Министерство науки и высшего образования Российской Федерации

Амурский государственный университет

Математическое моделирование

Учебно-методическое пособие

Благовещенск

Издательство АмГУ

2019

УДК 519.8

Рекомендовано
учебно-методическим советом университета

Рецензенты:

Бризицкий Р.В., с. н. с. Института прикладной математики ДВО РАН, канд. физ.-мат. наук,

Максимова, Н.Н.

Математическое моделирование. Учебно-методическое пособие / сост. Н.Н. Максимова. – Благовещенск: Изд-во АмГУ, 2019. – 88 с.

Учебно-методическое пособие содержит теоретические сведения и примеры задач по курсу «Математическое моделирование». Подробно рассмотрены моделирование задач линейного программирования, транспортных задач, динамического программирования, систем массового обслуживания. Учебный материал позволяет получить теоритические знания и выработать практические навыки построения и исследования рассмотренных в пособии типов моделей.

Учебно-методическое пособие предназначено для студентов, обучающихся по специальности 09.02.03 – «Программирование в компьютерных системах» (в рамках изучения дисциплины «Математическое моделирование»), а также будет полезно для студентов, обучающихся по направлению 01.03.02 – «Прикладная математика и информатика» (в рамках изучения дисциплины «Теория игр и исследование операций»).

ВВЕДЕНИЕ

Целью данного пособия является изложение теоретических сведений приведение и примеров по курсу «Математическое моделирование». Подробно рассмотрены моделирование задач линейного программирования, транспортных задач, динамического программирования, систем массового обслуживания.

Пособие состоит из введения, шести глав и библиографического списка. В первых двух главах основные понятия, связанные с моделирование вообще и математическим моделированием в частности. В третьей главе приводится понятие моделей линейного программирования, в четвертой – моделей транспортного типа, в пятой – моделей динамического программирования, в шестой – моделей систем массового облуживания. В каждой задаче представлены примеры моделей, а так рассмотрены методы их исследования.

Учебно-методическое пособие предназначено для студентов, обучающихся по специальности 09.02.03 – «Программирование в компьютерных системах» (в рамках изучения дисциплины «Математическое моделирование»), а также будет полезно для студентов, обучающихся по направлению 01.03.02 – «Прикладная математика и информатика» (в рамках изучения дисциплины «Теория игр и исследование операций»). Пособие может быть полезно студентам всех форм и ступеней обучения. Для более глубокого изучения методов математического моделирования рекомендован весьма обширный список литературы.

ГЛАВА 1. Введение в дисциплину «Математическое моделирование»

1.1. Метод моделирования

При познании окружающего мира человек имеет дело не непосредственно с реальными объектами, а с их образами, сформированными с помощью органов чувств, измерительных приборов, аналитического оборудования, оргтехники и абстрактного мышления.

Все предметы, явления, процессы окружающего мира, на которые направлена человеческая деятельность, называются ***объектами***. Объекты материального мира существуют вне нашего сознания и взаимодействуют между собой и внешней средой. Познание осуществляется посредством ощущений, которые создают в нашем сознании образы материальных объектов.

Объекты материального мира сложны и многообразны. Отражение всех их свойств в создаваемых, изучаемых и используемых образах весьма затруднительно, да и не нужно. Важно, чтобы образ объекта содержал черты, наиболее важные для его использования.

В узком смысле ***моделирование*** – это метод научного исследования (познания) окружающего нас мира, заключающийся в подмене реальных объектов или явлений их заведомо упрощенными образами (моделями) с целью изучения этих образов и последующего переноса полученных результатов и выводов на объекты и явления реального мира.

В широком смысле ***моделирование*** представляет собой научную дисциплину, в рамках которой изучаются методы построения и использования моделей для познания реального мира.

Методом моделирования называется замена объекта-оригинала объектом-заместителем, обладающим определенным сходством с оригиналом, с целью получения новой информации об оригинале.

Моделью называется объект-заместитель объекта-оригинала, предназначенный для получения информации об оригинале.

Связь (отношение) между объектом реального мира и его моделью можно проиллюстрировать графически с помощью укрупненного цикла моделирования (рисунок 1.1).

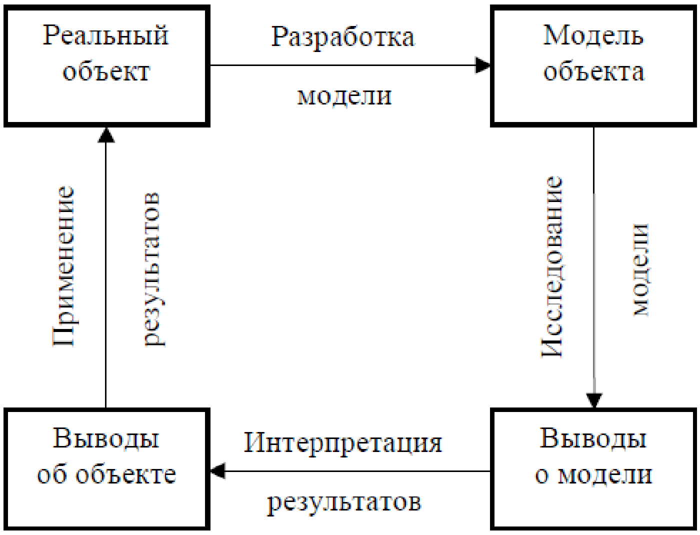

Рисунок 1.1 – Укрупненный цикл моделирования

1.2. Системный подход к моделированию

Изучаемые объекты, как правило, очень сложны. Для упрощения их изучения удобно разбить их на части, и изучать каждую из частей отдельно. При этом важно учитывать, что части находятся во взаимодействии и не являются независимыми друг от друга. Для правильного описания поведения взаимодействующих объектов используется ***системных подход***, заключающийся представления сложного объекта в виде системы взаимодействующих элементов.

Системой называется совокупность взаимодействующих элементов, объединенных наличием общей цели. Элементы системы имеют связи как между собой, так и с внешней средой (объектами, не принадлежащими к системе).

В системе могут быть выделены подсистемы. ***Подсистемой*** называется часть системы, имеющая собственную (локальную) цель, согласованную с целью системы. ***Элементом*** называется неделимая часть системы. Иерархии подсистем соответствует иерархия целей, или ***дерево целей*** (схема, отражающая иерархию целей, напоминает перевернутое дерево).

Сущность системного подхода к моделированию заключается в единстве процессов декомпозиции и композиции.

Декомпозицией называется метод, основанный на использовании структуры системы и позволяющий заменить решение одной большой задачи решением нескольких более простых задач. Декомпозиция системы производится путем последовательного применения структурного и функционального подходов. ***Структурный подход*** заключается в моделировании структуры системы, т.е. разбиении ее на подсистемы и элементы; ***функциональный подход*** предполагает построение модели каждого элемента на основе анализа его поведения без использования информации о структуре.

Композицией называется моделирование связей подсистем и элементов между собой и с внешней средой. Связь между элементами осуществляется через множество параметров, которые для одних элементов являются входными (влияющими на их функционирование), а для других – выходными (описывающими результат их функционирования).

1.3. Классификация видов моделирования

Классификацию видов моделирования и, соответственно, моделей (от лат. *modulus* – мера, образец) можно проводить по разным признакам: по сфере приложения (области применения), по характеру моделируемых объектов, по степени подробности моделей и т.д. В этом случае различают две большие группы моделей, относящихся соответственно к материальному и идеальному моделированию (рисунок 1.2).

Материальное моделирование предполагает наличие связи, имеющей материальный характер, между моделью и исследуемым объектом. В материальном моделировании можно условно выделить три основные группы методов: пространственное, физическое и аналоговое моделирование.

В *пространственном моделировании* используются модели, предназначенные для воспроизведения или отображения пространственных (геометрических) свойств изучаемых объектов. В качестве примеров такой группы моделей можно назвать макеты разнообразных типов (зданий, устройств и т.д.).

В *физическом моделировании* используются модели, предназначенные для воспроизведения динамики процессов, происходящих в изучаемых объектах, причем общность процессов, происходящих в объекте исследования и модели, основывается на сходстве их физической природы. Наиболее известным примером физического моделирования является исследование летательных аппаратов на основе экспериментов в аэродинамической трубе.

В *аналоговом моделировании* используются материальные модели, физическая природа которых отличается от природы исследуемых объектов, но, вместе с тем, они описываются сходными математическими соотношениями, т.е. связь между моделью и объектом основывается на аналогии их математического описания.

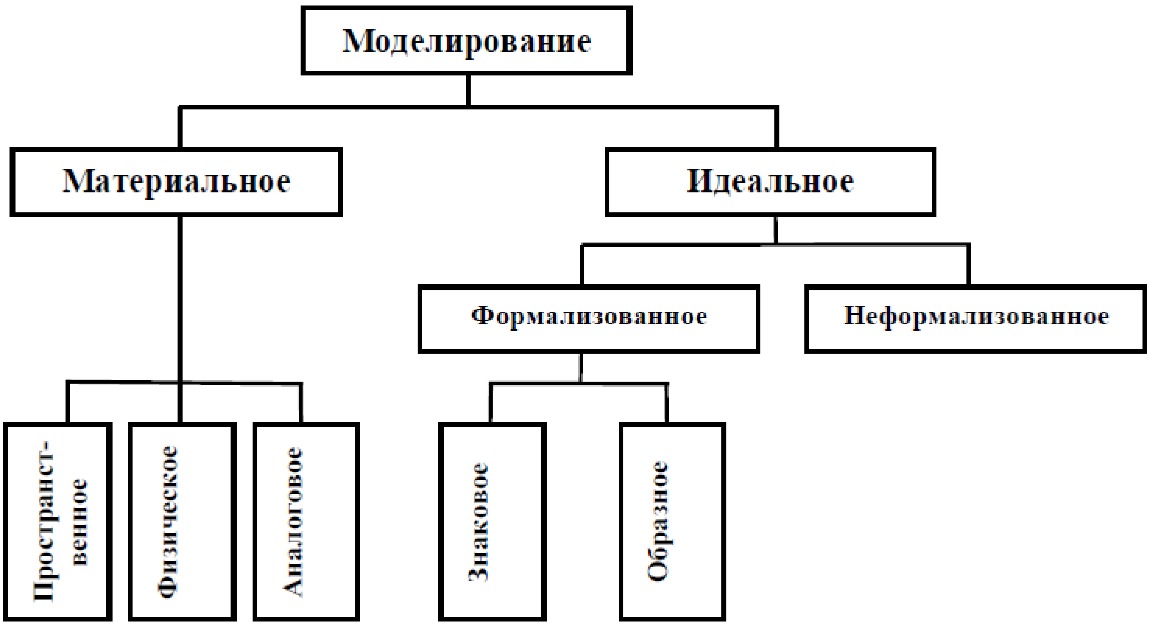

Рисунок 1.2 – Классификация видов моделирования

Необходимо отметить, что во всех случаях материального моделирования модель – это материальное отражение изучаемого (исходного) объекта. Исследование состоит в материальном воздействии на нее, т.е. в эксперименте с моделью. Таким образом, материальное моделирование по своей природе является экспериментальным методом.

Идеальное моделирование принципиально отличается от материального, поскольку оно основывается не на материальной аналогии между моделью и изучаемым объектом, а на идеальной, т.е. мыслимой связи между ними. В *формализованном моделировании* моделями служат системы знаков или образов, вместе с которыми задаются правила их преобразования и интерпретации. В *знаковом (символьном) моделировании* в качестве моделей используются системы знаков, которые могут существенно отличаться друг от. Важнейшим видом знакового моделирования является *математическое моделирование*, при использовании которого модель записывается в виде совокупности формул, преобразуемых на основе правил логики и математики.

В *образном моделировании* при построении модели используются такие наглядные элементы, как упругие шары, потоки жидкости, траектории движения тел.

Неформализованное моделирование – это анализ проблем разнообразного типа, когда модель не формулируется, а вместо нее используется некоторое, не зафиксированное точно, мысленное ощущение реальности, служащее основой для рассуждения и принятия решений.

1.4. Классификация математических моделей

Математические модели относятся к символьным моделям и представляют собой описание объектов в виде математических символов, формул, выражений. При наличии достаточно точной математической модели можно путем математических расчетов прогнозировать результаты функционирования объекта при различных условиях, выбрать из множества возможных вариантов тот, который дает наилучшие результаты

Рассмотрим классификацию математических моделей.

1. По способу построения модели подразделяются на ***аналитические (теоретические), статистические (эмпирические) и комбинированные.***

Аналитические модели строятся на основе информации, содержащейся в известных законах природы, например, законах сохранения энергии, массы, импульса, электрического заряда, Ома, Кирхгофа, Архимеда и т.п. Объект, для которого строится аналитическая модель, должен быть хорошо изучен. Как правило, при построении аналитических моделей используются различные допущения и упрощения, снижающие точность моделирования. Основным достоинством аналитических моделей является их универсальность.

Статистические модели строятся на основании обработки экспериментальных данных. Проводится ряд экспериментов, при которых фиксируются значения входных и выходных параметров, после чего производится статистическая обработка результатов экспериментов, на основании которой подбирается математическое выражение, описывающее экспериментальные данные с достаточной точностью. Основным достоинством статистических моделей является простота их построения, основным недостатком – низкая универсальность.

Наилучший результат дают *комбинированные модели,* сочетающие достоинства аналитических и статистических моделей.

2. ***Одномерные и многомерные модели*** различают по количеству входных переменных, входящих в модель.

3. ***Линейные и нелинейные модели.*** Для линейных моделей справедлив принцип суперпозиции, согласно которому реакция объекта на суммарное воздействие равна сумме реакций объекта на элементарные воздействия:

$$Y\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} Y(x_i). \qquad (1.1)$$

4. ***Статические и динамические модели.*** В *динамических моделях* переменные зависят от времени, в *статических* – не зависят. Динамические модели проще статических, и при необходимости они иногда преобразуются в статические. Статические модели описывают также установившиеся режимы, когда переходные процессы закончились.

5. ***Стационарные и нестационарные модели***

Стационарные модели описывают процессы, инвариантные относительно времени начала процесса. *Нестационарные модели* описывают процессы, течение которых зависит от времени их начала. Статические модели являются частными случаями стационарных моделей и описывают функционирование системы в установившихся условиях.

6. ***Модели с параметрами, сосредоточенными или распределенными в пространстве.*** В моделях с распределенными параметрами переменные зависят от пространственных координат, в моделях с сосредоточенными параметрами – не зависят. Динамические модели с сосредоточенными параметрами описываются обыкновенными дифференциальными уравнениями. Динамические модели с распределенными параметрами описываются дифференциальными уравнениями с частными производными.

7. ***Модели, дискретные и непрерывные*** (во времени). В дискретных моделях время принимает фиксированные значения, в непрерывных – любые значения. При использовании дискретных моделей ось времени разбивается на интервалы, границы которых называются опорными моментами времени. Расчет параметров модели производится для опорных моментов времени.

8. ***Детерминированные и стохастические (вероятностные) модели.*** Детерминированные модели являются воспроизводимыми: при одинаковых условиях модель всегда дает один и тот же результат. В стохастических моделях некоторые параметры являются случайными величинами, и результаты моделирования при каждой реализации отличаются друг от друга.

1.5. Свойства математических моделей и требования к ним

При разработке математической модели устанавливается ряд требований к ее свойствам, выполнение которых необходимо для ее эффективного использования. Рассмотрим основные из них.

Целенаправленность модели. В модели должны фигурировать параметры, описывающие цель объекта, а так же параметры, с помощью управления которыми можно добиться достижения цели.

Точность модели определяется величинами погрешности, с которыми рассчитываются выходные параметры. Погрешности подразделяются на систематические и случайные. Систематическая погрешность характеризует среднее отклонение между вычисленными и экспериментальными значениями выходного параметра, а случайная (среднеквадратичная) погрешность σ – среднеквадратичное отклонение экспериментальных значений от вычисленных.

Непротиворечивость модели характеризует отсутствие абсурдных ответов и выводов при использовании модели. Модель проверяется также на противоречия между выводами, которые можно сделать из модели и из экспериментальных данных.

Реалистичность модели оценивается путем также расчета типовых примеров, для которых заранее известен результат (точный или ориентировочный).

Устойчивостью модели называется слабая чувствительность к погрешностям ее параметров. Неустойчивость модели является ее свойством и не всегда свидетельствует о неустойчивости описываемых ею объектов.

Удобство использования является одним из основных свойств математических моделей, что обусловлено самим методом моделирования. Это требование, в частности, должно предусматривать удобство реализации в виде компьютерных программ.

Универсальность модели обеспечивает описание с помощью нее как можно более широкий класс объектов.

Адаптивность и возможность изменения. Модели, обладающие этими свойствами можно корректировать при изменении окружающих условий и совершенствовать для улучшения ее свойств.

Экономичность, простота, физический смысл. Требование экономичности модели подразумевает минимизацию затрат на ее разработку и реализацию (в частности, время, необходимое для компьютерных расчетов). Наличие физического смысла полезно для изучения модели с целью избежания возможных ошибок. Принцип простоты заключается в том, что из нескольких моделей с одинаковыми другими свойствами нужно выбрать наиболее простую.

Адекватность математической модели является ее интегральным свойствам, объединяющим другие наиболее важные свойства. Если свойства модели удовлетворяют требованиям, говорят, что она адекватна (оригиналу), в противном случае – не адекватна.

ГЛАВА 2. Математическое моделирование и вычислительный эксперимент

2.1. Предмет математического моделирования

Математическая модель – это приближенное описание какого-либо класса явлений или объектов реального мира на языке математики.

Основная цель моделирования – исследовать эти объекты и предсказать результаты будущих наблюдений. Однако моделирование – это еще и метод познания окружающего мира, дающий возможность управлять им.

Математическое моделирование и связанный с ним компьютерный эксперимент незаменимы в тех случаях, когда натурный эксперимент невозможен или затруднен по тем или иным причинам.

Отметим некоторые существенные для исследований особенностях механических систем и процессов.

Во-первых, факторы, определяющие такие объекты, характеризуются, как измеримые величины – параметры.

Во-вторых, в основе таких моделей лежат уравнения, описывающие фундаментальные законы природы (механики), не нуждающиеся в пересмотре и уточнении.

В-третьих, наибольшую трудность при разработке моделей механических систем и процессов представляет описание недостоверно известных характеристик объекта, как функциональных, так и числовых.

В-четвертых, современные требования к таким моделям приводят к необходимости учета множества факторов, влияющих на поведение объекта, не только таких, которые связаны известными законами природы.

2.2. Этапы построения математических моделей: содержательная, концептуальная и математическая постановка задачи моделирования

1. Содержательная постановка

Необходимость в новой модели может появиться в связи с проведением научных исследований (особенно – на стыке различных областей знания), выполнением проектных и конструкторских работ на производстве, созданием систем автоматического управления, планирования и контроля. Человека или организацию, заинтересованных в разработке новой математической модели, для краткости будем называть **заказчиком**. После принятия решения о необходимости построения новой математической модели заказчик ищет **исполнителя** своего заказа.

