые важные для данного исследования типичные его черты. Модель должна отражать важнейшие черты явления, т. е. в ней должны быть учтены все факторы, от которых в наибольшей степени зависит успех исследования. Вместе с тем модель должна быть по возможности простой — без мелких, второстепенных факторов, так как их учет усложняет математический анализ и делает трудно обозримым результат изысканий.

Различают несколько приемов моделирования, которые можно условно объединить в две большие группы: материальное (предметное) и идеальное моделирование.

Материальное моделирование основано на исследовании модели, воспроизводящей основные геометрические, физические, динамические и функциональные характеристики изучаемого объекта. Разновидностями материального моделирования являются физическое моделирование и аналоговое моделирование.

Физическим принято называть моделирование, при котором реальному объекту противопоставляется его увеличенная или уменьшенная копия, допускающая исследование с помощью последующего перенесения свойств изучаемых процессов и явлений с модели на объект на основе теории подобия.

Аналоговое моделирование основано на аналогии процессов и явлений, имеющих различную физическую природу, но одинаково описываемых формально (одними и теми же математическими уравнениями, логическими схемами и т. п.).

Идеальное моделирование основано не на материальной аналогии объекта и модели, а на аналогии идеальной, мыслимой.

Идеальное моделирование разделяют на два типа: интуитивное и знаковое.

Под интуитивным понимается моделирование, основанное на интуитивном представлении об объекте исследования, не поддающемся формализации либо не нуждающемся в ней. В этом смысле, например, жизненный опыт каждого человека может считаться его интуитивной моделью окружающего мира.

Знаковым называется моделирование, использующее в качестве моделей знаковые преобразования какого-либо вида (схемы, графики, чертежи, формулы, наборы символов и т. д.), а также включающее совокупность законов, по которым можно оперировать с выбранными знаковыми образованиями и их элементами.

Важнейшим видом знакового моделирования является математическое моделирование, при котором исследование объекта осуществляется посредством модели, сформированной на языке математики, и использованием тех или иных математических методов.

При построении математической модели реальное явление всегда упрощается, схематизируется, и полученная схема описывается с помощью того или иного математического аппарата.

Использование математических методов для решения социально-экономических задач требует предварительного качественного анализа исследуемой системы, глубокого изучения ее сущности, выяснения направления целесообразного ее изменения. На основе данных такого анализа социально-экономическая система описывается некоторыми математическими соотношениями с помощью независимых переменных, параметров и функций, отображающих интересующие нас свойства объекта или процесса, т. е. создается математическая модель. Эта модель, как правило, представляет собой математическое выражение в виде системы уравнений, неравенств, отображающих взаимосвязь между явлениями в каком-нибудь реальном экономическом процессе, либо функцию, которая описывает некоторую выявленную закономерность, содержащуюся в заданной исходной количественной информации об объекте.

Существует много разнообразных математических подходов, которые достаточно хорошо описывают различные ситуации, требующие принятия тех или иных управленческих решений. Из них выделим три класса: детерминированные, стохастические и игровые.

При использовании детерминированных методов исходят из того, что основные факторы, характеризующие ситуацию, вполне определены и известны. Обычно ставится задача оптимизации некоторой величины (например, максимизация прибыли).

Стохастические методы применяются в тех случаях, когда некоторые факторы носят неопределенный, случайный характер.

Наконец, при учете наличия союзников либо противников, имеющих собственные интересы, необходимо применять теоретико-игровые методы.

Отметим, что сам термин «управление» можно понимать по-разному. Это и организация, в том числе и технологическая, той или иной осмысленной деятельности для достижения каких-либо целей (в качестве математического обеспечения используются преимущественно детерминированные и стохастические методы), это и изучение моделей поведения взаимодействующих сторон (применяются игровые методы).

Часть 1. ДЕТЕРМИНИРОВАННЫЕ МЕТОДЫ

Глава 1. Методы линейного программирования

§ 1. Формулировка классической задачи линейного программирования

Если все ограничения на переменные в математической модели представляют собой линейные алгебраические неравенства (равенства), а целевая функция является линейным функционалом относительно переменных модели, то такая математическая модель относится по классификации к теории линейного программирования.

В задачах линейного программирования (ЛП) требуется найти экстремум (максимум или минимум в зависимости от содержательной постановки задачи) линейной функции

$$F(X) = \sum_{j=1}^{n} c_j x_j.$$

Эта функция, называемая целевой, неограниченна, и поэтому искать ее максимум или минимум, не налагая никаких ограничений на область изменения вектора X с компонентами $(x_1, x_2, \ldots, x_n)$, бессмысленно. Интерес представляет задача максимизации этой функции при условии, что вектор X принадлежит некоторому множеству, причем выбор значений вектора X определяется условиями типа линейных неравенств.

В общем случае ***задача линейного программирования формулируется*** следующим образом: найти величину вектора X, доставляющего максимум линейной функции

$$F(X) = \sum_{j=1}^{n} c_j x_j$$

на множестве значений $(x_1, x_2, \ldots, x_n)$, которые удовлетворяют ограничениям вида

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1,\\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m,\\
&x_1 \ge 0;\ x_2 \ge 0;\ \ldots, x_n \ge 0.
\end{aligned}$$

В более компактной форме это можно записать в виде

$$\sum_{j=1}^{n} c_j x_j \to \max$$

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i,\ i = 1, 2, \ldots, m,$$
$$x_j \ge 0,\ j = 1, 2, \ldots, n.$$

Вектор $b^T = (b_1, b_2, ..., b_m)$ называется вектором ограничений для задачи линейного программирования. Вектор $X^T = (x_1, x_2, ..., x_n)$, удовлетворяющий линейным ограничениям, называется допустимым решением задачи линейного программирования. Допустимый вектор X, доставляющий максимум целевой функции, называется оптимальным решением задачи линейного программирования. Коэффициенты a_{ij}, c_j, b_i составляют условия задачи и трактуются исходя из содержательной постановки проблемы.

§ 2. Задача о выпуске продукции

Рассмотрим некоторую производственную систему, в которой m ресурсов S_1, S_2, …, S_m участвуют в производственном процессе (сырье, топливо, материалы, инструменты и т. д.). Продукты производятся или потребляются с помощью линейных или технологических процессов. Линейная модель производства, в которой участвуют m ресурсов, состоит из набора технологических процессов P_1, P_2, …, P_n. Она полностью описывается числами a_{ij}, где a_{ij} — количество ресурса S_i, потребляемого при $a_{ij} > 0$ или производимого при $a_{ij} < 0$ при единичной интенсивности использования процесса P_j. Под интенсивностью процесса P_j понимается такое число x_j, при котором потреблено $a_{ij}\ x_j$ единиц ресурса S_i. Через c_j обозначим доход от реализации единицы продукции, получаемой при использовании процесса P_j с единичной интенсивностью.

В более простой интерпретации в качестве P_1, P_2, …, P_n можно рассматривать список производимой продукции. Запасы ресурсов каждого вида ограничены и равны b_1, b_2, …, b_m.

Требуется определить, в каком количестве выпускать продукцию каждого вида, чтобы доход от реализации этой продукции был максимален с учетом ограничений на ресурсы. Другими словами, нужно определить производственный план. Общий доход от реализации выпущенной продукции будет равен

$$F(X) = \sum_{j=1}^{n} c_j x_j.$$

Ограничениями в этой задаче являются требования, согласно которым расход ресурса вида S_i на выпуск всех видов продукции не превышал имеющегося запаса

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \; i = 1, 2, \ldots, m.$$

Пример 1. Заводу требуется составить оптимальный производственный план выпуска двух видов изделий при определенных возможностях трех типов машин. План выпуска должен быть таким, чтобы от реализации выпущенной по этому плану продукции завод получил бы наибольшую прибыль. Эти виды изделий последовательно обрабатываются данными машинами (табл. 1). Завод от реализации одного изделия первого вида получает 4 тыс. руб., а от реализации одного изделия второго вида — 6 тыс. руб. прибыли.

Таблица 1

Вид машин	Вид изделия		Время работы машин
	1	2	
1	0,5	1,0	12,0
2	2,0	1,0	30,0
3	0,0	1,0	9,0

Построим математическую модель этой задачи.

Пусть x_1 — число изделий первого вида, а x_2 — число изделий второго вида.

$$F(x_1, x_2) = 4x_1 + 6x_2 \to \max$$

$$\begin{cases} 0,5x_1 + \ x_2 \le 12, \\ 2x_1 + \ x_2 \le 30, \\ x_2 \le 9, \\ x_1, x_2 \ge 0. \end{cases}$$

Вместе с исходной (прямой) задачей рассматривается двойственная к ней задача, которую можно сформулировать следующим образом.

Предположим, что организация решила закупить ресурсы фирмы S_1, S_2,…, S_m. Необходимо определить цены y_1, y_2, …, y_m на эти ресурсы. Естественно, что покупающая организация заинтересована в том, чтобы затраты на покупку ресурсов в объеме b_1, b_2, …, b_m по ценам y_1, y_2,…, y_m были минимальны, т. е.

$$Z(Y) = b_1 y_1 + b_2 y_2 + \ldots + b_m y_m \to \min.$$

Предприятие, продающее ресурсы, заинтересовано в том, чтобы полученная выручка была не меньше той суммы, которую предприятие может получить при переработке ресурсов в готовую продукцию. На изготовление единицы продукции j-го вида сырье S_1 расходуется в объеме a_{1j}, сырье S_2 — в объеме a_{2j}…, сырье S_m — в объеме a_{mj} по цене $y_1, y_2,…, y_m$ соответственно, т. е. затраты на изготовление продукции P_j должны быть не меньше чем цена ее реализации. Приходим к ограничениям следующего вида:

$$a_{1j}y_1 + a_{2j}y_2 + \ldots + a_{mj}y_m \geq c_j,\ j=1, 2, \ldots, n.$$

Учитывая условия неотрицательности цены единицы i-го ресурса, приходим к следующей задаче.

Найдем такой вектор $Y = (y_1, y_2, \ldots, y_m)$ — цен ресурсов, удовлетворяющий системе неравенств

$$\begin{cases} a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_m \geq c_1, \\ a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_m \geq c_2, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ a_{1j}y_1 + a_{2j}y_2 + \ldots + a_{mj}y_m \geq c_j, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m \geq c_n, \\ y_i \geq 0,\ i = 1, 2, \ldots, m, \end{cases}$$

при котором функция $Z(Y) = b_1y_1 + b_2y_2 + \ldots + b_my_m$ принимает минимальное значение.

Цены ресурсов имеют различные названия: учетные, неявные, теневые, объективно-обусловленные оценки (о-о оценки). Смысл данных названий состоит в том, что это условные (ненастоящие) цены. Они определяются в ходе решения задачи и могут называться ценностью ресурсов.

Пример 2. Построим двойственную задачу к прямой задаче из примера 1.

Пусть y_1 — ценность первого ресурса, y_2 — ценность второго ресурса, y_3 — ценность третьего ресурса. Тогда

$$Z(y_1, y_2, y_3) = 12y_1 + 30y_2 + 9y_3 \to \min$$
$$\begin{cases} 0{,}5y_1 + 2y_2 \geq 4, \\ \ y_1 + \ y_2 + y_3 \geq 6, \\ \ y_1, y_2, y_3 \geq 0. \end{cases}$$

Сопоставим общие представления прямой и двойственной задач (табл. 2). При этом прямая задача — это задача модели распределения ограниченных ресурсов.

Таблица 2

Прямая задача	Двойственная задача
Максимизировать $F=\sum_{j=1}^{n} c_j x_j$ при ограничениях $\sum_{j=1}^{n} a_{ij} x_j \le b_i$; $i = 1, 2, \ldots, m$ $x_j \ge 0$, $j = 1, 2, \ldots, n$	Минимизировать $Z=\sum_{i=1}^{m} b_i y_i$ при ограничениях $\sum_{i=1}^{m} a_{ij} y_i \ge c_j$; $j = 1, 2, \ldots, n$ $y_i \ge 0$, $i = 1, 2, \ldots, m$

При сравнении примеров видно, что двойственная задача по отношению к исходной составляется согласно следующим *правилам*:

1. Если первая задача имеет размеры $m \times n$ (m–ограничений с n неизвестными), то вторая — размеры $n \times m$.
2. Матрицы из коэффициентов при неизвестных в левых частях ограничений обеих задач являются взаимно транспонированными.
3. В правых частях ограничений в каждой задаче стоят коэффициенты при неизвестных в целевой функции другой задачи.
4. В прямой задаче все ограничения представляют собой неравенства типа «≤», причем в этой задаче требуется достичь max F. Напротив, в двойственной задаче все ограничения суть неравенства типа «≥», причем требуется достичь min Z.

Если прямая и двойственная задачи удовлетворяют четвертому требованию, то говорят, что задача ЛП записана в симметричной квадратичной форме. Если в ограничениях прямой задачи часть ограничений имеет знак « = », то это общая формулировка и двойственная задача записывается по другим правилам.

Если в задаче линейного программирования число переменных равно двум, то ее можно решить графическим методом. Суть метода состоит в следующем: решением системы ограничений задачи является совокупность точек плоскости, координаты которых удовлетворяют каждому из неравенств системы. Эта совокупность может образовывать некоторую фигуру на плоскости, в качестве которой может быть некоторый многоугольник, либо некоторая неограниченная область, либо вообще пустое множество. Она называется допустимой областью. На этой области мы должны найти экстремум целевой функции. Известно, что этот экстремум достигается на границе области. Поиск этой граничной точки и составляет суть метода. Поясним это на примере.

Пример 3. Решим *графическим методом* задачу ЛП из примера 1.

$$F(x_1, x_2) = 4x_1 + 6x_2 \to \max$$

$$\begin{cases} 0,5x_1 + \ x_2 \le 12, \\ \ 2x_1 + \ x_2 \le 30, \\ \qquad\quad x_2 \le 9, \\ \qquad x_1, x_2 \ge 0. \end{cases}$$

Решением линейного алгебраического неравенства относительно двух переменных является совокупность точек плоскости, координаты которых удовлетворяют этому неравенству, графически — это полуплоскость. Решением системы таких неравенств является их пересечение. Для нахождения полуплоскости необходимо знак неравенства заменить на знак равенства, построить соответствующую этому уравнению прямую на плоскости и выбрать из двух образовавшихся полуплоскостей нужную.

В прямоугольной декартовой системе координат (рис. 1) построим прямую $0,5x_1 + x_2 = 12$, соответствующую ограничению (1) по двум точкам, например, (24; 0) и (0; 12). Находим, какая из полуплоскостей, на которые эта прямая делит всю координатную плоскость, является областью решений неравенства (1). Для этого достаточно координаты какой-либо точки, не лежащей на прямой, подставить в неравенство. Так как прямая не проходит через начало координат, подставим координаты точки *O* (0; 0) в первое ограничение $0,5 \cdot 0 + 1 \cdot 0 \le 12$. Получим строгое неравенство $0 \le 12$. Следовательно, точка *O* лежит в полуплоскости решений. Штриховкой отмечаем полуплоскость, содержащую точку *O*. Аналогично построим прямую $2x_1 + x_2 = 30$ по двум точкам (0; 30) и (15; 0) и определим область решений ограничения (2). Третья прямая будет параллельна оси Ox_1.

Построим на плоскости *область допустимых решений*. Находим общую часть полуплоскостей решений, учитывая при этом условия неотрицательности переменных.

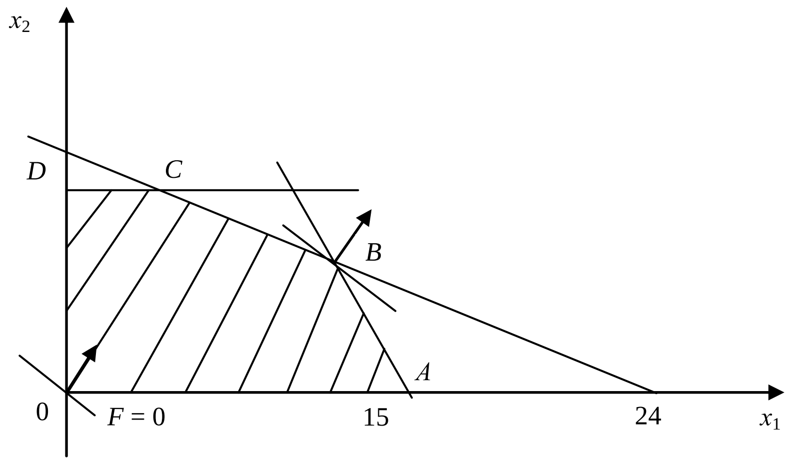


Рис. 1. Графическое решение задачи

Множество допустимых решений представляет собой многоугольник *OABCD* (см. рис. 1). Среди точек многоугольника *OABCD* найдем оптимальную.

Построим нормаль линий уровня $\vec{n}\,(4,\ 6)$. Так как решается задача на отыскание максимума целевой функции, то линию уровня $4x_1 + 6x_2 = C$ перемещаем в направлении нормали до угловой точки *B*, которая расположена на прямой, называемой опорной. Эта опорная прямая проходит через точку *B* пересечения прямых, ограничивающих область допустимых решений и соответствующих неравенствам (1) и (2), перпендикулярно нормали.

В этой точке пересекаются две прямые:

$$0,5x_1 + x_2 = 12$$
$$2x_1 + x_2 = 30.$$

Найдем координаты точки B:

$$\begin{cases} x_2 = 30 - 2x_1, \\ 0,5x_1 + 30 - 2x_1 = 12. \end{cases} \Rightarrow \quad -1,5x_1 = -18 \quad \Rightarrow \quad x_1 = 12, \quad x_2 = 6.$$

Полученное решение будет оптимальным производственным планом, дающим заводу максимальную прибыль, а именно: $F_{\max}(X^*) = 4 \cdot 12 + 6 \cdot 6 = 84.$

Двойственная задача из примера 2 уже имеет три неизвестных величины. Поэтому такую задачу ЛП мы уже не можем решить графическим методом. В этом случае, имея решение прямой задачи, мы можем воспользоваться следующей теоремой.

Теорема равновесия. Пусть X^* — какое-нибудь допустимое решение прямой задачи, а Y^* — какое-нибудь допустимое решение двойственной задачи. Для одновременной оптимальности этих решений необходимо и достаточно выполнение равенств:

$$x_j^* \left(\sum_{i=1}^{m} a_{ij} y_i^* - c_j\right) = 0, \qquad j = 1, 2, \ldots, n,$$
$$y_i^* \left(\sum_{j=1}^{n} a_{ij} x_j^* - b_i\right) = 0, \qquad i = 1, 2, \ldots, m.$$

Применим теорему равновесия для решения задачи из примера 2.

Пример 4. Запишем прямую и двойственную задачи (табл. 3).

Таблица 3

Прямая задача	Двойственная задача
$F(x_1, x_2) = 4x_1 + 6x_2 \to \max$ $\begin{cases} 0,5x_1 + x_2 \le 12 \\ 2x_1 + x_2 \le 30 \\ x_2 \le 9 \\ x_1, x_2 \ge 0 \end{cases}$	$Z(y_1, y_2, y_3) = 12y_1 + 30y_2 + 9y_3 \to \min$ $\begin{cases} 0,5y_1 + 2y_2 \ge 4 \\ y_1 + y_2 + y_3 \ge 6 \\ y_1, y_2, y_3 \ge 0 \end{cases}$

Тогда теорему равновесия запишем в виде

$$\begin{cases} x_1^* \ (0,5y_1^* + 2y_2^* - 4) = 0, \\ x_2^* \ (y_1^* + y_2^* + y_3^* - 6) = 0, \\ y_1^* \ (0,5x_1^* + x_2^* - 12) = 0, \\ y_2^* \ (2x_1^* + x_2^* - 30) = 0, \\ y_3^* \ (x_2^* - 9) = 0. \end{cases}$$

Третье ограничение прямой задачи на оптимальном плане выполняется как строгое неравенство. Следовательно, соответствующая переменная двойственной задачи равна нулю: $y_3^* = 0$.

Первое и второе ограничения прямой задачи обращаются в точные равенства. Следовательно, соответствующие переменные двойственной задачи положительны: $y_1^* > 0$ и $y_2^* > 0$.

Условия равновесия можно записать в виде равенств

$$\begin{cases} 12\ (0,5y_1^* + 2y_2^* - 4) = 0, \\ 6\ (y_1^* + y_2^* + y_3^* - 6) = 0. \end{cases} \Rightarrow \begin{cases} 0,5y_1^* + 2y_2^* - 4 = 0, \\ y_1^* + y_2^* + y_3^* - 6 = 0. \end{cases}$$

Подставляя $y_3^* = 0$ в последнюю систему уравнений, получим систему:

$$\begin{cases} 0,5y_1^* + 2y_2^* = 4, \\ y_1^* + y_2^* = 6. \end{cases}$$

Откуда получим, что $y_1^* = \frac{16}{3}$, $y_2^* = \frac{2}{3}$.

Оптимальное решение двойственной задачи $Y^* = (\frac{16}{3};\ \frac{2}{3};\ 0)$. При этом минимум целевой функции двойственной задачи $Z_{min} = Z\ (Y^*) = 12\ \frac{16}{3} + 30\ \frac{2}{3} = 84$.

Существует взаимосвязь между значениями целевых функций прямой и двойственной задач на допустимых решениях. Она определяется следующими теоремами теории двойственности.

Теорема (основное неравенство). Пусть X — какое-нибудь допустимое решение прямой задачи, Y — какое-нибудь допустимое решение двойственной задачи. Тогда справедливо неравенство

$$F(X) \le Z(Y).$$

Следствие 1 (***достаточный признак оптимальности***). Если для каких-то допустимых решений X^* и Y^*, соответственно, прямой и двойственной задач выполняется равенство

$$F(X^*) = Z(Y^*),$$

то X^* — оптимальное решение прямой задачи, Y^* — оптимальное решение двойственной задачи.

Следствие 2. Если в одной из пары двойственных задач целевая функция не ограничена с соответствующей стороны (т. е. $\max F = +\infty$ — в прямой задаче или $\min Z = -\infty$ — в двойственной задаче), то другая задача не имеет допустимых решений.

Основная теорема. Если разрешима одна из пары двойственных задач, то разрешима и другая задача, причем

$$\max F = \min Z.$$

Действительно, для разобранных задач получили соотношение

$$F_{\max}(X^*) = F(12{,}6) = 84 = Z_{\min}(Y^*) = Z(\frac{16}{3}, \frac{2}{3}, 0).$$

§ 3. Задача о выборе рациона питания

Пусть имеется n типов питательных веществ, содержащихся в m продуктах. Из продуктов нужно составить диету (определить количество каждого продукта, которое нужно потреблять ежегодно каждому индивиду), которая содержит определенное количество питательных веществ. Потребность человека в год в j-м питательном веществе равна c_j. Стоимость единицы i-го продукта равна b_i. Обозначим через a_{ij} количество j-го питательного вещества, содержащегося в единице i-го продукта. Обозначив через y_i количество i-го продукта в годовом рационе, получим следующую задачу линейного программирования об отыскании наиболее экономной диеты:

$$\nu \quad y \quad = \sum_{i=1}^{m} b_i y_i \to \min$$

$$\sum_{i=1}^{m} a_{ij} y_i \ge c_j, \; j = 1, 2, \ldots, n,$$

$$y_i \ge 0, \; i = 1, 2, \ldots, m.$$

Пример 5. Рассмотрим простейший вариант задачи о диете. Изучается вопрос о рационе кормления некоего живого существа. Пусть на данном этапе развития этого существа наиболее важными для него являются три питательных вещества (например, витамины A, B и C).

В нашем распоряжении находится два продукта, в которых содержатся эти витамины. Содержание витаминов в единице продуктов и суточная потребность в витаминах заданы в табл. 4.

Таблица 4

Витамин	Продукт		Суточная потребность
	1	2	
A	5	2	10
B	3	4	12
C	1	5	5

Стоимость единицы продуктов, соответственно, равна 13 и 8 денежным единицам. Обозначим через y_1 количество первого продукта, включаемого в ежедневный рацион, а через y_2 — количество второго продукта. Они являются переменными в оптимизационной задаче.

Построим целевую функцию, которая заключается в определении стоимости диеты. Стоимость диеты задается выражением $13y_1 + 8y_2$.

Математическая формулировка задачи о диете составляется следующим образом: определить вектор $y = (y_1, y_2)$ такой, что

$$
\begin{aligned}
&v \;\; y \;\; = 13y_1 + 8y_2 \to \min\\
&5y_1 + 2y_2 \ge 10,\\
&3y_1 + 4y_2 \ge 12,\\
&y_1 + 5y_2 \ge 5,\\
&y_1 \ge 0,\ y_2 \ge 0.
\end{aligned}
$$

§ 4. Классическая транспортная задача

Рассмотрим простейшую задачу, возникающую при моделировании транспортных потоков. Для простоты рассматривается вопрос об организации перевозок лишь какого-то одного продукта от пунктов, в которых он производится (или складируется), к потребителям, расположенным в других пунктах.

Пусть имеется m пунктов производства и n пунктов потребления. Известно, что в i-м пункте производства имеется a_i единиц рассматриваемого продукта, а объем его потребления в j-м пункте потребления составляет b_j единиц. Для простоты предположим, что затраты на перевозку продукта из i-го пункта производства в j-й пункт потребления пропорциональны объему поставки, а значит, легко вычисляются, если заданы величины c_{ij}, указывающие стоимость перевозки единицы продукта. Эти величины известны.

Требуется спланировать перевозки так, чтобы суммарные затраты на реализацию предлагаемого плана были как можно меньше. Ясно, что задачу имеет смысл рассматривать лишь в предположении, что суммарный объем производства не меньше суммарного объема потребления, т. е. $\sum_{i=1}^{m} a_i \geq \sum_{j=1}^{n} b_j$. Обозначим через y_{ij} объем планируемой поставки продукта из i-го пункта производства в j-й пункт потребления. Тогда суммарный объем вывоза из i-го пункта производства будет равен $\sum_{j=1}^{n} y_{ij}$, а суммарный объем поставки в j-й пункт потребления составит $\sum_{i=1}^{m} y_{ij}$. Затраты на реализацию всех перевозок будут задаваться формулой $\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} y_{ij}$.

В результате получим следующую классическую транспортную задачу:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} y_{ij} \to \min$$

$$\begin{cases} \sum_{j=1}^{n} y_{ij} \leq a_i, \ i = 1, 2, \ldots, m, \\ \sum_{i=1}^{m} y_{ij} \geq b_j, \ j = 1, 2, \ldots, n, \\ y_{ij} \geq 0, \ i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n. \end{cases}$$

Первое неравенство означает, что количество продукта, вывезенного со склада, не должно превышать запаса, а второе — потребности каждого пункта должны быть удовлетворены. Несмотря на кажущееся полное внешнее отличие получаемой математической задачи от той, которую мы получили при рассмотрении рационов кормления, на самом деле это задачи одного класса: требуется минимизировать некоторую линейную функцию при линейных ограничениях на переменные.

Если выполнено условие $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$, то это гарантирует, что все заявки будут выполнены полностью и из каждого пункта отправления будет вывезен весь груз. В этом случае транспортная задача называется *закрытой* (*сбалансированной*). Это условие обеспечивает существование оптимального решения задачи. В противном случае задача называется *открытой*.

Математическая формулировка закрытой транспортной задачи будет иметь вид

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij} y_{ij} \quad \to \quad \min$$

$$\begin{cases} \sum_{j=1}^{n} y_{ij} = a_i, \quad i = \overline{1, m}, \\ \sum_{i=1}^{m} y_{ij} = b_j, \quad j = \overline{1, n}, \\ y_{ij} \ge 0 \quad \text{для всех} \ i, j. \end{cases}$$

Пример 6. Имеются два склада A_1 и A_2 и три магазина B_1, B_2 и B_3.

Наличие товара на складах и запросы магазинов, а также стоимость доставки указаны в следующей таблице:

a_i	b_j		
	50	20	30
50	10	15	25
60	20	30	30

Составим экономико-математическую модель транспортной задачи.

Проверим условие сбалансированности: 50 + 60 ≠ 50 + 20 + 30, т. е. задача будет открытой. Исходя из этого транспортная задача имеет вид

$$10y_{11} + 15y_{12} + 25y_{13} + 20y_{21} + 30y_{22} + 30y_{23} \to \min$$

$$\begin{cases} y_{11} + y_{12} + y_{13} \le 50, \\ y_{21} + y_{22} + y_{23} \le 60, \\ y_{11} + y_{21} \ge 50, \\ y_{12} + y_{22} \ge 20, \\ y_{13} + y_{23} \ge 30, \\ y_{ij} \ge 0, i = 1, 2; j = 1, 2, 3. \end{cases}$$

Вектор неизвестных транспортной задачи имеет вид

$$Y^T = (y_{11}, y_{12}, y_{13}, y_{21}, y_{22}, y_{23}).$$

Пример 7. Представим два склада A_1 и A_2 и два магазина B_1 и B_2.

Наличие товара на складах и запросы магазинов, а также стоимость доставки указаны в следующей таблице:

a_i	b_j	
	20	20
25	3	4
15	5	2

Составим экономико-математическую модель транспортной задачи.

Проверим условие сбалансированности: 20 + 20 = 25 + 15 = 40, т. е. задача будет закрытой. Исходя из этого транспортная задача имеет вид

$$3y_{11} + 4y_{12} + 5y_{21} + 2y_{22} \to \min$$

$$\begin{cases} y_{11} + y_{12} = 25, \\ y_{21} + y_{22} = 15, \\ y_{11} + y_{21} = 20, \\ y_{12} + y_{22} = 20, \\ y_{ij} \geq 0,\ i = 1, 2;\ j = 1, 2. \end{cases}$$

Вектор неизвестных транспортной задачи имеет вид

$$Y^T = (y_{11}, y_{12}, y_{21}, y_{22}).$$

Из приведенных примеров видно, что вектор неизвестных имеет размерность больше чем 2, и такие задачи мы не можем решать графическим методом. Для решения закрытых транспортных задач существуют специальные методы (например, метод потенциалов), которые учитывают специфику транспортной задачи. Изучение этих методов не входит в содержание пособия.

Для решения задач ЛП произвольной размерности рекомендуем воспользоваться компьютером и надстройкой «Поиск решения» для Microsoft Office Excel. На страницах Microsoft Office Online доступны сведения об установке этой надстройки и примеры использования.

§ 5. Задача о контракте

Предположим, что предприниматель обдумывает вопрос о заключении контракта на определенный заказ. Для выполнения всего объема работ нужно привлечь специалистов разного профиля. Допустим, нужно n различных специалистов. Пронумеруем их и будем отождествлять каждого специалиста с его номером.

Таким образом, $\{1, 2, \ldots, n\}$ — множество всех требуемых специалистов. Известно, что специалисты могут объединяться в различные творческие коллективы — коалиции. Каждая такая коалиция $S \subset \{1, 2, \ldots, n\}$ может сама организовывать свое дело и получать при этом доход $V(S)$. Будем для простоты считать, что любая коалиция допустима, но если коалиция S «плохая», то $V(S) = 0$.

Ясно, что если наш предприниматель предложит участникам коалиции S такие оклады, что в сумме они получат меньше, чем $V(S)$, то они не будут участвовать в деле предпринимателя, а организуют свое дело. Предпринимателю необходимо знать, какая минимальная сумма контракта позволит заинтересовать требуемых специалистов.

Обозначим через x_j величину оклада, который предприниматель намерен предложить j-му специалисту. Тогда задача о контракте примет вид

$$\sum_{j=1}^{n} x_j \to \min$$

$$\sum_{j \in S} x_j \ge V(S).$$

$$\text{Для всех } S \subset \{1, 2, \ldots, n\}.$$

Итак, в главе 1 мы познакомились с примерами постановок задач из разных предметных областей. Решение этих задач сводится к математической модели линейного программирования.

Контрольные вопросы и задания

1. Какие задачи линейного программирования решаются графическим методом?
2. Что называется областью допустимых решений?
3. Как найти область допустимых решений?
4. С какой целью определяется градиент целевой функции?
5. Как определяется положение линии уровня?
6. Как составляются целевая функция, ограничения и условия для двойственных задач?
7. Сформулируйте теорему равновесия.
8. Сформулируйте задачу о диете.
9. Сформулируйте классическую транспортную задачу.
10. В чем заключается отличие между открытой и закрытой транспортной задачами?

Задания для самостоятельного решения

1. С Курского вокзала Москвы ежедневно отправляются скорые и пассажирские поезда.

Постройте модель задачи, с помощью которой можно определить оптимальное количество поездов двух типов, обеспечивающих максимальное количество ежедневно отправляемых пассажиров с вокзала. Пассажировместимость и количество вагонов железнодорожного депо станции отправления указаны в следующей таблице:

Тип вагона	Количество вагонов		Пассажировместительность, человек	Парк вагонов
	скорый	пассажирский		
Багажный	1	1	—	10
Почтовый	1	0	—	8
Плацкартный	8	5	58	80
Купейный	4	6	36	70
Мягкий	1	3	18	30

2. Правление банка сочло возможным инвестировать капитал суммой 300 тыс. долл. в 6 конкретных проектов. Эксперты оценили годовую эффективность каждого проекта на 2 года следующим образом:

Год	Номер проекта					
	1	2	3	4	5	6
1	0,12	0,14	0,15	0,10	0,18	0,25
2	0,10	0,10	0,12	0,18	0,12	0,15

Менеджер по инвестициям считает, что не стоит вкладывать в проект 5 более 40 тыс. долл., а в проекты 4 и 6 — более 25 % от общей денежной суммы ввиду высокой вероятности риска, соответствующего этим проектам. В то же время не менее 40 % денежных средств желательно поместить в проекты 1 и 2.

Найдите план инвестиций в каждый проект с целью максимизации дохода.

3. Фирма производит и продает столы и шкафы из древесины хвойных и лиственных пород.

Расход каждого вида в кубических метрах на каждое изделие задан в следующей таблице:

Изделие	Расход древесины, м^3		Цена изделия, тыс. руб.
	Хвойные	Лиственные	
Стол	0,15	0,2	0,8
Шкаф	0,3	0,1	1,5
Запасы древесины, м^3	75	40	—

Определите оптимальное количество столов и шкафов, которое следует поставлять на продажу для получения максимального дохода фирмы.

4. Кондитерская фабрика для производства карамели использует три вида основного сырья: сахарный песок, патоку и фруктовое пюре.

Нормы расходов сырья каждого вида на производство 1 кг карамели каждого вида приведены в следующей таблице:

Сырье	Норма расхода, кг/кг карамели		
	«Сахарная»	«Детская»	«Фруктовая»
Сахар-песок	0,8	0,6	0,5
Патока	0,2	0,3	0,4
Фруктовое пюре	0,0	0,1	0,1

Фабрика в сутки получает 800 кг сахара-песка, 600 кг патоки и 120 кг фруктового пюре. Изучение покупательского спроса показало, что суточное производство «Детской» карамели не должно превышать 300 кг. Доход от реализации 1 кг карамели составляет 27 руб. для «Сахарной» карамели, 32 руб. — для «Детской», 28 руб. — для «Фруктовой» карамели.

Составьте задачу линейного программирования о производстве с целью получения максимального дохода.

5. Предприятие может выпускать три вида сувенирной продукции, сбыт любого количества которой обеспечен. Для изготовления используются трудовые ресурсы, полуфабрикаты и станочное оборудование.

Расход каждого ресурса на единицу выпускаемой продукции приведен в следующей таблице:

Ресурс	Норма расхода на единицу продукции		
	«Сибирь»	«Наш край»	«Новосибирск»
Затраты труда, ч	4	2	2
Сырье, кг	2	10	6
Станочное оборудование, ч	1	0	2

В течение 1 недели предприятие может задействовать 4 800 ч труда работников, 2 400 кг сырья и 1 500 ч работы станков. Доход от реализации сувенира составляет 65, 70 и 60 руб. соответственно.

Составьте задачу линейного программирования о производстве с целью получения максимального дохода.

6. Фирма специализируется на производстве шкафов трех моделей.

Затраты труда на различных стадиях производства отражены в следующей таблице:

Участок	Затраты труда на единицу продукции, ч		
	«Классический»	«Новый»	«Антикварный»
Лесопильный	1	2	4
Сборочный	2	4	2
Отделочный	1	1	2

В течение 1 недели можно планировать работу на лесопилке на 360 чел.-ч, в сборочном цехе — на 520, в отделочном цехе — на 220 чел.-ч. Доход от продажи шкафа каждой из моделей составляет 9 000, 11 000 и 15 000 руб. соответственно.

Составьте задачу линейного программирования о производстве с целью получения максимального дохода.

7. Караван Марко Поло использует для перевозки сухого инжира из Багдада в Мекку дромадеров (одногорбых верблюдов) и обычных (двугорбых) верблюдов. Верблюд может нести 1 000 фунтов груза, а дромадер — 500 фунтов. За время пути верблюд потребляет 3 тюка сена и 100 галлонов воды, а дромадер — 4 тюка сена и 80 галлонов воды. На пути каравана Марко Поло встречаются пункты снабжения, расположенные в оазисах. Общая емкость запасов на этих участках — 1 600 галлонов воды и 60 тюков сена. Путешественники арендуют верблюдов и дромадеров у пастуха около Багдада. Стоимость аренды верблюда составляет 11 монет, а дромадера — 5 монет. Караван должен доставить из Багдада в Мекку не менее 10 000 фунтов инжира.

Составьте задачу линейного программирования о минимальных издержках на аренду верблюдов и дромадеров. Решите задачу графическим методом. Сколько потребуется верблюдов и дромадеров, чтобы арендная плата пастуху была минимальной?

8. Рацион кормления домашнего животного может состоять из трех имеющихся концентратов: «Вкусного», «Полезного» и «Питательного».

Содержание полезных веществ (белки, кальций и витамины) в 1 кг каждого концентрата приведено в следующей таблице:

Полезное вещество	Затраты труда на единицу продукции, ч		
	«Вкусный»	«Полезный»	«Питательный»
Белок	50	70	180
Кальций	10	6	3
Витамины	2	3	1

Суточные нормы потребления белка, кальция и витаминов составляют 20 г, 2,1 г и 87 мг соответственно. Цена 1 кг концентрата равна 30, 40 и 120 руб. соответственно.

Составьте задачу линейного программирования о рационе с минимальной стоимостью.

9. Средства очистки пола оценивают по трем показателям: очищающие свойства, дезинфицирующие свойства, раздражающее воздействие на кожу. Каждый из этих показателей измеряется по линейной шкале от 0 до 100 единиц.

Смесь для очистки пола составлена из трех основных очистителей (*А*, *Б*, *В*), характеристики которых приведены в следующей таблице:

Свойство	*А*	*Б*	*В*
Очищающее	90	65	45
Дезинфицирующее	70	50	10
Раздражающее	30	85	70

Смесь должна иметь по крайней мере 60 единиц очищающих свойств и по крайней мере 60 единиц дезинфицирующих свойств по соответствующей шкале. При этом раздражающее воздействие на кожу должно быть минимальным.

Сформулируйте задачу линейного программирования об оптимальном составе смеси.

