Федеральное агентство по образованию

Государственное образовательное учреждение
высшего профессионального образования

Владимирский государственный университет

А.А. ГАЛКИН

МАТЕМАТИЧЕСКАЯ ЭКОНОМИКА

*Допущено Министерством образования и науки
Российской Федерации в качестве учебника
для студентов высших учебных заведений, обучающихся
по специальности «Прикладная информатика (в экономике)»*

Владимир 2006

УДК 330.45 : 519.85
ББК 65 В 631
Г16

Рецензенты:

Доктор технических наук, профессор зав. кафедрой
автоматизированных информационных и управляющих систем
Тульского государственного университета
В.А. Фатуев

Доктор технических наук, профессор
зав. кафедрой информационных систем
Тверского государственного технического университета
Б.В. Палюх

Доктор экономических наук, профессор
зав. кафедрой экономики и управления на предприятиях
Владимирского государственного университета
В.Ф. Архипова

Доктор физико-математических наук, профессор
зав. кафедрой алгебры и геометрии
Владимирского государственного университета
Н.И. Дубровин

Печатается по решению редакционно-издательского совета
Владимирского государственного университета

Галкин, А. А.

Г16 Математическая экономика : учебник / А. А. Галкин ; Владим. гос. ун-т. – Владимир : Изд-во Владим. гос. ун-та, 2006. – 304 с. – ISBN 5-89368-624-1.

Рассматривается широкий круг типовых оптимизационных задач, возникающих в экономике, и алгоритмов, позволяющих решать эти задачи. Даны методика формализации указанных задач и их классификация. Представлены методы решения детерминированных задач статической и динамической оптимизации. По каждому типу задач и алгоритмов приведены примеры, демонстрирующие технику практического использования этих алгоритмов, а также набор задач для самостоятельного решения.

Предназначен для студентов вузов, обучающихся по специальности 080801 – прикладная информатика (в экономике), а также студентов, магистрантов и аспирантов смежных специальностей очного, заочного обучения, лиц, получающих второе высшее образование, а также специалистов-практиков.

Табл. 80. Ил. 60. Библиогр.: 39 назв.

УДК 330.45 : 519.85
ББК 65 В 631

ISBN 5-89368-624-1

О Г Л А В Л Е Н И Е

Список принятых сокращений

ЦФ – целевая функция
ОДР – область допустимых решений
ЛП – линейное программирование
ЗЛП – задача ЛП
КЗЛП – каноническая ЗЛП
ТЗ – транспортная задача
ПО – пункты отправления, ПН – пункты назначения в ТЗ
МСЗУ – метод северо-западного угла
МЗС – метод золотого сечения
ДП – динамическое программирование
ВИ – вариационное исчисление
ПМ – принцип максимума;
ДУ – дифференциальное уравнение

ПРЕДИСЛОВИЕ

В подготовке студентов различных технических и экономических специальностей и направлений значительное место занимает изучение типичных для соответствующей предметной области математических моделей и методов, позволяющих, оперируя этими моделями, объяснять поведение рассматриваемых систем, оценивать их характеристики, обоснованно принимать конструктивные, технологические, экономические, организационные и другие решения.

Освоение этих моделей и методов основывается на фундаменте, заложенном в довольно универсальной классической дисциплине, обычно называемой «Высшая математика». Математический аппарат, позволяющий решать типовые и наиболее важные для соответствующей сферы приложений задачи, изучается в специальных дисциплинах.

Для студентов, обучающихся по специальности «Прикладная информатика (в экономике)», одной из таких дисциплин является «Математическая экономика». В соответствии с действующим государственным образовательным стандартом (ГОС) в программу этой дисциплины включен большой объем учебного материала, связанного с проведением математических расчетов в сфере экономики. Этот материал делится на две части.

В первой части изучаются задачи финансового анализа, которые в ГОС предшествующего поколения рассматривались в специальной дисциплине – «Финансовая математика».

Вторая часть программы содержит с точки зрения математики более сложные задачи и методы, связанные с отысканием наилучших, т.е. оптимальных, решений различных задач, встречающихся в области прикладной экономики. Ранее студенты осваивали этот материал при изучении дисциплины «Теория оптимального управления в экономических системах».

Учебная программа дисциплины «Математическая экономика» содержит широкий спектр довольно сложных для изучения вопросов. Поскольку объем времени, выделенного для аудиторных занятий по этой дисциплине, довольно небольшой, особое значение приобретает самостоятельная работа студентов с учебной литературой.

Следует отметить, что за последние 30 лет в нашей стране было издано много различных монографий, учебников и учебных пособий по математическим методам, применяемым в экономике. Однако при работе с ними у студентов возникают серьезные затруднения. Во-первых, многие из этих книг сейчас практически недоступны для студентов, так как либо отсутствуют в библиотеках вузов, либо имеются в единичных экземплярах. Во-вторых, для изучения всего предусмотренного программой материала одного учебника недостаточно, а в разных книгах, как правило, используются разный стиль изложения, разные обозначения. Нередко уровень изложения материала недоступен «реальному» студенту. В-третьих, при организации учебного процесса по дисциплинам математического характера принципиально важное значение имеет приобретение студентами практических навыков в использовании изучаемых методов, а для этого необходимы задачи для самостоятельного решения. Большинство учебных пособий по рассматриваемой тематике содержит примеры и задачи для иллюстрации техники применения излагаемых методов, но их недостаточно для того, чтобы выдать всем студентам обычной учебной группы индивидуальные задания.

Предлагаемый учебник предназначен для изучения второй, более сложной части дисциплины «Математическая экономика», в которой рассматриваются оптимизационные задачи, возникающие в экономике, и алгоритмы их решения. Он подготовлен с учетом изложенных выше обстоятельств.

В книге приведены формулировки типовых оптимизационных задач, возникающих в экономической сфере, осуществлена их формализация, изложена сущность методов и алгоритмов, позволяющих выполнять решение с иллюстрацией техники этих алгоритмов на конкретных примерах. Кроме того, по каждой теме представлен достаточно большой набор задач для самостоятельного решения, позволяющий каждому студенту дать свое индивидуальное задание.

Из огромного разнообразия возможных оптимизационных задач и предлагаемых современной наукой методов для включения в этот учебник выбраны детерминированные задачи и алгоритмы статической и динамической оптимизации. Из-за ограниченного объема книги задачи оптимизации с неопределенностями, в том числе вероятностно-статистические, интервальные, нечеткие и другие задачи и модели, а также задачи векторной оптимизации, не рассматриваются.

Книга включает девять глав. В первой даны примеры оптимизационных задач экономического характера, на которых продемонстрирована методика формализации, т.е. получения математической модели решаемой задачи, приведена классификация оптимизационных задач.

Главы вторая, третья и четвертая посвящены линейным задачам статической оптимизации. В второй главе изложены задачи и методы линейного программирования, отдельно в третьей – рассмотрены транспортные задачи, а в четвертой – оптимизационные задачи, которые интерпретируются на графах. Для каждой задачи представлен наиболее эффективный метод (алгоритм) решения и дан пример, демонстрирующий технику практического использования этого алгоритма. В пятой главе изложены аналитические и численные методы решения нелинейных задач статической оптимизации при отсутствии и наличии ограничений.

Динамические задачи оптимизации, обычно называемые задачами оптимального управления, рассмотрены в главах с шестой по девятую. В шестой главе дано общее представление о динамических системах непрерывного и дискретного типа, сформулирована классическая задача об оптимальном управлении и динамическом программировании (ДП), изложена сущность ДП и на различных примерах экономического характера показана техника его практического применения. В седьмой главе изложены основы вариационного исчисления, в восьмой – принцип максимума для непрерывных систем, а в девятой – для дискретных систем. В каждой из этих глав большое внимание уделено анализу различных частных задач и примеров, иллюстрирующих методику практического использования расчетных соотношений.

В конце каждой из глав с первой по шестую приведены задачи для самостоятельного решения. В конце девятой главы даны задачи для самостоятельного решения, посвященные методам оптимального динамического управления.

Особой проблемой, для решения которой автору в процессе работы над книгой потребовались значительные усилия, явилось то, что некоторые методы и алгоритмы в оригинальной литературе изложены так, что студентам нематематического, а информационно-экономического профиля разобраться в них довольно трудно. Поэтому необходимо было найти возможности для адаптации соответствующего теоретического материала к реальному уровню подготовки студентов, на которых ориентирована книга.

Кроме того, автор стремился при изложении большого количества существенно отличающихся задач и методов в максимальной степени выдержать единый стиль, характер, систему изложения материала. Хотелось бы надеяться, что это в определенной мере удалось осуществить.

При подготовке учебника был использован материал лекций и практических занятий по дисциплинам «Методы оптимизации», «Теория управления», «Теория оптимального управления в экономических системах» и «Математическая экономика», которые автор преподавал в течение 25 лет во Владимирском государственном университете (ВлГУ). На этих занятиях большая часть теоретического материала и задач для самостоятельного решения прошла апробацию. Электронная версия учебника включена в информационные ресурсы электронной библиотеки ВлГУ.

Несмотря на то что учебник подготовлен для студентов специальности «Прикладная информатика (в экономике)», несомненно, он может оказаться полезен студентам, магистрантам, аспирантам и специалистам других профилей, поскольку оптимизационные задачи возникают всюду. Не случайно говорят, что «в природе нет ничего, в чем нельзя было бы усмотреть смысл какого-либо максимума или минимума».

Автор выражает искреннюю признательность рецензентам за внимательное отношение к рукописи и ряд ценных советов и замечаний.

Он будет благодарен всем тем, кто воспользуется книгой и сообщит свое мнение о ее содержании, возможно, о недостатках или неточностях. Для этого можно воспользоваться e_mail: aagalkin@vpti.vladimir.ru.

Работа над книгой с некоторыми перерывами велась около 10 лет, но она могла затянуться на неопределенный срок, если бы не оперативная и высококвалифицированная помощь в работе над рукописью, которую оказала аспирант И.В. Лагерь. За это автор выражает ей особую благодарность.

ВВЕДЕНИЕ

Экономика традиционно является областью широкого практического применения математики. В первую очередь это касается бухгалтерского учета и финансового планирования, предусматривающих подготовку смет, калькуляций, сводных отчетов, балансов и т.д. Безусловно, это самая широкая сфера применения математических расчетов в экономике, хотя с точки зрения математики и сравнительно несложная, поскольку требует применения лишь самых простых математических операций – сложения, умножения, вычитания, деления, начисления процентов.

В 30-е годы XX века сформировалось и стало активно развиваться научное направление, связанное с применением в экономике более сложных математических методов, основанных на моделировании производственных, финансовых и других процессов и использовании соответствующих математических методов для наиболее рационального, т.е. оптимального использования имеющихся ресурсов.

Это направление основывалось на классических результатах математического анализа и вариационного исчисления, сформулированных Ньютоном, Эйлером, Лагранжем, и новых результатах, полученных при решении задач, возникших в сфере практической экономики – оптимального планирования производства, раскроя материалов, составления графиков замены оборудования, организации наиболее рациональных перевозок и т.д.

Следует отметить, что многие основополагающие работы в области современных экономико-математических методов были получены нашими соотечественниками, среди которых, прежде всего, следует назвать академика Л.В. Канторовича, чьи достижения были отмечены Нобелевской премией, и академика Л.С. Понтрягина, который за заслуги в теории оптимального управления был удостоен самой престижной в СССР Ленинской премии.

Теоретические исследования в области оптимального планирования и управления в экономике активно стимулировались как плановой экономикой СССР, так и западной рыночной экономикой.

Принципиально новые возможности для развития, а главное внедрения в практику экономико-математических методов открылись с появлением, стремительным совершенствованием и широким распространением вычислительной техники.

В последние годы использование математических методов и, в частности, методов оптимизации в экономике стало особенно актуальным по нескольким причинам:

– в экономике все чаще стали реализовываться крупные проекты, в которых задействованы значительные ресурсы, а это требует взвешенного, по возможности оптимального их использования;

– деятельность и больших, и малых предприятий любого профиля (производственных, финансовых, транспортных, страховых и т.д.) осуществляется в условиях жесткой конкуренции, в которой успеха добиваются лишь те, кто наиболее эффективно использует ресурсы;

– стала доступной практически для всех вычислительная техника, эксплуатационные характеристики которой (производительность, объем памяти и т.п.) 15 – 20 лет назад казались недостижимой фантастикой, эта техника дает возможность оперировать с очень большими объемами информации и реализовывать алгоритмы вычислений практически любой сложности.

Для внедрения математических методов и информационных технологий в практическую деятельность нужны специалисты, которые, с одной стороны, достаточно глубоко разбираются в сущности экономических проблем и способны формализовать возникающие задачи, а с другой, – профессионально владеют математическими методами и соответствующим программным обеспечением.

Подготовка таких профессионалов, осуществляемая по двум направлениям: экономическому и информационно-математическому, в вузах России ведется по специальностям «Прикладная информатика (в экономике)» и «Математические методы в экономике». Государственные образовательные стандарты этих специальностей предусматривают изучение студентами широкого спектра задач, математических моделей и методов, а также программного обеспечения и информационных технологий, востребованных современной экономикой.

Особо следует отметить, что реальные задачи, возникающие на практике, в том числе в экономической сфере, как правило, имеют большую размерность и их решение можно осуществить только с помощью компьютеров. При этом возможны два подхода:

– можно самостоятельно составить программу, выбрав соответствующий решаемой задаче алгоритм, тщательно изучив его и воспользовавшись тем или иным языком программирования;

– можно использовать готовые программы или пакеты прикладных программ, в том числе известные математические пакеты Matlab, MathCAD, Matematica и другое ПО (см., например [38]). Следует отметить, что некоторые возможности для решения оптимизационных задач заложены и в широко распространенную программу Excel, входящую в состав пакета Microsoft Office.

Может показаться, что во втором случае вовсе не обязательно знать «механизм», заключенный в программу, и можно обращаться с ней как с «черным ящиком» – достаточно ввести свои исходные данные и получить искомый результат. Однако опыт – «сын ошибок трудных» - показывает, что это не так. В [32] отмечается: «мнение о том, что исследователю, использующему методы оптимизации, нет необходимости изучать соответствующий математический аппарат потому, что «обеспечить» решение задачи может компьютер, – опасное заблуждение. Следует всегда помнить, что ЭВМ решает такую задачу, которую сформулировал разработчик и так, как он предписал ей это делать в выбранной программе. Если исследователь не осведомлен об ограниченных возможностях использованной модели или выбранного метода, то это отразится на качестве решения».

Поэтому специалист, применяющий математические методы для решения прикладных, в том числе экономических, задач с помощью вычислительной техники и тех или иных программных средств, в любом случае должен знать сущность используемых алгоритмов, ограничения на их применение, возможные проблемы и «подводные камни», которые могут возникнуть при их реализации. Хотя, разумеется, во втором случае, когда используются готовые программные продукты, специалист освобождается от выполнения большого объема высококвалифицированной работы по программированию, поэтому этот путь, как правило, является предпочтительным.

Таким образом, использование компьютерных программ не освобождает специалиста от необходимости тщательного изучения алгоритмов решения соответствующих задач, в том числе от рутинного анализа примеров и самостоятельного решения вручную тестовых задач небольшой размерности.

Данный учебник дает студентам теоретический и практический материал, необходимый для изучения сущности оптимизационных задач и алгоритмов, понимания техники их практического применения. После изучения этого материала открываются возможности для разработки собственных компьютерных прикладных программ или грамотного использования готовых программных продуктов при решении реальных практических задач оптимизационного характера.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО РАБОТЕ С УЧЕБНИКОМ

Для наиболее эффективного усвоения материала, представленного в учебнике, рекомендуется следующая методика работы с ним.

1. При изучении каждой темы, освещаемой в соответствующем разделе учебника, сначала нужно внимательно прочитать теоретическую часть этого раздела, в которой излагаются сущность рассматриваемого класса задач и методы (алгоритмы) их решения. Следует отметить, что большинство задач довольно просто формулируется и их усвоение, как правило, не вызывает трудностей.

2. С методами решения ситуация иная. Многие из них обычно не сразу становятся понятными для студентов, хотя и предполагают использование довольно простых математических операций.

В этих случаях следует после прочтения методики решения задачи перейти к разбору демонстрационного примера, который приведен после изложения общего алгоритма решения рассмотренного типа задач. Реализуемые при этом действия нужно соотносить с операциями, изложенными в общей методике.

Как показывает опыт, обычно после тщательного изучения иллюстративного примера сущность алгоритма становится понятной.

3. Для закрепления изученного материала и приобретения навыков и умений следует самостоятельно решить указанную преподавателем задачу из списка задач небольшой размерности, помещенного в конце соответствующего раздела.

4. Завершающим этапом в освоении материала является разработка и отладка компьютерной программы, реализующей рассмотренный алгоритм и позволяющей решать задачи соответствующего типа произвольной размерности для вводимых исходных данных.

В качестве задачи, используемой при отладке программы, можно взять ту же задачу небольшой размерности, которая была решена «вручную».

5. Создав такую программу, рекомендуется поэкспериментировать с ней, варьируя исходные данные. Например, реализовав тот или иной алгоритм поиска экстремума нелинейной функции, интересно исследовать процессы поиска экстремума для разных целевых функций, обладающих той или иной особенностью, разных начальных условий.

Можно сопоставить работу алгоритма, реализованного в Вашей программе, с работой другого алгоритма, использованного Вашим коллегой, на одних и тех же задачах. В каждом конкретном случае Вы можете проявить творческую инициативу.

6. При изложении теоретического материала по методам оптимизации в данном учебнике предполагалось, что студент владеет необходимыми знаниями в пределах традиционного вузовского курса математики. Приступая к работе с учебником, рекомендуется повторить следующие разделы этого курса:

1. Линейная алгебра, матрицы и операции с ними.
2. Математический анализ (дифференциальное исчисление и исследование функций на экстремум).
3. Обыкновенные дифференциальные уравнения.

Для повторения указанного материала можно обратиться к литературе [17, 19, 20].

Глава 1. **ПОСТАНОВКА, ФОРМАЛИЗАЦИЯ И КЛАССИФИКАЦИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ В ЭКОНОМИЧЕСКИХ СИСТЕМАХ**

§ 1.1. Содержательная постановка задач оптимизации и их формализация

Рассмотрим некоторую практическую оптимизационную задачу. Обычно в исходном виде ее суть раскрывается в словесной (содержательной) форме. В ней описывается данная система, указываются изменяемые и подлежащие выбору факторы, задается показатель, использующийся для оценки (сравнения) различных вариантов функционирования системы, а также все условия, признаваемые как существенные в этой задаче.

Чтобы использовать математические методы при решении рассматриваемой задачи, нужно от подобной содержательной постановки перейти к математической (формализованной) модели этой задачи. Для этого необходимо:

1) ввести обозначения для всех величин, фигурирующих в задаче;

2) с их помощью записать в виде математических соотношений (функций, уравнений, неравенств) критерий оптимальности и все заданные условия.

Пример 1. Содержательная формулировка задачи. Из листа металла необходимо изготовить цилиндрическую емкость. Она должна иметь заданный объем. Нужно выбрать геометрические размеры емкости такими, чтобы количество израсходованного металла на ее изготовление было минимальным.

Формализация задачи. Введем обозначения: R и H – радиус основания и высота емкости; S – площадь полной поверхности емкости; V_0 – заданный объем. Запишем с их помощью соотношения, определяющие объем и площадь поверхности цилиндра:

$$S = 2\pi R^2 + 2\pi RH\,; \qquad (1.1)$$

$$V_0 = \pi R^2 H\,. \qquad (1.2)$$

Тогда поставленная задача получает следующую математическую формулировку: найти численные значения переменных $R \geq 0, H \geq 0,$ при которых величина S, определяемая функцией (1.1), принимала бы наименьшее значение, и при этом выполнялось бы равенство (1.2).

Пример 2. Содержательная постановка задачи. Проектируется распределенная информационная система, позволяющая собирать в диспетчерский пункт (ДП) информацию от заданных объектов, расположение которых на местности фиксировано и известно. Требуется найти такое расположение ДП, при котором общая длина проводной линии, связывающей объекты с ДП, была бы минимальной.

Формализация задачи. Введем систему координат $x0y$ (рис. 1.1).

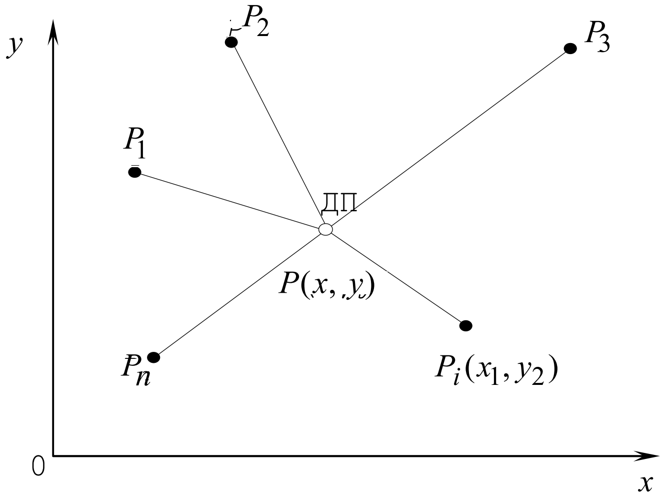


Рис. 1.1

Обозначим координаты каждого из n заданных объектов P_i через (x_i, y_i), $(i=\overline{1,n})$, а координаты ДП P через (x, y). При этом расстояние между P и P_i будет равно:

$$l_i = \sqrt{(x_i - x)^2 + (y_i - y)^2}.$$

Тогда общая длина линий связи рассматриваемой системы (см. рис. 1.1) будет определяться выражением

$$L = \sum_{i=1}^{n} \sqrt{(x_i - x)^2 + (y_i - y)^2}\,. \tag{1.3}$$

Таким образом, необходимо найти значения переменных x и y, при которых величина L, определяемая (1.3), была бы минимальна.

Пример 3. Содержательная постановка задачи. Оборудование предприятия позволяет выпускать два вида продукции. На ее изготовление идет пять видов сырья, запасы которых ограничены и заданы. Расход сырья на выпуск единицы продукции каждого вида, а также прибыль от ее реализации, известны. Требуется найти такой план выпуска продукции каждого вида, при котором общая прибыль от реализации всей продукции была бы максимальна.

Формализация задачи. Обозначим виды продукции P_1 и P_2, сырья – $S_1,\ldots, S_5$, а его запасы – $Z_1,\ldots, Z_5$. Доход от реализации единицы продукции P_1 обозначим $d_1 = 7$, а для P_2 соответственно $d_2 = 4$. Данные об удельных расходах сырья и его запасах сведем в табл. 1.1.

Т а б л и ц а 1.1

S_i \ P_i	P_1	P_2	Z_j
S_1	5	2	19
S_2	3	7	27
S_3	1	8	14
S_4	4	0	25
S_5	9	5	31

Допустим, что планом предусматривается выпуск x_1 изделий вида P_1 и x_2 – вида P_2 . Тогда общая прибыль y от реализации продукции составит

$$y = 7x_1 + 4x_2 .$$

Нужно найти такие значения $x_1 \geq 0$ и $x_2 \geq 0$, при которых величина y была бы наибольшей. При этом должна учитываться ограниченность ресурсов, которую в соответствии с принятыми обозначениями и данными табл. 1.1 можно записать в виде:

$$\begin{aligned} &5x_1 + 2x_2 \leq 19; \quad 4x_1 \leq 25; \\ &3x_1 + 7x_2 \leq 27; \quad 9x_1 + 5x_2 \leq 31; \\ &x_1 + 8x_2 \leq 14. \end{aligned} \tag{1.4}$$

Таким образом, задача свелась к отысканию неотрицательных значений переменных x_1 и x_2, при которых функция y принимала бы наибольшее значение при выполнении системы неравенств (1.4).

Пример 4. Содержательная постановка задачи. Рассматриваются искусственный спутник Земли и задача о переориентации его в плоскости путем управления тягой пары реактивных двигателей РД (рис. 1.2).

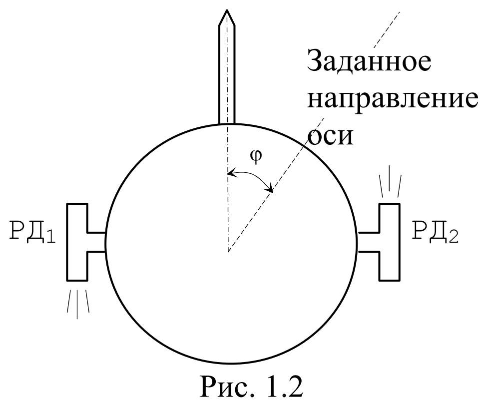


Рис. 1.2

Требуется найти закон изменения тяги двигателей, при котором спутник из произвольного начального направления будет переведен в заданное направление за минимальное время [31].

Формализация задачи. Обозначим через φ текущее значение угла отклонения оси спутника от заданного направления (см. рис. 1.2); M – тяга, развиваемая двигателями и ограниченная величиной M_0; J – момент инерции спутника. Тогда связь между управляющим воздействием и управляемой величиной можно описать уравнением [31]:

$$J\varphi''(t) = M. \tag{1.5}$$

Нужно найти такой закон изменения величины M:

$$M = M(t); \quad 0 \le t \le t_К; \quad |M(t)| \le M_0, \qquad (1.6)$$

при котором функция $\varphi(t)$, входящая в уравнение (1.5), удовлетворяла бы условиям: $\varphi(0) = \varphi_Н$; $\varphi'(0) = \varphi'_Н; \varphi(t_К) = 0$; $\varphi'(t_К) = 0$; и при этом критерий оптимальности, определяющий время, затраченное на управление $T_У = t_К$, принимал бы наименьшее значение.

§ 1.2. Классификация задач оптимизации

Многообразие оптимизационных задач можно разбить на следующие основные типы.

Задача, в которой необходимо найти численные значения варьируемых факторов (переменных) x_i, при которых критерий, являющийся заданной функцией этих факторов, принимает наименьшее (или наибольшее) значение, называется *задачей статической оптимизации*. Примерами таких задач служат задачи 1 – 3, рассмотренные в предыдущем параграфе.

Задачу, в которой нужно найти закон изменения варьируемых факторов, т.е. функции $x_i = x_i(t)$, при которых критерий, являющийся заданным функционалом от этих функций $x_i(t)$, принимает наименьшее (или наибольшее) значение, называют *задачей динамической оптимизации*. К этой группе задач относится, например, изложенная выше задача об оптимальном по быстродействию управлении спутником (см. пример 4).

Следует подчеркнуть, что методам динамической оптимизации посвящено специальное научное направление, называемое обычно теорией оптимального управления.

Рассмотрим далее более подробно задачи статической оптимизации. Пусть задача имеет следующую математическую формулировку. Найти численные значения переменных: $x_1 = x_1^0, ..., x_n = x_n^0$, при которых заданная функция, называемая целевой,

$$y = f(x_1, ..., x_n) \qquad (1.7)$$

принимает наименьшее значение, и выполняются условия:

$$\left.\begin{aligned} h_1(x_1, ..., x_n) = 0, \\ h_m(x_1, ..., x_n) = 0; \end{aligned}\right\} \qquad (1.8)$$

$$\left.\begin{aligned} g_i(x_1, ..., x_n) \ge 0, \\ g_k(x_1, ..., x_n) \ge 0. \end{aligned}\right\} \qquad (1.9)$$

Если ввести в рассмотрение матрицы-столбцы (векторы)

$$\bar{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}; \quad \bar{h} = \begin{bmatrix} h_1 \\ \dots \\ h_m \end{bmatrix}; \quad \bar{g} = \begin{bmatrix} g_1 \\ \dots \\ g_k \end{bmatrix},$$

то образующие задачу оптимизации компоненты (1.7 – 1.9) можно записать в краткой векторно-матричной форме:

$$y = f(\bar{x}); \quad \bar{h}(\bar{x}) = 0; \quad \bar{g}(\bar{x}) \geq 0.$$

Таким образом, статическую задачу оптимизации образуют в общем случае целевая функция (ЦФ) (1.7), определяющая критерий оптимальности, система ограничений-равенств (1.8) и ограничений-неравенств (1.9). В частных случаях ограничения равенства или неравенства могут отсутствовать (см. примеры 1 – 3 из § 1.1).

Изложенная формулировка задачи статической оптимизации является классической, хотя и не охватывает всех возможных задач подобного рода. К такой формулировке можно свести большое количество прикладных задач, возникающих на практике.

Входящие в выражения (1.7) – (1.9) функции f, h_i, g_i далее всюду будут предполагаться непрерывными, а в некоторых случаях и дифференцируемыми. Отметим, что ограничения (1.8) – (1.9) определяют в пространстве R^n переменных x_i некоторое подмножество *D*. Если оно представляет собой замкнутую ограниченную область, называемую областью допустимых решений (ОДР), то при непрерывности функции *f* сформулированная задача оптимизации имеет решение в соответствии с теоремой Вейерштрасса [17].

В зависимости от наличия и характера ограничений (1.8) и (1.9), а также вида ЦФ (1.7) различают следующие основные классы задач статической оптимизации.

1. Задача об отыскании переменных x_i, при которых целевая функция принимает наименьшее значение при отсутствии ограничений на переменные вида (1.8) и (1.9). Ее решение, если оно существует, совпадает с одним из локальных экстремумов ЦФ. Так, наименьшее значение функции будет совпадать с наименьшим из локальных минимумов, называемым глобальным минимумом. В связи с указанными обстоятельствами статическую задачу оптимизации при отсутствии ограничений называют ***задачей на безусловный экстремум*** [26].

2. Задача оптимизации, в которой наряду с целевой функцией задано одно или несколько ограничений-равенств вида (1.8), называется ***задачей на условный экстремум*** [26].

3. Задача, в которой наряду с целевой функцией задано хотя бы одно ограничение-неравенство (независимо от наличия или отсутствия ограничений-равенств), называется ***задачей математического программирования*** [14, 15].

Задачи математического программирования разбиваются на следующие основные типы.

3.1. ***Задача линейного программирования***. Все функции f, h_i, g_i. определяющие критерий оптимальности (1.7) и ограничения (1.8), (1.9), линейны (см. пример 3 из §1.1).

3.2. ***Задача выпуклого программирования***. Целевая функция является выпуклой (или вогнутой), все функции $g_i(\bar{x})$, определяющие ограничения-неравенства в форме (1.9), – вогнутыми, а функции $h_i(\bar{x})$, определяющие ограничения-равенства (1.8) (если они имеются), – линейными [26, 37].

3.3. ***Задача квадратичного программирования***. ЦФ представляет собой сумму линейной и выпуклой (или вогнутой) квадратичной формы:

$$f(x_1,...,x_n) = \sum_{i=1}^{n} d_i x_i + \sum_{j=i}^{n}\sum_{k=1}^{n} c_{jk} x_j x_k,$$

и все ограничения линейны [33, 37].

3.4. ***Задача геометрического программирования***. Как целевая функция, так и функции, входящие в ограничения, имеют следующий вид:

$$\alpha(\bar{x}) = \sum_{j=1}^{l} c_j z_j, \tag{1.10}$$

где $z_j = \prod_{i=1}^{n} x_i^{\gamma_{ij}}$, причем в (1.10) коэффициенты c_j положительны, а показатели степени γ_{ij} могут быть как положительными, так и отрицательными. Функции вида (1.10) называют позиномами [12, 33].

3.5. ***Задача сепарабельного программирования***. Как целевая функция, так и функции, входящие в ограничения, являются сепарабельными, т.е. имеют вид [33]:

$$\beta(\bar{x}) = \beta(x_1,...,x_n) = \beta_1(x_1) + \beta_2(x_2) + ... + \beta_n(x_n), \tag{1.11}$$

например, $\beta = 4x_1^2 + 7x_2^{-2x_2} + 1{,}8x_3$.

3.6. ***Задача целочисленного программирования***. В отличие от рассмотренных выше задач в ней оговаривается, что все или некоторые из переменных x_i могут принимать только целые значения. Это дополнительное требование, как правило, обусловлено физической неделимостью объектов, характеризуемых этими переменными.

В рассмотренных задачах был задан или выбран один показатель, оценивающий качество (эффективность) системы или решения. Такие задачи называют ***однокритериальными.*** Существует немало прикладных задач, в которых было бы желательно добиться экстремального значения двух или более показателей, например, получить наибольшую прибыль при наименьших затратах ресурсов. Такие задачи называют ***многокритериальными*** или ***задачами векторной оптимизации***. Довольно часто их сводят к однокритериальным задачам, выбрав один критерий в качестве оптимального (например прибыль), а на остальные показатели накладывают ограничения – «не более чем…» или «не менее чем…». Однако существуют и специальные подходы к анализу задач векторной оптимизации (см. например [27]).