Основной целью данного этапа является подготовка содержательной постановки задачи моделирования. Перечень сформулированных в содержательной (словесной) форме основных вопросов об объекте моделирования, интересующих заказчика, составляет **содержательную постановку задачи моделирования**.

2. Концептуальная постановка задачи моделирования

Концептуальная постановка задачи моделирования – это сформулированный в терминах конкретных дисциплин (физики, химии, биологии и т.д.) перечень основных вопросов, интересующих заказчика, а также совокупность гипотез относительно свойств и поведения объекта моделирования. Наибольшие трудности при формулировке концептуальной постановки приходится преодолевать в моделях, находящихся на “стыке” различных дисциплин.

3. Математическая постановка задачи моделирования

Законченная концептуальная постановка позволяет сформулировать математическую постановку задачи моделирования, включающую совокупность различных математических соотношений, описывающих поведение и свойства объекта моделирования. **Математическая постановка задачи моделирования** – это совокупность математических соотношений, описывающих поведение и свойства объекта моделирования. Совокупность математических соотношений определяет вид оператора модели.

Для создания математических моделей сложных систем и процессов, применимых для широкого класса реальных задач требуется привлечение большого объема знаний, накопленных в рассматриваемой дисциплине (а в некоторых случаях и в смежных областях).

Математические модели основываются на математическом описании объекта. В математическое описание, прежде всего, входят, и это естественно, взаимосвязи параметров объекта, что характеризует его особенности функционирования. Такие связи могут представляться в виде:

– вектор-функций $\bar{y} = f(\bar{x}, t)$,

– неявных функций $F(\bar{x}, \bar{y}, t) = 0$,

– обыкновенных дифференциальных уравнений $F(\bar{x}, \bar{x}', \bar{x}'', \ldots, \bar{x}^{(m)}, t) = 0$,

– дифференциальных уравнений с частными производными $F(\bar{x}, \bar{y}, t, \frac{\partial \bar{x}}{\partial t}, \frac{\partial \bar{y}}{\partial t}, \ldots) = 0$,

– интегро-дифференциальных уравнений $\int_{\Omega} F(\bar{x}, \bar{y}, t, \frac{\partial \bar{x}}{\partial t}, \frac{\partial \bar{y}}{\partial t}, \ldots) d\Omega = 0$,

– оптимизационных задач $\underset{\bar{x} \in X}{extr}\, f(\bar{x})$,

– вычислительного алгоритма,

– вероятностного (стохастического) описания.

Под математическим описанием понимается полная совокупность данных, функций и методов вычисления, позволяющая получать результат.

Простейшая классификация математических моделей в зависимости от природы объекта, решаемых задач и применяемых методов, происходит следующим образом:

– *линейные или нелинейные* (описываемые функциями, которые содержат основные параметры только в степени 0 и 1, или любыми видами функций),

– *стационарные или нестационарные* (независящие или зависящие от времени),

– *непрерывные или дискретные*,

– *детерминированные или стохастические* (точные, однозначные или вероятностные),

– *четкие или нечеткие* (примеры нечетких множеств: около 10; глубоко или мелко; хорошо или плохо).

2.3. Алгоритм научных исследований с помощью математического моделирования

Математическое моделирование – мощное современное средство научных исследований и его применение требует соблюдения определенной строгости во избежание получения неверных выводов.

Процесс моделирования состоит из следующих этапов:

1) изучение оригинала: выявление основных факторов, особенностей, диапазонов исследуемых параметров, условий и задач исследования, постановка (формулировка) задачи исследования, оценка требуемой точности (содержательная постановка);

2) феноменологическое описание оригинала («физическое» описание): поиск аналогий и функциональных зависимостей на основе предыдущего этапа и достижений в различных областях науки (концептуальная постановка);

3) математическое описание оригинала (математическая постановка);

4) разработка алгоритмического и программного обеспечения для реализации математического описания с помощью ЭВМ;

5) проведение контрольного вычислительного эксперимента (воспроизводящего реальный известный случай поведения оригинала в конкретных условиях);

6) оценка адекватности результатов контрольного вычислительного эксперимента реальному случаю; при необходимости – повторение алгоритма с пункта 3, 2 или 1;

7) планирование вычислительного эксперимента в целях исследования;

8) проведение вычислительного эксперимента в целях исследования, обработка его результатов;

9) анализ результатов вычислительного эксперимента, сравнение с результатами изучения оригинала (при необходимости – повторение алгоритма с пункта 7 или 1);

10) формулировка выводов исследования.

2.4. Понятие вычислительного эксперимента

Широкое применение ЭВМ в математическом моделировании, достаточно мощная теоретическая и экспериментальная база позволяют говорить о вычислительном эксперименте как о новой технологии и методологии в научных и прикладных исследованиях.

Вычислительный эксперимент – это эксперимент над математической моделью объекта на ЭВМ, который состоит в том, что по одним параметрам модели вычисляются другие её параметры и на этой основе делаются выводы о свойствах явления, описываемого математической моделью.

При проведении вычислительного эксперимента можно убедиться в необходимости и полезности последнего, особенно в случаях, когда провести натуральный эксперимент затруднительно или невозможно. Вычислительный эксперимент, по сравнению с натурным, значительно дешевле и доступнее, его подготовка и проведение требует меньшего времени, его легко переделывать, он даёт более подробную информацию. Кроме того, в ходе вычислительного эксперимента выявляются границы применимости математической модели, которые позволяют прогнозировать эксперимент в естественных условиях.

2.5. Проверка адекватности модели

Под **адекватностью математической модели** будет пониматься степень соответствия результатов, полученных по разработанной модели, данным эксперимента или тестовой задачи. Прежде чем переходить к проверке адекватности модели, необходимо убедиться в правильном комплексном функционировании всех алгоритмов и программ модели, выполнить независимое тестирование и отладку всех отдельных алгоритмов (например, используемых программных модулей, реализующих используемый численный метод).

Проверка разработанной математической модели выполняется путем сравнения с имеющимися экспериментальными данными о реальном объекте или с результатами других, созданных ранее и хорошо себя зарекомендовавших моделей. В первом случае говорят о проверке путем сравнения с экспериментом, во втором – о сравнении с результатами решения тестовой задачи.

Решение вопроса о точности моделирования зависит от требований, предъявляемых к модели, и ее назначения. При этом должна учитываться точность получения экспериментальных результатов или особенности постановок тестовых задач. В моделях, предназначенных для выполнения оценочных и прикидочных расчетов, удовлетворительной считается точность 10-15%. В моделях, используемых в управляющих и контролирующих системах, требуемая точность может быть 1-2% и даже более.

При возникновении проблем, связанных с адекватностью модели, ее корректировку требуется начинать с последовательного анализа всех возможных причин, приведших к расхождению результатов моделирования и результатов эксперимента.

2.6. Применение инструментальных средств пакета Matlab для решения прикладных задач

MATLAB – это высокоуровневый язык технических расчетов, интерактивная среда разработки алгоритмов и современный инструмент анализа данных.

Ядро MATLAB позволяет максимально просто работать с матрицами реальных, комплексных и аналитических типов данных и со структурами данных и таблицами поиска.

Язык MATLAB является высокоуровневым языком программирования, включающим основанные на матрицах структуры данных, широкий спектр функций, интегрированную среду разработки, объектно-ориентированные возможности и интерфейсы к программам, написанным на других языках программирования.

Программы, написанные на MATLAB, бывают двух типов – функции и скрипты.

Функции имеют входные и выходные аргументы, а также собственное рабочее пространство для хранения промежуточных результатов вычислений и переменных.

Скрипты же используют общее рабочее пространство. Как скрипты, так и функции не интерпретируются в машинный код и сохраняются в виде текстовых файлов.

Основной особенностью языка MATLAB является его широкие возможности по работе с матрицами, которые создатели языка выразили в лозунге думай векторно (англ. Think vectorized).

MATLAB предоставляет пользователю большое количество (несколько сотен) функций для анализа данных, покрывающие практически все области математики, в частности: матрицы и линейная алгебра, многочлены и интерполяция, математическая статистика и анализ данных, обработка данных, дифференциальные уравнения, разреженные матрицы, целочисленная арифметика.

MATLAB предоставляет удобные средства для разработки алгоритмов, включая высокоуровневые с использованием концепций объектно-ориентированного программирования. В нём имеются все необходимые средства интегрированной среды разработки, включая отладчик и профайлер. Функции для работы с целыми типами данных облегчают создание алгоритмов для микроконтроллеров и других приложений, где это необходимо.

В составе пакета MATLAB имеется большое количество функций для построения графиков, в том числе трёхмерных, визуального анализа данных и создания анимированных роликов. Встроенная среда разработки позволяет создавать графические интерфейсы пользователя с различными элементами управления, такими как кнопки, поля ввода и другими. С помощью компонента MATLAB Compiler эти графические интерфейсы могут быть преобразованы в самостоятельные приложения.

ГЛАВА 3. Модели линейного программирования

3.1. Постановка задачи линейного программирования

Задачей линейного программирования называется задача минимизации или максимизации линейной функции при линейных ограничениях.

Будем рассматривать *общую задачу* линейного программирования в форме

$$f(x) = \sum_{j=1}^{n} c_j x_j \to \max/\min, \tag{3.1}$$

при ограничениях

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \ i = \overline{1,k}, \tag{3.2}$$

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i, \ i = \overline{k+1,m}, \tag{3.3}$$

$$x_j \ge 0, \ \ j = \overline{1,n}. \tag{3.4}$$

Ограничения типа (3.2) называются *ограничениями-равенствами*, ограничения типа (3.3) – *ограничениями-неравенствами*, ограничения типа (3.4) – *прямыми ограничениями*. Если в условии задачи линейного программирования не содержаться ограничения-неравенства, то есть в (3.3) $k = m$, что она называется задачей линейного программирования в *каноническом виде*. Любую задачу линейного программирования можно представить в каноническом виде.

Совокупность чисел $x = (x_1, x_2, \ldots, x_n)$, удовлетворяющих условиям задачи (3.1)-(3.4) называют *допустимым решением* и *множеством (областью) допустимых значений (решений)*, и обозначается

$$X = \left\{ x \in R^n : \sum_{j=1}^{n} a_{ij} x_j = b_i, i = \overline{1,k}, \sum_{j=1}^{n} a_{ij} x_j \le b_i, i = \overline{k+1,m}, \ x_j \ge 0, j = \overline{1,n} \right\}.$$

Вектор, при котором целевая функция $f(x)$ достигает своего минимального или максимального значения на допустимом множестве X, называется *оптимальным*, и обозначается $x^{\max} = \left(x_1^{\max}, x_2^{\max}, \ldots, x_n^{\max}\right)$ или

$x^{\min} = \left(x_1^{\min}, x_2^{\min}, \ldots, x_n^{\min}\right)$, а соответствующее значение целевой функции на оптимальном векторе называется *оптимальным* (*максимальным* $f^{\max} = f\left(x^{\max}\right)$ или *минимальным* $f^{\min} = f\left(x^{\min}\right)$).

3.2. Примеры моделей линейного программирования

Рассмотрим классический пример – задачу о рационе.

Пример 3.1. Для поддержания нормальной жизнедеятельности человеку ежедневно необходимо потреблять белки, жиры, углеводы, минеральные соли. Количество питательных веществ, содержащихся в 1 кг имеющихся продуктов питания, а также их стоимость и нормы суточной потребности питательных веществ изображены в виде матрицы (таблица 3.1).

Требуется составить дневной рацион продуктов питания, содержащий не менее суточной нормы потребности человека в необходимых питательных веществах и обеспечивающий минимальную общую стоимость продуктов.

Таблица 3.1

Стоимость продуктов и нормы суточной потребности питательных веществ

Питательные вещества	Содержание питательных веществ в 1 кг продуктов, г/кг							Норма в сутки, г
	Мясо	Рыба	Молоко	Масло	Сыр	Крупа	Картофель	
Белки, г	180	190	30	70	260	130	21	118
Жиры, г	20	3	40	865	310	30	2	56
Углеводы, г	0	0	50	6	20	650	200	500
Минеральные соли, г	9	10	7	12	60	20	70	8
Цена 1 кг продукта, руб./кг	1,9	1,0	0,28	3,4	2,9	0,5	0,1	

Решение. Составим математическую модель задачи и введем переменные x_j, $j = 1,\ldots,7$, – вес соответствующего продукта в диете. Ограничениям в задачи связаны с суточной потребностью организма в питательных веществах, показатель эффективности – минимизация затрат на покупку продуктов. Получим следующую математическую модель:

найти минимум целевой функции (функция затрат)

$$f(x) = 1{,}9x_1 + 1{,}0x_2 + 0{,}28x_3 + 3{,}4x_4 + 2{,}9x_5 + 0{,}5x_6 + 0{,}1x_7 \to \min$$

при ограничениях

$$\begin{cases} 180x_1 + 190x_2 + 30x_3 + 70x_4 + 260x_5 + 130x_6 + 21x_7 \ge 118, \\ 20x_1 + 3x_2 + 40x_3 + 865x_4 + 310x_5 + 30x_6 + 2x_7 \ge 56, \\ 0x_1 + 0x_2 + 50x_3 + 6x_4 + 20x_5 + 650x_6 + 200x_7 \ge 500, \\ 9x_1 + 10x_2 + 7x_3 + 12x_4 + 60x_5 + 20x_6 + 70x_7 \ge 8, \\ x_j \ge 0,\ j = 1,\ldots,7. \end{cases}$$

Пример 3.2. Инвестор, располагающий суммой в 300 тыс. ден. ед., может вложить свой капитал в акции автомобильного концерна *A* и строительного предприятия *B*. Чтобы уменьшить инвестиционные риски, акций концерна *A* должно быть приобретено не меньше, чем акций строительного предприятия *B*, причем последних можно купить не более чем на 100 тыс. ден. ед. Дивиденды по акциям *A* составляют 8%, а по акциям *B* – 10% в год. Определить, какую максимальную прибыль может получить инвестор в первый год.

Решение. Построим экономико-математическую модель задачи. Введем обозначения: x_1 и x_2 – количество денежных средств (в тыс. ден. ед.), которые будут вложены в автомобильный концерн и в строительное предприятие соответственно. Критерий оптимальности в задачи – максимум прибыли от вложений.

Получаем следующую задачу: найти максимум целевой функции (функция прибыли)

$$f(x) = 0{,}08x_1 + 0{,}1x_2 \to \max$$

при ограничениях

$$\begin{cases} x_1 + x_2 \le 300, \\ x_1 - x_2 \ge 0, \\ x_2 \le 100, \\ x_j \ge 0,\ j = 1, 2. \end{cases}$$

Пример 3.3. Кондитерская фабрика для производит карамель трех видов и использует три вида основного сырья: сахарный песок, патоку и фруктовое пюре. Нормы расхода сырья на каждого вида на производство 1 т карамели данного вида, а также запасы сырья каждого вида и прибыль от реализации 1 т карамели каждого вида приведены в табл. 3.2.

Требуется составить план производства карамели, обеспечивающий максимум прибыли от ее реализации.

Таблица 3.2

Нормы расхода сырья, запасы сырья и прибыль от реализации

Вид сырья	Нормы расхода сырья на 1 т карамели, т			Запасы сырья, т
	A	B	C	
Сахарный песок	0,8	0,5	0,6	800
Патока	0,2	0,4	0,3	600
Фруктовое пюре	0	0,1	0,1	120
Прибыль от реализации 1 т продукции, тыс. руб.	108	112	126	

Решение. Составим экономико-математическую модель задачи. Обозначим через $x_j, j = 1, 2, 3$, – количество карамели (в т) каждого вида. Цель задачи – максимизации прибыли от реализации при ограничениях на запасы сырья. Тогда получаем следующую задачу: найти максимум целевой функции (функция прибыли)

$$f(x) = 108x_1 + 112x_2 + 126x_3 \to \max$$

при ограничениях

$$\begin{cases} 0{,}8x_1 + 0{,}5x_2 + 0{,}6x_3 \le 800, \\ 0{,}2x_1 + 0{,}4x_2 + 0{,}3x_3 \le 600, \\ 0x_1 + 0{,}1x_2 + 0{,}1x_3 \le 120, \\ x_j \ge 0,\ j = 1, 2, 3. \end{cases}$$

3.3. Графический метод решения задачи линейного программирования

Пусть задача линейного программирования содержит только две переменные, и в ее условии нет ограничений равенств, то есть, имеем задачу вида

$$f(x) = c_1x_1 + c_2x_2 \to \max/\min, \tag{3.5}$$

$$a_{i1}x_1 + a_{i2}x_2 \le b_i,\ i = \overline{1,m}, \tag{3.6}$$

$$x_1, x_2 \ge 0. \tag{3.7}$$

Допустимое множество X в задаче (3.5)-(3.7) является пересечением первого квадранта и полуплоскостей, соответствующих неравенствам (3.6). Чтобы построить каждую из полуплоскостей, соответствующую неравенствам (3.6), необходимо построить граничную прямую $a_{i1}x_1 + a_{i2}x_2 = b_i$, выбрать на плоскости любую точку, не лежащую на данной прямой и подставить в соответствующее неравенство $a_{i1}x_1 + a_{i2}x_2 \le b_i$. Если неравенство выполняется, то решением будет все множество точек координатной плоскости, лежащее по ту же сторону, что и выбранная точка, в неравенство не выполняется, то множество точек на противоположной стороне.

При этом множество допустимых значений X может представлять собой ограниченный многоугольник в координатной плоскости (в этом случае задача разрешима, и решений может быть одно или бесконечное количество), пустое множество (иначе – система ограничений несовместна, и задача не имеет решение), неограниченный многоугольник (решение может существовать или не существовать, если функция неограниченна сверху или снизу на допустимом множестве).

Для решения задачи (3.5)-(3.7) рассмотрим семейство линий уровня целевой функции из (3.5)

$$c_1x_1 + c_2x_2 = C,\ C = const, \tag{3.8}$$

которые являются параллельными прямыми. Градиент $f'(x) = (c_1, c_2)$ и антиградиент $-f'(x) = (-c_1, -c_2)$ перпендикулярны прямым (3.8) и указывают направление возрастания и убывания целевой функции. Если перемещать параллельно самой себе произвольную прямую (3.8), проходящую через допустимое множество X, в направлении градиента или антиградиента до тех пор, пока эта прямая будет иметь хотя бы одну общую точку с множеством X, то в своем крайнем положении указанная прямая пройдет через точку множества X, в которой целевая функция $f(x)$ принимает максимальное или минимальное на X значение.