10. Решите задачу линейного программирования графическим методом:

$$4x_1 + 2x_2 \to \max$$

$$\begin{cases} -x_1 + 3x_2 \le 9, \\ 2x_1 + 3x_2 \le 18, \\ 2x_1 - x_2 \le 10, \\ x_1 \ge 0, \\ x_2 \ge 0. \end{cases}$$

11. Решите задачу линейного программирования графическим методом:

$$x_1 + x_2 \to \max$$

$$\begin{cases} 3x_1 - 2x_2 \le 12, \\ -x_1 + 2x_2 \le 8, \\ 2x_1 + 3x_2 \ge 6, \\ x_1 \ge 0, \\ x_2 \ge 0. \end{cases}$$

12. Решите задачу линейного программирования с помощью теории двойственности:

$$2x_1 + 4x_2 + 23x_3 + 4x_4 \to \min$$

$$\begin{cases} -x_1 + 3x_3 - x_4 \ge 2, \\ x_1 + x_2 + 2x_3 - 3x_4 \ge -2, \\ x_i \ge 0. \end{cases}$$

13. Решите задачу линейного программирования с помощью теории двойственности:

$$x_1 + 3x_2 + 4x_3 + 2x_4 \to \min$$

$$\begin{cases} x_1 - x_2 + 4x_3 + 5x_4 \ge 27, \\ 2x_1 + 3x_2 - 2x_3 + 4x_4 \ge 24, \\ x_i \ge 0. \end{cases}$$

14. Решите задачу линейного программирования с помощью теории двойственности:

$$27x_1 + 30x_2 + 32x_3 \to \min$$

$$\begin{cases} 3x_1 + 3x_2 + 2x_3 \ge 4, \\ 2x_1 + 3x_2 + 5x_3 \ge 5, \\ x_i \ge 0. \end{cases}$$

15. Три предприятия выпускают товары в количествах, равных 200 т, 250 т и 350 т соответственно. Эти товары следует доставить на четыре базы, потребности которых составляют 170 т, 120 т, 280 т и 230 т соответственно.

Тарифы перевозок товаров с каждого предприятия в соответствующие пункты назначения заданы матрицей

$$\begin{pmatrix} 5 & 8 & 7 & 3 \\ 4 & 2 & 6 & 5 \\ 7 & 3 & 5 & 9 \end{pmatrix}.$$

Составьте экономико-математическую модель транспортной задачи.

16. На трех элеваторах находится зерно в количествах 225 т, 250 т и 25 т соответственно. Зерно необходимо доставить в четыре фермерских хозяйства, заявки которых составляют 120 т, 150 т, 110 т и 135 т соответственно.

Стоимость доставки зерна от элеваторов к соответствующим хозяйствам задана матрицей тарифов

$$\begin{pmatrix} 9 & 16 & 8 & 15 \\ 10 & 14 & 11 & 17 \\ 15 & 18 & 14 & 16 \end{pmatrix}.$$

Составьте экономико-математическую модель транспортной задачи.

Библиографический список

Основная литература

1. *Красс, М. С.* Математика для экономистов : учеб. пособие / М. С. Красс, Б. П. Чупрынов. — Санкт-Петербург : Питер, 2004. — С. 357—421; 449—457.
2. *Осипов, А. Л.* Математика : учеб. пособие для дистанц. обучения и самост. работы / А. Л. Осипов, Е. А. Рапоцевич ; СибАГС. — Новосибирск, 2005. — 276 с.; *То же* [Электронный ресурс]. — Доступ из Б-ки электрон. изданий / Сиб. ин-т упр. — филиал РАНХиГС. — Режим доступа : http://www.sapanet.ru, требуется авторизация (дата обращения: 19.03.2013). — Загл. с экрана.
3. *Сборник* задач по высшей математике для экономистов : учеб. пособие / под ред. В. И. Ермакова. — 2-е изд., испр. — Москва : ИНФРА-М, 2007. — С. 412—497.
4. *Шикин, Е. В.* Математические методы и модели в управлении : учеб. пособие / Е. В. Шикин, А. Г. Чхартишвили. — 2-е изд., испр. — Москва : Дело, 2002. — С. 65—79.

Дополнительная литература

1. *Бычков, А. Г.* Сборник задач по теории вероятностей, математической статистике и методам оптимизации : учеб. пособие / А. Г. Бычков. — Москва : ФОРУМ, 2008. — 224 с.
2. *Волошин, Г. Я.* Методы оптимизации в экономике : учеб. пособие / Г. Я. Волошин. — Москва : Дело и Сервис, 2004. — 320 с.
3. *Розен, В. В.* Математические модели принятия решений в экономике : учеб. пособие / В. В. Розен. — Москва : Высш. шк., 2002. — 288 с.
4. *Соболь, Б. В.* Методы оптимизации : практикум / Б. В. Соболь, Б. Ч. Месхи, Г. И. Каныгин. — Ростов-на-Дону : Феникс, 2009. — 380 с.
5. *Стрикалов, А. И.* Экономико-математические методы и модели : пособие к решению задач / А. И. Стрикалов, И. А. Печенежская. — Ростов-на-Дону : Феникс, 2008. — 348 с.

Глава 2. Методы прогнозирования

§ 1. Анализ временных рядов

Динамические процессы, происходящие в социально-экономических системах, обычно представляются в виде ряда значений некоторого показателя, последовательно расположенных в хронологическом порядке. Изменение этого показателя отражает ход развития изучаемого процесса.

Последовательность наблюдений одного показателя, упорядоченная в зависимости от последовательно возрастающих или убывающих значений другого показателя, называется динамическим рядом, или рядом динамики. Если в качестве признака, в зависимости от которого происходит упорядочивание, берется время, то такой динамический ряд называется временным рядом.

Одной из важнейших социально-экономических задач является изучение показателей, изменяющихся во времени, — временных рядов, или рядов динамики.

Под ***временным рядом*** (***динамическим рядом***, или ***рядом динамики***) в экономике подразумевается последовательность наблюдений некоторого признака (случайной величины) y в последовательные моменты времени. Отдельные наблюдения называются *уровнями* ряда, которые будем обозначать $y_t (t = 1, 2, \ldots, n)$, где n — число уровней.

Временные ряды, в которых заданы значения социально-экономического показателя, относящиеся к определенным моментам времени, называются *моментными* (например, остатки на счете на первое число каждого месяца). Если уровни временного ряда образуются суммированием, усреднением за некоторый промежуток времени, то такие ряды называются *интервальными временными рядами* (например, объем произведенной продукции по месяцам).

Под *длиной временного ряда* понимают время, прошедшее от начального момента наблюдения до конечного, или число уровней ряда.

Каждый уровень временного ряда формируется под воздействием большого числа факторов, которые условно можно подразделить на следующие группы: факторы, формирующие тенденцию ряда, сезонные колебания ряда, циклические колебания ряда, а также случайные факторы.

Если во временном ряду проявляется длительная закономерность изменения уровней, то говорят, что имеет место *тренд*. Таким образом, тренд определяет общее направление развития социально-экономического процесса. Соответствующая математическая модель называется ***трендовой моделью***. Для выявления тренда временных рядов, а также для построения и анализа трендовых моделей используется аппарат теории вероятностей и математической статистики.

В большинстве случаев фактический уровень временного ряда можно представить как сумму или произведение перечисленных компонент.

Модель, в которой временной ряд представлен как сумма перечисленных компонент, называется ***аддитивной моделью*** временного ряда.

При исследовании экономического временного ряда y_t аддитивная модель имеет следующий *вид*:

$$y_t = u_t + v_t + c_t + \text{e}_t \quad (t = 1, 2, \ldots, n),$$

где u_t — *тренд*, плавно меняющаяся компонента, описывающая чистое влияние долговременных факторов, т. е. длительную тенденцию изменения признака (например, рост населения, экономическое развитие, изменение структуры потребления);

v_t — *сезонная компонента*, отражающая повторяемость экономических процессов в течение не очень длительного периода (года, иногда месяца: например, объем продаж товаров или перевозок пассажиров в различные времена года, уровень безработицы в курортных городах в зимний период выше по сравнению с летним);

c_t — *циклическая компонента*, отражающая повторяемость экономических процессов в течение длительных периодов (например, влияние волн экономической активности Кондратьева, демографических ям, циклов солнечной активности);

ε_t — *случайная компонента*, отражающая влияние не поддающихся учету и регистрации случайных факторов.

В отличие от ε_t первые три составляющие, или компоненты u_t, v_t, c_t, являются *закономерными, неслучайными.* Если они выделены правильно, то случайная (остаточная) компонента будет обладать следующими свойствами:

— случайностью изменения значений;

— соответствием нормальному закону распределения;

— равенством нулю математического ожидания;

— независимостью значений уровня друг от друга, т. е. отсутствием автокорреляции.

Модель, в которой временной ряд представлен как произведение перечисленных компонент, называется ***мультипликативной моделью*** временного ряда:

$$y_t = u_t v_t c_t \text{e}_t.$$

Существует также и смешанная модель вида:

$$y_t = u_t v_t c_t + \text{e}_t.$$

Следует иметь в виду принципиальные отличия временного ряда $y_t (t = 1, 2, \ldots, n)$ от последовательности наблюдений $y_1, y_2, \ldots, y_n$, образующих случайную выборку. Во-первых, в отличие от элементов случайной выборки члены временного ряда, как правило, не являются статистически независимыми. Во-вторых, члены временного ряда не являются одинаково распределенными. Кроме того, уровни временного ряда упорядочены во времени и их перемешивать недопустимо. Элементы статистической совокупности не являются упорядоченными, и их перемешивание не изменяет значения статистических показателей.

Отметим основные *этапы* анализа временных рядов:

— графическое представление и описание поведения временного ряда;

— сглаживание и фильтрация (удаление низко- или высокочастотных составляющих временного ряда);

— выделение и удаление закономерных (неслучайных) составляющих временного ряда (тренда, сезонных и циклических составляющих);

— исследование случайной составляющей временного ряда, построение и проверка адекватности математической модели ее описания;

— прогнозирование развития изучаемого процесса на основе имеющегося временного ряда;

— исследование взаимосвязи между различными временными рядами.

Предварительная обработка временных рядов состоит в выявлении аномальных значений ряда и его сглаживании. Аномальные значения временного ряда не отвечают потенциалу исследуемой системы и их использование для построения трендовой модели может сильно исказить получаемые результаты. Причинами появления аномальных уровней могут быть технические ошибки при сборе, обработке и передаче информации. Такие ошибки называют ошибками первого рода. Эти ошибки можно выявить, устранить и принять меры к их недопущению. Кроме того, аномальные уровни могут возникать из-за воздействия факторов, имеющих объективный характер, но действующих эпизодически. Это ошибки второго рода, их невозможно устранить, но можно исключить из рассмотрения, заменив значение на среднеарифметическое соседних уровней.

Для выявления аномальных значений ряда используется *критерий Ирвина*, согласно которому аномальной считается точка y_t, отстоящая от предыдущей точки y_{t-1} на величину, большую среднеквадратического отклонения:

$$л_t = \frac{|y_t - y_{t-1}|}{у},$$

где $л_t$ — критерий Ирвина;

σ — среднеквадратическое отклонение.

$$у = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \bar{y})^2}{n-1}}.$$

Точка считается аномальной, если $л_t > л_{таб}$. Табличные значения $л_{таб}$ уменьшаются с ростом длины ряда.

Их значения приведены в следующей таблице:

n	10	20	30	50	100
$л_{таб}$	1,5	1,3	1,2	1,1	1,0

Сглаживание временного ряда позволяет отфильтровать мелкие случайные колебания и выявить основную тенденцию изменения исследуемой величины. При механическом сглаживании выравнивание отдельных уровней производится с использованием значений соседних уровней.

Этот прием обработки называется ***методом скользящей средней*** (эмпирическое выравнивание), который заключается в следующем. Берется определенное количество начальных уровней ряда динамики, складывается и таким образом рассчитывается их средний уровень. Далее со второго уровня берется такое же количество показателей, складывается и рассчитывается средний уровень и т. д. В основе приема лежит свойство средней погашать случайные колебания и выявлять тенденцию развития.

Достоинство метода состоит в том, что он позволяет сохранять форму развития динамики явления в сглаженном виде и довольно прост для исчисления. Другими словами, метод скользящей средней имеет ряд преимуществ перед другими методами:

— скользящая средняя дает функцию тенденции, значения которой самые близкие к значениям исследуемого ряда, поскольку для отдельных частей ряда выбирается наилучшая тенденция;

— к исследуемому ряду могут быть прибавлены новые уровни;

— нахождение тенденции требует небольшого труда.

Недостаток метода состоит в том, что при сглаживании по большому количеству уровней теряются начальные и конечные показатели ряда динамики, что недопустимо при некоторых приемах анализа временных рядов (например, спектральный анализ). Число теряемых уровней равно

$$h = k - 1,$$

где k — число уровней, по которым проводили сглаживание.

Скользящие средние находим по формуле

$$\tilde{y}_t = \frac{\sum_{i=t-p}^{t+p} y_i}{m},$$

когда $m = 2p + 1$ — нечетное число. При $m = 3$ находим, что $p = 1$.

Наиболее часто используется сглаживание по 5 точкам:

$$\tilde{y}_t = \frac{y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2}}{5}.$$

Взвешенную (средневзвешенную) скользящую среднюю определим по формуле

$$\tilde{y}_t = \frac{\sum_{i=t-p}^{t+p} c_i y_i}{\sum_{i=t-p}^{t+p} c_i}.$$

В этом методе каждая из точек входит в общую сумму с весовым коэффициентом ρ_i. Для сглаживания по 5 точкам используют весовые коэффициенты (– 3, 12, 17, 12, – 3).

При сглаживании временного ряда по $2p + 1$ соседним точкам p точек в начале и в конце ряда остаются несглаженными. Эти точки следует либо исключить из рассмотрения, либо использовать для них специальные формулы сглаживания для крайних точек. В частности, для сглаживания по трем точкам можно использовать формулы

$$\tilde{y}_1 = \frac{5y_1 + 2y_2 - y_3}{6}, \quad \tilde{y}_n = \frac{5y_n + 2y_{n-1} - y_{n-2}}{6}.$$

Одним из наиболее распространенных способов моделирования тенденции временного ряда является построение аналитической функции, характеризующей зависимость уровней ряда от времени, или тренда, — это ***аналитическое выравнивание временного ряда***. Подбор аналитической функции, в наибольшей степени выражающей тенденцию развития, имеет следующие *этапы*:

— с помощью графика или методом перебора определяется вид зависимости, которой подчинено развитие явления;

— через систему уравнений из метода наименьших квадратов рассчитываются параметры уравнения;

— рассчитываются значения временной функции, которые определяют тенденцию развития;

— при необходимости проводится экстраполяция или интерполяция тенденции развития.

Под *интерполяцией* понимается прием заполнения отсутствующих данных по выявленной тенденции развития.

Для определения наличия тренда временного ряда используется ***метод проверки разностей средних уровней***.

Исходный ряд из n точек делится на два с примерно одинаковым числом точек n_1 и n_2 ($n = n_1 + n_2$).

Для каждой из частей вычислим средние значения и дисперсию:

$$\overline{y}_1 = \frac{\sum_{i=1}^{n_1} y_i}{n_1}, \ \text{у}_1^2 = \frac{\sum_{i=1}^{n_1} (y_i - \overline{y}_1)^2}{n_1 - 1};$$

$$\overline{y}_2 = \frac{\sum_{i=1}^{n_2} y_i}{n_2}, \ \text{у}_2^2 = \frac{\sum_{i=1}^{n_2} (y_i - \overline{y}_2)^2}{n_2 - 1}.$$

Гипотеза об однородности дисперсий частей ряда проверяется с помощью *критерия Фишера*:

$$F_{\text{наб}} = \begin{cases} \text{у}_1^2 / \text{у}_2^2, \text{если } \text{у}_1^2 > \text{у}_2^2; \\ \text{у}_2^2 / \text{у}_1^2, \text{если } \text{у}_2^2 > \text{у}_1^2. \end{cases}$$

Если $F_{\text{наб}} < F_{\text{таб}}$, то гипотеза об однородности дисперсии принимается и мы переходим к следующему этапу проверки; если больше или равно, то гипотеза об однородности дисперсии отклоняется и метод не дает ответа на вопрос о наличии или отсутствии тренда. Табличное значение определяется так: $F_{\text{таб}} = F(1 - \text{б}/2, n_1 - 1, n_2 - 1)$ — квантиль распределения Фишера.

Окончательная проверка гипотезы об отсутствии тренда проводится с использованием *критерия Стьюдента*:

$$t_{\text{наб}} = \frac{\left|\bar{y}_1 - \bar{y}_2\right|}{\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}},$$

где $\sigma = \sqrt{\dfrac{(n_1 - 1)\,\sigma_1^2 + (n_2 - 1)\,\sigma_2^2}{n_1 + n_2 - 2}}$.

Если $t_{\text{наб}} < t_{\text{таб}}$, то гипотеза принимается и тренда нет, в противном случае тренд есть. Здесь $t_{\text{таб}} = t\,(1 - \alpha/2, n - 2)$ — квантиль распределения Стьюдента, где α — уровень значимости (обычно α = 0,05).

Изучая показатели абсолютных приростов, коэффициентов или темпов роста (прироста)[1], можно высказать гипотезу о наиболее адекватной форме модели тренда.

Полезно придерживаться, например, следующих рекомендаций:

— если ряд динамики имеет постоянные или почти постоянные ежегодные (ежеквартальные, ежемесячные и т. п.) абсолютные приросты, то для отображения тренда следует воспользоваться линейной моделью:

$$\hat{y}_t = a + bt;$$

— если ряд динамики описывает процесс развития объекта и имеет ежегодно возрастающий абсолютный прирост, то можно выдвинуть гипотезу о показательной форме модели тренда:

$$\hat{y}_t = ak^t.$$

Для нахождения коэффициентов уравнения тренда используется ***метод наименьших квадратов*** (МНК).

Линейный тренд имеет уравнение вида

$$\hat{y}_i = a + bt_i,$$

где $\hat{y}_i$ — уровень тренда для периода или момента с номером t_i;

a — свободный член уравнения, равный среднему уровню тренда для периода (момента) с нулевым номером t_i;

b — главный параметр линейного тренда, который показывает среднее абсолютное изменение за принятую в ряду единицу времени.

[1] Абсолютный прирост — это разность между сравниваемым уровнем и уровнем более раннего периода, принятым за базу сравнения. Если эта база — непосредственно предыдущий уровень, то показатель называется цепным.

Коэффициент роста — это отношение между сравниваемым и предыдущим уровнями.

Темп роста — это коэффициент роста, выраженный в процентах.

Величина параметров a и b определяется по МНК путем приравнивания частных первых производных функции $f(a,b)=\sum_{i=1}^{n}(y_i-\tilde{y}_i)^2=\sum_{i=1}^{n}(y_i-a-bt_i)^2$ к нулю. Имеем

$$\frac{\partial f}{\partial a}=2\sum_{i=1}^{n}(y_i-a-bt_i)(-1)=0,$$
$$\frac{\partial f}{\partial b}=2\sum_{i=1}^{n}(y_i-a-bt_i)(-t_i)=0.$$

После алгебраических преобразований получим два «нормальных уравнения» МНК для прямой:

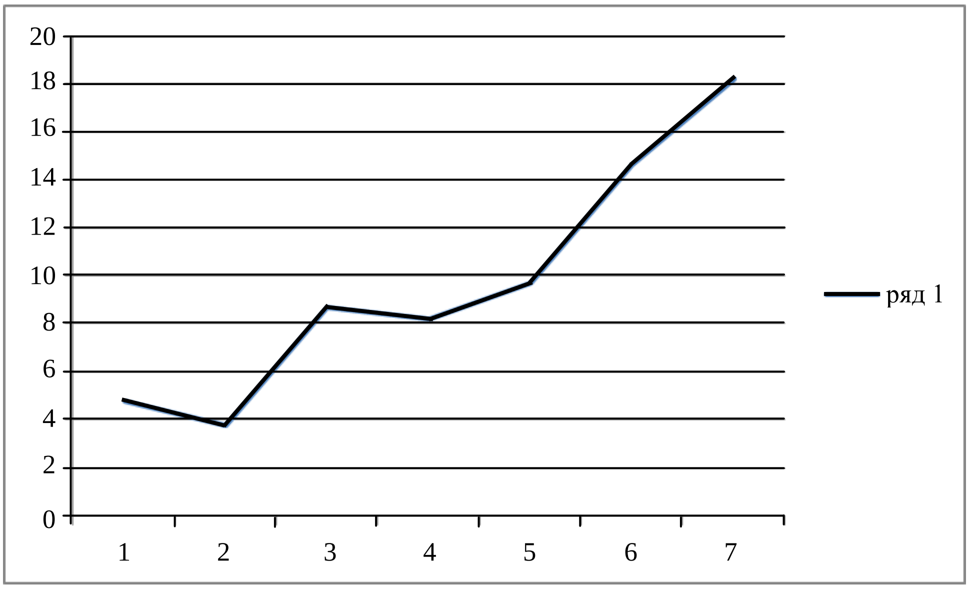
$$na+b\sum_{i=1}^{n}t_i=\sum_{i=1}^{n}y_i,$$
$$a\sum_{i=1}^{n}t_i+b\sum_{i=1}^{n}t_i^2=\sum_{i=1}^{n}(y_it_i).$$

Решая эти уравнения с двумя неизвестными по данным фактического временного ряда $y_i(i=1,2,\ldots,n)$, получим значения a и b.

Пример 1. В таблице дан временной ряд, характеризующий динамику затрат на рекламу по месяцам.

Месяц	1	2	3	4	5	6	7
Затраты	4,8	3,8	8,7	8,2	9,7	14,6	18,2

График временного ряда имеет следующий вид:



Построим линейный тренд. Для этого составим вспомогательную таблицу:

t	*y*	*t*^2	*yt*
1	4,8	1	4,8
2	3,8	4	7,6
3	8,7	9	26,1
4	8,2	16	32,8
5	9,7	25	48,5
6	14,6	36	87,6
7	18,2	49	127,4
Итого			
28	**68**	**140**	**334,8**

Тогда система уравнений МНК будет иметь вид

$$\begin{pmatrix} 7 & 28 \\ 28 & 140 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 68 \\ 334,8 \end{pmatrix}.$$

Решая эту систему, используя знания линейной алгебры или средства MS Excel, находим значения: a = 0,7429; b = 2,2429. Итоговое уравнение линейного тренда имеет вид

$$\hat{y}_i = 0,7429 + 2,2429\ t_i.$$

Полиномиальный (*параболический*) *тренд* имеет уравнение вида

$$\hat{y}_i = a + bt_i + ct_i^2\ .$$

Для вычисления параметров a, b, c по МНК три частные производные функции $f(a,b,c) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ приравниваются к нулю. После преобразований получим систему трех уравнений с тремя неизвестными:

$$na + b\sum_{i=1}^{n} t_i + c\sum_{i=1}^{n} t_i^2 = \sum_{i=1}^{n} y_i,$$

$$a\sum_{i=1}^{n} t_i + b\sum_{i=1}^{n} t_i^2 + c\sum_{i=1}^{n} t_i^3 = \sum_{i=1}^{n} y_i t_i,$$

$$a\sum_{i=1}^{n} t_i^2 + b\sum_{i=1}^{n} t_i^3 + c\sum_{i=1}^{n} t_i^4 = \sum_{i=1}^{n} y_i t_i^2.$$

Пример 2. Построим полиномиальный тренд для задачи из предыдущего примера. Для этого составим вспомогательную таблицу:

t	*y*	*t*^2	*yt*	*t*^3	*t*^4	*yt*^2
1	4,8	1	4,8	1	1	4,8
2	3,8	4	7,6	8	16	15,2
3	8,7	9	26,1	27	81	78,3
4	8,2	16	32,8	64	256	131,2
5	9,7	25	48,5	125	625	242,5
6	14,6	36	87,6	216	1 296	525,6
7	18,2	49	127,4	343	2 401	891,8
Итого						
28	**68,0**	**140**	**334,8**	**784**	**4 676**	**1 889,4**

Тогда система уравнений МНК будет иметь вид

$$\begin{pmatrix} 7 & 28 & 140 \\ 28 & 140 & 784 \\ 140 & 784 & 4\,676 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 68 \\ 334,8 \\ 1889,4 \end{pmatrix}.$$

Решив эту систему, получим следующий результат: a = 4,6; b = – 0,3286; c = 0,3214. Окончательное уравнение тренда запишем в виде

$$\hat{y}_i = 4,6 - 0,3286\, t_i + 0,3214\, t_i^2 .$$

Гиперболический тренд имеет уравнение вида $\hat{y}_i = a + \dfrac{b}{t_i}$, т. е. отличается от линейного уравнения тем, что вместо t_i в первой степени включает номера периодов времени (моментов) в минус первой степени: $\dfrac{1}{t_i}$. Соответственно нормальные уравнения МНК будут иметь вид

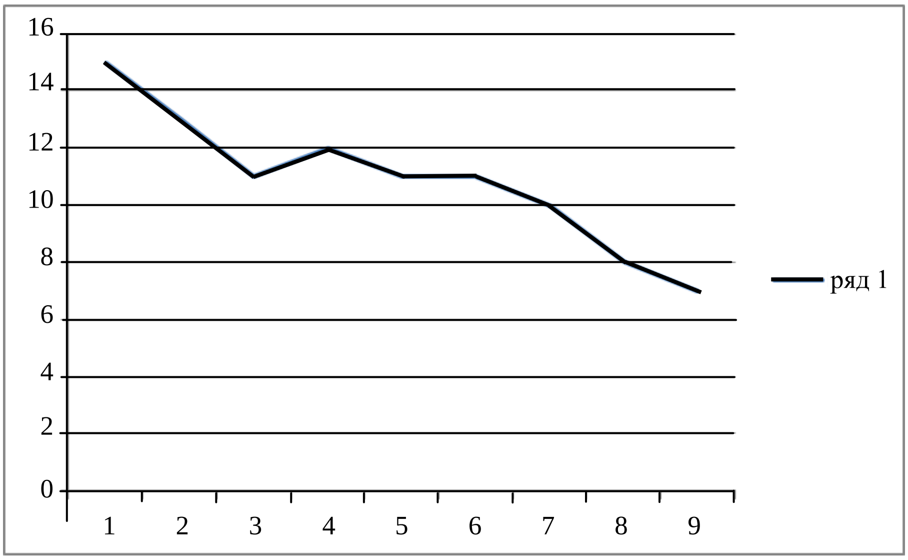
$$na + b\sum_{i=1}^{n} \frac{1}{t_i} = \sum_{i=1}^{n} y_i;$$

$$a\sum_{i=1}^{n} \frac{1}{t_i} + b\sum_{i=1}^{n} \frac{1}{t_i^2} = \sum_{i=1}^{n} \frac{y_i}{t_i}.$$

Пример 3. Имеются данные за 9 месяцев об уровне безработицы y_t (к общему числу трудоспособного населения области, %):

t	1	2	3	4	5	6	7	8	9
y_t	15	13	11	12	11	11	10	8	7

График временного ряда имеет следующий вид:



Построим гиперболический тренд. Для этого составим вспомогательную таблицу:

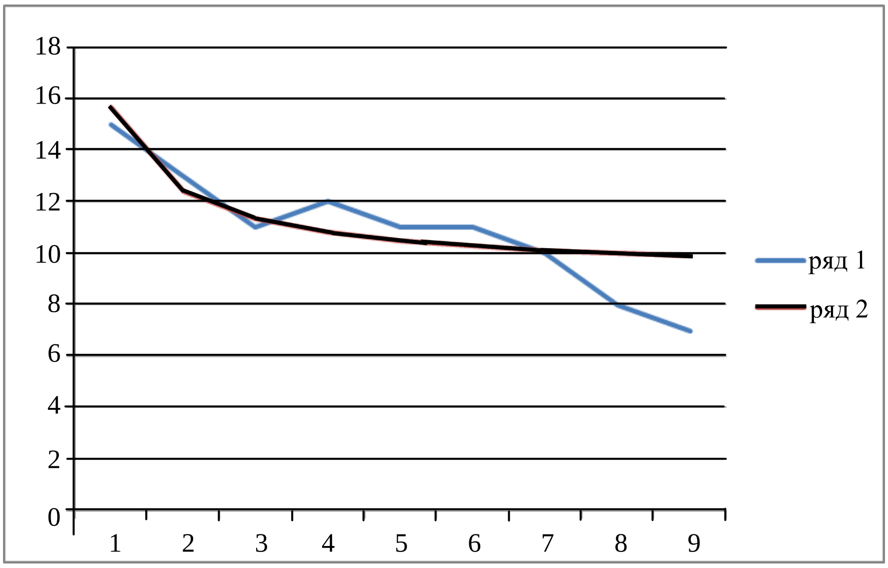
t	y	$1/t$	$1/t$^2	y/t
1	15	1	1	15
2	13	0,5	0,25	6,5
3	11	0,333333	0,111111	3,666667
4	12	0,25	0,0625	3
5	13	0,2	0,04	2,6
6	11	0,166667	0,027778	1,833333
7	11	0,142857	0,020408	1,571429
8	8	0,125	0,015625	1
9	7	0,111111	0,012346	0,777778
Итого				
45	**101**	**2,828968**	**1,539768**	**35,94921**

Тогда система уравнений МНК будет иметь вид

$$\begin{pmatrix} 9 & 2,829 \\ 2,829 & 1,5398 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 101 \\ 35,949 \end{pmatrix}.$$

Решая эту систему, используя знания линейной алгебры или средства MS Excel, находим значения: a = 9,1919; b = 6,4591. Итоговое уравнение гиперболического тренда имеет вид $\tilde{y}_i = 9,1919 + 6,4591/\ t_i$.

Результат представлен на следующем рисунке:



Показательный тренд имеет уравнение вида

$$\tilde{y}_i = ak^{t_i}.$$

Для нахождения параметров a и k уравнение прологарифмируем:

$$\ln \tilde{y}_i = \ln a + t_i \ln k .$$

В такой форме, т. е. для логарифмов, уравнение соответствует линейному. Следовательно, МНК дает для логарифмов a и k нормальные уравнения, аналогичные таковым для параметров a и b линейного тренда:

$$n \ln a + \ln k \sum_{i=1}^{n} t_i = \sum_{i=1}^{n} \ln y_i;$$

$$\ln a \sum_{i=1}^{n} t_i + \ln k \sum_{i=1}^{n} t_i^2 = \sum_{i=1}^{n} t_i \ln y_i .$$

Линейно-логарифмический тренд имеет уравнение вида

$$\hat{y}_i = a + b\ln t_i.$$

Особенность этого тренда заключается в том, что логарифмировать необходимо номера периодов (моментов) времени.

Следовательно, все номера должны быть положительными числами. Построение уравнения тренда происходит аналогично разобранным примерам.

Если поиск уравнения трендов осуществляется с помощью компьютера и MS Excel, то в приложении «добавить линию тренда» нужно указать опции «показать уравнение на диаграмме» и «поместить на диаграмму величину достоверности аппроксимации (R^2)». Последний параметр есть не что иное, как коэффициент детерминации. О нем подробно мы поговорим в теме «Линейная регрессия». Здесь же заметим, что если, используя указанное приложение, мы построили по заданному графику пять предлагаемых трендов, то по наибольшему значению коэффициента детерминации (R^2) можно определить наилучший тренд из рассмотренных.

Сезонными называются колебания, которые связаны со сменой времен года и которые поэтому повторяются ежегодно.

Часто объемы производства и потребления имеют существенные сезонные колебания (они происходят в течение года). Так, от времени года существенно зависят потребление топлива, производство сельскохозяйственных продуктов и т. д.

Для изучения сезонных колебаний обычно используется *метод одногодичных или многолетних средних*.

Метод одногодичных средних используется в том случае, если нет резких перепадов в уровнях ряда, чаще всего помесячной динамики в течение одного года.

Исследование сезонности сводится к расчету индексов (коэффициентов) сезонности для каждого месяца (квартала) внутри года по формуле

$$J_t = \frac{y_t}{\bar{y}} 100\ \%$$

(темп роста по отношению к y, за базу взято среднее значение $\bar{y} = \frac{1}{n}\sum y_t$),

где n — число временных интервалов в году (если n = 12, то речь идет о месячных колебаниях; если n = 4, то речь идет о квартальных колебаниях).

Индексы сезонности являются таким образом отношением уровней каждого месяца к среднемесячному уровню за весь год, которые обычно выражаются в процентах.

Виды сезонных колебаний по форме являются выпуклыми, когда явление возрастает к середине года и снижается к началу или концу (строительство, добыча полезных ископаемых открытым способом, торговля и т. д.), или вогнутыми, когда явление снижается к середине года и возрастает к концу (торговля товарами зимнего ассортимента, мясная промышленность, производство сахара, электроника, потребление тепла и т. д.).

Виды сезонных колебаний по промежутку времени — это колебание внутри года, месяца, квартала, рабочего дня.

Расчеты сезонных колебаний часто проводят с помощью *ряда Фурье*:

$$y_t = a_0 + \sum_{k=1}^{m} (a_k \cos kt + b_k \sin kt),$$

где y_t — выравненные уровни ряда;

a_0, a_k, b_k — параметры уравнения;

cos t, sin t — тригонометрические функции, соответствующие каждому месяцу года;

k — номер гармоники;

m — взятое количество гармоник.

Параметры уравнения ряда Фурье для m = 2 рассчитываются по формулам

$$a_0 = \frac{\sum_{t=1}^{n} y_t}{n}; \ a_1 = \frac{\sum_{t=1}^{n} y_t \cos t}{6}; \ a_2 = \frac{\sum_{t=1}^{n} y_t \cos 2t}{6};$$

$$b_1 = \frac{\sum_{t=1}^{n} y_t \sin t}{6}; \ b_2 = \frac{\sum_{t=1}^{n} y_t \sin 2t}{6}.$$

Для исследования *случайной компоненты* ε_t на нормальность используют *R*/*S* *критерий*. *R*/*S* критерий — один из самых простых критериев проверки случайной величины на нормальность. Он рассчитывается по формуле

$$R/S = \frac{\max_i \varepsilon_i - \min_i \varepsilon_i}{S_e},$$

где $S_e = \sqrt{\frac{\sum_{i=1}^{n} \varepsilon_i^2}{n-1}}$ — стандартная ошибка.

Значения критерия сравниваются с табличными нижней и верхней границами отношения. Если вычисленное значение критерия попадает в табличный интервал, то гипотеза о нормальности распределения принимается, в противном случае гипотеза отвергается (критические значения границ критерия приведены в табл. 1 приложения).

В результате построенная модель временного ряда используется для прогнозирования будущих значений (экстраполяция) либо для восстановления отсутствующих в исходных данных внутренних значений (интерполяция).

§ 2. Качественные методы прогнозирования

Если отсутствуют количественные данные или количественная модель получается слишком сложной и громоздкой, то используются ***качественные методы прогнозирования***, которые строятся на основе экспертных оценок.

Сущность *метода экспертных оценок* заключается в проведении экспертами интуитивно-логического анализа проблемы с количественной оценкой суждений и формальной обработкой результатов. Полученное в результате обработки обобщенное мнение экспертов принимается как решение проблемы. Комплексное использование интуиции, логического мышления и количественных оценок с их формальной обработкой позволяет получить эффективное решение проблемы.

Метод экспертных оценок используется для решения различных сложных неформализуемых проблем. Все множество плохо формализуемых проблем можно условно разделить на два класса.

К первому классу относятся проблемы, в отношении которых имеется достаточный информационный потенциал.

Ко второму классу относятся проблемы, в отношении которых информационный потенциал знаний недостаточен.

Основные трудности применения экспертной оценки к решению проблем первого класса заключаются в подборе экспертов, построении рациональных процедур опроса и выборе оптимальных методов обработки его результатов. При этом метод опроса и обработки основывается на использовании «хорошего измерителя», т. е. выполняются следующие условия:

— эксперт располагает большим объемом рационально обработанной информации, и поэтому он может рассматриваться как качественный источник информации (своего рода «информационный измеритель с небольшими погрешностями»);

— групповое мнение экспертов близко к истинному решению проблемы.

Если эти условия выполняются, то для построения процедур опроса и алгоритмов обработки можно использовать теорию измерений и математическую статистику.

При решении проблем второго класса экспертов уже нельзя рассматривать как «хороших измерителей». Поэтому необходимо очень осторожно проводить обработку результатов экспертизы. Применение методов осреднения, справедливых для «хороших измерителей», в данном случае может привести к большим ошибкам. Например, мнение одного эксперта, сильно отличающееся от мнения остальных экспертов, может оказаться правильным. В связи с этим для проблем второго класса в основном должна применяться качественная обработка.

Область применения метода экспертных оценок обширна и охватывает неформализуемые проблемы первого и второго классов.

Перечислим *типовые задачи*, решаемые методом экспертных оценок:

— составление перечня возможных событий в различных областях за определенный промежуток времени;

— определение наиболее вероятных интервалов времени наступления совокупности событий;

— определение целей и задач управления, упорядочение их по степени важности;

— определение альтернативных вариантов решения задачи, оценка их предпочтения;

— альтернативное распределение ресурсов для решения задач, оценка их предпочтительности;

— выбор альтернативных вариантов принятия решений в определенной ситуации, оценка их предпочтительности.

Можно ориентировочно наметить следующие основные *этапы проведения экспертизы*, последовательность и содержание которых будут изменяться в зависимости от реальных условий и ограничений:

— формулирование цели экспертизы;

— формирование группы специалистов-аналитиков;

— отбор и формирование группы экспертов;

— проведение опроса;

— анализ и обработка информации экспертов;

— синтез объективной (статистической) информации и информации, полученной в результате экспертизы, с целью приведения их в форму, удобную для принятия решения.

Подготовка экспертизы включает три этапа. Большое значение имеет четкое определение цели (целей) экспертизы. Основой для выбора целей является описание предыстории и текущего состояния проблемы. Выбор целей и характер экспертизы в значительной степени обусловливаются существом проблемы, предполагаемыми конечными результатами и возможными способами их представления.

После определения цели (целей) экспертизы формируется группа специалистов-аналитиков, важнейшими задачами которой являются разработка методов опроса, отбор экспертов, проведение опроса, анализ и обобщение информации. Большой объем, сложность и разнообразие задач, возлагаемых на группу аналитиков, требуют включения в ее состав высококвалифицированных специалистов как в области исследуемой проблемы, так и в смежных областях деятельности, а также специалистов по экспертным методам — математиков, психологов и социологов.

Группа аналитиков, разрабатывая метод опроса, подготавливает перечень (множество) оцениваемых событий и устанавливает совокупность устойчивых факторов, характеризующих эти события. Определение совокупности факторов зависит от специфики и целей экспертизы и может быть выполнено на разном уровне детализации. Наметим следующие уровни:

— их качественное описание;

— составление перечня оцениваемых событий;

— описание устойчивых факторов для каждого события;

— выделение числа различимых уровней для каждого события;

— выделение числа различимых уровней для каждого фактора;

— описание набора устойчивых значений факторов для каждого уровня событий и т. д.