До сих пор нами предполагалось, что все исходные данные, фигурирующие в задаче, например, нормы расхода материалов, удельная доходность от реализации продукции и т.д., однозначно определены и известны. Эти задачи называют ***детерминированными***. Однако на практике такие ситуации встречаются далеко не всегда. Зачастую в формулировке задачи подобной четкой определенности нет. Многие исходные условия оказываются неопределенными, например, спрос на продукцию, которая запланирована к выпуску. Такие задачи называются ***задачами с неопределенностями***. При их формализации и анализе используют специальные подходы, в частности вероятностно-статистические, интервальные и нечеткие модели [9, 11, 18].

В заключение отметим, что в теории оптимизации отсутствует какой-либо универсальный метод, позволяющий одинаково эффективно решать различные типы оптимизационных задач. Развитие этой теории идет по пути разработки алгоритмов, обеспечивающих достаточно эффективное решение определенного класса задач, например, задач линейного программирования. Поэтому при решении конкретной оптимизационной задачи важно, прежде всего, правильно формализовать эту задачу и определить ее тип, что позволит далее перейти к выбору соответствующего алгоритма решения.

Контрольные вопросы

1. Какие действия необходимо осуществить для перехода от содержательной постановки задачи к математической модели?
2. В чем состоит задача статической оптимизации, каковы особенности задач динамической оптимизации?
3. Что такое целевая функция, какую роль она играет в оптимизационной задаче?
4. Как формулируется задача безусловной оптимизации?
5. Как формулируется задача условной оптимизации?
6. Как формулируется задача математического программирования?
7. Какой смысл имеет термин «программирование» в теории оптимизации?
8. В чем состоит сущность задач векторной оптимизации?
9. Какими особенностями обладают детерминированные задачи оптимизации?
10. Какие подходы используются при формализации и анализе оптимизационных задач с неопределенностями?

Задачи для самостоятельного решения

Требуется формализовать приведенные ниже задачи и определить их тип.

1. Найти размеры прямоугольника, имеющего заданный периметр и максимально возможную площадь.

2. Из всех треугольников данного периметра p найти тот, который имеет наибольшую площадь.

3. Из проволоки заданной длины l необходимо изготовить равносторонний треугольник и квадрат. Требуется найти такие размеры этих фигур, чтобы их суммарная площадь была максимальна.

4. Малое предприятие изготавливает два вида красок: для внутренних K_1 и наружных K_2 работ. Продукция обоих видов поступает в оптовую продажу. Для производства этих красок используются два исходных продукта $П_1$ и $П_2$, их максимально возможные суточные запасы со-

ставляют 6 и 8 т соответственно. Расходы продуктов $П_1$ и $П_2$ на изготовление 1 т соответствующих красок известны. Они приведены в табл. 1.2. Изучение рынка сбыта показало, что суточный спрос на краску K_1 никогда не превышает спрос на краску K_2 более чем на 1 т. Кроме того, спрос на K_1 никогда не превышает 2 т в сутки. Оптовые цены одной тонны красок равны: 3 тыс. у.е. для K_1 и 2 тыс. у.е. для K_2. Какое количество краски каждого вида должно производить предприятие, чтобы доход от реализации обеих красок был наибольшим? (Решение аналогично примеру 3 из § 1.1, оно приведено в [32]).

Таблица 1.2

Исходные продукты / Виды продукции	Расход исходных продуктов, на изготовление 1 т краски		Максимально возможный запас, т
	K_1	K_2	
$П_1$	1	2	6
$П_2$	2	1	8

5. Производственная мощность цеха сборки некоторого изделия составляет 120 шт. типа P_1 и 360 шт. типа P_2 в смену. Технический контроль может пропустить в сутки не более 200 изделий того и другого типа. Доход от реализации изделий P_2 в 4 раза выше, чем от реализации P_1. Определить план выпуска изделий, при котором будет обеспечена наибольшая прибыль.

6. Предприятие выпускает две модели электронных блоков $Б_1$ и $Б_2$, причем каждая модель производится на отдельной технологической линии. Первая линия может выпускать 60 изделий в сутки, вторая – 75. В производстве обеих моделей используются однотипные узлы, суточный запас которых ограничен – он не может превышать 1000 штук. Причем на один блок 1-й модели расходуется 10 узлов указанного типа, а для 2-й модели – 8 узлов. Прибыль от реализации одного блока 1-й модели равна 30 у.е., а второй – 20 у.е. Определить оптимальные суточные объемы производства.

7. Фирма имеет возможность рекламировать свою продукцию через радио- и телевизионную сеть. Для этих целей из бюджета фирмы выделяется 1000 у.е. в месяц. Каждая минута радиорекламы стоит 5 у.е., а телерекламы – 100 у.е. Опыт показывает, что объем сбыта, который обеспечивает каждая минута телерекламы, в 25 раз больше сбыта, обеспечиваемого

одной минутой радиорекламы. Определить оптимальное распределение средств между радио- и телерекламой.

8. На предприятии изготавливаются изделия двух видов. Процесс изготовления каждого из них осуществляется последовательно в трех цехах. Время обработки и прибыль от продажи одного изделия каждого вида приведены в табл. 1.3. Рабочее время составляет 8 ч в сутки. Найти оптимальные объемы производства изделий каждого вида.

Т а б л и ц а 1.3

Тип изделия	Время обработки одного изделия, мин			Удельная прибыль, у.е.
	в цехе № 1	в цехе № 2	в цехе № 3	
1-й	10	6	8	2
2-й	5	20	15	3

9. Формализовать задачу об оптимальном составе смеси. В ней требуется определить такие концентрации веществ в смеси, чтобы стоимость смеси была минимальна, а показатели, оценивающие ее свойства (плотность, прочность, теплоемкость и т.п.), были не меньше заданных значений. Примером такой задачи может служить задача об оптимальной диете (рационе). При откормке животных каждое из них ежедневно должно получить не менее 60 ед. питательного вещества A, не менее 50 ед. вещества В и не менее 12 ед. вещества C. Использоваться могут три вида корма. Содержание питательных веществ в каждом из них задается в табл. 1.4.

Составить дневной рацион, обеспечивающий получение необходимого количества питательных веществ при минимальных денежных затратах, если цена кормов 1-го, 2-го и 3-го видов составляет соответственно 9, 12 и 10 у.е.

Т а б л и ц а 1.4

Питательные вещества	Количество питательных веществ в 1 кг корма		
	1-го вида	2-го вида	3-го вида
A	1	3	4
B	2	4	2
C	1	4	3

10. Для производства чугунного литья используется n различных исходных шихтовых материалов (чугун различных марок, стальной лом,

феррофосфор и др.). Химический состав чугунного литья определяется содержанием в нем химических элементов (кремния, марганца, фосфора и др.). Готовый чугун должен иметь строго определенный химический состав, который задается величинами H_j, представляющими собой доли (%) j-го химического элемента в готовом продукте. При этом известны величины: h_{ij} – содержание (%) j-го химического элемента в i-м исходном шихтовом материале; c_i – цена единицы каждого i-го шихтового материала. Определить состав шихты, обеспечивающей получение литья заданного качества при минимальной общей стоимости используемых шихтовых материалов.

11. Имеются заготовки в виде листов материала определенного размера. Из них нужно накроить детали четырех видов: А – 6 шт., Б – 15 шт., В – 25 шт., Г – 10 шт. Известны три способа раскроя заготовки. Количество деталей каждого вида, получаемых при этих способах раскроя, приведено в табл. 1.5.

Таблица 1.5

Способ раскроя	Количество получаемых деталей из одной заготовки			
	А	Б	В	Г
1-й	3	2	1	1
2-й	2	4	1	2
3-й	2	1	1	4

Найти оптимальный план раскроя заготовки, т.е. количество заготовок, при котором должно быть получено заданное количество деталей каждого вида из минимального количества заготовок.

12. Имеется n = 2 предприятий A_1 и A_2, в которых производится однородная продукция. Количество ежедневно производимой продукции в каждом из них задано: a_1 = 120 и a_2 = 90. Эту продукцию нужно доставить в m = 4 пунктов потребления: B_1, B_2, B_3 и B_4. Их ежедневные объемы потребления заданы: $в_1 = 40$, $в_2 = 20$, $в_3 = 80$, $в_4 = 70$. Требуется найти оптимальный план перевозок, определяющий количество продукции, перевозимой из пункта A_i в пункт B_j (i = 1,2; j = 1,..., 4). Этот план

должен обеспечивать минимальные общие затраты на перевозки. При этом стоимость c_{ij} перевозки единицы продукции из пункта A_i в пункт B_j задается табл. 1.6.

Т а б л и ц а 1.6

A_i \ B_j	B_1	B_2	B_3	B_4
A_1	2,3	4,7	1,8	1,1
A_2	3,7	4,6	8,2	12,3

Примечание. В рассматриваемой задаче объем производства равен объему потребления $\sum_i a_i = \sum_j в_j$, т.е. система «производство – потребление» сбалансирована. В противном случае продукция либо «затоварилась» бы в пунктах потребления, либо возникал дефицит в ней.

Пусть x_{ij} – количество единиц продукции, перевозимой из A_i в B_j. Тогда подлежащая минимизации общая стоимость перевозок будет определяться выражением

$$y = \sum_{i=1}^{2} \sum_{j=1}^{4} c_{ij} x_{ij} ,$$

где c_{ij} – стоимости из табл. 1.6.

$$y = 2{,}3x_{11} + 4{,}7x_{12} + 1{,}8x_{13} + 1{,}1x_{14} + 3{,}7x_{21} + 4{,}6x_{22} + 8{,}2x_{23} + 12{,}3x_{24}. \quad (1.12)$$

План должен быть составлен так, чтобы обеспечивался полный вывоз продукции из каждого пункта A_i и обеспечивалось полное удовлетворение спроса в каждом пункте потребления B_j. Это приводит к следующим двум группам условий: $\forall i : \sum_{j=1}^{4} x_{ij} = a_i ; \quad \forall j : \sum_{i=1}^{2} x_{ij} = b_j$:

$$\left.\begin{aligned} x_{11} + x_{12} + x_{13} + x_{14} = 120, \\ x_{21} + x_{22} + x_{23} + x_{24} = 90; \end{aligned}\right\} \quad \left.\begin{aligned} x_{11} + x_{21} = 40, \\ x_{12} + x_{22} = 20; \end{aligned}\right\} \quad \left.\begin{aligned} x_{13} + x_{23} = 80, \\ x_{14} + x_{24} = 70. \end{aligned}\right\} \quad (1.13)$$

Таким образом, данная задача сводится к отысканию nm переменных $x_{ij} \geq 0$, при которых выполняется система из $n+m$ ограничений-равенств (1.13), а целевая функция (1.12) принимает наименьшее значение. Эта задача является задачей линейного программирования, причем она имеет существенную особенность: все переменные x_{ij} входят в ограничения-равенства с единичными коэффициентами. Заметим, что в отличие от предыдущих задач при формализации рассмотренной транспортной задачи оказалось удобным использовать обозначение искомых величин в виде переменных с двойной индексацией, хотя можно было использовать и одноиндексные переменные $x_k \quad (k = 1,...,nm)$.

13. Задача о назначениях (о распределении работ). В цехе имеется два участка A_1 и A_2 по контролю и настройке производимой электронной аппаратуры четырех разновидностей: $B_1,...,B_4$. Известно, что оборудование участков и квалификация работающих там специалистов таковы, что на участке A_i на наладку одного изделия вида B_j требуется время C_{ij} (ч) (табл. 1.7).

Т а б л и ц а 1.7

A_i \ B_j	B_1	B_2	B_3	B_4
A_1	0,8	1,2	0,7	0,5
A_2	0,9	1,0	0,8	0,4

Поступило задание: провести настройку 7 изделий вида B_1, 10 изделий – B_2, 14 изделий – B_3 и 9 изделий – B_4. Как распределить эти изделия по участкам, чтобы общие затраты времени на их настройку были минимальными?

Глава 2. **ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ**

§ 2.1. Общая и каноническая задачи линейного программирования

Как указывалось в § 1.2, оптимизационную задачу называют задачей линейного программирования (ЛП), если она состоит в отыскании значений переменных $x_1,...,x_n$, при которых целевая функция, являющаяся линейной формой этих переменных:

$$y = c_1x_1 + ... + c_nx_n, \tag{2.1}$$

принимает наименьшее (или наибольшее) значение и выполняется система линейных ограничений-равенств:

$$\left.\begin{array}{c} a_{11}x_1 + ... + a_{1n}x_n = b_1, \\ ... \\ a_{m1}x_1 + ... + a_{mn}x_n = b_m \end{array}\right\} \tag{2.2}$$

и система линейных ограничений-неравенств:

$$\left.\begin{array}{c} d_{11}x_1 + ... + d_{1n}x_n \geq e_1, \\ ... \\ d_{k1}x_1 + ... + d_{kn}x_n \geq e_k. \end{array}\right\} \tag{2.3}$$

Примечание 1. При формализации конкретной задачи среди ограничений-неравенств могут встретиться неравенства вида

$$d_{i1}x_1 + ... + d_{in}x_n \leq e_i, \tag{2.4}$$

но они легко сводятся к виду (2.3) умножением обеих частей неравенства на (–1).

Примечание 2. Из физического смысла переменных, фигурирующих в задаче, довольно часто следует, что все либо часть этих переменных должны быть неотрицательны (см. пример 3 из § 1.1). Это приводит к появлению особой группы условий:

$$x_i \geq 0, \quad i = 1,..., p, \tag{2.5}$$

хотя ясно, что они непосредственно вписываются в систему ограничений (2.3).

Выделим частный случай сформулированной общей задачи ЛП, в которой находятся наименьшее значение линейной целевой функции (2.1) при выполнении системы ограничений-равенств (2.2) и требования неотрицательности всех переменных. Такая задача называется канонической задачей линейного программирования (КЗЛП) [1, 26], или основной (ОЗЛП) [6, 26], или стандартной [32].

Отметим, что с помощью матриц

$$\bar{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}; \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}; \quad \bar{b} = \begin{bmatrix} b_1 \\ \dots \\ b_m \end{bmatrix}; \quad \bar{c} = \begin{bmatrix} c_1 \\ \dots \\ c_n \end{bmatrix}$$

соотношения (2.1), (2.2) и (2.5) записываются в краткой форме:

$$y = \bar{c}^T \bar{x}; \quad A\bar{x} = \bar{b}; \quad \bar{x} \geq 0. \tag{2.6}$$

(Последнее условие означает, что все компоненты столбца $\bar{x}$ должны быть неотрицательны).

Общая задача ЛП отличается от канонической тем, что в ней может находиться как наименьшее, так и наибольшее значение целевой функции, а среди ограничений-неравенств кроме простейших условий $x_i \geq 0$ могут присутствовать неравенства, в которые входят несколько переменных, например, $2x_1 + x_2 - x_3 + 4x_5 \geq 10$. В теории ЛП рассматривается главным образом более простая КЗЛП. При этом учитывается то, что общую задачу ЛП всегда можно свести к канонической.

Если необходимо найти наибольшее значение целевой функции y, то для сведения этой задачи к канонической следует перейти к вспомогательной функции $\tilde{y} = -y$. Ясно, что при этом наименьшему значению $\tilde{y}$ будет соответствовать наибольшее значение y и наоборот.

Допустим, что при формализации конкретной практической задачи возникло ограничение-неравенство:

$$a_{i1}x_1 + \dots + a_{in}x_n \geq b_i. \tag{2.7}$$

Введем новую дополнительную переменную x_{n+1}, полагая:

$$x_{n+1} = a_{i1}x_1 + \dots + a_{in}x - b_i. \tag{2.8}$$

Тогда неравенству (2.7) будут эквивалентны два условия:

$$a_{i1}x_1 + \dots + a_{in}x - x_{n+1} = b_i; \quad x_{n+1} \geq 0, \tag{2.9}$$

соответствующие условиям канонической задачи.

Пример. Формализация задачи 14 о планировании производства из § 1.3 приводит к следующей задаче ЛП: найти неотрицательные значения x_1 и x_2, при которых целевая функция, выражающая прибыль предприятия от реализации продукции

$$y = 3x_1 + 2x_2, \tag{2.10}$$

будет наибольшей при наличии ограничений, выражаемых системой неравенств:

$$\begin{aligned} &x_1 + 2x_2 \leq 6; \quad x_1 - x_2 \leq 1; \\ &2x_1 + x_2 \leq 8; \qquad x_1 \leq 2. \end{aligned} \tag{2.11}$$

Требуется свести эту задачу к КЗЛП.

Решение. Преобразуем каждое из неравенств системы (2.11) к виду

$$\sum_{j=1}^{n} a_{ij} x_j - b_j \geq 0. \qquad (2.12)$$

В результате получаем:

$$\begin{aligned} &6 - x_1 - 2x_2 \geq 0; \quad 1 - x_1 + x_2 \geq 0; \\ &8 - 2x_1 - x_2 \geq 0; \qquad 2 - x_1 \geq 0. \end{aligned} \qquad (2.13)$$

Вводим дополнительные переменные:

$$\begin{aligned} &x_3 = 6 - x_1 - 2x_2; \quad x_5 = 1 - x_1 + x_2; \\ &x_4 = 8 - 2x_1 - x_2; \quad x_6 = 2 - x_1. \end{aligned} \qquad (2.14)$$

Преобразуя эту систему к стандартному виду (2.2), т.е. перенося все члены уравнений, содержащие переменные x_i, в левую часть, а свободные члены – в правую, вместо исходной системы неравенств (2.11) получаем систему уравнений

$$\begin{aligned} &x_1 + 2x_2 + x_3 = 6; \quad x_1 - x_2 + x_5 = 1; \\ &2x_1 + x_2 + x_4 = 8; \quad\; x_1 + x_6 = 2 \end{aligned} \qquad (2.15)$$

и требования $x_i \geq 0, \quad i = \overline{1,6}.$

Вводя вспомогательную функцию

$$\tilde{y} = -3x_1 - 2x_2, \qquad (2.16)$$

получаем вместо исходной задачи об отыскании наибольшего значения функции (2.10) при выполнении неравенств (2.11) каноническую задачу об отыскании наименьшего значения функции (2.16) при выполнении системы линейных алгебраических уравнений (2.15) и требовании неотрицательности всех переменных $x_i \quad (i = \overline{1,6})$.

Будем в дальнейшем рассматривать задачи ЛП в канонической форме, считая, что в том случае, когда исходная задача окажется сформулированной в общем виде, она указанным выше образом будет сведена в КЗЛП.

Отметим, что каждый набор переменных $(x_1,...,x_n)$, удовлетворяющий системе ограничений (2.2) и требованию неотрицательности $x_i \geq 0$ $(i = 1,...,n)$, называется допустимым решением рассматриваемой задачи ЛП, а множество всех таких наборов называется ***областью допустимых решений*** (ОДР). Ясно, что для осуществления решения задачи ЛП необходимо (хотя и не достаточно), чтобы ОДР не была пустым множеством, а для этого нужно прежде всего, чтобы система уравнений (2.2) была совместна. Для этого (в соответствии с теоремой Кронекера – Капелли [17]) необходимо и достаточно, чтобы ранг матрицы системы A был равен рангу расширенной матрицы этой системы A_p, где

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}; \; A_p = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdot & \cdot & \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix}.$$

Напомним, что ранг матрицы – это наибольший порядок отличных от нуля миноров данной матрицы. Это означает, что если ранг матрицы равен r, то в матрице имеется хотя бы один отличный от нуля минор порядка r, но всякий минор порядка большего чем r равен нулю. При этом ранг матрицы равен наибольшему числу линейно независимых строк (или столбцов) матрицы. Ясно, что r не превышает наименьшего из значений n и m.

В дальнейшем будем считать, что указанное выше условие совместности системы (2.2) выполнено и ранг матрицы $A = A_P$ обозначать r.

Для совместной системы возможны два случая.

1. Ранг системы равен числу неизвестных, т.е. $r = m$. Тогда система (2.2) имеет единственное решение $\bar{x} = \bar{x}^{\mathrm{o}}$. Если все компоненты x_i^{o} этого решения неотрицательны, то ОДР состоит из одной точки $\bar{x}^{\mathrm{o}}$. Она и является решением рассматриваемой КЗЛП. Этот случай интереса не представляет.

2. В случае, когда $r < n$, система (2.2) имеет бесконечное множество решений. Его образуют наборы значений переменных $(x_1,...,x_n)$, в которых $(n - r)$ переменных, называемых свободными, принимают любые значения, а остальные r переменных, называемых базисными, определяются в виде линейной комбинации соответствующих значений свободных переменных.

В указанном бесконечном множестве решений системы (2.2) нужно выделить образующее ОДР подмножество тех решений, в которых все переменные неотрицательны. Если оно не окажется пустым, то в нем и следует искать решение $\bar{x}$, при котором целевая функция (2.1) принимает наименьшее значение.

Учитывая изложенное, будем далее рассматривать только КЗЛП, в которых число r линейно независимых уравнений-ограничений меньше числа n переменных x_i.

§ 2.2. Графическое решение задач ЛП

В тех случаях, когда в поставленной задаче $n = r + 1$ или $n = r + 2$, ее можно решить довольно просто, используя наглядные геометрические представления.

Пример. Найти значения переменных $x_1,...,x_5$, при которых целевая функция

$$y = x_1 + x_2 + x_3 + 4x_4 \qquad (2.17)$$

принимает наименьшее значение и выполняются условия:

$$\begin{aligned} &-x_1 + 2x_2 + x_4 = 2; \qquad x_2 + x_5 = 6; \\ &-x_2 + x_3 + x_4 = 4; \qquad x_i \geq 0 \; (i = 1,...,5). \end{aligned} \qquad (2.18)$$

Решение. Данная задача является КЗЛП, в ней $n = 5, r = 3$. Так как $n - r = 2$, то две переменные нужно выбрать в качестве свободных, а остальные выразить через них. В данном случае удобно взять в качестве свободных переменных x_2 и x_4, тогда остальные переменные непосредственно выражаются через них из системы уравнений (2.18):

$$x_1 = -2 + 2x_2 + x_4; \quad x_3 = 4 + x_2 - x_4; \quad x_5 = 6 - x_2. \qquad (2.19)$$

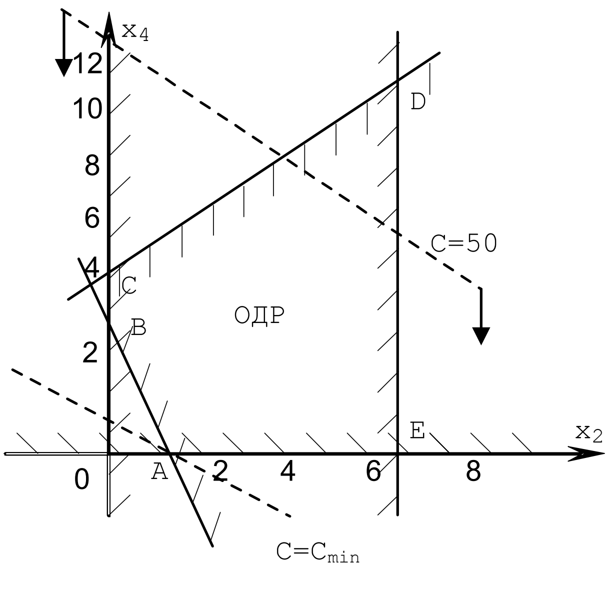


Рис. 2.1

Рассмотрим плоскость с координатами x_2, x_4 (рис. 2.1). На ней можно наглядно интерпретировать условия-ограничения задачи.

Условию $x_4 \geq 0$ соответствует множество точек верхней полуплоскости, а $x_2 \geq 0$ – правой полуплоскости; обоим условиям соответствует 1-й квадрант этой плоскости. Отметим его штриховкой на рис. 2.1.

Рассмотрим теперь условие $x_1 \geq 0$. Используя первое уравнение из системы (2.19), получаем: $-2 + 2x_2 + x_4 \geq 0$ или $x_4 \geq 2 - 2x_2$.

Равенству $x_4 = 2 - 2x_2$ соответствует прямая, проходящая через точки ($x_2 = 0$, $x_4 = 2$) и ($x_4 = 0$, $x_2 = 1$). Неравенству $x_4 \geq 2 - 2x_2$ соответствует множество точек, лежащих на этой прямой и выше ее, т.е. полуплоскость. Отметим ее штриховкой.

Аналогично рассматриваем остальные условия:

$$x_3 \geq 0; \quad 4 + x_2 - x_4 \geq 0; \quad x_4 \leq 4 + x_2;$$

$$x_5 \geq 0; \qquad 6 - x_2 \geq 0; \qquad\qquad x_2 \leq 6.$$

Таким образом, каждому неравенству $x_i \geq 0$ на плоскости $x_2 0 x_4$ соответствует множество точек некоторой полуплоскости. По условиям задачи должны выполняться все неравенства $x_i \geq 0$, i = 1, 2,…, 5. Значит, на плоскости нужно выделить те точки (если они существуют), которые принадлежат всем полученным полуплоскостям. В рассматриваемом случае множество таких точек ограничено многоугольником *ABCDE* (см. рис. 2.1).

В каждой точке, принадлежащей границе или внутренней области *ABCDE*, свободные переменные x_2 и x_4, а также определяемые соотношениями (2.19) базисные переменные x_1, x_3, x_5 удовлетворяют всем условиям – ограничениям данной КЗЛП и, следовательно, они образуют область допустимых решений (ОДР) (см. рис. 2.1).

Для решения поставленной задачи теперь нужно в полученной ОДР отыскать такую точку, в которой целевая функция принимает наименьшее значение. Для этого подставим выражения (2.19) в ЦФ (2.17). В результате получаем

$$y = (-2 + 2x_2 + x_4) + x_2 + (4 + x_2 - x_4) + 4x_4 = 2 + 4x_2 + 4x_4.$$

Рассмотрим множество точек (x_2, x_4), для которых целевая функция принимает одно и то же значение y = C. Геометрическое место таких точек образует линию равного уровня ЦФ. Найдем уравнение этой линии:

$$C = 2 + 4x_2 + 4x_4 \text{ или } x_4 = \frac{C-2}{4} - x_2.$$

Оно показывает, что линии равного уровня ЦФ в рассматриваемом случае представляют собой прямые. Такая линия для C = 50 пунктиром изображена на рис. 2.1. При различных значениях C наклон прямых не меняется, причем уменьшению C соответствует смещение этих прямых вниз.

Отыскание в ОДР точки, в которой целевая функция принимает наименьшее значение, сводится к перемещению прямой y = C без изменения наклона в сторону убывания y (в рассматриваемом случае – вниз) до тех пор, пока эта прямая не достигнет крайней точки (в данном случае такой является точка *A* (см. рис. 2.1)). Координаты этой точки соответствуют искомым значениям свободных переменных: $x_2^0 = 1$; $x_4^0 = 0$. Используя их в выражениях (2.19), получим искомые значения остальных (базисных) переменных: $x_1^0 = 0$; $x_3^0 = 5$; $x_5^0 = 5$.

Геометрическое истолкование рассмотренной задачи ЛП подсказывает следующие основные свойства решений КЗЛП общего вида (строгое доказательство этих свойств см., например, в [3, 15, 26]).

1. Решение КЗЛП, если оно существует, всегда будет на границе ОДР, представляющей собой выпуклый многогранник в (n – r)-мерном пространстве свободных переменных. Это решение обычно получается в одной из вершин ОДР, тогда оно единственное. При этом по крайней мере n – r переменных равны нулю. В вырожденном случае вершине ОДР может соответствовать более чем n – r переменных, равных нулю. Такой случай иллюстрируется рис. 2.2 (т. A).

2. Искомое оптимальное решение КЗЛП может оказаться не единственным. Такой случай показан на рис. 2.3. Здесь прямая, соответствующая y = C = const, параллельна грани AB. При этом каждая точка грани AB будет давать одно и то же значение ЦФ, являющееся наименьшим возможным значением для данной задачи.

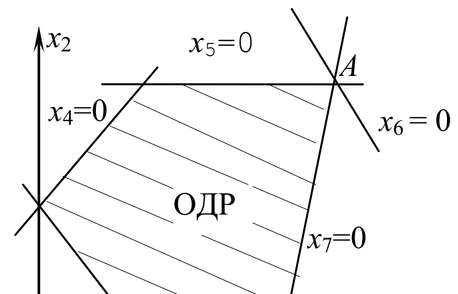


Рис. 2.2

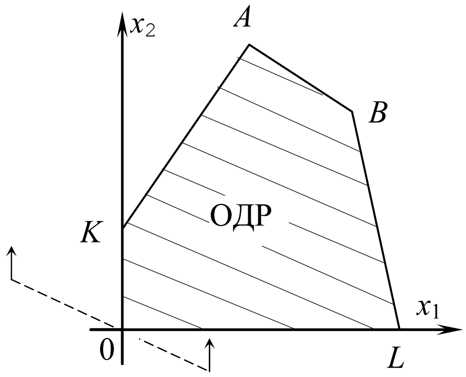


Рис. 2.3

3. КЗЛП может не иметь решения в двух случаях:

а) когда ОДР – пустое множество (получающаяся система неравенств оказывается противоречивой (рис. 2.4, a);

б) ОДР не ограничена и оказывается возможным неограниченное уменьшение целевой функции (рис. 2.4, $б$).

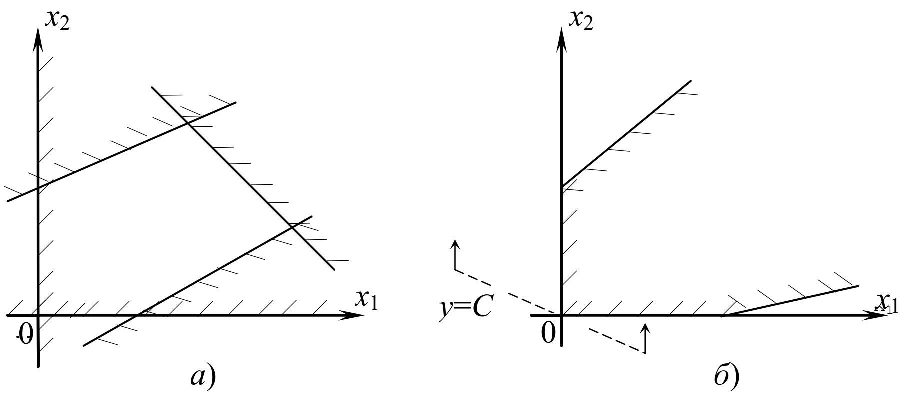


Рис. 2.4

Рассмотренные свойства решений ЗЛП лежат в основе алгебраического подхода к решению задач ЛП, реализуемого в так называемом симплекс-методе.

§ 2.3. Алгебраическое решение задач ЛП. Сущность симплекс-метода

Рассмотрим некоторую задачу ЛП вида (2.6). Например, найти $x_1,...,x_5$, при которых функция

$$y = 63 + x_1 + 2x_2 + 2x_3 + 4x_4 + x_5 \qquad (2.20)$$

принимает наименьшее значение и выполняются условия:

$$\begin{aligned} &x_1 + 4x_2 + 4x_4 = 24; \qquad && x_1 + x_3 + x_5 = 9; \\ &2x_1 + 3x_2 + x_3 + 4x_4 = 28; && x_i \geq 0. \end{aligned} \qquad (2.21)$$

Выше было установлено, что решение такой задачи получается в одной из вершин ОДР. Кроме того, отмечалось, что любой вершине ОДР соответствует комбинация переменных $x_1,...,x_n$, в которой $(n-r)$ переменных равны нулю, а остальные переменные неотрицательны. Такая комбинация называется опорным решением или ***опорным планом***. Найдем одну из таких комбинаций для рассматриваемой задачи. В ней $n=5$, $r=3$, поэтому выбираем $n-r=2$ переменных, например, x_1 и x_2 в качестве свободных, а остальные выражаем через них из системы (2.21). Для этого x_4 выражаем из 1-го уравнения: $x_4 = 6 - \frac{1}{4}x_1 - x_2$, полученное соотношение подставляем во 2-е уравнение и получаем $x_3 = 4 - x_1 + x_2$. Используя этот результат, из последнего уравнения системы (2.21) находим $x_5 = 5 - x_2$.

Подставив выражения для x_3, x_4, x_5:

$$x_3 = 4 - x_1 + x_2; \quad x_4 = 6 - \frac{1}{4}x_1 - x_2; \quad x_5 = 5 - x_2 \qquad (2.22)$$

в (2.20), получаем

$$y = 100 - 2x_1 - x_2. \qquad (2.23)$$

Положим, что в (2.22) свободные переменные равны нулю, в результате найдем следующую комбинацию значений переменных:

$$x_1 = 0; \quad x_2 = 0; \quad x_3 = 4; \quad x_4 = 6; \quad x_5 = 5. \qquad (2.24)$$

Поскольку в ней $n - r$ переменных равны нулю, а остальные неотрицательны, то (2.24) является опорным решением рассматриваемой КЗЛП. Возникает вопрос – является ли оно оптимальным решением. Анализируя выражение для целевой функции (2.23), видим, что соответствующее решению (2.24) значение этой функции $y = 100$ не является наименьшим возможным значением, так как заменяя $x_1 = 0$ или $x_2 = 0$ некоторыми положительными величинами, можно уменьшить это значение.