Пример 3.4. Вернемся к примеру 3.2 и запишем математическую модель задачи:

$$f(x) = 0{,}08x_1 + 0{,}1x_2 \to \max$$

$$\begin{aligned} &x_1 + x_2 \le 300, \\ &x_1 - x_2 \ge 0, \\ &x_2 \le 100, \\ &x_{1,2} \ge 0. \end{aligned}$$

Найдем решение данной задачи графическим методом.

Решение. Изобразим на координатной плоскости допустимое множество X. Проведем граничные прямые $x_1 + x_2 = 300$ (l_1), $x_1 - x_2 = 0$ (l_2), $x_2 = 100$ (l_3) и определим полуплоскости, соответствующие ограничениям-неравенствам. Для этого для каждого ограничения выберем точку (например, начало координат (0, 0)), не лежащую на соответствующей граничной прямой, и проверим выполнение неравенства. Например, для ограничения $x_1 + x_2 \le 300$ подставим координаты (0, 0) и получим, что неравенство выполняется, т.е. это ограничение описывает множество точек, лежащих ниже (левее) относительно прямой $x_1 + x_2 = 300$. В результате получаем следующее множество допустимых решений, изображенное на рисунке 3.1.

Построим линию уровня целевой функции $0{,}08x_1 + 0{,}1x_2 = 0$ и вычислим градиент $grad\, f = (0{,}08,\ 0{,}1)$. Для удобства построений, можно умножить целевую функцию на 50,чтобы получить целые значения коэффициентов при переменных. Следует также отметить, что необязательно строить градиент согласно вычисленным значениям, достаточно лишь определить его направление.

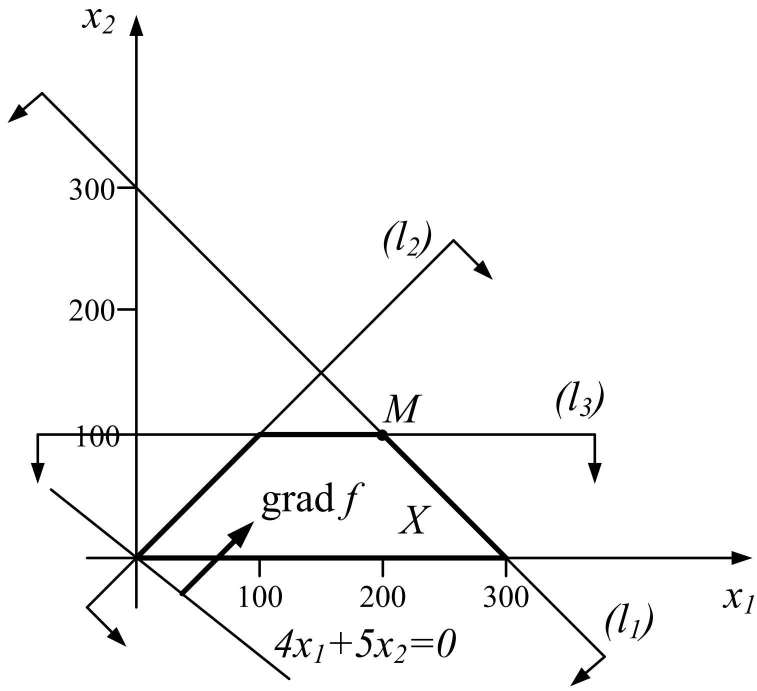

Рисунок 3.1 – Графическая иллюстрация к примеру 3.4

Совершая параллельный перенос линии уровня в направлении вектора *grad f*, находим ее крайнее положение. В этом положении прямая проходит через точку M – точку пересечения граничным прямых (l_1) и (l_3). Таким образом, целевая функция достигает максимум в точке $x^{\max} = (200, 100)$, и максимальное значение равное $f^{\max} = 26$.

Это означает, что инвестору следует вложить 200 тыс. ден. ед. в автомобильный концерн и 100 тыс. ден. ед. в строительное предприятие, при этом максимальная прибыль в первый год составит 26 тыс. ден. ед.

Пример 3.5. Решить следующую задачу линейного программирования графически:

$$
\begin{aligned}
&f(x) = -3x_1 - 2x_2 \to \min,\\
&x_1 + 2x_2 \le 7,\\
&2x_1 + x_2 \le 8,\\
&\qquad x_2 \le 3,\\
&x_1, x_2 \ge 0.
\end{aligned}
$$

Решение. Изобразим на координатной плоскости допустимое множество X. Проведем граничные прямые $x_1 + 2x_2 = 7$ (l_1), $2x_1 + x_2 = 8$ (l_2), $x_2 = 3$ (l_3) и определим полуплоскости, соответствующие ограничениям-неравенствам. Для этого для каждого ограничения выберем точку (например, начало координат (0, 0)), не лежащую на соответствующей граничной прямой, и проверим выполнение неравенства. Например, для ограничения $x_1 + 2x_2 \le 7$ подставим координаты (0, 0) и получим, что неравенство выполняется, т.е. это ограничение описывает множество точек, лежащих ниже (левее) относительно прямой $x_1 + 2x_2 = 7$. В результате получаем следующее множество допустимых решений (рисунок 3.2).

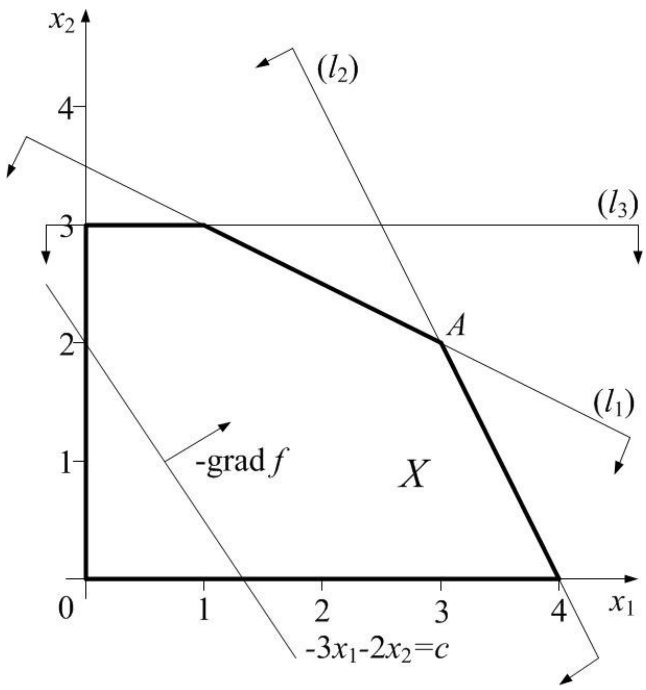

Рисунок 3.2 – Графическая иллюстрация к примеру 3.5

Построим линию уровня целевой функции $-3x_1 - 2x_2 = c$ и вычислим антиградиент $-grad\ f = (3, 2)$. Совершая параллельный перенос линии уровня в направлении вектора $-grad\ f$, находим ее крайнее положение. В этом положении прямая проходит через точку $A(3, 2)$. Таким образом, целевая функция принимает минимальное значение $f^{min} = -13$ в точке $x^{min} = (3, 2)$.

Графический метод используют также и для решения задачи линейного программирования в каноническом виде с произвольным числом переменных, если число свободных переменных системы ограничений-равенств не превосходит двух. В этом случае, исходя из каких-либо соображений, одна или две переменные выбираются в качестве свободных, остальные переменные (базисные) и целевая функция выражаются через эти свободные переменные. Пользуясь свойством неотрицательности переменных из выражений для базисных переменных, получим систему ограничений неравенств, определяющих допустимое множество. А далее алгоритм решения повторяет метод, описанный выше.

Пример 3.6. Решить графически следующую задачу линейного программирования:

$$f(x) = 3x_1 + 4x_2 + 7x_3 - x_4 \to extr,$$
$$x_1 + 3x_2 + 2x_3 + 2x_4 = 7,$$
$$2x_1 + 4x_2 + x_3 + 3x_4 = 10,$$
$$x_j \geq 0,\ j = \overline{1,4}.$$

Решение. Запишем расширенную матрицу системы ограничений и приведем ее к треугольному виду

$$A = \begin{pmatrix} 1 & 3 & 2 & 2 & 7 \\ 2 & 4 & 1 & 3 & 10 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & 2 & 2 & 7 \\ 0 & -2 & -3 & -1 & -4 \end{pmatrix}.$$

В качестве базисных переменных выберем x_1, x_4 и выразим их через свободные переменные x_2, x_3. Имеем

$$x_1 = x_2 + 4x_3 - 1,\ x_4 = -2x_2 - 3x_3 + 4.$$

Из условия неотрицательности переменных получаем систему ограничений неравенств

$$\begin{cases} x_1 = x_2 + 4x_3 - 1 \geq 0, \\ x_4 = -2x_2 - 3x_3 + 4 \geq 0, \\ x_{2,3} \geq 0 \end{cases} \qquad \begin{cases} x_2 + 4x_3 \geq 1, & (l_1) \\ 2x_2 + 3x_3 \leq 4, & (l_2) \\ x_{2,3} \geq 0. \end{cases}$$

Выразим целевую функцию через свободные переменные и рассчитаем градиент и антиградиент:

$f(x) = 3x_1 + 4x_2 + 7x_3 - x_4 = 3(x_2 + 4x_3 - 1) + 4x_2 + 7x_3 - (-2x_2 - 3x_3 + 4)$,

$f(x) = 9x_2 + 22x_3 - 7$, $gradf = (9,22)$, $-gradf = (-9,-22)$.

Построим область допустимых значений (рисунок 3.3) и определим решение задачи.

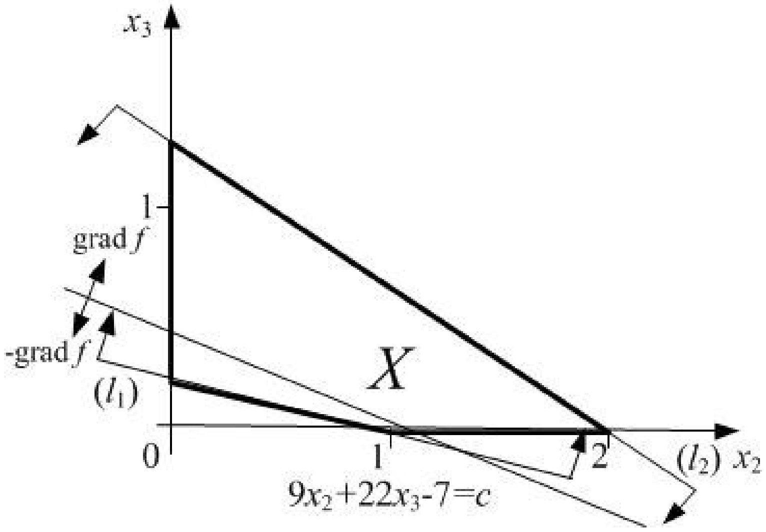

Рисунок 3.3 – Графическая иллюстрация к примеру 3.6

Передвигая линию уровня $9x_2 + 22x_3 - 7 = c$ в направлении градиента, получаем, что максимум достигается в точке $(x_2, x_3) = (0, \frac{4}{3})$, откуда вычисляем $x_1 = x_2 + 4x_3 - 1 = \frac{13}{3}$, $x_4 = -2x_2 - 3x_3 + 4 = 0$, $f(x) = 9x_2 + 22x_3 - 7 = \frac{67}{3}$. Минимум достигается в точке $(x_2, x_3) = (0, \frac{1}{4})$, откуда вычисляем $x_1 = x_2 + 4x_3 - 1 = 0$, $x_4 = -2x_2 - 3x_3 + 4 = \frac{13}{4}$, $f(x) = 9x_2 + 22x_3 - 7 = -\frac{3}{2}$.

Окончательно получаем точки экстремума $x^{\max} = (\frac{13}{3}, 0, \frac{4}{3}, 0)$, $f^{\max} = \frac{67}{3}$; $x^{\min} = (0, 0, \frac{1}{4}, \frac{13}{4})$, $f^{\min} = -\frac{3}{2}$.

ГЛАВА 4. Моделирование транспортных задач

В современных условиях большие транспортные расходы связаны с простоями в ожидании обслуживания на погрузочно-разгрузочных работах, порожними пробегами, встречными и нерациональными перевозками, затратами на бензин, техническое обслуживание и заработную плату водителей. В связи с этим необходимо решать задачи оптимального планирования перевозок грузов в коммерческой деятельности из пунктов отправления (баз, станций, фабрик, совхозов, заводов) в пункты назначения (магазины, склады) методами, позволяющими оптимизировать план по какому-либо экономическому показателю, например, финансовых затрат или времени на перевозку грузов.

Для решения подобного рода задач в линейном программировании существуют специально разработанные методы, а задачи такого рода называются транспортными задачами.

4.1. Математическая модель транспортной задачи

Пусть имеется *n* пунктов отправления (поставщиков) грузов $A_1, A_2, ..., A_m$, на которых сосредоточены запасы какого-либо однородного груза в объемах соответственно $a_1, a_2, ..., a_m$. Величины a_i определяют максимально возможные размеры вывоза груза с пунктов отправления. Суммарный запас груза поставщиков составляет $\sum_{i=1}^{m} a_i$.

Кроме того, имеется *n* пунктов назначения $B_1, B_2, ..., B_n$, которые подали заявки на поставку грузов в объемах соответственно $b_1, b_2, ..., b_n$. Суммарная величина заявок составляет $\sum_{j=1}^{n} b_j$.

Стоимость перевозки одной единицы груза от поставщика A_i к потребителю B_j обозначим через c_{ij} (транспортный тариф), образующих матрицу транспортных издержек $C = \{c_{ij}\}$. В качестве критерия оптимальности выбираем суммарные издержки по перевозке грузов.

Тогда транспортная задача формулируется следующим образом: необходимо составить оптимальный план, т.е. найти такие значения объема перевозок грузов $X = \{x_{ij}\}$ от поставщиков A_i к потребителям B_j, чтобы вывести все грузы от поставщиков; удовлетворить заявки каждого потребителя и обеспечить минимальные транспортные расходы на перевозку груза.

Все исходные данные транспортной задачи можно записать в виде табл. 4.1, которая называется *транспортной*. Матрица $X = \{x_{ij}\}$, определяющая решение задачи, называется *матрицей перевозок*. Очевидно, что элементы x_{ij}, определяющие объем перевозимого от поставщиков A_i к потребителям B_j груза, должны быть неотрицательны. Столбец u_i и строка v_j в таблице 4.1 являются вспомогательными при решении задачи. О правилах их заполнения будет сказано ниже.

Таблица 4.1

$a_i \setminus b_j$	b_1	b_2	…	b_j	…	b_n	u_i
a_1	c_{11} x_{11}	c_{12} x_{12}	…	c_{1j} x_{1j}	…	c_{1n} x_{1n}	
a_2	c_{21} x_{21}	c_{22} x_{22}	…	c_{2j} x_{2j}	…	c_{2n} x_{2n}	
…	…	…	…	…	…	…	
a_i	c_{i1} x_{i1}	c_{i2} x_{i2}		c_{ij} x_{ij}		c_{in} x_{in}	
…	…	…	…	…	…	…	
a_m	c_{m1} x_{m1}	c_{m2} x_{m2}	…	c_{mj} x_{mj}	…	c_{mn} x_{mn}	
v_j							

Задача заключается в определении плана перевозок – матрицы $X = \{x_{ij}\}$ ($i = \overline{1,m},\ \ j = \overline{1,n}$), которая удовлетворяет следующим условиям: обеспечивает минимальное значение целевой функции

$$f(X) = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \to \min \tag{4.1}$$

и удовлетворяет следующим условиям

$$\sum_{j=1}^{n} x_{ij} = a_i, \ i = \overline{1,m}, \tag{4.2}$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \ j = \overline{1,n}, \tag{4.3}$$

$$x_{ij} \ge 0, \ i = \overline{1,m}, \ j = \overline{1,n}. \tag{4.4}$$

В таком виде экономико-математическая постановка транспортной задачи считается законченной.

Целевая функция задачи $f(X)$ выражает требование обеспечить минимум суммарных затрат на перевозку всех грузов. Первая группа из уравнений ограничений, записанных в общем виде (4.2), выражает требование, что запасы всех m поставщиков вывозятся полностью, а вторая группа из уравнений ограничений, записанных в общем виде (4.3) означает, полностью должны удовлетворятся запросы всех n потребителей. Последнее неравенство (4.4) является условием неотрицательности всех переменных.

В рассмотренной математической модели транспортной задачи предполагается, что суммарные запасы поставщиков равны суммарным запросам потребителей, т.е.

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j . \tag{4.5}$$

Такая задача называется *сбалансированной*, а её модель *закрытой*. Если же это равенство не выполняется, то задача называется *несбалансированной* (с неправильным балансом), а её модель – *открытой*.

Для того чтобы транспортная задача линейного программирования имела решение, необходимо, чтобы суммарные запасы поставщиков равнялись суммарным запросам потребителей, т.е. задача должна быть сбалансированной.

В случае превышения запаса над потребностями, т.е. если $\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$, вводится фиктивный ($n$+1)-й пункт назначения с потребностью $b_{n+1} = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j$ и равными нулю тарифами перевозок $c_{i(n+1)} = 0 \ (i = \overline{1,m})$.

Аналогично при превышении потребностей над запасами, т.е. при $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$, вводится фиктивный ($m$+1)-й поставщик с запасами $a_{m+1} = \sum_{j=1}^{n} b_j - \sum_{i=1}^{m} a_i$ и нулевыми тарифами перевозок $c_{(m+1)j} = 0 \ (j = \overline{1,n})$.

Число переменных x_{ij} в транспортной задачу с m пунктами отправления и n пунктами назначения равно mn, а число уравнений в системе (4.2), (4.3) равно (m+n).

Решение транспортной задачи разбивается на следующие этапы:

1) определение начального (опорного) решения;

2.1) проверка данного решения на оптимальность;

2.2) в случае не оптимальности решения его улучшение.

Для определения начального (опорного) решения в транспортной задаче существует несколько методов (метод северо-западного угла, метод минимального элемента и метод аппроксимации Фогеля).

Для проверки плана на оптимальность используется методы потенциалов.

4.2. Построение опорного плана транспортной задачи

4.2.1. Метод северо-западного угла

В данном методе запасы очередного поставщика используются для обеспечения запросов очередных потребителей до тех пор, пока не будут исчерпаны полностью, после чего используются запасы следующего по номеру поставщика.

Заполнение таблицы транспортной задачи начинается с левого верхнего угла и состоит из ряда однотипных шагов. На каждом шаге, исходя из запасов очередного поставщика и запросов очередного потребителя, заполняется только одна клетка и соответственно исключается из рассмотрения один поставщик или потребитель. Назначение перевозок и перемещение по транспортной таблице происходит из левого верхнего угла в правый нижний. Осуществляется это таким образом.