От уровня детализации в существенной степени зависит достоверность результатов экспертизы, причем с увеличением степени детализации согласованность экспертных оценок, как правило, увеличивается.

Анкетирование экспертов является наиболее эффективным и самым распространенным видом опроса, который заключается в заполнении экспертами опросных листов (анкет).

Серьезного внимания требует подбор вопросов, которые желательно включить в анкету. Различают три вида вопросов, по которым дается экспертная оценка:

— вопросы, ответы на которые содержат количественную оценку;

— требующие содержательного ответа в сжатой форме;

— требующие содержательного ответа в развернутой форме.

Чтобы уточнить содержание анкеты — формулировка вопросов (исключаются вопросы, допускающие двойное толкование), их последовательность, полнота перечня (выясняется, не нужно ли включить дополнительные и исключить «неработающие» вопросы), — проводится ее проверка. Для этого подбираются «разноплановые» эксперты (по стажу работы и специализации). Проверка осуществляется в форме личного интервью.

Исследуемые объекты, проблемы или явления можно опознавать или различать на основе присущих им свойств, выражаемых факторами.

Фактор — это свойство объекта, которое позволяет на множестве, состоящем, по крайней мере, из двух элементов, отражать различные уровни некоторых подлежащих рассмотрению величин.

Каждый фактор, выражая определенное свойство объекта, одновременно является оценкой отношения к данному свойству со стороны лица, принимающего решение.

Уровень одних факторов может быть выражен количественно (в рублях, процентах, тоннах и т. д.), такие факторы называются количественными.

Уровень же других нельзя выразить с помощью числа, они обычно называются качественными.

При решении многих практических задач некоторые факторы, определяющие конечные результаты, не поддаются непосредственному измерению, и тогда используют *ранжирование*.

Ранжирование может применяться в следующих ситуациях:

— когда необходимо упорядочить какие-либо явления (объекты) во времени или пространстве. Это ситуация, когда интересуются не сравнением степени выраженности какого-либо их качества, а лишь взаимным пространственным или временным расположением этих объектов (явлений);

— когда нужно упорядочить объекты в соответствии с каким-либо качеством, но при этом не требуется производить его точное измерение;

— когда какое-либо качество в принципе измеримо, однако в настоящий момент не может быть измерено по причинам практического или теоретического характера.

Под ранжированием понимается процедура расположения факторов x_i ($i = 1,\ldots, n$) в порядке их существенности: на первом месте стоит самый существенный, следом за ним — менее существенный, но самый важный из оставшихся, и т. д.

Ранжированную последовательность можно составить с помощью специалистов-экспертов, имеющих представление о будущем алгоритме управления этим объектом. С помощью экспертов составляется последовательность рангов k_1, k_2, …, k_n, где k_i — ранг i-го входа x_i. Построить из нее ранжированный ряд не представляет труда. Например, при $n = 5$ последовательность рангов k_1, k_2, …, k_n может иметь вид 3, 1, 5, 4, 2.

При ранжировании эксперт должен расположить объекты (альтернативы) в порядке, который представляется ему наиболее рациональным, и приписать каждому из них числа натурального ряда — ранги. Так, ранг 1 — наиболее предпочтительная альтернатива, а ранг N — наименее предпочтительная. Следовательно, порядковая шкала, получаемая в результате ранжирования, должна удовлетворять условию равенства числа рангов N числу ранжируемых объектов n.

При анализе оценок, полученных от экспертов, часто возникает задача определения общей меры согласованности экспертных оценок. Для этого используют *коэффициент конкордации* (W), который является числовым критерием согласованности мнений экспертов в рассматриваемой группе. Поскольку каждый объект характеризуется совокупностью рангов, полученных от m экспертов, то в основу статистической меры согласованности, очевидно, может быть положена средняя сумма рангов одного объекта и отклонения от нее.

Формула для коэффициента согласованности имеет следующий вид:

$$W = \frac{12\sum_{i=1}^{n} S_i^2}{m^2(n^3 - n)},$$

где n — число объектов;

m — число экспертов;

S_i — отклонение суммы рангов i-го объекта от средней их суммы для всех объектов.

Если оценки всех экспертов совпадают, то $W = 1$.

Если же оценки экспертов полностью не совпадают, то коэффициент $W < 1$. Наименьшее возможное его значение равно нулю. В случае $W < 0{,}4$ говорят о слабой согласованности экспертов, а большие величины — $W > 0{,}7$ — свидетельствуют о сильной согласованности экспертов.

Слабая согласованность обычно является следствием следующих причин: в группе экспертов действительно отсутствует общность мнений или внутри группы существуют коалиции с высокой согласованностью мнений, однако обобщенные мнения коалиций противоположны.

Пример 4. Пусть пять дегустаторов выразили предпочтения следующим образом относительно вкусовых качеств продукта, выпускаемого пятью заводами:

Дегустатор	Завод					**Итого**
	1	2	3	4	5	
1	1	3	4	2	5	—
2	1	2	5	3	4	—
3	2	2	5	3	4	—
4	1	1	4	5	3	—
5	3	1	4	2	5	—
Сумма рангов	8	9	22	15	21	**75**
S_i	–7	–6	7	0	6	—
S_i^2	49	36	49	0	36	**170**

Величины S_i здесь получены как отклонения суммы рангов от средней, равной 15.

Найдем коэффициент согласованности:

$$W = \frac{12 \cdot 170}{25\,(125 - 5)} = 0,68.$$

Полученная величина коэффициента конкордации ($W = 0{,}68$) показывает среднюю степень согласованности мнений экспертов.

В настоящее время распространены экспертные, маркетинговые, социологические и иные опросы, в которых опрашиваемых просят выставить баллы объектам, изделиям, технологическим процессам, предприятиям, проектам, заявкам на выполнение научно-исследовательских работ, идеям, проблемам, программам, политикам и т. п., а затем рассчитывают средние баллы и рассматривают их как интегральные оценки, выставленные коллективом опрошенных.

Для вычисления средних величин обычно рассчитывают среднеарифметическое. Известно, что такой способ некорректен, поскольку баллы обычно измеряются в порядковой шкале. Поэтому обоснованным является использование медиан в качестве средних баллов.

Однако полностью игнорировать среднеарифметические не следует из-за их распространенности, поэтому целесообразно использовать одновременно оба метода — и метод среднеарифметических рангов, и метод медианных рангов.

Такая рекомендация находится в согласии с концепцией устойчивости, позволяющей использовать различные методы для обработки одних и тех же данных с целью получить выводы, подтверждаемые всеми методами.

При использовании *метода среднеарифметических рангов* в первую очередь подсчитывается сумма рангов. Затем эта сумма делится на число экспертов и в результате рассчитывается среднеарифметический ранг (именно эта операция дала название методу). По средним рангам строится итоговая ранжировка (в другой терминологии — упорядочение) исходя из принципа: чем меньше средний ранг, чем лучше проект.

При использовании *метода медиан* все ранги по каждой оцениваемой позиции располагаются в порядке неубывания. Вычисляется медиана рангов. В итоге получается совокупность медиан, которая снова ранжируется. Лучшей считается позиция с наименьшим итоговым рангом.

Следующим этапом является сравнение ранжировок по методу среднеарифметических и методу медиан.

Пример 5. По данным голосования пяти избирателей при выборе трех кандидатов по десятибалльной шкале экспертам следует найти наилучшего кандидата с помощью медианного метода и метода средних оценок (таблица).

Для получения группового мнения экспертов воспользуемся методом среднеарифметических рангов.

Эксперт	Кандидат		
	A	*B*	*C*
1	2	4	6
2	1	2	5
3	6	8	3
4	9	4	7
5	5	3	2

Подсчитаем сумму рангов, присвоенных кандидатам, разделим на число экспертов и в результате рассчитаем среднеарифметический ранг. По средним рангам построим итоговую ранжировку исходя из принципа: чем меньше средний ранг, чем лучше кандидат.

Наименьший средний ранг (4,2) имеет кандидат *B*. Следовательно, в итоговой ранжировке он получит ранг 1.

Кандидаты *A* и *C* имеют одинаковые суммы, равные 4,6. Значит, с точки зрения экспертов, они равноценны (при рассматриваемом способе сведения вместе мнений экспертов), а потому они должны бы стоять на 2-м и 3-м местах и получить средний балл (2 + 3) / 2 = 2,5.

Итак, ранжировка по суммам рангов имеет следующий вид: $B < \{A, C\}$.

Далее воспользуемся методом медиан рангов. Возьмем ответы экспертов, соответствующие, например, кандидату *A* — это ранги 2, 1, 6, 9, 5. Расположим их в порядке возрастания и получим последовательность: 1, 2, 5, 6, 9. На центральном месте стоит 5. Следовательно, медиана равна 5.

Кандидат *B* имеет ранги: 4, 2, 8, 4, 3. Расположим их в порядке возрастания и получим следующую последовательность: 2, 3, 4, 4, 8. На центральном месте стоит 4. Следовательно, медиана равна 4.

Кандидат *C* имеет ранги: 6, 5, 3, 7, 2. Расположим их в порядке возрастания и получим последовательность: 2, 3, 5, 6, 7. На центральном месте стоит 5. Следовательно, медиана равна 5.

Эксперт	Кандидат		
	A	*B*	*C*
1	2	4	6
2	1	2	5
3	6	8	3
4	9	4	7
5	5	3	2
Медиана рангов	5	4	5
Итоговый ранг по медианам	**2,5**	**1**	**2,5**

Ранжировка по медианам имеет вид $B < \{A, C\}$.

Таким образом, два метода при выборе наилучшего кандидата показали одинаковые результаты. По мнению экспертов, наилучшим является кандидат B, так как он на первом месте, а кандидаты A и C — равноценны и хуже кандидата B.

Контрольные вопросы и задания

1. В чем заключается суть временного ряда?
2. Как провести предварительный анализ временного ряда?
3. Для каких целей используется критерий Ирвина?
4. Как происходит выравнивание временного ряда?
5. Определите понятие линейного тренда.
6. Определите понятие гиперболического тренда.
7. Что означают сезонные колебания? Как они рассчитываются?
8. В чем заключается суть коэффициента согласованности?
9. Зачем используется процедура ранжирования?
10. Изложите суть метода среднеарифметических рангов и метода медиан.

Задания для самостоятельного решения

1. Дан временной ряд, характеризующий динамику прибыли некоторой продукции:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Прибыль	48	80	97	24	78	57	84	41	60	35	35	48	47	68	66

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

2. Дан временной ряд, характеризующий динамику издержек некоторой продукции:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Издержки	40	82	97	24	78	57	84	41	60	35	35	48	47	48	56

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

3. Дан временной ряд, характеризующий динамику прибыли некоторой продукции:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Прибыль	24	28,2	29,7	32,4	27,8	25,7	38,4	40,1	36	33,5	35	48,1	47,9	48,3	45,6

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

4. Данные о состоянии динамики урожайности представлены в следующей таблице:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Урожайность	16	21	18	8	16	17	20	18	22	23	20	24	19	17	26

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

5. Данные о состоянии уровня безработицы в нашем городе за последние 15 месяцев представлены в следующей таблице:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Уровень безработицы	54	63	65	58	65	75	81	90	103	107	112	125	120	115	128

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

6. Дан временной ряд, характеризующий динамику продаж некоторой продукции:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Динамика продаж	23	20	29	32	27	25	38	40	36	33	46	48	47	48	45

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

7. Дан временной ряд, характеризующий динамику выпуска продукции предприятием:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Динамика выпуска	40	80	87	96	104	113	124	135	149	145	148	151	152	154	160

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

8. Дан временной ряд, характеризующий динамику численности занятых в сфере услуг фирмы по месяцам:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Динамика численности	34	36	39	44	52	55	59	65	67	73	82	86	92	93	98

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

9. Дан временной ряд, характеризующий месячный объем товарооборота фирмы:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
y_t	38	35	39	42	52	53	59	68	67	78	82	85	92	96	98

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

10. Данные о состоянии динамики урожайности представлены в следующей таблице:

Год	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Состояние	16	21	18	8	16	17	20	18	22	23	20	24	19	17	26

Проведите сглаживание исходных данных по трем, пяти и 5* точкам, постройте графики и визуально выберете лучший. К выбранному примените критерий Ирвина и определите наличие тренда по методу разностей средних уровней. Выберете лучший тренд средствами EXCEL.

11. Данные о состоянии уровня преступности в процентах за последние 15 месяцев нашего города представлены в следующей таблице:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Уровень преступности	59	60	62	60	58	65	75	81	15	103	103	112	116	122	125

Постройте оптимальный тренд и на его основе получите прогноз изменения преступности в процентах за 16-й и 17-й месяцы.

12. Дан временной ряд, характеризующий фонд заработной платы работников фирмы по месяцам:

Месяц	Январь	Февраль	Март	Апрель	Май	Июнь
Фонд заработной платы, ден. ед.	1 520	1 590	1 650	1 710	1 780	1 890

Определите оптимальный тренд и рассчитайте точечный прогноз фонда заработной платы на последующие 3 месяца.

13. Дан временной ряд, характеризующий динамику продаж некоторой продукции:

Год	1	2	3	4	5	6	7	8	9	10
Динамика продаж	139	101	121	145	150	148	157	164	160	168

Определите оптимальный тренд и рассчитайте точечный прогноз продаж на последующие 5 лет.

14. Данные о состоянии уровня заболеваемости нашего города в процентах за последние 15 месяцев представлены в следующей таблице:

Месяц	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Уровень заболеваемости	56	62	63	60	59	65	75	81	15	103	103	115	114	118	124

Постройте оптимальный тренд и на его основе получите прогноз изменения заболеваемости в процентах за 16-й и 17-й месяцы.

15. Финансовый директор фирмы рассматривает целесообразность ежемесячного финансирования инвестиционного проекта со следующими объемами нетто-платежей y_t, тыс. руб.:

t	y_t
1	45
2	40
3	43
4	48
5	42
6	47
7	51
8	55
9	50
10	57
11	62
12	62

Определите наилучший тренд и прогнозные значения данного показателя на следующие 3 месяца.

16. Данные, которые характеризуют динамику спроса на товары потребления за 15 месяцев, представлены в следующей таблице:

№ п/п	y	№ п/п	y	№ п/п	y
1	18,3	6	26,3	11	37,8
2	19,4	7	29,4	12	39,3
3	21,6	8	31,8	13	41,8
4	23,4	9	33,4	14	44,3
5	25,8	10	35,6	15	47,4

Определите наилучший тренд и прогнозные значения данного показателя на следующие 3 месяца.

17. Три эксперта фирмы, рассмотрев шесть возможных инвестиционных проектов, выставили следующие балльные оценки:

Эксперт	Инвестиционный проект					
	1	2	3	4	5	6
1	7	3	5	8	6	9
2	5	4	6	4	7	6
3	4	5	6	4	7	6

Определите лучший проект с помощью метода среднеарифметических рангов и метода медиан рангов.

18. Пять экспертов оценили результаты работы шести объектов.

Исходные данные представлены в следующей таблице:

Объект	Эксперт				
	1	2	3	4	5
1	1	2	1,5	1	2
2	2,5	2	1,5	2,5	1
3	2,5	2	3	2,5	3
4	4	5	4,5	4,5	4
5	5	4	4,5	4,5	5,5
6	6	6	6	6	5,5

С помощью коэффициента конкордации оцените степень согласованности оценок, принятых экспертами.

19. Десять участниц конкурса красоты были ранжированы тремя судьями.

Получены следующие результаты:

Судья	Ранг девушки									
	1	2	3	4	5	6	7	8	9	10
1	1	6	5	10	3	2	4	9	7	8
2	3	5	8	4	7	10	2	1	6	9
3	6	4	9	8	1	2	3	10	5	7

С помощью коэффициента конкордации оцените степень согласованности оценок, принятых судьями.

Библиографический список

Основная литература

1. *Кремер, Н. Ш.* Эконометрика : учеб. для вузов / Н. Ш. Кремер, Б. А. Путко. — Москва : ЮНИТИ, 2008. — 310 с.

2. *Осипов, А. Л.* Электрометрика : учеб. пособие / А. Л. Осипов, В. Н. Храпов ; СибАГС. — Новосибирск : Изд-во СибАГС, 2004. — 276 с.

3. *Практикум* по эконометрике : учеб. пособие / И. И. Елисеева, С. В. Курышева, Н. М. Горденко и др. ; под ред. И. И. Елисеевой. — Москва : Финансы и статистика, 2001. — 192 с.

4. *Тюрин, Ю. Н.* Анализ данных на компьютере : учеб. пособие / Ю. Н. Тюрин, А. А. Макаров. — Москва : ФОРУМ, 2008. — 366 с.

5. *Шикин, Е. В.* Математические методы и модели в управлении : учеб. пособие / Е. В. Шикин, А. Г. Чхартишвили. — 2-е изд., испр. — Москва : Дело, 2002. — С. 65—79.

6. *Эконометрика* : учеб. для вузов / [И. И. Елисеева и др.] ; под ред. И. И. Елисеевой. — Москва : Финансы и статистика, 2007. — 575 с.

Дополнительная литература

1. *Большаков, А. А.* Методы обработки многомерных данных и временных рядов : учеб. пособие / А. А. Большаков, Р. Н. Каримов. — Москва : Горячая линия — Телеком, 2007. — 522 с.

2. *Бычков, А. Г.* Сборник задач по теории вероятностей, математической статистике и методам оптимизации : учеб. пособие / А. Г. Бычков. — Москва : ФОРУМ, 2008. — 224 с.

3. *Розен, В. В.* Математические модели принятия решений в экономике : учеб. пособие / В. В. Розен. — Москва : Высш. шк., 2002. — 288 с.

4. Статистика : учеб. / В. Г. Минашкин [и др.] ; под ред. В. Г. Минашкина. — Москва : ТК Велби : Проспект, 2006. — 272 с.

5. *Тимофеев, В. С.* Эконометрика : учеб. / В. С. Тимофеев, А. В. Фаддеенков, В. Ю. Щеколдин. — Новосибирск : Изд-во НГТУ, 2013. — 340 с.

6. *Методы* и модели анализа временных рядов : метод. указания к лаб. работам / сост. С. И. Татаренко. — Тамбов : Изд-во Тамб. гос. техн. ун-та, 2008. — 32 с.

Глава 3. Многокритериальные методы

§ 1. Анализ Парето

В большинстве практических задач принятия решения (ЗПР) исходы, в качестве которых выступают реальные объекты и явления, оцениваются, как правило, не по одному, а по нескольким критериям. Например, при выборе кандидата на должность важнейшими критериями оценки являются квалификация, образование, эрудиция, возраст, коммуникабельность и т. п. В экономических задачах основными критериями служат экономическая эффективность и стоимость. При этом каждый из критериев, в свою очередь, может быть подразделен на более частные критерии.

Если исходы оцениваются по $m > 1$ критериям, то такая ***задача принятия решений называется многокритериальной***.

Основная сложность логического анализа многокритериальных задач состоит в том, что в них в отличие от однокритериальных задач появляется *эффект несравнимости исходов*. Например, если исходы оцениваются по двум критериям, несводимым один к другому, и исход первый лучше исхода второго по первому критерию, но хуже по второму, то эти исходы будут несравнимыми между собой. Выбор между несравнимыми исходами является сложной проблемой и составляет основное содержание многокритериальной оптимизации.

Математическая модель ЗПР при многих критериях может быть представлена в виде

$$\langle D; f_1, f_2, \ldots, f_m \rangle,$$

где D — множество допустимых исходов;

f_j — числовая функция, заданная на множестве D; при этом $f_j(a)$ есть оценка исхода $a \in D$ по j-му критерию.

Такая модель соответствует задаче принятия решения в условиях определенности, в которой множество альтернатив отождествляется с множеством допустимых исходов, а оценочная структура задается вектором $(f_1, f_2, \ldots, f_m)$.

Критерий f_j называется *позитивным*, если лицо, принимающее решение (ЛПР), стремится к его увеличению, и *негативным*, если стремится к его уменьшению. При рассмотрении многокритериальных ЗПР в общем виде будем, если не оговорено противное, предполагать, что все имеющиеся критерии являются позитивными. В многокритериальной ЗПР с позитивными критериями цель принимающего решения — получение исхода, имеющего как можно более высокие оценки по каждому критерию.

Пусть Y_j — множество значений функции f_j, т. е. множество всех оценок по j-му критерию. Тогда множество $Y = Y_1 \times Y_2 \times \ldots \times Y_m$, состоящее из всевозможных упорядоченных наборов по критериям 1, 2, …, m, называется множеством векторных оценок. Любой элемент $y \in Y$ представляет собой вектор $y = (y_1, y_2, \ldots, y_m)$, где

$y_j \in Y_j$. Для всякого исхода $a \in D$ набор его оценок по всем критериям, т. е. набор $\left(f_1(a), f_2(a), \ldots, f_m(a)\right)$, есть *векторная оценка исхода* a. Векторная оценка исхода содержит полную информацию о ценности (полезности) этого исхода для принимающего решение, и сравнение любых двух исходов заменяется сравнением их векторных оценок. Основное отношение, по которому производится сравнение векторных оценок, а значит, и сравнение исходов, — это ***отношение доминирования по Парето***.

Говорят, что векторная оценка $y = (y_1, y_2, \ldots, y_m)$ доминирует по Парето *векторную оценку* $y' = (y'_1, y'_2, \ldots, y'_m)$, если для всех $j = 1, 2, \ldots, m$ выполняется неравенство $y_j \geq y'_j$. Причем, по крайней мере, для одного индекса $j = 1, 2, \ldots, m$ неравенство должно быть строгим.

Пусть $Q \subseteq Y$ — некоторое множество векторных оценок. ***Векторная оценка*** $y^* \in Q$ называется ***Парето-оптимальной*** в Q, если она является максимальным элементом множества Q относительно Парето-доминирования, т. е. если в множестве Q не существует такой векторной оценки y, которая доминирует по Парето векторную оценку y^*.

Говорят, что исход a_1 доминирует по Парето ***исход*** a_2, если векторная оценка исхода a_1 доминирует по Парето векторную оценку исхода a_2.

Содержательно это означает, что исход a_1 не хуже, чем исход a_2 по любому из рассматриваемых критериев. Причем, по крайней мере, по одному из этих критериев a_1 лучше, чем a_2. В этом случае ЛПР, безусловно, отдаст предпочтение исходу a_1.

Исход $a^* \in D$ называется ***Парето-оптимальным*** исходом в множестве D, если он не доминируется по Парето никаким другим исходом из множества D, т. е. если векторная оценка исхода a^* является Парето-оптимальной в множестве векторных оценок $Q = \left\{\left(f_1(a), f_2(a), \ldots, f_m(a)\right) : a \in D\right\}$.

Парето-оптимальность исхода a^* означает, что он не может быть улучшен ни по одному из критериев без ухудшения по какому-нибудь другому критерию. Итак, кандидатом на оптимальное решение многокритериальной ЗПР может являться только Парето-оптимальный исход. Однако Парето-оптимальных исходов может быть несколько, а в случае непрерывности — бесконечное множество. Дать однозначный ответ на вопрос, какой же из Парето-оптимальных исходов следует считать оптимальным для общего случая, не имея дополнительной информации о критериях, невозможно. Дело в том, что любые два Парето-оптимальных исхода несравнимы относительно доминирования по Парето. Для любых двух Парето-оптимальных исходов a_1 и a_2 всегда найдутся такие два критерия j_1 и j_2, что a_1 лучше a_2 по критерию j_1, но хуже по критерию j_2. Если нет информации об относительной важности критериев j_1 и j_2, то рациональный выбор между a_1 и a_2 сделать невозможно. Отметим, что нельзя сделать рационального выбора и в такой ситуации, когда, например, имеется всего 10 критериев, причем a_1 лучше a_2 по одному критерию, но хуже по девяти остальным: понятно, что в некоторых реальных случаях превосходство по одному критерию может перевесить превосходство по всем остальным.

Общая методика исследования задач принятия решений на основе математического моделирования для многокритериальных ЗПР может быть реализована в рамках одного из следующих **подходов**.

Первый подход. Для заданной многокритериальной ЗПР находится множество ее Парето-оптимальных исходов, а выбор конкретного оптимального исхода из множества Парето-оптимальных предоставляется ЛПР.

Второй подход. Производится сужение множества Парето-оптимальных исходов с помощью некоторых формализованных процедур, что облегчает окончательный выбор исхода для ЛПР. Отметим, что такое сужение может быть произведено только при наличии дополнительной информации о критериях или о свойствах оптимального решения.

Рассмотрим некоторые *простейшие способы сужения Парето-оптимального множества*. При этом внимание следует акцентировать на необходимой дополнительной информации.

1) Указание нижних границ критериев. Дополнительная информация об оптимальном исходе $a^* \in D$ в этом случае имеет следующий вид:

$$f_j(a^*) \geq \gamma_j,\ j = 1, 2,\ \ldots,\ m.$$

Число γ_j рассматривается здесь как нижняя граница по j-му критерию. Набор оценок $(\gamma_1, \gamma_2,\ \ldots,\ \gamma_m)$ представляет собой дополнительную информацию, полученную от ЛПР. При указании нижних границ критериев оптимальным может считаться только такой Парето-оптимальный исход, для которого оценка по каждому из критериев j не ниже назначенной γ_j, что и дает сужение Парето-оптимального множества. Ясно, что при увеличении значений γ_j Парето-оптимальное множество сокращается. Окончательный выбор Парето-оптимального исхода производится из суженного Парето-оптимального множества ЛПР на основе субъективных соображений. Основной недостаток метода состоит в том, что оптимальное решение становится в данном случае субъективным, так как зависит, во-первых, от величин назначаемых нижних границ критериев, во-вторых, от окончательного выбора, совершаемого ЛПР.

2) Субоптимизация. В этом случае выделяют один из критериев, а по всем остальным критериям назначают нижние границы. Оптимальным при этом считается исход, максимизирующий выделенный критерий на множестве исходов, оценки которых по остальным критериям не ниже назначенных. Пусть, например, f_1 — выделенный критерий; γ_j — нижняя граница для j-го критерия, где $j = 2,\ \ldots,\ m.$ Тогда оптимальным считается тот исход $a^* \in D,$ на котором достигает максимума функция f_1, рассматриваемая на множестве

$$D_1 = \{a \in D\colon f_j(a^*) \geq \gamma_j,\ j = 2,\ \ldots,\ m\}.$$

С помощью метода субоптимизации задача многокритериальной оптимизации превращается в задачу скалярной оптимизации на суженном допустимом множестве. Выделение одного из критериев, а также указание нижних границ для остальных критериев основано на дополнительной информации, получаемой от ЛПР. Следовательно, окончательное решение также имеет субъективный характер.

3) Лексикографическая оптимизация. Она основана на упорядочении критериев по их относительной важности. После этого процедуру нахождения оптимального решения проводят следующим образом. На первом шаге отбирают исходы, которые имеют максимальную оценку по важнейшему критерию. Если такой исход единственный, то его и считают оптимальным. Если же таких исходов несколько, то среди них отбирают те, которые имеют максимальную оценку по следующему за важнейшим критерию и т. д. В результате такой процедуры всегда остается (по крайней мере, в случае конечного множества исходов) единственный исход — он и будет оптимальным.

Основными недостатками метода лексикографической оптимизации являются следующие: при практическом применении данного метода возникают содержательные трудности в установлении полной упорядоченности критериев по их относительной важности; фактически при использовании этого метода принимается во внимание только первый — важнейший — критерий, а следующий за ним по важности критерий учитывается только тогда, когда важнейший критерий достигает максимума на нескольких исходах.

Пример 1. Предположим, что предстоит выбрать место работы из девяти вариантов. В качестве основных критериев для выбора работы были взяты заработная плата (З), длительность отпуска (Д), время поездки на работу (В).

Вариант	Критерий		
	заработная плата, руб.	длительность отпуска, дни	время поездки, мин
1	900	20	– 60
2	500	30	– 20
3	700	36	– 40
4	800	40	– 50
5	400	60	– 15
6	600	30	– 10
7	900	35	– 60
8	600	24	– 10
9	650	35	– 40

Так как критерий В имеет характер потерь, оценки по этому критерию берутся со знаком минус.

Какой из девяти вариантов является оптимальным?

Выделим вначале Парето-оптимальные варианты (используя первый подход). Здесь вариант 3 доминирует по Парето вариант 9, вариант 6 доминирует по Парето вариант 2, вариант 6 доминирует по Парето вариант 8 и вариант 7 доминирует по Парето вариант 1. Отбрасывая доминируемые по Парето варианты 1, 2, 8, 9, получим Парето-оптимальное множество 3, 4, 5, 6, 7.

При отсутствии информации об относительной важности рассматриваемых критериев, а также о каких-либо дополнительных свойствах оптимального решения дальнейшее сужение Парето-оптимального множества произвести нельзя. Тогда формальный анализ заканчивается указанием Парето-оптимального множества, и окончательный выбор оптимального варианта производится ЛПР из этих вариантов на основе каких-то дополнительных соображений.

Рассмотрим теперь второй подход, который приводит к сужению Парето-оптимального множества на основе дополнительной информации, получаемой от ЛПР.

Указание нижних границ критериев. Наложим, например, следующие ограничения на оптимальное решение: заработная плата — не менее 600 руб.; длительность отпуска — не менее 30 дней; время поездки — не более 40 мин. Варианты, удовлетворяющие этим дополнительным ограничениям, — 3, 6, 9. Из них оптимальными по Парето являются варианты 3, 6, остается сделать окончательный выбор между ними.

Субоптимизация. Пусть в качестве выделенного критерия выступает критерий З; ограничения: длительность отпуска — не менее 30 дней, время поездки — не более 40 мин. Отбросим варианты, которые не удовлетворяют данным ограничениям. Остаются варианты 2, 3, 5, 6, 9. Из них максимальную заработную плату имеет вариант 3. Этот вариант и будет оптимальным.

Лексикографическая оптимизация. Упорядочим критерии по относительной важности, например: З > В > Д (важнейший критерий — заработная плата, следующий за ним по важности — время поездки, наименее важный критерий — длительность отпуска). Максимальное значение по критерию З имеют варианты 1 и 7. Далее сравним эти варианты по второму по важности критерию — В. Так как время поездки для этих вариантов одинаково, переходим к третьему критерию — Д. По критерию Д лучшим является вариант 7, который, соответственно, и является оптимальным. При упорядочении В > Д > З оптимальным будет вариант 6, а при упорядочении Д > З > В оптимальным становится вариант 5. В этом примере наглядно проявляется недостаток лексикографической оптимизации — фактический учет одного важнейшего критерия. Например, в последнем случае (Д > З > В) в качестве оптимального выступает вариант 5, который имеет самую низкую оценку по критерию З.

§ 2. Обобщенный критерий в задачах принятия решения

Под ***построением обобщенного критерия в многокритериальной*** ЗПР понимается процедура, которая синтезирует набор оценок по заданным критериям в единую численную оценку, выражающую итоговую полезность этого набора оценок для ЛПР. Формально обобщенный критерий для ЗПР вида $\langle D; f_1, f_2, \ldots, f_m \rangle$ представляет собой функцию

$$б : Y_1 \times Y_2 \times \cdots \times Y_m \to R,$$

где Y_j — множество оценок по j-му критерию.

Если обобщенный критерий φ построен, то для каждого допустимого исхода $a \in D$ может быть найдена численная оценка его полезности: $f(a) = φ\ (f_1(a),\ f_2(a),\ \ldots,\ f_m(a))$. Таким образом, задание обобщенного критерия сводит задачу многокритериальной оптимизации к задаче однокритериальной оптимизации с целевой функцией φ.

Наиболее распространенным обобщенным критерием является взвешенная сумма частных критериев, которая превращает векторную оценку $y = (y_1, y_2,\ \ldots,\ y_m)$ в скалярную:

$$φ(y) = б_1 y_1 + \cdots + б_m y_m,$$

где $б_j \geq 0,\ \ j = 1, 2, ..., m$ и $\sum_{j=1}^{m} б_j = 1.$

Числа $б_j$ в этом случае называют *весовыми коэффициентами*. При этом весовой коэффициент $б_j$ интерпретируется как показатель относительной важности j-го критерия: чем больше $б_j$, тем больший вклад дает оценка по j-му критерию в итоговую оценку φ (y).

Справедливы следующие *правила*.

Правило 1. Пусть $Q \subseteq Y$ — произвольное множество векторных оценок. Если векторная оценка $y^* = (y_1^*, y_2^*,\ \ldots,\ y_m^*)$ доставляет максимум функции $φ(y) = б_1 y_1 + \cdots + б_m y_m$, где все $б_j > 0$, то векторная оценка y^* является Парето-оптимальной в множестве Q.

Правило 2. Пусть $Q \subseteq Y$ — выпуклое множество, $y^* \in Q$ — Парето-оптимальная векторная оценка на множестве Q. Тогда найдутся такие неотрицательные числа $б_j \geq 0,\ \ j = 1, 2,\ \ldots,\ m,$ что функция $φ(y) = б_1 y_1 + \cdots + б_m y_m$ достигает максимума на множестве Q в точке y^*.

Правила 1 и 2 указывают способ перебора Парето-оптимальных точек заданного множества Q: зафиксировав положительный вектор весов $б = (б_1,\ \ldots,\ б_m)$ и найдя максимум взвешенной суммы $φ(y) = \sum_{j=1}^{m} б_j y_j$, получим некоторую точку Парето-оптимального множества; в случае выпуклого множества Q все Парето-оптимальные точки множества могут быть получены таким способом при некоторых $б_j \geq 0$, $j = 1, 2,\ \ldots,\ m.$

Обратимся теперь к проблеме построения обобщенного критерия в виде взвешенной суммы частных критериев $\varphi(y) = \sum_{j=1}^{m} б_j y_j$. Предложено множество различных способов нахождения весовых коэффициентов, однако ни один из них не может претендовать на роль универсального.

Рассмотрим в качестве примера следующий способ нахождения весовых коэффициентов.

$$б_j = \frac{1}{M_j},$$

где $M_j = \max_{a \in D} \left| f_j(a) \right|$.

В этом случае $f(a) = \sum_{j=1}^{m} \frac{f_j(a)}{M_j}$, т. е. итоговой численной оценкой исхода a является сумма нормализованных оценок по всем критериям. На первый взгляд обобщенный критерий $f(a) = \sum_{j=1}^{m} \frac{f_j(a)}{M_j}$ представляется вполне разумным, однако следующий пример выявляет один его существенный недостаток.

Пример 2. Предположим, что требуется сравнить два альтернативных варианта места работы A и B. Их векторные оценки приведены в следующей таблице:

Вариант	Заработная плата, руб.	Длительность отпуска, дни	Время поездки, мин
A	900	20	–60
B	500	30	–40

Здесь $M_1 = 900$, $M_2 = 30$, $M_3 = 60$, откуда $f(A) = \frac{900}{900} + \frac{20}{30} - \frac{60}{60} = \frac{2}{3}$, а $f(B) = \frac{500}{900} + \frac{30}{30} - \frac{40}{60} = \frac{8}{9}$. Так как $f(B) > f(A)$, то альтернатива B более предпочтительна, чем альтернатива A. Пусть теперь наряду с альтернативами A и B появилась еще одна альтернатива — C, которая характеризуется векторной оценкой вида $(400, 60, -100)$. В этом случае $M_1' = 900$, $M_2' = 60$, $M_3' = 100$. Откуда $f(A) = \frac{900}{900} + \frac{20}{60} - \frac{60}{100} = \frac{22}{30}$, $f(B) = \frac{500}{900} + \frac{30}{60} - \frac{40}{100} = \frac{59}{90}$, а $f(C) = \frac{400}{900} + \frac{60}{60} - \frac{100}{100} = \frac{4}{9}$.

В этом случае альтернатива A стала предпочтительнее, чем альтернатива B, т. е. порядок предпочтения альтернатив A и B получился обратным. Итак, наличие еще одной альтернативы C меняет предпочтения между альтернативами A и B. Это парадоксальное свойство называется нарушением независимости предпочтений относительно посторонних альтернатив. Следует заметить, что дополнительная альтернатива C здесь не конкурирует ни с альтернативой A, ни с альтернативой B.

Контрольные вопросы и задания

1. Какие задачи принятия решений относятся к многокритериальным?
2. Определите позитивные и негативные критерии.
3. В чем заключается смысл отношения доминирования по Парето?
4. Какие оценки называются Парето-оптимальными?
5. Что означают субоптимизация и лексикографическая оптимизация?
6. Определите понятие обобщенного критерия в задачах принятия решений.

Задания для самостоятельного решения

1. При выборе квартиры в качестве существенных критериев взяты: p_1 — метраж (м^2), p_2 — время поездки на работу (мин), p_3 — время поездки в зону отдыха (мин). При этом критерий p_1 рассматривается как позитивный, а критерии p_2 и p_3 — как негативные.

Сравните по предпочтительности семь вариантов, представленных в следующей таблице:

Вариант	p_1	p_2	p_3
1	60	50	30
2	50	45	25
3	45	30	20
4	60	40	30
5	42	20	10
6	45	30	15
7	48	45	25

2. Используя в качестве обобщенного критерия для предыдущего задания критерий $\varphi(y) = \sum_{j=1}^{m} б_j y_j$, постройте полное ранжирование вариантов места работы.

Соответствует ли полученное ранжирование вашим предпочтениям?

Библиографический список

Основная литература

1. *Осипов, А. Л.* Экономико-математические методы в управлении : учеб.-метод. комплекс для дистанц. обучения / А. Л. Осипов, Е. А. Рапоцевич ; СибАГС. — Новосибирск : Изд-во СибАГС, 2006. — 144 с.

2. *Тюрин, Ю. Н.* Анализ данных на компьютере : учеб. пособие / Ю. Н. Тюрин, А. А. Макаров. — Москва : ФОРУМ, 2008. — 366 с.

3. *Шикин, Е. В.* Математические методы и модели в управлении : учеб. пособие / Е. В. Шикин, А. Г. Чхартишвили. — 2-е изд., испр. — Москва : Дело, 2002. — С. 65—79.

Дополнительная литература

1. *Розен, В. В.* Математические модели принятия решений в экономике : учеб. пособие / В. В. Розен. — Москва : Высш. шк., 2002. — 288 с.

2. *Экономико*-математические методы в управлении : практикум / СибАГС ; сост. : А. Л. Осипов, Е. А. Рапоцевич. — Новосибирск : Изд-во СибАГС, 2007. — 160 с.

Глава 4. Иерархии и приоритеты

§ 1. Приоритеты

Проблема сравнения возникает и при измерении физических величин, и при оценке совершенных поступков. Для получения хороших результатов в сравнениях требуется правильно выбрать шкалу сравнения и адекватно определить степень согласованности наших суждений.

Попарные сравнения позволяют повысить согласованность оценок. Количественные оценки, получаемые при парных сравнениях, опираются на выбор подходящей шкалы. Шкала должна быть проста и естественна. Мы будем использовать шкалу от 1 до 9. Этот выбор основывается на способности человека производить качественные разграничения, хорошо представленные пятью определениями: слабый, равный, сильный, очень сильный, абсолютный. Для большей точности пользуются промежуточными определениями. На этих принципах построим шкалу отношений (табл. 5).