Это говорит о том, что по крайней мере одна переменная из выбранных нами в качестве свободных переменных должна быть переведена в разряд базисных, а на ее место должна соответственно встать другая переменная, которая до сих пор была базисной. Пусть в базисные переменные переводится переменная x_1 (она теперь будет больше нуля), а x_2 пока остается свободной (она будет по-прежнему равной нулю).

Возникает вопрос – какую переменную из x_3, x_4, x_5 перевести в свободные и насколько можно увеличить переменную x_1. Так как при увеличении x_1 ЦФ пропорционально x_1 уменьшается, то желательно было бы придать x_1 как можно большее значение. Но изменение x_1 приводит к изменению базисных переменных в соответствии с (2.22).

Получаемые из (2.22) соотношения

$$x_3 = 4 - x_1; \quad x_4 = 6 - \frac{1}{4}x_1; \quad x_5 = 5 \qquad (2.25)$$

показывают, что при чрезмерном увеличении x_1 переменные x_3 и x_4 могут принять недопустимые отрицательные значения. Анализируя (2.25), видим, что предельно допустимое увеличение x_1, при котором x_3 уменьшается до нуля, равно $x_1 = 4/1 = 4$, аналогично для x_4 получаем «критическое» значение $x_1 = 6/(1/4) = 24$.

Увеличивая x_1 от 0 до 4, находим вместо (2.24) новое опорное решение:

$$x_1 = 4; \quad x_2 = 0; \quad x_4 = 5; \quad x_3 = 0; \quad x_5 = 5, \qquad (2.26)$$

при этом $y = 100 - 2 + 4 = 92$.

Найдем соотношения, связывающие отличные от нуля переменные x_1, x_4, x_5 с переменными x_2 и x_3, получившими нулевые значения. Для этого выразим из 1-го уравнения системы (2.22) x_1 и полученное соотношение подставим в остальные уравнения (2.22) и ЦФ (2.23). В результате получим:

$$x_1 = 4 + x_2 - x_3; \qquad x_5 = 5 - x_2;$$
$$x_4 = 5 - \frac{5}{4}x_2 + \frac{1}{4}x_3; \quad y = 92 - 3x_2 + 2x_3. \tag{2.27}$$

Анализируя (2.27), видим, что возможно дополнительное уменьшение ЦФ за счет увеличения x_2, так как коэффициент при x_2 является отрицательным. Рассматривая соотношения, входящие в (2.27), устанавливаем, что увеличение x_2 «неопасно» для x_1, так как переменная x_1 увеличивается, а для x_4 и x_5 «опасно», так как коэффициенты при x_2 в уравнениях для x_4 и x_5 отрицательны. Определяя отношения свободных членов к коэффициентам при x_2 в этих уравнениях, находим допустимое увеличение переменной x_2:

$$\text{по } x_4\text{: } x_2^{\text{доп}} = 5/(5/4) = 4; \text{ по } x_5\text{: } x_2^{\text{доп}} = 5/1 = 5. \tag{2.28}$$

Принимая наименьшее из этих значений, увеличиваем x_2 от 0 до 4. В результате вместо опорного решения (2.26) получаем:

$$x_1 = 8; \quad x_2 = 4; \quad x_3 = 0; \quad x_4 = 0; \quad x_5 = 1. \tag{2.29}$$

Выразим из 2-го уравнения системы (2.27) переменную x_2, ставшую ненулевой, и полученное соотношение подставим в остальные уравнения этой системы:

$$x_2 = 4 - \frac{4}{5}x_4 + \frac{1}{5}x_3; \; x_1 = 8 - \frac{4}{5}x_3 - \frac{4}{5}x_4; \quad x_5 = 1 + \frac{4}{5}x_4 - \frac{1}{5}x_3, \tag{2.30}$$

а также в целевую функцию $y = 80 + \frac{7}{5}x_3 + \frac{12}{5}x_4$.

Выражение для ЦФ показывает, что значение функции y = 80, получаемое для опорного решения (2.29), является наименьшим, так как коэффициенты при x_2 и x_4 положительны, и увеличение этих переменных привело бы к увеличению y. Таким образом, опорное решение (2.29) является искомым решением рассматриваемой КЗЛП.

Отметим, что в рассматриваемой задаче $n-r=2$, поэтому решение этой задачи можно наглядно интерпретировать на плоскости подобно тому, как это было изложено в § 2.2. Такая геометрическая интерпретация представлена на рис. 2.5.

Первому выбранному по существу наугад опорному решению (2.24) на рис. 2.5 соответствует вершина A_1 в многоугольнике $A_1A_2A_3A_4A_5$, являющемся ОДР и построенном по уравнениям (2.22) и условиям $x_i \geq 0$.

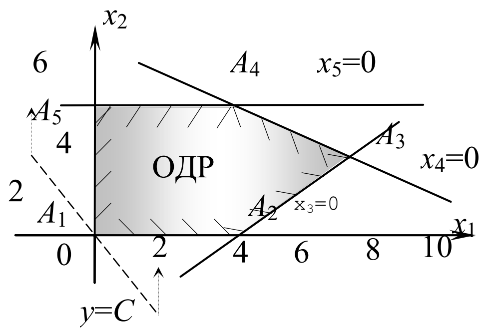


Рис. 2.5

Найденному затем решению (2.26) соответствует вершина A_2, являющаяся точкой пересечения прямых $x_2 = 0$ и $x_3 = 0$. Полученное вслед за ним решение (2.29) оказалось искомым. Ему соответствует вершина A_3. Это же решение получается в результате перемещения линии равных значений ЦФ (пунктирная прямая на рис. 2.5). Таким образом, начав с вершины A_1, выбранной наугад в ОДР, мы в соответствии с рассмотренным алгоритмом последовательно переходили от вершины к вершине в ОДР, все время уменьшая значения ЦФ и приближаясь к ее наименьшему значению.

Рассмотренная для наглядности на конкретном примере методика решения КЗЛП применима к задачам любой размерности и сводится к выполнению следующих операций:

1. Разбиваем список переменных $x_1,\ldots,x_n$ на $m=n-r$ свободных переменных, например, $x_1,x_2,\ldots,x_m$ и r базисных – соответственно $x_{m+1},x_{m+2},\ldots,x_n$. Пользуясь методом исключения, выражаем выбранные базисные переменные через свободные. В результате получаем соотношения вида:

$$x_i = \beta_i - (\alpha_{i1}x_1 + \alpha_{i2}x_2 + \ldots + \alpha_{is}x_s + \ldots + \alpha_{im}x_m), \qquad (2.31)$$

где $i = m+1,\ldots,n$. Пользуясь этими соотношениями, ЦФ тоже выражаем через свободные переменные:

$$y = \gamma_0 - (\gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_s x_s + \ldots + \gamma_m x_m). \qquad (2.32)$$

Примечание. Вслед за [6] будем при записи выражений базисных переменных и целевой функции через свободные переменные пользоваться формой (2.31) и (2.32), в которой свободные переменные объединены в скобки, перед которыми вынесен знак “минус”.

Предположим, что все свободные члены β_i неотрицательны.

2. Анализируем выражение для ЦФ (2.32) и выясняем, имеется ли среди коэффициентов $\gamma_1,...,\gamma_m$ при свободных переменных хотя бы один неотрицательный (с учетом знака "минус" за скобкой). Если таких коэффициентов нет, то решение

$$x_1 = x_2 = ... = x_s = ...x_m = 0; \quad x_i = \beta_i \text{ для } i = m+1,...,n \qquad (2.33)$$

является искомым решением данной КЗЛП.

В противном случае опорное решение (2.32) оптимальным не является, и нужно перейти к следующей операции.

3. Выбираем одну из переменных, например, x_s с неотрицательным коэффициентом γ_s. Если таких коэффициентов несколько, можно выбирать, например, переменную с наибольшим по модулю коэффициентом.

4. Среди коэффициентов α_{is} при этой переменной x_s выбираем положительные коэффициенты (опять же с учетом знака «минус» за скобками), для каждого из них находим отношение β_i / α_{ik} и выделяем ту переменную, например x_k, для которой это отношение оказалось наименьшим.

5. Из уравнения для выделенной базисной переменной, входящего в систему (2.31) $\underline{\underline{x}}_k = \beta_k - (\alpha_{k1}x_1 + ... + \alpha_{ks}\underline{\underline{x}}_s + ... + \alpha_{km}x_m)$, выражаем выделенную свободную переменную

$$\underline{\underline{x}}_s = \frac{\beta_k}{\alpha_{ks}} - (\frac{\alpha_{k1}}{\alpha_{ks}} x_1 + ... + \frac{1}{\alpha_{ks}} \underline{\underline{x}}_k + ... + \frac{\alpha_{km}}{\alpha_{ks}} x_m). \qquad (2.34)$$

6. Подставляем (2.34) в каждое из остальных уравнений системы (2.31) и в выражение для ЦФ (2.32); после приведения подобных членов получаем выражения:

$$x_i = \left(\beta_i - \beta_k \frac{\alpha_{is}}{\alpha_{ks}}\right) - \left[\left(\alpha_{i1} - \alpha_{k1}\frac{\alpha_{is}}{\alpha_{ks}}\right)x_1 + \left(\alpha_{i2} - \alpha_{k2}\frac{\alpha_{is}}{\alpha_{ks}}\right)x_2 + \right.$$
$$\left. + ... + \left(\frac{\alpha_{is}}{\alpha_{ks}}\right)x_k + ... + \left(\alpha_{im} - \alpha_{km}\frac{\alpha_{is}}{\alpha_{ks}}\right)x_m\right]; \qquad (2.35)$$

$$y_i = \left(\gamma_0 - \beta_k \frac{\gamma_s}{\alpha_{ks}}\right) - \left[\left(\gamma_1 - \alpha_{k1}\frac{\gamma_s}{\alpha_{ks}}\right)x_1 + \left(\gamma_2 - \alpha_{k2}\frac{\gamma_s}{\alpha_{ks}}\right)x_2 + \right.$$
$$\left. + ... + \left(-\frac{\gamma_s}{\alpha_{ks}}\right)x_k + ... + \left(\gamma_m - \alpha_{km}\frac{\gamma_s}{\alpha_{ks}}\right)x_m\right]. \qquad (2.36)$$

Уравнения (2.35) и (2.36) по виду аналогичны (2.31) и (2.32). Однако в списке свободных и базисных переменных, фигурирующих в этих уравнениях, переменные x_k и x_s поменялись местами.

7. С полученными уравнениями (2.35) и (2.36) переходим к шагу 2.

Перечисленные операции, входящие в алгебраическое решение КЗЛП (симплекс-метод), можно выполнять с помощью так называемой симплекс-таблицы. Одна из возможных форм ее записи представлена табл. 2.1 [6].

Таблица 2.1

	β	x_1	…	x_s	…	x_m	β/α
x_{m+1}	β_{m+1} $-\beta_k \frac{\alpha_{m+1,s}}{\alpha_{ks}}$	$\alpha_{m+1,1}$ $-\alpha_{k1} \frac{\alpha_{m+1,s}}{\alpha_{ks}}$	…	$\alpha_{m+1,s}$ $-\frac{\alpha_{m+1,s}}{\alpha_{ks}}$	…	$\alpha_{m+1,m}$ $-\alpha_{km} \frac{\alpha_{m+1,s}}{\alpha_{ks}}$	$\beta_{m+1}/\alpha_{m+1,s}$
…	…	…	…	…	…	…	…
x_k	β_k $\frac{\beta_k}{\alpha_{ks}}$	α_{k1} $\frac{\alpha_{k1}}{\alpha_{ks}}$	…	α_{ks} $\frac{1}{\alpha_{ks}}$	…	α_{km} $\frac{\alpha_{km}}{\alpha_{ks}}$	$\beta_k/\alpha_{k,s}$
…	…	…	…	…	…	…	…
x_n	β_n $-\beta_k \frac{\alpha_{ns}}{\alpha_{ks}}$	α_{n1} $-\alpha_{k1} \frac{\alpha_{ns}}{\alpha_{ks}}$	…	α_{ns} $-\frac{\alpha_{ns}}{\alpha_{ks}}$	…	α_{nm} $-\alpha_{km} \frac{\alpha_{ns}}{\alpha_{ks}}$	$\beta_n/\alpha_{k,s}$
y	γ_0 $-\beta_k \frac{\gamma_s}{\alpha_{ks}}$	γ_1 $-\alpha_{k1} \frac{\gamma_s}{\alpha_{ks}}$	…	 $-\frac{\gamma_s}{\alpha_{ks}}$	…	γ_m $-\alpha_{km} \frac{\alpha_{ns}}{\alpha_{ks}}$	

Другие формы ее записи и соответствующие правила ее преобразования представлены в [1, 3, 14, 32].

В клетки этой таблицы (в левый верхний угол) записываются соответствующие коэффициенты из системы ограничений-равенств и выражения для ЦФ, записанные в форме (2.31) и (2.32).

Выполнение операций № 2 и 3 сводится к анализу коэффициентов при переменных в строке для ЦФ. Если все они отрицательны, то данная таблица соответствует оптимальному решению. При этом оптимальные значения свободных переменных равны нулю, а базисные – соответствующим значениям из столбца β. Искомое наименьшее значение целевой функции $y_{\inf} = \gamma_0$. Заметим, что при отыскании наибольшего значения ЦФ искомому решению будет соответствовать симплекс-таблица, в которой все коэффициенты $\gamma_1,...,\gamma_m$ положительны. Если среди коэффициентов $\gamma_1,...,\gamma_m$ в строке для y имеется хотя бы один неотрицательный, то следует преобразовать таблицу. Для этого нужно:

1) выделить столбец, содержащий такой коэффициент (если их несколько, то выбирают наибольший по модулю); этот столбец называют ***разрешающим***;

2) анализируя коэффициенты α, стоящие в этом столбце, выделить из них положительные и вычислить отношения β/α в соответствующих строках; строку, для которой это отношение минимально, называют ***разрешающей***;

3) выделить элемент α_{ks}, стоящий в разрешающей строке и разрешающем столбце; его называют ***разрешающим*** (или генеральным, или ведущим);

4) в клетке, содержащей разрешающий элемент, в правом нижнем углу записать величину, обратную ему, т.е. $\lambda = 1/\alpha_{ks}$;

5) все остальные элементы разрешающей строки умножить на λ, и результат записать в нижней части тех же клеток;

6) все элементы разрешающего столбца (кроме α_{ks}) умножить на ($-\lambda$) и результат записать в нижней части тех же клеток;

7) выделить (подчеркнуть или обвести рамкой) в разрешающей строке верхние числа, а в разрешающем столбце – все нижние числа (за исключением разрешающего элемента);

8) в каждой клетке, не принадлежащей разрешающей строке или столбцу, записать в нижнюю часть произведение выделенных чисел, стоящих в том же столбце и в той же строке;

9) переписать таблицу, заменив:

– “бывшую” свободную переменную на базисную и наоборот ($x_s \leftrightarrow x_k$);

– элементы разрешающей строки и столбца – числами, записанными в нижние части тех же клеток;

– каждый из остальных элементов – суммой чисел, стоящих в верхних и нижних частях тех же клеток.

Пример. Найти $x_i \geq 0$, при которых $y = 10 + x_1 - 2x_2 + 4x_3$ принимает наименьшее значение и выполняются условия:

$$x_1 + 3x_2 - x_3 + x_4 = 9; \qquad -x_1 - 2x_2 + x_3 + x_6 = 5;$$
$$-2x_1 + 6x_2 - 3x_3 + x_5 = 6; \quad -2x_1 + x_2 - 4x_3 + x_7 = 3.$$

Решение. Представим систему уравнений в виде:

$$x_4 = 9 - x_1 - 3x_2 + x_3; \qquad x_6 = 5 + x_1 + 2x_2 - x_3;$$

$$x_5 = 6 + 2x_1 - 6x_2 + 3x_3; \quad x_7 = 3 + 2x_1 - x_2 + 4x_3.$$

Тем самым переменные $x_1,...,x_7$ разделены на свободные (x_1, x_2, x_3) и базисные (x_4, x_5, x_6, x_7). Отметим, что это разбиение соответствует начальному плану, так как все свободные коэффициенты в полученной системе положительны.

Эту систему уравнений и выражение для ЦФ представляем в виде табл. 2.2. В ней выделяем ведущий столбец, соответствующий положительному коэффициенту в выражении для *у*. Этот столбец соответствует переменной x_2. Находим для элементов этого столбца отношение β/α и выбираем наименьшее из них. Получаем ведущую строку, соответствующую переменной x_5. Заполняем части клеток таблицы по описанному выше алгоритму и переходим к новой табл. 2.3, в которой x_2 и x_5 поменялись местами, элементы разрешающей строки и столбца заменены нижними числами тех же клеток, а остальные элементы получены как сумма верхних и нижних чисел.

Т а б л и ц а 2.2

СП \ БП	Св.ч	x_1	x_2	x_3	β/α
x_4	9 / –3	1 / 1	3 / –1/2	–1 / 3/2	9/3=3
x_5	6 / 1	–2 / –1/3	6 / 1/6	-3 / –1/2	6/6=1
x_6	5 / 2	–1 / –2/3	–2 / 1/3	1 / –1	*
x_7	3 / –1	–2 / 1/3	1 / –1/6	– 4 / 1/2	3/1=3
y	10 / –2	–1 / 2/3	2 / –1/3	– 4 / 1	

Т а б л и ц а 2.3

СП \ БП	Св.ч	x_1	x_5	x_3
x_4	6	2	–1/2	1/2
x_2	1	–1/3	1/6	–1/2
x_6	7	–5/3	1/3	0
x_7	2	–5/3	–1/6	–7/2
y	8	–1/3	–1/3	–3

В полученной табл. 2.3 все коэффициенты при переменных x_i в строке для *у* отрицательны. Это означает оптимальность решения, соответствующего табл. 2.3: $x_1 = x_3 = x_5 = 0$, $x_4 = 6$, $x_2 = 1$, $x_6 = 7$, $x_7 = 2$, $y_{\inf} = 8$.

§ 2.4. Отыскание начального опорного решения методом искусственного базиса

Для использования симплекс-метода при решении КЗЛП нужно из переменных $x_1,...,x_n$ какие-либо m переменных выбрать в качестве базисных, а затем уравнения системы представить в виде (2.31), причем свободные члены β в получаемых выражениях должны быть неотрицательны. Такое представление соответствует некоторому опорному плану, который можно использовать в качестве начального для симплекс-метода. Для его отыскания можно применить метод искусственных переменных, реализуемый в одной из двух возможных модификаций [1, 3, 32].

***M*-метод (метод штрафов)**

Поясним его сущность на следующем примере. Найти наименьшее значение функции $y = 5 - x_1 - 2x_2$ при ограничениях

$$x_1 + x_2 - x_3 = 2; \quad 2x_1 + x_2 + x_4 = 6. \tag{2.37}$$

В этой задаче $m = 2$; $n = 4$; $(n - m) = 2$. Отметим, что данную систему легко разрешить относительно x_3, x_4:

$$x_3 = -2 + x_1 + x_2; \quad x_4 = 6 - 2x_1 - x_2,$$

но получаемое опорное решение ($x_1 = 0, x_2 = 0, x_3 = -2, x_4 = 4$) оказывается недопустимым, так как $x_3 < 0$.

Перепишем систему ограничений так, чтобы в каждом уравнении свободный член был неотрицательным, и введем m вспомогательных переменных $\tilde{x}_5$, $\tilde{x}_6$, добавляя их в левую часть получаемых уравнений:

$$x_1 + x_2 - x_3 + \tilde{x}_5 = 2; \quad 2x_1 + x_2 + x_4 + \tilde{x}_6 = 6. \tag{2.38}$$

Кроме того, запишем новую целевую функцию

$$\tilde{y} = (5 - x_1 - 2x_2) + M(\tilde{x}_5 + \tilde{x}_6), \tag{2.39}$$

где M – достаточно большое положительное число.

Уравнения-ограничения (2.38) легко сводятся к виду (2.31):

$$\tilde{x}_5 = 2 - (x_1 + x_2 - x_3); \; \tilde{x}_6 = 6 - (2x_1 + x_2 + x_4) \tag{2.40}$$

и, следовательно, для (2.39) и (2.40) можно непосредственно использовать симплекс-метод. При этом $\tilde{n} = 6$; $\tilde{m} = 2$, $\tilde{n} - \tilde{m} = 4$, и

$$\tilde{y} = (5 - 8M) - [(1 + 3M)x_1 + (2 + 2M)x_2 - Mx_3 + Mx_4] \; . \tag{2.41}$$

Идея рассмотренного подхода состоит в том, что при минимизации $\tilde{y}$ из-за очень большого числа M любые решения, получающиеся при $\tilde{x}_i > 0$, будут хуже, чем решение, в котором $\tilde{x}_i = 0$.

Допустим, что в результате применения симплекс-метода получено некоторое решение, в котором $(\tilde{n}-\tilde{m})$ переменных равны нулю, и среди них $\tilde{x}_5$, $\tilde{x}_6$. Тогда это решение совпадает с решением исходной задачи, так как при $\tilde{x}_i = 0$ система (2.40) совпадает с (2.37), а функция $\tilde{y} = y$. Если же в этом решении хотя бы одна переменная $\tilde{x}_i \neq 0$, то исходная задача допустимого решения не имеет.

Процесс отыскания решения приведен в табл. 2.4 – 2.7.

Таблица 2.4

БП \ СП	Св.ч	x_1	x_2	x_3	x_4
$\tilde{x}_5$	2 2	1 1	1 1	–1 –1	0 0
$\tilde{x}_6$	6 –4	2 –2	1 –2	0 2	1 0
$\tilde{y}$	5+8M –2(1+3M)	1+3M –(1+3M)	2+2M –(1+3M)	-M (1+3M)	M 0

Таблица 2.5

БП \ СП	Св.ч	$\tilde{x}_5$	x_2	x_3	x_4
x_1	2 1	1 –1	1 –1/2	– 1 1/2	0 1/2
$\tilde{x}_6$	2 1	–2 –1	–1 –1/2	2 1/2	1 1/2
$\tilde{y}$	3+2M – (1+2M)	– (1+3M) (1+2M)	1 – M (1+2M)/2	1+2M – (1+2M)/2	M – (1+2M)/2

Таблица 2.6

БП \ СП	Св.ч	$\tilde{x}_5$	x_2	$\tilde{x}_6$	x_4
x_1	3 6	0 0	1/2 2	1/2 1	1/2 1
x_3	1 3	–1 0	–1/2 1	1/2 1/2	1/2 1/2
$\tilde{y}$	2 –9	–M 0	3/2 –3	–(1+2M)/2 –3/2	–1/2 –3/2

Таблица 2.7

БП \ СП	Св.ч	$\tilde{x}_5$	x_1	$\tilde{x}_6$	x_4
x_2	6	0	2	1	1
x_3	4	–1	1	1	1
$\tilde{y}$	–7	–M	–3	–2 – M	–2

Заключительной таблице, в которой все коэффициенты при переменных в строке для $\tilde{y}$ отрицательны, соответствует искомое оптимальное решение, в котором все вспомогательные переменные оказались равными нулю, а остальные переменные принимают значения $x_i = 0$; $x_2 = 6$; $x_3 = 4$; $x_4 = 0$, образующие решение исходной задачи.

Двухэтапный метод

При реализации рассмотренного выше M-метода на ЦВМ необходимо задать и ввести в программу некоторое конкретное большое число, например M = 100 000. При этом вклад коэффициентов исходной целевой функции, например, 1 и 3 в коэффициенты, определяющие вспомогательную функцию $\tilde{y}$, например $2M + 1 = 200\,001$ и $2M + 3 = 200\,003$, будет мал. В таком случае из-за ошибок округления, возникающих при использовании ЦВМ, процесс вычислений может оказаться нечувствительным к значениям исходных коэффициентов, что может привести к ошибочным результатам [32]. Этих трудностей можно избежать, используя ту же идею искусственного базиса, но решая задачу в два этапа.

На первом этапе систему ограничений-равенств с введенными вспомогательными переменными $\tilde{x}_i$ рассматривают совместно с целевой функцией $\tilde{y}$, равной сумме всех вспомогательных переменных. Если в результате решения этой задачи получается наименьшее значение $\tilde{y} = 0$ и, следовательно, все $\tilde{x}_i = 0$, то исходная задача имеет допустимое решение. Получаемое при этом распределение основных переменных x_i на свободные и базисные соответствует некоторому опорному плану.

На втором этапе, используя этот план в качестве начального, решают исходную задачу. При этом вспомогательные переменные уже не нужны. Проиллюстрируем применение этой методики на рассмотренном выше примере.

На 1-м этапе нужно решить задачу об отыскании наименьшего значения функции $\tilde{y} = \tilde{x}_5 + \tilde{x}_6$ при наличии ограничений:

$$\tilde{x}_5 = 2 - (x_1 + x_2 - x_3); \quad \tilde{x}_6 = 6 - (2x_1 + x_2 + x_4).$$

Этим условиям соответствует симплекс-таблица (табл. 2.8). Отметим, что в ней элементы строки $\tilde{y}$ получаются суммированием элементов в соответствующих столбцах, так как $\tilde{y} = \tilde{x}_5 + \tilde{x}_6$. Результаты симплекс-преобразований этой таблицы представлены в табл. 2.8 – 2.10. Полученное в табл. 2.10 минимальное значение $\tilde{y}$ оказалось равным нулю, что означает наличие допустимого решения исходной задачи. При этом мы получили распределение переменных на базисные x_1, x_3 и свободные x_2, x_4. Связь между ними определяется табл. 2.11, в которой вспомогательные переменные удалены, а исходная функция y соответствующей подстановкой выражена через x_2, x_4. Применяя к ней симплекс-преобразования, получаем результат, представленный в табл. 2.12. Он совпадает с решением, полученным выше с помощью *М*-метода.

1-й этап

Таблица 2.8

БП \ СП	С в.ч	$\tilde{x}_5$	x_2	$\tilde{x}_6$	x_4
$\tilde{x}_5$	2 2	1 1	1 1	– 1 – 1	0 0
$\tilde{x}_6$	6 – 4	2 –2	1 – 2	0 2	1 0
$\tilde{y}$	8 – 6	3 – 3	2 – 3	– 1 3	1 0

Таблица 2.9

БП \ СП	С в.ч	$\tilde{x}_5$	x_2	x_3	x_4
x_1	2 1	1 –1	1 –1/2	–1 1/2	0 1/2
$\tilde{x}_6$	2 1	– 2 – 1	–1 – 1/2	2 1/2	1 1/2
$\tilde{y}$	2 – 2	– 3 2	–1 1	2 –1	1 –1

2-й этап

Таблица 2.10

БП \ СП	Св.ч	$\tilde{x}_5$	x_2	$\tilde{x}_6$	x_4
x_1	3	0	1/2	1/2	1/2
x_3	1	–1	–1/2	1/2	1/2
$\tilde{y}$	0	–1	0	–1	0

Таблица 2.11

БП \ СП	Св.ч		x_2		x_4
x_1	3 6		1/2 2		1/2 1
x_3	1 3		–1/2 1		1/2 1/2
y	2 – 9		3/2 – 3		–1/2 –3/2

Таблица 2.12

БП \ СП	Св.ч		x_1		x_4
x_2	6		2		1
x_3	4		1		1
$\tilde{y}$	–7		–3		–2

Следует отметить, что количество необходимых итераций *М*-метода и двухэтапного метода всегда одинаково [32] .

§ 2.5. Двойственные задачи линейного программирования

Рассмотрим две связанные между собой задачи ЛП.

Задача 1. Найти n переменных $x_i \ge 0\ (i = 1,2,...n),$ при которых ЦФ

$$y = \bar{c}^T \bar{x} \to sup \tag{2.42}$$

принимает наибольшее значение и выполняются m условий-ограничений:

$$A\bar{x} \le \bar{b}, \tag{2.43}$$

где $A = \begin{bmatrix} a_{11} \cdots a_{1n} \\ \cdots \\ a_{m1} \cdots a_{mn} \end{bmatrix};\quad \bar{b} = \begin{bmatrix} b_1 \\ ... \\ b_m \end{bmatrix};\quad \bar{c} = \begin{bmatrix} c_1 \\ ... \\ c_n \end{bmatrix};\quad \bar{x} = \begin{bmatrix} x_1 \\ ... \\ x_n \end{bmatrix}.$

Задача 2. Найти m переменных $u_i \ge 0 \quad (i = 1,...,m)$, при которых ЦФ

$$z = \bar{b}^{\,T} \bar{u} \to inf \tag{2.44}$$

принимает наименьшее значение и выполняются n условий-ограничений:

$$A^T \bar{u} \ge \bar{c}. \tag{2.45}$$

Указанные задачи образуют так называемую двойственную пару. При этом вторая задача называется двойственной по отношению к исходной (прямой) задаче. Она формулируется согласно следующим правилам.

1. Число переменных u_i в двойственной задаче (ДЗ) равно числу ограничений (2.43) в прямой задаче (ПЗ). Число ограничений (2.45) в ДЗ равно числу переменных в ПЗ.

2. Коэффициенты при неизвестных в ЦФ исходной задачи являются правыми частями неравенств-ограничений (2.45) ДЗ, а правые части системы неравенств исходной задачи входят в качестве коэффициентов в целевую функцию ДЗ.

Пример.

Прямая задача:

Найти $x_i \ge 0,\ (i = \overline{1,3})$, при которых

$y = 4x_1 + x_3 \to sup.$

$$\begin{cases} 2x_1 + 5x_2 + x_3 \le 10, \\ x_1 + 9x_3 \le 5, \\ 7x_1 + x_2 \le 8, \\ x_1 - 4x_2 + 3x_3 \le 6. \end{cases}$$

Двойственная задача:

Найти $u_i \ge 0,\ (i = \overline{1,4})$, при которых

$z = 10u_1 + 5u_2 + 8u_3 + 6u_4 \to inf.$

$$\begin{cases} 2u_1 + u_2 + 7u_3 + u_4 \ge 4, \\ 5u_1 + u_3 - 4u_4 \ge 0, \\ u_1 + 9u_2 + 3u_4 \ge 1. \end{cases}$$

Примечание. Требование $x_i \ge 0$ в исходной задаче может распространяться не на все переменные, а в систему ограничений (2.43) могут входить как неравенства, так и строгие равенства. Это отражается на формулировке двойственной задачи следующим образом. Если $x_i \geq 0$, то i-е условие в системе (2.45), входящей в двойственную задачу, является неравенством вида "≥". Если же x_i может принимать любые значения, то i-е соотношение в (2.45) представляет собой равенство. Если j-е ограничение в исходной задаче является неравенством "≤", то в двойственной задаче $u_j \geq 0$. В противном случае u_j может принимать любые значения [1].

Пример.

Прямая задача:

Найти x_i, $(i = \overline{1,3})$, при которых

$$y = 2x_1 + 7x_2 + 4x_3 \to sup.$$

$$\begin{cases} 2x_1 - x_2 + x_3 \leq 9, \\ x_1 + 3x_2 + x_3 = 15, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

Двойственная задача:

найти u_i, $(i = \overline{1,2})$, при которых

$$z = 9u_1 + 15u_2 \to inf.$$

$$\begin{cases} 2u_1 + u_2 \geq 2, \quad u_1 \geq 0; \\ -u_1 + 3u_2 \geq 7, \quad u_2 - \text{любое}; \\ u_1 + u_2 = 4. \end{cases}$$

Между решением прямой и двойственной задач существует тесная связь, которая проявляется, в частности, в следующем.

1. Если одна из пары двойственных задач имеет оптимальное решение, то и другая имеет оптимальное решение, причем

$$y^* = \max\ [y(\overline{x})] = z^* = \min[z(\overline{u})].$$

Если же $\overline{x}$ – произвольный допустимый план исходной задачи, а $\overline{y}$ – произвольный план двойственной задачи, то

$$y(\overline{x}) \leq z(\overline{u}).$$

2. Если одна из пары двойственных задач не имеет решения из-за того, что ЦФ не ограничена (для исходной задачи – сверху и для двойственной – снизу), то другая задача тоже не имеет решения из-за отсутствия допустимых решений.

Имея решение исходной задачи, можно довольно легко найти решение двойственной задачи. Рассмотрим сущность необходимых для этого операций, опираясь для наглядности на конкретный пример.

Пусть исходная задача решена симплекс-методом, для чего она предварительно сведена к канонической форме, а затем для нее найден допустимый опорный план, по которому составлена начальная симплекс-таблица, например табл. 2.13. Выполняя соответствующие симплекс-преобразования, получаем некоторую конечную таблицу, например табл. 2.14, из которой и находится оптимальное решение исходной задачи.