Изначально полагаем, что все перевозки равны нулю. На первом шаге:

1) если $a_1 < b_1$, то назначается объем перевозок $x_{11} = a_1$ и из рассмотрения исключается первый перевозчик, как исчерпавший свой ресурс, при этом остальные перевозки от первого поставщика равны нулю, т.е. $x_{1j} = 0,\ j = \overline{2,n}$ (причем эти нули в транспортную таблицу ставить не следует), для назначения следующей перевозки следует переместиться на одну клетку вниз;

2) если $a_1 > b_1$, то назначается объем перевозок $x_{11} = b_1$ и из рассмотрения исключается первый потребитель, как удовлетворивший свои потребности, при этом остальные перевозки к первому потребителю равны нулю, т.е. $x_{i1} = 0,\ i = \overline{2,m}$ (эти нули в транспортную таблицу ставить также не следует), для назначения следующей перевозки следует переместиться на одну клетку вправо;

3) если $a_1 = b_1$, то назначается объем перевозок $x_{11} = a_1 = b_1$ и из рассмотрения исключается либо первый поставщик (см. п. 1), либо первый потребитель (см. п. 2).

Далее перевозка x_{ij} в произвольную клетку $(i, j),\ i = \overline{1,m},\ j = \overline{1,n}$, ставится исходя из следующих соображений:

1) если $a_i - \sum_{l=1}^{j-1} x_{il} < b_j - \sum_{k=1}^{i-1} x_{kj}$, то назначается объем перевозок $x_{ij} = a_i - \sum_{l=1}^{j-1} x_{il}$ и из рассмотрения исключается i-й перевозчик, как исчерпавший свой ресурс, при этом следующие перевозки от этого поставщика равны нулю, т.е. $x_{il} = 0,\ l = \overline{j+1,n}$ (эти нули в транспортную таблицу ставить не следует),

для назначения следующей перевозки следует переместиться на одну клетку вниз;

2) если $a_i - \sum_{l=1}^{j-1} x_{il} > b_j - \sum_{k=1}^{i-1} x_{kj}$, то назначается объем перевозок $x_{ij} = b_j - \sum_{k=1}^{i-1} x_{kj}$ и из рассмотрения исключается j-й потребитель, как удовлетворивший свои потребности, при этом следующие перевозки к этому потребителю равны нулю, т.е. $x_{kj} = 0,\ k = \overline{i+1, m}$ (эти нули в транспортную таблицу ставить не следует), для назначения следующей перевозки следует переместиться на одну клетку вправо;

3) если $a_i - \sum_{l=1}^{j-1} x_{il} = b_j - \sum_{k=1}^{i-1} x_{kj}$, то назначается объем перевозок $x_{ij} = a_i - \sum_{l=1}^{j-1} x_{il} = b_j - \sum_{k=1}^{i-1} x_{kj}$ и из рассмотрения исключается либо первый поставщик (см. п. 1), либо первый потребитель (см. п. 2).

В последней клетке должно выполниться равенство $a_m - \sum_{l=1}^{n-1} x_{ml} = b_n - \sum_{k=1}^{m-1} x_{kn}$.

Нулевые перевозки принято заносить в таблицу только тогда, когда они попадают в клетку (i, j), подлежащую заполнению. Если в очередную клетку таблицы (i, j) требуется поставить перевозку, а i-й поставщик или j-й потребитель имеет нулевые запасы или запросы, то в клетку ставится перевозка, равная нулю (базисный нуль), и после этого, как обычно, исключается из рассмотрения соответствующий поставщик или потребитель. Таким образом, в таблицу заносят только базисные нули, остальные клетки с нулевыми перевозками остаются пустыми.

Во избежание ошибок после построения начального опорного решения необходимо проверить, что число занятых клеток равно (m+n–1) и условия (4.2), (4.3) выполнены.

4.2.2. Метод минимального элемента (наименьшей стоимости)

Метод минимальной стоимости прост, он позволяет построить опорное решение, достаточно близкое к оптимальному, так как использует матрицу стоимостей транспортной задачи $C = \{c_{ij}\}$. Как и метод северо-западного угла, он состоит из ряда однотипных шагов, на каждом из которых заполняется только одна клетка таблицы, соответствующая минимальной стоимости $\min\{c_{ij}\}$, и исключается из рассмотрения только одна строка (поставщик) или один столбец (потребитель).

Очередную клетку, соответствующую $\min\{c_{ij}\}$, заполняют по тем же правилам, что и в методе северо-западного угла. Поставщик исключается из рассмотрения, если его запасы использованы полностью. Потребитель исключается из рассмотрения, если его запросы удовлетворены полностью. На каждом шаге исключается либо один поставщик, либо один потребитель. При этом если поставщик, еще не исключен, но его запасы равны нулю, то на том шаге, когда от данного поставщика требуется поставить груз, в соответствующую клетку таблицы заносится базисный нуль, и лишь затем поставщик исключается из рассмотрения. Аналогично с потребителем.

Если клеток с минимальной стоимостью несколько, то произвольно выбирается любая из них.

4.3. Метод потенциалов

Для каждого опорного решения X следует вычислить значения переменных u_i и v_j следующим образом:

1) для всех базисных клеток создать систему уравнений вида $u_i + v_j = c_{ij}$;

2) задать произвольно какое-либо значение двойственной переменной, например, $u_1 = 0$ или $u_1 = \max\{c_{ij}\}$;

3) рассчитать остальные значения двойственных переменных из составленной в п. 1 системе уравнений.

Полученные значения записываются в столбец u_i и строку v_j в табл. 4.1.

Переменные u_i и v_j называются *потенциалами*.

Условия оптимальности имеют вид: для свободных клеток должны выполнятся неравенства $u_i^* + v_j^* - c_{ij} \le 0$ (или величина $\Delta_{ij} = c_{ij} - u_i^* - v_j^* \ge 0$). Если хотя бы для одной из свободных клеток это условие не выполняется, то построенное опорное решение не является оптимальным.

Для улучшения опорного решения следует пользоваться следующим алгоритмом. Начиная с клетки с наименьшим отрицательным Δ_{ij}, строим цикл (или многоугольник) перераспределения. Цикл перераспределения – это замкнутая ломанная линия, одна вершина которой расположена в клетке с наименьшим отрицательным Δ_{ij}, остальные – в базисных клетках, а звенья параллельны строкам и столбцам таблицы. При правильном построении опорного решения для любой свободной клетки можно построить только один цикл, причем число вершин цикла четно.

Далее каждой вершине цикла приписывается определенный знак: в свободной клетке ставится знак «+», в следующей – знак «–», далее знаки чередуются. Среди элементов x_{ij}, находящихся в базисных клетках цикла со знаком «–», находим наименьший; эту величину назовем шагом перераспределения и обозначим *h*. Затем к элементам x_{ij}, расположенных в клетках со знаком «+», прибавляем *h* (в свободной клетке цикла проставляем число *h*), от элементов x_{ij}, расположенных в клетках со знаком «–», вычитаем *h*.

В результате получаем новое опорное решение, которое также следует проверить на оптимальность, как было указано выше.

Следует отметить, что после пересчета новое опорное решение может оказаться вырожденным. Это возможно, если при в базисных клетках со знаком «–» имеется два или более одинаковых минимальных числа x_{ij}. В этом случае следует освобождать только одну клетку, в остальных клетках оставить нули и считать их базисными.

Кроме того, если в циклах для всех свободных клеток с отрицательными Δ_{ij} получаем шаг перераспределения h, равный нулю, то решение улучшить нельзя, т.е. построенный план является оптимальным.

Рассмотрим следующий пример.

Пример 4.1. На трех базы A_1, A_2, A_3 имеется однородный груз в количествах, соответственно равных 60, 80, 100 ед. Этот груз требуется перевезти в четыре магазина B_1, B_2, B_3, B_4 соответственно в количествах 40, 60, 90, 70 ед. Стоимость доставки единицы груза из каждого пункта отправления в соответствующие пункты назначения задана матрицей тарифов (тыс. руб. за ед. груза):

$$A = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 6 & 3 & 8 & 5 \\ 4 & 7 & 6 & 2 \end{pmatrix}.$$

Требуется составить план перевозок однородного груза с минимальными транспортными издержками.

Решение. Составим математическую модель задачи. Проверим необходимое и достаточное условие разрешимости:

$$\sum_{i=1}^{3} a_i = 60 + 80 + 100 = 240,$$

$$\sum_{j=1}^{4} b_j = 40 + 60 + 90 + 70 = 260.$$

Как видно, суммарная потребность груза в пунктах назначения превышает запасы груза на трех базах. Следовательно, модель исходной транспортной задачи является открытой. Чтобы получить закрытую модель, введем дополнительную (фиктивную) базу A_4 с запасом груза, равным $a_4 = \sum_{j=1}^{4} b_j - \sum_{i=1}^{3} a_i = 260 - 240 = 20$ ед. Тарифы перевозки единицы груза из базы A_4 во все магазины полагаем равны нулю, т.е. $c_{4j} = 0,\ j = \overline{1,4}$.

Решение будем искать в виде матрицы перевозок

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix},$$

где x_{ij} – количество ед. груза, перевозимого с *i*-го склада в *j*-й магазин.

Математическая модель задачи имеет вид: найти минимум суммарной стоимости перевозок

$$f(X) = 1x_{11} + 2x_{12} + 4x_{13} + 3x_{14} + 6x_{21} + 3x_{22} + 8x_{23} + 5x_{24} +$$
$$+4x_{31} + 7x_{32} + 6x_{33} + 2x_{34} + 0x_{41} + 0x_{42} + 0x_{43} + 0x_{44} \to \min$$

при условии вывоза всего груза со складов

$$x_{11} + x_{12} + x_{13} + x_{14} = 60,$$
$$x_{21} + x_{22} + x_{23} + x_{24} = 80,$$
$$x_{31} + x_{32} + x_{33} + x_{34} = 100,$$
$$x_{41} + x_{42} + x_{43} + x_{44} = 20,$$

при удовлетворении потребностей всех магазинов

$$x_{11} + x_{21} + x_{31} + x_{41} = 40,$$
$$x_{12} + x_{22} + x_{32} + x_{42} = 60,$$
$$x_{13} + x_{23} + x_{33} + x_{43} = 90,$$
$$x_{14} + x_{24} + x_{34} + x_{44} = 70,$$

и при условиях неотрицательности переменных

$$x_{ij} \ge 0,\ i = \overline{1,4},\ j = \overline{1,4}.$$

Найдем решение транспортной задачи методом потенциалов, находя первый опорный план двумя описанными алгоритмами.

1. Метод северо-западного угла

Проставим перевозку в клетку (1, 1): $x_{11} = \min\{a_1, b_1\} = \min\{60, 40\} = 40$. Поскольку потребности первого потребителя удовлетворены, то его исключаем из рассмотрения и передвигаемся на одну клетку вправо.

Ставим перевозку в клетку (1, 2): $x_{12} = \min\{a_1 - x_{11}, b_2\} = \min\{20, 60\} = 20$. Поскольку весь груз из первого склада вывезен, то его исключаем из рассмотрения и перемещаемся на одну клетку вниз.

Ставим перевозку в клетку (2, 2): $x_{22} = \min\{a_2, b_2 - x_{12}\} = \min\{80, 40\} = 40$. Поскольку все потребности второго магазина удовлетворены, то его исключаем из рассмотрения и перемещаемся на одну клетку вправо.

Ставим перевозку в клетку (2, 3): $x_{23} = \min\{a_2 - x_{22}, b_3\} = \min\{40, 90\} = 40$. Поскольку весь груз из второго склада вывезен, то его исключаем из рассмотрения и перемещаемся на одну клетку вниз.

Ставим перевозку в клетку (3, 3): $x_{33} = \min\{a_3, b_3 - x_{23}\} = = \min\{100, 50\} = 50$. Поскольку все потребности третьего магазина удовлетворены, то его исключаем из рассмотрения и перемещаемся на одну клетку вправо.

Ставим перевозку в клетку (3, 4): $x_{34} = \min\{a_3 - x_{33}, b_4\} = \min\{50, 70\} = 50$. Поскольку весь груз из третьего склада вывезен, то его исключаем из рассмотрения и перемещаемся на одну клетку вниз.

Ставим перевозку в клетку (4, 4): $x_{44} = \min\{a_4, b_4 - x_{34}\} = \min\{20, 20\} = 20$. Все потребности четвертого магазина удовлетворены и весь груз из четвертого склада вывезен.

Таким образом, построен первый опорный план (табл. 4.2).

Рассчитаем стоимость такого плана перевозок:

$F = 1 \cdot 40 + 2 \cdot 20 + 3 \cdot 40 + 8 \cdot 40 + 6 \cdot 50 + 2 \cdot 50 + 0 \cdot 20 = 920$.

Количество базисных клеток равно 7 и равно $m+n-1$=4+4-1=7, следовательно, план не вырожден.

Проверим план на оптимальность. Для базисных клеток построим следующую систему для расчета потенциалов:

$$\begin{cases} u_1 + v_1 = 1, \quad u_1 + v_2 = 2, \\ u_2 + v_2 = 3, \; u_2 + v_3 = 8, \\ u_3 + v_3 = 6, \; u_3 + v_4 = 2, \\ u_4 + v_4 = 0. \end{cases}$$

Зададим значение $u_1 = 0$ и из построенной системы рассчитаем:

$$\begin{cases} v_1 = 1 - u_1 = 1, \\ v_2 = 2 - u_1 = 2, \\ u_2 = 3 - v_2 = 1, \\ v_3 = 8 - u_2 = 7, \\ u_3 = 6 - v_3 = -1, \\ v_4 = 2 - u_3 = 3, \\ u_4 = 0 - v_4 = -3. \end{cases}$$

Все вычисления удобно проводить в транспортной таблице (таблица 4.2).

Таблица 4.2

$a_i \setminus b_j$	40	60	90	70	u_i
60	1 **40**	2 **20**	4	3	0
80	6	3 **40**	8 **40**	5	1
100	4	7	6 – **50**	2 + **50**	–1
20	0	0	0 +	0 – **20**	–3
v_j	1	2	7	3	

Для свободных клеток вычислим разности:

$\Delta_{13} = 4 - 0 - 7 = -3 < 0$, $\Delta_{14} = 3 - 0 - 3 = 0$,

$\Delta_{21} = 6 - 1 - 1 = 4 > 0$, $\Delta_{24} = 5 - 1 - 3 = 1 > 0$,

$\Delta_{31} = 4 + 1 - 1 = 4 > 0$, $\Delta_{32} = 7 + 1 - 2 = 6 > 0$,

$\Delta_{41} = 0 + 3 - 1 = 2 > 0$, $\Delta_{42} = 0 + 3 - 2 = 1 > 0$,

$\Delta_{43} = 0 + 3 - 7 = -4 < 0$.

Поскольку среди Δ_{ij} есть отрицательные, то план не оптимален. Для клетки (4, 3) с наименьшей отрицательной разностью строим цикл перераспределения: это четырехугольник с вершинами в клетках (4, 3), (4, 4), (3, 4), (3, 3). Расставим, чередуя, знаки «+» и «–», начиная с клетки (4, 3). Далее вычислим шаг пересчета $h=\min\{50,20\}=20$ и строим новый план перевозок (таблица 4.3).

Таблица 4.3

$a_i \setminus b_j$	40	60	90	70	u_i
60	1 **40**	2 − **20**	4 +	3	0
80	6	3 + **40**	8 − **40**	5	1
100	4	7	6 **30**	2 **70**	–1
20	0	0	0 **20**	0	–7
v_j	1	2	7	3	

Рассчитаем стоимость нового плана перевозок:

$F=1\cdot 40+2\cdot 20+3\cdot 40+8\cdot 40+6\cdot 30+2\cdot 70+0\cdot 20=840.$

Видно, что на новом плане значение стоимости перевозок уменьшилось.

Число базисных клеток равно 7, следовательно, план не вырожден. Аналогично рассчитаем потенциалы и проверим план на оптимальность:

$\Delta_{13}=4-0-7=-3<0$, $\Delta_{14}=3-0-3=0$,

$\Delta_{21}=6-1-1=4>0$, $\Delta_{24}=5-1-3=1>0$,

$\Delta_{31}=4+1-1=4>0$, $\Delta_{32}=7+1-2=6>0$,

$\Delta_{41}=0+7-1=6>0$, $\Delta_{42}=0+7-2=5>0$,

$\Delta_{43}=0+7-3=4>0$.

Поскольку среди $\Delta_{13}<0$, то план не оптимален. Для клетки (1, 3) строим цикл перераспределения: это четырехугольник с вершинами в клетках (1, 3), (2, 3), (2, 2), (1, 2). Расставим, чередуя, знаки «+» и «–», начиная с клетки (1, 3). Далее вычислим шаг пересчета $h=\min\{40,20\}=20$ и строим новый план перевозок (таблица 4.4).

Таблица 4.4

$a_i \setminus b_j$	40	60	90	70	u_i
60	1 **40**	2	4 **20**	3	0
80	6	3 **60**	8 **20**	5	4
100	4	7	6 **30**	2 **70**	2
20	0	0	0 **20**	0	–4
v_j	1	–1	4	0	

Число базисных клеток равно 7, следовательно, план не вырожден. Рассчитаем потенциалы и проверим план на оптимальность:

$\Delta_{12} = 2 - 0 + 1 = 3 > 0$, $\Delta_{14} = 3 - 0 - 0 = 3 > 0$,

$\Delta_{21} = 6 - 4 - 1 = 1 > 0$, $\Delta_{24} = 5 - 4 - 0 = 1 > 0$,

$\Delta_{31} = 4 - 2 - 1 = 1 > 0$, $\Delta_{32} = 7 - 2 + 1 = 6 > 0$,

$\Delta_{41} = 0 + 4 - 1 = 3 > 0$, $\Delta_{42} = 0 + 4 + 1 = 5 > 0$,

$\Delta_{43} = 0 + 4 - 0 = 4 > 0$.

Поскольку все разности Δ_{ij} неотрицательны, то построенный план является оптимальным.

2. Метод минимального элемента

Среди всех стоимостей перевозок выберем минимальную. Это будут нули, расположенные в четвертой, фиктивной, строке. Остановимся на первом столбце и поставим перевозку в клетку (4, 1): $x_{41} = \min\{a_4, b_1\} = \min\{20, 40\} = 20$. Поскольку весь груз из четвертого склада вывезен, то его исключаем из рассмотрения.

Среди оставшихся клеток выбираем клетку (1, 1) с наименьшей стоимостью и назначаем перевозку $x_{11} = \min\{a_1, b_1 - x_{41}\} = \min\{60, 20\} = 20$. Поскольку все потребности первого магазина удовлетворены, то его исключаем из рассмотрения.