Таблица 5

Степень значимости	Определение	Объяснение
1	Одинаковая значимость	Два действия вносят одинаковый вклад в достижение цели
3	Некоторое преобладание значимости одного действия над другим (слабая значимость)	Существуют соображения в пользу предпочтения одного из действий. Однако эти соображения недостаточно убедительны
5	Существенная, или сильная, значимость	Имеются надежные данные или логические суждения для того, чтобы показать предпочтительность одного из действий
7	Очевидная, или очень сильная, значимость	Убедительное свидетельство в пользу одного действия перед другим
9	Абсолютная значимость	Свидетельства в пользу предпочтения одного действия другому в высшей сте-

		пени убедительны
2, 4, 6, 8	Промежуточные значения между двумя соседними суждениями	Ситуации, когда необходимо компромиссное решение

Введем далее несколько определений.

Пусть A — квадратная положительная матрица порядка n

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Матрица A называется обратно-симметричной, если для любых i и k выполняется соотношение

$$a_{ki} = \frac{1}{a_{ik}}.$$

Матрица A называется согласованной, если для любых i, k, j имеет место равенство

$$a_{ik} \cdot a_{kj} = a_{ij}.$$

Положительная обратно-симметричная матрица является согласованной тогда и только тогда, когда порядок матрицы и ее наибольшее собственное значение совпадают:

$$л_{\max} = n.$$

Пример 1. Пусть заданы три объекта, которым поставлены в соответствие оценки некоторой характеристики: $щ_1 = 2,\ щ_2 = 3,\ щ_3 = 5.$

Составим матрицу, элементы которой вычислим по формуле $a_{ij} = \frac{щ_i}{щ_j}$. Такая матрица называется идеальной матрицей сравнения. В результате получим

$$A = \begin{pmatrix} \frac{2}{2} & \frac{2}{3} & \frac{2}{5} \\ \frac{3}{2} & \frac{3}{3} & \frac{3}{5} \\ \frac{5}{2} & \frac{5}{3} & \frac{5}{5} \end{pmatrix} = \begin{pmatrix} 1 & \frac{2}{3} & \frac{2}{5} \\ \frac{3}{2} & 1 & \frac{3}{5} \\ \frac{5}{2} & \frac{5}{3} & 1 \end{pmatrix}.$$

Причем $a_{ik}a_{kj} = \frac{\text{щ}_i}{\text{щ}_k}\frac{\text{щ}_k}{\text{щ}_j} = \frac{\text{щ}_i}{\text{щ}_j} = a_{ij}$. Следовательно, полученная матрица является обратно-симметричной и согласованной.

Если взять произвольную положительную обратно-симметричную матрицу, то она не обязательно будет согласованной. Для оценки степени ее отклонения от согласованной используется *индекс однородности Т. Саати*, который вычисляется по следующей формуле:

$$\text{ИО} = \frac{\text{л}_{\max} - n}{n-1}$$

и является показателем близости этой матрицы к согласованной.

Довольно естественно встает вопрос о том, как находить наибольшее собственное значение $\text{л}_{\max}$.

Опишем способ приближенного вычисления собственного столбца или собственного вектора.

1. Перемножим элементы каждой строки и запишем полученные результаты в столбец.
2. Извлечем корень n-й степени из каждого элемента найденного столбца.
3. Сложим элементы этого столбца.
4. Разделим каждый из этих элементов на полученную сумму.

Найдем собственное значение, соответствующее собственному столбцу x. Для этого выполним следующие этапы:

— умножим исходную матрицу на этот столбец, получим $Ax = y$;

— разделим элементы полученного столбца y на соответствующие элементы столбца x: $\frac{y_1}{x_1}, \frac{y_2}{x_2}, \cdots, \frac{y_n}{x_n}$, и если $\frac{y_1}{x_1} = \frac{y_2}{x_2} = \cdots = \frac{y_n}{x_n}$, то это отношение и есть собственное значение матрицы A, отвечающее данному столбцу x. Если же хотя бы одно из равенств нарушается, то столбец x не является собственным столбцом матрицы A.

В этом случае в качестве приближенного собственного значения выбирается

$$\lambda_{\max} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i}.$$

§ 2. Метод анализа иерархий

Метод анализа иерархий (***МАИ***) вырос в настоящее время в обширный междисциплинарный раздел науки, имеющий строгие математические и психологические обоснования и многочисленные приложения.

Основное применение метода — поддержка принятия решений посредством иерархической композиции задачи и рейтингования альтернативных решений. Метод позволяет учесть это обстоятельство с помощью построения дополнительной модели для согласования различных мнений посредством определения их приоритетов.

Перечислим возможности метода: он позволяет организовать обсуждение проблемы, провести сбор данных по проблеме; проанализировать проблему и минимизировать ее; оценить важность учета каждого решения и каждого фактора, влияющего на приоритеты решений; оценить устойчивость принимаемого решения; провести синтез проблемы принятия решения.

Преимуществом этого метода выступает то, что он позволяет учитывать человеческий фактор при подготовке принятия решения.

МАИ является системной процедурой для иерархического представления элементов, определяющих суть любой проблемы. Решение проблемы есть процесс поэтапного установления приоритетов, который включает:

1) определение и выделение проблемы;
2) декомпозицию проблемы в иерархию;
3) построение матриц парных сравнений;
4) вычисление приоритетов, наибольшего собственного значения матрицы суждений, индекса согласованности и отношения согласованности;
5) вычисление глобальных приоритетов.

Система поддержки и принятия решений, основанная на МАИ, является простым и удобным средством, которое помогает структурировать проблему, строить набор альтернатив, выделять характеризующие их факторы, задавать значимость этих факторов, оценивать альтернативы по каждому из них, находить неточности и противоречия в суждениях лица, принимающего решения, или эксперта, ранжировать альтернативы, проводить анализ решения и обосновывать полученные результаты.

МАИ может использоваться при решении таких типовых задач, как:

— оценка качества организационных, проектных и конструкторских решений;
— определение политики инвестиций в различных областях;
— размещение объектов (выбор места расположения вредных и опасных производств, пунктов обслуживания);
— распределение ресурсов;
— проведение анализа проблемы по критерию «стоимость-эффективность»;
— стратегическое планирование;
— проектирование и выбор оборудования, товаров;
— выбор профессии, места работы, подбор кадров и т. д.

МАИ используется для решения слабоструктурированных и неструктурированных проблем. Методология решения таких проблем опирается на системный подход, при котором проблема рассматривается как результат взаимодействия и, соответственно, взаимозависимости множества разнородных объектов, а не просто как их изолированная и автономная совокупность.

Человеку присущи два характерных признака аналитического мышления: умение наблюдать и анализировать наблюдения; способность устанавливать отношения между наблюдениями, оценивать уровень взаимосвязей, а затем синтезировать эти отношения в общее восприятие наблюдаемого. На основе этих свойств человеческого мышления были сформулированы *три принципа*, реализация которых и является содержанием МАИ:

— принцип идентичности и декомпозиции;

— дискриминации и сравнительных суждений;

— синтеза.

Принимая решение, группа экспертов производит декомпозицию сложной проблемы, т. е. определяет ее компоненты и отношения между ними. Получается модель реальной действительности, построенная в виде *иерархии*. Вершина иерархии — общая цель, далее располагаются подцели, затем силы, которые влияют на эти подцели, люди, их цели, политики, стратегии и, наконец, исходы, являющиеся результатами стратегий.

На следующем этапе решения сравниваются уже отдельные компоненты иерархии между собой. В результате может быть выражена относительная степень интенсивности взаимодействия элементов в иерархии. Затем эти суждения выражаются численно.

Далее проводятся процедуры синтеза множественных суждений, *приоритетности критериев* и нахождения альтернативных решений.

Таким образом, можно представить следующие *этапы принятия решения* с помощью *МАИ*:

— построение иерархии рассматриваемой проблемы;

— парное сравнение компонент иерархии;

— математическая обработка полученных суждений.

Опишем *методологию МАИ*.

Она предполагает рассмотрение иерархий с одинаковым числом и функциональным составом альтернатив, расположенных под критериями, и применение метода попарного сравнения элементов иерархии. Построение иерархии начинается с очерчивания проблемы исследования. Далее строится собственно иерархия с целью, расположенной в ее вершине, промежуточными уровнями (например, критериями) и альтернативами, формирующими самый нижний иерархический уровень.

При принятии решений группа экспертов использует шкалу отношений и матрицу парных сравнений.

Допустим, что самая простая иерархия состоит из цели, уровня критериев и уровня альтернатив. На следующем шаге мы должны попарно сравнить между собой все критерии, используя приведенную шкалу (см. табл. 5), и результаты записать в соответствующую матрицу парных сравнений. Полученные суждения выражаются в целых числах с учетом девятибалльной шкалы.

Заполнение квадратных матриц парных сравнений осуществим по следующему правилу. Если критерий K_1 доминирует над критерием K_2, то клетка матрицы, соответствующая строке K_1 и столбцу K_2 заполняется целым числом, а клетка, соответствующая строке K_2 и столбцу K_1, заполняется обратным к нему числом. Если элемент K_2 доминирует над K_1, то целое число ставится в клетку, соответствующую строке K_2 и столбцу K_1, а дробь проставляется в клетку, соответствующую строке K_1 и столбцу K_2. Если критерии K_1 и K_2 равнопредпочтительны, то в обе позиции матрицы ставятся единицы.

Далее составим матрицы парных сравнений альтернатив по отношению к каждому критерию.

При проведении попарных сравнений следует ответить на следующие вопросы: какой из двух сравниваемых элементов важнее или имеет большее воздействие, какой более вероятен и какой предпочтительнее. При сравнении критериев обычно спрашивают, какой из критериев более важен; при сравнении альтернатив по отношению к критерию — какая из альтернатив более предпочтительна или более вероятна.

В практических задачах количественная (кардинальная) и транзитивная (порядковая) однородность (согласованность) нарушается, поскольку человеческие ощущения нельзя выразить точной формулой. Для оценки качества проведенных парных сравнений служит отношение однородности $\text{ОО} = \dfrac{\text{ИО}}{\text{М(ИО)}}$, где М(ИО) — среднее значение индекса однородности случайным образом составленной матрицы парных сравнений, которое основано на представленном экспериментальном материале:

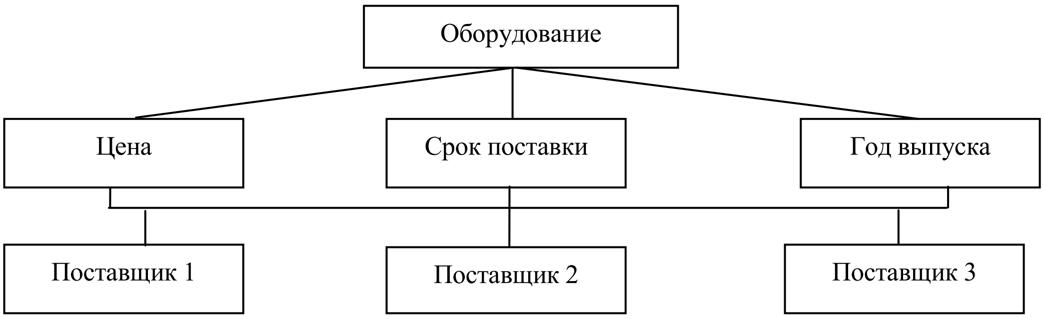
Порядок матрицы (n)	М (ИО)
1	0,00
2	0,00
3	0,58
4	0,90
5	1,12
6	1,24
7	1,32
8	1,41
9	1,45
10	1,49
11	1,51
12	1,48
13	1,56
14	1,57
15	1,59

В качестве допустимого используется значение $\text{ОО} \le 0{,}1$. Если для матрицы парных сравнений отношение однородности $\text{ОО} > 0{,}1$, то это свидетельствует о существенном нарушении логики суждений, допущенном экспертом при заполнении матрицы, поэтому эксперту предлагается пересмотреть данные, использованные для построения матрицы, чтобы улучшить однородность.

Иерархический синтез используется для взвешивания собственных векторов матриц парных сравнений альтернатив весами критериев (элементов), имеющихся в иерархии, а также для вычисления суммы по всем соответствующим взвешенным компонентам собственных векторов нижележащего уровня иерархии.

Пример 2. Из трех поставщиков оборудования требуется выбрать лучшего. Оценим поставщиков по трем критериям: цена оборудования, срок поставки, год выпуска оборудования.

Декомпозиция проблемы имеет следующий вид: на первом уровне находится цель (оборудование), на втором — три фактора, уточняющих цель, и, наконец, на последнем уровне — три кандидата, которые должны быть оценены по отношению к критериям второго уровня:



Матрица парных сравнений для оценки критериев выглядит следующим образом:

Критерий	1	2	3	Произведение	Корень	КОВ
Цена	1	3	3	9	2,080	0,594
Срок поставки	0,333333	1	2	0,6666667	0,874	0,249
Год выпуска	0,333333	0,5	1	0,1666667	0,550	0,157
Сумма по столбцам	1,666667	4,5	6	9,8333333	3,504	
Произведение сумм на КОВ	0,99	1,12	0,94			
л$_{max}$	3,05					
М (ИО)	0,58					
ИО	0,026804					
ОО	0,046215					

Векторы приоритетов, или КОВ, вычислим следующим приближенным способом. Сначала вычислим столбец из 3 чисел, каждое из которых равняется произведению элементов соответствующей строки матрицы критериев. Затем вычислим столбец чисел, каждое из которых равняется корню 3-й степени из элементов предыдущего столбца: в общем случае для элемента i-й строки $b_i = \sqrt[n]{a_{i1}a_{i2}\cdots\ a_{in}}$, где n — количество критериев. После этого для всех n строк $(b_1,\ b_2,\ \ldots,\ b_n)$ произведем их нормализацию путем деления всех элементов на $\sum_{i=1}^{n} b_i$.

Таким образом, получим нормализованный собственный вектор, который является вектором приоритетов, или КОВ.

ИО вычислим по схеме:

1. Суммируем каждый столбец суждений.
2. Сумму первого столбца умножим на величину первой компоненты нормализованного вектора приоритетов (КОВ), сумму второго столбца — на вторую компоненту и т. д.
3. Полученные числа суммируем. Их сумму обозначим $л_{max}$.
4. $ИО = \frac{(л_{max} - n)}{(n-1)}$, где n — число сравниваемых критериев. В нашем примере n = 3.

Согласно формуле отношение согласованности $ОО = \frac{ИО}{М(ИО)}$. Как видим, ОО составляет порядка 4 % для нашей задачи, что указывает на допустимость суждений по данной матрице.

В данном примере критерии второго уровня необходимо сравнить попарно по отношению к общей цели первого уровня. Для сравнения поставщиков потребуется уже не одна, а три матрицы, поскольку необходимо сравнить поставщиков относительно друг друга по каждому критерию.

Коэффициенты приоритетов (КОВ), ИО, ОО для поставщиков по каждому из критериев вычислим аналогично предыдущему:

Цена	П1	П2	П3	Произведение	Корень	КОВ
Поставщик 1 (П1)	1	0,5	1	0,5	0,794	0,260
Поставщик 2 (П2)	2	1	1	2	1,260	0,413
Поставщик 3 (П3)	1	1	1	1	1,000	0,327
Сумма по столбцам	4	2,5	3	3,5	3,0536216	
Произведение сумм на КОВ	1,04	1,03	0,98			
$л_{max}$	3,05					
М (ИО)	0,58					
ИО	0,026811					
ОО	0,046225					

Срок поставки	П1	П2	П3	Произведение	Корень	КОВ
Поставщик 1 (П1)	1	0,5	3	1,5	1,145	0,332
Поставщик 2 (П2)	2	1	3	6	1,817	0,528
Поставщик 3 (П3)	0,333333	0,333333	1	0,1111111	0,481	0,140
Сумма по столбцам	3,333333	1,833333	7	7,6111111	3,4425847	
Произведение сумм на КОВ	1,11	0,97	0,98			
$л_{max}$	3,05					

М (ИО)	0,58					
ИО	0,026811					
ОО	0,046225					

Год выпуска	П1	П2	П3	Произведение	Корень	КОВ
Поставщик 1 (П1)	1	0,5	2	1	1,000	0,297
Поставщик 2 (П2)	2	1	3	6	1,817	0,540
Поставщик 3 (П3)	0,5	0,333333	1	0,1666667	0,550	0,163
Сумма по столбцам	3,5	1,833333	6	7,1666667	3,3674418	
Произведение сумм на КОВ	1,04	0,99	0,98			
$л_{max}$	3,01					
М (ИО)	0,58					
ИО	0,004601					
ОО	0,007933					

Приоритеты синтезируем, начиная со второго уровня вниз. Локальные приоритеты умножим на приоритет соответствующего критерия на вышестоящем уровне и суммируем по каждому элементу в соответствии с критериями, на которые воздействует этот элемент (каждый элемент второго уровня умножим на единицу, т. е. на вес единственной цели самого верхнего уровня). Это дает составной, или глобальный, приоритет того элемента, который затем используется для взвешивания локальных приоритетов элементов, сравниваемых с ним как с критерием и расположенных уровнем ниже. Процедуру продолжим до самого нижнего уровня.

В задаче выбора поставщика на втором уровне расположены критерии качеств, характеризующих работу поставщика (элементы их вектора приоритета умножим на единицу).

Третий уровень иерархии — перечень поставщиков. Каждый элемент этого уровня (относительный вес каждого поставщика по сравниваемому качеству) умножим на приоритет данного качества среди прочих, затем полученные произведения сложим.

В результате получим сводную таблицу:

	Срок поставки	Год выпуска	Сумма по столбцам	Глобальные приоритеты
КОВ критериев качества	0,594	0,249	0,157	
Поставщик 1 (П1)	0,260	0,332	0,297	0,2837
Поставщик 2 (П2)	0,413	0,528	0,540	**0,4616**
Поставщик 3 (П3)	0,327	0,140	0,163	0,2547

Глобальный приоритет первого поставщика является результатом вычислений:

$$0{,}594 \cdot 0{,}260 + 0{,}249 \cdot 0{,}332 + 0{,}157 \cdot 0{,}297 = 0{,}2837.$$

Аналогично вычисляются глобальные приоритеты других поставщиков. В результате делаем вывод о предпочтительности поставщика под номером 2.

Контрольные вопросы и задания

1. Каковы основные принципы метода анализа иерархий?
2. Назовите недостатки метода анализа иерархий. Как их можно устранить?
3. Опишите основные этапы принятия решений с помощью метода анализа иерархий.
4. Какие показатели используются для оценки однородности суждений эксперта?
5. Опишите алгоритм иерархического синтеза.
6. Какую шкалу использует метод анализа иерархий?

Задания для самостоятельного решения

1. Проведите выбор секретаря из девушек, подавших резюме, по пяти критериям:
— знание делопроизводства;
— внешний вид;
— знание английского языка;
— знание компьютера;
— умение разговаривать по телефону.
Собеседование прошли пять девушек:
— Ольга;
— Елена;
— Светлана;
— Галина;
— Жанна;
После собеседования были получены следующие описания девушек:

Ольга. Приятная внешность. Отличное знание английского языка. Хорошее поведение. Нет навыков работы на компьютере, посредственное общение по телефону.

Елена. Красивая, приятная внешность, хорошее умение общаться по телефону. Незнание английского языка, нет навыков работы на компьютере, делопроизводство знает весьма плохо.

Светлана. Очень хорошее знание делопроизводства, хорошие навыки работы на компьютере, достаточно хорошо общается по телефону, очень исполнительная. Не очень приятная внешность, посредственное знание английского языка.

Галина. Достаточно хорошо знает делопроизводство, неплохие навыки работы на компьютере, по телефону общается на высоком уровне, достаточно хорошее поведение. Плохое знание английского языка, неприятная внешность.

Жанна. Приятная внешность, очень хорошее поведение, неплохие навыки работы на компьютере, достаточно хорошее знание английского языка. По телефону общается плохо, не знает делопроизводство.

Оцените «важность» критериев по своему усмотрению. Используя метод анализа иерархий, проранжируйте кандидаток.

2. Примените метод анализа иерархий для поддержки принятия решений во внешнеэкономической сфере.

3. Примените метод анализа иерархий для решения вопроса о распределении ограниченного объема средств бюджета среди трех основных отраслей деятельности членов сообщества: наука, культура и образование; промышленное производство; сельскохозяйственное производство. Уровень развития каждой из указанных отраслей оказывает самое непосредственное влияние на каждую из трех основных компонент, определяющих благосостояние общества: материальное положение населения; его здоровье; безопасность общества.

4. Разрешите проблему по заполнению вакансии, на которую претендуют три человека, используя МАИ. В качестве критериев выбраны квалификационные (I), профессиональные (II), личностные качества (III), результативность работы (IV), удовлетворенность условиями, которые готова предоставить организация (V), и прочие аспекты личности (VI).

5. Разрешите проблему распределения доверия народных масс в стране между тремя ее крупнейшими партиями: Единая Россия, СПС и КПРФ. Целями, по отношению к которым оцениваются эти партии, являются вклад в развитие демократии, вклад в борьбу с коррупцией и вклад в национальную безопасность. Общая цель — благоприятное социальное и политическое развитие страны.

6. Разрешите проблему распределения энергии в стране между тремя ее крупнейшими потребителями: быт, промышленность и транспорт. Целями, по отношению к которым оценивается распределение энергии, являются вклад в развитие экономики, вклад в развитие сельского хозяйства и вклад в национальную безопасность. Общая цель — благоприятное социальное и политическое развитие страны.

Библиографический список

Основная литература

1. *Осипов, А. Л.* Экономико-математические методы в управлении : учеб.-метод. комплекс для дистанц. обучения / А. Л. Осипов, Е. А. Рапоцевич ; СибАГС. — Новосибирск : Изд-во СибАГС, 2006. — 144 с.

2. *Тюрин, Ю. Н.* Анализ данных на компьютере : учеб. пособие / Ю. Н. Тюрин, А. А. Макаров. — Москва : ФОРУМ, 2008. — 366 с.

3. *Шикин, Е. В.* Математические методы и модели в управлении : учеб. пособие / Е. В. Шикин, А. Г. Чхартишвили. — 2-е изд., испр. — Москва : Дело, 2002. — С. 65—79.

Дополнительная литература

1. *Осипов, А. Л.* Информационные технологии в управлении : учеб. пособие / А. Л. Осипов ; Новосиб. гос. ун-т экономики и управления. — Новосибирск, 2012. — 216 с.
2. *Саати, Т. Л.* Принятие решений при зависимостях и обратных связях. Аналитические сети / Т. Л. Саати. — Москва : ЛКИ, 2008. — 360 с.
3. *Саати, Т. Л.* Принятие решений. Метод анализа иерархий / Т. Л. Саати. — Москва : Радио и связь, 1993. — 314 с.
4. *Синюк, В. Г.* Использование информационно-аналитических технологий при принятии управленческих решений : учеб. пособие / В. Г. Синюк, А. В. Шевырев. — Москва : Экзамен, 2004. — 160 с.

Часть 2. СТОХАСТИЧЕСКИЕ МЕТОДЫ

Глава 5. Регрессионный анализ

§ 1. Парный линейный регрессионный анализ

Предположим, что для исследуемых социально-экономических переменных x и y имеется n выборочных наблюдений $x_1, x_2, \ldots, x_n$, $y_1, y_2, \ldots, y_n$. На рис. 2 в системе прямоугольных координат нанесено поле рассеяния, точки которого соответствуют парам чисел (x_i, y_i), $i = 1, \ldots, n$. На основе анализа поля рассеяния выдвигаем гипотезу о том, что зависимость y от x описывается линейной моделью вида

$$y = \text{б}_0 + \text{б}_1 x + \text{е},$$

где α_0 и α_1 — неизвестные постоянные коэффициенты;

ε — случайная переменная (случайное возмущение), отражающая влияние неучтенных факторов и погрешностей измерений.

Для данной модели задача состоит в получении уравнения регрессии $\hat{y} = a_0 + a_1 x$ (см. прямую на рис. 2), в котором коэффициенты a_0 и a_1 есть оценки неизвестных параметров α_0 и α_1. Нахождение оценок a_0 и a_1 основывается на применении метода наименьших квадратов (суть МНК была изложена в главе «Методы прогнозирования», в которой описывалось построение линейного тренда).

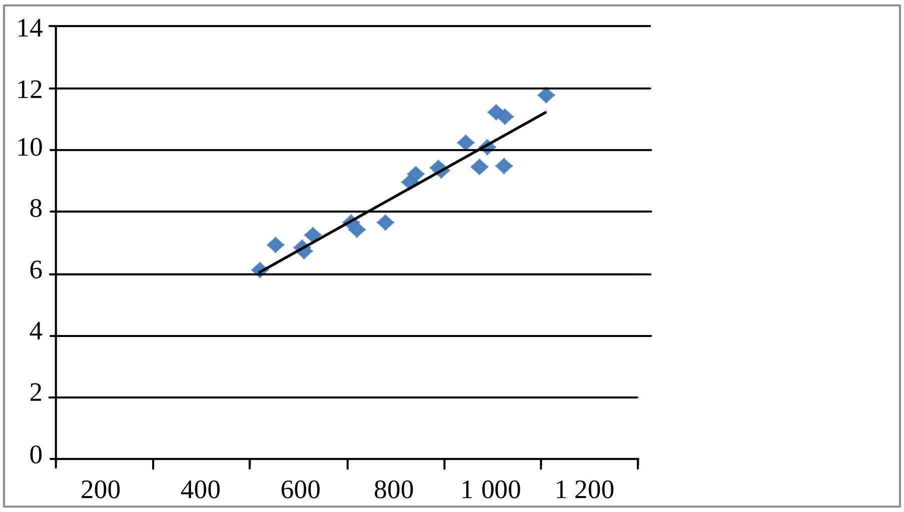


Рис. 2. Построение уравнения регрессии

Различие заключается в том, что вместо графика исходными данными является поле рассеяния, вместо переменной *t* используется переменная *x*.

Напомним, что нахождение оценок a_0, a_1 неизвестных параметров α_0 и α_1 сводится к решению экстремальной задачи для функции двух переменных $F\,(a_0, a_1)$:

$$F(a_0, a_1) \equiv \sum (y_i - a_0 - a_1 x_i)^2 \to \min.$$

Необходимые условия минимума функции $F\,(a_0,\ a_1)$ — равенство нулю частных производных:

$$\frac{\partial F(a_0, a_1)}{\partial a_0} = 0;$$
$$\frac{\partial F(a_0, a_1)}{\partial a_1} = 0.$$

Вычислим эти частные производные:

$$\frac{\partial F}{\partial a_0} = -2\sum (y_i - a_0 - a_1 x_i);$$
$$\frac{\partial F}{\partial a_1} = -2\sum (y_i - a_0 - a_1 x_i)\, x_i.$$

Приравняем их нулю и после элементарных преобразований получим систему двух линейных уравнений с двумя неизвестными a_0 и a_1:

$$\left.\begin{aligned} a_0 n + a_1 \sum x_i &= \sum y_i, \\ a_0 \sum x_i + a_1 \sum x_i^2 &= \sum x_i y_i \end{aligned}\right\}.$$

Эту систему уравнений называют системой нормальных уравнений. Ее решение может быть получено, например, по правилу Крамера:

$$\left.\begin{aligned} a_0 &= \frac{\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i}{n\sum x_i^2 - (\sum x_i)^2} \\ a_1 &= \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \end{aligned}\right\}.$$

Суммирование проводят для $i = 1, 2, \ldots, n$.

Пример 1. Для определения выручки от продажи подсолнечного масла в магазинах города в течение дня было исследовано 20 магазинов.

В результате получили следующие данные:

Число посетителей магазина	Выручка магазина
907	11,20
926	11,05
506	6,84
741	9,21
789	9,42
889	10,08
874	9,45
510	6,73
529	7,24
420	6,12
679	7,63
872	9,43
924	9,46
607	7,64
452	6,92
729	8,95
794	9,33
844	10,23
1 010	11,77
621	7,41

Далее построим регрессионную модель зависимости выручки магазина от числа посетителей.

Поле рассеяния для этой задачи приведено на рис. 2. Найдем коэффициенты линейного уравнения регрессии. Для этого составим вспомогательную таблицу.

x	y	x^2	$x\,y$
907	11,20	822 649	10 158,40
926	11,05	857 476	10 232,30
506	6,84	256 036	3 461,04
741	9,21	549 081	6 824,61
789	9,42	622 521	7 432,38
889	10,08	790 321	8 961,12
874	9,45	763 876	8 259,30

Окончание таблицы

x	y	x^2	$x\,y$
510	6,73	260 100	3 432,30
529	7,24	279 841	3 829,96
420	6,12	176 400	2 570,40
679	7,63	461 041	5 180,77
872	9,43	760 384	8 222,96
924	9,46	853 776	8 741,04
607	7,64	368 449	4 637,48
452	6,92	204 304	3 127,84
729	8,95	531 441	6 524,55
794	9,33	630 436	7 408,02
844	10,23	712 336	8 634,12
1 010	11,77	1 020 100	11 887,70
621	7,41	385 641	46 01,61
Итого			
14 623	**176,11**	**11 306 209**	**134 127,9**

Тогда систему уравнений МНК запишем в виде

$$\left.\begin{aligned} 20a_0 + 14\,623a_1 &= 176{,}11 \\ 14\,523a_0 + 11\,306\,209a_1 &= 134\,127{,}9 \end{aligned}\right\}.$$

Решая ее по методу Крамера, получим

$$a_0 = \frac{\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{176{,}11 \cdot 11\,306\,209 - 134\,127{,}9 \cdot 14623}{20 \cdot 11\,306\,209 - 14\,623^2} = 2{,}423004,$$

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{20 \cdot 134\,127{,}9 - 14\,623 \cdot 176{,}11}{20 \cdot 11\,306\,209 - 14\,623^2} = 0{,}008729.$$

Уравнение линейной парной регрессии запишем в виде

$$\hat{y} = a_0 + a_1 x = 2{,}423004 + 0{,}008729x.$$

Для оценки качества полученного уравнения регрессии существует ряд характеристик.

В первую очередь рассмотрим коэффициент парной корреляции.

Коэффициент парной корреляции r_{yx} характеризует тесноту линейной зависимости между x и y. Он находится по формуле

$$r_{yx} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

Суммирование проводят для $i = 1, 2, \ldots, n$.

Для коэффициента парной корреляции выполняется соотношение $-1 \le r_{yx} \le 1$. Чем ближе значение $|r_{yx}|$ к единице, тем теснее линейная связь между x и y. Если $|r_{yx}| = 1$, то между x и y существует функциональная зависимость вида

$$y = a_0 + a_1 x.$$

Если величина $|r_{yx}|$ близка к нулю, то это свидетельствует об отсутствии линейной зависимости между x и y и не исключает наличия нелинейной взаимосвязи между x и y. Близость значения коэффициента корреляции к нулю или единице носит относительный характер. Действительно, если $r_{yx} = 0,99$, то можно с уверенностью говорить о близости значения к единице и достаточно сильной линейной взаимосвязи между x и y. Но если r_{yx} равен, например, 0,7, то говорить о его близости к единице оснований значительно меньше, а если $r_{yx} = 0,5$, то можно с равными основаниями говорить как о близости к нулю, так и о близости к единице.

Для того чтобы с большей уверенностью полагаться на значение коэффициента корреляции, т. е. с большей уверенностью делать вывод о наличии или отсутствии линейной взаимосвязи между переменными y и x, разработан критерий проверки того, насколько существенно отличие коэффициента корреляции от нуля или, как говорят, насколько значимо значение коэффициента корреляции. Если в результате проверки выясняется, что коэффициент корреляции существенно отличается от нуля, то, несмотря даже на не очень близкое значение коэффициента к единице, делается вывод о наличии линейной взаимосвязи между переменными y и x. Если же подтверждается несущественное отличие r_{yx} от нуля, то, несмотря на возможно достаточно большое значение коэффициента, делается вывод об отсутствии линейной взаимосвязи между переменными.

Проверка существенности отличия коэффициента парной корреляции от нуля (его значимости) проводится по схеме проверки статистических гипотез. Выдвигается нулевая гипотеза — коэффициент парной корреляции не является статистически значимым $(H_0 : r_{yx} = 0)$. Вместе с ней выдвигается альтернативная гипотеза — существует положительная корреляционная зависимость $(H_1 : r_{yx} > 0)$. Для проверки гипотезы в качестве статистического критерия используется статистика t-Стьюдента с $\nu = n - 2$ степенями свободы: $t_{\text{наб}} = |r_{yx}|\sqrt{\dfrac{n-2}{1-r_{yx}^2}}$. Если $t_{\text{наб}} > t_{1-б/2}(n-2)$, то гипотеза о существенном отличии коэффициента парной корреляции от нуля принимается, в противном случае — отвергается. В формуле $t_{1-б/2}(n-2)$ — квантиль порядка $(1-\alpha/2)$ распределения Стьюдента с $(n-2)$ степенями свободы (табл. 2 приложения).

Пример 2. Следует найти коэффициент парной корреляции для задания из предыдущего примера и проверить его значимость.

Для этого во вспомогательной таблице поместим еще один столбец, в котором рассчитаем величину $\sum y_i^2 = 1\,602,097$. Тогда вычисления будут иметь вид

$$r_{yx} = \frac{20\cdot 134\,127,9 - 14\,623\cdot 175,11}{\sqrt{20\cdot 11306\,209 - 14\,623^2}\cdot\sqrt{20\cdot 1\,602,097 - 176,11^2}} = 0,954613.$$

Проверим гипотезу о значимости коэффициента парной корреляции при $\alpha = 0,05$. Для этого вычислим $t_{\text{наб}} = |r_{yx}|\sqrt{\dfrac{n-2}{1-r_{yx}^2}} = 13,64$ и $t_{1-б/2}(n-2) = t_{0,975}(18) = 2,101$. Наблюдаемое значение критерия больше табличного. Следовательно, коэффициент корреляции значим.

Коэффициент парной корреляции r_{yx} связан с коэффициентом a_1 уравнения регрессии $\hat{y} = a_0 + a_1 x$ следующим образом:

$$r_{yx} = a_1 \frac{S_x}{S_y},$$

где S_x, S_y — выборочные среднеквадратические отклонения случайных переменных x и y соответственно, рассчитывающиеся по формулам

$$S_x = \sqrt{S_x^2}, \qquad S_x^2 = \frac{1}{n}\sum (x_i - \bar{x})^2,$$

$$S_y = \sqrt{S_y^2}, \qquad S_y^2 = \frac{1}{n}\sum (y_i - \bar{y})^2.$$

Следующей важной характеристикой качества подбора уравнения регрессии является ***коэффициент детерминации***, обозначаемый R^2. Определение R^2 и его содержательный смысл основаны на следующей формуле:

$$\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2,$$

где $\bar{y}$ — выборочное среднее;

y_i — выборочные значения зависимой переменной y;

$\hat{y}_i$ — значения зависимой переменной, вычисленные по уравнению регрессии $\hat{y}_i = a_0 + a_1 x_i$.

Приведенная формула имеет глубокий содержательный смысл. Действительно, левая ее часть, т. е. $\sum (y_i - \bar{y})^2$, интерпретируется как мера общего разброса, или рассеивания, переменной y относительно ее среднего значения $\bar{y}$. Эта мера раскладывается на две составляющие. Первая часть $\sum (\hat{y}_i - \bar{y})^2$ — это мера разброса, «объясненная» уравнением регрессии. Вторая часть $\sum (y_i - \hat{y}_i)^2$ — это мера разброса, «не объясненная» уравнением регрессии. Слова «объясненный» и «не объясненный» взяты в кавычки, так как объяснение, в сущности, может оказаться мнимым. В действительности y может зависеть от какой-то другой переменной z, и x может действовать как величина, заменяющая z.

Коэффициент детерминации R^2 определяется по формуле

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}, \quad \text{или} \quad R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}.$$

Очевидно, что $0 \le R^2 \le 1$. Значение R^2 характеризует ту долю дисперсии переменной y, которая обусловливается, или которую можно «объяснить», уравнением регрессии $\hat{y} = a_0 + a_1 x$.

Таким образом, чем ближе значение R^2 к единице, тем точнее уравнение регрессии отражает имеющуюся зависимость между переменными y и x. Максимальное значение коэффициента детерминации, равное единице, достигается только тогда, когда линия регрессии точно соответствует всем наблюдениям, так что $y_i = \hat{y}_i$ для всех наблюдений и все остатки равны нулю. Если же в выборке отсутствует видимая связь между y и x, то R^2 будет близок к нулю.

Коэффициенты корреляции и детерминации для уравнения парной регрессии связаны между собой простым соотношением:

$$R^2 = r_{yx}^2.$$

Качество уравнения регрессии оценивает и ***F*-тест**. Он основан на проверке гипотезы H_0 о статистической незначимости уравнения регрессии и показателя тесноты связи. Для этого выполняется сравнение фактического $F_{\text{наб}}$ и критического (табличного) $F_{\text{таб}}$ значений F-статистики Фишера. Если $F_{\text{таб}} < F_{\text{наб}}$, то гипотеза H_0 отклоняется и признается статистическая значимость уравнения регрессии. Если $F_{\text{таб}} > F_{\text{наб}}$, то гипотеза H_0 не отклоняется и признается статистическая незначимость, ненадежность уравнения регрессии.

$F_{\text{наб}}$ определяется как отношение объясненной суммы квадратов в расчете на одну независимую переменную к остаточной сумме квадратов в расчете на одну степень свободы:

$$F_{\text{наб}} = \frac{\sum (\hat{y}_i - \bar{y})^2 / m}{\sum (y_i - \bar{y})^2 / (n - m - 1)} = \frac{R^2\ (n - m - 1)}{(1 - R^2)\ m}.$$

Для парной регрессии m = 1, поэтому

$$F_{\text{наб}} = \frac{R^2\ (n - 2)}{(1 - R^2)}.$$

F-распределение Фишера зависит от степеней свободы $df1$ и $df2$ и от уровня значимости α. Количество степеней свободы $df1$ равно числу m объясняющих переменных модели. Количество степеней свободы $df2$ определяется объемом выборки n за вычетом числа объясняющих переменных модели $df1$ минус единица: $df2 = n - df1 - 1$.

$F_{\text{таб}}$ можно вычислить как по статистической таблице (табл. 3 приложения), так и с помощью статистической функции из приложения MS Excel.

Пример 3. Следует проверить значимость полученного в предыдущем задании уравнения регрессии.

Найдем $R^2 = r_{yx}^2 = 0{,}911859.$ Проверим гипотезу о значимости коэффициента детерминации.

$$F_{\text{наб}} = \frac{R^2\ (n - 2)}{(1 - R^2)} = 186{,}2188.$$

$$F_{\text{таб}} = F(0{,}95; 1; 18) = 4{,}41.$$

Наблюдаемое значение критерия больше табличного. Следовательно, коэффициент детерминации значим, что свидетельствует о хорошем качестве найденного уравнения регрессии.