Т а б л и ц а 2.13

СП / БП	$C_{в.ч}$	x_1	x_3	x_4	x_7
x_2					
x_5					
x_6					
y		3	–1	5	2

Т а б л и ц а 2.14

СП / БП	$C_{в.ч}$	x_6	x_3	x_2	x_7
x_4		3	1	7	5
x_5		–2	0	2	–1
x_1		7	8	4	0
y					

Для получения решения двойственной задачи необходимо:

– составить список переменных x_i, которые вошли как базисные в конечную симплекс-таблицу; для рассматриваемого примера это $[x_4, x_5, x_1]$;

– записать строку коэффициентов при этих переменных в исходной ЦФ, фигурирующей в начальной симплекс-таблице $\overline{d}^T = [5\ 0\ 3]$ (если некоторой переменной, например x_5, не оказалось среди свободных переменных, определяющих ЦФ в начальной таблице, то соответствующий коэффициент в указанной строке принимается равным нулю);

– составить список переменных, которые вошли как базисные в начальную симплекс-таблицу, в рассматриваемом случае $[x_2, x_5, x_6]$;

– для каждой из этих m переменных записать столбец из m чисел по следующему правилу: если эта переменная в конечной таблице оказалась свободной переменной, то нужно выписать соответствующий ей столбец из этой таблицы; если же переменная (в данном примере x_5) оказалась в конечной таблице в списке базисных, то записывается столбец, в котором в строке, соответствующей этой базисной переменной, проставляется 1, а остальные позиции заполняются нулями; в результате получается квадратная матрица

$$\begin{array}{cc} & \begin{array}{ccc} x_2 & x_5 & x_6 \end{array} \\ Q = & \begin{bmatrix} 7 & 0 & 3 \\ 2 & 1 & -2 \\ 4 & 0 & 7 \end{bmatrix}; \end{array}$$

– искомое решение двойственной задачи получается перемножением строки $\overline{d}^T$ и матрицы Q:

$$\overline{u}^* = \overline{d}^T Q = [5\ \ 0\ \ 3] \begin{bmatrix} 7 & 0 & 3 \\ 2 & 1 & -2 \\ 4 & 0 & 7 \end{bmatrix} = [47 \quad 0 \quad 36].$$

Пример. Исходная задача состоит в отыскании переменных $x_1 \geq 0$ и $x_2 \geq 0$, при которых ЦФ $y = 2x_1 + 7x_2$ принимает наибольшее значение, и выполняются условия: $-2x_1 + 3x_2 \leq 14;\quad x_1 + x_2 \leq 8$. Необходимо сформулировать и решить двойственную для нее задачу.

Решение

ДЗ будет состоять в отыскании значений $u_1 \geq 0,\ u_2 \geq 0$, при которых ЦФ $z = 14u_1 + 8u_2$ принимает наименьшее значение и выполняются условия $-2u_1 + u_2 \geq 2;\ \ 3u_1 + u_2 \geq 7$.

Отметим, что в рассматриваемой двойственной паре $n = m = 2$, поэтому обе задачи можно легко решить графически. Соответствующие построения представлены на рис. 2.6. В результате получается:

$$x_1^* = 2,\ x_2^* = 6,\ y_{\max} = 46;\quad u_1^* = 1,\ u_2^* = 4,\ z_{\min} = 46.$$

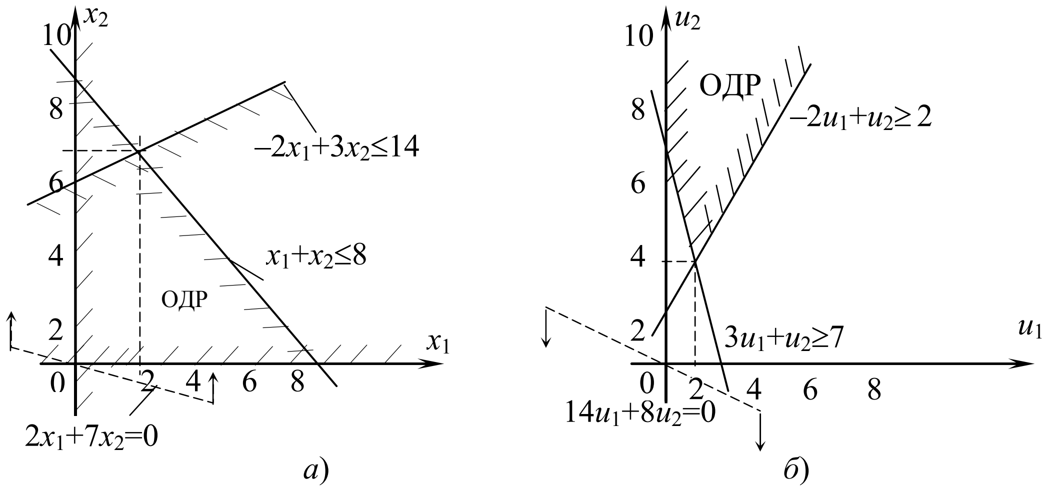


Рис. 2.6

Найдем теперь решение ПЗ симплекс-методом. Для этого, используя дополнительные переменные, сводим ее к канонической задаче об отыскании $x_i \geq 0\ (i = \overline{1,\ 4})$, при которых

$$\tilde{y} = -(2x_1 + 7x_2) \tag{2.46}$$

принимает наименьшее значение и выполняются условия:

$$\tilde{x}_3 = 14 - (-2x_1 + 3x_2);\quad \tilde{x}_4 = 8 - (x_1 + x_2). \tag{2.47}$$

По (2.46) и (2.47) составляем начальную симплекс-таблицу (табл. 2.15), которая после преобразований сводится к конечной табл. 2.17.

Т а б л и ц а 2.15

СП / БП	Св.ч	x_1	x_2
$\tilde{x}_3$	$\boxed{14}$ 14/3	$\boxed{-2}$ –2/3	3 1/3
$\tilde{x}_4$	8 –14/3	1 2/3	1 $\boxed{-1/3}$
$\tilde{y}$	0 –98/3	2 14/3	7 $\boxed{-7/3}$

Т а б л и ц а 2.16

СП / БП	Св.ч	x_1	$\tilde{x}_3$
x_2	14/3 4/3	–2/3 $\boxed{2/5}$	1/3 –2/15
$\tilde{x}_4$	$\boxed{10/3}$ 2	5/3 3/5	$\boxed{-1/3}$ –1/5
$\tilde{y}$	–98/3 –40/3	20/3 $\boxed{-4}$	–7/3 4/3

Т а б л и ц а 2.17

СП / БП	Св.ч	$\tilde{x}_4$	$\tilde{x}_3$
x_2	6	2/5	1/5
x_1	2	3/5	– 1/5
$\tilde{y}$	– 46	– 4	– 1

Из нее видно, что оптимальное решение ПЗ образуют $x_2^* = 6, x_1^* = 2$, причем $\tilde{y} = -46$, следовательно y = 46. Для получения оптимального решения ДЗ выделяем базисные переменные в конечной таблице (это x_2 и x_1) и записываем строку $\overline{d}^{\,T}$ коэффициентов при этих переменных в строке ЦФ $\tilde{y}$ из начальной таблицы, т.е.

$$\overline{d}^{\,T} = \begin{array}{cc} x_2 & x_1 \\ [\,7 & 2\,]. \end{array}$$

Затем выделяем базисные переменные в начальной таблице – это x_3 и x_4 и записываем матрицу из соответствующих им столбцов конечной таблицы:

$$Q = \begin{bmatrix} x_3 & x_4 \\ 1|5 & 2|5 \\ -1/5 & 3/5 \end{bmatrix}.$$

В результате получаем:

$$u^* = [7 \quad 2]\begin{bmatrix} 1/5 & 2/5 \\ -1/5 & 3/5 \end{bmatrix} = [1 \quad 4],$$

т.е. $u_1^* = 1$ и $u_2^* = 4$.

Связь между решениями задач, образующих двойственную пару, позволяет использовать для решения данной конкретной задачи ЛП переход к двойственной для нее задаче. При этом вычисления, необходимые для решения ДЗ, могут оказаться проще, чем для ПЗ. Отметим, что трудоемкость вычислений при решении задач ЛП в большей степени зависит от числа ограничений, чем от количества переменных. Поэтому, если в ДЗ ограничений меньше, чем в ПЗ, обычно целесообразнее решать ДЗ и полученный результат использовать для отыскания оптимального решения ПЗ.

Пример. Необходимо найти значения переменных u_1, u_2, удовлетворяющих условиям

$$3u_1 + 3u_2 \geq 27; \quad u_1 + 3u_2 \geq 15; \quad u_1 \geq 0;$$
$$2u_1 + u_2 \geq 10; \quad 2u_1 + 4u_2 \geq 28; \quad u_2 \geq 0,$$

при которых ЦФ $z = 2u_1 + 5u_2$ принимает наименьшее значение.

Решение. Данная задача является двойственной по отношению к задаче об отыскании переменных x_1, x_2, x_3, x_4, удовлетворяющих условиям

$$\begin{cases} 3x_1 + 2x_2 + x_3 + 2x_4 \leq 2; \\ 3x_1 + x_2 + 3x_3 + 4x_4 \leq 5; \end{cases} \quad x_i \geq 0,$$
$$y = 27x_1 + 10x_2 + 15x_3 + 28x_4 \to sup.$$

В этой задаче всего два ограничения, в то время как в исходной их четыре. Используя вспомогательные переменные $\tilde{x}_5, \tilde{x}_6$, эту задачу сводим к канонической с ограничителями-равенствами:

$$\tilde{x}_5 = 2 - (3x_1 + 2x_2 + x_3 + 2x_4); \quad \tilde{x}_6 = 5 - (3x_1 + x_2 + 3x_2 + 4x_4).$$

Ее решение симплекс-методом представлено в табл. 2.18 – 2.20.

Таблица 2.18

БП \ СП	Св.ч	x_1	x_2	x_3	x_4
$\tilde{x}_5$	2 1	3 3/2	2 1	1 1/2	2 1/2
$\tilde{x}_6$	5 – 4	3 –6	1 –4	3 –2	4 –2
$\tilde{y}$	0 –28	27 – 42	10 –28	15 –14	28 –14

Таблица 2.19

БП \ СП	Св.ч	x_1	x_2	x_3	$\tilde{x}_5$
x_4	1 –1/2	3/2 3/2	1 3/2	1/2 –1/2	1/2 1
$\tilde{x}_6$	1 1	–3 –3	–3 –3	1 1	–2 –2
$\tilde{y}$	–28 –1	–15 3	–18 3	1 –1	–14 2

Таблица 2.20

БП \ СП	Св.ч	x_1	x_2	$\tilde{x}_6$	$\tilde{x}_5$
x_4	1/2	3	5/2	–1/2	3/2
x_3	1	–3	–3	1	–2
$\tilde{y}$	–29	–12	–15	–1	–12

Используя начальную и конечную таблицы, по сформулированному выше правилу находим решение поставленной задачи:

$$\bar{u}^* = \begin{array}{cc} x_4 & x_3 \\ [28 & 15\,] \end{array} \begin{bmatrix} \tilde{x}_5 & \widetilde{x_6} \\ 3/2 & -1/2 \\ -2 & 1 \end{bmatrix} = [12 \;\; 1]\,;\; u_1^* = 12,\; u_2^* = 1,\; z_{\min} = 29.$$

В заключение отметим, что существует довольно интересная с практической точки зрения экономическая интерпретация двойственности в задачах ЛП. С ней можно ознакомиться, например, в [19, 20, 32].

§ 2.6. Целочисленные задачи линейного программирования

Рассмотрим задачу ЛП, в которой переменные x_i должны удовлетворять еще одному условию – они должны быть целыми. Обычно это требование обусловлено тем, что эти переменные по физическому смыслу не могут быть дробными, например, количество станков, приборов, работников и т.п. Один из методов решения такой задачи – метод отсекающих плоскостей (дробный алгоритм Гомори), который сводится к следующему [32].

1. Решается симплекс-методом поставленная задача без учета требования целочисленности (так называемая задача с ослабленными ограничениями).

2. Анализируются значения переменных в найденном решении. Если все они оказались целыми, то решение данной ЦЗЛП найдено, иначе нужно перейти к следующей операции.

3. Составляется дополнительное условие-ограничение, добавление которого к имеющейся системе ограничений было бы эквивалентно отсечению той части ОДР, в которой находится полученное неподходящее решение и не содержится ни одной точки с целочисленными коэффициентами.

4. Решается полученная задача с дополнительным ограничением, после чего осуществляется переход к п. 2.

Рассмотрим вопрос о том, как следует формировать указанное в п. 3 дополнительное ограничение. Выделим в симплекс-таблице строку, соответствующую базисной переменной x_i, для которой свободный член β_i оказался дробным. Представим его в виде $\beta_i = [\beta_i] + \{\beta_i\}$, где $[\beta_i]$ – наибольшее целое число, не превышающее β_i, а $\{\beta_i\} = \beta_i - [\beta_i]$ – дробная часть числа $(0 \le \{\beta_i\} < 1)$.

Аналогично представим все коэффициенты при свободных переменных в рассматриваемой строке: $\alpha_i = [\alpha_i] + \{\alpha_i\}$.

Тогда в соответствии с [32] дополнительное ограничение можно записать в виде

$$\sum \{\alpha_{ij}\} x_j \geq \{\beta_i\}, \qquad (2.48)$$

где суммирование осуществляется по всем свободным переменным x_j.

Для того чтобы использовать полученное ограничение, необходимо свести его к канонической форме, вводя дополнительную переменную

$$u_1 = -\{\beta_i\} + \sum \{\alpha_{ij}\}\, x_j, \quad u_1 \geq 0,$$

и включить полученное уравнение в симплекс-таблицу.

Таблица 2.21

СП / БП	Свч	x_3	x_4
x_1	7/2	–1/2	1/2
x_2	3	1	0
y	–9/2	–3/2	–1/2

Пример. Найти целые значения переменных $x_i \geq 0\,(i = \overline{1,\ 4})$, при которых ЦФ $y = 5 - x_1 - 2x_2$ принимает наименьшее значение и выполняются условия: $2x_1 + x_2 + x_4 = 10$; $x_2 + x_3 = 3$.

Решение. Сформулированная задача является канонической. Принимая x_3 и x_4 в качестве базисных переменных, легко находим опорное решение: $x_3 = 3 - x_2$; $x_4 = 10 - (2x_1 + x_2)$; $y = 5 - (x_1 + 2x_2)$.

Выполнив симплекс-преобразования, получаем табл. 2.21, соответствующую оптимальному решению без учета целочисленности переменных:

$$x_1 = \beta_1 = 7/2, \quad x_2 = \beta_2 = 3.$$

Таблица 2.22

СП / БП	Св.ч	x_3	x_4	x_5
x_1	7/2 –1/2	–1/2 – 1/2	1/2 $\boxed{-1}$	0 1
x_2	3 0	1 0	0 $\boxed{0}$	0 0
$\tilde{x}_6$	$\boxed{1/2}$ 1	$\boxed{1/2}$ 1	1/2 2	$\boxed{-1}$ –2
$\tilde{y}$	(M–9)/2 (1–M)/2	(M–3)/2 (1–M)/2	(M – 1)/2 (1–M)	–M M–1

Таблица 2.23

СП / БП	Св.ч	x_3	$\tilde{x}_6$	x_5
x_1	3	–1	–1	1
x_2	3	1	0	0
x_4	1	1	2	–2
$\tilde{y}$	– 4	–1	1–M	–1

Как видим, одна из переменных $-x_1$ оказалась дробной. В соответствии с алгоритмом Гомори сформируем дополнительное ограничение. Для этого коэффициенты, стоящие в строке для x_1, представим в виде целой и дробной частей:

$$7/2 = 3 + 1/2; \quad -1/2 = -1 + 1/2; \quad 1/2 = 0 + 1/2$$

и, используя дробные части этих коэффициентов, запишем условие (2.48):

$$\frac{1}{2}x_3 + \frac{1}{2}x_4 \ge \frac{1}{2}$$

или

$$x_5 = \frac{1}{2}x_3 + \frac{1}{2}x_4 - \frac{1}{2}; \quad x_5 \ge 0, \tag{2.49}$$

где x_5 – дополнительная переменная. Так как свободный член в выражении для x_5 отрицателен, то имеющееся распределение переменных на базисные (x_1, x_2, x_5) и свободные (x_3, x_4) не соответствует допустимому опорному плану. Для решения задачи в этой ситуации воспользуемся *М*-методом (см. § 2.4). Для этого преобразуем уравнение (2.49), используя дополнительную переменную $\tilde{x}_6$: $\tilde{x}_6 = \frac{1}{2} - \frac{1}{2}x_3 - \frac{1}{2}x_4 + x_5$ и введем новую целевую функцию

$$\tilde{y} = -\frac{9}{2} + \frac{3}{2}x_3 + \frac{1}{2}x_4 + M\tilde{x}_6 = -\frac{9}{2} + \frac{3}{2}x_3 + \frac{1}{2}x_4 + M\left(\frac{1}{2} - \frac{1}{2}x_3 - \frac{1}{2}x_4 + x_5\right) =$$
$$= \frac{M-9}{2} - \left[\frac{M-3}{2}x_3 + \frac{M-1}{2}x_4 - Mx_5\right],$$

считая *M* большим числом.

Полученные исходные условия представлены в симплекс-таблице (табл. 2.22). В результате ее преобразований получаем табл. 2.23, в которой все коэффициенты при переменных, стоящие в строке $\tilde{y}$, отрицательны, что соответствует решению, при котором функция $\tilde{y}$ принимает наименьшее значение:

$$x_1 = 3, \quad x_2 = 3, \quad x_4 = 1,$$
$$x_3 = 0, \quad x_5 = 0, \quad \tilde{x}_6 = 0.$$

Поскольку все переменные оказались целыми, полученное решение является искомым оптимальным решением, при котором $y_{\inf} = -4$.

§ 2.7. Замечания

1. При использовании симплекс-метода широкое распространение получили две различные формы симплекс-таблиц. Одна из них была описана и использована выше. Правила составления и использования таблицы иного вида см., например, в [1, 14, 32].

2. При решении некоторых задач ЛП могут возникать особые случаи: вырожденность, альтернативные оптимальные решения, неограниченные решения, отсутствие допустимых решений. Они достаточно подробно рассмотрены в [32]. Здесь же изложена методика анализа чувствительности решений задач ЛП к изменению исходных данных.

3. Существуют классы задач ЛП, например транспортные задачи, для решения которых разработаны специальные алгоритмы. Их применение оказывается обычно более эффективным (с точки зрения затрат времени, необходимого объема памяти и т.п.) по сравнению с симплекс-методом. Такие алгоритмы рассмотрены в гл. 3 и 4.

4. Для решения задач целочисленного ЛП часто рекомендуют вместо алгоритма отсечений Гомори использовать метод ветвей и границ. Его сущность описана, например, в [22, 32].

5. В [1] сформулированы задачи параметрического и дробно-рационального ЛП и изложены методики их решения.

6. Как было показано выше, в основе решения различных модификаций задачи ЛП лежит анализ коэффициентов при переменных в выражении для ЦФ и преобразование системы линейных алгебраических уравнений-ограничений. Эти операции по ходу решения приходится выполнять несколько раз. При решении задач «вручную» даже при невысокой их размерности указанные операции оказываются довольно трудоемкими и их выполнение желательно автоматизировать (с помощью ЭВМ). Во многих задачах ЛП, возникающих на практике, особенно при оптимальном планировании производства, содержатся сотни, а нередко и тысячи переменных и ограничений. Ясно, что в таких случаях использование ЭВМ является единственным средством для получения результата.

Контрольные вопросы

1. Как формулируется общая задача линейного программирования в развернутом и кратком векторно-матричном виде?
2. В чем состоит особенность канонической задачи ЛП?
3. Каким образом осуществляется переход от общей ЗЛП к канонической?
4. Что называется областью допустимых решений ЗЛП?
5. Как формулируется условие совместности системы линейных алгебраических уравнений (ограничений-равенств)?
6. Каким образом осуществляется графическое решение задачи ЛП, при каких условиях оно может быть осуществлено?
7. Что называется опорным решением ЗЛП?
8. При каких условиях симплекс-преобразования, осуществляемые при алгебраическом решении ЗЛП, приостанавливаются, а получаемое при этом решение является искомым оптимальным решением?
9. Каким условием определяется выбор разрешающего столбца в симплекс-таблице?
10. Как выбирается разрешающая строка в симплекс-таблице?
11. Как называется элемент симплекс-таблицы, стоящий в разрешающем столбце и разрешающей строке?
12. В чем состоит проблема отыскания начального опорного плана задачи ЛП?
13. В чем состоит сущность М-метода (метода штрафов) для отыскания начального опорного решения?
14. Какие трудности могут возникнуть в ходе выполнения расчетов на ЦВМ при решении ЗЛП методом штрафов?
15. В чем состоит сущность двухэтапного метода решения ЗЛП с отысканием начального опорного решения методом искусственного базиса?
16. Как формулируется двойственная задача ЛП по прямой ЗЛП?
17. Каким образом связаны между собой решения прямой и двойственной задач ЛП?
18. В чем состоит особенность целочисленной задачи ЛП?
19. В чем состоит сущность метода отсекающих плоскостей (дробного алгоритма Гомори) для решения целочисленной ЗЛП?

Задачи для самостоятельного решения

Найти графическим методом решение следующих задач ЛП:

1. $y = x_1 + 3x_2 \to \inf$

$$\begin{cases} x_1 + x_2 \le 12 \\ x_1 - x_2 \le 4 \\ -\frac{1}{2}x_1 + x_2 \le 6 \\ 2x_1 + x_2 \ge 4 \\ x_1, x_2 \ge 0 \end{cases}$$

2. $y = 5 + x_1 - x_2 \to \inf$

$$\begin{cases} x_1 - 3x_2 \le 3 \\ -2x_1 + x_2 \le 2 \\ 0 \le x_1 \le 6 \\ 0 \le x_2 \le 6 \end{cases}$$

3. $y = x_1 + x_2 \to \sup$

$$\begin{cases} -x_1 + x_2 \le 2 \\ x_1 + 2x_2 \le 16 \\ x_1 - 2x_2 \le 4 \\ x_1, x_2 \ge 0 \end{cases}$$

4. $y = x_1 + x_2 \to \sup$

$$\begin{cases} x_1 + x_2 \le 6 \\ 0 \le x_2 \le 6 \\ 2x_1 - x_2 \le 6 \\ x_1 \ge 0 \end{cases}$$

5. $y = x_1 + x_2 \to \sup$

$$\begin{cases} x_1 + 2x_2 \le 14 \\ -5x_1 + 3x_2 \le 15 \\ 4x_1 + 6x_2 \le 24 \\ 2x_1 + x_2 \ge 4 \\ x_1, x_2 \ge 0 \end{cases}$$

6. $y = x_1 + 2x_2 \to \inf$

$$\begin{cases} 4x_1 - 2x_2 \le 12 \\ -x_1 + 3x_2 \le 6 \\ 2x_1 + 4x_2 \ge 16 \\ x_1, x_2 \ge 0 \end{cases}$$

7. $y = x_1 - 2x_2 \to \inf$

$$\begin{cases} x_1 - x_2 \le 1 \\ x_1 + x_2 \ge 2 \\ x_1 - 2x_2 \le 0 \\ x_1, x_2 \ge 0 \end{cases}$$

8. $y = 5x_1 + 3x_2 \to \sup$

$$\begin{cases} 3x_1 + 5x_2 \le 15 \\ 5x_1 + 2x_2 \le 10 \\ x_1, x_2 \ge 0 \end{cases}$$

9. $y = 2x_1 + 3x_2 \to \inf$

$$\begin{cases} 3x_1 + 2x_2 \ge 6 \\ x_1 + 4x_2 \ge 4 \\ x_1, x_2 \ge 0 \end{cases}$$

10. $y = x_1 + x_2 \to \sup$

$$\begin{cases} x_1 + 2x_2 \le 10 \\ x_1 + 2x_2 \ge 2 \\ 2x_1 + x_2 \le 10 \\ x_1, x_2 \ge 0 \end{cases}$$

11. $y = 3x_1 + 4x_2 \to \sup$

$$\begin{cases} -1 \le -x_1 + x_2 \le 1 \\ 4x_1 - x_2 \le 6 \\ x_1 \ge 0 \\ x_2 \le 2 \end{cases}$$

12. $y = 5x_1 + 5x_2 \to \inf$

$$\begin{cases} x_1 - 2x_2 \le 1 \\ x_1 + 2x_2 \le 4 \\ x_1 + x_2 \ge 1 \\ -x_1 + (1/3)x_2 \le 1 \\ x_2 \ge 0 \end{cases}$$

13. $y = x_1 + 2x_2 \to \sup$

$$\begin{cases} x_1 - x_2 \le 1 \\ -2x_1 + x_2 \le 1 \\ 2x_1 + x_2 \ge 1 \\ x_1, x_2 \ge 0 \end{cases}$$

14. $y = 7x_1 + 12x_2 \to \sup$

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

15. $y = -x_1 + 4x_2 + 2x_4 - x_5 \to \sup$

$$\begin{cases} x_1 - 5x_2 + x_3 = 5 \\ -x_1 + x_2 + x_4 = 4 \\ x_1 + x_2 + x_5 = 8 \\ x_1, \dots, x_5 \ge 0 \end{cases}$$

16. $y = x_1 + 2x_3 + x_5 \to \inf$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 5 \\ x_2 + x_3 + x_4 - x_5 = 2 \\ x_3 - x_4 + x_5 = 1 \\ x_i \ge 0, i = \overline{1,5} \end{cases}$$

17. $y = -5x_1 + x_2 - x_3 \to \sup$

$$\begin{cases} 3x_1 - x_2 - x_3 = 4 \\ x_1 - x_2 + x_3 - x_4 = 1 \\ 2x_1 + x_2 + 2x_3 + x_5 = 7 \\ x_i \ge 0, i = \overline{1,5} \end{cases}$$

18. $y = \frac{1}{3} + x_1 + 2x_2 + x_3 - x_4 \to \inf$

$$\begin{cases} -x_1 + 5x_2 + x_3 + x_4 + x_5 = 10 \\ 2x_1 - x_2 + x_3 - 3x_4 = 6 \\ 10x_2 + x_3 + 2x_4 + 3x_5 = 25 \\ x_i \ge 0, i = \overline{1,5} \end{cases}$$

19. $y = 3x_1 - 2x_2 + 3x_4 + x_5 \to \inf$

$$\begin{cases} 3x_1 - 2x_2 - x_3 + x_4 = 2 \\ 4x_1 - x_2 + x_4 + x_5 = 21 \\ 4x_1 - x_2 - x_4 + x_5 = 13 \\ x_1 + x_2 - x_6 = 3 \\ x_i \ge 0, i = \overline{1,6} \end{cases}$$

20. $y = 4x_1 - 3x_2 - x_4 + x_5 \to \inf$

$$\begin{cases} 2x_1 - x_2 - x_3 = -1 \\ x_1 - 3x_2 - x_4 = -13 \\ 4x_1 + x_2 + x_5 = 26 \\ x_1 - 3x_2 + x_6 = 0 \\ x_i \ge 0, i = \overline{1,6} \end{cases}$$

Преобразовать следующие задачи ЛП к канонической форме и решить их симплекс-методом. (При решении всех нижеследующих задач можно использовать ЦВМ с соответствующими программами, реализующими симплекс-метод, в частности программы из [3]).

21. $y = -2x_1 + 4x_2 - 8x_3 \to \inf$

$$\begin{cases} x_1 + x_2 + x_3 \le 10 \\ -x_1 - x_2 + x_3 \le 5 \\ 2x_2 - x_3 \le 5 \\ x_1 \le 1 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

22. $y = 10 - 2x_1 + x_2 - 3x_3 \to \inf$

$$\begin{cases} 3x_1 - x_3 \le 8 \\ -x_1 + x_2 + 4x_3 \le 1 \\ 2x_1 + x_2 - 3x_3 \le 6 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

23. $y = x_1 - 3x_2 + x_3 \to \inf$

$$\begin{cases} 3x_1 - x_2 + 2x_3 \le 1 \\ -2x_1 + 4x_2 \le 12 \\ -4x_1 + 3x_2 + 8x_3 \le 10 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

24. $y = x_2 - 3x_3 - 2x_5 \to \sup$

$$\begin{cases} x_1 + 3x_2 - x_3 + 2x_5 = 7 \\ -2x_1 + 4x_3 + x_4 = 12 \\ -4x_2 + 3x_3 + 8x_5 + x_6 = 10 \\ x_i \ge 0, \quad i = \overline{1,6} \end{cases}$$

25. $y = w \to \inf$

$$\begin{cases} x_{11} + x_{21} = 50 \\ x_{12} + x_{22} = 30 \\ x_{13} + x_{23} = 45 \\ w - 4x_{11} - 10x_{12} - 10x_{13} = z_1 \\ w - 6x_{21} - 8x_{22} - 20x_{23} = z_2 \\ w \ge 0, \quad x_{ij} \ge 0, \quad z_i \ge 0 \quad i, j = 1,2 \end{cases}$$

26. $y = 2x_1 - 6x_2 + 5x_5 \to \sup$

$$\begin{cases} -2x_1 + x_2 + x_3 + x_5 = 20 \\ -x_1 - 2x_2 + x_4 + 3x_5 = 24 \\ -3x_1 - x_2 - 12x_5 + x_6 = 18 \\ x_i \ge 0, \quad i = \overline{1,6} \end{cases}$$

Используя метод искусственного базиса и симплекс-метод, решить следующие задачи ЛП.

27. $y = 2x_1 + x_2 - x_3 - x_4 \to \inf$

$$\begin{cases} x_1 + x_2 + 2x_3 - x_4 = 2 \\ 2x_1 + x_2 - 3x_3 + x_4 = 6 \\ x_1 + x_2 + x_3 + x_4 = 7 \\ x_i \ge 0, \quad i = \overline{1,4} \end{cases}$$

28. $y = -3x_1 + x_2 + 3x_3 - 34x_4 \to \inf$

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 0 \\ 2x_1 - 2x_2 + 3x_3 + 3x_4 = 9 \\ x_1 - x_2 + 2x_3 - x_4 = 6 \\ x_i \ge 0, \quad i = \overline{1,4} \end{cases}$$

29. $y = x_1 + x_2 + x_3 \to \inf$

$$\begin{cases} x_1 - x_4 - 2x_6 = 5 \\ x_2 + 2x_4 - 3x_5 + x_6 = 3 \\ x_3 + 2x_4 - 5x_5 + 6x_6 = 5 \\ x_i \ge 0, \quad i = \overline{1,6} \end{cases}$$

30. $y = x_1 + 2x_2 \to \sup$

$$\begin{cases} 2x_1 + 3x_2 \le 6 \\ 2x_1 + x_2 \le 4 \\ x_1 \le 1, \quad x_{1,2} \ge 0 \\ x_1 - x_2 \ge -1 \\ 2x_1 + x_2 \ge 1 \end{cases}$$

31. $y = x_1 + 2x_2 + x_3 + x_4 + x_5 \to \sup$

$$\begin{cases} x_1 + 4x_2 + 2x_3 + 2x_4 + x_5 = 8 \\ x_1 - 2x_2 + 2x_3 - 2x_4 - x_5 = -6 \\ x_1 + 2x_2 + 2x_4 - x_5 = 2 \\ x_i \ge 0, \quad i = \overline{1,5} \end{cases}$$

32. $y = x_1 + x_2 + x_3 - x_4 - x_5 - x_6 \to \sup$

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + 3x_5 + 3x_6 = 3 \\ x_1 + 2x_2 + 3x_3 + 4x_4 + 6x_5 + 5x_6 = 6 \\ x_1 + 8x_2 + 27x_3 + 64x_4 + 36x_5 + 17x_6 = 36 \\ x_i \ge 0, \quad i = \overline{1,6} \end{cases}$$

33. $y = 100 - 2x_1 + 3x_2 - 6x_3 - x_4 \to \inf$

$$\begin{cases} 2x_1 + x_2 - 2x_3 + x_4 = 24 \\ x_1 + 2x_2 + 4x_3 \le 22 \\ x_1 - x_2 + 2x_3 \ge 10 \\ x_i \ge 0, \quad i = \overline{1,4} \end{cases}$$

34. $y = 2x_1 - x_2 - x_4 \to \inf$

$$\begin{cases} x_1 - 2x_2 + x_3 = 10 \\ -2x_1 - x_2 - 2x_4 \ge 18 \\ 3x_1 + 2x_2 + x_4 \ge 36 \\ x_i \ge 0, \quad i = \overline{1,4} \end{cases}$$

35. $y = 2x_1 - 3x_2 + 6x_3 + x_4 \to \sup$

$$\begin{cases} x_1 + 2x_2 - 4x_3 \le 20 \\ x_1 - x_2 + 2x_3 \ge 10 \\ 2x_1 + x_2 - 2x_3 + x_4 = 24 \\ x_i \ge 0, \quad i = \overline{1,4} \end{cases}$$

36. $y = 10 - 3x_1 + 2x_2 + x_3 \to \inf$

$$\begin{cases} 2x_1 - x_2 + x_3 \le 4 \\ -x_1 + 3x_2 + 3x_3 \ge 14 \\ x_1 + 2x_2 + x_3 = 8 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

Для следующих задач ЛП сформулировать двойственные задачи. Решив одну из них симплекс-методом, получить решение другой.