Далее аналогично ставим следующие перевозки:

$x_{12} = \min\{a_1 - x_{11}, b_2\} = \min\{40, 60\} = 40$, исключаем из рассмотрения первую строку;

$x_{34} = \min\{a_3, b_4\} = \min\{100, 70\} = 70$, исключаем из рассмотрения четвертый столбец;

$x_{22} = \min\{a_2, b_2 - x_{12}\} = \min\{80, 20\} = 20$, исключаем из рассмотрения второй столбец;

$x_{33} = \min\{a_3 - x_{34}, b_3\} = \min\{30, 90\} = 30$, исключаем из рассмотрения третью строку;

$x_{23} = \min\{a_2 - x_{22}, b_3 - x_{33}\} = \min\{60, 60\} = 60$.

Весь груз вывезен со складов, потребности всех магазинов удовлетворены. Получили первый опорный план (таблица 4.5), который является невырожденным, поскольку число базисных клеток равно 7.

Таблица 4.5

$a_i \setminus b_j$	40	60	90	70	u_i
60	1 + **20**	2 – **40**	4	3	0
80	6	3 + **20**	8 – **60**	5	1
100	4	7	6 **30**	2 **70**	–1
20	0 – **20**	0	0 +	0	–1
v_j	1	2	7	3	

Рассчитаем стоимость такого плана перевозок:

$F = 1 \cdot 20 + 2 \cdot 40 + 3 \cdot 20 + 8 \cdot 60 + 6 \cdot 30 + 2 \cdot 70 + 0 \cdot 20 = 900$.

Рассчитаем потенциалы и проверим план на оптимальность:

$\Delta_{13} = 4 - 0 - 7 = -3 < 0$, $\Delta_{14} = 3 - 0 - 3 = 0$,

$\Delta_{21} = 6 - 1 - 1 = 4 > 0$, $\Delta_{24} = 5 - 1 - 3 = 1 > 0$,

$\Delta_{31} = 4 + 1 - 1 = 4 > 0$, $\Delta_{32} = 7 + 1 - 2 = 6 > 0$,

$\Delta_{42} = 0 + 1 - 2 = -1 < 0$, $\Delta_{43} = 0 + 1 - 7 = -6 < 0$,

$\Delta_{44} = 0 + 1 - 3 = -2 < 0$.

Поскольку среди Δ_{ij} есть отрицательные, то план не оптимален. Для клетки (4, 3) с наименьшей отрицательной разностью строим цикл перераспределения: это шестиугольник с вершинами в клетках (4, 3), (2, 3), (2, 2), (1, 2), (1, 1), (4, 1). Расставим, чередуя, знаки «+» и «–», начиная с клетки (4, 3). Далее вычислим шаг пересчета $h = \min\{20, 40, 60\} = 20$ и строим новый план перевозок (таблица 4.6).

Таблица 4.6

$a_i \setminus b_j$	40	60	90	70	u_i
60	1 **40**	2 **20**	4	3	
80	6	3 **40**	8 **40**	5	
100	4	7	6 **30**	2 **70**	
20	0	0	0 **20**	0	
v_j					

План, соответствующий таблице 4.6, совпадает с планом таблицы 4.3, который не является оптимальным. Дальнейший процесс решения полностью совпадает с описанным выше (в случае метода северо-западного угла). В итоге получаем такое же решение.

Окончательно получаем оптимальный план перевозок

$$X = \begin{pmatrix} 40 & 0 & 20 & 0 \\ 0 & 60 & 20 & 0 \\ 0 & 0 & 30 & 70 \\ 0 & 0 & 20 & 0 \end{pmatrix},$$

состоящий в перевозке 40 ед. груза с первого склада в первый магазин, 20 ед. груза с первого склада в третий магазин, 60 ед. груза со второго склада во второй магазин, 20 ед. груза со второго склада в третий магазин, 30 ед. груза с третьего склада в третий магазин, 70 ед. груза с третьего склада в четвертый магазин, в третий магазин не довезут 20 ед. груза. Стоимость такого плана перевозок составит

$$F = 1 \cdot 40 + 4 \cdot 20 + 3 \cdot 60 + 8 \cdot 20 + 6 \cdot 30 + 2 \cdot 70 + 0 \cdot 20 = 780 \text{ тыс. руб.}$$

ГЛАВА 5. Модели динамического программирования

5.1. Предмет динамического программирования

Динамическое программирование (ДП) представляет собой математический аппарат, который подходит к решению некоторого класса задач путем их разложения на части, небольшие и менее сложные задачи. При этом отличительной особенностью является решение задач по этапам, через фиксированные интервалы, промежутки времени, что и определило появление термина «динамическое программирование». Следует заметить, что методы динамического программирования успешно применяются и при решении задач, в которых фактор времени не учитывается. В целом математический аппарат можно представить как пошаговое или поэтапное программирование. Решение задач методами динамического программирования проводится на основе сформулированного Р.Э. Беллманом принципа оптимальности: *оптимальное поведение обладает тем свойством, что какими бы ни были первоначальное состояние системы и первоначальное решение, последующее решение должно определять оптимальное поведение относительно состояния, полученного в результате первоначального решения.*

Из этого следует, что планирование каждого шага должно проводиться с учетом общей выгоды, получаемой по завершении всего процесса, что и позволяет оптимизировать конечный результат по выбранному критерию.

Таким образом, динамическое программирование в широком смысле представляет собой оптимальное управление процессом посредством изменения управляемых параметров на каждом шаге, и, следовательно, воздействовать на ход процесса, изменяя на каждом шаге состояние системы.

Вместе с тем ДП свойственны и недостатки. Прежде всего, в нем нет единого универсального метода решения. Практически каждая задача, решаемая этим методом, характеризуется своими особенностями и требует проведения поиска наиболее приемлемой совокупности методов для ее решения. Кроме того, большие объемы и трудоемкость решения многошаговых задач, имеющих множество состояний, приводят к необходимости отбора задач малой размерности либо использования сжатой информации. Последнее достигается с помощью методов анализа вариантов и переработки списка состояний.

ДП применяется для решения задач, в которых поиск оптимума возможен при поэтапном подходе, например распределение дефицитных капитальных вложений между новыми направлениями их использования; разработка правил управления спросом или запасами, устанавливающими момент пополнения запаса и размер пополняющего заказа; разработка принципов календарного планирования производства и выравнивания занятости в условиях колеблющегося спроса на продукцию; составления календарных планов текущего и капитального ремонтов оборудования и его замены; поиск кратчайших расстояний на транспортной сети; формирование последовательности развития коммерческой операции и т.д .

5.2. Постановка задачи динамического программирования

Рассмотрим постановку задачи ДП на примере инвестирования, связанного с распределением средств между несколькими предприятиями. В результате управления инвестициями система последовательно переводится из начального состояния S_0 в конечное S_n. Предположим, что управление можно разбить на *n* шагов и решение принимается последовательно на каждом шаге, а управление представляет собой совокупность *n* пошаговых управлений. На каждом шаге необходимо определить два типа переменных – переменную состояния системы S_k и переменную управления x_k. Переменная S_k определяет, в каких состояниях может оказаться система на рассматриваемом *k*-м шаге. В зависимости от состояния S_k на этом шаге можно применить некоторые управления, которые характеризуются переменной x_k, удовлетворяющей определенным ограничениям, и называются допустимыми.

Допустим $\overline{X} = (x_1, x_2, \ldots, x_k, \ldots, x_n)$ арифметический вектор – управление, переводящий систему из состояния S_0 в состояние S_n, а S_k – промежуточное состояние системы на k-м шаге управления. Тогда последовательность состояний системы можно представить в виде графа (рисунок 5.1).

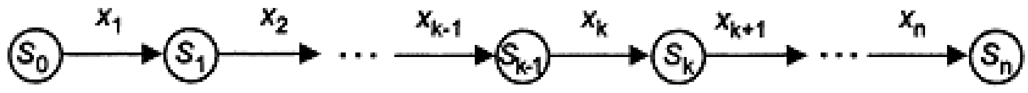

Рисунок 5.1 – Граф состояний системы.

Применение управляющего воздействия x_k на каждом шаге переводит систему в новое состояние S_k и приносит некоторый результат $\varphi_k(S_{k-1}, x_k)$. Для каждого возможного состояния на каждом шаге среди всех возможных управлений выбирается оптимальное управление x_k^* – такое, чтобы результат, который достигается за шаги с k-го по последний n-й, оказался бы оптимальным. Числовая характеристика этого результата называется функцией Беллмана $F_k(S_k)$и зависит от номера шага k и состояния системы S_{k-1}.

Задача динамического программирования формулируется следующим образом: требуется определить такое управление $\overline{X}^*$, переводящее систему изначального состояния S_0 в конечное состояние S_n, при котором целевая функция принимает наибольшее (наименьшее) значение $F\left(S_0, \overline{X}^*\right) \to extr$.

Особенности математической модели динамического программирования заключаются в следующем:

1) задача оптимизации формулируется как конечный многошаговый процесс управления;

2) показатель эффективности или критерий оптимальности операции определяется целевой функцией, которая является аддитивной от каждого шага оптимизации:

$$F\left(\overline{X}\right) = \sum_{k=1}^{n} \varphi_k(S_{k-1}, x_k);$$

3) выбор управления x_k на каждом шаге зависит только от состояния системы к этому шагу S_{k-1} и не влияет на предшествующие шаги (нет обратной связи);

4) состояние системы S_k после каждого шага управления зависит только от предшествующего состояния системы S_{k-1} и управляющего воздействия x_k (отсутствие последействия) и может быть записано в виде уравнения состояния системы:

$$S_k = f_k(S_{k-1}, x_k), k = \overline{1,n};$$

5) на каждом шаге управление x_k зависит от конечного числа управляющих переменных, а состояние системы S_k зависит от конечного числа параметров;

6) оптимальное управление представляет собой арифметический вектор $\overline{X}^*$, определяемый последовательностью оптимальных пошаговых управлений $\overline{X}^* = (x_1, x_2, ..., x_k, ..., x_n)$, число которых и определяет количество шагов задачи.

5.3. Принцип оптимальности и математическое описание динамического процесса управления

В основе метода ДП лежит принцип оптимальности, впервые сформулированный в 1953 г. американским математиком Р.Э. Беллманом: *каково бы ни было состояние системы в результате какого-либо числа шагов, на ближайшем шаге нужно выбирать управление так, чтобы оно в совокупности с оптимальным управлением на всех последующих шагах приводило к оптимальному выигрышу на всех оставшихся шагах, включая выигрыш на данном шаге*. При решении задачи на каждом шаге выбирается управление, которое приводит к оптимальному выигрышу. Если считать все шаги независимыми, тогда оптимальным управлением будет то управление, которое обеспечит максимальный выигрыш на каждом шаге.

Кроме того, при выборе управления на данном шаге следует учитывать возможные варианты состояния предыдущего шага. Например, при определении количества средств, вкладываемых в предприятие в i-м году, необходимо знать, сколько средств осталось в наличии к этому году и какой доход получен в предыдущем (i–1)-м году. Таким образом, при выборе шагового управления необходимо учитывать следующие требования:

1) возможные исходы предыдущего шага S_{k-1};

2) влияние управления x_k на все оставшиеся до конца процесса шаги (i–k).

В задачах динамического программирования первое требование учитывают, делая на каждом шаге условные предположения о возможных вариантах окончания предыдущего шага и проводя для каждого из вариантов условную оптимизацию. Выполнение второго требования обеспечивается проведением безусловной оптимизации в обратном порядке.

Условная оптимизация. На первом этапе решения задачи, называемом условной оптимизацией, определяются функция Беллмана и оптимальные управления для всех возможных состояний на каждом шаге, начиная с последнего в соответствии с алгоритмом обратной прогонки. На последнем, n-м, шаге оптимальное управление x_n^* определяется функцией Беллмана

$$F(S) = \max\{\varphi_n(S, x_n)\},$$

в соответствии с которой максимум выбирается из всех возможных значений x_n, причем $x_n \in X$.

Дальнейшие вычисления проводятся согласно рекуррентному соотношению, связывающему функцию Беллмана на каждом шаге с этой же функцией, но вычисленной на предыдущем шаге. В общем виде это соотношение имеет вид

$$F_k(S) = \max_{x_k \in X}\{\varphi_k(S_{k-1}, x_k) + F_{k+1}(f_k(S_{k-1}, x_k))\} = \max_{x_k \in X}\{\varphi_k(S_{k-1}, x_k) + F_{k+1}(S_k)\}.$$

Этот максимум (или минимум) определяется по всем возможным для k и S значениям переменной управлениях $x_k \in X$.

Безусловная оптимизация. После того как функция Беллмана и соответствующие оптимальные управления найдены для всех шагов с *n*-го по первый, осуществляется второй этап решения задачи, называемый безусловной оптимизацией, проводимой в обратном порядке.

Пользуясь тем, что на первом шаге (k=1) состояние системы известно – это ее начальное состояние S_0, можно найти оптимальный результат за все n шагов и оптимальное управление на первом шаге x_1^*, которое этот результат доставляет. После применения этого управления система перейдет в другое состояние $S_1 = f_1\left(S_0, x_1^*\right)$, зная которое можно, пользуясь результатами условной оптимизации, найти оптимальное управление на втором шаге x_2^* и так далее до последнего *n*-го шага.

Вычислительную схему динамического программирования можно строить на сетевых моделях, а также по алгоритмам прямой прогонки (от начала) и обратной прогонки (от конца к началу). Рассмотрим примеры решения различных по своей природе задач, содержание которых требует выбора переменных состояния и управления.

5.4. Оптимальное распределение инвестиций

Требуется распределить имеющиеся B единиц средств среди n предприятий, доход $g_j(x_i)$ от которых в зависимости от количества вложенных средств x_i определяется матрицей (mxn), приведенной в таблице 5.1, так, чтобы суммарный доход со всех предприятий был бы максимальным. Состояние системы перед каждым шагом определяется числом еще не вложенных средств.

Запишем математическую модель задачи.

Определить $\overline{X}^* = \left(x_1^*, x_2^*, \ldots, x_i^*, \ldots, x_n^*\right)$, удовлетворяющий условиям

$$\begin{aligned} &\sum_{i=1}^{n} x_i = B, \\ &x_i \geq 0,\ i = \overline{1, n}, \end{aligned} \tag{5.1}$$

и обеспечивающий максимум целевой функции

$$F\left(\overline{X}\right)=\sum_{i=1}^{n} g_i(x_i) \to \max . \quad (5.2)$$

Таблица 5.1

$x \setminus g_i$	g_1	g_2	…	g_j	…	g_n
x_1	$g_1(x_1)$	$g_2(x_1)$	…	$g_j(x_1)$	…	$g_n(x_1)$
x_2	$g_1(x_2)$	$g_2(x_2)$	…	$g_j(x_2)$	…	$g_n(x_2)$
…	…	…	…	…	…	…
x_i	$g_1(x_i)$	$g_2(x_i)$	…	$g_j(x_i)$	…	$g_n(x_i)$
…	…	…	…	…	…	…
x_m	$g_1(x_m)$	$g_2(x_m)$	…	$g_j(x_m)$	…	$g_n(x_m)$

Теперь необходимо записать рекуррентное соотношение связи между шагами управления $F_k(x)$ и $F_{k+1}(x)$. Очевидно, эта задача может быть решена простым перебором всех возможных вариантов распределения *B* единиц средств по *n* предприятиям, на пример на сетевой модели. Однако ее можно решить более эффективным методом, который заключается в замене сложной многовариантной задачи многократным решением простых задач с малым количеством исследуемых вариантов.

С этой целью разобьем процесс оптимизации на *n* шагов и будем на каждом *k*-м шаге оптимизировать инвестирование не всех предприятий, а только предприятий с *k*-го по *n*-е. При этом естественно считать, что в остальные предприятия (с первого по (*k*-1)-е) тоже вкладываются средства, и поэтому на инвестирование предприятий с *k*-го по *n*-е остаются не все средства, а некоторая меньшая сумма $c_k \le B$. Эта величина и будет являться переменной состояния системы. Переменной управления на *k*-м шаге назовем величину x_k средств, вкладываемых в *k*-е предприятие. В качестве функции Беллмана $F_k\left(c_k\right)$ на *k*-м шаге можно выбрать максимально возможный доход, который можно получить с предприятий с *k*-го по *n*-е при условии, что на их инвестирование осталось c_k средств. Очевидно, что при вложении в *k*-е предприятие x_k

средств будет получена прибыль $g_k(x_k)$, а система к (k+1)-му шагу перейдет в состояние S_{k+1} и, следовательно, на инвестирование предприятий с (k+1)-го до n-го останется средств $c_{k+1} = c_k - x_k$.

Таким образом, на **первом шаге условной оптимизации** при k=n функция Беллмана представляет собой прибыль только с n-го предприятия. При этом на его инвестирование может остаться количество средств $c_n, 0 \le c_n \le B$. Чтобы получить максимум прибыли с этого предприятия, можно, например, вложить в него все эти средства, т.е. $F_n(c_n) = g_n(c_n)$ и $x_n = c_n$.

На каждом последующем шаге для вычисления функции Беллмана необходимо использовать результаты предыдущего шага. Пусть на k-м шаге для инвестирования предприятий с k-го по n-е осталось c_k средств ($0 \le c_k \le B$). Тогда от вложения в k-е предприятие x_k средств будет получена прибыль $g_k(x_k)$, а на инвестирование остальных предприятий (с k-го по n-е) останется $c_{k+1} = c_k - x_k$ средств. Максимально возможный доход, который может быть получен с предприятий (с k-го по n-е), будет равен:

$$F_k(c_k) = \max_{x_k \le c_k}\{g_k(x_k) + F_{k+1}(c_k - x_k)\}, \;\; k = \overline{1, n}. \tag{5.3}$$

Максимум выражения (5) достигается на некотором значении x_k^*, которое является оптимальным управлением на k-м шаге для состояния системы S_k. Действуя таким образом, можно определить функцию Беллмана и оптимальные управления последовательно вплоть до шага k=1.

Значение функции Беллмана $F_1(c_1)$ представляет собой максимально возможный доход со всех предприятий, а значение x_1^*, на котором достигается максимум дохода, является оптимальным количеством средств, вложенных в первое предприятие. Затем на этапе безусловной оптимизации для всех последующих шагов вычисляется величина $c_k = c_{k-1} - x_{k-1}$ и оптимальным управлением на k-м шаге является то значение x_k, которое обеспечивает максимум дохода при соответствующем состоянии системы S_k.

Пример 5.1. Для инвестирования в три предприятия выделено 5 млн. руб. Известна эффективность капитальных вложений в каждое предприятие, заданная значением нелинейной функции $g_j(x_i)$, представленной в таблице 5.2. Необходимо распределить выделенные средства между предприятиями таким образом, чтобы получить максимальный суммарный доход.

Для упрощения расчетов предполагаем, что распределение средств осуществляется в целых числах $x_i = \{0, 1, 2, 3, 4, 5\}$ млн. руб.