§ 2. Множественный линейный регрессионный анализ

Парная регрессия может дать хороший результат при моделировании, если влиянием других факторов, воздействующих на объект исследования, можно пренебречь. Если же этим влиянием пренебречь нельзя, то следует попытаться выявить влияние других факторов, введя их в модель, т. е. построить ***уравнение множественной регрессии***:

$$y = f(x_1, x_2, \ldots, x_m),$$

где y — зависимая переменная (результативный признак);

x_i — независимые, или объясняющие, переменные (признаки-факторы).

Множественная регрессия широко используется в решении проблем спроса, доходности акций, при изучении функции издержек производства, в макроэкономических расчетах и целом ряде других вопросов моделирования социально-экономических процессов.

В настоящее время множественная регрессия — один из наиболее распространенных инструментов моделирования.

Основная *цель* множественной регрессии — построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Построение уравнения множественной регрессии начинается с решения вопроса о спецификации модели. Он включает в себя два круга вопросов: отбор факторов и выбор вида уравнения регрессии.

Включение в уравнение множественной регрессии того или иного набора факторов связано прежде всего с представлением исследователя о природе взаимосвязи моделируемого показателя с другими экономическими явлениями. Данные факторы должны отвечать следующим *требованиям*:

1) они должны быть количественно измеримы. Если необходимо включить в модель качественный фактор, не имеющий количественного измерения, то ему нужно придать количественную определенность;

2) не должны быть интеркоррелированы и тем более не должны находиться в точной функциональной связи;

3) включение в модель факторов с высокой интеркорреляцией может привести к нежелательным последствиям: система нормальных уравнений может оказаться плохо обусловленной и повлечь за собой неустойчивость и ненадежность оценок коэффициентов регрессии;

4) если между факторами существует высокая корреляция, то нельзя определить их изолированное влияние на результативный показатель, и параметры уравнения регрессии оказываются неинтерпретируемыми.

Включаемые во множественную регрессию факторы должны объяснить вариацию независимой переменной. Если строится модель с набором m факторов, то для нее рассчитывается показатель детерминации, который фиксирует долю объясненной вариации результативного признака за счет рассматриваемых в регрессии m факторов.

При дополнительном включении в регрессию $m + 1$ фактора коэффициент детерминации должен возрастать, а остаточная дисперсия уменьшаться:

$$R^2_{m+1} \geq R^2_m \qquad \text{и} \qquad S^2_{m+1} \leq S^2_m.$$

Если же этого не происходит и данные показатели практически не отличаются друг от друга, то включаемый в анализ $m + 1$ фактор не улучшает модель и практически является лишним.

Несмотря на то что теоретически регрессионная модель позволяет учесть любое число факторов, практически в этом нет необходимости. Отбор факторов производится на основе качественного теоретико-экономического анализа. Однако теоретический анализ часто не позволяет однозначно ответить на вопрос о количественной взаимосвязи рассматриваемых признаков и целесообразности включения фактора в модель. Поэтому отбор факторов обычно осуществляется в две стадии: на первой подбираются факторы исходя из сущности проблемы; на второй — на основе матрицы показателей корреляции определяют статистики для параметров регрессии.

Коэффициенты интеркорреляции (т. е. корреляции между объясняющими переменными) позволяют исключать из модели дублирующие факторы. Считается, что две переменные явно коллинеарны, т. е. находятся между собой в линейной зависимости, если $r_{x_i x_j} \geq 0,7.$ Если факторы явно коллинеарны, то они дублируют друг друга и один из них рекомендуется исключить из регрессии. Предпочтение при этом отдается не фактору, более тесно связанному с результатом, а тому фактору, который при достаточно тесной связи с результатом имеет наименьшую тесноту связи с другими факторами. В этом требовании проявляется специфика множественной регрессии как метода исследования комплексного воздействия факторов в условиях их независимости друг от друга.

По величине парных коэффициентов корреляции обнаруживается лишь явная коллинеарность факторов. Наибольшие трудности в использовании аппарата множественной регрессии возникают при наличии *мультиколлинеарности* факторов, когда более чем два фактора связаны между собой линейной зависимостью, т. е. имеет место совокупное воздействие факторов друг на друга. Наличие мультиколлинеарности факторов может означать, что некоторые факторы будут всегда действовать в унисон. В результате вариация в исходных данных перестает быть полностью независимой и нельзя оценить воздействие каждого фактора в отдельности.

Включение в модель мультиколлинеарных факторов нежелательно в силу следующих *последствий*:

1) затрудняется интерпретация параметров множественной регрессии как характеристик действия факторов в «чистом» виде, ибо факторы коррелированы; параметры линейной регрессии теряют экономический смысл;

2) оценки параметров ненадежны, меняются с изменением объема наблюдений (не только по величине, но и по знаку), а также обнаруживаются большие стандартные ошибки, что делает модель непригодной для анализа и прогнозирования.

Для оценки мультиколлинеарности факторов может использоваться определитель матрицы парных коэффициентов корреляции между факторами.

Если бы факторы не коррелировали между собой, то матрица парных коэффициентов корреляции между факторами была бы единичной матрицей, поскольку все недиагональные элементы $r_{x_i x_j}$ $i \neq j$ были бы равны нулю.

Так, для уравнения, включающего три объясняющих переменных $\hat{y} = a + b_1 x_1 + b_2 x_2 + b_3 x_3,$ матрица коэффициентов корреляции между факторами имела бы определитель, равный единице:

$$\mathrm{Det}\, R = \begin{vmatrix} r_{x_1x_1} & r_{x_1x_2} & r_{x_1x_3} \\ r_{x_2x_1} & r_{x_2x_2} & r_{x_2x_3} \\ r_{x_3x_1} & r_{x_3x_2} & r_{x_3x_3} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Если же между факторами существует полная линейная зависимость и все коэффициенты корреляции равны единице, то определитель такой матрицы будет равен нулю:

$$\mathrm{Det}\ R = \begin{vmatrix} r_{x_1x_1} & r_{x_1x_2} & r_{x_1x_3} \\ r_{x_2x_1} & r_{x_2x_2} & r_{x_2x_3} \\ r_{x_3x_1} & r_{x_3x_2} & r_{x_3x_3} \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

Чем ближе к нулю определитель матрицы межфакторной корреляции, тем сильнее мультиколлинеарность факторов и ненадежнее результаты множественной регрессии. И, наоборот, чем ближе к единице определитель матрицы межфакторной корреляции, тем меньше мультиколлинеарность факторов.

Возможны разные *виды уравнений множественной регрессии*: линейные и нелинейные. Ввиду четкой интерпретации параметров наиболее широко используется линейная функция. В линейной множественной регрессии

$$\hat{y}_x = a + b_1x_1 + b_2x_2 + \ldots + b_mx_m$$

параметры при x называются *коэффициентами регрессии*. Они характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизмененном значении других факторов, закрепленных на среднем уровне.

Классический подход к оцениванию параметров линейной модели множественной регрессии основан на МНК. Этот метод позволяет получить такие оценки параметров, при которых сумма квадратов отклонений фактических значений результативного признака y от расчетных $\hat{y}$ минимальна:

$$\sum_i \left(y_i - \hat{y}_{x_i}\right)^2 \to \min.$$

Как известно из курса математического анализа, для того чтобы найти экстремум функции нескольких переменных, надо вычислить частные производные первого порядка по каждому из параметров и приравнять их к нулю. Имеем функцию $m + 1$ аргумента:

$$S\left(a,\, b_1,\, b_2,\, \ldots,\, b_m\right) = \sum \left(y - a - b_1x_1 - b_2x_2 - \ldots - b_mx_m\right)^2.$$

Найдем частные производные первого порядка:

$$\begin{cases} \dfrac{\partial S}{\partial a} = -2\sum \left(y - a - b_1 x_1 - b_2 x_2 - \ldots - b_m x_m \right) = 0; \\ \dfrac{\partial S}{\partial b_1} = -2\sum x_1 \left(y - a - b_1 x_1 - b_2 x_2 - \ldots - b_m x_m \right) = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial S}{\partial b_m} = -2\sum x_m \left(y - a - b_1 x_1 - b_2 x_2 - \ldots - b_m x_m \right) = 0. \end{cases}$$

После элементарных преобразований приходим к системе линейных нормальных уравнений для нахождения параметров линейного уравнения множественной регрессии:

$$\begin{cases} na + b_1 \sum x_1 + b_2 \sum x_2 + \ldots + b_m \sum x_m = \sum y; \\ a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 + \ldots + b_m \sum x_1 x_m = \sum y x_1; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a \sum x_m + b_1 \sum x_1 x_m + b_2 \sum x_2 x_m + \ldots + b_m \sum x_m^2 = \sum y x_m. \end{cases}$$

Для двухфакторной модели данная система будет иметь вид

$$\begin{cases} na + b_1 \sum x_1 + b_2 \sum x_2 = \sum y; \\ a \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 = \sum y x_1; \\ a \sum x_2 + b_1 \sum x_1 x_2 + b_2 \sum x_2^2 = \sum y x_2. \end{cases}$$

Пример 4. Пусть имеются следующие данные (условные) о сменной добыче угля на 1 рабочего y (т), мощности пласта x_1 (м) и уровне механизации работ x_2 (%), характеризующие процесс добычи угля в 10 шахтах.

Номер шахты	1	2	3	4	5	6	7	8	9	10

x_1	8	11	12	9	8	8	9	9	8	12
x_2	5	8	8	5	7	8	6	4	5	7
y	5	10	10	7	5	6	6	5	6	8

Предполагая, что между переменными y, x_1, x_2 существует линейная корреляционная зависимость, найдем уравнение регрессии y по x_1 и x_2. Для удобства дальнейших вычислений составим таблицу ($e = y - \hat{y}_x$):

Номер шахты	x_1	x_2	y	x_1^2	x_2^2	y^2	x_1x_2	x_1y	x_2y	$\hat{y}_x$	e^2
1	8	5	5	64	25	25	40	40	25	5,13	0,016
2	11	8	10	121	64	100	88	110	80	8,79	1,464
3	12	8	10	144	64	100	96	120	80	9,64	0,127
4	9	5	7	81	25	49	45	63	35	5,98	1,038
5	8	7	5	64	49	25	56	40	35	5,86	0,741
6	8	8	6	64	64	36	64	48	48	6,23	0,052
7	9	6	6	81	36	36	54	54	36	6,35	0,121
8	9	4	5	81	16	25	36	45	20	5,61	0,377
9	8	5	6	64	25	36	40	48	30	5,13	0,762
10	12	7	8	144	49	64	84	96	56	9,28	1,631
Итого	**94**	**63**	**68**	**908**	**417**	**496**	**603**	**664**	**445**	**68,00**	**6,329**
Среднее значение	**9,4**	**6,3**	**6,8**	**90,8**	**41,7**	**49,6**	**60,3**	**66,4**	**44,5**	**—**	**—**

Для нахождения параметров уравнения регрессии в данном случае необходимо решить следующую систему нормальных уравнений:

$$\begin{cases} 10a + 94b_1 + 63b_2 = 68, \\ 94a + 908b_1 + 603b_2 = 664, \\ 63a + 603b_1 + 417b_2 = 445. \end{cases}$$

Откуда получим, что $a = -3{,}54$, $b_1 = 0{,}854$, $b_2 = 0{,}367$, т. е. получим следующее уравнение множественной регрессии:

$$\hat{y}_x = -3{,}54 + 0{,}854x_1 + 0{,}367x_2.$$

Оно показывает, что при увеличении только мощности пласта x_1 (при неизменном x_2) на 1 м добыча угля на 1 рабочего y увеличится в среднем на 0,854 т, а при увеличении только уровня механизации работ x_2 (при неизменном x_1) на 1 % — в среднем на 0,367 т.

Контрольные вопросы и задания

1. В чем заключается суть метода наименьших квадратов?
2. Дайте определение понятия коэффициента парной корреляции.
3. В каких пределах этот коэффициент изменяется?
4. Как можно проверить его значимость?
5. Что означает коэффициент детерминации?
6. В каких пределах он изменяется?
7. Как связаны между собой коэффициенты парной корреляции и детерминации?
8. Перечислите основные свойства факторов, включаемых в множественную регрессию.
9. В чем заключается суть понятия мультиколлинеарности?
10. Как оценить мультиколлинеарность?
11. В чем заключается суть статистической значимости коэффициентов регрессии?
12. Верно ли утверждение о том, что для парной линейной регрессии коэффициент корреляции превосходит коэффициент детерминации.
13. Как проводится проверка значимости регрессионной модели?
14. Как связан коэффициент детерминации с критерием Фишера?
15. Поясните смысл коэффициента регрессии.
16. Как проводится проверка значимости уравнения множественной регрессии?
17. Что характеризуют коэффициенты регрессии в модели множественной регрессии?

Задания для самостоятельного решения

1. Определите функцию спроса (зависимость сбыта Q от цены товара P) по следующим данным:

Цена P, тыс. руб.	54	50	55	59	60	58	64
Объем сбыта Q, шт.	570	600	580	510	480	500	450

Постройте линейное уравнение регрессии и проверьте его значимость. Осуществите прогноз при цене, равной 68 тыс. руб.

2. В таблице представлены данные о рентах и свободных площадях для 10 городов.

Город	1	2	3	4	5	6	7	8	9	10
Доля площадей	10	2	7	8	4	6	7	8	9	10
Месячная рента 1 фут^2, долл.	5,0	2,5	4,75	4,5	3,0	4,5	4,0	3,0	3,25	2,75

Постройте линейное уравнение регрессии и проверьте его значимость. Выполните прогноз доли свободной площади для города с месячной рентой за 1 фут^2, равной 3,50 долл.

3. По 20 предприятиям региона изучается зависимость выработки продукции на 1 работника y (тыс. руб.) от ввода в действие новых основных фондов x_1 (от стоимости фондов на конец года, %) и от удельного веса рабочих высокой квалификации в общей численности рабочих x_2 (%).

Постройте линейную модель множественной регрессии по следующим данным.

Вариант 1

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	6	3,6	9	11	9	6,3	21
2	6	3,6	12	12	11	6,4	22
3	6	3,9	14	13	11	7,0	24
4	7	4,1	17	14	12	7,5	25
5	7	3,9	18	15	12	7,9	28
6	7	4,5	19	16	13	8,2	30
7	8	5,3	19	17	13	8,0	30
8	8	5,3	19	18	13	8,6	31
9	9	5,6	20	19	14	9,5	33
10	10	6,8	21	20	14	9,0	36

Вариант 2

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	6	3,5	10	11	10	6,3	21
2	6	3,6	12	12	11	6,4	22
3	7	3,9	15	13	11	7,0	23
4	7	4,1	17	14	12	7,5	25
5	7	4,2	18	15	12	7,9	28
6	8	4,5	19	16	13	8,2	30
7	8	5,3	19	17	13	8,4	31
8	9	5,3	20	18	14	8,6	31
9	9	5,6	20	19	14	9,5	35
10	10	6,0	21	20	15	10,0	36

Вариант 3

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,7	9	11	11	6,3	22
2	7	3,7	11	12	11	6,4	22

3	7	3,9	11	13	11	7,2	23
4	7	4,1	15	14	12	7,5	25
5	8	4,2	17	15	12	7,9	27
6	8	4,9	19	16	13	8,1	30
7	8	5,3	19	17	13	8,4	31
8	9	5,1	20	18	13	8,6	32
9	10	5,6	20	19	14	9,5	35
10	10	6,1	21	20	15	9,5	36

Вариант 4

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,5	9	11	10	6,3	22
2	7	3,6	10	12	10	6,5	22
3	7	3,9	12	13	11	7,2	24
4	7	4,1	17	14	12	7,5	25
5	8	4,2	18	15	12	7,9	27
6	8	4,5	19	16	13	8,2	30
7	9	5,3	19	17	13	8,4	31
8	9	5,5	20	18	14	8,6	33
9	10	5,6	21	19	14	9,5	35
10	10	6,1	21	20	15	9,6	36

Вариант 5

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,6	9	11	10	6,3	21
2	7	3,6	11	12	11	6,9	23
3	7	3,7	12	13	11	7,2	24
4	8	4,1	16	14	12	7,8	25
5	8	4,3	19	15	13	8,1	27
6	8	4,5	19	16	13	8,2	29
7	9	5,4	20	17	13	8,4	31
8	9	5,5	20	18	14	8,8	33
9	10	5,8	21	19	14	9,5	35
10	10	6,1	21	20	14	9,7	34

Вариант 6

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,5	9	11	10	6,3	21
2	7	3,6	10	12	10	6,8	22
3	7	3,8	14	13	11	7,2	24
4	7	4,2	15	14	12	7,9	25

5	8	4,3	18	15	12	8,1	26
6	8	4,7	19	16	13	8,3	29
7	9	5,4	19	17	13	8,4	31
8	9	5,6	20	18	13	8,8	32
9	10	5,9	20	19	14	9,6	35
10	10	6,1	21	20	14	9,7	36

Вариант 7

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,8	11	11	10	6,8	21
2	7	3,8	12	12	11	7,4	23
3	7	3,9	16	13	11	7,8	24
4	7	4,1	17	14	12	7,5	26
5	7	4,6	18	15	12	7,9	28
6	8	4,5	18	16	12	8,1	30
7	8	5,3	19	17	13	8,4	31
8	9	5,5	20	18	13	8,7	32
9	9	6,1	20	19	13	9,5	33
10	10	6,8	21	20	14	9,7	35

Вариант 8

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,8	9	11	11	7,1	22
2	7	4,1	14	12	11	7,5	23
3	7	4,3	16	13	12	7,8	25
4	7	4,1	17	14	12	7,6	27
5	8	4,6	17	15	12	7,9	29
6	8	4,7	18	16	13	8,1	30
7	9	5,3	20	17	13	8,5	32
8	9	5,5	20	18	14	8,7	32
9	11	6,9	21	19	14	9,6	33
10	10	6,8	21	20	15	9,8	36

Вариант 9

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,9	12	11	11	7,1	22
2	7	4,2	13	12	12	7,5	25
3	7	4,3	15	13	13	7,8	26

4	7	4,4	17	14	12	7,9	27
5	8	4,6	18	15	13	8,1	30
6	8	4,8	19	16	13	8,4	31
7	9	5,3	19	17	13	8,6	32
8	9	5,7	20	18	14	8,8	32
9	10	6,9	21	19	14	9,6	34
10	10	6,8	21	20	14	9,9	36

Вариант 10

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,6	12	11	10	7,2	23
2	7	4,1	14	12	11	7,6	25
3	7	4,3	16	13	12	7,8	26
4	7	4,4	17	14	11	7,9	28
5	7	4,5	18	15	12	8,2	30
6	8	4,8	19	16	12	8,4	31
7	8	5,3	20	17	12	8,6	32
8	8	5,6	20	18	13	8,8	32
9	9	6,7	21	19	13	9,2	33
10	10	6,9	22	20	14	9,6	34

Библиографический список

Основная литература

1. *Доугерти, К.* Введение в эконометрику : учеб. для экон. специальностей вузов : пер. с англ. / К. Доугерти. — Москва : ИНФРА-М, 2010. — С. 53—69; 134—158.

2. *Кремер, Н. Ш.* Эконометрика : учеб. для вузов / Н. Ш. Кремер, Б. А. Путко. — Москва : ЮНИТИ, 2008. — 310 с.

3. *Практикум* по эконометрике : учеб. пособие / И. И. Елисеева, С. В. Курышева, Н. М. Гордеенко и др. ; под ред. И. И. Елисеевой. — Москва : Финансы и статистика, 2001. — 192 с.

4. *Тюрин, Ю. Н.* Анализ данных на компьютере : учеб. пособие / Ю. Н. Тюрин, А. А. Макаров. — Москва : ФОРУМ, 2008. — 366 с.

5. *Осипов, А. Л.* Электрометрика : практикум / СибАГС ; сост. : А. Л. Осипов, Е. А. Рапоцевич. — Новосибирск : Изд-во СибАГС, 2008. — С. 19—42; 59—71.

6. *Эконометрика* : учеб. для вузов / [И. И. Елисеева и др.] ; под ред. И. И. Елисеевой. — Москва : Финансы и статистика, 2007. — 575 с.

Дополнительная литература

1. *Большаков, А. А.* Методы обработки многомерных данных и временных рядов : учеб. пособие / А. А. Большаков, Р. Н. Каримов. — Москва : Горячая линия — Телеком, 2007. — 522 с.

2. *Статистика* : учеб. / В. Г. Минашкин [и др.] ; под ред. В. Г. Минашкина. — Москва : ТК Велби : Проспект, 2006. — 272 с.

3. *Тимофеев, В. С.* Эконометрика : учеб. / В. С. Тимофеев, А. В. Фаддеенков, В. Ю. Щеколдин. — Новосибирск : Изд-во НГТУ, 2013. — 340 с.

4. *Яновский, Л. П.* Введение в эконометрику : учеб. пособие / Л. П. Яновский, А. Г. Буховец. — Москва : КноРус, 2007. — 255 с.

Глава 6. Анализ зависимостей в слабых шкалах

§ 1. Основные типы слабых шкал

В основе любого исследования лежат измерения тех или иных показателей. Но каждое измерение производится по определенной шкале, и поэтому необходимо понять, с какой шкалой следует иметь дело в каждом конкретном случае и какие операции можно совершать с данными.

Мы привыкли работать с вещественными числами (абсолютная шкала) и выполнять с ними арифметические действия. Подсознательно эти аналогии переносим к объектам другой природы, что часто приводит к неверным, а иногда и курьезным выводам.

Измерение — это алгоритмическая операция, которая ставит в соответствие наблюдаемому состоянию объекта, процесса, явления определенное обозначение: число, номер или символ.

Такое соответствие обеспечивает то, что результаты измерений содержат информацию о наблюдаемом объекте, количество же информации зависит от полноты этого соответствия и разнообразия вариантов. Нужная нам информация получается из результатов измерения с помощью их преобразований, или, как еще говорят, с помощью обработки экспериментальных данных.

Мы будем рассматривать только такие объекты, про любые два состояния которых можно сказать, различимы они или нет, и только такие алгоритмы измерения, которые различимым состояниям ставят в соответствие разные обозначения, а неразличимым состояниям — одинаковые обозначения.

Введем ***классификацию слабых шкал***.

Предположим, что число различимых состояний (математический термин — число классов эквивалентности) конечно. Каждому классу эквивалентности поставим в соответствие обозначение, отличное от обозначений других классов. Теперь измерение будет состоять в том, чтобы, проведя эксперимент над объектом, определить принадлежность результата к тому или иному классу эквивалентности и записать это с помощью символа, обозначающего данный класс. Такое измерение называется измерением в ***шкале наименований*** (иногда эту шкалу называют номинальной); указанное множество символов и образует шкалу. Разумеется, ни логические, ни математические операции к элементам этой шкалы неприменимы. Назначение измерений в шкале наименований — различать объекты. Так, в шкале наименований измерены номера телефонов, автомашин, паспортов, студенческих билетов.

В тех случаях, когда наблюдаемый (измеряемый) признак состояния имеет природу, не только позволяющую отождествить состояние с одним из классов эквивалентности, но и дающую возможность в каком-то отношении сравнивать разные классы, для измерений можно выбрать более сильную шкалу, чем номинальная, что дает возможность получить дополнительную полезную информацию о наблюдаемом объекте.

Порядковая шкала применяется в том случае, если классы эквивалентности можно упорядочить. Номер объекта в ранжированном ряду называется *рангом объекта*. Отсюда происходит другое название порядковых шкал — ***ранговые***.

Если часть наблюдений совпадает (в статистике такая группа наблюдений называется связкой) и все члены связки получают одинаковый (старший для них) ранг, то имеет место слабый порядок. В этом случае либо прибегают к присвоению ранга, среднего для данной связки, либо присваивают ранги от младшего до старшего случайным образом.

При измерениях в порядковых шкалах обработка данных должна основываться только на допустимых для этих шкал операциях — ***вычислении рангов и относительных частот***. Примерами шкал порядка могут служить шкалы силы ветра, силы землетрясения, сортности товаров в торговле, различные социологические шкалы.

В порядковой шкале числа используются для установления порядка между объектами. Простейшим примером являются оценки знаний учащихся. Так, в средней школе применяются оценки 2, 3, 4, 5, а в высшей школе тот же смысл выражается словесно — «неудовлетворительно», «удовлетворительно», «хорошо», «отлично». Этим подчеркивается нечисловой характер оценок знаний учащихся. Оценки экспертов следует считать измеренными в порядковой шкале.

Если упорядочивание объектов можно выполнить настолько точно, что известны расстояния между любыми двумя из них, то измерение окажется заметно сильнее, чем в шкале порядка. Такие шкалы называются ***интервальными***. Естественно выражать все расстояния в единицах, хотя и произвольных, но одинаковых по всей длине шкалы. Это означает, что объективно равные интервалы измеряются одинаковыми по длине отрезками шкалы, где бы они на ней ни располагались. Следствием такой равномерности шкал этого класса является независимость отношения двух интервалов от того, в какой из шкал эти интервалы измерены (т. е. какова единица длины интервала и какое значение принято за начало отсчета).

Примерами величин, которые по физической природе либо не имеют абсолютного нуля, либо допускают свободу выбора в установлении начала отсчета и поэтому измеряются в интервальных шкалах, являются температура, время, высота местности.

Другим примером измерения в интервальной шкале может служить признак «дата совершения события», поскольку для измерения времени в конкретной шкале необходимо фиксировать масштаб и начало отсчета. Григорианский и мусульманский календари — две конкретизации шкал интервалов.

Над интервалами можно выполнять любые арифметические операции, а вместе с ними использовать подходящие способы статистической и иной обработки данных.

§ 2. Корреляционный анализ в слабых шкалах

В случае, когда изучаются не количественные, а качественные признаки, обычные коэффициенты корреляции для выявления взаимозависимости оказываются непригодными. Однако если удается как-то упорядочить наблюдения в отношении определенного показателя, т. е. приписать им порядковые номера — ранги, то задача может быть решена.

Анализ изменения рангов соответствующих наблюдений двух признаков позволяет делать выводы о наличии или отсутствии зависимости между этими признаками. Так, если при увеличении ранга одного признака ранг другого в среднем тоже увеличивается, то говорят о прямой зависимости. А при обратной связи увеличение ранга одного признака соответствует уменьшению ранга другого.

Для изучения силы зависимости между качественными признаками применяются ***ранговые коэффициенты корреляции Спирмена и Кендэла***.

Коэффициентом ранговой корреляции Спирмена называется величина

$$r_c = 1 - \frac{6S}{n^3 - n},$$

где n — объем выборки; $S = \sum_{i=1}^{n} \left(R_1(i) - R_2(i)\right)^2$; $R_1(i), R_2(i)$ — ранги i-го наблюдения по первому и второму признакам соответственно.

Если при определении рангов некоторые наблюдения неразличимы, то они получают одинаковые ранги. Отсюда ранговый коэффициент корреляции обладает всеми свойствами парного коэффициента корреляции, т. е. он по модулю не превосходит единицу. Чем коэффициент Спирмена по модулю ближе к единице, тем связь между признаками сильнее, чем он ближе к нулю, тем связь слабее.

Для оценки этой близости используют аппарат проверки статистических гипотез.

Например, выдвигаются нулевая гипотеза — коэффициент корреляции не является статистически значимым $(H_0 : r_c = 0)$, а также альтернативная гипотеза — существует положительная корреляционная зависимость $(H_1 : r_c > 0)$. Для проверки гипотезы в качестве статистического критерия используется статистика t-Стьюдента с $\nu = n - 2$ степенями свободы: $t = |r_c|\sqrt{\frac{n-2}{1-r_c^2}}$. Выбирается уровень значимости α. По нему по табл. 2 приложения распределения Стьюдента находится квантиль $t_{\text{таб}} = t_{1-б}(n-2)$. Далее по формуле вычисляется наблюдаемое значение критерия $t_{\text{наб}}$.

Если $t_{\text{наб}} < t_{\text{таб}}$, то принимается нулевая гипотеза, в противном случае принимается альтернативная гипотеза и тогда связь между признаками есть.

Пример 1. Преподавателю и студенту было предложено расположить 10 профессий в порядке их общественной значимости.

Оценка преподавателя $R_1(i)$	Профессия	Оценка студента $R_2(i)$
3	Профессор	2
1	Врач	1
4	Учитель школы	7
2	Директор магазина	4
8	Бухгалтер	5

Окончание таблицы

Оценка преподавателя $R_1(i)$	Профессия	Оценка студента $R_2(i)$
6	Банкир	3
9	Водитель	9
5	Журналист	8
10	Диджей	10
7	Программист	6

Какова корреляция рангов между двумя рядами оценок? Одинаковое ли мнение имеют преподаватель и студент по этому вопросу?

Определим разности рангов, их квадраты и суммы:

$R_1(i)-R_2(i)$	1	0	–3	–2	3	3	0	–3	0	1	$\sum=0$
$(R_1(i)-R_2(i))^2$	1	0	9	4	9	9	0	9	0	1	$\sum=42$

Представим $r_c=1-\dfrac{6S}{n^3-n}=1-\dfrac{6\cdot 42}{1000-10}=0,7456.$ Проверим, существует ли положительная корреляционная связь между мнением преподавателя и мнением студента. При уровне значимости $б=0,05$ для односторонней критической области $t_{\text{кр}}=t_{0,95;\,8}=1{,}86$, $t_{\text{расч}}=0,7456\sqrt{\dfrac{10-2}{1-0,7456^2}}=2,8284.$ Итак, 2,8284 > 1,86 ($t_{\text{расч}}$ > $t_{\text{кр}}$). Следовательно, связь между мнением преподавателя и мнением студента является статистически значимой при 5 %-м уровне значимости.

М. Кендэл предложил другой коэффициент ранговой корреляции — τ. Он вычисляется по двумерной выборке рангов следующим образом.

Столбцы $\begin{pmatrix} R_1(i) \\ R_2(i) \end{pmatrix}$ переставляются так, чтобы ранги $R_1(i)$ образовали возрастающую последовательность 1, 2, …, *n*. Теперь $R_1(i)$ = *i*. Для каждого ранга $R_2(i)$ через p_i обозначим число рангов $R_2(k) > R_2(i)$, причем $k > i$, а через q_i обозначим число рангов $R_2(k) < R_2(i)$, причем $k > i$. Пусть $P = \sum_i p_i$, $Q = \sum_i q_i$, $S = P - Q$.

Коэффициент *ранговой корреляции Кендэла* (τ) вычисляется по одной из эквивалентных формул: $\tau = \frac{2S}{n\,(n-1)} = 1 - \frac{4Q}{n\,(n-1)} = \frac{4P}{n\,(n-1)} - 1$.

Нетрудно заметить, что $P + Q = \frac{n\,(n-1)}{2}$. Число τ лежит в пределах $-1 \le \tau \le 1$. Причем $\tau = 1$, если $R_1(i) = R_2(i)$, $i = 1, 2, \ldots, n$, и $\tau = -1$, если $R_1(i) + R_2(i) = n + 1$, $i = 1, 2, \ldots, n$.

Пример 2. Измерения длины головы и длины грудного плавника у десяти окуней дали следующие результаты:

Длина головы	66	61	67	73	51	59	48	47	45	44
Длина плавника	38	31	36	43	29	35	28	25	26	23

Следует вычислить коэффициент ранговой корреляции Кендэла. Для этого сначала определим ранги элементов этой выборки:

$R_1(i)$	8	7	9	10	5	6	4	3	2	1
$R_2(i)$	9	6	8	10	5	7	4	2	3	1

Вычисления для коэффициента τ приведены в следующей таблице:

Номер элемента выборки	Ранг		p_i	q_i
	$R_1(i)$	$R_2(i)$		
10	1	1	9	0
9	2	3	7	1
8	3	2	7	0
7	4	4	6	0
5	5	5	5	0
6	6	7	3	1
2	7	6	3	0
1	8	9	1	1
3	9	8	1	0
4	10	10	0	0
Итого	**55**	**55**	***P* = 42**	***Q* = 3**

Итак, $\tau = \frac{2 \cdot 39}{90} = 1 - \frac{3 \cdot 4}{90} = \frac{42 \cdot 4}{90} - 1 = 0{,}867.$

Коэффициенты Спирмена и Кендэла никак не связаны между собой.

Значимость коэффициента ранговой корреляции Кендэла при $n \geq 10$ проверяется при помощи статистики:

$$\lambda = \frac{\tau}{\sqrt{2\,(2n+5)\,/\,9n\,(n-1)}},$$

которую можно считать приближенно нормально распределенной со средним — нулем и дисперсией, равной единице.

Для нашего примера с уровнем значимости $\alpha = 0{,}05$ по табл. 4 приложения нормального распределения находим, что $\lambda_{\text{таб}} = 1{,}65$. Между тем $\lambda_{\text{эксп}} = \dfrac{0{,}867}{\sqrt{2\,(2 \cdot 10+5)\,/\,90\,(10-1)}} = 3{,}49$. Значение 3,49 > 1,65. Нулевую гипотезу следует отвергнуть, т. е. коэффициент корреляции Кендэла является значимым.

§ 3. Обработка таблиц сопряженности

Если измерения признаков проводятся в номинальной шкале, то ***выявление*** факта ***взаимозависимости признаков*** проводится ***с помощью таблиц сопряженности и критерия*** χ^2.

Допустим, что имеется два признака, измеряемых в номинальной шкале, и пусть первый признак может принимать M различных значений, а второй — K различных значений. Через N обозначим общее количество имеющихся наблюдений, а количество наблюдений (частоту) с одновременным появлением i-го значения первого признака и j-го значения второго признака обозначим как n_{ij}. Очевидно, что должно выполняться равенство $\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{K} n_{ij} = N$.

Частоты всех возможных пар значений рассматриваемых признаков принято записывать в виде *таблицы сопряженности*:

Признак 1	Признак 2			
	1	2	…	K
1	n_{11}	n_{12}	…	n_{1K}
2	n_{21}	n_{22}	…	n_{2K}
…	…	…	…	…
M	n_{M1}	n_{M2}	…	n_{MK}

Анализ таблиц сопряженности может состоять в сравнении наблюдаемых частот с частотами, рассчитанными в соответствии с некоторыми предположениями (гипотезами) о характере зависимости между признаками или ожидаемыми (теоретическими) частотами. Теоретические частоты рассчитываются в предположении об отсутствии зависимости между признаками.

Это делают следующим образом. Исходную таблицу заменяют таблицей, в которой добавляется сумма элементов по строке $б_i$, $i = \overline{1, M}$ и сумма элементов по столбцу $в_j$, $j = \overline{1, K}$.

Признак 1	Признак 2				
	1	2	…	K	
1	n_{11}	n_{12}	…	n_{1K}	α_1
2	n_{21}	n_{22}	…	n_{2K}	α_2
…	…	…	…	…	…
M	n_{M1}	n_{M2}	…	n_{MK}	α_M
…	ß$_1$	ß$_2$	…	ß$_K$	S

В таблице S — это общая сумма. Тогда теоретические частоты будем рассчитывать по формуле

$$\tilde{n}_{ij} = \frac{б_i в_j}{S}, \quad i = \overline{1, M}; \quad j = \overline{1, K}.$$

Степень соответствия наблюдаемых частот теоретическим можно определить с помощью *критерия* χ^2. Проверочная статистика рассчитывается на основе разностей между наблюдаемыми и теоретическими частотами:

$$\chi^2 = \sum_{i=1}^{M} \sum_{j=1}^{K} \frac{(n_{ij} - n_{ij})^2}{n_{ij}},$$

где n_{ij} — теоретическая частота одновременного появления i-го значения первого признака и j-го значения второго признака.

При справедливости сделанных предположений статистика χ^2 имеет распределение χ^2 («хи-квадрат») с $\nu = (K - 1)(M - 1)$ степенями свободы. Сделанные предположения отвергаются, если

$$\chi^2 > \chi^2_{кр}((1 - \alpha), \nu),$$

где $\chi^2_{кр}((1 - \alpha), \nu)$ — критическое значение, определяемое по табл. 5 приложения квантилей распределения χ^2.

В противном случае, если $\chi^2 \le \chi^2_{\text{кр}}((1-\alpha), \nu)$, то предположения не отвергаются и зависимости между признаками нет.

Пример 3. Пусть имеются результаты опроса тысячи покупателей прохладительных напитков. Каждого покупателя просили выбрать из двух типов прохладительных напитков («Фанта» и минеральная вода) один.

Результаты опроса с разбивкой по возрастным группам представим в виде таблицы сопряженности:

Покупатель	«Фанта»	Минеральная вода	Всего
Дети	254	124	378
Взрослые	167	130	297
Пенсионеры	123	202	325
Всего	544	456	1 000

Было выдвинуто следующее предположение (гипотеза): предпочтение «Фанты» или минеральной воды не зависит от возраста опрашиваемого. Если бы это предположение было справедливо, то ожидаемые частоты в таблице сопряженности совпали бы с частотами, вычисленными в следующей таблице:

Покупатель	«Фанта»	Минеральная вода	Всего
Дети	378 · 544 /1 000	378 · 456 /1 000	378
Взрослые	297 · 544 /1 000	297 · 456 /1 000	297
Пенсионеры	325 · 544 /1 000	325 · 456 /1 000	325
Всего	544	456	1 000

Для проверки выдвинутой гипотезы рассчитаем статистику $\chi^2 = \sum_{i=1}^{M} \sum_{j=1}^{K} \frac{(n_{ij} - n_{ij})^2}{n_{ij}} = 61,25.$ Число степеней свободы в этом случае ν = (2 – 1) (3 – 1) = 2. Соответствующее критическое значение при вероятности ошибки α = 0,05 (табл. 5 приложения) $\chi^2_{\text{кр}}(0,95;\, 2) = 5,99.$ Поскольку $\chi^2 = 61,25 > \chi^2_{\text{кр}}(0,95;\, 2) = 5,99,$ то выдвинутая гипотеза отвергается. В этом случае можно сделать вывод о наличии зависимости между возрастом покупателя и его предпочтением к «Фанте» или минеральной воде.

Контрольные вопросы и задания

1. Для чего проводится проверка значимости коэффициента корреляции?

2. Когда следует использовать ранговый коэффициент корреляции?

3. В каких случаях используется таблица сопряженности?

4. Что означает ранжирование?

5. В каких пределах изменяются ранговые коэффициенты корреляции?

6. Каковы свойства парного коэффициента корреляции?

7. Какие статистики используются для проверки значимости коэффициентов корреляции?

8. Чем вызвана необходимость построения ранговых коэффициентов корреляции?

9. Опишите меры связи номинальных признаков в таблицах сопряженности.

Задания для самостоятельного решения

1. С помощью таблицы сопряженности ответьте на вопрос: являются ли конфликтные ситуации» фактором гипертонической болезни.

Конфликтные ситуации на работе	Больные гипертонией	Здоровые	Всего
Есть	28	7	35
Нет	17	38	55
Итого	45	45	90

2. Используя данные в таблице сопряженности, оцените взаимосвязь между смертностью населения различных расовых групп и местом их рождения.

Раса	Количество умерших за 1 год, человек		Итого
	Европа	Африка	
Негроидная	1 050	600	1 650
Европеоидная	750	1 300	2 050
Итого	1 800	1 900	3 700

3. По мнению врачей, прием некоторого витамина как-то сказывается на профилактике простудных заболеваний.