37. $y = 6x_1 + 8x_2 + x_3 \to \sup$

$$\begin{cases} x_1 + x_2 + x_3 \le 3 \\ x_1 + 2x_2 \le 4 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

38. $y = 27x_1 + 10x_2 + 15x_3 + 28x_4 \to \sup$

$$\begin{cases} 3x_1 + 2x_2 + x_3 + 2x_4 \le 2 \\ 3x_1 + x_2 + 3x_3 + 4x_4 \le 5 \\ x_i \ge 0, \quad i = \overline{1,4} \end{cases}$$

39. $y = 3x_1 + 2x_2 - 6x_3 \to \sup$

$$\begin{cases} 2x_1 - 3x_2 + x_3 \le 18 \\ -3x_1 + 2x_2 - 2x_3 \le 24 \\ x_1 + 3x_2 - 4x_3 \le 36 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

40. $y = 3x_1 - 7x_2 - 4x_3 \to \inf$

$$\begin{cases} -2x_1 - 3x_2 - 2x_3 \le 12 \\ -4x_1 - 4x_2 - 3x_3 \le 24 \\ 5x_1 + 5x_2 + 3x_3 \le 15 \\ x_i \ge 0, \quad i = \overline{1,5} \end{cases}$$

41. $y = 4x_1 + 2x_2 \to \sup$

$$\begin{cases} -x_1 + x_2 \le 6 \\ x_1 + x_2 \le 9 \\ 3x_1 - x_2 \le 15 \\ x_{1,\,2} \ge 0 \end{cases}$$

42. $y = 2x_1 + 3x_2 \to \sup$

$$\begin{cases} -3x_1 + 2x_2 \le 6 \\ x_1 - 4x_2 \le 2 \\ x_1 - x_2 \le 5 \\ x_{1,\,2} \ge 0 \end{cases}$$

43. $y = x_1 - 2x_2 \to \sup$

$$\begin{cases} -3x_1 + 2x_2 \le 6 \\ x_1 - 4x_2 \le 2 \\ x_1 - x_2 \le 5 \\ x_{1,\,2} \ge 0 \end{cases}$$

44. $y = 2x_1 + x_2 \to \inf$

$$\begin{cases} -2x_1 + 3x_2 \le 6 \\ x_1 + x_2 \ge 5 \\ x_i \ge 0 \end{cases}$$

45. $y = 2x_1 + 3x_2 \to \sup$

$$\begin{cases} x_1 + x_2 \le 60 \\ x_1 + 2x_2 \le 80 \\ 0 \le x_1 \le 40 \\ 0 \le x_2 \le 30 \end{cases}$$

46. $y = 8x_1 + 5x_2 + 3x_3 \to \inf$

$$\begin{cases} x_1 + x_2 + x_3 \ge 2 \\ 2x_1 + 2x_2 \ge 1 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

47. $y = 3x_1 + x_2 + 4x_3 \to \inf$

$$\begin{cases} 3x_1 + x_2 + 2x_3 \ge 8 \\ x_1 + 2x_2 + 2x_3 \ge 6 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

48. $y = 2x_1 - x_2 + x_3 \to \sup$

$$\begin{cases} x_1 + x_2 + x_3 \ge 6 \\ 2x_1 - x_2 + x_3 \le 2 \\ x_i \ge 0, \quad i = \overline{1,3} \end{cases}$$

Найти целочисленное решение следующих задач ЛП (все $x_i \ge 0$ и целые).

49. $y = x_1 + 4x_2 \to \sup$

$$\begin{cases} -x_1 + 2x_2 + x_3 = 2 \\ 3x_1 + 2x_2 + x_4 = 6 \end{cases}$$

50. $y = 3x_1 + 4x_2 \to \sup$

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 8 \\ x_1 + 4x_2 + x_4 = 10 \end{cases}$$

51. $y = 3x_1 + 3x_2 \to \sup$

$$\begin{cases} x_1 + 3x_2 \ge 6 \\ 3x_1 + 2x_2 \le 36 \\ x_2 \le 13 \end{cases}$$

52. $y = 3x_1 + 4x_2 \to \sup$

$$\begin{cases} 3x_1 + 2x_2 \le 8 \\ x_1 + 4x_2 \le 10 \end{cases}$$

53. $y = x_1 + x_2 \to \sup$

$$\begin{cases} 3x_1 + 2x_2 \le 5 \\ x_2 \le 2 \end{cases}$$

54. $y = x_1 \to \sup$

$$\begin{cases} x_1 + 3x_2 + x_3 = 12 \\ 3x_1 - 8x_2 + x_4 = 24 \end{cases}$$

55. $y = 3x_1 + x_2 \to \inf$

$$\begin{cases} -4x_1 + x_2 \le 29 \\ 3x_1 - x_2 \le 15 \\ 5x_1 + 2x_2 \ge 38 \end{cases}$$

56. $y = 5x_1 + 7x_2 \to \inf$

$$\begin{cases} -3x_1 + 14x_2 \le 78 \\ 5x_1 - 6x_2 \le 26 \\ x_1 + 4x_2 \ge 25 \end{cases}$$

57. $y = 3x_1 + 2x_2 \to \sup$

$$\begin{cases} x_1 + x_2 + x_3 = 13 \\ x_1 - x_2 + x_4 = 6 \\ -3x_1 + x_2 + x_5 = 9 \end{cases}$$

58. $y = x_1 + 4x_2 \to \sup$

$$\begin{cases} 2x_1 + x_2 \le 19/3 \\ x_1 + 3x_2 \le 4 \end{cases}$$

Глава 3. **ТРАНСПОРТНЫЕ ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ**

§ 3.1. Формулировка классической транспортной задачи (ТЗ)

Среди задач математической экономики значительное место занимают задачи оптимального планирования транспортных перевозок. Существует довольно много различных частных задач такого рода (см., например, в [32, 6]). Они имеют определенные отличия и в содержательной постановке и в математической формализованной модели. Характерно, что, несмотря на эти отличия, большая часть из них относится к задачам линейного программирования с различными особенностями. Поэтому ТЗ можно считать задачами ЛП специального типа.

Рассмотрим следующую ТЗ, которую можно считать классической. Пусть в m пунктах $A_1, A_2, \ldots, A_m$ (на предприятиях или складах) имеется однородный продукт или товар, запасы которого составляют $a_1, a_2, \ldots, a_m$ единиц. Этот продукт должен быть доставлен в n пунктов потребления $B_1, B_2, \ldots, B_n$, например в магазины. Потребности каждого пункта B_j в указанном продукте заданы величинами b_j $(j = 1, \ldots, n)$.

Из каждого пункта A_i возможна транспортировка продукта в любой пункт потребления B_j. Известны связанные с этим транспортные издержки – удельные стоимости c_{ij} транспортировки единицы продукта, например одной тонны груза из A_i в B_j.

Требуется определить план перевозок, при котором запросы всех потребителей продукта были бы удовлетворены, а суммарные транспортные издержки минимальны.

В классической формулировке эта задача рассматривается как сбалансированная – общие запасы продукта в пунктах отправления (ПО) A_i предполагаются равными по величине потребностям в нем в пунктах назначения (ПН) B_j.

Обозначим через x_{ij} количество продукта, вывозимого из ПО A_i в ПН B_j. Тогда задача сводится к отысканию nm неотрицательных величин x_{ij}, при которых общие транспортные затраты на все перевозки, выражаемые величиной

$$y = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij}\,, \tag{3.1}$$

были бы минимальны и при этом для каждого ПО A_i и каждого ПН B_j выполнялись бы условия:

$$\forall i: \; x_{i1} + x_{i2} + ... + x_{in} = a_i \quad (i = 1,...,m), \tag{3.2}$$

$$\forall j: \; x_{1j} + x_{2j} + ... + x_{mj} = b_j \quad (j = 1,...,n). \tag{3.3}$$

Следует отметить, что для сбалансированной задачи выполняется еще одно условие

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j \,. \tag{3.4}$$

Анализируя полученную математическую модель рассматриваемой ТЗ, нетрудно убедиться в том, что входящие в нее целевая функция (3.1) и все $n+m$ ограничений (3.2) и (3.3) являются линейными. Следовательно, эта задача является задачей линейного программирования. При этом ТЗ (3.1 – 3.4) имеет важную особенность по сравнению с общей ЗЛП, которая состоит в том, что коэффициенты при неизвестных x_{ij} в ограничениях (3.2) и (3.3) принимают значения 0 или 1 и каждая переменная входит только в два ограничения.

На первый взгляд может показаться, что по смыслу задачи ограничения в ней должны иметь вид неравенств, в частности для ПО A_i количество вывозимого продукта не должно превышать имеющийся запас a_i, т.е.

$$\sum_{j=1}^{n} x_{ij} \le a_i \,.$$

Однако в действительности из-за введенного условия (3.4) для удовлетворения всех заявок все ресурсы в ПО должны быть исчерпаны и поэтому все ограничения должны быть равенствами.

Условия рассматриваемой ТЗ удобно представить в виде табл. 3.1.

Таблица 3.1

ПН / ПО	B_1	…	B_n	Запасы a_i
A_1	c_{11} x_{11}	…	c_{1n} x_{1n}	a_1
…	…	…	…	…
A_m	c_{m1} x_{m1}	…	c_{mn} x_{mn}	a_m
Заявки b_j	b_1	…	b_n	$\sum a_i = \sum b_j$

Содержащиеся в табл. 3.1 величины a_i, b_j, c_{ij} – заданы, а x_{ij} – подлежащие определению неизвестные. При этом сумма всех x_{ij} по каждой строке должна равняться соответствующей величине a_i, сумма всех x_{ij} по каждому столбцу должна равняться b_j, а сумма произведений $c_{ij}\ x_{ij}$ по всем клеткам таблицы будет выражать суммарные транспортные затраты на перевозки, и эта величина должна быть минимизирована.

§ 3.2. Решение классической транспортной задачи

Как отмечалось, рассматриваемая задача является задачей ЛП, поэтому она обладает свойствами решений ЗЛП (см. гл. 2) и в принципе может быть решена с помощью симплекс-метода. Однако специфика ТЗ дает возможность решать ее специальными более эффективными методами.

Как отмечалось, для решения ЗЛП важно определить ранг системы ограничений. Можно показать [6], что из-за наличия уравнения баланса (3.4) система (3.2 – 3.3), состоящая из $(n+m)$ уравнений, будет иметь ранг $r = m + n - 1$.

Это означает, что при решении ТЗ r переменных должны быть выделены как базисные, а остальные переменные, количество которых равно $k = nm - (m + n - 1) = (m - 1)(n - 1)$, будут свободными. В оптимальном решении они должны быть равны нулю.

Как правило, решение ТЗ как и общей ЗЛП, осуществляется в два этапа. Сначала находится некоторый план перевозок, при котором k величин x_{ij} равно нулю. При этом ЦФ (3.1) пока не рассматривается. На втором этапе полученный план последовательно (итерационно) улучшается. При этом на каждой последующей итерации находится план, имеющий меньшую величину суммарных затрат. Процесс вычисления прекращается, когда получится план, который улучшить невозможно.

Существует несколько методов, реализующих эту концепцию отыскания решения ТЗ [6, 32]. Рассмотрим одну из наиболее известных методик.

§ 3.3. Отыскание начального опорного плана методом северо-западного угла (МСЗУ)

Идею метода рассмотрим на примере ТЗ, представленной в табл. 3.2.

Т а б л и ц а 3.2

ПН / ПО	B_1	B_2	B_3	B_4	B_5	Запасы
A_1	18	27	3			48
A_2			30			30
A_3			9	12	6	27
A_4					20	20
Заявки	18	27	42	12	26	125

В этой задаче $m=4$, $n=5$, $r=m+n-1=8$, $k=(m-1)(n-1)=12$. Будем последовательно удовлетворять заявки ПН B_1, B_2 и т.д., используя для этого последовательно ресурсы, имеющиеся в ПО A_1, A_2 и т.д. При этом будем заполнять клетки таблицы, начиная с левого верхнего угла. Сначала выполним заявку потребителя B_1 в размере 18 ед. за счет A_1. Ресурс ПО – 48 ед. – позволяет это сделать. Переходим к ПН B_2. Его заявку в 27 ед. тоже можно выполнить за счет A_1. Для удовлетворения заявки B_3 в 42 ед. ПО A_1 может выделить только оставшиеся 3 ед., поэтому переходим к A_2. Используя весь его ресурс в 30 ед. полностью заявку ПН B_3 удовлетворить нельзя, поэтому переходим к A_3. Процесс продолжаем до тех пор, пока заявка последнего в списке ПН B_5 в 26 ед. не будет полностью удовлетворена.

Заметим, что при этом все заявки ПН выполнены, а ресурсы ПО исчерпаны, так как задача сбалансирована. При этом $r=8$ клеток оказались заполненными – их можно назвать базисными, а $k=12$ клеток – остались пустыми, для них $x_{ij}=0$ и эти клетки соответствуют свободным переменным в ЗЛП.

Обратим внимание на то, что заполнение таблицы в соответствии с рассматриваемым методом осуществляется с левого верхнего угла. Это объясняет его название – метод северо-западного угла.

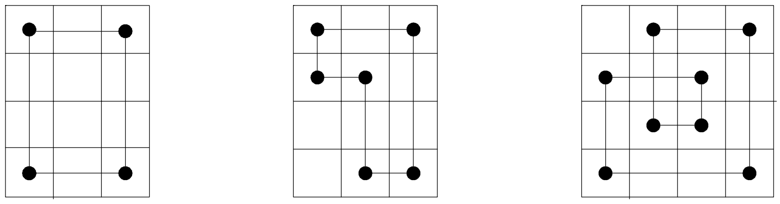
Примером иного подхода к отысканию начального опорного плана в ТЗ может служить метод минимальной стоимости и алгоритм Фогеля [32].

§ 3.4. Улучшение плана перевозок методом потенциалов

Этот метод впервые был предложен советскими учеными Л.В. Канторовичем и М.К. Гавуриным в 1949 году.

Следуя, например [6], введем понятие цикла. Это несколько клеток транспортной таблицы, соединенных замкнутой линией, в которой линии проходят по строкам и столбцам, а повороты осуществляются на 90^0.

Примеры циклов приведены на рис. 3.1.



а) *б*) *в*)

Рис. 3.1

Количество вершин в цикле четное. Вершины цикла снабжаются чередующимися знаками <+> и <–>. При этом отрицательные вершины могут стоять только в базисных клетках, где $x > 0$.

Можно доказать, что для любой свободной клетки транспортной таблицы всегда существует цикл (и при этом единственный), одна из вершин которого лежит в этой свободной клетке, а все остальные вершины находятся в базисных клетках [6].

В пределах цикла осуществляется перераспределение величин планируемых перевозок. При этом в клетках, имеющих знак <+>, x увеличивается на величину некоторую α, а в клетках со знаком <–> x уменьшается на ту же величину. Важно, что при этом перераспределении сохраняется выполнение условий (3.2 – 3.4).

Назовем ценой цикла величину γ, на которую изменяется стоимость транспортировки единицы продукта по этому циклу. Она равна алгебраической сумме удельных стоимостей c_{ij} (у.е./ед.), стоящих в клетках рассматриваемого цикла, взятых с соответствующими знаками.

Для улучшения плана нужно находить циклы с отрицательной ценой. Если такой цикл найден, то по нему следует осуществить перераспределение перевозимого продукта на величину α. При этом общая стоимость перевозок уменьшится на величину $\alpha\gamma$.

Количество продукта α, которое можно перераспределить по сформированному циклу, определяется минимальным значением x из клеток этого цикла с отрицательными вершинами. Например, для цикла, показанного в табл. 3.3, цена цикла равна $\gamma = -5 + 4 - 7 + 6 = -2$ [у.е./ед.]. По этому циклу можно перераспределить $\alpha = \min\{23, 20\} = 20$ единиц продукта. В результате такого перераспределения получается табл. 3.4, для которой транспортные издержки меньше на $\alpha\gamma = 40$ у.е.

Т а б л и ц а 3.3

$c = 5$; $-$; $x = 23$	$c = 4$; $+$
$c = 6$; $+$; $x = 18$	$c = 7$; $-$; $x = 20$

Т а б л и ц а 3.4

$c = 5$; $x = 3$	$c = 4$; $x = 20$
$c = 6$; $x = 38$	$c = 7$

Заметим, что при таком перераспределении не меняется сумма x по строкам и по столбцам. Кроме того, как и при симплекс-преобразованиях для ЗЛП, одна свободная клетка, где $x = 0$, становится базисной, и одна базисная клетка, где было $x > 0$, становится свободной.

Если получена транспортная таблица, в которой не удается найти ни одного цикла с отрицательной ценой, то соответствующий этой таблице план улучшить нельзя, следовательно, он является оптимальным.

Находить циклы с отрицательной ценой или убедиться в их отсутствии помогает ***метод потенциалов***. Для пояснения идеи метода допустим, что каждый ПО A_i вносит перевозчику за транспортировку единицы продукта u_i у.е./ед. и каждый ПН B_j также вносит ему плату υ_j у.е./ед. Тогда за доставку единицы продукта из A_i в B_j перевозчик получит

$$\tilde{c}_{ij} = u_i + \upsilon_j \text{ [у.е.].}$$

Можно доказать [6], что если для всех базисных клеток $\tilde{c}_{ij} = c_{ij}$, а для всех свободных клеток $\tilde{c}_{ij} \le c_{ij}$, то соответствующий план является оптимальным. Если же существует хотя бы одна свободная ячейка, в которой $\tilde{c}_{ij} > c_{ij}$, то план не является оптимальным и его можно улучшить. Для этого следует принять эту клетку за основу для построения цикла с отрицательной ценой и перераспределить по нему соответствующим образом перевозки.

Используя этот результат, можно сформулировать следующий алгоритм решения ТЗ.

1. По исходным условиям находим начальный опорный план, например используя МСЗУ.

2. Для этого плана находим условные платежи u_i и υ_j. При этом формируется, а затем решается система уравнений

$$u_i + \upsilon_j = c_{ij} \tag{3.5}$$

для всех базисных клеток. Таких уравнений получается $r = m + n - 1$, а неизвестных имеется m для u_i и n для υ_j. Значит неизвестных на одно больше, чем уравнений, поэтому одну из них, например u_1, можно взять любой, обычно полагают $u_1 = 0$.

3. Получив из системы (3.5) величины u_i и υ_j, находим псевдостоимости $\tilde{c}_{ij} = u_i + \upsilon_j$ для свободных клеток.

4. Сравниваем полученные значения $\tilde{c}_{ij}$ с имеющимися значениями c_{ij} в каждой из свободных клеток. Если для всех этих клеток $\tilde{c}_{ij} \le c_{ij}$, то план оптимален. Если нашлась свободная клетка, в которой $\tilde{c}_{ij} > c_{ij}$, то план не является оптимальным и его можно улучшить.

5. Для улучшения плана нужно взять клетку, в которой оказалось $\tilde{c}_{ij} > c_{ij}$, и сформировать цикл с отрицательной ценой. По нему следует перераспределить перевозки. Для полученного нового плана перейти к п. 2.

Пример. Рассмотрим ТЗ, исходные данные для которой представлены в табл. 3.5. В ней заданы запасы a_i перевозимого продукта и заявки b_j на его поставки, а также удельные стоимости перевозок c_{ij}, указанные в правом верхнем углу клеток. Используя МСЗУ, найдем начальный опорный план перевозок. Соответствующие значения $x_{ij} \ne 0$ проставлены в левом нижнем углу клеток. Эти клетки являются базисными. Отсутствие чисел в клетках таблицы соответствует $x_{ij} = 0$. Такие клетки являются свободными.

Для каждой базисной клетки запишем уравнение (3.5). В результате получаем:

$$\left.\begin{aligned} u_1 + \upsilon_1 &= 1, \\ u_1 + \upsilon_2 &= 2, \\ u_2 + \upsilon_2 &= 3; \end{aligned}\right\} \qquad \left.\begin{aligned} u_2 + \upsilon_3 &= 1, \\ u_2 + \upsilon_4 &= 5, \\ u_3 + \upsilon_4 &= 4. \end{aligned}\right\}$$

Принимая $u_1 = 0$ из этой системы, последовательно находим:

$$\left.\begin{aligned} \upsilon_1 &= 1, \\ \upsilon_2 &= 2, \\ u_2 &= 1; \end{aligned}\right\} \qquad \left.\begin{aligned} \upsilon_3 &= 0, \\ \upsilon_4 &= 4, \\ u_3 &= 0. \end{aligned}\right\}$$

Эти значения для удобства внесены в дополнительный столбец и дополнительную строку табл. 3.5.

Затем рассматриваем свободные клетки и находим для них $\widetilde{c}_{ij} = u_i + \upsilon_j$, сравнивая эти значения с c_{ij}:

Т а б л и ц а 3.5

ПН / ПО	B_1	B_2	B_3	B_4	Запасы a_i	u_i
A_1	1 / 30	2 / 20 −	4	1 / +	50	0
A_2	2	3 / 10 +	1 / 10	5 / 5 −	25	1
A_3	3	2	4	4 / 15	15	0
Заявки b_j	30	30	10	20	90	
υ_i	1	2	0	4		

$$\widetilde{c}_{13} = 0 < c_{13} = 4; \qquad \widetilde{c}_{14} = 4 > c_{14} = 1;$$

$$\widetilde{c}_{21} = 2 = c_{21}; \qquad \widetilde{c}_{31} = 1 < c_{13} = 3;$$

$$\widetilde{c}_{32} = 2 = c_{32}; \qquad \widetilde{c}_{33} = 0 < c_{33} = 4.$$

Так как среди найденных псевдостоимостей $\widetilde{c}_{ij}$ имеется $\widetilde{c}_{14} > c_{14}$, то план, соответствующий табл. 3.5, не является оптимальным.

Используя свободную клетку (1, 4), формируем цикл, представленный в табл. 3.5. Его цена равна

$$\alpha = +1 - 5 + 3 - 2 = -3.$$

Величина продукта (груза), который следует перераспределить по этому циклу,

$$\gamma = \min\{5;\ 20\} = 5.$$

В результате такого перераспределения получаем новый план, представленный в табл. 3.6. Общая стоимость перевозок по этому плану составляет 180 у.е. Этот результат на $\alpha\gamma = 15$ у.е. лучше, чем стоимость, соответствующая плану из табл. 3.5.

Т а б л и ц а 3.6

ПН / ПО	B_1	B_2	B_3	B_4	Запасы a_i	u_i
A_1	$\boxed{1}$ 30	− $\boxed{2}$ 15	0 $\boxed{4}$	1 + $\boxed{1}$ 5	50	0
A_2	$\boxed{2}$	$\boxed{3}$ 15	$\boxed{1}$ 10	2 $\boxed{5}$	25	1
A_3	4 $\boxed{3}$	5 $\boxed{2}$ +	3 $\boxed{4}$	$\boxed{4}$ 15 −	15	3
Заявки b_j	30	30	10	20	90	
υ_i	1	2	0	1		

Для базисных клеток этой таблицы составляем систему уравнений:

$$\left.\begin{aligned} u_1+\upsilon_1&=1,\\ u_1+\upsilon_2&=2,\\ u_1+\upsilon_4&=1; \end{aligned}\right\} \qquad \left.\begin{aligned} u_2+\upsilon_2&=3,\\ u_2+\upsilon_3&=1,\\ u_3+\upsilon_4&=4. \end{aligned}\right\}$$

Из системы находим:

$$\left.\begin{aligned} u_1&=0,\\ \upsilon_1&=1,\\ \upsilon_2&=2,\\ \upsilon_4&=1; \end{aligned}\right\} \qquad \left.\begin{aligned} u_2&=1,\\ \upsilon_3&=0,\\ u_3&=3. \end{aligned}\right\}$$

Получаемые по этим значениям псевдостоимости $\tilde{c}_{ij}$ проставлены в табл. 3.6 в левом верхнем углу свободных клеток. Сравнивая их с c_{ij}, видим, что в двух клетках (3, 1) и (3, 2) получается $\tilde{c}_{ij} > c_{ij}$.

Примем за основу клетку (3, 2) и для нее сформируем цикл, представленный в табл. 3.6.

Заметим, что его цена $\alpha = +2 - 2 + 1 - 4 = -3$ совпадает с разницей $\tilde{c}_{ij} - c_{ij} = 5 - 2 = 3$. По этому циклу следует перераспределить 15 ед. продукта. В результате получаем табл. 3.7.

Т а б л и ц а 3.7

ПН / ПО	B_1	B_2	B_3	B_4	Запасы a_i	u_i
A_1	1 30	2 0	0 4	1 20	50	0
A_2	2 2	3 15	1 10	2 5	25	1
A_3	1 3	2 15	0 4	1 4	15	0
Заявки b_j	30	30	10	20	90	
υ_i	1	2	0	1		

Для этой таблицы составляем уравнения:

$$\left.\begin{array}{l} u_1+\upsilon_1=1, \\ u_1+\upsilon_2=2, \\ u_1+\upsilon_4=1; \end{array}\right\} \qquad \left.\begin{array}{l} u_2+\upsilon_2=3, \\ u_2+\upsilon_3=1, \\ u_3+\upsilon_2=2. \end{array}\right\}$$

Откуда получаем:

$$\left.\begin{array}{l} u_1=0, \\ \upsilon_1=1, \\ \upsilon_2=2, \\ \upsilon_4=1; \end{array}\right\} \qquad \left.\begin{array}{l} u_2=1, \\ \upsilon_3=0, \\ u_3=0. \end{array}\right\}$$

Во всех клетках табл. 3.7 выполняется условие $\tilde{c}_{ij} \le c_{ij}$, значит соответствующий ей план перевозок является оптимальным.

Для него общая стоимость перевозок $y = y_{\min} = 30 + 20 + 45 + 10 + 30 = 135$ [у.е.].

Замечание 1. В процессе решения ТЗ может возникнуть ситуация, когда при перераспределении перевозок по циклу не одна, а две или более базисных переменных обратятся в нуль. В рассмотренном примере такая ситуация возникает на последнем этапе. Данное решение называется вырожденным. Однако вырожденность не требует никаких специальных мер, в соответствующей базисной (вырожденной) клетке следует проставить 0 в отличие от свободных клеток, где $x = 0$, но по принятым правилам нулевое значение x не проставляется. С нулевыми базисными переменными оперируют точно так же, как и с переменными, имеющими положительное значение [3, 32].

Замечание 2. Если в исходных данных все a_i и b_j целые, то в искомом оптимальном решении все переменные тоже будут целыми [3].

§ 3.5. Неклассические транспортные задачи

В рассмотренной выше формулировке ТЗ предполагается, что во всех ПО сосредоточен однотипный продукт, в котором нуждаются ПН. Из каждого ПО груз перевозится непосредственно в ПН. Суммарные запасы продукта в ПО совпадают с суммарными заявками в ПН. В реальных условиях каждое из этих предположений может не выполняться. Рассмотрим особенности ТЗ в подобных случаях и приемы, позволяющие решать такие задачи.

Несбалансированная транспортная задача

В классической формулировке (3.1 – 3.4) содержалось важное условие (3.4) – предполагалось, что суммарные запасы продукта в ПО совпадают с суммарными заявками в ПН. Однако этот баланс может не выполняться. Получающаяся ТЗ называется несбалансированной [32], ТЗ с неправильным балансом [6] или открытой ТЗ [14].

Рассмотрим ТЗ с избытком запасов, т.е. задачу, в которой

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

В этом случае вместо ТЗ (3.1 – 3.4) получаем следующую задачу.

Найти $x_{ij} \geq 0$, при которых

$$y = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}\, x_{ij} \to \inf ,$$

$$\left.\begin{aligned} &\sum_{j=1}^{n} x_{ij} \leq a_i \qquad && (i = 1,\ldots, m), \\ &\sum_{i=1}^{m} x_{ij} = b_j \qquad && (j = 1,\ldots,n). \end{aligned}\right\}$$

Такую задачу можно свести к классической сбалансированной ТЗ, а затем решать рассмотренными выше методами, если ввести еще один дополнительный, фиктивный пункт назначения B^*, которому будет приписана фиктивная заявка

$$b^* = \sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j .$$

При этом удельные стоимости перевозок из всех ПО в фиктивный ПН полагаются равными нулю [6, 32].

Следует отметить, что при решении такой модифицированной задачи для каждого ПО A_i будет получаться некоторая величина x_i^*, которая должна быть отправлена в фиктивный ПН B^*. Однако в действительности это будет означать, что такое количество груза x_i^* будет оставаться в A_i неотправленным.

Пример несбалансированной ТЗ с избытком запасов дан в табл. 3.8 [6].

Т а б л и ц а 3.8

ПН / ПО	B_1	B_2	B_3	Запасы a_i
A_1	5	7	6	50
A_2	6	6	5	40
A_3	8	4	5	20
Заявки b_j	18	21	33	110 / 72

Благодаря введению фиктивного ПН B^* с заявкой в 38 единиц эта задача сведена к сбалансированной (табл. 3.9). Процесс ее решения представлен в табл. 3.9 – 3.12. Финальную таблицу, соответствующую оптимальному плану, нужно интерпретировать следующим образом:

– из 50 ед. груза, имеющихся в ПО A_1, 18 ед. направляются в ПН B_1, а 32 ед. остаются в A_1 и не перевозятся;

– из 40 ед. груза, имеющихся в ПО A_2, 1 ед. груза направляется в ПН B_2, 33 ед. – в ПН B_3, а 6 ед. остаются не вывезенными;

– все 20 ед. груза, имеющихся в ПО A_3, направляются в ПН B_2.

Т а б л и ц а 3.9

ПН / ПО	B_1	B_2	B_3	$B_ф$	Запасы a_i	Платежи u_i
A_1	5 $\boxed{5}$ 18	7 $\boxed{7}$ 21	6 $\boxed{6}$ 11 −	1 $\boxed{0}$ +	50	0
A_2	4 $\boxed{6}$	6 $\boxed{6}$	5 $\boxed{5}$ 22 +	0 $\boxed{0}$ 18 −	40	-1
A_3	4 $\boxed{8}$	6 $\boxed{4}$	5 $\boxed{5}$	0 $\boxed{0}$ 20	20	-1
Заявки b_j	18	21	33	38	110	
Платежи υ_i	5	7	6	1		

Т а б л и ц а 3.10

ПН / ПО	B_1	B_2	B_3	$B_ф$	Запасы a_i	Платежи u_i
A_1	5 $\boxed{5}$ 18	7 $\boxed{7}$ 21 −	5 $\boxed{6}$	0 $\boxed{0}$ 11 +	50	0
A_2	5 $\boxed{6}$	7 $\boxed{6}$	5 $\boxed{5}$ 33	0 $\boxed{0}$ 7	40	0
A_3	5 $\boxed{8}$	7 $\boxed{4}$ +	5 $\boxed{5}$	0 $\boxed{0}$ 20 −	20	0
Заявки b_j	18	21	33	38	110	
Платежи υ_i	5	7	5	0		

Т а б л и ц а 3.11

ПН / ПО	B_1	B_2	B_3	$B_ф$	Запасы a_i	Платежи u_i
A_1	5 $\boxed{5}$ 18	7 $\boxed{7}$ 1 −	5 $\boxed{6}$	0 $\boxed{0}$ 31 +	50	0
A_2	5 $\boxed{6}$	7 $\boxed{6}$ +	5 $\boxed{5}$ 33	0 $\boxed{0}$ 7 −	40	0
A_3	2 $\boxed{8}$	4 $\boxed{4}$ 20	2 $\boxed{5}$	-3 $\boxed{0}$	20	-3
Заявки b_j	18	21	33	38	110	
Платежи υ_i	5	7	5	0		

Т а б л и ц а 3.12

ПН / ПО	B_1	B_2	B_3	$B_ф$	Запасы a_i	Платежи u_i
A_1	5 $\boxed{5}$ 18	6 $\boxed{7}$	5 $\boxed{6}$	0 $\boxed{0}$ 32	50	0
A_2	5 $\boxed{6}$	6 $\boxed{6}$ 1	5 $\boxed{5}$ 33	0 $\boxed{0}$ 6	40	0
A_3	3 $\boxed{8}$	4 $\boxed{4}$ 20	3 $\boxed{5}$	-2 $\boxed{0}$	20	-2
Заявки b_j	18	21	33	38	110	
Платежи υ_i	5	6	5	0		

В условиях реальной практики дисбаланс в ТЗ может быть вызван и избытком заявок, когда

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

В такой ТЗ имеющихся в ПО запасов недостаточно для удовлетворения всех заявок в ПН. Следуя [6], можно указать два подхода к решению такой задачи.

В первом случае при решении задачи исходят из того, что нужно вывести из ПО все имеющиеся там запасы и минимизировать общую стоимость перевозок, не заботясь о том, в какой мере заявки различных ПН окажутся удовлетворенными. В этом случае для решения задачи вводят фиктивный ПО A^* с запасом:

$$a^* = \sum_{j=1}^{n} b_j - \sum_{i=1}^{m} a_i ,$$

а стоимости перевозок из A^* в каждый ПН полагают равными нулю.