Таблица 5.2

x_i	g_1	g_2	g_3
0	0	0	0
1	2,2	2	2,8
2	3	3,2	5,4
3	4,1	4,8	6,4
4	5,2	6,2	6,6
5	5,9	6,4	6,9

Решение. Условная оптимизация (*k* = 3, 2, 1)

Шаг 1. $k = 3$. Предположим, что все средства в количестве $x_3 = 5$ млн руб. отданы третьему предприятию. В этом случае максимальный доход, как это видно из таблицы 5.3, составит $g_3(x_3) = 6{,}9$ тыс. руб., следовательно: $F_3(c_3) = g_3(x_3)$.

Таблица 5.3

$c_3 \setminus x_3$	0	1	2	3	4	5	$F_3(c_3)$	x_3^*
0	0	-	-	-	-	-	0	0
1	-	2,8	-	-	-	-	2,8	1
2	-	-	5,4	-	-	-	**5,4**	2
3	-	-	-	6,4	-	-	6,4	3
4	-	-	-	-	6,6	-	6,6	4
5	-	-	-	-	-	6,9	6,9	5

Шаг 2. k = 2. Определяем оптимальную стратегию при распределении денежных средств между вторым и третьим предприятиями. При этом рекуррентное соотношение Беллмана имеет вид

$$F_2(c_2) = \max_{x_2 \le c_2} \{g_2(x_2) + F_3(c_2 - x_2)\},$$

на основе которого составлена таблица 5.4 по данным таблиц 5.2 и 5.3.

Таблица 5.4

$c_2 \setminus x_2$	0	1	2	3	4	5	$F_2(c_2)$	x_2^*
0	0+0	-	-	-	-	-	0	0
1	0+2,8	2+0	-	-	-	-	2,8	0
2	0+5,4	2+2,8	3,2+0	-	-	-	5,4	0
3	0+6,4	2+5,4	3,2+2,8	4,8+0	-	-	7,4	1
4	0+6,6	2+6,4	**3,2+5,4**	4,8+2,8	6,2+0	-	**8,6**	2
5	0+6,9	2+6,6	3,2+6,4	4,8+5,4	6,2+2,8	6,4+0	10,2	3

Шаг 3. k = 1. Определяем оптимальную стратегию при распределении денежных средств между первым и двумя другими предприятиями, используя следующую формулу для расчета суммарного дохода:

$$F_1(c_1) = \max_{x_1 \le c_1} \{g_1(x_1) + F_2(c_1 - x_1)\},$$

на основе которого составлена таблица 5.5.

Таблица 5.5

$c_1 \setminus x_1$	0	1	2	3	4	5	$F_1(c_1)$	x_1^*
0	0+0	-	-	-	-	-	0	0
1	0+2,8	2,2+0	-	-	-	-	2,8	0
2	0+5,4	2,2+2,8	3+0	-	-	-	5,4	0
3	0+7,4	2,2+5,4	3+2,8	4,1+0	-	-	7,6	1
4	0+8,6	2,2+7,4	3+5,4	4,1+2,8	5,2+0	-	9,6	1
5	0+10,2	**2,2+8,6**	3+7,4	4,1+5,4	5,2+2,8	5,9+0	**10,8**	1

Безусловная оптимизация (k = 1, 2, 3)

Определяем компоненты оптимальной стратегии.

Шаг 1: k = 1. По данным из таблицы 5.5 максимальный доход при распределении 5 млн. руб. между тремя предприятиями составляет: $c_1 = 5,\ F_1(c_1) = 10{,}8$. При этом первому предприятию нужно выделить $x_1^* = 1$ млн. руб.

Шаг 2: k = 2. Определяем величину оставшихся денежных средств, приходящуюся на долю второго и третьего предприятий: $c_2 = c_1 - x_1^* = 4$ млн. руб. По данным таблицы 4.4 находим, что оптимальный вариант распределения денежных средств размером 4 млн. руб. между вторым и третьим предприятиями составляет: $F_2(4) = 8{,}6$ при выделении второму предприятию $x_2^* = 2$ млн. руб.

Шаг 3: k = 3. Определяем величину оставшихся денежных средств, приходящуюся на долю третьего предприятия: $c_3 = c_2 - x_2^* = 2$ млн. руб. По данным таблицы 4.3 находим: $F_3(2) = 5{,}4$ и $x_3^* = 2$ млн. руб.

Таким образом, оптимальный план инвестирования предприятий:

$x_1^* = 1$ млн. руб., $x_2^* = 2$ млн. руб., $x_3^* = 2$ млн. руб.

Данный план обеспечит максимальный доход, равный

$F(5) = g_1(1) + g_2(2) + g_3(2) = 2{,}2 + 3{,}2 + 5{,}4 = 10{,}8$ млн. руб.

5.5. Выбор оптимального маршрута перевозки грузов

Математический аппарат ДП, основанный на методологии пошаговой оптимизации, может быть использован при нахождении кратчайших расстояний, например, на географической карте, представленной в виде сети. Решение задачи по определению кратчайших расстояний между пунктами отправления и пунктами получения продукции по существующей транспортной сети является исходным этапом при решении таких экономических задач, как оптимальное прикрепление потребителей к поставщикам, повышение производительности транспорта за счет сокращения непроизводительного пробега и др.

Пусть транспортная сеть состоит из 10 узлов. На рисунке 5.2 показаны сеть дорог и стоимость перевозки единицы груза между пунктами сети. Ребра являются вариантами выбора решения. Необходимо определить маршрут доставки груза из пункта 1 в пункт 10, обеспечивающий наименьшие транспортные расходы.

В задаче имеется ограничение – двигаться по изображенным на схеме маршрутам можно только слева направо. Эта особенность транспортной сети дает право отнести каждый из десяти пунктов к одному из поясов. Будем считать, что пункт принадлежит k-му поясу, если из него попасть в конечный пункт ровно за k шагов, т.е. с заездом ровно в $(k-1)$-й промежуточный пункт. Таким образом, пункты 7, 8 и 9 принадлежат к первому поясу, 5 и 6 – ко второму, 2, 3 и 4 – к третьему и 1 – к четвертому. Тогда на k-м шаге будем находить оптимальные маршруты перевозки груза из пунктов k-го пояса до конечного пункта. Оптимизацию будем производить с конца процесса, и потому, дойдя до k-то шага, неизвестно, в каком из пунктов k-го пояса окажется груз, перевозимый из первого пункта.

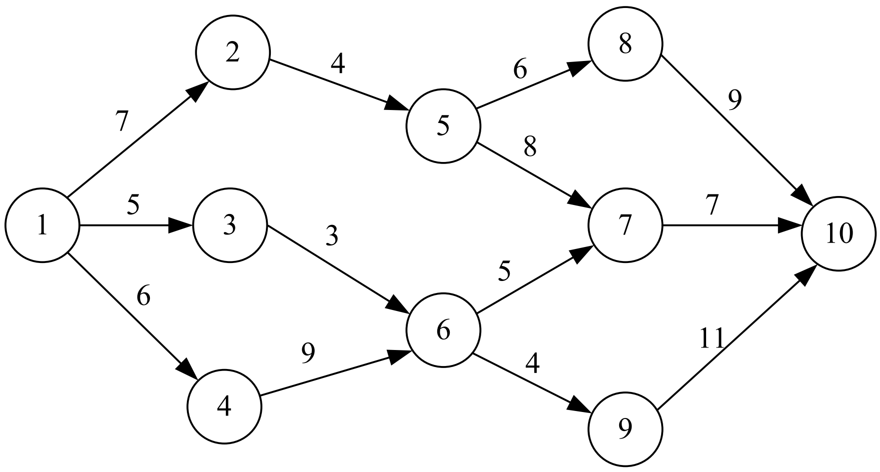

Рисунок 5.2 – Модель транспортной сети

Введем обозначения:

k – номер шага ($k = 1, 2, 3, 4$);

i – пункт, из которого осуществляются перевозки (i = 1,2, ..., 9);

j – пункт, в который доставляется груз (j = 2, 3, ..., 10);

c_{ij} – стоимость перевозки груза из пункта i в пункт j.

$F_k(i)$ – минимальные затраты на перевозку груза на k-м шаге решения задачи из пункта i до конечного пункта.

Очевидно, что минимум затрат на перевозку груза из пунктов k-то пояса до пункта 10 будет зависеть от того, в каком пункте этого пояса мы оказались. Номер i пункта, узел, принадлежащий k-му поясу, будет являться переменной состояния системы на k-м шаге. Поскольку оптимизация осуществляется с конца процесса, то, находясь в некотором пункте i k-то пояса, принимается решение о перемещении груза в один из пунктов (k-1)-го пояса, а направление дальнейшего движения известно из предыдущих шагов. Номер j пункта (k – 1)-го пояса будет переменной управления на k-м шаге.

Для первого шага управления (k = 1) функция Беллмана представляет собой минимальные затраты на перевозку груза из пунктов 1-го пояса в конечный пункт, т.е. $F_1(i) = C_{i10}$. Для последующих шагов затраты складываются из двух слагаемых – стоимости перевозки груза C_{ij} из пункта k-го пояса в пункт (k – 1)-го пояса и минимально возможных затрат на перевозку из пункта j до конечного пункта, т.е. $F_{k-1}(j)$, Таким образом, функциональное уравнение Беллмана будет иметь вид

$$F_k(i) = \min_j \left\{ C_{ij} + F_{k-1}(j) \right\}. \qquad (5.8)$$

Минимум затрат достигается на некотором значении j^*, которое является оптимальным направлением движения из пункта i в конечный пункт.

На четвертом шаге попадаем на 4-й пояс; состояние системы становится определенным i = 1. Функция $F_4(1)$ представляет собой минимально возможные затраты по перемещению груза из 1-го пункта в 10-й. Оптимальный маршрут определяется в результате анализа всех шагов в обратном порядке, а выбор некоторого управления j на k-м шаге приводит к тому, что состояние системы на (k – 1)-м шаге становится определенным.

Пример 5.3. Решим сформулированную выше задачу, исходные данные которой приведены на рис. 5.2.

Условная оптимизация

1-й шаг: $k = 1$.

На первом шаге в пункт 10 груз может быть доставлен из пунктов 7, 8 или 9 (таблица 5.9).

Таблица 5.9

$i \setminus j$	10	$F_1(i)$	j^*
7	7	7	10
8	9	9	10
9	11	11	10

2-й шаг: $k = 2$.

Функциональное уравнение на втором шаге принимает вид

$$F_2(i) = \min_j \{C_{ij} + F_1(j)\}.$$

Все возможные перемещения груза на втором шаге и результаты расчета приведены в таблице 5.10.

Таблица 5.10

$i \setminus j$	7	8	9	$F_2(i)$	j^*
5	8+7	6+9	–	15	7, 8
6	5+7	–	4+11	12	7

3-й шаг: $k = 3$.

Функциональное уравнение на втором шаге принимает вид

$$F_3(i) = \min_j \{C_{ij} + F_2(j)\}.$$

Все возможные перемещения груза на втором шаге и результаты расчета приведены в таблице 5.11.

Таблица 5.11

$i \setminus j$	5	6	$F_3(i)$	j^*
2	4+15	–	19	5
3	–	3+12	15	6
4	–	9+12	21	6

4-й шаг: $k = 4$.

Функциональное уравнение на втором шаге принимает вид

$$F_4(i) = \min_j \{C_{ij} + F_3(j)\}.$$

Все возможные перемещения груза на втором шаге и результаты расчета приведены в таблице 5.12.

Таблица 5.12

$i \backslash j$	2	3	4	$F_3(i)$	j^*
1	7+19	5+15	6+21	20	3

Безусловная оптимизация

На этапе условной оптимизации получено, что минимальные затраты на перевозку груза из пункта 1 в пункт 10 составляют $F_4(1) = 20$. Данный результат достигается при движении груза из 1-го пункта в 3-й. По данным табл. 5.11, из пункта 3 необходимо двигаться в пункт 6, затем – в пункт 7 (см. табл. 5.10) и из него – в конечный пункт (см. табл. 5.9).

Таким образом, оптимальный маршрут доставки груза: $1 \to 3 \to 6 \to 7 \to 10$. На рисунке 5.3 он показан жирными стрелками.

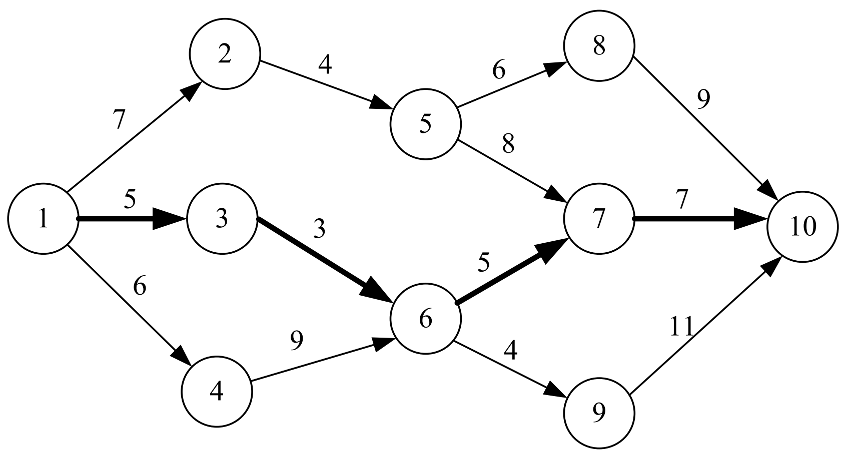

Рисунок 5.3 – Транспортная сеть с оптимальным маршрутом

ГЛАВА 6. Моделирование систем массового обслуживания

6.1. Основные понятия систем массового обслуживания

Системы массового обслуживания (СМО) – это такие системы, в которые в случайные моменты времени поступают заявки на обслуживание, при этом поступившие заявки обслуживаются с помощью имеющихся в распоряжении системы каналов обслуживания.

С позиции моделирования процесса массового обслуживания ситуации, когда образуются очереди заявок (требований) на обслуживание, возникают следующим образом. Поступив в обслуживающую систему, требование присоединяется к очереди других (ранее поступивших) требований. Канал обслуживания выбирает требование из находящихся в очереди, с тем, чтобы приступить к его обслуживанию. После завершения процедуры обслуживания очередного требования канал обслуживания приступает к обслуживанию следующего требования, если таковое имеется в блоке ожидания.

Цикл функционирования системы массового обслуживания подобного рода повторяется многократно в течение всего периода работы обслуживающей системы. При этом предполагается, что переход системы на обслуживание очередного требования после завершения обслуживания предыдущего требования происходит мгновенно, в случайные моменты времени.

Примерами систем массового обслуживания могут служить:

• магазины;

• банки;

• ремонтные мастерские;

• почтовые отделения;

• посты технического обслуживания автомобилей, посты ремонта автомобилей;

• персональные компьютеры, обслуживающие поступающие заявки или требования на решение тех или иных задач;

• аудиторские фирмы;

• отделы налоговых инспекций, занимающиеся приемкой и проверкой текущей отчетности предприятий;

• телефонные станции и т. д.

Основными компонентами системы массового обслуживания любого вида являются:

– входной поток поступающих требований или заявок на обслуживание;

– дисциплина очереди;

– механизм обслуживания.

Раскроем содержание каждого из указанных выше компонентов.

Входной поток требований. Для описания входного потока требуется задать вероятностный закон, определяющий последовательность моментов поступления требований на обслуживание и указать количество таких требований в каждом очередном поступлении. При этом, как правило, оперируют понятием «вероятностное распределение моментов поступления требований». Здесь могут поступать как единичные, так и групповые требования (требования поступают группами в систему). В последнем случае обычно речь идет о системе обслуживания с параллельно-групповым обслуживанием.

Дисциплина очереди – это важный компонент системы массового обслуживания, он определяет принцип, в соответствии с которым поступающие на вход обслуживающей системы требования подключаются из очереди к процедуре обслуживания.

Чаще всего используются дисциплины очереди, определяемые следующими правилами:

– первым пришел – первый обслуживаешься (FIFO);

– пришел последним – обслуживаешься первым (LIFO);

– случайный отбор заявок;

– отбор заявок по критерию приоритетности;

– ограничение времени ожидания момента наступления обслуживания (имеет место очередь с ограниченным временем ожидания обслуживания, что ассоциируется с понятием «допустимая длина очереди»).

Механизм обслуживания определяется характеристиками самой процедуры обслуживания и структурой обслуживающей системы. К характеристикам процедуры обслуживания относятся:

– продолжительность процедуры обслуживания

– количество требований, удовлетворяемых в результате выполнения каждой такой процедуры.

Для аналитического описания характеристик процедуры обслуживания оперируют понятием «вероятностное распределение времени обслуживания требований».

Следует отметить, что время обслуживания заявки зависит от характера самой заявки или требований клиента и от состояния и возможностей обслуживающей системы. В ряде случаев приходится также учитывать вероятность выхода обслуживающего прибора по истечении некоторого ограниченного интервала времени.

Структура обслуживающей системы определяется количеством и взаимным расположением каналов обслуживания (механизмов, приборов и т. п.). Прежде всего, следует подчеркнуть, что система обслуживания может иметь не один канал обслуживания, а несколько; система такого рода способна обслуживать одновременно несколько требований. В этом случае все каналы обслуживания предлагают одни и те же услуги, и, следовательно, можно утверждать, что имеет место параллельное обслуживание.

Система обслуживания может состоять из нескольких разнотипных каналов обслуживания, через которые должно пройти каждое обслуживаемое требование, т.е. в обслуживающей системе процедуры обслуживания требований реализуются последовательно. Механизм обслуживания определяет характеристики выходящего (обслуженного) потока требований.

Рассмотрев основные компоненты систем обслуживания, можно констатировать, что **функциональные возможности любой системы массового обслуживания определяются следующими основными факторами:**

– вероятностным распределением моментов поступлений заявок на обслуживание (единичных или групповых);

– вероятностным распределением времени продолжительности обслуживания;

– конфигурацией обслуживающей системы (параллельное, последовательное или параллельно-последовательное обслуживание);

– количеством и производительностью обслуживающих каналов;

– дисциплиной очереди;

– мощностью источника требований.

В качестве основных критериев эффективности функционирования систем массового обслуживания, в зависимости от характера решаемой задачи, могут выступать:

– вероятность немедленного обслуживания поступившей заявки;

– вероятность отказа в обслуживании поступившей заявки;

– относительная и абсолютная пропускная способность системы;

– средний процент заявок, получивших отказ в обслуживании;

– среднее время ожидания в очереди;

– средняя длина очереди;

– средний доход от функционирования системы в единицу времени и т.п.

Предметом теории массового обслуживания является установление зависимости между факторами, определяющими функциональные возможности системы массового обслуживания, и эффективностью ее функционирования. В большинстве случаев все параметры, описывающие системы массового обслуживания, являются случайными величинами или функциями, поэтому эти системы относятся к стохастическим системам.

Независимо от характера процесса, протекающего в системе массового обслуживания, различают **два основных вида СМО:**

– системы с отказами, в которых заявка, поступившая в систему в момент, когда все каналы заняты, получает отказ и сразу же покидает очередь;

– системы с ожиданием (очередью), в которых заявка, поступившая в момент, когда все каналы обслуживания заняты, становится в очередь и ждет, пока не освободится один из каналов.