Проведен следующий эксперимент: 200 человек случайным образом были разделены на две равные группы: одной группе дали витамин, другой — «пустышку», но всем 200 исследуемым было сказано, что им дали витамин.

Результаты обследования приведены в таблице:

Группа	Меньше простудных заболеваний	Больше простудных заболеваний	Без изменений
Контрольная	39	21	40
Принимавшая витамин	51	20	29

Проверьте на 5 %-м уровне значимости гипотезу о независимости простудных заболеваний от приема витамина.

4. Таблица показывает число сотрудников фирмы: мужчин и женщин с высшим и средним образованием. Содержимое некоторых ячеек неизвестно, но известно, что процент лиц с высшим образованием в общем числе сотрудников не зависит от пола сотрудника:

Образование	Мужчина	Женщина	Всего
Высшее		30	40
Среднее			60

Восстановите содержимое ячейки «мужчины со средним образованием».

5. Универмаг решил проанализировать сроки погашения кредита для различных категорий своих покупателей. Выборка, включающая 1 200 платежей, дала следующие результаты:

Время, сут	Категория покупателей		
	рабочие	священники	служащие
До 30	380	220	120
От 30 до 90	220	200	60

Есть ли существенная разница между отдельными категориями покупателей с точки зрения сроков погашения кредита?

6. Представлены результаты опроса общественного мнения с точки зрения поддержки четырех кандидатов избирателями южных и северных районов некоторой страны:

Район	Кандидат			
	1	2	3	4
Север	200	156	128	116
Юг	100	104	92	104

Имеется ли существенное различие в степени поддержки кандидатов избирателями каждого из регионов на 5 %-м уровне значимости?

7. Комплектующие изделия одного наименования поступают с трех предприятий (A, B и C).

Результаты проверки изделий приведены в следующей таблице:

Результат проверки	Поставщик		
	A	*B*	*C*
Годные	29	38	53
Негодные	1	2	7

Можно ли считать, что качество изделий не зависит от поставщика? Принять б = 0,1.

8. Утверждается, что результат действия лекарства зависит от способа его применения (A, B, C).

Проверьте это утверждение при $б = 0,05$ по следующим данным:

Результат	Способ применения		
	A	*B*	*C*
Неблагоприятный	11	17	16
Благоприятный	20	23	19

9. Отношение зрителей к включению одной из телепередач в программу выразилось следующими данными:

Пол	Отношение		
	положительное	безразличное	отрицательное
Мужчины	14	24	2
Женщины	29	36	15

Можно ли считать, что отношение к включению данной передачи в программу не зависит от пола зрителя?

10. Изменение производительности труда на предприятии при проведении мероприятий A, B и C выражается следующими данными:

Производительность	Мероприятие		
	A	*B*	*C*
Увеличилась	14	47	16
Не изменилась	22	37	7
Уменьшилась	20	25	2

Можно ли считать, что проведение этих мероприятий не влияет на производительность труда? Принять $б = 0,1$.

11. Представлены результаты опроса 100 студентов первых трех курсов, которые отвечали на вопрос: считаете ли вы, что курение мешает учебе.

Ответ	Курс		
	I	II	III
Нет	15	10	—
Не знаю	8	5	7
Да	—	30	25

Подтверждают ли эти данные предположение о том, что отношение к курению студентов разных курсов различно? Принять $б = 0,01$.

12. Для определения зависимости цвета волос жителей от их местожительства были обследованы три группы людей из районов *A*, *B* и *C*:

Район	Цвет волос		
	рыжий	светлый	темный
A	2	9	9
B	3	6	21
C	15	15	20

Свидетельствуют ли приводимые результаты обследования о зависимости цвета волос жителей от их местожительства? Принять б = 0,05.

13. Содержание никотина (мг) для двух марок сигарет характеризуется следующими данными:

Марка сигарет	Содержание никотина			
A	24	26	25	22
B	27	28	25	29

Указывают ли эти результаты на различие в содержании никотина в сигаретах марок *A* и *B*? Принять б = 0,1.

14. Исследователь хочет сравнить выживаемость при различных типах шока по таблице сопряженности при б = 0,01.

Тип шока	Исход	
	выжили	умерли
Гиповолемический	7	8
Кардиогенный	11	11
Неврогенный	10	6
Септический	9	7
Эндокринный	3	5

15. С помощью критерия χ^2 («хи-квадрат») оцените значимость связи признаков «завод» и «отрасль» с видом заболеваемости:

Вид заболевания	Отрасль 1		Отрасль 2	Всего
	Электросигнал	Экран	СЭТМ	
Гипертония	83	102	53	238
Остеохондроз	60	108	34	202
Болезни желудочно-кишечного тракта	58	84	106	248
Итого	201	294	193	688

Сделайте выводы.

16. Найдите выборочные коэффициенты корреляции Кендэла и Спирмена. Проверьте значимость полученного результата при $б = 0,05$:

Ранг оценки	Студент								
	1	2	3	4	5	6	7	8	9
По математике	9	3	1	4	2	8	5	6	7
По истории	6	7	3	2	1	8	5	4	9

17. По данным об объеме строительно-монтажных работ, выполненных собственными силами, и численности работающих в 10 строительных компаниях одного из городов определите зависимость между этими признаками с помощью коэффициентов Кендела и Спирмена:

Номер строительной компании	Объем работ, тыс. руб.	Численность работающих, человек
1	3 998	66
2	2 821	71
3	4 121	73
4	3 583	59
5	3 646	52
6	3 008	50
7	3 973	61
8	2 973	70
9	2 911	38
10	3 114	54

18. Эксперты оценивали вкусовые качества разных вин. Получены следующие суммарные оценки:

Марка вина	Оценка, балл	Цена, усл. ед.
1	11	1,57
2	14	1,60
3	17	2,00
4	15	2,10
5	13	1,70
6	13	1,85
7	18	1,80
8	10	1,15
9	19	2,30
10	25	2,40

Определите зависимость между оценкой и ценой с помощью коэффициентов Кендела и Спирмена.

19. На конкурсе красоты 12 участниц были проранжированы по двум признакам: X — артистизм, Y — красота.

Ранг X	3	11	4	10	1	8	9	2	12	6	7	5
Ранг Y	4	11	1	12	6	2	10	5	9	7	8	3

Найдите выборочные коэффициенты ранговой корреляции Спирмена и Кендела. Проверьте значимость полученного результата при $б = 0,05$.

20. Рейтинг девяти банков был оценен тремя экспертами. Данные о рейтинге приведены в следующей таблице:

Эксперт	Номер банка								
	1	2	3	4	5	6	7	8	9
1	3	2	1	4	5	6	7	8	9
2	2	3	1	4	7	9	8	5	6
3	1	2	5	3	4	6	9	7	8

С помощью коэффициента ранговой корреляции найдите пары экспертов, оценки которых наиболее близко соответствуют друг другу. Оцените значимость различий в оценке рейтинга банков экспертами.

Библиографический список

Основная литература

1. *Кремер, Н. Ш.* Эконометрика : учеб. для вузов / Н. Ш. Кремер, Б. А. Путко. — Москва : ЮНИТИ, 2008. — 310 с.

2. *Осипов, А. Л.* Экономико-математические методы в управлении : учеб.-метод. комплекс для дистанц. обучения / А. Л. Осипов, Е. А. Рапоцевич ; СибАГС. — Новосибирск : Изд-во СибАГС, 2006. — 144 с.

3. *Практикум* по эконометрике : учеб. пособие / И. И. Елисеева, С. В. Курышева, Н. М. Гордеенко и др. ; под ред. И. И. Елисеевой. — Москва : Финансы и статистика, 2001. — 192 с.

4. *Тюрин, Ю. Н.* Анализ данных на компьютере : учеб. пособие / Ю. Н. Тюрин, А. А. Макаров. — Москва : ФОРУМ, 2008. — 366 с.

5. *Эконометрика* : учеб. для вузов / [И. И. Елисеева и др.] ; под ред. И. И. Елисеевой. — Москва : Финансы и статистика, 2007. — 575 с.

6. *Экономико*-математические методы в управлении : практикум / СибАГС ; сост. : А. Л. Осипов, Е. А. Рапоцевич. — Новосибирск : Изд-во СибАГС, 2007. — С. 5—27.

Дополнительная литература

1. *Большаков, А. А.* Методы обработки многомерных данных и временных рядов : учеб. пособие / А. А. Большаков, Р. Н. Каримов. — Москва : Горячая линия — Телеком, 2007. — 522 с.

2. *Статистика* : учеб. / В. Г. Минашкин [и др.] ; под ред. В. Г. Минашкина. — Москва : ТК Велби : Проспект, 2006. — 272 с.

3. *Тимофеев, В. С.* Эконометрика : учеб. / В. С. Тимофеев, А. В. Фаддеенков, В. Ю. Щеколдин. — Новосибирск : Изд-во НГТУ, 2013. — 340 с.

4. *Фаддеев, М. А.* Элементарная обработка результатов эксперимента : учеб. пособие / М. А. Фаддеев. — Санкт-Петербург : Лань, 2008. — 117 с.

5. *Яновский, Л. П.* Введение в эконометрику : учеб. пособие / Л. П. Яновский, А. Г. Буховец. — Москва : КноРус, 2007. — 255 с.

Часть 3. ИГРОВЫЕ МЕТОДЫ

Глава 7. Матричные игры

§ 1. Элементы теории игр

Предметом теории игр являются ситуации, в которых важную роль играют конфликты и совместные действия.

В отличие от линейного программирования в теории игр рассматриваются ситуации, в которых решения принимают несколько действующих лиц.

Ситуация называется ***конфликтной***, если в ней участвуют стороны, интересы которых полностью или частично противоположны.

Классическим примером конфликта является ситуация, в которой есть один продавец и один покупатель. Другой пример конфликтной ситуации — это существование на рынке нескольких производителей одного и того же товара, обладающих достаточной силой для воздействия на цену этого товара (ситуация олигополии).

Конфликт может проявиться не только в результате сознательных действий различных участников, но и как результат действия «стихийных сил». Подобные ситуации называют «играми с природой».

Лица, принимающие участие в рассматриваемом конфликте, называются ***игроками***. Игроки могут различаться по своим именам (например, Продавец и Покупатель) или по номерам в случае конечного числа игроков (например, 1-й и 2-й игрок при игре в «орлянку»).

Любая игра ведется по определенным правилам. Правилами игры называются допустимые действия игроков, направленные на достижение ими некоторой цели. В соответствии с этими правилами игроки предпринимают некоторые действия, делают ходы. Однозначное описание действий игрока в каждой из возможных ситуаций, в которой он должен сделать личный ход, называется ***стратегией игрока***.

Количественная оценка результатов игры называется ***платежом***, или ***выигрышем***, того или иного игрока. Соответствие между возможным набором возникающих в игре ситуаций и выигрышем игрока в каждой из этих ситуаций называется ***функцией выигрыша*** или ***платежной функцией***.

Итак, формальное описание игры должно включать:

1) перечень игроков;

2) список стратегий каждого игрока;

3) описание функций выигрыша каждого игрока.

Игры можно классифицировать по видам, основываясь на разных принципах: по числу игроков, по числу стратегий, по свойствам функций выигрыша, по возможности предварительных переговоров и взаимодействий между игроками в ходе игры.

В зависимости от числа игроков выделяют игры двух, трех и т. д. лиц. Возможны игры с бесконечным количеством игроков.

По количеству стратегий выделяют конечные и бесконечные игры. В конечных играх игроки располагают конечным числом стратегий (например, при игре в «орлянку» игроки могут выбрать «орла» или «решку», т. е. они имеют по две стратегии). Стратегии в конечных играх часто называют ***чистыми стратегиями*** (смысл этого специального термина будет понятен позже). В бесконечных играх игроки имеют бесконечное число возможных стратегий. Например, в игре «Продавец — Покупатель» каждый из них может назвать любую устраивающую его цену и количество продаваемого или покупаемого товара.

С точки зрения свойств платежных функций важным классом являются так называемые ***антагонистические игры***, или ***игры с нулевой суммой***. Игра называется игрой с нулевой суммой, если выигрыш одного игрока равен проигрышу другого. Возможны также ситуации, когда игроки и выигрывают, и проигрывают одновременно, так что им выгодно действовать сообща. Такие игры называются ***играми с постоянной разностью***. Кроме того, существуют ***игры с ненулевой суммой***, в которых есть и конфликты, и совместные действия игроков.

В дальнейшем мы будем рассматривать конечные игры двух лиц с нулевой суммой.

Пусть есть два игрока, первый из которых (игрок A) имеет m возможных чистых стратегий, а второй (игрок B) — n чистых стратегий. Пусть при выборе первым игроком некоторой i-й стратегии $(i=\overline{1,m})$, а вторым игроком j–й стратегии $(j=\overline{1,n})$ первый игрок получает выигрыш, равный a_{ij}, а второй игрок эту величину проигрывает. Заметим, что в данном предположении нет какой-либо дискриминации по отношению ко второму игроку. Если величина a_{ij} отрицательна, то первый игрок получает отрицательный выигрыш, т. е. проигрывает, а второй игрок при этом выигрывает. Из величин a_{ij} составим матрицу, строки которой будут соответствовать чистым стратегиям первого игрока, а столбцы — чистым стратегиям второго игрока.

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Матрица $A=\left\|a_{ij}\right\|,\ i=\overline{1,m};\ j=\overline{1,n}$ называется *платежной матрицей*. Используя подобное представление конечных игр двух лиц с нулевой суммой, их часто называют ***матричными играми***.

Итак, матричная игра характеризуется следующими позициями:

1) в ней участвует два игрока;

2) сколько один выигрывает, столько проигрывает другой игрок;

3) каждый игрок имеет конечное число стратегий.

Цель игры — это поиск оптимальной стратегии для каждого игрока.

Оптимальной называется ***стратегия***, которая при многократном повторении игры обеспечивает первому игроку максимально возможный выигрыш, а второму игроку — минимально возможный проигрыш.

При каждом ходе первого игрока второй игрок выбирает стратегию, минимизирующую выигрыш первого игрока. Соответственно при каждом ходе второго игрока первый игрок выбирает стратегию, максимизирующую проигрыш второго игрока. На этом основаны принципы минимакса и максимина.

Введем числа $б_i = \min\limits_j a_{ij}$.

Число $б_i$ представляет собой гарантированный уровень стратегии *i*, т. е. выбрав стратегию *i*, первый игрок имеет полную гарантию того, что его выигрыш независимо от действий второго игрока будет не меньше чем $б_i$. Тогда наибольший гарантированный выигрыш первого игрока будет равен $v_1 = \max\limits_i б_{ij} = \max\limits_i \min\limits_j a_{ij}$.

Число v_1 называется ***нижней ценой игры*** (или ***максимином***), а стратегия, максимизирующая минимальный возможный выигрыш игрока 1, т. е. стратегия i^*, для которой $б_{i^*} = v_1$, — ***максиминной стратегией*** игрока 1.

Введем числа $в_i = \max\limits_i a_{ij}$.

Число $в_i$ представляет собой гарантированный уровень стратегии *j*, т. е. выбрав стратегию *j*, второй игрок имеет полную гарантию, что его проигрыш независимо от действий первого игрока будет не больше чем $в_i$. Тогда наименьший гарантированный проигрыш $v_2 = \min\limits_j в_j$.

Стратегия j^*, которая минимизирует максимальные возможные потери игрока 2, называется его ***минимаксной стратегией***, а число $v_2 = \min\limits_j в_j = \min\limits_j \max\limits_i a_{ij}$ — ***верхней ценой игры*** (или ***минимаксом***).

Правило. В матричной игре нижняя цена игры не превосходит верхней цены: $v_1 \le v_2$. Если нижняя и верхняя цена игры совпадает, то число $v = v_1 = v_2$ называется ценой игры.

Действительно, при любых $i = \overline{1, m}$; $j = \overline{1, n}$ выполняется условие $б_i \le a_{ij}$ и $в_j \ge a_{ij}$, откуда $б_i \le в_j$. Пусть i^* — максиминная стратегия игрока 1, j^* — минимаксная стратегия игрока 2. Тогда $v_1 = б_{i^*} \le в_{j^*} = v_2$.

Практическое *правило* для нахождения нижней и верхней цены матричной игры состоит в следующем. Добавим к платежной матрице *A* столбец $б_i \; {}_{i=\overline{1,m}}$ строчных минимумов и строку $в_j \; {}_{j=\overline{1,n}}$ столбцовых максимумов. Наибольшее из чисел $б_i$ будет нижней ценой, а наименьшее из чисел $в_j$ — верхней ценой. Этот способ продемонстририруем на примерах.

Пример 1. Исходя из заданной платежной матрицы необходимо определить нижнюю и верхнюю цену игры:

	B_1	B_2	B_3	б$_i$
A_1	1	9	3	1
A_2	5	3	4	3
A_3	4	2	8	2
в$_j$	5	9	8	$v_1 = 3$ $v_2 = 5$

В данной игре $v_1 < v_2$.

Пример 2. Исходя из заданной платежной матрицы следует определить нижнюю и верхнюю цену игры.

	B_1	B_2	B_3	B_4	б$_i$
A_1	4	5	9	3	3
A_2	7	6	8	9	6
A_3	8	2	9	6	2
в$_j$	8	6	9	9	$v_1 = 6$ $v_2 = 6$

В этой игре выбор ситуации (2, 2), первой компонентой которой является максиминная стратегия игрока 1, а второй — минимаксная стратегия игрока 2, устойчив, так как ни одному из игроков невыгодно одностороннее отклонение от нее. Учитывая, что данная игра в отличие от предыдущей имеет цену, можно связать наличие устойчивых ситуаций в матричной игре с наличием у нее цены.

Определение. Пусть $Г_A$ — матричная игра с платежной матрицей $A = \left\| a_{ij} \right\|$. Ситуация i_0, j_0 называется ***седловой точкой*** в игре $Г_A$, если при всех $i = \overline{1, m}$; $j = \overline{1, n}$ выполняется двойное неравенство: $a_{ij_0} \le a_{i_0 j_0} \le a_{i_0 j}$.

Правило. Ситуация i_0, j_0 — седловая точка игры $Г_A$ тогда и только тогда, когда элемент $a_{i_0 j_0}$ является одновременно наименьшим элементом своей строки и наибольшим элементом своего столбца.

Ясно, что устойчивые в матричной игре ситуации — это и есть ее седловые точки: одностороннее отклонение от седловой точки невыгодно ни одному из игроков.

Теорема (связь седловой точки с ценой игры)

1. Пусть i_0, j_0 — седловая точка игры $Г_A$.

Тогда:

а) i_0 является максиминной стратегией игрока 1;

б) j_0 является минимаксной стратегией игрока 2;

в) игра $Г_A$ имеет цену, причем исход в седловой точке совпадает с ценой игры: $a_{i_0 j_0} = v$.

2. Пусть игра $Г_A$ имеет цену v. Тогда любая ситуация i_0, j_0 , где i_0 — максиминная стратегия игрока 1, а j_0 — минимаксная стратегия игрока 2, является седловой точкой игры $Г_A$.

Отметим некоторые следствия теоремы.

Следствие 1. В матричной игре наличие седловой точки эквивалентно наличию цены игры.

Следствие 2. Исходы в седловых точках матричной игры совпадают между собой и совпадают с ценой игры.

Таким образом, если матричная игра имеет цену, то в ней реализуются два принципа оптимальности: принцип оптимальности стратегий (в качестве оптимальных ситуаций выступают максиминные стратегии игрока 1 и минимаксные стратегии игрока 2) и принцип оптимальности ситуаций (в качестве оптимальных ситуаций выступают седловые точки). При этом оба принципа оптимальности оказываются согласованными между собой: если игроки выбирают оптимальные стратегии, то возникающая ситуация является оптимальной (т. е. седловой точкой); другими словами, компонентами любой седловой точки служат оптимальные стратегии игроков. Кроме того, исход в любой седловой точке равен цене игры.

К сожалению, не всякая матричная игра имеет цену, что является главным препятствием для рекомендации применения указанных принципов оптимальности в приложениях теории игр. Однако существует общий метод, который позволяет в рамках матричной игры обеспечить цены в некотором обобщенном смысле. Этот метод, состоящий в переходе к смешанным стратегиям, впервые был обоснован в теории игр фон Нейманом.

§ 2. Графическое решение матричных игр

Если матричная игра не имеет седловой точки, то у нее нет решения в чистых стратегиях. В этом случае ее решение нужно искать в так называемых ***смешанных стратегиях***, которые определяются следующим образом.

Предположим, что игра является многоходовой, т. е. состоит из многих партий. Тогда каждый игрок будет в разных партиях применять свои разные чистые стратегии, причем одни из них чаще, а другие реже.

Смешанной стратегией первого игрока называется вектор $p = (p_1, ..., p_m)$, каждая компонента $p_i (i = \overline{1, m})$ которого — это вероятность применения первым игроком его i-й чистой стратегии в многоходовой игре.

Из определения смешанной стратегии следует, что компоненты вектора p удовлетворяют следующим соотношениям:

$$\sum_{i=1}^{m} p_i = 1, \;\; p_i \geq 0, \;\; i = \overline{1, m}.$$

Аналогично определяется смешанная стратегия второго игрока. Смешанной стратегией второго игрока называется вектор $q = (q_1, ..., q_n)$, каждая компонента $q_j \ (j = \overline{1, n})$ которого — это вероятность применения вторым игроком его j-й чистой стратегии в многоходовой игре. Компоненты вектора q удовлетворяют условиям

$$\sum_{j=1}^{n} q_j = 1, \quad q_j \geq 0, \quad j = \overline{1, n}.$$

Заметим, что чистые стратегии игроков являются частным случаем его смешанных стратегий. Например, i-я чистая стратегия игрока может быть представлена в виде такой смешанной стратегии, у которой i-я компонента равна единице, а остальные — нулю.

Если игроки используют свои смешанные стратегии, то функция

$$E(p, q) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} p_i q_j$$

называется ***платежной***.

Оптимальная смешанная стратегия первого игрока p^* — это решение следующей задачи:

$$б^* = \max_p \min_q E(p, q).$$

Оптимальная смешанная стратегия второго игрока q^* — это решение следующей задачи:

$$в^* = \min_q \max_p E(p, q).$$

Если $p^* = (p_1^*, ..., p_m^*)$ и $q^* = (q_1^*, ..., q_n^*)$ — оптимальные смешанные стратегии игроков, то число

$$\upsilon = E(p^*, q^*) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} p_i^* q_j^*$$

называется ***ценой игры***.

Рассмотрим *теорему фон Неймана*.

Для любой матричной игры $Г_A$ выполняется соотношение $\max\limits_p \min\limits_q E(p, q) =$ $= \min\limits_q \max\limits_p E(p, q)$.

Другими словами, в смешанном расширении любой матричной игры нижняя и верхняя цена игры совпадает.

Существует несколько простых правил, позволяющих упростить нахождение решения матричной игры. Одно из них — это *правило отбрасывания недоминируемых стратегий*. При нахождении решения матричной игры в смешанных стратегиях доминируемые стратегии игроков могут быть отброшены.

В игре с платежной матрицей $A = \left\| a_{ij} \right\|$, $i = \overline{1,m}$; $j = \overline{1,n}$ для стратегий i_1, i_2 игрока 1 условие доминирования означает, что при всех $j = \overline{1,n}$ выполняется $a_{i_1 j} \geq a_{i_2 j}$, а для стратегий j_1, j_2 игрока 2 условие доминирования означает, что при всех $i = \overline{1,m}$ $a_{ij_1} \leq a_{ij_2}$.

Применение этого правила позволяет уменьшить размерность матрицы игры.

Пример 3. Пусть задана матрица игры

$$A = \begin{pmatrix} 4 & 2 & 3 & -1 \\ -4 & 2 & 2 & 2 \end{pmatrix}.$$

У первого игрока две стратегии, у второго — четыре. Если сравнить вторую и четвертую стратегии второго игрока, то видно, что вторая заведомо хуже и ее можно вычеркнуть. Аналогично, сравнивая третью и четвертую стратегии, вычеркиваем третью. То есть для второго игрока, если все элементы столбика *не меньше* элементов другого столбика, то эта стратегия заведомо хуже и вычеркивается. Аналогично для первого игрока сравниваются строчки и худшая имеет элементы *не больше* соответствующих элементов сравниваемой строки. В результате получим матрицу игры $A' = \begin{pmatrix} 4 & -1 \\ -4 & 2 \end{pmatrix}$.

Пример 4. Пусть задана матрица игры

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 2 & 0 & 2 & 1 & 1 \\ 4 & 3 & 1 & 3 & 4 & 4 \\ 4 & 3 & 7 & -5 & 1 & 2 \\ 4 & 3 & 4 & -1 & 2 & 2 \\ 4 & 3 & 3 & -2 & 2 & 2 \end{pmatrix}.$$

Вычеркнем заведомо худшие стратегии у первого и второго игроков.

$$A=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 2 & 0 & 2 & 1 & 1 \\ 4 & 3 & 1 & 3 & 4 & 4 \\ 4 & 3 & 7 & -5 & 1 & 2 \\ 4 & 3 & 4 & -1 & 2 & 2 \\ 4 & 3 & 3 & -2 & 2 & 2 \end{pmatrix}.$$

В результате получим матрицу игры $A'=\begin{pmatrix} 1 & 3 \\ 7 & -5 \\ 4 & -1 \end{pmatrix}$, которую можно решить графическим методом.

Графический метод применяется в том случае, когда одна из размерностей матрицы игры равна двум.

Поясним его действие на следующих примерах.

Пример 5. Найдем оптимальные стратегии игроков в игре, заданной платежной матрицей

$$A=\begin{pmatrix} 4 & 9 & 5 & 6 \\ 7 & 3 & 9 & 9 \end{pmatrix}.$$

Сначала проверим, есть ли в данной игре седловая точка. Нижняя цена игры равна $б=\max\ \{4,\ 3\}$ = 4. Верхняя цена этой игры равна $в=\min\ \{7,\ 9,\ 9,\ 9\}$ = 7. Поскольку $б\neq в$, то седловой точки у данной игры нет и решение нужно искать в смешанных стратегиях.

Графически решаются те матричные игры, в которых хотя бы у одного из игроков есть лишь две чистые стратегии. Задача именно этого игрока и решается графически. В задаче у первого игрока две чистых стратегии, а у второго — четыре, поэтому будем решать графически задачу первого игрока.

Смешанная стратегия первого игрока задается вектором $p=(p_1, p_2)$. Построения осуществляются следующим образом. На горизонтальной прямой откладывается отрезок единичной длины, характеризующий вероятность применения чистых стратегий первым игроком. Каждой точке этого отрезка сопоставляется смешанная стратегия первого игрока по следующему *правилу*: расстояние от точки до правого конца отрезка задает величину p_1, а расстояние до левого его конца — величину p_2 (рис. 3). Для определенных таким образом величин p_1 и p_2 выполняются соотношения $p_1+p_2=1$, $p_1\geq 0$, $p_2\geq 0$, поэтому вектор $p=(p_1, p_2)$ задает смешанные стратегии первого игрока.

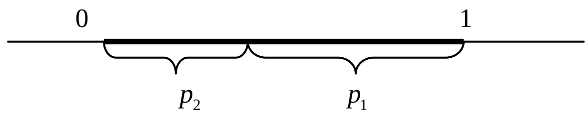


Рис. 3. Построение смешанных стратегий первого игрока

Тогда точка 0 задает вектор (1, 0), т. е. первую чистую стратегию первого игрока, а точка 1 задает вектор (0, 1), т. е. вторую чистую стратегию первого игрока. Далее через концы единичного отрезка проводятся вертикальные линии. На этих линиях откладываются выигрыши первого игрока при применении вторым игроком его различных чистых стратегий. При этом выигрыши в случае применения первым игроком его первой чистой стратегии располагаются на левой вертикальной линии, а соответствующие второй чистой стратегии первого игрока — на правой вертикали. Точки левой и правой вертикали, соответствующие одной и той же чистой стратегии второго игрока, соединяются отрезками (рис. 4).

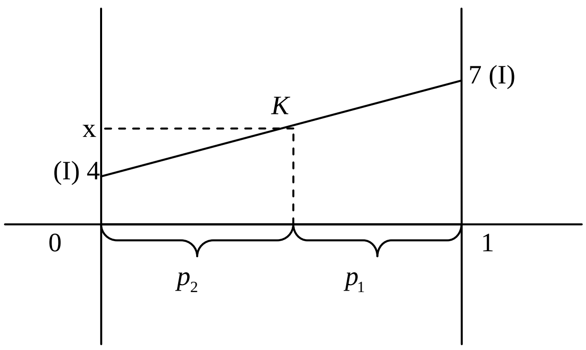


Рис. 4. Интерпретация построений при графическом решении матричной игры

На рис. 4 изображены выигрыши первого игрока при применении вторым игроком первой чистой стратегии. Римские цифры указывают, что второй игрок применяет именно первую чистую стратегию.

Любая точка *K* этого отрезка с координатами p_1, p_2 и x показывает, что если первый игрок будет применять свою смешанную стратегию (p_1, p_2), а второй игрок — свою первую чистую стратегию, то средний выигрыш первого игрока будет равен x.

Аналогичные построения выполняются для остальных чистых стратегий второго игрока (рис. 5). Римские цифры указывают на номер его чистой стратегии.

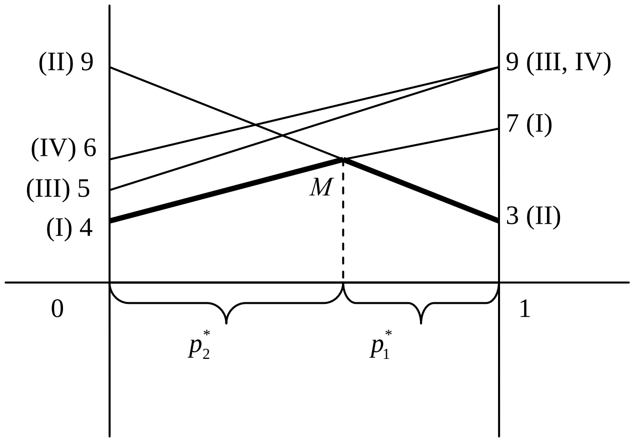


Рис. 5. Графическое решение задачи первого игрока

Жирным шрифтом выделена ломаная, соответствующая нижней границе выигрыша первого игрока, т. е. дающая его средний гарантированный выигрыш. В точке *M* находится наибольший гарантированный выигрыш первого игрока (так как *M* — наивысшая точка ломаной). Эта точка является пересечением отрезков, соответствующих первой и второй чистым стратегиям второго игрока. Эти ***стратегии*** называются ***активными***. Второй игрок будет использовать их в своей оптимальной смешанной стратегии с ненулевой вероятностью. Отрезки, соответствующие третьей и четвертой чистым стратегиям второго игрока, не проходят через точку *M*, поэтому эти стратегии в оптимальную смешанную стратегию второго игрока войдут с нулевыми вероятностями, так как их реализация приведет к большему проигрышу второго игрока. Такие ***стратегии*** называют ***пассивными***.

Обозначим смешанную стратегию первого игрока через $S_A = \begin{Bmatrix} A_1 & A_2 \\ p_1 & p_2 \end{Bmatrix}$. Пусть $p_2 = x$, тогда $p_1 = 1 - x$.

Графическое решение задачи состоит в нахождении уравнений прямых, соответствующих активным стратегиям, и определении точки их пересечения.

Найдем уравнение первой прямой:

$$4\,(1 - x) + 7x = 4 + 3x.$$

Найдем уравнение второй прямой:

$$9\,(1 - x) + 3x = 9 - 6x.$$

Приравняем эти уравнения и найдем значение $x=\frac{5}{9}$. Тогда смешанная стратегия первого игрока будет $S_A=\begin{Bmatrix} A_1 & A_2 \\ \frac{4}{9} & \frac{5}{9} \end{Bmatrix}$. Для того чтобы найти цену игры, подставим найденное значение *x* в любое из уравнений: $\text{x}=4+3\frac{5}{9}=\frac{17}{3}$.

Для нахождения смешанной стратегии для второго игрока, учитывая то, что первые две активные, имеем: $S_B=\begin{Bmatrix} B_1 & B_2 & B_3 & B_4 \\ q_1 & q_2 & 0 & 0 \end{Bmatrix}$. Вероятность использования стратегии B_2 можно вычислить по формуле

$$q_2=\frac{-(a_{21}-a_{11})}{(a_{22}-a_{12})-(a_{21}-a_{11})}=\frac{-(7-4)}{(3-9)-(7-4)}=\frac{1}{3}.$$

Следовательно, $q_1=1-q_2=\frac{2}{3}$. В итоге $S_B=\begin{Bmatrix} B_1 & B_2 & B_3 & B_4 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \end{Bmatrix}$.

Если активные стратегии игрока *B* имеют номера j_1 и j_2 соответственно, то общая формула расчета вероятности будет иметь вид

$$q_{j_1}=\frac{-(a_{2j_2}-a_{1j_2})}{(a_{2j_1}-a_{1j_1})-(a_{2j_2}-a_{1j_2})}.$$

Решение игры в смешанных стратегиях можно найти и другим способом. Для этого при определении оптимальной смешанной стратегии первого игрока и цены игры будем использовать следующее утверждение.

Утверждение 1. Если один из игроков применяет свои оптимальные смешанные стратегии, то его выигрыш будет равен цене игры независимо от того, с какими вероятностями применяет другой игрок свои стратегии, вошедшие в оптимальную смешанную стратегию.

Платежная функция в примере 1 имеет вид

$$E(p,q)=4p_1q_1+9p_1q_2+5p_1q_3+6p_1q_4+7p_2q_1+3p_2q_2+9p_2q_3+9p_2q_4.$$

Пусть оптимальная смешанная стратегия первого игрока равна $p^* = (p_1^*, p_2^*)$. По утверждению он получит выигрыш, равный цене игры х, с какими бы вероятностями не применял второй игрок свои активные стратегии. Мы рассмотрим случаи, когда второй игрок применяет свои активные чистые стратегии, т. е. либо первую, либо вторую.

Итак, пусть $p = p^*$, $q = (1, 0, 0, 0)$. Тогда $E(p,q) = 4p_1^* + 7p_2^*$.

Пусть $p = p^*$, $q = (0, 1, 0, 0)$. Тогда $E(p,q) = 9p_1^* + 3p_2^*$. Приравняем эти значения к цене игры х и добавим уравнение $p_1^* + p_2^* = 1$. Получим систему уравнений:

$$\begin{cases} 4p_1^* + 7p_2^* = \text{х}, \\ 9p_1^* + 3p_2^* = \text{х}, \\ p_1^* + p_2^* = 1. \end{cases}$$

Решив эту систему, получим $p_1^* = \frac{4}{9}$, $p_2^* = \frac{5}{9}$, $\text{х} = \frac{17}{3}$.

Далее найдем оптимальную смешанную стратегию второго игрока. Пусть она задается вектором $q^* = (q_1^*, q_2^*, 0, 0)$. Здесь учитывается тот факт, что третья и четвертая чистые стратегии второго игрока являются пассивными. Выпишем величину проигрыша второго игрока, если он применяет свою оптимальную смешанную стратегию, а первый игрок — свои чистые стратегии.

Пусть $p = (1, 0)$, $q = q^* = (q_1^*, q_2^*, 0, 0)$. Тогда $E(p,q) = 4q_1^* + 9q_2^*$.

Пусть $p = (0, 1)$, $q = q^* = (q_1^*, q_2^*, 0, 0)$. Тогда $E(p,q) = 7q_1^* + 3q_2^*$.

Применяем утверждение 1, учитывая, что цена игры найдена и равна $\text{х} = \frac{17}{3}$, получим систему уравнений:

$$\begin{cases} 4q_1^* + 9q_2^* = \dfrac{17}{3}, \\ 7q_1^* + 3q_2^* = \dfrac{17}{3}. \end{cases}$$

Решив эту систему, получим $q_1^* = \frac{2}{3}$, $q_2^* = \frac{1}{3}$. Итак, оптимальная смешанная стратегия второго игрока задается вектором $q^* = (\frac{2}{3}, \frac{1}{3}, 0, 0)$.

Ответ к данной задаче запишем в следующем виде:

$$p^* = \left(\frac{4}{9}, \frac{5}{9}\right), \quad q^* = \left(\frac{2}{3}, \frac{1}{3}, 0, 0\right), \quad \text{х} = \frac{17}{3}.$$

Пример 6. Рассмотрим матричную игру, платежная матрица которой является транспонированной к матрице предыдущей задачи, т. е. игра задается матрицей

$$A=\begin{pmatrix} 4 & 7 \\ 9 & 3 \\ 5 & 9 \\ 6 & 9 \end{pmatrix}.$$

В новой игре первый игрок имеет четыре чистые стратегии, второй — две. Нижняя цена игры $б = \max\{4, 3, 5, 6\} = 6,$ а верхняя цена игры $в = \min\{9, 9\} = 9$. Так как $б \neq в$, то у этой игры нет седловой точки, поэтому ее решение нужно искать в смешанных стратегиях.

Пусть $p = (p_1, p_2, p_3, p_4)$ — вектор смешанных стратегий первого игрока, а $q = (q_1, q_2)$ — вектор смешанных стратегий второго игрока. Платежная функция данной игры равна:

$$E(p,q) = 4p_1q_1 + 7p_1q_2 + 9p_2q_1 + 3p_2q_2 + 5p_3q_1 + 9p_3q_2 + 6p_4q_1 + 9p_4q_2.$$

Второй игрок имеет две чистые стратегии, поэтому *графически* будет решаться задача второго игрока. Построения выполняются аналогично предыдущему примеру, если поменять местами первого и второго игроков (рис. 6).

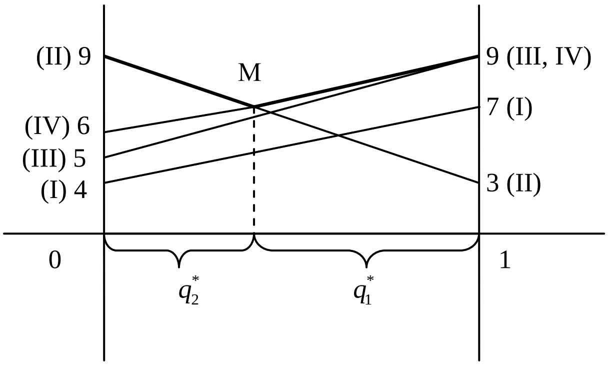


Рис. 6. Графическое решение задачи второго игрока

Цель второго игрока, согласно его осторожному поведению, состоит в минимизации его возможного риска. Риск второго игрока (т. е. максимально возможный проигрыш второго игрока при применении им той или иной смешанной стратегии) показан жирной линией (см. рис. 6). Точка *M* обозначает минимальный риск второго игрока. Она лежит на пересечении отрезков, соответствующих второй и четвертой чистым стратегиям первого игрока.

То есть, решая задачу графически, мы должны найти нижнюю точку (точка *M*) верхней огибающей на построенном рисунке (см. рис. 6). Из этих соображений определяются активные стратегии первого игрока (вторая и четвертая).