Во втором случае несбалансированную исходную задачу сводят к сбалансированной ТЗ, корректируя заявки ПН. Можно, например, следуя [6], умножить каждую заявку на один и тот же коэффициент:

$$K = \sum_{i=1}^{m} a_i / \sum_{j=1}^{n} b_i .$$

В этом случае потребности каждого ПН будут удовлетворяться в равной доле.

Можно исправлять заявки по-разному в зависимости от важности того или иного ПН. В каждом из указанных выше случаев исходная задача сводится к сбалансированной ТЗ, а затем решается изложенными выше методами.

Многопродуктовая ТЗ

Рассмотрим ситуацию, когда в пунктах отправления A_i сосредоточены различные виды продукции и каждый пункт назначения B_j нуждается в своем ассортименте продукции. Пусть, например, имеется четыре вида продукции $P_1,\dots,P_4$, которые выпускаются на трех заводах (ПО) A_1, A_2, A_3, и необходимо развести эту продукцию в два магазина (ПН) B_1 и B_2 [14, 32].

Информация о выпускаемых видах продукции на предприятиях, имеющихся объемах, а также о потребностях в ней в каждом из магазинов приведена в табл. 3.13.

Удельные затраты на перевозки единицы продукции (1 шт., 1 т и т.п.) представлены в табл. 3.14. Предполагается, что от вида продукции они не зависят.

Требуется составить оптимальный план перевозок, минимизирующий общие затраты.

Таблица 3.13

ПН / ПО	Виды продукции P_1	P_2	P_3	P_4
Заводы				
A_1	-	-	70	30
A_2	50	60	-	40
A_3	80	40	-	-
Магазины				
B_1	70	50	50	60
B_2	60	50	20	10

Таблица 3.14

ПН / ПО	B_1	B_2
A_1	8	21
A_2	10	12
A_3	11	7

Эту задачу можно свести к классической (однопродуктовой) задаче двумя способами.

Можно каждый ПО (завод) рассматривать условно как несколько заводов, выпускающих один вид продукции. Соответственно каждый ПН (магазин) будем заменять на несколько ПН, нуждающихся в одном виде продукции. В результате получим транспортную таблицу, которую называют полной [32] (табл. 3.15).

В этой таблице интерес представляют выделенные клетки, а остальные клетки соответствуют недопустимым связям, так как продукция различных марок не может заменять друг друга. В таких клетках можно проставить заведомо очень высокую стоимость перевозок [32].

Таблица 3.15

ПН / ПО		B_1 P_1	P_2	P_3	P_4	B_2 P_1	P_2	P_3	P_4	Запасы
A_1	P_3									70
	P_4									30
A_2	P_1									50
	P_2									60
	P_4									40
A_3	P_1									80
	P_2									40
Заявки		70	50	50	60	60	50	20	10	

Однако многопродуктовую задачу необязательно описывать одной моделью (полной таблицей). Можно рассмотреть задачу об организации поставок по каждому виду продукции в отдельности. В этом случае для анализируемого примера вместо одной полной таблицы нужно будет использовать четыре частные таблицы (табл. 3.16 – 3.19).

Т а б л и ц а 3.16

Продукция P_1

ПН / ПО	B_1	B_2	
A_2			50
A_3			80
	70	60	

Т а б л и ц а 3.17

Продукция P_2

ПН / ПО	B_1	B_2	
A_2			60
A_3			40
	50	50	

Т а б л и ц а 3.18

Продукция P_3

ПН / ПО	B_1	B_2	
A_1			70
	50	20	

Т а б л и ц а 3.19

Продукция P_4

ПН / ПО	B_1	B_2	
A_1			30
A_2			40
	60	10	

Изложенные приемы позволяют многопродуктовую задачу свести к одной (для полной таблицы) или нескольким (по каждому виду продукции) задачам классического типа, к которым можно применять рассмотренные в § 3.2 – 3.4 методы решения.

ТЗ с промежуточными пунктами

В классической транспортной задаче предполагается, что продукция из ПО A_i по прямому маршруту отправляется в выбранный ПН B_j. На практике может оказаться целесообразной перевозка продукции (полностью или частично) транзитом через другие ПО или ПН. Например, чтобы удовлетворить заявку ПН B_4 в размере b_4 груз отправляется из A_2 сначала в A_5, там он объединяется с имеющимся запасом a_5 и потом объединенный груз отправляется в ПН B_4. Аналогично товар может быть отправлен сначала в один из ПН, а затем передан в другой ПН.

Возможно использование и специальных промежуточных пунктов, например, центров распределения, специальных перевалочных баз и т.п.

Следуя [32], в этих случаях ПН нужно рассматривать как возможные ПО, поэтому в формируемой для решения транспортной таблице нужно список ПО дополнить списком ПН и наоборот. Подробнее см. в [32].

ТЗ с минимизацией времени перевозок

Выше рассматривались задачи, в которых критерием оптимальности была общая стоимость перевозок. Однако в некоторых случаях, например при перевозках скоропортящихся продуктов, может оказаться оправданным планирование перевозок, при которых обеспечивалось бы минимально возможное время окончания всех поставок.

В этом случае для каждой пары (A_i, B_j) задается время t_{ij}, за которое можно осуществить соответствующую перевозку. Предполагается, что оно не зависит от количества перевозимого груза x_{ij}. При этом в модели ТЗ (3.1 – 3.4) критерий оптимальности (3.1) следует заменить на условия

$$T = \max_{x_{ij}>0} t_{ij} \to \inf . \tag{3.6}$$

Этот критерий представляет собой максимальное время из всех клеток транспортной таблицы, содержащих ненулевые перевозки. Сами величины перевозок назначаются так, чтобы удовлетворить все заявки и исчерпать все ресурсы (3.2 – 3.3).

Следует отметить, что величина T, определяемая соотношением (3.6), не является линейной функцией перевозок x_{ij}, поэтому рассматриваемая задача не является задачей линейного программирования.

Методики и примеры решения такой задачи см. в [6, 14].

§ 3.6. Задачи о назначениях и распределительные задачи

Существуют прикладные задачи, которые не связаны с какими-либо перевозками и в этом смысле не являются транспортными. Однако их формализация приводит к модели, совпадающей с транспортной моделью (3.1 – 3.4), хотя и с некоторой спецификой.

Примером такой задачи является так называемая ***задача о назначениях*** [3, 32]. Ее можно сформулировать следующим образом.

Задан перечень работ $A_1,...,A_m$. Для их выполнения можно использовать n станков (или другого оборудования) $B_1,...,B_n$, причем одна работа может выполняться одним станком. Выполнение работы A_i на станке B_j связано с затратами c_{ij}. Если какая либо работа A_i не может быть выполнена на станке B_j, то соответствующую величину c_{ij} будем считать равной достаточно большому числу.

Необходимо так распределить заданный комплекс работ $\{A_1,...,A_m\}$ по имеющемуся парку оборудования $\{B_1,...,B_n\}$, чтобы общие затраты были минимальными.

Для формализации этой задачи введем переменную x_{ij}, полагая:

$$x_{ij} = \begin{cases} 1, & \text{если } i\text{-я работа выполняется на } j\text{-м станке;} \\ 0, & \text{если } i\text{-я работа не выполняется на } j\text{-м станке.} \end{cases}$$

Тогда задача сводится к отысканию $n\,m$ чисел x_{ij}, при которых

$$y = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\, x_{ij} \to \inf \tag{3.7}$$

и выполняются условия

$$\forall i:\ \sum_{j=1}^{n} x_{ij} = 1,\ \ i = 1,...,m, \tag{3.8}$$

$$\forall j:\ \sum_{i=1}^{m} x_{ij} = 1,\ \ j = 1,...,n, \tag{3.9}$$

$$x \in \{0,1\}. \tag{3.10}$$

Эту задачу можно условно интерпретировать как транспортную, если работы рассматривать как исходные пункты или пункты отправления (ПО), а станки (или другое используемое оборудование) как пункты назначения (ПН). В результате решения каждый ПО (работа A_i) должен быть связан с определенным ПН (B_j). Из смысла этой задачи и требования, по которому суммы, стоящие в (3.8; 3.9), должны быть равны 1, следует, что в левых частях каждого из уравнений (3.8 – 3.9) одна какая-то переменная будет равна 1, а остальные – равны 0.

Возможна несколько иная формулировка задачи о назначениях [3]. Имеется m человек (исполнителей) $A_1,...,A_m$ и n заданий $B_1,...,B_n$. У исполнителей разный уровень квалификации, поэтому на выполнение им потребуется разное время: исполнитель A_i может выполнить работу B_j за время c_{ij}. Требуется спланировать распределение исполнителей по заданиям так, чтобы общие затраты времени были минимальными.

Введем переменную x_{ij}, которая будет равна 1, если исполнитель A_i назначен на работу B_j, и примем эту переменную равной 0, если этот исполнитель не будет выполнять работу B_j. В таком случае мы получаем модель (3.7 – 3.10).

Рассматриваемую задачу о назначениях можно представить в виде транспортной табл. 3.20.

Ясно, что при $m \neq n$ задача будет несбалансированной. Изложенными выше способами ее можно искусственно свести к сбалансированной. В этом случае табл. 3.20 будет квадратной размерами $n \times n$ (или $m \times m$). В каждом столбце этой таблицы в одной из клеток x должно быть равно 1, а в остальных клетках $x = 0$, аналогичная ситуация будет в каждой строке. Выбор клеток с $x = 1$ необходимо осуществлять так, чтобы сумма c_{ij} для всех выбираемых клеток оказалась минимальной по сравнению с любой другой комбинацией указанного типа.

Таблица 3.20

ПН / ПО		Задания (работы)				Заявки
		B_1	…	…	B_n	
Исполнители	A_1	c_{11}			c_{1n}	1
	. . .					1
	A_m	c_{m1}			c_{mn}	1
Запасы		1	1	1	1	n / m

Таким образом, для решения сбалансированной задачи необходимо отыскать (выбрать) n клеток в указанной таблице $n \times n$. Может показаться, что эта задача довольно простая и ее можно решить простым перебором различных вариантов. Однако можно показать [3], что при таком подходе количество перебираемых вариантов будет определяться величиной $n!$, и для задачи большой размерности это количество вариантов может оказаться неприемлемо большим даже для современных быстродействующих компьютеров.

Для решения рассматриваемых задач о назначениях можно использовать методы решения транспортных задач. Однако они оказываются довольно неэффективными. Поэтому для решения задач о назначениях были разработаны специальные методы. С их сущностью и примерами применения можно ознакомиться, например в [14, 32].

Более сложные задачи о назначениях (распределительные задачи) рассмотрены в [14].

Контрольные вопросы

1. Как формулируется классическая транспортная задача?
2. В чем состоит особенность классической ТЗ по сравнению с общей задачей ЛП?
3. Как формируется транспортная таблица?
4. Что представляет собой начальный опорный план для ТЗ?
5. Для чего предназначен и в чем состоит метод северо-западного угла?
6. Для чего предназначен метод потенциалов?
7. Что называется циклом в транспортной таблице, как определяется его цена?
8. Как определяется величина продукта (груза), который можно перераспределить в пределах выделенного цикла?
9. При выполнении какого условия решение транспортной задачи, получаемое методом потенциалов, является оптимальным?
10. Каким образом несбалансированную ТЗ можно свести к классической формулировке?

Задачи для самостоятельного решения

Сбалансированные задачи

Дана классическая транспортная задача с тремя ПО A_i и четырьмя ПН B_j. Информация о запасах перевозимого груза в ПО, заявках на него в ПН и удельных стоимостях перевозок представлена в приведенных ниже транспортных таблицах. Для соответствующего варианта найти оптимальный план перевозок.

Т а б л и ц а 1

ПО \ ПН	B_1	B_2	B_3	B_4	a_i
A_1	2	7	1	3	48
A_2	5	6	2	4	23
A_3	6	3	1	2	29
b_j	21	14	37	28	100

Т а б л и ц а 2

ПО \ ПН	B_1	B_2	B_3	B_4	a_i
A_1	5	2	3	4	43
A_2	7	3	1	2	38
A_3	9	3	8	7	19
b_j	15	23	34	28	100

Т а б л и ц а 3

ПО \ ПН	B_1	B_2	B_3	B_4	a_i
A_1	4	3	5	2	33
A_2	5	1	2	7	48
A_3	8	2	1	3	69
b_j	44	42	38	26	150

Т а б л и ц а 4

ПО \ ПН	B_1	B_2	B_3	B_4	a_i
A_1	11	4	7	1	42
A_2	8	6	5	3	29
A_3	9	10	2	4	79
b_j	37	55	28	30	150

Т а б л и ц а 5

ПО \ ПН	B_1	B_2	B_3	B_4	a_i
A_1	8	7	1	4	84
A_2	5	6	3	2	59
A_3	7	4	5	1	57
b_j	48	36	25	91	200

Т а б л и ц а 6

ПО \ ПН	B_1	B_2	B_3	B_4	a_i
A_1	4	9	7	6	64
A_2	5	3	6	4	44
A_3	7	9	2	1	92
b_j	61	57	48	34	200

Т а б л и ц а 7

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	6	4	2	5	21
A_2	8	9	3	7	43
A_3	4	5	1	9	61
b_j	33	28	47	17	125

Т а б л и ц а 8

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	3	2	3	4	55
A_2	7	2	1	9	42
A_3	8	5	2	6	28
b_j	45	38	29	13	125

Т а б л и ц а 9

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	4	2	3	1	57
A_2	8	11	9	6	24
A_3	5	10	4	7	59
b_j	41	34	22	43	140

Т а б л и ц а 10

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	10	13	8	6	41
A_2	11	6	5	9	34
A_3	8	6	4	3	65
b_j	35	21	19	65	140

Т а б л и ц а 11

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	5	11	13	4	27
A_2	6	4	10	8	33
A_3	7	2	3	9	60
b_j	20	32	19	49	120

Т а б л и ц а 12

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	9	10	8	5	37
A_2	7	4	9	3	34
A_3	7	11	13	18	49
b_j	43	21	18	38	120

Таблица 13

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	10	4	8	5	49
A_2	11	5	7	4	32
A_3	6	12	9	1	29
b_j	20	17	31	42	110

Таблица 14

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	6	1	5	4	42
A_2	3	2	7	8	35
A_3	6	9	11	1	33
b_j	18	27	40	25	110

Таблица 15

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	1	2	1	7	61
A_2	3	8	4	5	28
A_3	6	7	8	5	71
b_j	34	43	51	32	160

Таблица 16

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	9	7	8	4	56
A_2	6	8	3	5	52
A_3	4	2	4	3	52
b_j	27	49	43	41	160

Таблица 17

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	5	2	4	6	46
A_2	3	9	8	4	37
A_3	9	5	6	10	87
b_j	41	32	55	42	170

Таблица 18

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	4	5	4	3	62
A_2	2	7	6	5	29
A_3	6	9	7	5	79
b_j	34	39	58	39	170

Таблица 19

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	3	7	8	2	51
A_2	5	6	1	3	77
A_3	4	8	7	6	52
b_j	46	32	54	48	180

Таблица 20

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	6	5	8	4	62
A_2	4	6	3	2	29
A_3	1	9	7	3	79
b_j	34	39	58	39	170

Таблица 21

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	2	1	2	5	54
A_2	6	5	3	2	42
A_3	7	8	4	9	32
b_j	29	27	55	19	130

Таблица 22

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	3	2	7	5	63
A_2	6	3	4	6	41
A_3	5	1	6	9	26
b_j	18	29	17	66	130

Таблица 23

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	5	3	2	6	68
A_2	4	8	4	3	77
A_3	6	5	5	2	45
b_j	48	51	42	49	190

Таблица 24

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	1	1	4	7	52
A_2	5	6	2	3	67
A_3	9	7	4	8	51
b_j	36	49	54	51	190

Таблица 25

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	5	3	2	6	75
A_2	4	7	4	3	61
A_3	2	5	5	1	64
b_j	63	58	44	35	200

Таблица 26

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	2	1	2	5	73
A_2	4	3	2	3	38
A_3	9	5	7	4	89
b_j	47	59	45	49	200

Таблица 27

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	6	4	7	8	27
A_2	5	2	3	5	80
A_3	9	3	4	6	48
b_j	29	33	46	47	155

Таблица 28

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	8	5	9	6	65
A_2	10	7	5	8	54
A_3	5	3	2	9	26
b_j	61	42	25	27	155

Таблица 29

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	4	9	7	8	59
A_2	2	6	1	3	46
A_3	5	4	8	9	70
b_j	56	45	37	37	175

Таблица 30

ПН / ПО	B_1	B_2	B_3	B_4	a_i
A_1	6	4	2	5	71
A_2	9	3	4	7	68
A_3	5	7	9	10	36
b_j	28	39	45	63	175

Несбалансированные задачи

Найти решение несбалансированной ТЗ, представленной соответствующей таблицей из предыдущего задания, если запас груза в ПО A_1 увеличен на 5 ед., а заявки в ПН B_2 уменьшены на 8 ед.

Глава 4. **ЗАДАЧИ ОПТИМИЗАЦИИ, ПРЕДСТАВЛЯЕМЫЕ НА ГРАФАХ**

§ 4.1. Основные понятия теории графов

Рассмотрим некоторое конечное множество X с элементами $x_1, x_2,..., x_n$, а также множество X^2, элементы которого представляют собой всевозможные пары (x_i, x_j), составленные из элементов *X*. Пусть в X^2 выделено некоторое подмножество U. Пару множеств $G = (X, U)$ называют графом. При этом элементы $x_i \in X$ называют вершинами графа, а элементы $u_k = (x_i, x_j) \in U$ – дугами этого графа. Если порядок элементов в парах (x_i, x_j) не принимается во внимание, то граф называется неориентированным. Входящие в него пары называются ребрами. В противном случае, когда граф определяется упорядоченными парами элементов и, следовательно, пары (x_i, x_j) и (x_j, x_i) различны, граф называется ориентированным.

Граф можно представить диаграммой, в которой вершины изображаются точками или кружками, а ребра – отрезками линий, соединяющими соответствующие пары вершин (точек). При этом дуги ориентированного графа изображаются линиями со стрелками, указывающими на порядок элементов в соответствующих парах. На рис. 4.1 приведены примеры неориентированного (*а*) и ориентированного (*б*) графов.

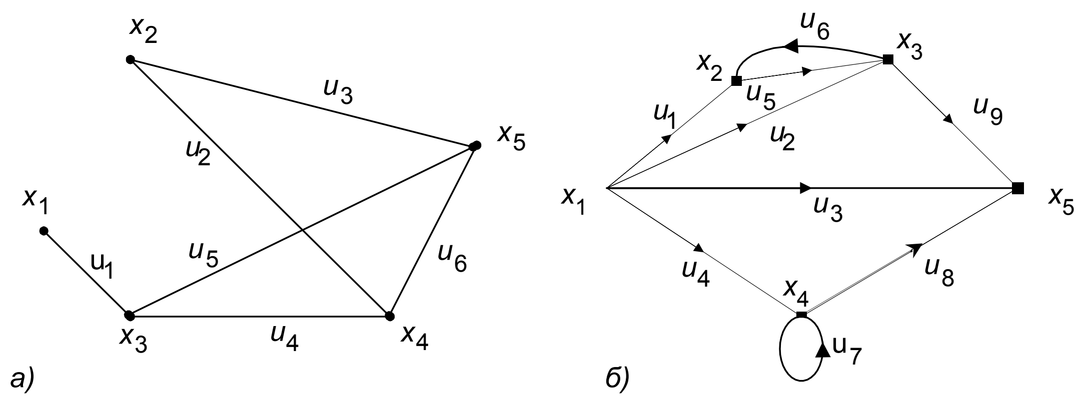


Рис. 4.1

Используя подобные диаграммы, можно приведенному выше теоретико-множественному определению графа дать простую и наглядную интерпретацию: граф – это множество точек, называемых вершинами, и множество линий, соединяющих отдельные пары этих точек и называемых ребрами или дугами.

Вершины x_i и x_j, определяющие ребро $u_k = (x_i, x_j)$, называются концевыми вершинами этого ребра. При этом в ориентированном графе одна из концевых вершин дуги $u_k = (x_i, x_j)$ называется начальной (первая – x_i), а другая $-x_j$ – конечной. Две вершины называются смежными, если они являются концевыми вершинами некоторого ребра. Два ребра называются смежными, если они имеют общую концевую вершину. Ребро называется инцидентным некоторой вершине x_i, если она является концевой точкой этого ребра. Отметим, что концевые вершины ребра в графе не обязательно различны. Если $u_k = (x_i, x_i) \in U$, то ребро u_k называют петлей. Иногда рассматривают графы, в которых имеется более одного ребра с одинаковыми концевыми вершинами. Такие ребра называют параллельными, а соответствующий граф – мультиграфом. Граф, не содержащий параллельных ребер и петель, называется простым.

Используя упомянутые понятия смежности и инцидентности, графы можно задавать в удобной матричной форме. Так, например, любой граф $G = (X, U)$, имеющий n вершин, однозначно определяется матрицей A с элементами a_{ij}, равными числу дуг, идущих от x_i к x_j (при их отсутствии $a_{ij} = 0$). Эта матрица называется матрицей смежности вершин графа. Для графов, изображенных на рис. 4.1, она имеет вид

$$A = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array}, \quad A = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \!\!\begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}.$$

Очевидно, что для неориентированного графа матрица A будет симметрической.

Другие виды матриц, задающих графы, см., например в [28].

При решении многих задач на графах на базе исходных графов образуют новые графы. Широко используются, в частности, следующие разновидности получаемых при этом графов.

Подграф исходного графа $G = (X, U)$ – это граф, в который входит некоторая часть $X_0 \subset X$ вершин графа G вместе с дугами из U, соединяющими эти вершины. Частичным графом по отношению к исходному графу G называют граф, содержащий некоторую часть дуг $U_0 \subset U$ вместе с соответствующими вершинами из X [18].

Для примера на рис. 4.2 приведены G – исходный граф с вершинами $\{x_1, x_2, x_3, x_4\}$ и дугами $\{u_1, u_2, u_3, u_4, u_5, u_6\}$ (см. рис. 4.2, *а*), подграф G_{x_0}, соответствующий подмножеству $X_0 = \{x_1, x_2, x_4\}$ (см. рис. 4.2, *б*), частичный граф G_{U_0}, соответствующий подмножеству дуг $U_0 = \{u_1, u_2, u_5, u_6\}$ (см. рис. 4.2, *в*).

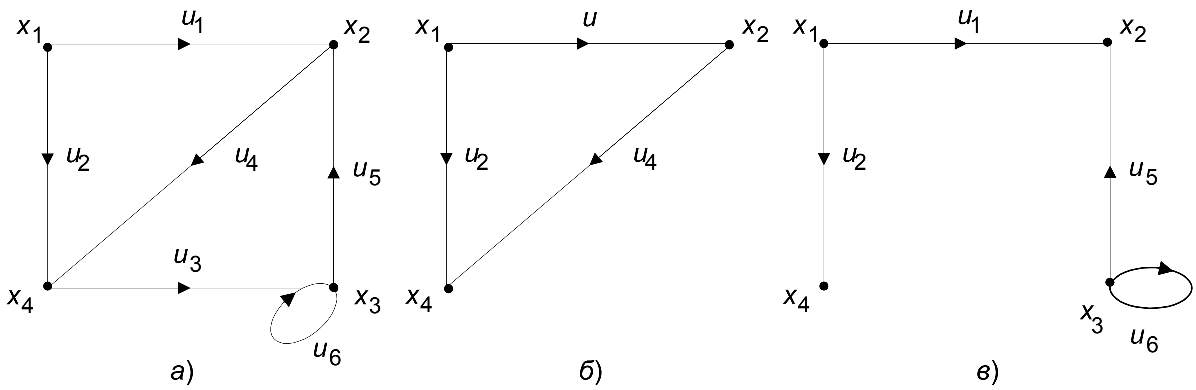


Рис. 4.2

Частичный граф некоторого ориентированного графа G, порождаемый множеством дуг $\{u_1 = (x_s, x_p),\ u_2 = (x_p, x_r),\ u_3 = (x_r, x_q), \ldots,\ u_{n-1} = (x_i, x_h),\ u_k = (x_h, x_t)\}$, образующих последовательность, в которой конец предыдущей дуги совпадает с началом последующей, называется путем в этом графе.

Примером пути в графе G на рис. 4.2, *а* может служить последовательность (u_1, u_4, u_3).

Путь, в котором никакая дуга не встречается дважды, называется простым. Путь, в котором никакая вершина не встречается дважды, называется элементарным. Конечный путь, образованный некоторой последовательностью дуг $(u_1, u_2, \ldots, u_k)$ и проходящий через вершины $(x_s, x_p, x_r, x_q, \ldots, x_f, x_h, x_t)$, называется контуром, если в нем начальная вершина x_s совпадает с конечной x_t. В графе G на рис. 4.2, *а* контур образуют дуги (u_4, u_3, u_5).

Введенным для ориентированных графов понятиям пути и контура в неориентированных графах соответствуют аналогичные понятия цепи и цикла [18].

Важной характеристикой графов является связность. В неориентированном графе она означает возможность соединения любых двух вершин некоторой цепью. Если граф не удовлетворяет этому условию (не является связным), то его всегда можно разбить на связные подграфы, называемые компонентами связности. Аналогично определяют связность ориентированного графа, игнорируя при этом направленность дуг. Более жестким является условие так называемой сильной связности ориентированного графа, предполагающее существование пути, идущего от x_s и x_t для любых вершин $x_s, x_t \in X$ [28].

Рассмотренные выше графы находят широкое применение для математического представления разнообразных, в том числе и экономических систем. Однако понятие графов становится гораздо содержательнее, а сфера их применения намного шире, если вершинам или (и) ребрам графа приписать некоторые количественные или качественные характеристики, называемые весами. В этом случае графы называются взвешенными. В данном разделе рассматривается применение различных типов взвешенных графов для решения оптимизационных задач.

§ 4.2. Задача о кратчайшем пути в графе

Пусть дан некоторый связный ориентированный граф $G = (X, U)$, каждой дуге $u_k = (x_i, x_j) \in U$ которого приписано число (вес) $c_{ij} \geq 0$, называемое обобщенной длиной этой дуги. При этом граф можно характеризовать и задавать матрицей *C*, в которой на пересечении *i*-й строки и *j*-го столбца записывается соответствующее число (длина) c_{ij} при наличии дуги, направленной от x_i к x_j, и записывается ∞ при отсутствии такой дуги.

Пусть в графе заданы две вершины *s* и *t*. Рассмотрим некоторый путь из *s* к *t*. Длиной этого пути называют число, равное суммарной длине дуг, входящих в этот путь. Путей из *s* в *t* может быть несколько. Путь, имеющий наименьшую длину из всех путей, ведущих от *s* к *t*, называется кратчайшим путем из *s* в *t* , и его длина называется расстоянием от *s* до *t*. Отыскание такого пути в заданном графе называется задачей о кратчайшем пути [18, 21, 29, 36].

Эта задача имеет большое практическое значение. К ней, в частности, сводятся многие задачи выбора наиболее экономичного, с точки зрения расстояния либо стоимости, либо времени, маршрута, связывающего два заданных пункта на имеющейся карте дорог. Такую задачу приходится решать, например диспетчеру, при выборе наиболее рационального маршрута доставки нужного товара из складского помещения к заданному потребителю.

Одним из наиболее эффективных методов решения задачи о кратчайшем пути в графе является так называемый метод (алгоритм) Дейкстры [21, 29, 36]. Он состоит в приписывании вершинам графа специальных весов, называемых метками.

Метка $l(x_i)$ произвольной вершины x_i представляет собой число, оценивающее верхнюю границу длины пути от заданной начальной вершины s к вершине x_i. Это означает, что истинная длина пути $d(s, x_i)$ не превышает $l(x_i)$. Алгоритм состоит в изменении меток вершин на каждой итерации по определенному правилу. От итерации к итерации метки уменьшаются либо остаются неизменными. На каждой итерации ровно одна метка принимает значение, уменьшить которое нельзя. Она объявляется постоянной и далее уже не меняется. Эта метка дает точную длину кратчайшего пути от s к вершине с этой меткой. Когда все вершины получат постоянные метки, процесс вычислений заканчивается. Полученные метки вершин дают длины кратчайших путей от заданной вершины s до каждой из остальных вершин графа.

Расстановка и преобразование меток осуществляется следующим образом [21].

Шаг 1 (присвоение меткам начальных значений). Начальной вершине s приписывается метка $l(s) = 0$, и она объявляется постоянной. Всем остальным вершинам $x_i \neq s$ приписываются временные метки $l(x_i) = \infty$. Полагаем $p = s$ и переходим к следующему шагу.

Шаг 2 (обновление меток). Для всех вершин x_i, связанных дугами (p, x_i) с вершиной p и метки которых временные, изменяются метки по следующему правилу: вместо метки $l(x_i)$ записывается метка, равная наименьшему из двух чисел: $l(x_i)$ и $l(p) + c(p, x_i)$. Кратко эту замену метки можно записать следующим образом:

$$l(x_i) \leftarrow \min[l(x_i);\ \ l(p) + c(p, x_i)]. \tag{4.1}$$

Шаг 3 (превращение метки в постоянную). Среди всех вершин x_i с временными метками выбирается такая вершина x_i^*, для которой $l(x_i^*) = \min[l(x_i)]$. Метка этой вершины становится постоянной, а $p = x_i^*$.

Шаг 4. Если p совпадает с заданной конечной вершиной t, то полученная постоянная метка этой вершины $l(t)$ дает длину искомого кратчайшего пути из s в t. Решение закончено (останов). Если же $p \neq t$, то следует перейти к шагу 2.

Замечание 1. Иногда требуется найти кратчайшие пути от заданной вершины s к каждой из остальных вершин графа. В этом случае решение сводится к выполнению всех тех же шагов, кроме последнего, который несколько меняется.

Шаг 4, а (проверка наличия временных меток). Если все вершины имеют постоянные метки, то эти метки дают длины искомых путей. Решение закончено (останов). Если же имеется хотя бы одна временная метка, то следует перейти к шагу 2.

Как указывалось выше, получаемая постоянная метка каждой вершины x_i равна длине кратчайшего пути от s к x_i. Сами пути можно найти, пользуясь соотношением

$$l(x_i') + c(x_i', x_i) = l(x_i), \tag{4.2}$$

где x_i' – вершина, непосредственно предшествующая вершине x_i в кратчайшем пути от s к x_i, а $l(x_i')$ и $l(x_i)$ – полученные постоянные метки этих вершин. Полагая сначала $x_i = t$ и пользуясь (4.2), можно найти предшествующую t некоторую вершину $x_i' = p$. Затем, положив $x_i = p$ и пользуясь (4.2), аналогично отыскивается предшествующая ей некоторая вершина $x_i' = q$ и т.д. до $x_i' = s$.

Аналогично рассмотренной задаче об отыскании кратчайшего пути в ориентированном графе можно сформулировать задачу об отыскании кратчайшей цепи в неориентированном графе, а также в смешанном графе, часть ребер которого ориентирована, а часть неориентирована. В этих случаях практически без изменений можно использовать тот же алгоритм Дейкстры.

Для решения рассматриваемой задачи с помощью алгоритма Дейкстры, особенно для графов с большим количеством вершин и дуг, целесообразно использовать ЦВМ. Описанные выше операции (шаги 1 – 4) легко программируются на соответствующем алгоритмическом языке. Листинг программы, реализующей данный алгоритм, приведен в [36]. При решении задачи «вручную» описанный итерационный процесс изменения меток удобно свести в таблицу.

Пример [21]. Неориентированный граф с девятью вершинами $x_1,...,x_9$ задан матрицей весов, приведенной в табл. 4.1. Необходимо найти кратчайшие пути из вершины $x_1 = s$ в каждую из остальных вершин этого графа.

Т а б л и ц а 4.1

C =

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1		10					3	6	12
x_2	10		18				2		13
x_3		18		25		20			7
x_4			25		5	16	4		
x_5				5		10		23	
x_6			20	16	10		14	15	9
x_7	3	2		4		14			24
x_8	6				23	15			5
x_9	12	13	7			9	24	5	

Примечание 1. Заданный граф является неориентированным, поэтому его матрица является симметрической. Учитывая это, в матрицу C можно было записать элементы только верхней треугольной «половины» матрицы весов. Нижняя половина ей симметрична.

Примечание 2. Незаполненные элементы матрицы C указывают на отсутствие ребер между соответствующими вершинами графа. Как отмечалось, эти элементы можно было принять равными ∞.

Решение. По матрице C строим граф (рис. 4.3), записав рядом с каждым ребром его вес (длину).