Системы массового обслуживания с ожиданием делятся на системы с ограниченным ожиданием и системы с неограниченным ожиданием.

В системах с **ограниченным ожиданием** может ограничиваться:

– длина очереди;

– время пребывания в очереди.

В системах с **неограниченным ожиданием** заявка, стоящая в очереди, ждет обслуживание неограниченно долго, т.е. пока не подойдет очередь.

Все системы массового обслуживания различают **по числу каналов обслуживания**:

– одноканальные системы;

– многоканальные системы.

Приведенная классификация СМО является условной. На практике чаще всего системы массового обслуживания выступают в качестве смешанных систем. Например, заявки ожидают начала обслуживания до определенного момента, после чего система начинает работать как система с отказами.

Определим характеристики основных систем массового обслуживания.

6.2. Одноканальная СМО с отказами

Рассмотрим следующую задачу. Имеется один канал, на который поступает поток заявок с интенсивностью λ. Поток обслуживаний имеет интенсивность μ. Найти предельные вероятности состояний системы и показатели ее эффективности.

Среднее время обслуживания $\overline{T}_{об} = 1/\mu$.

Среднее время простоя канала $\overline{T}_{пр} = 1/\lambda$.

Система S (СМО) имеет два состояния: S_0 – канал свободен, S_1 – канал занят. Тогда ***предельные вероятности состояний*** равны

$$p_0 = \frac{\mu}{\lambda+\mu}, \; p_1 = \frac{\lambda}{\lambda+\mu}.$$

Предельные вероятности состояний можно выразить через средние времена простоя канала и обслуживания одной заявки:

$$p_0 = \frac{\overline{T}_{пр}}{\overline{T}_{об} + \overline{T}_{пр}}, \; p_1 = \frac{\overline{T}_{об}}{\overline{T}_{об} + \overline{T}_{пр}}.$$

Относительная пропускная способность Q системы и ***вероятность отказа*** $P_{отк}$ равны

$$Q = p_0, \; P_{отк} = p_1.$$

Абсолютная пропускная способность равна

$$A = \lambda Q = \frac{\lambda\mu}{\lambda+\mu}.$$

Пример 6.1. В фирму поступает простейший поток заявок на телефонные переговоры с интенсивностью λ=90 вызовов в час, а средняя продолжительность разговора по телефону $\overline{T}_{об} = 2$ мин. Определить показатели эффективности работы СМО (телефонной связи) при наличии одного телефонного номера.

Решение. Интенсивность потока обслуживаний

$$\mu = 1/\overline{T}_{об} = 0{,}5\,(1/мин.) = 30\,(1/ч.).$$

Относительная пропускная способность СМО $Q = \frac{30}{90+90} = 0{,}25$, т.е. в среднем только 25% поступающих заявок осуществят переговоры по телефону. Соответственно вероятность отказа составит $P_{отк} = 1 - Q = 1 - 0{,}25 = 0{,}75$. Абсолютная пропускная способность СМО $A = \lambda Q = 90 \cdot 0{,}25 = 22{,}5$, т.е. в среднем в час будут обслужены 22,5 заявки на переговоры. Очевидно, что при наличии только одного телефонного номера СМО будет плохо справляться с потоком заявок.

6.3. Многоканальная система с отказами (задача Эрланга)

Здесь мы рассмотрим одну из первых по времени, «классических» задач теории массового обслуживания; эта задача возникла из практических нужд телефонии и была решена в 1909 г. датским инженером-математиком А.К. Эрлангом. Задача ставится так: имеется n каналов (линий связи), на которые поступает поток заявок с интенсивностью λ. Поток обслуживаний каждого канала имеет интенсивность μ. Найти предельные вероятности состояний системы и показатели ее эффективности.

Система S (СМО) имеет следующие состояния (нумеруем их по числу заявок, находящихся в системе): $S_0, S_1, S_2, \ldots, S_n$, где S_k – состояние системы, когда в ней находится k заявок, т.е. занято k каналов.

Предельная вероятность состояния равна

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2!\,\mu^2} + \ldots + \frac{\lambda^k}{k!\,\mu^k} + \ldots + \frac{\lambda^n}{n!\,\mu^n}\right)^{-1},$$

где члены разложения $\frac{\lambda}{\mu}, \frac{\lambda^2}{2!\,\mu^2}, \ldots, \frac{\lambda^n}{n!\,\mu^n}$ – коэффициенты при p_0 в выражениях для предельных вероятностей $p_1, p_2, \ldots, p_n$.

Заметим, что в формулу для p_0 интенсивности λ и μ входят не по отдельности, а только в виде отношения λ/μ . Обозначим $\frac{\lambda}{\mu} = \rho$ и будем называть величину ρ ***приведенной интенсивностью потока заявок*** или ***интенсивностью нагрузки канала***. Она выражает среднее число заявок, приходящих за среднее время обслуживания одной заявки. Пользуясь этим обозначением, перепишем формулу для p_0 в виде:

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + \ldots + \frac{\rho^k}{k!} + \ldots + \frac{\rho^n}{n!}\right)^{-1}.$$

При этом

$$p_1 = \rho\, p_0, \quad p_2 = \frac{\rho^2}{2!}\, p_0, \quad p_n = \frac{\rho^n}{n!}\, p_0.$$

Вероятность отказа СМО есть предельная вероятность того, что все n каналов системы будут заняты, т.е.

$$P_{отк} = p_n = \frac{\rho^n}{n!} p_0 .$$

Отсюда находим ***относительную пропускную способность*** – вероятность того, что заявка будет обслужена:

$$Q = 1 - P_{отк} = 1 - \frac{\rho^n}{n!} p_0 .$$

Абсолютную пропускную способность получим, умножая интенсивность потока заявок λ на Q:

$$A = \lambda Q = \lambda\left(1 - \frac{\rho^n}{n!} p_0\right).$$

Осталось только найти ***среднее число занятых каналов*** $\bar{k}$. Эту величину можно было бы найти «впрямую», как математическое ожидание дискретной случайной величины с возможными значениями 0,1,..., n и вероятностями этих значений $p_0, p_1, \ldots, p_n$:

$$\bar{k} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \ldots + n \cdot p_n = \sum_{k=0}^{n} k p_k .$$

Однако среднее число занятых каналов можно найти проще, если учесть, что абсолютная пропускная способность A системы есть не что иное, как интенсивность потока обслуженных системой заявок (в единицу времени). Так как каждый занятый канал обслуживает в среднем μ заявок (в единицу времени), то среднее число занятых каналов

$$\bar{k} = \frac{A}{\mu} = \rho\left(1 - \frac{\rho^n}{n!} p_0\right).$$

Пример 6.2. В условиях предыдущего примера 6.5 определить оптимальное число телефонных номеров в фирме, если условием оптимальности считать удовлетворение из каждых 100 заявок на переговоры в среднем не менее 90 заявок.

Решение. Интенсивность нагрузки канала ρ=90/30=3, т.е. за время среднего (по продолжительности) телефонного разговора $\overline{T}_{об} = 2$ мин. поступает в среднем 3 заявки на переговоры.

Будем постепенно увеличивать число каналов (телефонных номеров) $n = 2, 3, 4,...$ и определим для получаемой n-канальной СМО характеристики обслуживания. Значения характеристик СМО сведем в таблицу 6.1.

Таблица 6.1

n	p_0	p_1	p_2	p_3	p_4	p_5	p_6	$P_{отк}$	Q	A	$\overline{k}$
1	0,25	0,75						0,75	0,25	22,5	
2	0,118	0,353	0,529					0,529	0,471	42,353	1,412
3	0,077	0,231	0,346	0,346				0,346	0,654	58,846	1,962
4	0,061	0,183	0,275	0,275	0,206			0,206	0,794	71,45	2,382
5	0,054	0,163	0,245	0,245	0,183	0,11		0,11	0,89	80,095	2,67
6	0,052	0,155	0,232	0,232	0,174	0,104	0,052	0,052	0,948	85,306	2,844

По условию оптимальности $Q \geq 0{,}9$, следовательно, в фирме необходимо установить 6 телефонных номеров (в этом случае Q = 0,948). При этом в час будут обслуживаться в среднем 85 заявок (A = 85,306), а среднее число занятых телефонных номеров (каналов) $\overline{k} = 2{,}844$.

Тут уже проглядывает некоторый намек на *оптимизацию*. В самом деле, содержание каждого канала в единицу времени обходится в какую-то сумму. Вместе с тем, каждая обслуженная заявка приносит какой-то доход (если речь идет о СМО, для которых этот доход можно оценить). Умножая этот доход на среднее число заявок A, обслуживаемых в единицу времени, мы получим средний доход от СМО в единицу времени. Естественно, при увеличении числа каналов этот доход растет, но растут и расходы, связанные с содержанием каналов. Что перевесит – увеличение доходов или расходов? Это зависит от условий операции, т.е. от «платы за обслуживание заявки» и от стоимости содержания канала. Зная эти величины, можно найти оптимальное число каналов, наиболее экономически эффективное.

6.4. Одноканальная СМО с ожиданием и ограничением на длину очереди

На практике довольно часто встречаются одноканальные СМО с очередью (врач, обслуживающий пациентов; кассир, выдающий зарплату; телефон-автомат на улице и т.д.). В теории массового обслуживания одноканальные СМО с очередью также занимают особое место: именно к таким СМО относится большинство полученных до сих пор аналитических формул для немарковских систем.

Рассмотрим одноканальную СМО, на вход которой поступает простейший поток заявок с интенсивностью λ. Предположим, что поток обслуживаний также простейший с интенсивностью μ. Это означает, что непрерывно занятый канал обслуживает в среднем μ заявок в единицу времени. Заявка, поступившая в СМО в момент, когда канал занят, в отличие от СМО с отказами, не покидает систему, а становится в очередь и ожидает обслуживания.

Далее предполагаем, что в данной системе имеется ограничение на длину очереди, под которой понимается максимальное число мест в очереди, а именно, предполагаем, что в очереди могут находиться максимум $m \geq 1$ заявок. Поэтому заявка, пришедшая на вход СМО, в момент, когда в очереди уже стоят m заявок, получает отказ и покидает систему необслуженной.

Таким образом, рассматриваемая СМО относится к системам *смешанного типа с ограничением на длину очереди*.

Пронумеруем состояния СМО по числу заявок, находящихся в системе, т.е. под обслуживанием и в очереди:

S_0 – канал свободен (следовательно, очереди нет);

S_1– канал занят и очереди нет, т.е. в СМО находится (под обслуживанием) одна заявка;

S_2 – канал занят и в очереди стоит одна заявка;

…

S_{m+1} – канал занят и в очереди m заявок.

При наличии только одного канала обслуживания все интенсивности потоков обслуживания равны μ.

Для описания предельного режима работы СМО можно воспользоваться изложенными ранее правилами и формулами. Запишем сразу выражения, определяющие ***предельные вероятности состояний***:

$$p_0 = \left(1 + \rho + \rho^2 + \ldots + \rho^k + \ldots + \rho^{m+1}\right)^{-1}.$$

$$p_k = \rho^k \, p_0, \; k = 1, \ldots, m+1,$$

где $\rho = \lambda/\mu$ – интенсивность нагрузки канала.

Если $\lambda = \mu$, то получаем $p_0 = p_1 = \ldots = p_{m+1} = 1/(m+2)$.

Пусть теперь $\lambda \neq \mu$ ($\rho \neq 1$). Выражение для p_0 можно в данном случае записать проще, пользуясь тем, что в знаменателе стоит сумма $m+2$ членов геометрической прогрессии со знаменателем ρ:

$$p_0 = \frac{1-\rho}{1-\rho^{m+2}}.$$

Определим основные характеристики одноканальной СМО с ожиданием:

относительная пропускная способность

$$P_{отк} = p_{m+1} = \begin{cases} \dfrac{\rho^{m+1}(1-\rho)}{1-\rho^{m+2}}, & если \;\; \rho \neq 1; \\ \dfrac{1}{m+2}, & если \;\; \rho = 1; \end{cases}$$

относительная пропускная способность, или доля обслуживаемых заявок, поступающих в единицу времени

$$Q = 1 - P_{отк} = \begin{cases} \dfrac{1-\rho^{m+1}}{1-\rho^{m+2}}, & если \;\; \rho \neq 1; \\ \dfrac{m+1}{m+2}, & если \;\; \rho = 1; \end{cases}$$

абсолютная пропускная способность системы

$A = \lambda\, Q$;

среднее число заявок (***средняя длина очереди)***

$$L_{оч} = \begin{cases} \dfrac{\rho^2\left(1-\rho^m(m+1-m\rho)\right)}{(1-\rho)\left(1-\rho^{m+2}\right)}, & если \quad \rho \neq 1; \\ \dfrac{m(m+1)}{2(m+2)}, & если \quad \rho = 1; \end{cases}$$

среднее время ожидания заявки

$$\overline{T}_{оч} = \frac{L_{оч}}{\lambda};$$

среднее время пребывания заявки в системе

$$\overline{T}_{СМО} = \overline{T}_{оч} + \overline{T}_{об}.$$

Пример 6.3. На автозаправочной станции (АЗС) имеется одна колонка. Площадка при станции, на которой машины ожидают заправку, может вместить не более трех машин одновременно, и если она занята, то очередная машина, прибывшая к станции, в очередь не становится, а проезжает на соседнюю АЗС. В среднем машины прибывают на станцию каждые 2 мин. Процесс заправки одной машины продолжается в среднем 2,5 мин. Определить основные характеристики системы.

Решение. Математической моделью данной АЗС является одноканальная СМО с ожиданием и ограничением на длину очереди (m = 3). Предполагается, что поток машин, подъезжающих к АЗС для заправки, и поток обслуживаний – простейшие.

Поскольку машины прибывают в среднем через каждые 2 мин, то интенсивность входящего потока равна λ =1/2 = 0,5 (машин в минуту). Среднее время обслуживания одной машины равно 2,5 мин, следовательно, интенсивность потока обслуживаний μ =1/2,5 = 0,4 (машины в минуту).

Определяем интенсивность нагрузки канала: $\rho = \lambda/\mu = 0{,}5/0{,}4 = 1{,}25$.

Вычисляем вероятность отказа $P_{отк} = \dfrac{\rho^4(1-\rho)}{1-\rho^5} = \dfrac{1{,}25^4(1-1{,}25)}{1-1{,}25^5} \approx 0{,}297$,

откуда относительная пропускная способность $Q = 1 - P_{отк} = 1 - 0{,}297 \approx 0{,}703$ и абсолютная пропускная способность $A = \lambda Q \approx 0{,}5 \cdot 0{,}703 \approx 0{,}352$.

Среднее число машин, ожидающих в очереди на заправку

$$L_{оч}=\frac{\rho^2\left(1-\rho^3(4-3\rho)\right)}{(1-\rho)\left(1-\rho^5\right)}=\frac{1{,}25^2\left(1-1{,}25^3(4-3\cdot 1{,}25)\right)}{(1-1{,}25)\left(1-1{,}25^5\right)}\approx 1{,}559.$$

Среднее время ожидания машины в очереди

$$\overline{T}_{оч}=\frac{L_{оч}}{\lambda}=\frac{1{,}559}{0{,}5}=3{,}118 \text{ мин.}$$

Таким образом, из анализа работы СМО следует, что из каждых 100 подъезжающих машин 30 получают отказ, т.е. обслуживаются 70% заявок. Поэтому необходимо либо сократить время обслуживания одной машины (увеличить интенсивность потока обслуживаний), либо увеличить число колонок, либо увеличить площадку для ожидания. Оптимальное решение принимается с учетом затрат, связанных соответственно с увеличением штата обслуживающего персонала (увеличение производительности канала), с расширением площадки для ожидания или приобретением дополнительной колонки, и потерь, связанных с потерей заявок на обслуживание.

6.5. Одноканальная СМО с неограниченным ожиданием

Проанализируем работу одноканальной СМО с ожиданием без ограничений на длину очереди и на время ожидания в очереди. По-прежнему будем предполагать, что входящий поток и поток обслуживаний являются простейшими и имеют интенсивности λ и μ соответственно.

Такая система представляет собой предельный случай системы, рассмотренной в предыдущем пункте, при $m\to\infty$. Таким образом, длина очереди станет бесконечной и в соответствии с этим бесконечным станет число состояний СМО.

Если отказаться от ограничения на длину очереди, то случаи $\rho < 1$ и $\rho \geq 1$ начинают существенно различаться.

Если $\lambda > \mu$ ($\rho > 1$), т.е. среднее число заявок, поступивших в систему за единицу времени, больше среднего числа обслуживаемых заявок за то же время при непрерывно работающем канале, то очевидно, что **очередь неограниченно**

растет. В этом случае предельный режим не устанавливается и предельных вероятностей состояний не существует (точнее, они равны нулю).

В случае $\lambda = \mu$ ($\rho = 1$) только при условии, что входящий поток заявок и поток обслуживаний регулярные (т.е. заявки поступают в СМО через равные интервалы времени, и время обслуживания одной заявки является постоянным, равным интервалу времени между поступлениями заявок), **очереди вообще не будет** и канал будет обслуживать заявки непрерывно. Но как только входящий поток или поток обслуживаний перестает быть регулярным и приобретает элементы случайности, очередь начинает расти до бесконечности.

Поэтому далее при рассмотрении указанных систем будем предполагать, что $\lambda < \mu$, т.е. $\rho < 1$. При этом условии с течением времени устанавливается предельный режим, и предельные вероятности состояний существуют.

Устремляя m к бесконечности в формулах для вероятностей состояний (полученных для СМО с ограниченной длиной очереди при $\rho < 1$), находим выражения для ***предельных вероятностей состояний*** рассматриваемой СМО:

$$p_k = \lim_{m\to\infty} \rho^k\, p_0 = \rho^k \lim_{m\to\infty} \frac{1-\rho}{1-\rho^{m+2}} = \rho^k (1-\rho),\ k = 0,1,2,\ldots.$$

При отсутствии ограничений на очередь каждая заявка, поступившая в СМО, рано или поздно будет обслужена. Поэтому ***вероятность отказа*** равна нулю: $P_{отк} = 0$.

Следовательно, вероятность того, что поступившая заявка будет принята в систему, так же как и ***относительная пропускная способность*** Q, равна единице: $Q = 1 - P_{отк} = 1$.

Тогда для ***абсолютной пропускной способности*** A (и интенсивности выходящего потока) будем иметь: $A = \lambda Q = \lambda$, т.е. интенсивности входящего и выходящего потоков совпадают.