Найдем решение игры вторым способом. Обозначим оптимальную смешанную стратегию второго игрока: $q^* = (q_1^*, q_2^*)$. Для нахождения значений q_1^* и q_2^* воспользуемся утверждением 1. Активными стратегиями первого игрока являются вторая и четвертая.

Тогда:

1) $q = q^*$, $p = (0, 1, 0, 0)$ => $E(p, q) = 9q_1^* + 3q_2^*$;

2) $q = q^*$, $p = (0, 0, 0, 1)$ => $E(p, q) = 6q_1^* + 9q_2^*$.

Приравняем эти значения к цене игры x и добавим уравнение $q_1^* + q_2^* = 1$, получим систему уравнений

$$\begin{cases} 9q_1^* + 3q_2^* = \text{x}, \\ 6q_1^* + 9q_2^* = \text{x}, \\ q_1^* + q_2^* = 1. \end{cases}$$

Решив эту систему, получим $q_1^* = \frac{2}{3}$, $q_2^* = \frac{1}{3}$, $\text{x} = 7$.

Теперь найдем оптимальную смешанную стратегию первого игрока $p^* = (p_1^*, p_2^*, p_3^*, p_4^*)$. Так как его активные стратегии — вторая и четвертая, а первая и третья — пассивные, то $p_1^* = p_3^* = 0$. Следовательно, $p^* = (0, p_2^*, 0, p_4^*)$. Применяя утверждение и учитывая, что цена игры уже найдена, получим систему уравнений

$$\begin{cases} 9p_2^* + 6p_4^* = 7, \\ 3p_2^* + 9p_4^* = 7. \end{cases}$$

Решив эту систему, получим $p_2^* = \frac{1}{3}$, $p_4^* = \frac{2}{3}$. Итак, решение матричной игры задается векторами $p^* = (0, \frac{1}{3}, 0, \frac{2}{3})$, $q^* = (\frac{2}{3}, \frac{2}{3})$ и ценой игры $\text{x} = 7$:

$$p^* = (0, \frac{1}{3}, 0, \frac{2}{3}),\ q^* = (\frac{2}{3}, \frac{1}{3}),\ \text{x} = 7.$$

Глава 8. Игры с природой

§ 1. Задача о структуре посевов

Задачу о структуре посевов рассмотрим на следующем примере.

Агрофирма может выращивать любую из культур: или A, или B. Требуется установить, какими из этих культур и в каких пропорциях нужно засеять земли, принадлежащие фирме, чтобы в предстоящем сезоне после продажи урожая получить максимальную гарантированную в среднем выручку с 1 га используемых земель.

Достоверный прогноз погоды отсутствует, и неизвестно, будет ли предстоящее лето засушливым, дождливым или нормальным.

Средняя урожайность культур (ц/га) в зависимости от погоды, установленная на основе прошлого опыта, приведена в таблице:

Культура	Засушливое лето	Нормальное лето	Дождливое лето
A	18	22	32
B	17	9	7

Агрофирму можно считать первым игроком, имеющим две чистые стратегии (засеять всю площадь культурой A или культурой B), а природу — вторым игроком, имеющим три чистые стратегии (установить засушливое, нормальное или дождливое лето). Известно, что цены на продажу 1 ц культуры A и культуры B в следующем году прогнозируются на уровне 1 тыс. руб. и 2 тыс. руб. соответственно.

Требуется найти и пояснить оптимальную смешанную стратегию посевов агрофирмы, наиболее неблагоприятную для агрофирмы смешанную стратегию природы и цену полученной матричной игры.

Составим платежную матрицу игры, которая описывается заданной ситуацией.

У платежной матрицы A будет две строки и три столбца. Выигрыш агрофирмы — это ее выручка, полученная с 1 га от продажи выращенных культур. Умножив урожайность культуры A в тех или иных условиях на ожидаемую цену 1 ц данной культуры, получим ожидаемую выручку агрофирмы с 1 га пашни, засеянного культурой A. В случае засушливого лета ожидаемая выручка составит 1 тыс. руб/ц · 18 ц/га = 18 тыс. руб/га; в случае нормального лета — 1 тыс. руб/ц · 12 ц/га = 12 тыс. руб/га; в случае дождливого лета — 1 тыс. руб/ц · 32 ц/га = 32 тыс. руб/га.

Аналогично рассчитаем ожидаемую выручку в случае применения агрофирмой ее второй чистой стратегии.

Итак, игра задается следующей платежной матрицей:

$$A = \begin{pmatrix} 18 & 22 & 32 \\ 34 & 18 & 14 \end{pmatrix}.$$

Пусть p_1 — доля имеющейся у агрофирмы пашни, засеянная культурой *А*, а p_2 — доля имеющейся у агрофирмы пашни, засеянная культурой *В*. Если считать, что будет засеяна вся площадь, то $p_1 + p_2 = 1,\ \ p_1 \geq 0,\ \ p_2 \geq 0,$ т. е. вектор $p = (p_1, p_2)$ задает смешанную стратегию агрофирмы. Пусть q_1 — вероятность засушливого лета, q_2 — вероятность нормального лета, q_3 — вероятность дождливого лета. Тогда вектор $q = (q_1, q_2, q_3)$ задает смешанную стратегию природы. Обозначим через х ожидаемую выручку агрофирмы с 1 га пашни при применении ей своей смешанной стратегии. Найдем оптимальные смешанные стратегии агрофирмы и природы и цену игры.

Рассмотрим еще один способ решения матричных игр — сведение задач теории игр к *задачам линейного программирования*. Опишем сначала процедуру сведения в общем виде.

Рассмотрим игру, определяемую матрицей

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Утверждение 2. Для того чтобы число υ было ценой игры, а векторы $p^* = (p_1^*, \ldots, p_m^*)$ и $q^* = (q_1^*, \ldots, q_n^*)$ — оптимальными смешанными стратегиями первого и второго игроков соответственно, необходимо и достаточно выполнение неравенств

$$\sum_{i=1}^{m} a_{ij} p_i^* \geq \text{х}\ \ (j = \overline{1, n}),$$

$$\sum_{j=1}^{n} a_{ij} q_j^* \leq \text{х}\ \ (i = \overline{1, m}).$$

Согласно утверждению 2 для оптимальной стратегии первого игрока должно выполняться первое неравенство. Предположим для определенности, что $\text{х} > 0$. Этого всегда можно добиться, прибавив ко всем элементам матрицы A одно и то же число C. Подобная операция не изменит оптимальных стратегий игроков, а только увеличит цену игры на C.

Разделим обе части первого неравенства на х (так как $\text{х} > 0,$ то знак неравенства сохраняется):

$$\sum_{i=1}^{m} a_{ij} \frac{p_i^*}{\text{х}} \geq 1,\ \ (j = \overline{1, n}).$$

Положим, $p_i^* / \text{х} = u_i^*$.

Тогда

$$\sum_{i=1}^{m} a_{ij} u_i^* \geq 1, \ \ (j = \overline{1, n}).$$

Кроме того, так как по определению смешанных стратегий $p_i^* \geq 0, \ \ (i = \overline{1, m}),$ то и $u_i^* \geq 0, \ \ (i = \overline{1, m}).$ Сумма компонент вектора смешанных стратегий должна быть равна единице:

$$\sum_{i=1}^{m} p_i^* = 1 \ \Rightarrow \ \sum_{i=1}^{m} u_i^* = 1 / \text{х}.$$

Так как первый игрок стремится получить максимальный в среднем выигрыш, то он должен минимизировать величину $1/\text{х}$. Учитывая это, получим, что для нахождения оптимальной смешанной стратегии первого игрока необходимо решить задачу линейного программирования:

$$W = \sum_{i=1}^{m} u_i \to \min$$

$$\sum_{i=1}^{m} a_{ij} u_i \ \geq 1, \ \ (j = \overline{1, n}),$$

$$u_i \ \geq 0, \ \ (i = \overline{1, m}).$$

Аналогично можно показать, что для нахождения оптимальной смешанной стратегии второго игрока необходимо решить задачу линейного программирования:

$$Z = \sum_{j=1}^{n} x_j \to \max$$

$$\sum_{j=1}^{n} a_{ij} x_j \ \leq 1, \ \ (i = \overline{1, m}),$$

$$x_j \ \geq 0, \ \ (j = \overline{1, n}).$$

Здесь $x_j = q_j / \text{х}, \ \ j = \overline{1, n},$ где $q = (q_1, \ldots, q_n)$ — вектор смешанных стратегий второго игрока.

Заметим, что эти задачи являются двойственными друг другу.

Пусть $u^* = (u_1^*, \ldots, u_n^*)$ и $x^* = (x_1^*, \ldots, x_m^*)$ — оптимальные решения задач, $Z^* = W^*$ — оптимальные значения целевых функций. Формулы для определения оптимальных стратегий и цены игры имеют вид

$$p_i^* = \text{x} u_i^*, \;\; i = \overline{1, m}, \;\; q_j^* = \text{x} x_j^*, \;\; j = \overline{1, n}, \;\; \text{x} = 1/Z^* = 1/W^*.$$

Теперь выпишем задачи линейного программирования, соответствующие игре, сформулированной в примере:

Задача агрофирмы	Задача природы
$W = u_1 + u_2 \to \min$	$Z = x_1 + x_2 + x_3 \to \max$
$18u_1 + 34u_2 \geq 1,$	$18x_1 + 22x_2 + 32x_3 \leq 1,$
$22u_1 + 18u_2 \geq 1,$	$34x_1 + 18x_2 + 14x_3 \leq 1,$
$32u_1 + 14u_2 \geq 1,$	$x_1 \geq 0, \; x_2 \geq 0, \; x_3 \geq 0.$
$u_1 \geq 0, \; u_2 \geq 0.$	

Найдем оптимальное решение каждой из этих задач. Задача агрофирмы имеет две переменные, поэтому ее можно решить *графически* (рис. 7).

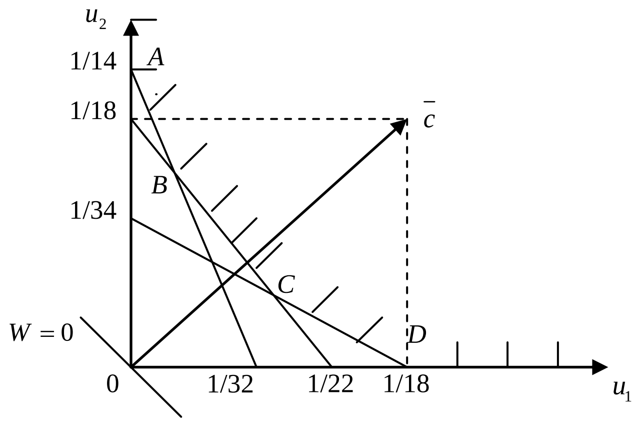


Рис. 7. Графическое решение задачи линейного программирования

Область допустимых решений задачи (см. рис. 7) представлена неограниченным четырехугольником *ABCD*, вектор $\bar{c} = (1/18; 1/18)$ дает направление вектора-градиента целевой функции, линия $W = 0$ изображает линию уровня целевой функции. Точкой минимума является точка C, лежащая на пересечении первой и второй граничных прямых. Ее координаты найдем, решив систему уравнений

$$\begin{cases} 18u_1^* + 34u_2^* = 1, \\ 22u_1^* + 18u_2^* = 1. \end{cases}$$

Отсюда

$$u_1^* = \frac{4}{106}, \quad u_2^* = \frac{1}{106}.$$

Тогда

$$W^* = u_1^* + u_2^* = \frac{5}{106}.$$

Оптимальное решение задачи природы найдем при использовании *условия «дополняющей нежесткости»*.

Первая группа условий: $x_j v_j = 0$, $j = \overline{1,3}$,

где $v_1 = 18u_1 + 34u_2 - 1,$
$v_2 = 22u_1 + 18u_2 - 1,$
$v_3 = 32u_1 + 14u_2 - 1.$

Тогда

$$x_1(18u_1 + 34u_2 - 1) = 0 \;\Rightarrow\; x_1 \cdot 0 = 0 \;\Rightarrow\; x_1 \geq 0,$$
$$x_2(22u_1 + 18u_2 - 1) = 0 \;\Rightarrow\; x_2 \cdot 0 = 0 \;\Rightarrow\; x_2 \geq 0,$$
$$x_3(328u_1 + 14u_2 - 1) = 0 \;\Rightarrow\; x_3 \frac{36}{106} = 0 \;\Rightarrow\; x_3 = 0.$$

Вторая группа условий: $u_i y_i = 0$, $i = \overline{1,2}$,

где $y_1 = 1 - 18x_1 - 22x_2 - 32x_3,$
$y_2 = 1 - 34x_1 - 18x_2 - 14x_3.$

Тогда

$$u_1(1 - 18x_1 - 22x_2 - 32x_3) = 0 \;\Rightarrow\; \frac{4}{106}(1 - 18x_1 - 22x_2 - 32x_3) = 0,$$
$$u_2(1 - 34x_1 - 18x_2 - 14x_3) = 0 \;\Rightarrow\; \frac{1}{106}(1 - 34x_1 - 18x_2 - 14x_3) = 0.$$

Отсюда получим

$$\begin{cases} 18x_1 + 22x_2 + 32x_3 = 1, \\ 34x_1 + 18x_2 + 14x_3 = 1, \\ \qquad\qquad\qquad x_3 = 0. \end{cases}$$

Решение этой системы: $x_1^* = \frac{1}{106}$, $x_2^* = \frac{4}{106}$, $x_3^* = 0.$ Далее, используя найденные оптимальные решения пары двойственных задач, найдем цену игры и оптимальные смешанные стратегии игроков.

$$\text{х} = 1/W^* = \frac{106}{5} = 21{,}2, \quad p_1^* = \frac{106}{5}\,\frac{4}{106} = \frac{4}{5}, \quad p_2^* = \frac{106}{5}\,\frac{1}{106} = \frac{1}{5},$$

$$q_1^* = \frac{106}{5}\,\frac{1}{106} = \frac{1}{5}, \; q_2^* = \frac{106}{5}\,\frac{4}{106} = \frac{4}{5}, \; q_3^* = \frac{106}{5}\,0 = 0.$$

Таким образом, оптимальная смешанная стратегия агрофирмы $p^* = (0{,}8;\, 0{,}2).$ Это означает, что для получения максимальной ожидаемой выручки агрофирме необходимо 80 % площади имеющейся пашни засеять культурой *A* и 20 % пашни засеять культурой *B*. Ожидаемая выручка при этом должна составить 21,2 тыс. руб/га.

Оптимальная смешанная стратегия природы $q^* = (0{,}2;\, 0{,}8;\, 0),$ т. е. наиболее неблагоприятной для фирмы погодой будет с вероятностью 20 % — засушливое лето, с вероятностью 80 % — нормальное лето, а с вероятностью 0 — дождливое лето.

§ 2. Задача об оптовой закупке при неопределенности розничной продажи

Рассмотрим данную задачу на следующем примере.

Торговая фирма на каждый предстоящий день решает вопрос об объемах оптовой закупки скоропортящихся продуктов двух наименований.

По статистике наблюдений за местным рынком этих товаров при холодной и дождливой погоде на каждые 2 единицы товара *A* в среднем реализуется 3 единицы товара *B*, а при жаркой солнечной погоде товары реализуются в усредненной пропорции 5 единиц товара *A* к 1 единице товара *B*.

Оптовая цена товара *A* равна 40 руб., товара *B* — 30 руб. Розничная цена товара *A* составляет 70 руб., товара *B* — 60 руб.

Товары, закупленные по оптовым ценам в расчете на холодную погоду с учетом известной пропорции на сумму 97 750 руб., полностью удовлетворяют спрос на эти товары при холодной погоде. Аналогичное замечание верно для солнечной погоды.

Ежедневные издержки на розничную реализацию продукции составляют 19 000 руб.

В конце дня нереализованный товар фирма сдает на пищевую переработку по ценам на 75 % меньше оптовой.

Администрацию интересует, в каких объемах следует сделать оптовые закупки этих товаров, чтобы максимизировать гарантированную в среднем прибыль фирмы в условиях полной неопределенности предстоящей погоды.

Требуется:

1) составить игровую математическую модель предложенной ситуации, рассчитав соответствующую платежную матрицу;

2) указать оптимальную стратегию закупок фирмы, наиболее неблагоприятную для фирмы погоду и гарантированный максимум прибыли фирмы, найдя решение полученной матричной игры;

3) определить, какова будет ожидаемая прибыль фирмы при выборе рассчитанной оптимальной стратегии, если учесть информацию о том, что вероятность завтрашнего дождя равна 40 %. Определить, каковы будут новая оптимальная стратегия и новая ожидаемая прибыль фирмы при учете этой информации.

Составим *платежную матрицу игры*, описывающую предложенную в задаче ситуацию. Эта задача является примером *игр с природой*.

Торговая фирма выступает в качестве первого игрока. У нее есть две чистые стратегии: делать закупки, рассчитывая на холодную и дождливую погоду (первая стратегия) или на жаркую солнечную (вторая стратегия).

Вторым игроком является природа. У нее также две чистые стратегии. Первая стратегия — установить холодную и дождливую погоду, вторая стратегия — установить жаркую солнечную погоду.

Элементы платежной матрицы — это прибыль, которую будет получать фирма в той или иной ситуации. Прибыль рассчитывается как разница между выручкой фирмы от розничной реализации продукции и ежедневными затратами фирмы.

Затраты, в свою очередь, есть сумма средств, направленных на оптовую закупку товаров, и издержек, направленных на розничную реализацию продукции.

Рассмотрим первую чистую стратегию фирмы. Пусть фирма планирует закупить товары A и B в количествах k_A и k_B соответственно. Предполагается, что завтра будет холодная погода. Исходя из статистических данных, на каждые 2 единицы товара A реализуется 3 единицы товара B. Следовательно, фирма должна закупить продукты таким образом, чтобы выполнялось соотношение на их количество. Это значит, что должно выполняться соотношение

$$\frac{k_A}{k_B} = \frac{2}{3}.$$

На закупки фирма предполагает тратить 97 750 руб. ежедневно. Поэтому, исходя из оптовых цен продуктов, она может купить продукты в количествах, для которых выполнено соотношение:

$$4k_A + 3k_B = 97\,750.$$

Итак, для нахождения количеств закупаемых в расчете на холодную погоду продуктов необходимо решить следующую систему уравнений:

$$\begin{cases} \dfrac{k_A}{k_B} = \dfrac{2}{3}, \\ 40k_A + 30k_B = 97\,750. \end{cases}$$

Отсюда получим, что k_A = 1 150 k_B = 1 725, т. е. необходимо закупить 1 150 единиц товара *A* и 1 725 единиц товара *B*.

Пусть фирма применяет свою вторую чистую стратегию, т. е. делает покупки в расчете на солнечную погоду. Это значит, что фирма предполагает на каждые 5 единиц товара *A* продать 1 единицу товара *B*. Следовательно, для количеств k_A и k_B закупаемых товаров должно выполняться соотношение

$$\frac{k_A}{k_B} = \frac{5}{1}.$$

Учитывая оптовые цены товаров и планируемую сумму трат на их закупку, получим следующую систему уравнений:

$$\begin{cases} \dfrac{k_A}{k_B} = \dfrac{5}{1}, \\ 40k_A + 30k_B = 97\,750. \end{cases}$$

Решение этой системы: k_A = 2125, k_B = 425. Итак, при второй своей чистой стратегии фирма должна закупить 2 125 единиц товара *A* и 425 единиц товара *B*.

Теперь рассчитаем ожидаемую прибыль фирмы при различных погодных условиях. Может быть четыре ситуации:

1) фирма применяла свою первую чистую стратегию, т. е. делала закупки в расчете на холодную погоду и угадала;

2) фирма делала закупки в расчете на холодную погоду (применяла первую чистую стратегию), а погода была жаркой;

3) фирма применяла свою вторую чистую стратегию, т. е. делала закупки в расчете на жаркую погоду, а погода была холодной;

4) фирма применяла свою вторую чистую стратегию, т. е. делала закупки в расчете на жаркую погоду, и погода была действительно жаркой.

Рассмотрим отдельно каждую ситуацию.

1. Фирма совершает покупки в расчете на холодную погоду, т. е. закупает 1 150 единиц товара *A* и 1 725 единиц товара *B*. При холодной погоде (т. е. первой чистой стратегии природы) продукты продаются в выбранном фирмой соотношении, т. е. фирма

продаст все продукты, которые закупила (товар A — в количестве 1 150 единиц, товар B — в количестве 1 725 единиц).

Учитывая издержки, связанные с розничной продажей, прибыль фирмы (П) в данном случае будет равна

$$П = 1\,150 \cdot 70 + 1\,725 \cdot 60 - 1\,150 \cdot 40 - 1\,725 \cdot 30 - 19\,000 = 67\,250 \text{ (руб.)}.$$

2. Фирма по-прежнему закупает 1 150 единиц товара A и 1 725 единиц товара B, рассчитывая на холодную погоду. При жаркой погоде товара A можно было бы продать 2 125 единиц, но фирма купила только 1 150. Товара B фирма купила 1 725 единиц, но при жаркой погоде можно продать только 425 единиц. Излишек товара B в количестве 1 725 – 425 = 1 300 единиц можно сдать на пищевую переработку по цене на 75 % меньше оптовой, т. е. по 0,25 · 30 = 7,5 (руб.). Итого прибыль фирмы в данной ситуации будет равна

$$П = 1\,150 \cdot 70 + 425 \cdot 60 - 1\,150 \cdot 40 - 1\,725 \cdot 30 + 1\,300 \cdot 7{,}5 - 19\,000 = -\,1\,000 \text{ (руб.)}.$$

Отрицательная прибыль фирмы означает, что фирма несет убытки, так как не угадала погоду на завтра.

3. Фирма применяет свою вторую чистую стратегию, т. е. закупает 2 125 единиц товара A и 425 единиц товара B, рассчитывая на жаркую погоду. Если в действительности погода оказалась холодной, то фирма сможет продать только 1 150 единиц товара A. Остаток в 2 125 – 1 150 = 975 единиц можно будет сдать на пищевую переработку по цене на 75 % меньше оптовой, т. е. по 0,25 · 40 = 10 руб.

Фирма сможет продать товара B только 425 единиц вместо возможных 1 725 единиц. Итак, если фирма применит свою вторую чистую стратегию против первой стратегии природы, то прибыль будет равна

$$П = 1\,150 \cdot 70 + 425 \cdot 60 - 2\,125 \cdot 40 - 425 \cdot 30 + 975 \cdot 10 - 19\,000 = -\,1\,000 \text{ (руб.)}.$$

4. Фирма совершает покупки в расчете на жаркую погоду, т. е. закупает 2 125 единиц товара A и 425 единиц товара B. Если погода угадана, то фирма продаст все продукты, которые закупила (товар A — в количестве 2 125 единиц, товар B — в количестве 425 единиц). Данная ситуация соответствует второй чистой стратегии природы. Учитывая издержки, связанные с розничной продажей, прибыль фирмы в данном случае будет равна

$$П = 2\,125 \cdot 70 + 425 \cdot 60 - 2\,125 \cdot 40 - 425 \cdot 30 - 19\,000 = 57\,500 \text{ (руб.)}.$$

Таким образом, платежная матрица игры имеет вид

$$A = \begin{pmatrix} 67\,250 & -1\,000 \\ -1\,000 & 57\,500 \end{pmatrix}.$$

Найдем решение полученной матричной игры *графическим способом*. Пусть вектор $p = (p_1, p_2)$ — вектор смешанных стратегий фирмы, а $q = (q_1, q_2)$ — вектор смешанных стратегий природы. Тогда платежная функция игры имеет вид

$$E(p,q) = 67\,250 p_1 q_1 - 1\,000 p_2 q_1 - 1\,000 p_1 q_2 + 57\,500 p_2 q_2.$$

Оптимальной смешанной стратегии фирмы соответствует точка *K* (рис. 8).

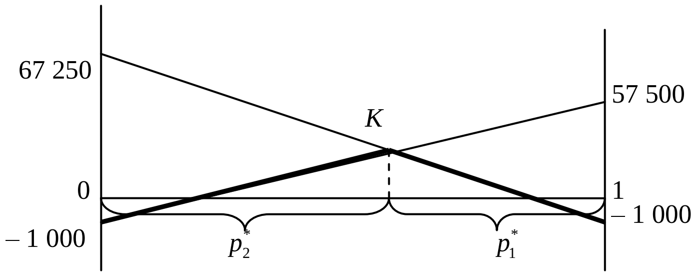


Рис. 8. Графическое решение задачи фирмы

Применим к данной игре утверждение 1. Обе чистые стратегии природы являются активными. Обозначим оптимальную смешанную стратегию фирмы $p^* = (p_1^*, p_2^*)$. Итак, пусть $p = p^*$, $q = (1,0)$. Тогда $E(p,q) = 67\,250 p_1^* - 1\,000 p_2^*$. Пусть $p = p^*$, $q = (0,1)$. Тогда $E(p,q) = -1\,000 p_1^* + 57\,500 p_2^*$. Эти значения приравняем к цене игры x и добавим уравнение $p_1^* + p_2^* = 1$. Получим следующую систему уравнений:

$$\begin{cases} 67\,250 p_1^* - 1\,000 p_2^* = \text{x}, \\ -1\,000 p_1^* + 57\,500 p_2^* = \text{x}, \\ p_1^* + p_2^* = 1. \end{cases}$$

Решение этой системы будет следующим: $p_1^* \approx 0,46$, $p_2^* \approx 0,54$, $\text{x} \approx 30\,500$.

Итак, оптимальная стратегия фирмы состоит в том, чтобы делать закупки, рассчитывая на завтрашний дождь в 46 % случаев, а на солнечную погоду — в 54 % случаев. Средняя ожидаемая выручка фирмы при этом должна составить 30 500 руб. ежедневно.

Снова применяя утверждение 1, составим систему для нахождения оптимальной смешанной стратегии природы:

$$\begin{cases} 67\,250 q_1^* - 1\,000 q_2^* = 30\,500, \\ -1\,000 q_1^* + 57\,500 q_2^* = 30\,500. \end{cases}$$

Отсюда $q_1^* \approx 0,46$, $q_2^* \approx 0,54$.

Предположим, что вероятность завтрашнего дождя известна и составляет 40 %. Тогда смешанная стратегия природы $q = (0,4;\, 0,6)$.

Найдем ожидаемую прибыль фирмы с учетом этой информации, если применяется оптимальная смешанная стратегия фирмы. Подставим оптимальную смешанную стратегию фирмы $p^* = (0,46;\, 0,54)$ и $q = (0,4;\, 0,6)$ в выражение для платежной функции игры:

$$E(p^*, q) = 67\,250 \cdot 0,46 \cdot 0,4 - 1\,000 \cdot 0,46 \cdot 0,6 - 1\,000 \cdot 0,54 \cdot 0,4 +$$
$$+\, 57\,500 \cdot 0,54 \cdot 0,6 = 30\,514 \text{ (руб.).}$$

Заметим, что согласно теореме, ожидаемая прибыль фирмы должна была оказаться равной цене игры, т. е. 30 500 руб. Отклонение рассчитанной величины от данного значения вызвано погрешностью вычисления оптимальной смешанной стратегии фирмы.

Найдем оптимальную стратегию фирмы на завтрашний день, если администрации известна вероятность завтрашнего дождя, равная 40 %. Возможные стратегии фирмы на завтрашний день — делать закупки в расчете либо на дождливую, либо на солнечную погоду. Смешанная стратегия природы $q = (0,4;\, 0,6)$.

Если фирма будет делать закупки в расчете на дождливую погоду, т. е. применять свою первую чистую стратегию $p^1 = (1, 0)$, то ее ожидаемая выручка составит

$$E(p^1, q) = 67\,250 \cdot 1 \cdot 0,4 - 100 \cdot 1 \cdot 0,6 = 26\,300 \text{ (руб.).}$$

Если фирма будет делать закупки в расчете на жаркую погоду, т. е. применять свою вторую чистую стратегию $p^2 = (0, 1)$, то ее ожидаемая выручка составит

$$E(p^2, q) = -1\,000 \cdot 1 \cdot 0,4 + 57\,500 \cdot 1 \cdot 0,6 = 34\,100 \text{ (руб.).}$$

Итак, поскольку $E(p^1, q) < E(p^2, q)$, то оптимальной стратегией фирмы на следующий день является ее вторая чистая стратегия, т. е. выполнение закупок в количествах 2 125 единиц товара *A* и 425 единиц товара *B* соответственно. При этом на следующий день ожидается прибыль, равная 34 100 руб.

Исходную задачу можно представить еще одним способом — в виде *задачи линейного программирования*. Вначале разберем этот подход в общем виде.

Пусть смешанная стратегия для первого игрока $S_A = \begin{Bmatrix} A_1 & \dots & A_m \\ p_1 & \dots & p_m \end{Bmatrix}$. Введем в рассмотрение новую переменную: p_{m+1} — нижняя цена игры.

Тогда прямая задача линейного программирования примет вид

$$p_{m+1} \to \max$$

$$-\sum_{i=1}^{m} a_{ij} p_i + p_{m+1} \le 0, \quad j = \overline{1,n},$$
$$\sum_{i=1}^{m} p_i = 1,$$
$$p_i \ge 0, \quad i = \overline{1,m}.$$

Пусть смешанная стратегия для второго игрока $S_B = \begin{Bmatrix} B_1 & \dots & B_n \\ q_1 & \dots & q_n \end{Bmatrix}$. Введем в рассмотрение новую переменную: q_{n+1} — верхняя цена игры. Тогда двойственная задача линейного программирования примет вид

$$q_{n+1} \to \min$$

$$-\sum_{j=1}^{n} a_{ij} q_j + q_{n+1} \le 0, \quad i = \overline{1,m},$$
$$\sum_{j=1}^{n} q_j = 1,$$
$$q_j \ge 0, \quad j = \overline{1,n}.$$

Решение задачи линейного программирования даст цену игры $\text{x} = p_{m+1} = q_{n+1}$.

В примере платежная матрица игры имела вид

$$A = \begin{pmatrix} 67\,250 & -1\,000 \\ -1\,000 & 57\,500 \end{pmatrix}.$$

Тогда соответствующую задачу линейного программирования запишем так:

$$p_3 \to \max \qquad\qquad q_3 \to \min$$

$$\begin{aligned} &-67\,250 p_1 + 1\,000 p_2 + p_3 \le 0, && -67\,250 q_1 + 1\,000 q_2 + q_3 \le 0, \\ &10\,00 p_1 - 57\,500 p_2 + p_3 \le 0, && 1\,000 q_1 - 57\,500 q_2 + q_3 \le 0, \\ &p_1 + p_2 = 1, && q_1 + q_2 = 1, \\ &p_i \ge 0, \ i = \overline{1,2}. && p_i \ge 0, \ i = \overline{1,2}. \end{aligned}$$

Возможным недостатком этого представления является тот факт, что большая размерность обеих переменных задачи не позволяет применить графический метод решения. Если же для решения применять приложение «Поиск решения» из Microsoft Excel, то данное представление, наоборот, будет более простым для понимания и мы сразу получим решение окончательной задачи.

Контрольные вопросы и задания

1. Дайте определение понятия матричной игры.
2. Что означает смешанная стратегия?
3. Какие матричные игры можно решать графическим способом?
4. Какие игры называются «играми с природой»?
5. Какие стратегии называются чистыми, активными и оптимальными?

Задания для самостоятельного решения

1. Фирма располагает сетью из n магазинов. В инспекцию поступили сведения о том, что в этих магазинах продается бракованный товар, причем его объем в магазине j равен q_j. Инспектирующий орган намерен провести инспекцию с целью обнаружения брака. Так как эти магазины территориально разбросаны, инспекции может быть подвергнут только один из них. Фирма узнает о готовящейся инспекции, и чтобы обезопасить себя, решает изъять бракованный товар. По техническим причинам такое изъятие может быть проведено только в одном магазине.

Какие действия инспектора и фирмы будут оптимальными?

2. Руководство универмага заказывает товар определенного вида. Известно, что спрос на товар данного вида лежит в пределах от 6 до 9 единиц. В случае, если заказанного товара окажется недостаточно для удовлетворения спроса, то имеется возможность срочно заказать и завезти недостающее количество. Если же спрос будет меньше наличного количества товара, то нереализованный товар придется хранить на складе универмага. Требуется определить такой объем заказа на товар, при котором дополнительные затраты, связанные с хранением и срочным завозом, были бы минимальными. Расходы на хранение единицы товара составляют 1 тыс. условных денежных единиц, а по срочному заказу и завозу — 2 тыс. условных единиц.

Составьте платежную матрицу. Найдите оптимальные решения игроков.

3. Решите матричную игру с платежной матрицей $\begin{pmatrix} 2 & 2 & 3 & -1 \\ 4 & 3 & 2 & 6 \end{pmatrix}$.

4. Решите матричную игру с платежной матрицей $\begin{pmatrix} 2 & 4 \\ 2 & 3 \\ 3 & 2 \\ -2 & 6 \end{pmatrix}$.

5. Решите матричную игру с платежной матрицей $\begin{pmatrix} 2 & 4 \\ 2 & 1 \\ 1 & 6 \\ 4 & 0 \end{pmatrix}$.

6. Решите матричную игру с платежной матрицей $\begin{pmatrix} 0 & 1/2 & 5/6 \\ 1 & 3/4 & 1/2 \end{pmatrix}$.

7. Решите матричную игру с платежной матрицей $\begin{pmatrix} 1 & 4/3 & 2 \\ 2 & 4/3 & 0 \end{pmatrix}$.

8. Задана матричная игра платежной матрицей $\begin{pmatrix} 4 & 5 & 3 \\ 6 & 7 & 4 \\ 5 & 2 & 3 \end{pmatrix}$.

Определите нижнюю и верхнюю цену игры. Сформулируйте соответствующую задачу линейного программирования. Упростите матрицу игры и найдите решение в смешанных стратегиях.

9. Задана матричная игра платежной матрицей $\begin{pmatrix} 8 & 3 & 9 & 4 \\ 6 & 5 & 8 & 7 \\ 3 & 4 & 5 & 6 \end{pmatrix}$.

Определите нижнюю и верхнюю цену игры. Сформулируйте соответствующую задачу линейного программирования. Упростите матрицу игры и найдите решение в смешанных стратегиях.

10. Задана матричная игра платежной матрицей $\begin{pmatrix} 3 & 4 & 2 \\ 7 & 9 & 3 \\ 1 & 3 & 1 \\ 1 & 6 & 4 \\ 5 & 2 & 7 \\ 8 & 1 & 20 \end{pmatrix}$.

Определите нижнюю и верхнюю цену игры. Сформулируйте соответствующую задачу линейного программирования. Упростите матрицу игры и найдите решение в смешанных стратегиях.

11. Две фирмы производят два конкурирующих товара, контролируя по 50 % рынка. Если одна из фирм будет рекламировать свои товары, то другая потеряет определенный процент клиентов. Известно, что 50 % получают информацию через телевидение, 30 % — через газеты, остальные 20 % — по радио.

Какие средства рекламы выбрать фирмам? Задачу следует решить методами теории игр.

12. Прибыль сельскохозяйственного предприятия (млн руб.) от выращивания двух культур K и L в зависимости от состояний погоды представлена в таблице:

Культура	Засушливое лето	Нормальное лето	Дождливое лето
K	8	5	3
L	2	3	6

Определите оптимальный план посева культур.

Библиографический список

Основная литература

1. *Осипов, А. Л.* Экономико-математические методы в управлении : учеб.-метод. комплекс для дистанц. обучения / А. Л. Осипов, Е. А. Рапоцевич. — Новосибирск : СибАГС, 2006. — 144 с.

2. *Сборник* задач по высшей математике для экономистов : учеб. пособие / под ред. В. И. Ермакова. — 2-е изд., испр. — Москва : ИНФРА-М, 2007. — С. 412—497.

3. *Шикин, Е. В.* Математические методы и модели в управлении : учеб. пособие / Е. В. Шикин, А. Г. Чхартишвили. — 2-е изд., испр. — Москва : Дело, 2002. — С. 65—79.

4. *Экономико*-математические методы в управлении : практикум / СибАГС ; сост. : А. Л. Осипов, Е. А. Рапоцевич. — Новосибирск : Изд-во СибАГС, 2007. — 160 с.

Дополнительная литература

1. *Лабскер, Л. Г.* Игровые методы в управлении экономикой и бизнесом : учеб. пособие / Л. Г. Лабскер, Л. О. Бабешко. — Москва : Дело, 2001. — 420 с.

2. *Надеев, А. Т.* Моделирование социально-политических и экономических процессов : учеб. / А. Т. Надеев. — Нижний Новгород : Изд-во ВВАГС, 2002. — 350 с.

3. *Синюк, В. Г.* Использование информационно-аналитических технологий при принятии управленческих решений : учеб. пособие / В. Г. Синюк, А. В. Шевырев. — Москва : Экзамен, 2004. — 160 с.

Часть 4. ПРИЛОЖЕНИЯ В СОЦИОЛОГИИ

Глава 9. Моделирование демографических процессов

Демографические процессы включают следующие процессы: рождаемость, продолжительность жизни, смертность, статистику абортов, детскую смертность, самоубийства, миграцию, интеллектуальную эмиграцию.

Демографические процессы развиваются под воздействием экономических, политических и социальных процессов. В свою очередь, демографические процессы оказывают влияние на ход различных общественных процессов. Например, низкий уровень рождаемости ведет к увеличению процентной доли пенсионеров в обществе и обострению проблемы «отцов и детей». Колебания уровня рождаемости через определенное время проявляются в соответствующих (или противоположных) колебаниях уровня занятости на рынке труда, уровня преступности, конкурсов между абитуриентами при поступлении в учебные заведения и т. п.

Демография в исследовании своего предмета — естественного воспроизводства населения — использует различные методы, основные из которых можно объединить по их характеру в три группы: ***статистические, математические, социологические методы***.

Объектами наблюдения в демографии являются не отдельные люди и события, а сгруппированные по определенным правилам, однородные в некотором отношении совокупности людей и событий. Такие совокупности называются статистическими фактами. Демография стремится установить и измерить объективно существующие взаимосвязи между статистическими фактами, имеющими отношение к ее предмету. Для этого она использует методы, также разработанные в статистике: например, ***методы корреляционного и факторного анализа, выборочный и индексный табличный методы, метод средних величин, методы выравнивания и др***.

Процессы воспроизводства населения связаны между собой иногда простыми, иногда довольно сложными количественными соотношениями, что обусловливает применение многих математических методов для измерения одних демографических характеристик по данным о других характеристиках.

В демографии широко используются математические модели населения, с помощью которых на основе фрагментарных и неточных данных можно получить достаточно полное и достоверное представление об истинном состоянии воспроизводства населения. К разряду ***математического моделирования*** в демографии относятся *вероятностные таблицы смертности*, а также *демографические прогнозы*, которые представляют собой один из видов математического моделирования.

Внутри демографии стихийно выделяется демографическая статистика — старейшая отрасль демографии. Ее частным предметом является изучение статистических закономерностей воспроизводства населения, а задачей — разработка методов статистического наблюдения и измерения демографических явлений и процессов, сбор и первичная обработка статистических материалов о воспроизводстве населения.