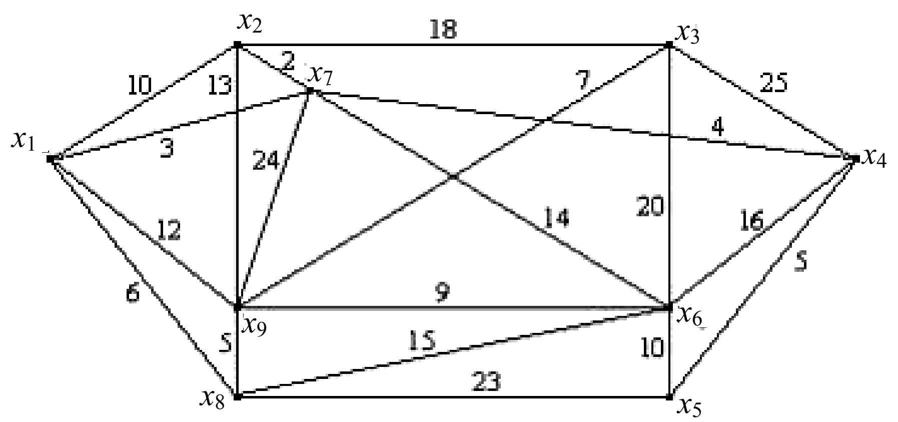


Рис. 4.3

Рассмотрим выполнение операций, предусмотренных алгоритмом Дейкстры, (шагов 1 – 4 а) применительно к данному графу. При этом метки, приписываемые вершинам на каждой итерации, будем записывать в строках табл. 4.2, выделяя постоянные метки прямоугольной рамкой.

Шаг 1. Принимаем $l(x_1) = 0,\quad l(x_2) = l(x_3) = ... = l(x_9) = \infty.$ Заносим эти метки в начальную строку табл. 4.2. Метку $l(x_1)$ выделяем рамкой. Полагаем $p = x_1$.

Т а б л и ц а 4.2

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
$\boxed{0}$	∞	∞	∞	∞	∞	∞	∞	∞
$\boxed{0}$	10	∞	∞	∞	∞	$\boxed{3}$	6	12
$\boxed{0}$	$\boxed{5}$	∞	7	∞	17	$\boxed{3}$	6	12
$\boxed{0}$	$\boxed{5}$	23	7	∞	17	$\boxed{3}$	$\boxed{6}$	12
$\boxed{0}$	$\boxed{5}$	23	$\boxed{7}$	29	17	$\boxed{3}$	$\boxed{6}$	11
$\boxed{0}$	$\boxed{5}$	23	$\boxed{7}$	12	17	$\boxed{3}$	$\boxed{6}$	$\boxed{11}$
$\boxed{0}$	$\boxed{5}$	18	$\boxed{7}$	$\boxed{12}$	17	$\boxed{3}$	$\boxed{6}$	$\boxed{11}$
$\boxed{0}$	$\boxed{5}$	18	$\boxed{7}$	$\boxed{12}$	$\boxed{17}$	$\boxed{3}$	$\boxed{6}$	$\boxed{11}$
$\boxed{0}$	$\boxed{5}$	$\boxed{18}$	$\boxed{7}$	$\boxed{12}$	$\boxed{17}$	$\boxed{3}$	$\boxed{6}$	$\boxed{11}$

Первая итерация

Шаг 2. Рассматриваем вершины, связанные ребрами с $p = x_1$ и имеющие временные метки. Это вершины x_2, x_7, x_9, x_8. Их метки равны ∞. Возьмем для начала x_2. Определим для этой вершины новую метку по правилу (4.1), в котором $l(x_i) = l(x_2) = \infty,\ l(p) = l(x_1) = 0, c(p, x_i) =$ $= c(x_1, x_2) = 10,\ l(x_2) = \min[\infty, 0 + 10] = 10.$

Аналогично получаем: $l(x_7) = 3,\ l(x_9) = 12,\ l(x_8) = 6.$

Заполняем вторую строку табл. 4.2, вписывая в нее найденные новые метки вершин x_2, x_7, x_8, x_9 и переписывая метки остальных вершин без изменения.

Шаг 3. Из всех временных меток выбираем наименьшую и объявляем ее постоянной. Такой меткой оказывается $l(x_7)=3$. Выделяем ее рамкой во второй строке табл. 4.2. Полагаем $p=x_7$.

Шаг 4. Поскольку имеются вершины с временными метками, переходим к шагу 2.

Вторая итерация

Шаг 2. Рассматриваем вершины, связанные ребрами с вершиной $p=x_7$ и имеющие временные метки. Это вершины x_2, x_9, x_6, x_4. Для вершины x_2, имеющей метку $l_{\text{ст}}(x_2)=10$, находим новую метку

$l_{\text{нов}}(x_2)=\min[l_{\text{ст}}(x_2);\ l(x_2)+c(x_1,x_2)]=\min(10,\ 3+2)=5.$

Аналогично получаем: $l(x_4)=7;\ l(x_9)=12;\ l(x_6)=17.$

Полученные новые метки записываем в 3-ю строку табл. 4.2, а метки остальных вершин переписываем без изменения.

Шаг 3. Из временных меток этой строки табл. 4.2 выбираем наименьшую и объявляем ее постоянной. Это метка $l(x_2)=5$. Отмечаем ее рамкой в 3-й строке табл. 4.2, полагаем $p=x_2$.

Шаг 4. Анализируя 3-ю строку табл. 4.2, убеждаемся в наличии временных меток, далее переходим к шагу 2 и выполняем следующую (3-ю) итерацию.

Продолжая этот процесс, будем последовательно заполнять табл. 4.2 новыми строками, отмечая каждый раз по одной новой постоянной метке до тех пор, пока не будет получена строка, все элементы которой отмечены как постоянные метки. Полученные в последней строке табл. 4.2 числа равны длинам кратчайших путей из x_1 в соответствующие вершины x_i. Сами пути можно получить, пользуясь соотношением (4.2).

Так, для вершины $t=x_5$ длина кратчайшего пути от $s=x_1$ оказалась равной 12. Этой вершине могут предшествовать вершины x_4, x_6, x_8 (см. рис. 4.3). Используя их постоянные метки (из последней строки табл. 4.2) и длины ребер, связывающие их с x_5, найдем левую часть соотношения (4.2) для каждой из вершин:

$$l(x_4)+c(x_4,x_5)=7+5=12;$$

$$l(x_6)+c(x_6,x_5)=17+10=27;$$

$$l(x_8)+c(x_8,x_5)=6+23=29.$$

Сравнивая полученные результаты с $l(x_5)=12$, видим, что в кратчайшем пути вершине x_5 предшествует x_4.

Аналогично находим, что x_4 предшествует x_7, а x_7 предшествует $x_1 = s$. Таким образом, кратчайший путь от x_1 к x_5 проходит через вершины x_7 и x_4.

Запишем это в виде $x_1 \xrightarrow{x_7, x_4} x_5$; $l = 12$.

Аналогично получаем $x_1 \xrightarrow{x_7} x_2$; $l = 5$; $x_1 \xrightarrow{x_8, x_9} x_3$; $l = 18$;

$x_1 \xrightarrow{x_7} x_4$; $l = 7$; $x_1 \xrightarrow{x_7} x_6$; $l = 17$; $x_1 \to x_7$; $l = 3$; $x_1 \to x_8$; $l = 6$;

$x_1 \xrightarrow{x_8} x_9$, $l = 11$.

Оформление решения

Для данной задачи в отчете необходимо привести ее формулировку и исходную матрицу весов с указанием заданной начальной вершины графа. Затем следует начертить соответствующий этой матрице граф. Процесс решения задачи оформляется в виде табл. 4.2. Результат решения записывается в виде последовательностей вершин, входящих в каждый из найденных кратчайших путей с указанием длины соответствующего пути.

§ 4.3. Задача о критическом пути в графе

Постановка задачи. Дан ориентированный граф, в котором отсутствуют контуры и каждой дуге (x_i, x_j) приписан вес $c_{ij} \geq 0$. Требуется найти путь от заданной начальной вершины *s* к заданной конечной вершине *t*, имеющий наибольшую длину, равную сумме весов дуг, входящих в этот путь. Такой путь называется критическим [21, 25].

К такой задаче сводится, в частности, задача о планировании работ по реализации некоторого проекта либо задания, предполагающего выполнение определенного набора операций (работ), часть из которых должна вестись последовательно, а часть может выполняться параллельно. В качестве примера можно привести задание на изготовление прибора. Оно разбивается на следующие операции: получение исходной документации, приобретение комплектующих изделий, изготовление нужного набора печатных плат, монтаж соответствующих узлов, их испытание, сборка блоков из узлов, сборка прибора, его испытание, оформление выходной документации, передача прибора и документации заказчику.

Совокупность операций, или работ, необходимых для выполнения задания, можно представить в виде ориентированного графа, в котором дуги изображают соответствующие операции с учетом их последовательности, а вершины – события. Говорят, что событие произошло (например, блок № 2 изготовлен), если все операции (например, изготовление и испытание узлов, проведение монтажа и т.д.), которые отображаются дугами, входящими в соответствующую этому событию вершину x_i, полностью завершены. Подобный граф, изображающий отношения предшествования между операциями проекта, называют сетевым графиком [25]. В качестве весовых характеристик c_{ij} дуг обычно рассматривают время, необходимое для выполнения изображаемых этими дугами операций. В этом случае сформулированная выше задача о критическом пути в графе приобретает следующий смысл – определить время, за которое будет выполнено задание при выбранном порядке операций, а также выявить те операции, которые определяют это время и не позволяют его сократить (критические операции).

Отметим, что задача о критическом пути сходна с задачей о кратчайшем пути в графе, поэтому ее можно решать, используя рассмотренный выше алгоритм Дейкстры, заменив в нем все операции min на max [21].

Однако специальная структура сетевого графика позволяет находить решение с помощью более простого алгоритма [21, 25]. Его можно представить в виде трех стадий.

Стадия А. Она заключается в специальной нумерации вершин графа, при которой каждая дуга ориентирована от вершины с меньшим номером i к вершине с большим номером j. При этом начальное событие получает номер 1, затем следующий номер присваивается любому неперенумерованному событию, для которого все предшествующие события перенумерованы (такое событие всегда существует благодаря отсутствию контуров). Эта операция повторяется до тех пор, пока все события не будут перенумерованы. При этом конечное событие всегда получит последний (наибольший) номер.

Стадия Б. Она состоит в последовательной расстановке меток перенумерованных вершин по следующему правилу. Вершине x_j приписывается метка $l(x_j)$, определяемая соотношением

$$l(x_j) = \max_{x_i} [l(x_i) + c_{ij}], \tag{4.3}$$

где максимум ищется по всем вершинам x_i, от которых идут дуги в x_j. Вслед за x_j метка по правилу (4.3) присваивается вершине, номер которой на единицу больше. Процесс расстановки меток начинается с x_1 (полагают $l(x_1) = 0$), а заканчивается тогда, когда последняя вершина $x_n = t$ получит соответствующую метку. Можно показать, что метка $l(x_j)$ любой вершины x_j, включая и x_n, равна длине самого длинного пути от x_1 до x_j.

Стадия В. Она состоит в отыскании дуг, образующих критический путь. Для этого используется условие, по которому дуга (x_i, x_j) входит в этот путь тогда и только тогда, когда

$$l(x_j) = l(x_i) + c_{ij}, \tag{4.4}$$

где $l(x_i)$ и $l(x_j)$ – метки, расставленные на стадии Б.

Процесс включения дуг в критический путь начинается с конечной вершины x_n и продолжается до тех пор, пока не будет достигнута начальная вершина x_1.

Пример [21]. Для графа, заданного матрицей весов, приведенной в табл. 4.3, нужно найти критический путь от x_1 к x_{13}.

Решение. Строим граф, соответствующий заданной матрице (рис. 4.4, *а*). Затем меняем нумерацию вершин, при этом во избежание путаницы изменяем обозначения – вместо x_k пишем y_m. Получаем граф, приведенный на рис. 4.4, *б*.

Таблица 4.3

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}
x_1		10				19	7	5					
x_2			30	26		11							
x_3												19	12
x_4			15										
x_5				13								17	
x_6				3									
x_7					11								
x_8					8		16		14		9		
x_9										18		21	
x_{10}													8
x_{11}										30			
x_{12}													14
x_{13}													

Затем, начиная с $x_1 = y_1$, расставляем последовательно метки по правилу (4.3), записывая их рядом с вершинами, например в кружках. Так, для вершины y_6 рассматриваем две дуги от y_2 и y_3 и, используя полученные ранее для них метки $l(y_2) = 5$ и $l(y_3) = 21$, получаем $l(y_6) = \max\,[5 + 8,\ 21 + 11] = 32$. Определив метку для y_{13}, расстановку меток прекращаем и, пользуясь правилом (4.4), находим критический путь. Он проходит через вершины $y_1, y_2, y_3, y_6, y_{10}, y_{11}, y_{12}, y_{13}$.

Оформление решения. Необходимо дать формулировку решаемой задачи и привести исходную числовую матрицу, определяющую сетевой график. По ней следует построить исходный граф, перенумеровать его вершины и расставить метки. С их помощью нужно выделить (например особым цветом) критический путь.

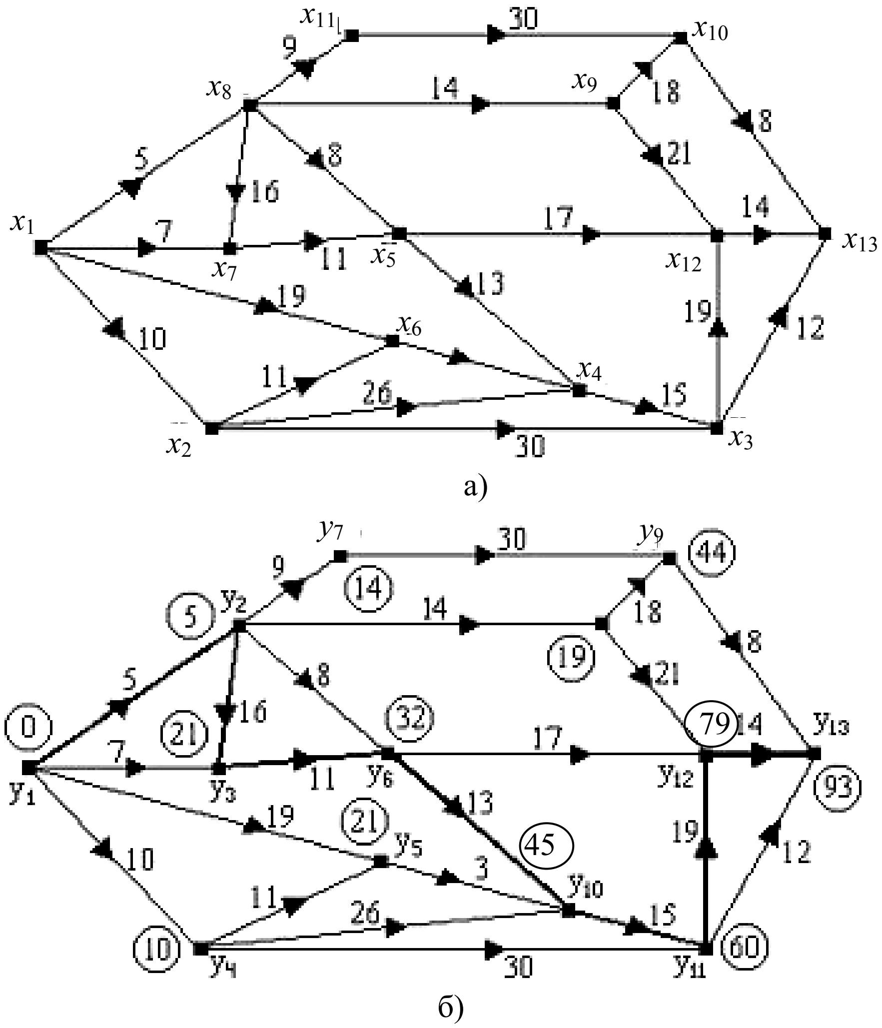


Рис. 4.4

Замечание. Если вес дуги c_{ij} в графе означает время выполнения операции, изображаемой дугой (x_i, x_j), то метка $l(x_i)$ определяет срок, раньше которого не может наступить событие x_j. При этом $l(x_n)$ дает наиболее ранний срок реализации всего задания (проекта). Аналогичным образом решается задача об определении наиболее позднего срока наступления событий, еще допускающего своевременное окончание всего проекта [25]. В этом случае в расчет закладывается заданное время L^0 завершения проекта (оно должно удовлетворять условию $L^0 \geq l(x_n)$. Для решения этой задачи нужно перенумеровать вершины (события) так, чтобы для каждой дуги (операции) (x_i, x_j) выполнялось условие $i < j$, а затем последовательно расставить метки $L(x_j)$ вершин. При этом сначала полагают $L(x_n) = L^0$, после чего для $x_{n-1}, x_{n-2}, ..., x_1$ вычисляют

$$L(x_i) = \min_{x_j}[L(x_j) - c_{ij}].$$

Пример расчета, а также дополнительные пояснения к изложенной задаче и ее решению см. в [25]. С другими задачами планирования и управления проектами можно познакомиться в [25, 36].

§ 4.4. Задача о графе минимальной длины

Постановка задачи. Дан неориентированный связный взвешенный граф, имеющий *n* вершин. Нужно построить новый связный граф, который содержал бы все вершины исходного графа и имел бы наименьшую сумму весов $c_{ij} \geq 0$, входящих в него ребер. К такой задаче сводится, например, задача о проектировании сети трубопроводов или электрической сети, содержащей заданное множество пунктов и имеющей при этом наименьшую суммарную длину труб либо проводов.

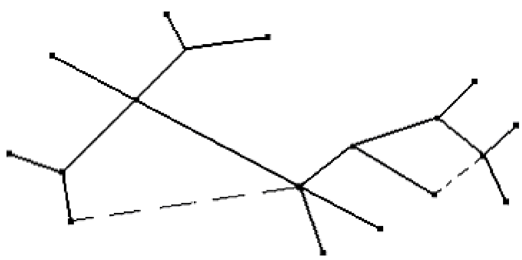


Рис. 4.5

Искомый граф, называемый графом наименьшей длины, должен быть деревом, т.е. не содержать циклов [18]. Иначе из графа можно было бы удалить некоторые «лишние» ребра, образующие циклы (например ребра, изображенные пунктиром на рис. 4.5); суммарный вес ребер при этом уменьшился бы, а связность графа не нарушилась.

Отметим, что дерево, содержащее все заданные вершины исходного графа, называют его остовным деревом (остовом) [21], или покрывающим деревом [25]. Можно показать, что остов, построенный на *n* вершинах, содержит (*n* – 1) ребер, а всего для полного графа можно построить n^{n-2} остовов [21]. Ясно, что с ростом *n* их количество стремительно возрастает.

Сформулированная задача состоит в том, чтобы из всех остовов выбрать кратчайший остов, соответствующий заданной матрице весов.

Для решения поставленной задачи можно использовать следующий алгоритм [18].

Шаг 1. Из всех ребер исходного графа выбираем ребро с минимальным весом. Принимаем его за начало строящегося графа.

Шаг 2. Рассматриваем ребра, инцидентные вершинам, вошедшим в полученный граф.

Шаг 3. Из этих ребер выбираем ребро, имеющее наименьший вес и присоединение которого к строящемуся графу не приведет к образованию цикла. Если при этом имеется несколько ребер с одинаковым наименьшим весом, то берем любое из них.

Шаг 4. Выбранное ребро присоединяем к строящемуся графу.

Шаг 5. Проверяем, имеются ли вершины, не вошедшие в полученный граф. Если таких вершин нет, то решение закончено (останов), полученный граф является искомым. Иначе переходим к шагу 2.

Определенный этим алгоритмом процесс образования искомого дерева можно вести непосредственно по весовой матрице, задающей исходный граф. Проиллюстрируем это на примере.

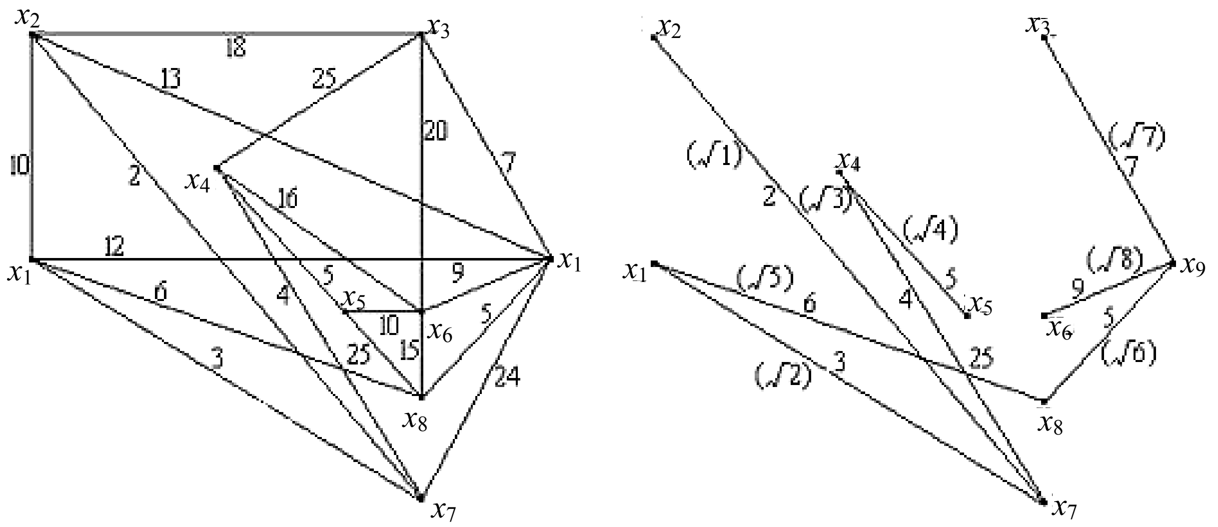
Пример. В табл. 4.4 приведена верхняя треугольная «половина» симметрической весовой матрицы, задающей исходный граф (рис. 4.6, *а*). Как и прежде отсутствие в этой матрице элемента на пересечении *i*-й строки и *j*-го столбца означает отсутствие ребра (связи) между x_i и x_j (или иначе $c_{ij=\infty}$). Требуется найти граф наименьшей длины.

Таблица 4.4

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
x_1		10					3	6	12
x_2			18				2		13
x_3						20			7
x_4					5	16	4		
x_5						10		23	
x_6							14	15	9
x_7									24
x_8									5
x_9									

Решение. Шаг 1. Выбираем в табл. 4.4 наименьший элемент (обводим его в кружок). Это $c_{27} = 2$, ему соответствует ребро с весом 2, соединяющее x_2 и x_7. Изобразим это ребро на рис. 4.6, *б*, на котором будем строить искомый граф.

Шаг 2. Выделяем (например вычеркивая карандашом) 2-й, 7-й столбцы и 2-ю, 7-ю строки в табл. 4.4 и рассматриваем элементы, стоящие в них. Они соответствуют ребрам, инцидентным x_2 и x_7.



а) *б*)

Рис. 4.6

Шаг 3. Ищем среди элементов, стоящих в выделенных строках и столбцах, наименьший элемент, не рассматривая при этом выделенный (в кружке) элемент. Таким элементом в табл. 4.4 является $c_{17} = 3$.

Шаг 4. На рис. 4.6, *б* к ребру, соединяющему x_2 и x_7, добавляем ребро, связывающее x_1 и x_7.

Шаг 5. Убеждаемся в наличии вершин, не вошедших пока в строящийся граф. Поэтому продолжаем построение, выполняя операции, предусмотренные шагами 2 – 5.

При выборе наименьшего элемента (шаг 3) следует рассматривать все элементы, стоящие в уже выделенных на предыдущих шагах строках и столбцах, за исключением элементов, обведенных кружками, и элементов, стоящих на пересечении выделенных строк и столбцов (они соответствуют ребрам, включение которых в строящийся граф привело бы к образованию цикла).

После того как $(n-1)$ ребро включено в строящийся граф, получаем искомое дерево, содержащее все заданные вершины. Оно приведено на рис. 4.6, *б*. В нем около каждого ребра для пояснения, кроме веса, в скобках указан порядковый номер, под которым это ребро вошло в искомый граф. Завершению построения данного графа соответствует табл. 4.4, в которой все строки и столбцы оказываются вычеркнутыми.

Оформление решения. Оно состоит в формулировке решаемой задачи, записи заданной матрицы весов подобной табл. 4.4, выполнении операций над элементами этой матрицы, вычеркивании строк и столбцов, а также выделении соответствующих элементов в этой таблице и параллельном построении искомого графа с указанием веса соответствующих ребер. Необходимо также подсчитать суммарный вес полученного графа.

§ 4.5. Задача о максимальном потоке в графе (сети)

Постановка задачи. Пусть $G=(X,U)$ – связный ориентированный граф. Для каждой его вершины $x_i \in X$ выделим множество $U^+_{x_i} \subset U$ всех дуг (x_i, x_j), выходящих из x_i; соответствующее множество всех вершин x_j, связанных с x_i дугами (x_i, x_j), обозначим $\Gamma(x_i)$. Аналогично выделим множество $U^-_{x_i}$ всех дуг (x_k, x_i), входящих в x_i, а соответствующее множество вершин x_k обозначим $\Gamma^{-1}(x_i)$. Так, для графа, изображенного на рис. 4.7, вершине x_3 соответствуют множества:

$$\Gamma(x_3)=(x_4,x_6) \quad \text{и} \quad \Gamma^{-1}(x_3)=\{x_1,x_2,x_5\};$$

$$U^+_{x_3}=\{(x_3,x_4),(x_3,x_6)\} \quad \text{и} \quad U^-_{x_3}=\{(x_2,x_3)(x_1,x_3)(x_5,x_3)\}.$$

Этот граф называют транспортной сетью [18], или просто сетью [36], если:

1) в этом графе нет петель;

2) в нем существует одна и только одна вершина $s \in X$, не имеющая входящих в нее дуг $(\Gamma^{-1}(s)=\varnothing)$, эту вершину называют входом сети, или истоком;

3) в нем существует одна и только одна вершина $t \in X$, не имеющая выходящих из нее дуг $(\Gamma(t)=\varnothing)$, эту вершину называют выходом сети, или стоком;

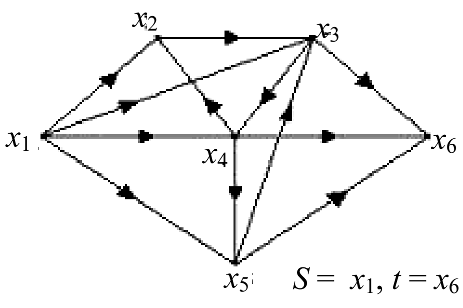


Рис. 4.7

4) каждой дуге $(x_i, x_j) \in U$ приписано число $c_{ij} \geq 0$, называемое пропускной способностью этой дуги;

5) каждой дуге приписано число p_{ij}, называемое потоком этой дуги и удовлетворяющее условиям:

а) $0 \le p_{ij} \le c_{ij}$;

б) $\forall x_i \ne s,t \quad \sum\limits_{x_j \in \Gamma(x_i)} p_{ij} = \sum\limits_{x_k \in \Gamma^{-1}(x_i)} p_{kj}$,

т.е. для каждой вершины, кроме стока и истока, сумма потоков, входящих в нее дуг, равна сумме потоков, выходящих из нее дуг;

в) $\sum\limits_{x_j \in \Gamma(s)} p_{sj} = \sum\limits_{x_k \in \Gamma^{-1}(t)} p_{kj} = P$,

т.е. суммарный поток, выходящий из истока *s,* равен суммарному потоку, входящему в сток *t*; при этом число *P* называется величиной потока сети.

Требуется найти такие значения потоков p_{ij}, при которых величина потока сети *P* была бы наибольшей для заданного графа и заданных пропускных способностей. К такой задаче сводятся многие прикладные задачи о распределении грузов по автомобильной, железнодорожной или другой сети, задачи, возникающие при планировании поставок и др. Примеры указанных задач см. в [36].

Решение. Одним из наиболее эффективных методов решения поставленной задачи является метод Форда и Фалкерсона, обоснование которого см., например, в [18, 21, 36]. Его суть состоит в выполнении итерационной процедуры, которая за конечное число шагов приводит к искомому результату. В ходе решения, осуществляемого по этому алгоритму, на каждой итерации находят некоторое распределение потока по дугам, причем от итерации к итерации величина потока сети увеличивается. Эти итерации выполняются до тех пор, пока дальнейшее увеличение потока станет невозможным из-за ограничений, определяемых пропускными способностями дуг.

Алгоритм использует вспомогательные метки вершин, которые указывают, вдоль каких дуг можно увеличить поток и насколько. Этот алгоритм можно представить в виде двух стадий, разбивающихся на несколько шагов [21].

Стадия А – расстановка меток.

Метка произвольной вершины x_i состоит из двух частей и имеет один из двух видов

$$(+x_j,\delta) \text{ или } (-x_j,\delta).$$

Часть x_j в $(+x_j,\delta)$ означает, что поток можно увеличить вдоль дуги (x_j,x_i). Часть $-x_j$ в $(-x_j,\delta)$ означает, что поток может быть уменьшен вдоль дуги (x_i, x_j). В обоих случаях δ задает максимальную величину дополнительного потока, который может проходить от истока s к вершине x_i.

Различают три возможных состояния вершин:

1) вершине приписана метка, и вершина просмотрена, т.е. все смежные с ней вершины снабжены соответствующими метками;

2) метка вершине приписана, но вершина не просмотрена, т.е. не все смежные с ней вершины получили метки;

3) вершина не имеет метки.

Сначала все вершины не имеют меток и все $p_{ij} = 0$.

Шаг 1. Вершине s присвоим метку $(+s, \delta(s) = \infty)$; при этом вершина s помечена и не просмотрена; все остальные вершины без меток. Примем $x_i = s$.

Шаг 2. Берем непросмотренную вершину x_i с меткой $(\pm x_k, \delta(x_i))$ и просматриваем ее, т.е. рассматриваем все вершины, связанные с x_i, подходящими или отходящими дугами. При этом различаем два случая:

Шаг 2 а. Рассматриваем непомеченные вершины $x_j \in \Gamma(x_i)$, связанные с x_i дугами, направленными от x_i к x_j и имеющими $p_{ij} < c_{ij}$; таким вершинам x_j присваиваем метки $(+x_i, \delta(xj))$, где $\delta(x_j) = \min[\delta(x_i),\ c_{ij} - p_{ij}]$.

Шаг 2 *б*. рассматриваем непомеченные вершины $x_j \in \Gamma^{-1}(x_i)$, связанные с x_i дугами, направленными от x_j к x_i и имеющими $p_{ji} > 0$; таким вершинам x_j присваиваем метки $(-x_i, \delta(x_j))$, где $\delta(x_j) = \min[\delta(x_i),\ p_{ji}]$.

В результате выполнения этого шага вершина x_i становится не только помеченной, но и просмотренной. При этом все смежные с ней вершины $x_j \in \Gamma(x_i)$ и $x_j \in \Gamma^{-1}(x_i)$ оказались помеченными.

Шаг 3. Повторяем шаг 2 до тех пор, пока либо вершина t (сток) окажется помеченной, тогда переходим к шагу 4, либо t будет не помечена и никаких других меток нельзя расставить; в этом случае алгоритм заканчивает работу, а полученное распределение потока является максимальным.

Стадия Б – использование меток для изменения потоков.

Шаг 4. Положим $x_i = t$ и переходим к следующему шагу.

Шаг 5. Для выбранной вершины x_i анализируем ее метку и в зависимости от ее вида меняем поток в одной из дуг, инцидентных x_i:

а) если x_i имеет метку $(+x_m, \delta(x_i))$, то изменяем поток вдоль дуги (x_m, x_i) с p_{mi} на $p_{mi} + \delta(t)$;

б) если x_i имеет метку $(-x_m, \delta(x_i))$, изменяем поток дуги (x_i, x_m) с p_{im} на $p_{im} - \delta(t)$.

Заметим, что какой бы ни была вершина x_i, в любом случае изменение потока проводится на величину $\delta(t)$, определяемую меткой, которая была приписана на данной итерации стоку *t*.

Шаг 6. Если $x_m = s$, то стираем все метки вершин и вновь их расставляем, начиная с шага 1, но при этом используем последнее распределение потока, полученное на шаге 5 предшествующей итерации. Если $x_m \neq s$, то полагаем $x_i = x_m$ и идем к шагу 5.

Примечание. При использовании данного алгоритма для удобства расстановки меток и изменения потоков можно ввести и использовать метки дуг вида $[p_{ij}, c_{ij}]$, где c_{ij} – заданная пропускная способность дуги (она в ходе решения не меняется), а p_{ij} – значение потока этой дуги на рассматриваемой итерации.

Пример. Сеть задана матрицей *C* пропускных способностей, представленной в табл. 4.5. При этом отсутствие элементов в *C* означает отсутствие дуги, идущей от x_i к x_j. Требуется найти максимальный поток, который можно пропустить через заданную сеть, и распределение этого потока по дугам.

Таблица 4.5

$C =$

	x_1	x_2	x_3	x_4
x_1		10	20	
x_2				12
x_3		8		17
x_4				

Решение. По матрице *C* строим граф (рис. 4.8). Как видим, в нем истоком является вершина $s = x_1$, а стоком $t = x_4$. Рядом с дугами графа запишем метки $[p_{ij}, c_{ij}]$, в которых сначала примем $p_{ij} = 0$, что соответствует нулевому начальному значению потока в сети. Приступаем к решению, используя описанный алгоритм.