Среднее число заявок в очереди равно

$$L_{оч} = \lim_{m\to\infty} \frac{\rho^2\left(1-\rho^m\left(m+1-m\rho\right)\right)}{\left(1-\rho\right)\left(1-\rho^{m+2}\right)} = \frac{\rho^2}{1-\rho}.$$

Среднее время ожидания заявки в очереди равно

$$\overline{T}_{оч} = \frac{L_{оч}}{\lambda} = \frac{\rho^2}{\lambda(1-\rho)} = \frac{\rho}{\mu(1-\rho)}.$$

Наконец, ***среднее время пребывания заявки*** в СМО $T_{СМО}$ складывается из среднего времени заявки в очереди и среднего времени обслуживания заявки:

$$\overline{T}_{СМО} = \overline{T}_{оч} + \overline{T}_{об} = \frac{\rho}{\mu(1-\rho)} + \frac{1}{\mu} = \frac{1}{\mu(1-\rho)} = \frac{\rho}{\lambda(1-\rho)}.$$

Пример 6.4. В парикмахерской работает только один мужской мастер. Среднее время стрижки одного клиента составляет 20 мин. Клиенты в среднем приходят каждые 25 мин. Средняя стоимость стрижки составляет 350 руб. Как в первую смену с 9 до 15, так и во вторую – с 15 до 21, работают по одному мастеру. Провести анализ работы системы обслуживания. Определить ежедневный «чистый» доход каждого мастера, если он получает только 30% от выручки (остальное уходит на оплату аренды помещения, налоги, амортизацию оборудования и проч.).

Решение. Интенсивность входящего потока λ=2,4 клиента/ч, интенсивность потока обслуживаний μ=3 клиента/ч. Находим:

интенсивность нагрузки (канала) мастера $\rho = \lambda/\mu = 0,8$;

долю времени (вероятность) простоя мастера $p_0 = 1 - \rho = 1 - 0,8 = 0,2$;

вероятность того, что мастер занят работой $p_{зан} = 1 - p_0 = 1 - 0,2 = 0,8$;

среднее число клиентов в очереди $L_{оч} = \dfrac{0,8^2}{1-0,8} = 3,2$ клиента;

среднее время ожидания в очереди $\overline{T}_{оч} = \dfrac{3,2}{2,4} \approx 1,33$ мин;

среднее время пребывания клиентов в парикмахерской $\overline{T}_{СМО} = 1,34 + 20 = 21,34$ мин.

Система работает вполне удовлетворительно. Поскольку $\rho < 1$, то режим работы системы устойчивый, 20% рабочего времени мастер не занят, а остальные 80% времени занят работой, длина очереди 3,2 клиента небольшая, а среднее время пребывания клиента в парикмахерской всего 21,34 мин.

Каждый мастер занимается обслуживанием клиентов в среднем ежедневно в течение 0,8· (15 − 9) = 4,8 *ч* = 288 *мин*.

За это время он обслужит 288 20 =14,4 клиента, поэтому ежедневная выручка в среднем составит 14,4·350 = 5040 руб. Ежедневный «чистый» доход каждого мастера в среднем составляет 5040·0,3 = 1512 руб.

6.6. Многоканальная СМО с ограниченной очередью

Рассмотрим n-канальную СМО с ожиданием, на которую поступает поток заявок с интенсивностью λ; интенсивность обслуживания μ (для одного канала); число мест в очереди m.

Состояния системы нумеруются по числу заявок, связанных системой:

нет очереди:

S_0 – все каналы свободны;

S_1 – занят один канал, остальные свободны;

S_k – заняты k каналов, остальные нет;

S_n – заняты все n каналов, свободных нет;

есть очередь:

S_{n+1} – заняты все n каналов; одна заявка стоит в очереди;

S_{n+r} – заняты все n каналов, r заявок в очереди;

S_{n+m} – заняты все n каналов, m заявок в очереди.

Напишем выражения для ***предельных вероятностей*** состояний, используя обозначение $\rho = \frac{\lambda}{\mu}$:

$$p_0 = \left(1+\rho+\frac{\rho^2}{2!}+\ldots+\frac{\rho^n}{n!}+\frac{\rho^{n+1}}{n\cdot n!}+\frac{\rho^{n+2}}{n^2\cdot n!}+\ldots+\frac{\rho^{n+m}}{n^m\cdot n!}\right)^{-1} =$$

$$= \left(1+\rho+\frac{\rho^2}{2!}+\ldots+\frac{\rho^n}{n!}+\frac{\rho^n}{n!}\cdot\frac{\frac{\rho}{n}-\left(\frac{\rho}{n}\right)^{m+1}}{1-\frac{\rho}{n}}\right)^{-1},$$

$$p_k = \frac{\rho^k}{k!} p_0 \quad (k = 1, \ldots, n),$$

$$p_{n+i} = \frac{\rho^{n+i}}{n^i \cdot n!} p_0 \quad (i = 1, \ldots, m).$$

Определим характеристики эффективности системы:

вероятность отказа $P_{отк} = p_{n+m} = \dfrac{\rho^{n+m}}{n^m \cdot n!} p_0$;

относительная пропускная способность $Q = 1 - P_{отк} = 1 - \dfrac{\rho^{n+m}}{n^m \cdot n!} p_0$;

абсолютная пропускная способность СМО $A = \lambda Q = \lambda \left(1 - \dfrac{\rho^{n+m}}{n^m \cdot n!} p_0\right)$;

среднее число занятых каналов $\bar{k} = \dfrac{A}{\mu} = \rho \left(1 - \dfrac{\rho^{n+m}}{n^m \cdot n!} p_0\right)$;

среднее число заявок в очереди $L_{оч} = \dfrac{\rho^{n+1} p_0}{n \cdot n!} \dfrac{1 - \left(\frac{\rho}{n}\right)^m \left(m + 1 - m\frac{\rho}{n}\right)}{\left(1 - \frac{\rho}{n}\right)^2}$ **или**

$L_{оч} = 1 \cdot p_{n+1} + 2 \cdot p_{n+2} + \ldots + m \cdot p_{n+m}$;

среднее число заявок в системе $K = L_{оч} + \bar{k}$;

среднее время ожидания заявки в очереди

$$\bar{T}_{оч} = \frac{L_{оч}}{\lambda} = \frac{\rho^n p_0}{n \cdot \mu \cdot n!} \frac{1 - \left(\frac{\rho}{n}\right)^m \left(m + 1 - m\frac{\rho}{n}\right)}{\left(1 - \frac{\rho}{n}\right)^2};$$

среднее время пребывания заявки в системе $\bar{T}_{СМО} = \bar{T}_{оч} + \bar{T}_{об}$.

Пример 6.5. На станцию техобслуживания с двумя подъёмниками для текущего ремонта поступает простейший поток заявок с плотностью $\lambda = 1{,}5$ маш./час. Во дворе могут находиться, дожидаясь обслуживания не более 3-х машин. Среднее время ремонта $T_{обс} = 2$ час. Найти основные характеристики работы станции.

Решение. Имеем марковскую с параметрами: m = 2, n = 3, λ = 1,5 , μ = $1/T_{обс}$ = 1/2, значит $\rho = \lambda/\mu = 3$.

В результате расчетов, получим следующие характеристики системы:

вероятность отказа $P_{отк} = 0{,}374$;

относительная пропускная способность $Q = 1 - P_{отк} = 0{,}626$;

абсолютная пропускная способность СМО $A = \lambda Q = 0{,}938$;

среднее число занятых каналов $\overline{k} = \dfrac{A}{\mu} = 1{,}877$;

среднее число заявок в очереди $L_{оч} = 1{,}789$;

среднее число заявок в системе $K = L_{оч} + \overline{k} = 3{,}666$;

среднее время ожидания заявки в очереди $\overline{T}_{оч} = \dfrac{L_{оч}}{\lambda} = 1{,}193$ час. или 72 мин.;

среднее время пребывания заявки в системе $\overline{T}_{СМО} = \overline{T}_{оч} + \overline{T}_{об} = 3{,}193$ час. или 192 мин.

Таким образом, получаем, что примерно треть поступающих машин получают отказ в обслуживании. В этом случае руководству станции техобслживания можно либо увеличить количество подъемников, либо сократить время обслуживания, либо увеличить площадку для ожидающих машин.

Например, если увеличить количество подъемников до четырех, то вероятность отказа составит $P_{отк} = 0{,}064$ или 6,4%.

6.7. Многоканальная СМО с неограниченной очередью

Мы рассмотрели n-канальную СМО с ожиданием, когда в очереди одновременно могут находиться не более m заявок. Так же, как и ранее, при анализе систем без ограничений необходимо рассмотреть полученные соотношения при

$m \to \infty$. Будем считать, что $\rho/n < 1$, в противном случае система не является устойчивой (очередь неограниченно растет).

Предельные вероятности состояний

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n!(n-\rho)}\right)^{-1},$$

$$p_k = \frac{\rho^k}{k!} p_0 \quad (k = 1, \ldots, n),$$

$$p_{n+i} = \frac{\rho^{n+i}}{n^i \cdot n!} p_0 \quad (i = 1, \ldots, m, \ldots);$$

вероятность отказа $P_{отк} = 0$;

относительная пропускная способность $Q = 1 - P_{отк} = 1$;

абсолютная пропускная способность СМО $A = \lambda Q = \lambda$;

среднее число занятых каналов $\bar{k} = \dfrac{A}{\mu} = \rho$;

среднее число заявок в очереди $L_{оч} = \dfrac{\rho^{n+1} p_0}{n \cdot n!} \cdot \dfrac{1}{\left(1 - \dfrac{\rho}{n}\right)^2}$;

среднее число заявок в системе $K = \bar{k} + L_{оч}$;

среднее время ожидания заявки в очереди $\bar{T}_{оч} = \dfrac{L_{оч}}{\lambda} = \dfrac{\rho^n p_0}{n \cdot \mu \cdot n!} \dfrac{1}{\left(1 - \dfrac{\rho}{n}\right)^2}$;

среднее время пребывания заявки в системе $\bar{T}_{СМО} = \bar{T}_{оч} + \bar{T}_{об}$.

Пример 6.6. Автозаправочная станция с двумя колонками (n=2) обслуживает поток машин с интенсивностью λ=0,8 (машин в минуту). Среднее время обслуживания одной машины $\bar{T}_{об} = 2$ мин. В данном районе нет другой АЗС, так что очередь машин перед АЗС может расти практически неограниченно. Найти характеристики СМО.

Решение. Имеем $\mu = 1/\overline{T}_{об} = 0{,}5$ (машин в минуту), $\rho = \lambda / \mu = 0{,}8/0{,}5 = 1{,}6$, $\rho / n = 0{,}8 < 1$, поэтому очередь не растет безгранично и имеет смысл говорить о предельном стационарном режиме работы СМО.

Находим характеристики системы:

$$p_0 = \left(1 + 1{,}6 + \frac{1{,}6^2}{2!} + \frac{1{,}6^3}{2!(2 - 1{,}6)} \right)^{-1} = (1 + 1{,}6 + 1{,}28 + 5{,}12)^{-1} = \frac{1}{9} \approx 0{,}11;$$

вероятность отказа $P_{отк} = 0$;

относительная пропускная способность $Q = 1 - P_{отк} = 1$;

абсолютная пропускная способность СМО $A = \lambda = 0{,}8$ (машин в минуту);

среднее число занятых каналов $\overline{k} = \rho = 1{,}6$ (каналов);

среднее число машин в очереди $L_{оч} = \dfrac{1{,}6^3 \cdot 0{,}11}{2 \cdot 2!} \cdot \dfrac{1}{(1 - 0{,}8)^2} \approx 2{,}844$ (машин);

среднее число машин на АЗС $K = 2{,}84 + 1{,}6 = 4{,}444$ (машин);

среднее время ожидания заявки в очереди $\overline{T}_{оч} = \dfrac{2{,}84}{0{,}8} = 3{,}556$ (мин.);

среднее время пребывания заявки в системе $\overline{T}_{СМО} = 3{,}55 + 2 = 5{,}556$ (мин.).

Как видим, автозаправочная станция довольно эффективно – среднее время пребывания машины на станции не более 6 минут, при этом среднее число машин в очереди не более трех.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1 Азарнова, Т.В. Методы оптимизации: Учебное пособие / Т.В. Азарнова, И.Л. Каширина, Г.Д. Чернышова. – Воронеж: Изд-во ВГУ, 2003. – 86 с.

2 Алексеев, В.М. Сборник задач по оптимизации: Теория. Примеры. Задачи: Учеб. пособие / В.М. Алексеев, Э.М. Галлеев, В.М. Тихомиров. – 2-е изд. – М.: ФИМАЛИТ, 2005. – 256 с.

3 Амос, Г. MATLAB. Теория и практика [Электронный ресурс] / Г. Амос; пер. с англ. Н.К. Смоленцев. – Электрон. дан. – Москва: ДМК Пресс, 2016. – 416 с. – Режим доступа: https://e.lanbook.com/book/82814

4 Ануфриев, И.Е. MATLAB 7 / И.Е. Ануфриев, И.Е. Смирнов, Е.Н. Смирнова – СПб.: БХВ-Петербург, 2005. – 1104 с.

5 Аттетков, А.В. Методы оптимизации / А.В. Аттетков, С.В. Галкин, В.С. Зарубин . – М.: Изд-во МГТУ им. Н.Э. Баумана, 2003. – 440 с.

6 Бережная, Е.В. Математические методы моделирования экономических систем: учеб. пособие: рек. УМО вузов / Е.В. Бережная, В.И. Бережной. – М.: Финансы и статистика, 2001, 2002, 2003, 2005. – 368 с.

7 Блинов, Ю.Ф. Методы математического моделирования, Ч. 1 [Электронное учебное пособие] / Ю.Ф. Блинов, В.В. Иванцов, П.В. Серба. – Таганрог: ТТИ ЮФУ, 2012. – 42 с.

8 Васильев, А.Н. MATLAB. Самоучитель. Практический подход [Электронный ресурс]: самоучитель / А.Н. Васильев. – Электрон. дан. – Санкт-Петербург: Наука и Техника, 2015. – 448 с. - Режим доступа: https://e.lanbook.com/book/69619

9 Васильев, О.В. Методы оптимизации в задачах и упражнениях / О.В. Васильев, А.В. Аргучинцев. – М.: ФИЗМАТЛИТ, 1999. – 208 с.

10 Васин, А.А. Исследование операций: учеб. пособие: рек. НМС / А.А. Васин, П.С. Краснощеков, В.В. Морозов. – М.: Академия, 2008. – 464 с.

11 Вентцель, Е.С. Исследование операций: задачи, принципы, методология / Е.С. Вентцель. – М.: Высшая школа, 2001. – 208 с.

12 Волков, И.К. Исследование операций: Учеб для вузов / И.К. Волков, Е.А. Загоруйко, под. ред. В.С. Зарубина, А.П. Крищенко. – М.: Изд-во МГТУ им. М.Э. Баумана, 2000. – 436 с.

13 Дьяконов, В.П. MATLAB. Полный самоучитель / В.П. Дьяконов – М.: ДМК Пресс, 2012. – С. 42-768.

14 Исследование операций в экономике: учеб. пособие: рек. Мин. обр. РФ / под ред. Н.Ш. Кремера. – М.: Маркет ДС, 2007. – 408 с.

15 Карманов, В.Г. Математическое программирование: Учеб. пособие / В.Г. Карманов. – 4-е изд., перераб. и доп. - М.: ФИМАЛИТ, 2000. – 264 с.

16 Климов, Г.П. Теория массового обслуживания [Электронный ресурс]: учебное пособие / Г. П. Климов. – Электрон. текстовые данные. – М.: Московский государственный университет имени М.В. Ломоносова, 2011. – 235 с. — Режим доступа: http://www.iprbookshop.ru/13316.html

17 Кубланов, М.С. Математическое моделирование. Методология и методы разработки математических моделей механических систем и процессов: Учебное пособие. Часть I. Третье издание / М.С. Кубланов. – М.: МГТУ ГА, 2004. – 108 с.

18 Максимова, Н.Н. Исследование операций. Модели линейного программирования: учебно-методическое пособие / Н.Н. Максимова. – Благовещенск: Изд-во АмГУ, 2016. – 144 с.

19 Морозов, В.В. Исследование операций в задачах и упражнениях: учеб. пособие: доп. Мин. обр. РФ / В.В. Морозов, А.Г. Сухарев, В.В. Федоров. – 2-е изд., испр. – М.: Либроком, 2009. – 287 с.

20 Мызникова, Т.А. Модели социально-эколого-экономических систем: Учебно-методическое пособие / Т.А. Мызникова. – Омск: Изд-во СибАДИ, 2008. – 62 с.

21 Пантелеев, А.В. Методы оптимизации в примерах и задачах / А.В. Пантелеев, Т.А. Летова. – М.: Высш.шк., 2002. – 544 с.

22 Самарский, А.А. Математическое моделирование: Идеи. Методы. Примеры. 2-е изд., испр. / А.А. Самарский, А.П. Михайлов. – М.: ФИЗМАТЛИТ, 2002. – 320 с.

23 Самусевич, Г.А. Основы теории массового обслуживания [Электронный ресурс]: практикум / Г.А. Самусевич. – Электрон. текстовые данные. – Екатеринбург: Уральский федеральный университет, ЭБС АСВ, 2014. – 44 с. — Режим доступа: http://www.iprbookshop.ru/68270.html

24 Стариков, А.В. Экономико-математическое и компьютерное моделирование: учеб. пособие / А.В. Стариков, И.С. Кущева. – Воронеж: ГОУ ВПО «ВГЛТА». – Воронеж, 2008. – 132 с.

25 Тарасова, Н.В. Системы массового обслуживания [Электронный ресурс]: методические указания к семинарским занятиям по дисциплине «Сервисная деятельность» / сост. Н. В. Тарасова. – Электрон. текстовые данные. – Липецк: Липецкий государственный технический университет, ЭБС АСВ, 2012. –24с.— Режим доступа: http://www.iprbookshop.ru/17695.html

26 Таха, Х.А. Введение в исследование операций, 7-е издание.: Пер. с англ. – М.: Издательский дом «Вильямс», 2005. – 912 с.

27 Фомин, Г.П. Математические методы и модели в коммерческой деятельности: учебник: рек. Мин. обр. РФ / Г.П. Фомин. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2005. – 616 с.

28 Шапошников, А.В. Теория систем массового обслуживания [Электронный ресурс]: учебное пособие / сост. А. В. Шапошников [и др.]. – Электрон. текстовые данные. – Ставрополь: Северо-Кавказский федеральный университет,2017.–134с.Режим доступа: http://www.iprbookshop.ru/75605.html

СОДЕРЖАНИЕ

Составитель

Надежда Николаевна Максимова,

доцент кафедры математического анализа и моделирования,

кандидат физико-математических наук

Математическое моделирование. Учебно-методическое пособие

Изд-во АмГУ. Подписано к печати __.__.2019. Формат 60x84/16.

Усл. печ. л. 5,11.

Тираж __. Заказ __.

Отпечатано в типографии АмГУ.