Математическая демография разрабатывает и применяет математические методы для изучения взаимосвязей демографических явлений и процессов, моделирования и прогнозирования.

Демографические модели включают вероятностные таблицы смертности, брачности, рождаемости, модели стационарного и стабильного населения, имитационные модели демографических процессов и т. п.

Историческая демография изучает состояние и динамику демографических процессов в истории стран и народов, а также историю развития самой демографической науки.

Этническая демография исследует этнические особенности воспроизводства населения. Этнические особенности бытового уклада жизни народов, обычаи, традиции, структура семейных отношений оказывают существенное влияние на уровень рождаемости, среднюю продолжительность жизни и состояние здоровья.

Экономическая демография исследует экономические факторы воспроизводства населения. Под экономическими факторами понимается вся совокупность экономических условий жизни общества, которые влияют на уровень роста населения, уровень рождаемости, смертности, брачности и т. д.

Социологическая демография изучает влияние социологических социально-психологических факторов на волевые, субъективные действия людей в демографических процессах.

Демографические прогнозы играют важную роль в комплексном долгосрочном социально-экономическом планировании. Практически очень трудно найти какую-либо область экономики и социальной жизни, где бы при долгосрочном планировании не использовались данные демографических прогнозов.

Демографические прогнозы имеют активный характер. Сравнивая полученные в результате перспективных исчислений величины и те параметры демографических процессов, например, численности и возрастно-полового состава населения в том или ином регионе, которые желательны с социально-экономических позиций для общества в перспективе, можно выявить степень расхождения желаемых и возможных характеристик демографических процессов. Если такие расхождения велики, то общество может принять меры для ликвидации или уменьшения возможных диспропорций этих процессов.

Следовательно, демографические прогнозы являются важным элементом в управлении общественными процессами. Они позволяют на основе знания перспектив его развития целенаправленно воздействовать на развитие социально-экономических явлений, корректировать их в необходимую для страны сторону.

В заключение еще раз отметим, что моделирование — это метод научного познания, основанный на изучении реальных объектов посредством изучения моделей этих объектов, т. е. посредством изучения более доступных для исследования и (или) вмешательства объектов-заместителей естественного или искусственного происхождения, обладающих свойствами реальных объектов.

Свойства любой модели не должны, да и не могут точно и полностью соответствовать абсолютно всем свойствам исследуемого реального объекта в любых ситуациях.

В математических моделях любой дополнительный параметр может привести к существенному усложнению решения соответствующей системы уравнений, к необходимости применения дополнительных допущений, отбрасывания малых членов и т. п.

Таким образом, при моделировании существенным является вопрос об оптимальной для данного конкретного исследования степени соответствия модели оригиналу по вариантам поведения исследуемой системы, по связям с другими объектами и внутренним связям этой системы. В зависимости от вопроса, на который хочет ответить исследователь, одна и та же модель одного и того же реального объекта может быть признана адекватной или абсолютно не отражающей реальность.

Глава 10. Распределение доходов и богатств в обществе

Одни экономисты трактуют понятие «***доход***» в более развернутом виде, другие — берут в расчет лишь некоторые аспекты.

Так, И. П. Николаева указывает на то, что доход — это часть произведенного продукта, полученная участником производства в зависимости от своего участия в нем.

А. С. Булатов считает, что доход — это сумма денежных средств и материальных благ, полученных или произведенных домашними хозяйствами за определенный промежуток времени.

Доходы *классифицируются*:

1. В зависимости от учета динамики уровня потребительских цен:

— на номинальный доход — количество денег, полученное в определенный период отдельным лицом; он также характеризует уровень денежных доходов независимо от налогообложения;

— реальный — количество товаров и услуг, которое можно купить на располагаемый доход в течение определенного периода;

— располагаемый — доход, который может быть использован на личное потребление и личные сбережения. Располагаемый доход меньше номинального на сумму налогов и обязательных платежей.

2. В зависимости от формы единицы дохода:

— на денежный — некоторые выплаты из социальных фондов: продукты, произведенные в личных подсобных хозяйствах, и услуги, оказываемые членами семьи в домашнем хозяйстве;

— натуральный — оплата труда, доходы от предпринимательской деятельности; пенсии, стипендии, пособия, социальные выплаты; поступления от собственности, в качестве процентов по вкладам; ценные бумаги; дивиденды; доходы от продажи продукции сельского хозяйства; страховые возмещения; сумма от продажи иностранной валюты и др.

3. В зависимости от государственного вмешательства:

— на первичный — доход, образованный под воздействием рыночного механизма;

— вторичный — доход, формирование которого связано с перераспределительной политикой государства.

Существует *два вида распределения доходов*: функциональное и вертикальное.

Функциональное распределение доходов — это распределение их между факторами (труд, капитал, природные ресурсы и предпринимательские способности). Функциональное распределение характеризует распределение дохода между факторами производства и прежде всего между трудом и капиталом. В результате функционального распределения доходов формируются такие первичные доходы, как заработная плата, процент, рента и прибыль. В системе факторов производства основная взаимосвязь касается капитала, поэтому для упрощения функциональное распределение можно представить как соотношение между доходами от труда и собственности.

Вертикальное распределение есть результат перераспределительного вмешательства государства в сферу доходов, благодаря чему доходами располагают даже такие группы (например, нетрудоспособные или безработные), которые не могли бы их иметь, если бы общество удовлетворялось только функциональным распределением.

Можно сказать, что доходы совершают в экономике своеобразный кругооборот. Доходы поступают в семью в обмен на факторы производства, такие как труд, земля, капитал, предпринимательская деятельность. Обмениваясь на товары и предоставляя семье услуги, они превращаются в доходы тех, кто произвел эти товары и предложил услуги. Часть средств пойдет на финансирование экономической деятельности, что обеспечит семью доходами в будущем. Свою долю доходов получит и государство. Оно поддержит экономически слабых и неимущих, создаст лучшие условия для функционирования экономики в целом. Это позволит семьям не только увеличить свои доходы, но и более щедро обеспечивать поступления средств в бюджет, создавая благоприятные условия для выполнения государством его функций, в том числе экономических.

Выделяют следующие *причины неравенства доходов*:

1. Различия в способностях. Все люди имеют разные интеллектуальные и физические способности. То есть врожденные таланты дают возможность некоторым индивидам внести свой вклад в совокупный продукт — вклад, который приносит очень высокие доходы, другие же оказываются в гораздо менее благоприятных условиях.

2. Образование и обучение. Люди существенно отличаются друг от друга по уровню полученного образования и профессиональной подготовки, а следовательно, и по своим возможностям зарабатывать доход. Отчасти эти различия являются результатом их свободного выбора.

3. Дискриминация. Простой анализ спроса и предложения показывает, как дискриминация — в данном случае имеется в виду дискриминация при найме на работу — является дополнительным источником неравенства доходов. Предположим, что ограничивается доступ женщин к таким профессиям, как секретарь или учитель, которые считались сугубо «женскими». Это означает, что предложение женской рабочей силы по сравнению со спросом на эти несколько профессий будет настолько большим, что заработная плата и доходы будут низкими.

4. Профессиональные вкусы и риск. Люди, которые готовы заниматься изнурительной, неприятной работой и очень интенсивно работать в течение многих часов, будут зарабатывать больше. Некоторые увеличивают свои доходы путем «совместительства». Люди также отличаются друг от друга готовностью рисковать (предпринимательский риск).

5. Распределение богатства. Богатство подразумевает наличие некоторого запаса, т. е. реальных материальных и финансовых активов, которые человек накопил за некоторый период времени.

6. Господство на рынке. Способность «искусственно вздувать цены на рынке» в своих интересах, безусловно, является главным фактором, определяющим неравенство доходов. Некоторые профсоюзы прибегают к таким мерам, которые ограничивают предложение оказываемых производственных услуг и тем самым повышают доходы своих членов.

7. Удача, связи и несчастные случаи.

Проблема неравенства граждан по уровню доходов исторически была одним из важнейших объектов экономической теории. Ее анализом занимались многие известные экономисты в силу высокой практической значимости данного вопроса. И все же единым мнением стало обоснование необходимости политики перераспределения доходов, активная роль в которой отводилась государству. Само устройство рыночной экономики делает неизбежным вмешательство государства в сферу доходов с целью их перераспределения. Благодаря этому правительство получает средства, необходимые для удовлетворения общих потребностей (оборона, экология, развитие производственной и социальной инфраструктуры), материальной поддержки временно не занятых в производстве, нетрудоспособных (престарелых и молодежи), а также малообеспеченных групп работников. Кроме того, общество ответственно за уровень доходов работников, занятых в «общественном» секторе экономики (бюджетных отраслях), чьи доходы (заработная плата и жалованье) носят фиксированный характер. Обычно это достигается законодательным установлением минимального уровня заработной платы как обязательной базы оплаты труда во всех сферах экономики. Размер минимальной заработной платы должен обеспечивать минимальный стандарт благосостояния.

Перераспределение доходов (переход от «функциональных» к «вертикальным») правительство осуществляет прямым и косвенным способами, включающими:

— «трансфертные платежи», т. е. пособия, выплачиваемые малообеспеченным группам, иждивенцам, инвалидам, престарелым и безработным;

— «регулирование цен» на социально важную продукцию;

— «индексацию» фиксированных доходов и трансфертных платежей при определенном законом проценте инфляции;

— «обязательный минимум заработной платы» как базы оплаты труда на всех предприятиях;

— «прогрессивное налогообложение», при котором налоговая ставка увеличивается по мере роста размеров номинального дохода.

Изменения в системе налогообложения и процентной ставке — это два мощных инструмента регулирования поведения доходополучателей в рыночной экономике, которыми располагает правительство. Налоги определяют размер реального личного дохода, а процентная ставка, влияя на величину сбережений, определяет размер «потребляемой» части дохода и тем самым величину действительного спроса. Важным элементом государственного регулирования доходов является определение и верхнего предела номинальной заработной платы. Такой предел должен препятствовать развертыванию инфляционной спирали «цена-заработная плата». Эта мера образует основной элемент «политики сдерживания», означая на практике «замораживание» заработной платы и цен (в противоположность «политике экспансии», когда стимулируется рост доходов населения).

Политика сдерживания ограничивает инфляционное превышение платежеспособного спроса над объемом реализуемого совокупного предложения. Осознавая особую социальную значимость перераспределения доходов для обеспечения стабильности рыночного общества, правительство стремится, однако, избежать двух крайностей: формирования иждивенческих настроений у малоимущих и подрыва у экономически активной части общества стремления к высокодоходной деятельности. Стремление к равенству доходов, воплощающему социальную справедливость, всегда сопровождается падением экономической эффективности, ибо незачем эффективно работать ни «бедному» (все равно общество поддержит), ни «богатому» (все равно общество отнимет). Неравенство же в доходах обеспечивает экономическую эффективность, но сопровождается социальной несправедливостью в виде значительной имущественной дифференциации общества.

Таким образом, выбор между равенством и неравенством доходов превращается в выбор между «социальной справедливостью» и «экономической эффективностью».

Рыночное распределение доходов несправедливо, но оно хотя бы в состоянии компенсировать эту несправедливость экономической эффективностью производства, обеспечивающей совокупный продукт в размерах, достаточных для поддержки малоимущих в виде трансфертных платежей и крупных социальных программ (это и есть социально ориентированное рыночное хозяйство).

Справедливое же распределение доходов означает (и это уже доказано практикой) подрыв стимулов к эффективной работе и завершается обычно тем, что справедливо распределять становится просто нечего. «Несправедливая экономическая эффективность» сегодня имеет объективное преимущество перед «неэффективной социальной справедливостью». И хотя их сближение составляет содержание социально-экономического прогресса, в обозримый исторический период названная альтернатива сохраняет свою жесткую однозначность.

Рассмотрим *параметры измерения фактического распределения доходов*: кривую Лоренца, коэффициент Джини и децильный коэффициент.

Кривая Лоренца — это метод графического изображения уровня концентрации явления. Для ее построения на обе оси координат наносят процентную масштабную шкалу (от 0 до 100 %). Для точек кривой абсциссами служат единицы совокупности, а ординатами — значения признака. Равномерное распределение признака будет представлено в таком случае диагональю, называемой «линией равномерного распределения», а неравномерное — «линией Лоренца», отклонение которой от диагонали и характеризует степень неравномерности. Таким образом, если принять величину дохода и численность населения за 100 %, то биссектриса покажет абсолютно равномерное распределение совокупного дохода между всеми группами населения (рис. 9).

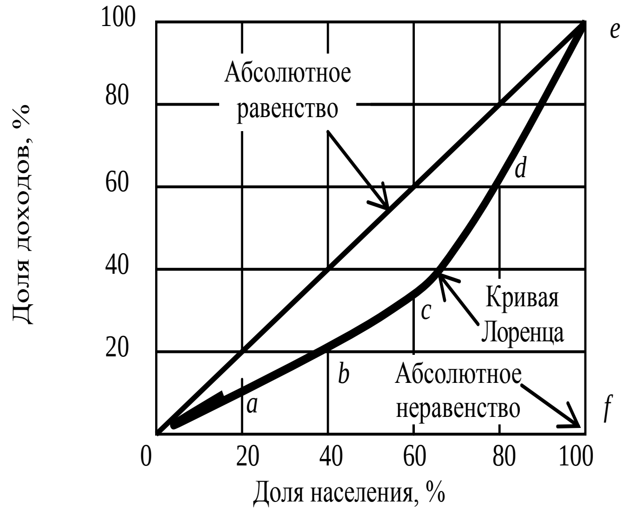


Рис. 9. Кривая Лоренца

Однако реальное распределение всегда будет характеризоваться отклонением от этой прямой. Абсолютно неравномерное распределение совпало бы с осями координат. Но поскольку «сверхбедные» и «сверхбогатые» всегда составляют незначительную часть рыночного общества, то перед нами будет некоторая кривая («кривая Лоренца»), отклонение которой от диагонали наглядно покажет степень неравномерного распределения доходов.

Для расчета конкретного уровня неравенства в распределении доходов поступают следующим образом. Площадь, образованную линиями равномерного и неравномерного распределения доходов (она на графике заштрихована), относят к площади треугольника под этой кривой *Oef*. Полученный результат и есть *коэффициент Джини*.

При коэффициенте, близком к нулю, общество находится в состоянии абсолютной «уравниловки», а при коэффициенте, равном единице, — в ситуации «нищего большинства и сверхбогатого меньшинства». Цивилизованная рыночная экономика исключает подобные крайности благодаря целенаправленному перераспределению доходов. Типичное значение индекса Джини в развитых странах колеблется между 0,2 (Скандинавские страны) и 0,35 (США); для развивающихся стран он составляет 0,4—0,5.

Еще одним параметром, характеризующим неравномерность распределения доходов, является *децильный коэффициент*. Он показывает разрыв в доходах между 10 % богатейших и 10 % беднейших. Мировая практика показывает, что коэффициент дифференциации доходов не должен превышать предельно критическое соотношение10 : 1.

Реально общество всегда живет в области между абсолютным равенством и абсолютным неравенством. Таким образом, неравенство доходов — это та цена, которую обществу приходится платить за ускорение роста общего уровня благосостояния всех граждан страны. Но необходимость такой «платы» никогда не вызывает у людей радости. Напротив, чем выше различия в уровнях жизни между богатыми и бедными, тем сильнее недовольство последних.

Важным условием социального мира в любой стране является предотвращение чрезмерной разницы в доходах наиболее богатых и самых бедных граждан. Для смягчения чрезмерной дифференциации доходов требуется вмешательство государства. Оно осуществляется с помощью прогрессивного налогообложения доходов и систем социальной поддержки. Механизм регулирования дифференциации доходов создан в развитых странах мира для разрешения противоречия между неравной одаренностью людей и размерами собственности, с одной стороны, и необходимостью обеспечить всем людям хотя бы минимально достойный образ жизни — с другой.

Контрольные вопросы и задания

1. Какие демографические модели вы знаете?
2. Приведите классификацию доходов.
3. Какие причины неравенства доходов вы знаете?
4. Объясните смысл кривой Лоренца.
5. Что показывает коэффициент Джинни?

Библиографический список

1. *Булатов, А. С.* Экономика / А. С. Булатов. — Москва : Юристь, 2000.
2. *Математическое* моделирование социальных процессов / под ред. В. И. Добренькова, А. А. Самарского. — Москва, 2000. — Вып. 2.
3. *Толстова, Ю. Н.* Анализ социологических данных / Ю. Н. Толстова. — Москва, 2000.
4. *Толстова, Ю. Н.* Математические методы в социологии / Ю. Н. Толстова // Социология в России / под ред. В. А. Ядова. — Москва, 1998.
5. *Экономическая* теория : учеб. / под ред. И. П. Николаевой. — Москва : Проспект, 2000.

Приложение

Таблица 1

Критические границы отношения *R*/*S*

Объем выборки *n*	Нижние границы						Верхние границы					
	Вероятность ошибки											
	0,000	**0,005**	**0,01**	**0,025**	**0,05**	**0,10**	**0,10**	**0,05**	**0,025**	**0,01**	**0,005**	**0,000**
3	1,732	1,735	1,737	1,745	1,758	1,782	1,997	1,999	2,000	2,000	2,000	2,000
4	1,732	1,83	1,87	1,93	1,98	2,04	2,409	2,429	2,439	2,445	2,447	2,449
5	1,826	1,98	2,02	2,09	2,15	2,22	2,712	2,753	2,782	2,803	2,813	2,828
6	1,826	2,11	2,15	2,22	2,28	2,37	2,949	3,012	3,056	3,095	3,115	3,162
7	1,821	2,22	2,26	2,33	2,40	2,49	3,143	3,222	3,282	3,338	3,369	4,465
8	1,821	2,31	2,35	2,43	2,50	2,59	3,308	3,399	3,471	3,543	3,585	3,742
9	1,897	2,39	2,44	2,51	2,59	2,68	3,449	3,552	3,634	3,720	3,772	4,000
10	1,897	2,46	2,51	2,59	2,67	2,76	3,57	3,685	3,777	3,875	3,935	2,243
11	1,915	2,53	2,58	2,66	2,74	2,84	3,68	3,80	3,903	4,012	4,079	4,472
12	1,915	2,59	2,64	2,72	2,80	2,90	3,78	3,91	4,02	4,134	4,208	4,690
13	1,927	2,64	2,70	2,78	2,86	2,96	3,87	4,00	4,12	4,244	4,325	4,899
14	1,927	2,70	2,75	2,83	2,92	3,02	3,95	4,09	4,21	4,34	4,431	5,099
15	1,936	2,74	2,80	2,88	2,97	3,07	4,02	4,17	4,29	4,44	4,53	5,292
16	1,936	2,79	2,84	2,93	3,01	3,12	4,09	4,24	4,37	4,52	4,62	5,477
17	1,944	2,83	2,88	2,97	3,06	3,17	4,15	4,31	4,44	4,60	4,70	5,657
18	1,944	2,87	2,92	3,01	3,10	3,21	4,21	4,37	4,51	4,67	4,78	5,831
19	1,949	2,90	2,96	3,05	3,14	3,25	4,27	4,43	4,57	4,74	4,85	6,000
20	1,949	2,94	2,99	3,09	3,18	3,29	4,32	4,49	4,63	4,80	4,91	6,164

Таблица 2

Квантили t_p распределения Стьюдента

k	p					
	0,900	**0,950**	**0,975**	**0,990**	**0,995**	**0,999**
1	3,078	6,314	12,706	31,821	63,657	318,00
2	1,886	2,920	4,303	6,965	9,925	22,300
3	1,638	2,353	3,182	4,541	5,841	10,200
4	1,533	2,132	2,776	3,747	4,604	7,173
5	1,476	2,015	2,571	3,365	4,032	5,893
6	1,440	1,943	2,447	3,143	3,707	5,208
7	1,415	1,895	2,365	2,998	3,499	4,785
8	1,397	1,860	2,306	2,896	3,355	4,501
9	1,383	1,833	2,262	2,821	3,250	4,297
10	1,372	1,812	2,228	2,764	3,169	4,144
11	1,363	1,796	2,201	2,718	3,106	4,025
12	1,356	1,782	2,179	2,681	3,055	3,930
13	1,350	1,771	2,160	2,650	3,012	3,852
14	1,345	1,761	2,145	2,624	2,977	3,787
15	1,341	1,753	2,131	2,602	2,947	3,733
16	1,337	1,746	2,120	2,583	2,921	3,686
17	1,333	1,740	2,110	2,567	2,898	3,646
18	1,330	1,734	2,101	2,552	2,878	3,610
19	1,328	1,729	2,093	2,539	2,861	3,579
20	1,325	1,725	2,086	2,528	2,845	3,552
21	1,323	1,721	2,080	2,518	2,831	3,527
22	1,321	1,717	2,074	2,508	2,819	3,505
23	1,319	1,714	2,069	2,500	2,807	3,485
24	1,318	1,711	2,064	2,492	2,797	3,467
25	1,316	1,708	2,060	2,485	2,787	3,450
26	1,315	1,706	2,056	2,479	2,779	3,435
27	1,314	1,703	2,052	2,473	2,771	3,421
28	1,313	1,701	2,048	2,467	2,763	3,408
29	1,311	1,699	2,045	2,462	2,756	3,396
30	1,310	1,697	2,042	2,457	2,750	3,385
40	1,303	1,684	2,021	2,423	2,704	3,307
60	1,296	1,671	2,000	2,390	2,660	3,232
100	1,290	1,660	1,984	2,364	2,626	3,174
∞	1,282	1,645	1,960	2,326	2,576	3,090

Таблица 3

Квантили распределения Фишера $f_{0,99}$ (k_1, k_2)

k_2	k_1									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	p = 0,99									
1	4052	4999,5	5403	5625	5764	5859	5928	5982	6022	6056
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32

Продолжение табл. 3

k_2	k_1								
	12	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**∞**
	$p = 0{,}99$								
1	6 106	6 157	6 209	6 235	6 261	6 287	6 313	6 339	6 350
2	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,50
3	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,22
4	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,64
8	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
26	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

Продолжение табл. 3

Квантили распределения Фишера $f_{0,975}$ (k_1, k_2)

k_2	k_1									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	p = 0,975									
1	647,80	799,50	864,20	899,60	921,80	937,10	948,20	956,70	963,30	968,60
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92
18	5,98	4,56	3,95	3,61	3,33	3,22	3,10	3,01	2,93	2,87
19	5,92	4,51	3,90	3,56	3,29	3,17	3,05	2,96	2,88	2,82
20	5,87	4,46	3,86	3,51	3,25	3,13	3,01	2,91	2,84	2,77
21	5,83	4,42	3,82	3,48	3,22	3,09	2,97	2,87	2,80	2,73
22	5,79	4,38	3,78	3,44	3,18	3,05	2,93	2,84	2,76	2,70
23	5,75	4,35	3,75	3,41	3,15	3,02	2,90	2,81	2,73	2,67
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,65	2,59
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,63	2,57
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,59	2,53
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05

Продолжение табл. 3

k_2	k_1								
	12	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**∞**
	p = 0,975								
1	976,7	984,9	993,1	997,2	1 001	1 006	1 010	1 014	1 018
2	39,41	39,43	39,45	39,46	39,46	39,47	39,48	39,49	40,50
3	14,34	14,25	14,17	14,12	14,08	14,04	13,99	13,95	13,09
4	8,75	8,66	8,56	8,51	8,46	8,41	8,36	8,31	8,26
5	6,52	6,43	6,33	6,28	6,23	6,18	6,12	6,07	6,01
6	5,37	5,27	5,17	5,12	5,07	5,01	4,96	4,90	4,84
7	4,67	4,57	4,47	4,42	4,36	4,31	4,25	4,20	4,14
8	4,20	4,10	4,00	3,95	3,89	3,84	3,78	3,73	3,67
9	3,87	3,77	3,67	3,61	3,56	3,51	3,45	3,39	3,33
10	3,62	3,52	3,42	3,37	3,31	3,26	3,20	3,14	3,08
11	3,43	3,33	3,23	3,17	3,12	3,06	3,00	2,94	2,88
12	3,28	3,18	3,07	3,02	2,96	2,91	2,85	2,79	2,72
13	3,15	3,05	2,95	2,89	2,84	2,78	2,72	2,66	2,60
14	3,05	2,95	2,84	2,79	2,73	2,67	2,61	2,55	2,49
15	2,96	2,86	2,76	2,70	2,64	2,59	2,52	2,46	2,40
16	2,89	2,79	2,68	2,63	2,57	2,51	2,45	2,38	2,32
17	2,82	2,72	2,62	2,56	2,50	2,44	2,38	2,32	2,25
18	2,77	2,67	2,56	2,50	2,44	2,38	2,32	2,26	2,19
19	2,72	2,62	2,51	2,45	2,39	2,33	2,27	2,20	2,13
20	2,68	2,57	2,46	2,41	2,35	2,29	2,22	2,16	2,09
21	2,64	2,53	2,42	2,37	2,31	2,25	2,18	2,11	2,08
22	2,60	2,50	2,39	2,33	2,27	2,21	2,14	2,08	2,00
23	2,57	2,47	2,36	2,30	2,24	2,18	2,11	2,04	1,96
24	2,54	2,44	2,33	2,27	2,21	2,15	2,08	2,01	1,94
25	2,51	2,41	2,30	2,24	2,18	2,12	2,05	1,98	1,91
26	2,49	2,39	2,28	2,22	2,16	2,09	2,03	1,95	1,88
27	2,47	2,36	2,25	2,19	2,13	2,07	2,00	1,93	1,85
28	2,45	2,34	2,23	2,17	2,11	2,05	1,98	1,91	1,83
29	2,43	2,32	2,21	2,15	2,09	2,03	1,96	1,89	1,81
30	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87	1,77
40	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72	1,64
60	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58	1,48
120	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43	1,31
∞	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27	1,00

Продолжение табл. 3

Квантили распределения Фишера $f_{0,95}(k_1, k_2)$

k_2	k_1									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	$p = 0,95$									
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9
2	15,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,05
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,90	2,90	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83

Продолжение табл. 3

k_2	k_1								
	12	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**∞**
	p = 0,95								
1	243,90	245,9	248,0	249,1	250,1	251,1	252,2	253,30	254,30
2	19,41	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	2,91	2,85	2,77	2,74	2,70	2,66	3,20	2,58	2,54
11	2,79	2,72,	2,65	2,61	2,57	2,53	3,00	2,45	2,40
12	2,69	2,62	2,54	2,51	2,47	2,43	2,85	2,34	2,30
13	2,60	2,53	2,46	2,42	2,38	2,34	2,72	2,25	2,21
14	2,53	2,46	2,39	2,35	2,31	2,27	2,61	2,18	2,13
15	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	2,12	2,04	1,96	1,84	1,87	1,82	1,77	1,71	1,65
29	2,10	2,03	1,94	1,81	1,85	1,81	1,75	1,70	1,64
30	2,09	2,01	1,93	1,78	1,84	1,79	1,74	1,68	1,62
40	2,00	1,92	1,84	1,76	1,74	1,69	1,64	1,58	1,51
60	1,92	1,84	1,75	1,88	1,65	1,59	1,53	1,47	1,39
120	1,83	1,75	1,66	1,76	1,55	1,50	1,43	1,35	1,25
∞	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Продолжение табл. 3

Квантили распределения Фишера $f_{0,90}(k_1, k_2)$

k_2	k_1									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	p = 0,90									
1	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44	59,86	60,19
2	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37	9,38	9,39
3	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25	5,24	5,23
4	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95	3,94	3,92
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32
11	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30	2,27	2,25
12	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24	2,21	2,19
13	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20	2,16	2,14
14	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15	2,12	2,10
15	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12	2,09	2,06
16	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09	2,06	2,03
17	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06	2,03	2,00
18	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04	2,00	1,98
19	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02	1,98	1,96
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94
21	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98	1,95	1,92
22	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97	1,93	1,90
23	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95	1,92	1,89
24	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94	1,91	1,88
25	2,92	2,53	2,32	2,18	2,09	2,02	1,97	1,93	1,89	1,87
26	2,91	2,52	2,31	2,17	2,08	2,01	1,96	1,92	1,88	1,86
27	2,90	2,51	2,30	2,17	2,07	2,00	1,95	1,91	1,87	1,85
28	2,89	2,50	2,29	2,16	2,06	2,00	1,94	1,90	1,87	1,84
29	2,89	2,50	2,28	2,15	2,06	1,99	1,93	1,89	1,86	1,83
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76
60	2,79	2,39	2,18	2,04	1,95	1,87	1,82	1,77	1,74	1,71
120	2,75	2,35	2,13	1,99	1,90	1,82	1,77	1,72	1,68	1,65
∞	2,71	2,30	2,08	1,94	1,85	1,77	1,72	1,67	1,63	1,60

Окончание табл. 3

k_2	k_1								
	12	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**∞**
	p = 0,90								
1	60,71	61,22	61,74	62,00	62,26	62,53	62,79	63,06	63,33
2	9,41	9,42	9,44	9,45	9,46	9,47	9,47	9,48	9,49
3	5,22	5,20	5,18	5,18	5,17	5,16	5,15	5,14	5,13
4	3,90	3,87	3,84	3,83	3,82	3,80	3,79	3,78	3,76
5	3,27	3,24	3,21	3,19	3,17	3,16	3,14	3,12	3,10
6	2,90	2,87	2,84	2,82	2,80	2,78	2,76	2,74	2,72
7	2,67	2,63	2,59	2,58	2,56	2,54	2,51	2,49	2,47
8	2,50	2,46	2,42	2,40	2,38	2,36	2,34	2,32	2,29
9	2,38	2,34	2,30	2,28	2,25	2,23	2,21	2,18	2,16
10	2,28	2,24	2,20	2,18	2,16	2,13	2,11	2,08	2,06
11	2,21	2,17	2,12	2,10	2,08	2,05	2,03	2,00	1,97
12	2,15	2,10	2,06	2,04	2,01	1,99	1,96	1,93	1,90
13	2,10	2,05	2,01	1,98	1,96	1,93	1,90	1,88	1,85
14	2,05	2,01	1,96	1,94	1,91	1,89	1,86	1,83	1,80
15	2,02	1,97	1,92	1,90	1,87	1,85	1,82	1,79	1,76
16	1,99	1,94	1,89	1,87	1,84	1,81	1,78	1,75	1,72
17	1,96	1,91	1,86	1,84	1,81	1,78	1,75	1,72	1,69
18	1,93	1,89	1,84	1,81	1,78	1,75	1,72	1,69	1,66
19	1,91	1,86	1,81	1,79	1,76	1,73	1,70	1,67	1,63
20	1,89	1,84	1,79	1,77	1,74	1,71	1,68	1,64	1,61
21	1,87	1,83	1,78	1,75	1,72	1,69	1,66	1,62	1,59
22	1,86	1,81	1,76	1,73	1,70	1,67	1,64	1,60	1,57
23	1,84	1,80	1,74	1,72	1,69	1,66	1,62	1,59	1,55
24	1,83	1,78	1,73	1,70	1,67	1,64	1,61	1,57	1,53
25	1,82	1,77	1,72	1,69	1,66	1,63	1,59	1,56	1,52
26	1,81	1,76	1,71	1,68	1,65	1,61	1,58	1,54	1,50
27	1,80	1,75	1,70	1,67	1,64	1,60	1,57	1,53	1,49
28	1,79	1,74	1,69	1,66	1,63	1,59	1,56	1,52	1,48
29	1,78	1,73	1,68	1,65	1,62	1,58	1,55	1,51	1,47
30	1,77	1,72	1,67	1,64	1,61	1,57	1,54	1,50	1,46
40	1,71	1,66	1,61	1,57	1,54	1,51	1,47	1,42	1,38
60	1,66	1,60	1,54	1,51	1,48	1,44	1,40	1,35	1,29
120	1,60	1,55	1,48	1,45	1,41	1,37	1,32	1,26	1,19
∞	1,55	1,49	1,42	1,38	1,34	1,30	1,24	1,17	1,00

Таблица 4

Значения функции распределения $F_{(0,1)}(x)$ нормального закона $N(0,1)$;

$$F_{(0,1)}(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$$

x	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5835	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,853 1	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,935 7	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,983 4	0,9838	0,9842	0,9846	0,9850	0.9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,994 0	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

$$F_{(0,1)}(-x)=1-F_{(0,1)}(x)$$

Квантили Z_p нормального распределения $N(0,1)$:

p	0,90	0,95	0,975	0,99	0,995	0,999	0,9995
Z_p	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Значения функции Лапласа вычисляются следующим образом:

$$\Phi_0(x)=F_{(0,1)}(x)-0,5.$$

Таблица 5

Квантили распределения χ^2(«хи-квадрат»)

k	*p*												
	0,005	**0,010**	**0,025**	**0,05**	**0,10**	**0,20**	**0,80**	**0,90**	**0,95**	**0,975**	**0,990**	**0,995**	**0,999**
1					0,02	0,06	1,64	2,71	3,84	5,02	6,63	7,88	10,8
2	0,010	0,020	0,051	0,10	0,21	0,45	3,22	4,61	5,99	7,38	9,21	10,6	13,8
3	0,072	0,115	0,216	0,35	0,58	1,00	4,64	6,25	7,81	9,35	11,3	12,8	16,3
4	0,207	0,297	0,484	0,71	1,06	1,65	5,99	7,78	9,49	11,1	13,3	14,9	18,5
5	0,412	0,554	0,831	1,15	1,61	2,34	7,29	9,24	11,1	12,8	15,1	16,7	20,5
6	0,676	0,872	1,24	1,64	2,20	3,07	8,56	10,6	12,6	14,4	16,8	18,5	22,5
7	0,989	1,24	1,69	2,17	2,83	3,82	9,80	12,0	14,1	16,0	18,5	20,3	24,3
8	1,34	1,65	2,18	2,73	3,49	4,59	11,0	13,4	15,5	17,5	20,1	22,0	26,1
9	1,73	2,09	2,70	3,33	4,17	5,38	12,2	14,7	16,9	19,0	21,7	23,6	27,9
10	2,16	2,56	3,25	3,94	4,87	6,18	13,4	16,0	18,3	20,5	23,2	25,2	29,6
11	2,60	3,05	3,82	4,57	5,58	6,99	14,6	17,3	19,7	21,9	24,7	26,8	31,3
12	3,07	3,57	4,40	5,23	6,30	7,81	15,8	18,5	21,0	23,3	26,2	28,3	32,9
13	3,57	4,11	5,01	5,89	7,04	8,63	17,0	19,8	22,4	24,7	27,7	29,8	34,5
14	4,07	4,66	5,63	6,57	7,79	9,47	18,2	21,1	23,7	26,1	29,1	31,3	36,1
15	4,60	5,23	6,26	7,26	8,55	10,3	19,3	22,3	25,0	27,5	30,6	32,8	37,7
16	5,14	5,81	6,91	7,96	9,31	11,2	20,5	23,5	26,3	28,8	32,0	34,3	39,3
17	5,70	6,41	7,56	8,67	10,1	12,0	21,6	24,8	27,6	30,2	33,4	35,7	40,8
18	6,26	7,01	8,23	9,39	10,9	12,9	22,8	26,0	28,9	31,5	34,8	37,2	42,3
19	6,84	7,63	8,91	10,1	11,7	13,7	23,9	27,2	30,1	32,9	36,2	38,6	43,8
20	7,43	8,26	9,59	10,9	12,4	14,6	25,0	28,4	31,4	34,2	37,6	40,0	45,3
21	8,03	8,90	10.3	11,6	13,2	15,4	26,9	29,6	32,7	35,5	38,9	41,4	46,8
22	8,64	9,54	11,0	12,3	14,0	16,3	27,3	30,8	33,9	36,8	40,3	42,8	48,3
23	9,26	10,2	11,7	13,1	14,8	17,2	28,4	32,0	35,2	38,1	41,6	44,2	49,7
24	9,89	10,9	12,4	13,8	15,7	18,1	29,6	33,2	36,4	39,4	43,0	45,6	51,2
25	10,5	11,5	13,1	14,6	16,5	18,9	30,7	34,4	37,7	40,6	44,3	46,9	52,6
26	11,2	12,2	13,8	15,4	17,3	19,8	31,8	35,6	38,9	41,9	45,6	48,3	54,1
27	11,8	12,9	14,6	16,2	18,1	20,7	32,9	36,7	40,1	43,2	47,0	49,6	55,5
28	12,5	13,6	15,3	16,9	18,9	21,6	34,0	37,9	41,3	44,5	48,3	51,0	56,9
29	13,1	14,3	16,0	17,7	19,8	22,5	35,1	39,1	42,6	45,7	49,6	52,3	58,3
30	13,8	15,0	16,8	18,5	20,6	23,4	36,3	40,3	43,8	47,0	50,9	53,7	59,7
35	17,2	18,5	20,6	22,5	24,8	27,8	41,8	46,1	49,8	53,2	57,3	60,3	66,6
40	20,7	22,2	24,4	26,5	29,1	32,3	47,3	51,8	55,8	59,3	63,7	66,8	73,4
41	21,4	22,9	25,2	27,3	29,9	33,3	48,4	52,9	56,9	60,6	64,9	68,1	74,7
42	22,1	23,6	26,0	28,1	30,8	32,2	49,5	54,1	58,1	61,8	66,2	69,3	76,1
43	22,9	24,4	26,8	29,0	31,6	35,1	50,5	55,2	59,3	63,0	67,5	70,6	77,4
44	23,3	25,,1	27,6	29,8	32,5	36,0	51,6	56,4	60,5	64,2	68,7	71,9	78,7
45	24,3	25,9	28,4	30,6	33,4	36,9	52,7	57,5	61,7	65,4	70,0	73,2	80,1
50	28,0	29,7	32,4	34,8	37,7	41,4	58,2	63,2	67,5	71,4	76,2	79,5	86,7
55	31,7	33,6	36,4	39,0	42,1	46,0	63,6	68,8	73,3	77,4	82,3	85,7	93,2
60	35,5	37,5	40,5	43,2	46,5	50,6	69,0	74,4	79,1	83,3	92,0	99,6	100,6
75	47,2	49,5	52,9	56,1	59,8	64,5	85,1	91,1	96,2	100,8	106.4	110,3	118,6
90	59,2	61,8	65,6	69,1	73,3	78,6	101	107,6	113,1	118,1	124,1	128,3	137,2
100	67,3	70,1	74,2	77,9	82,4	87,9	112	118,5	124,3	129,6	135,6	140,2	149,4

Учебное издание

Александр Леонидович Осипов
Евгений Алексеевич Рапоцевич

Основы математического моделирования социально-экономических процессов

УЧЕБНОЕ ПОСОБИЕ

для студентов всех форм обучения
по направлению 081100.62 — Государственное и муниципальное управление

Редактор В. В. Попова
Корректор Н. А. Внукова
Технический редактор О. А. Гладунова

Подписано в печать 28.03.2014. Бумага офсетная. Печать ОСЕ.
Гарнитура Times New Roman.
Формат 60х84 1/8. Уч.-изд. л. 10,29. Усл. п. л. 18,14. Тираж 122 экз. Заказ 29.
630102, г. Новосибирск, ул. Нижегородская, 6.
Сибирский институт управления — филиал РАНХиГС.