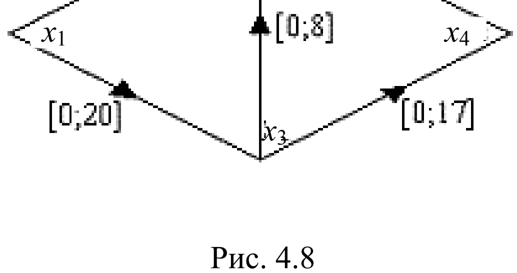


Рис. 4.8

Первая итерация (рис. 4.9).

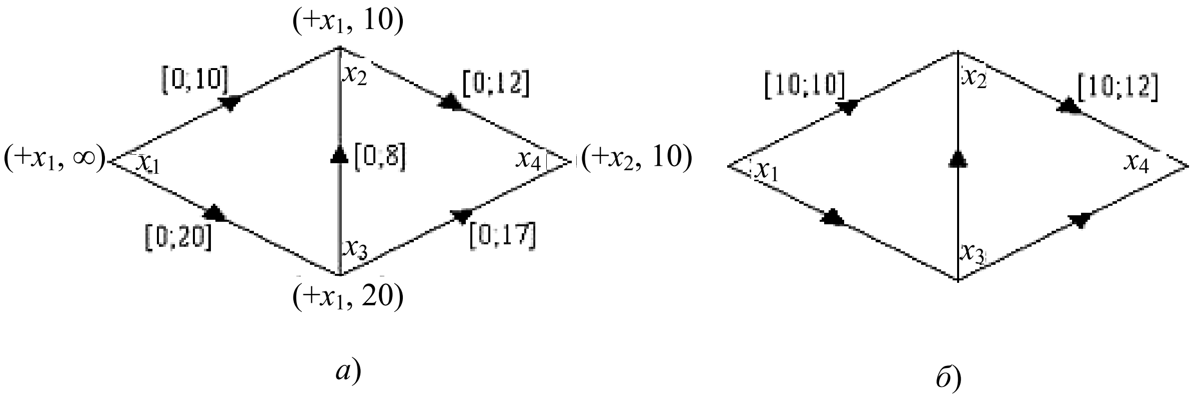


Рис. 4.9

Шаг 1. Вершине $s = x_1$ присвоим метку $(+x_1, \infty)$ (рис. 4.9, *а*), при этом x_1 помечена и не просмотрена, а остальные вершины не помечены.

Шаг 2. Просматриваем единственную пока помеченную вершину x_1. Для этого выделяем $\Gamma(x_1) = \{x_2,\ x_3\}$ и $\Gamma^{-1}(x_1) = \varnothing$. Далее берем x_2. Она не помечена, а дуга (x_1, x_2) имеет $p_{12} = 0 < c_{12} = 10$. Пользуясь правилом шага 2 *а*, вершине x_2 присваиваем метку:

$$x_2 \to (+x_1,\ \min\,[\infty,\ 10 - 0]) = (+x_1,\ 10).$$

Аналогично для вершины x_3 получаем

$$x_3 \to (+x_1, \min\,[\infty,\ 20 - 0]) = (+x_1,\ 20).$$

В результате выполнения этих операций x_1 стала помеченной и просмотренной, а x_2 и x_3 – помеченными.

Шаг 3. Выполняем шаг 2 для $x_i = x_2$:

а) $\Gamma(x_2) = \{x_4\}$, x_4 не помечена, $p_{24} = 0 < c_{24} = 12$. Приписываем вершине x_4 метку:

$$x_4 \to (+x_2, \min\,[\delta(x_2);\ c_{24} - p_{24}]) = (+x_2, \min[10;\ 12 - 0]) = (+x_2,\ 10).$$

б) $\Gamma^{-1}(x_2) = \{x_1, x_3\}$, но эти вершины уже помечены.

Итак, x_2 помечена и просмотрена.

Выполняем шаг 2 для $x_i = x_3$:

а) $\Gamma(x_3) = \{x_2,\ x_4\}$ – эти вершины помечены;

б) $\Gamma^{-1}(x_3) = \{x_1\}$ – эта вершина тоже помечена.

Таким образом, все вершины, в том числе сток $t = x_4$, помечены; переходим к стадии Б и изменяем потоки в дугах.

Шаг 4. Полагаем $x_i = t$.

Шаг 5. Вершине $x_i = t = x_4$ соответствует метка $(+x_2,\ 10)$. Поток в дуге $(x_2,\ x_4)$ меняем с $p_{24} = 0$ на $p_{24} = 0 + 10 = 10$.

Повторяем шаг 5 для $x_i = x_2$:

$x_2 \to (+x_1,\ 10)$ поток дуги (x_1, x_2) меняем с $p_{12} = 0$ на $p_{12} = 0 + 10 = 10$.

Шаг 6. Поскольку мы пришли к вершине $x_1 = s$, то стадию Б заканчиваем. Стираем все метки, перечерчиваем граф (рис. 4.9, *б*) без меток вершин, но с новыми метками дуг, содержащими полученные новые значения потоков в дугах (x_2, x_4) и (x_1, x_2); метки остальных дуг не меняем.

Вторая итерация (рис. 4.10).

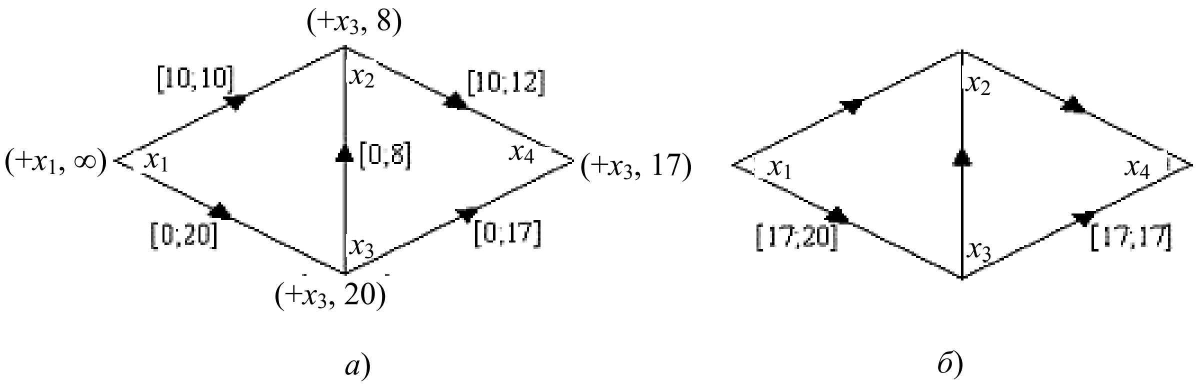


Рис. 4.10

Шаг 1. Для $s = x_1$ записываем метку $(+x_1, \infty)$ (рис. 4.10, *а*).

Шаг 2. Просматриваем дуги, инцидентные x_1. Выделяем вершины $\Gamma(x_1) = \{x_2, x_3\}$ и $\Gamma^{-1}(x_1) = \varnothing$.

Рассмотрим вершину x_2: в соответствии с шагом 2 *а* ей метка не ставится, так как $p_{12} = 10 = c_{12} = 10$ (для записи метки нужно, чтобы $p_{12} < c_{12}$).

Рассмотрим вершину x_3: так как $p_{13} = 0 < c_{13} = 20$, то по правилу 2 *а* приписываем этой вершине метку $x_3 \to (+x_1, \min[\infty,\ 20 - 0]) = (+x_1, 20)$. Берем непросмотренную вершину с меткой (x_3) и выделяем инцидентные ей дуги и соответствующие вершины:

$$\Gamma(x_3) = \{x_2, x_4\},\ \Gamma^{-1}(x_3) = \{x_1\}.$$

Вершина x_1 не рассматривается, так как она уже имеет метку. Для x_2 по правилу 2 *а* получаем метку

$x_2 \to (+x_3, \min[\delta(x_3);\ c_{32} - p_{32}] = (+x_3, \min[20;\ 8-0]) = (+x_3; 8).$

Для x_4 по правилу 2 *а* получаем $x_4 \to (+x_3; \min[20;\ 17-0]) = (+x_3;\ 17)$.

Так как сток $t = x_4$ получил метку, стадию А прекращаем и переходим к изменению потоков.

Шаг 4. Берем вершину $x_i = t$, анализируем ее метку $(+x_3, 17)$ и, используя $\delta = 17$, переходим к следующему шагу.

Шаг 5 а. Поток в дуге (x_3, x_4) меняем с $p_{34} = 0$ на $p_{34} = 0 + 17 = 17$. Для вершины x_3 по метке $(+x_1, 20)$ определяем, что поток дуги (x_1, x_3) нужно изменить с $p_{13} = 0$ на $p_{13} = 0 + 17 = 17$. Стираем метки вершин, перечерчиваем граф с новыми метками дуг (см. рис. 4.10, *б*).

Третья итерация (рис. 4.11, *а*).

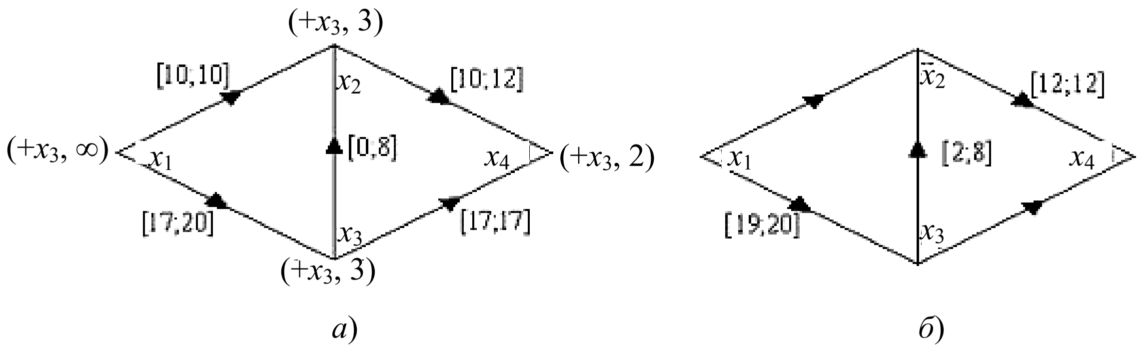


Рис. 4.11

Шаг 1. $s = x_1 \to (+x_1, \infty)$ (рис. 4.11, *а*).

Шаг 2. $\Gamma(x_1) = \{x_2, x_3\}$, $\Gamma^{-1}(x_1) = \varnothing$; x_2 метку не получает, так как $p_{12} = c_{12}$:

$x_3 \to (+x_1, \min[\infty,\ 20-17]) = (+x_1, 3)$;

$\Gamma(x_3) = \{x_2, x_4\}$, $\Gamma^{-1}(x_3) = \{x_1\}$;

$x_2 \to (+x_3, \min[3;\ 8-0]) = (+x_3, 3)$;

x_4 не рассматриваем, так как $p_{24} = c_{24} = 17$.

$\Gamma(x_2) = \{x_4\}$, $\Gamma^{-1}(x_2) = \{x_1, x_3\}$;

$x_4 \to (+x_2, \min[3,\ 12-10]) = (+x_2, 2)$;

x_1 и x_3 уже имеют метки.

Сток $x_4 = s$ получил метку, переходим к стадии Б, изменяя потоки на $\delta = 2$;

для дуги (x_2, x_4) меняем $p_{24} = 10$ на $p_{24} = 10 + 2 = 12$;

для (x_3, x_2) меняем $p_{32} = 0$ на $p_{32} = 0 + 2 = 2$;

для (x_1, x_3) меняем $p_{13} = 17$ на $p_{13} = 17 + 2 = 19$.

Перечерчиваем граф с новыми метками (рис. 4.11, *б*).

Четвертая итерация (рис. 4.12).

Шаг 1. $s = x_1 \to (+x_1, \infty)$.

Шаг 2. $\Gamma(x_1) = \{x_2, x_3\}$, $\Gamma^{-1}(x_1) = \varnothing$ – не рассматриваем, так как $p_{12} = c_{12} = 10$:

$x_3 \to (+x_1, \min[\infty;\ 20 - 19]) = (+x_1, 1)$;

$\Gamma(x_3) = \{x_2, x_4\}$, $\quad \Gamma^{-1}(x_3) = \{x_1\}$;

$x_2 \to (+x_3, \min[1,\ 8 - 2]) = (+x_3, 1)$;

x_4 – не рассматриваем, так как $p_{34} = c_{34} = 17$.

$\Gamma(x_2) = \{x_4\}$; $\Gamma^{-1}(x_2) = \{x_1, x_3\}$.

x_1, x_3 – не рассматриваем, так как они имеют метки;

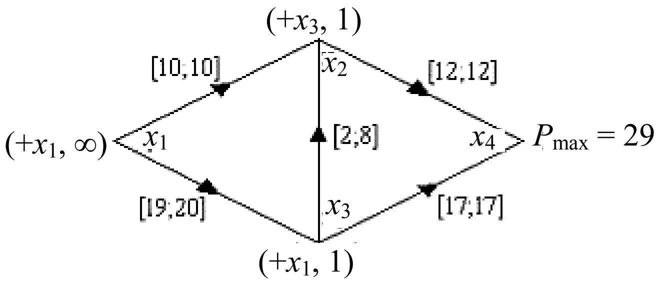


Рис. 4.12

x_4 – пометить нельзя, так как $p_{24} = c_{24} = 12$. Никаких других меток поставить нельзя. По условию, предусмотренному шагом 3, это означает конец решения.

Полученное на графе рис. 4.12 распределение потока соответствует искомой максимальной величине потока сети:

$$P = p_{12} + p_{13} = 10 + 19 = p_{24} + p_{34} = 12 + 17 = 29.$$

Оформление решения. Оно состоит в формулировке решаемой задачи, записи матрицы пропускных способностей, задающей транспортную сеть, построении соответствующего графа, выполнении нужного числа итераций по алгоритму Форда – Фалкерсона с изображением графов вида рис. 4.8 – 4.12, соответствующих каждой итерации, и кратких пояснений, касающихся расстановки меток вершин и изменений потоков дуг.

§ 4.6. Задача об оптимальном распределении заданного потока в транспортной сети

Важное прикладное значение имеет следующая задача. Для исходной транспортной сети с известной величиной максимального потока $P_{\max}$ найти распределение по дугам заданного потока $P = P_0 < P_{\max}$, наилучшее с точки зрения некоторого показателя. Речь идет, например, о составлении наиболее рационального плана перевозок какого-либо груза по имеющейся транспортной магистрали, при котором общая стоимость всех перевозок была бы минимальна. Такая задача называется транспортной задачей. Ей можно придать следующую математическую формулировку.

Дана некоторая сеть $G = (X, U)$, каждой дуге $(x_i, x_j) \in U$ которой помимо пропускной способности c_{ij} приписана еще одна числовая характеристика d_{ij}, трактуемая как стоимость прохождения единицы потока по соответствующей дуге (x_i, x_j). Требуется найти распределение p_{ij} заданного потока $P_0 < P_{\max}$ по этой сети, при котором $0 \le p_{ij} \le c_{ij}$, а общая стоимость прохождения потока была бы минимальной

$$Q = \sum d_{ij} p_{ij} \quad - \quad \min .$$

Решение. Для отыскания решения поставленной задачи можно использовать следующий итерационный алгоритм [18].

Шаг 1. Пусть заданы граф $G = G_0$ и поток $P = P_0$, который должен быть пропущен через сеть, определенную на G_0. Положим $G_k = G_0$; $P_k = P_0$.

Шаг 2. В графе G_k ищем кратчайший путь L_k от *s* к *t*, рассматривая в качестве длин дуг соответствующие стоимости d_{ij}.

Шаг 3. Определяем пропускную способность $c(L_k)$ этого пути L_k как $\min[c_{ij}]$ по всем дугам, вошедшим в L_k.

Шаг 4. По этому пути L_k пропускаем поток

$$P(L_k) = \begin{cases} P_k, & \text{если} \quad P_k \le c(L_k); \\ c(L_k), & \text{если} \quad P_k > c(L_k). \end{cases}$$

Шаг 5. Если $P_k \le c(L_k)$, то задача решена и передачу заданного потока P_k нужно осуществить по найденному пути L_k; после чего перейти к шагу 8. Если же $P_k > c(L_k)$, то переходим к следующему шагу.

Шаг 6. От графа G_k переходим к графу G_{k+1}, который получается из G_k заменой пропускных способностей дуг c_{ij} на c'_{ij} по следующему правилу:

$$c'_{ij} = \begin{cases} c_{ij} - c(L_k), & \text{если} \quad (x_i, x_j) \in L_k; \\ c_{ij}, & \text{если} (x_i, x_j) \notin L_k. \end{cases}$$

При этом дуги, для которых получается $c'_{ij} = 0$, можно исключить из графа.

Шаг 7. Рассматриваем в качестве G_k полученный граф G_{k+1} и, принимая новое значение P_{k+1} равным $P_k^{CTAP} - P(L_k)$, переходим к шагу 2.

Шаг 8. Суммируя для каждой дуги значения потоков, пропущенных через эту дугу на каждой итерации, находим искомое оптимальное распределение заданного потока.

Замечание. При отыскании кратчайших путей L_k в графах G_k можно использовать алгоритм, описанный в § 4.3.

Пример. Для пояснения изложенного алгоритма рассмотрим простой пример транспортной сети, заданной табл. 4.5 из предыдущего § 4.6, дополнив ее следующей матрицей стоимостей D:

$D =$

	x_1	x_2	x_3	x_4
x_1		26	20	
x_2				10
x_3		4		15
x_4				

Граф, соответствующий этой сети, изображен на рис. 4.13, *а*. Для каждой дуги (x_i, x_j) в этом графе указана пара (c_{ij}, d_{ij}). В § 4.5 была определена максимальная величина потока для данной сети $P_{\max} = 29$. Допустим, что требуется распределить заданный поток $P_0 = 22 < P_{\max}$ так, чтобы общая стоимость Q была минимальная.

Решение.

Первая итерация.

Шаг 1. Пусть G_k – исходный граф (рис. 4.13, *а*), а $P_k = P_0 = 22$.

Шаг 2. В G_k выбираем кратчайший путь от $s = x_1$ к $t = x_4$, используя в качестве длин дуг числа d_{ij}. В данном графе имеется всего три пути от s к t, они изображены на рис. 4.13, *б* – *г*. Определяя для каждого из них длину пути (см. рис. 4.13), выбираем кратчайший путь, приведенный на рис. 4.13, *б*. При этом $d(L_k) = 34$.

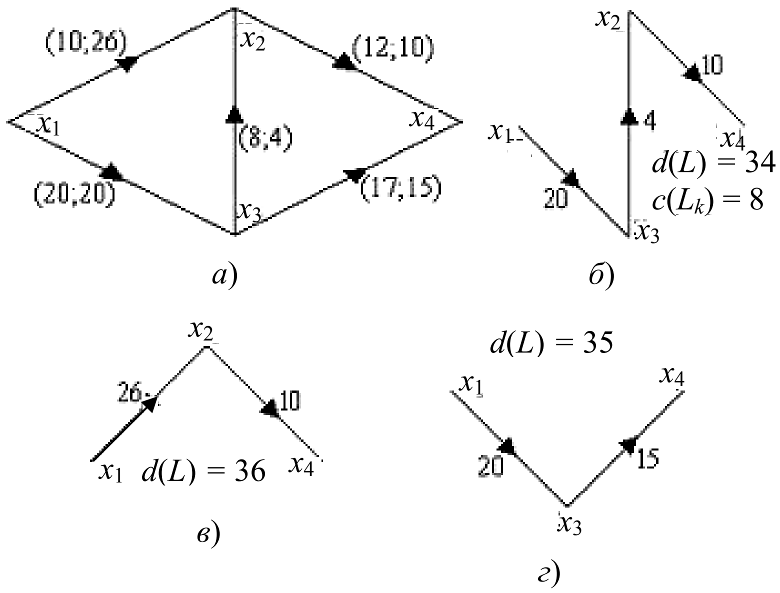


Рис. 4.13

Шаг 3. Определяем пропускную способность этого пути, используя числа c_{ij}, приписанные дугам, вошедшим в этот путь:

$$c(L_k) = \min\{20, 8, 12\} = 8.$$

Шаг 4. Так как $P_k = 22 > c(L_k) = 8$, то по пути L_k (см. рис. 4.13, *б*) пропускаем поток $P(L_k) = 8$.

Шаг 5. Поскольку $P_k > c(L_k)$, переходим к шагу 6.

Шаг 6. Строим новый граф (рис. 4.14, *а*), в котором пропускные способности дуг, вошедших в L_k, уменьшаем на $c(L_k) = 8$, при этом дугу (x_3, x_2), получившую c'_{32} = 0, исключаем из графа, а пропускные способности остальных дуг не меняем.

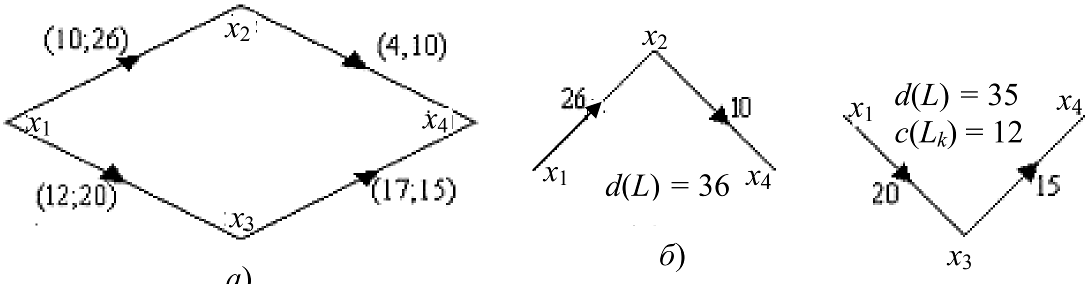
Шаг 7. Принимаем в качестве G_k полученный граф (рис. 4.14, *а*), полагаем $P_k = 22 - 8 = 14$ и переходим к шагу 2.

Вторая итерация.

Из двух возможных путей от s к t в графе (рис. 4.14. *б*, *в*) выбираем кратчайший путь L_k (рис. 4.14, *в*) и определяем его пропускную способность $c(L_k) = 12$.

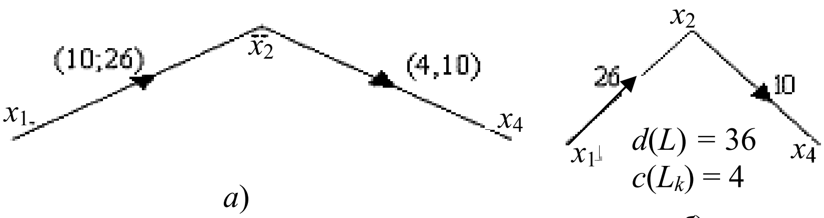
При этом $c(L_k) = 12 < P_k = 14$. Пропускаем по L_k поток $P(L_k) = 12$. Строим на базе графа рис. 4.14 новый граф, в котором пропускные способности дуг (x_1, x_3), $(x_3, x_4) \in L_k$ уменьшены на $c(L_k) = 12$, причем дуга (x_1, x_3), получившая c'_{13} = 0, исключена из графа. Принимаем в качестве

G_k полученный граф (рис. 4.15, *а*), полагаем $P_k = 14 - 12 = 2$ и переходим к следующей итерации.



а) *б*) *в*)

Рис. 4.14



а) *б*)

Рис. 4.15

Третья итерация.

Полученный граф (см. рис. 4.15, *а*) представляет собой единственный путь от s к t. Его пропускная способность $c(L_k) = \min\,[10, 4] = 4$. При этом $P_k = 2 < c(L_k)$. По условию шага 5 через этот путь пропускаем поток $P_k = 2$ и переходим к шагу 8.

Шаг 8. Получение оптимального распределения потока.

Суммируя потоки, пропущенные по дугам на каждой итерации, находим:

$p_{12} = 2$ (на 3-й итерации) $= 2$;
$p_{13} = 8$ (на 1-й итерации) + 12 (на 2-й итерации) $= 20$;
$p_{32} = 8$ (на 1-й итерации) $= 8$;
$p_{24} = 8$ (на 1-й итерации) + 2 (на 3-й итерации) $= 10$;
$p_{34} = 12$ (на 2-й итерации) $= 12$.

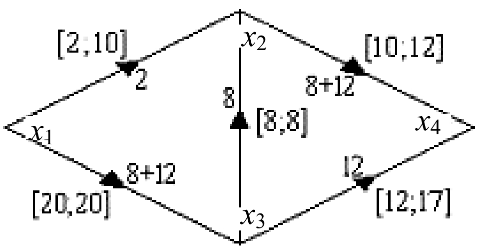


Рис. 4.16

Это распределение потока по дугам (x_i, x_j) представлено на рис. 4.16 вместе с соответствующими пропускными способностями $[p_{ij}, c_{ij}]$.

Контрольные вопросы

1. Что называется графом?
2. Какой граф называется ориентированным, а какой – неориентированным?
3. Каким образом можно граф задать с помощью матриц, как называются такие матрицы?
4. Какой граф называется связным?
5. Что представляет собой путь в графе, какой путь называется простым?
6. Какой граф называется взвешенным?
7. Как формулируется задача о кратчайшем пути в графе?
8. Какой метод используется для решения задачи о кратчайших путях в графе?
9. Что представляют собой метки, используемые при решении задачи о кратчайших путях, каков алгоритм, по которому они определяются?
10. Как формулируется задача о критическом пути в графе?
11. Что представляет собой сетевой график и какие задачи позволяет он решать?
12. Как формулируется задача о графе минимальной длины, какую структуру имеет такой граф?
13. Какой граф называется транспортной сетью?
14. Как формулируется задача о максимальном потоке в транспортной сети?
15. Какое содержание имеют и по какому правилу формируются метки в алгоритме Форда – Фалкерсона?
16. Для какой цели и каким образом используются метки в алгоритме Форда – Фалкерсона при решении задачи о максимальном потоке в транспортной сети?
17. Как формулируется задача об оптимальном распределении заданного потока в транспортной сети?

Задачи для самостоятельного решения

1. Задача о кратчайших путях в графе

Неориентированный граф с десятью вершинами $x1...x10$ задан верхней треугольной «половиной» матрицы весов. При этом отсутствие элемента c_{ij} в матрице указывает на отсутствие в графе ребра, связывающего вершины x_i и x_j.

Необходимо найти кратчайший путь из истока в каждую из остальных вершин графа.

№ 1-1

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		7	1	7	3					
x2			4						5	
x3				2					9	
x4					6	11		10		
x5							3			
x6							7	1	4	5
x7										3
x8									6	8
x9										4
x10										

Исток: x2

№ 1-2

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		7	4	12	5					
x2			2						5	
x3				4					9	
x4					7	5		18		
x5							14			
x6							2	8	1	13
x7										12
x8									9	3
x9										7
x10										

Исток: x1

№ 1-3

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		6	2	3	7					
x2			7						5	
x3				4					3	
x4					7	7		2		
x5							8			
x6							5	6	10	5
x7										10
x8									6	3
x9										10
x10										

Исток: x1

№ 1-4

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		6	2	4	5					
x2			1						2	
x3				3					9	
x4					5	4		7		
x5							6			
x6							6	2	8	3
x7										9
x8									2	2
x9										6
x10										

Исток: x6

№ 1-5

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		6	5	1	1					
x2			7						3	
x3				5					3	
x4					8	8		4		
x5							7			
x6							6	5	9	5
x7										8
x8									7	7
x9										10
x10										

Исток: x4

№ 1-6

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		5	4	3	2					
x2			5						7	
x3				10					4	
x4					4	5		3		
x5							8			
x6							6	6	4	7
x7										9
x8									10	5
x9										6
x10										

Исток: x1

№ 1-7

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		4	3	7	1					
x2			5						2	
x3				5					7	
x4					3	3		8		
x5							6			
x6							4	10	9	7
x7										11
x8									11	7
x9										5
x10										

Исток: x1

№ 1-8

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		9	5	4	6					
x2			7						5	
x3				6					5	
x4					12	7		8		
x5							4			
x6							3	7	12	1
x7										9
x8									1	6
x9										7
x10										

Исток: x4

№ 1-9

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		4	8	2	7					
x2			1						9	
x3				8					2	
x4					5	10		7		
x5							4			
x6							10	3	6	7
x7										5
x8									9	3
x9										8
x10										

Исток: x4

№ 1-10

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		6	2	7	5					
x2			3						3	
x3				9					6	
x4					11	3		2		
x5							11			
x6							12	3	8	4
x7										9
x8									2	1
x9										5
x10										

Исток: x2

№ 1-11

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		7	5	1	9					
x2			6						3	
x3				1					5	
x4					8	2		4		
x5							6			
x6							3	3	8	10
x7										4
x8									2	5
x9										5
x10										

Исток: x1

№ 1-12

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		9	4	1	3					
x2			5						1	
x3				7					3	
x4					5	8		2		
x5							6			
x6							12	8	4	12
x7										10
x8									7	8
x9										6
x10										

Исток: x6

№ 1-13

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		4	1	3	2					
x2			8						9	
x3				1					5	
x4					9	3		6		
x5							8			
x6							6	4	10	2
x7										9
x8									1	8
x9										11
x10										

Исток: x10

№ 1-14

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		12	7	8	2					
x2			6						4	
x3				5					1	
x4					7	3		8		
x5							6			
x6							7	4	9	12
x7										5
x8									5	4
x9										6
x10										

Исток: x5

№ 1-15

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		9	6	3	1					
x2			1						9	
x3				7					4	
x4					2	9		8		
x5							5			
x6							6	4	7	3
x7										11
x8									1	4
x9										6
x10										

Исток: x1

№ 1-16

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		8	1	3	2					
x2			9						11	
x3				17					6	
x4					9	6		3		
x5							8			
x6							4	10	12	2
x7										9
x8									4	8
x9										5
x10										

Исток: x1

№ 1-17

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		5	9	2	6					
x2			11						7	
x3				17					2	
x4					2	3		4		
x5							6			
x6							1	9	7	5
x7										8
x8									2	6
x9										4
x10										

Исток: x5

№ 1-18

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		5	13	9	4					
x2			6						11	
x3				2					7	
x4					8	7		4		
x5							2			
x6							5	5	2	1
x7										9
x8									4	7
x9										3
x10										

Исток: x9

№ 1-19

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		1	9	6	3					
x2			5						7	
x3				9					2	
x4					4	7		5		
x5							10			
x6							7	8	2	4
x7										8
x8									4	3
x9										6
x10										

Исток: x1

№ 1-20

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		9	6	2	3					
x2			1						5	
x3				3					5	
x4					8	4		1		
x5							6			
x6							1	5	7	2
x7										9
x8									3	4
x9										5
x10										

Исток: x10

№ 1-21

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		7	1	1	5					
x2			7						6	
x3				2					3	
x4					6	8		11		
x5							14			
x6							6	1	5	8
x7										4
x8									5	1
x9										7
x10										

Исток: x2

№ 1-22

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		4	2	9	5					
x2			2						4	
x3				10					6	
x4					12	8		6		
x5							4			
x6							9	3	2	8
x7										10
x8									4	3
x9										6
x10										

Исток: x7

№ 1-23

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		8	5	7	1					
x2			2						3	
x3				5					6	
x4					8	2		1		
x5							3			
x6							10	6	2	5
x7										9
x8									7	5
x9										11
x10										

Исток: x6

№ 1-24

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		4	7	1	3					
x2			2						5	
x3				4					8	
x4					6	2		4		
x5							2			
x6							5	9	1	6
x7										10
x8									6	3
x9										5
x10										

Исток: x1

№ 1-25

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		2	9	2	6					
x2			5						6	
x3				3					7	
x4					2	4		9		
x5							6			
x6							1	5	8	2
x7										9
x8									2	3
x9										8
x10										

Исток: x1

№ 1-26

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1		1	9	7	6					
x2			1						7	
x3				8					2	
x4					2	4		3		
x5							5			
x6							9	1	7	5
x7										8
x8									6	3
x9										1
x10										

Исток: x3

2. Задача о графе минимальной длины

Неориентированный связный взвешенный граф задан матрицей весовых коэффициентов.

Необходимо построить новый связный граф, который содержал бы все вершины исходного графа и имел бы наименьшую сумму весов входящих в него ребер.

Исходные данные выбрать в соответствии со своим вариантом предыдущего задания.

3. Задача о критическом пути в графе

Дан ориентированный граф, в котором отсутствуют контуры, и каждой дуге которого приписан вес $c(i, j) > 0$. Структура графа определяется следующим условием: он имеет дугу, связывающую вершину x_i с вершиной x_j, если в исходных условиях задан вес $c(i, j)$.

Необходимо найти путь от заданной начальной вершины (истока) к заданной конечной вершине (стоку), имеющий наибольшую длину, равную сумме весов дуг, входящих в этот путь.

№ 3-1

$c(1,2) = 5$; $c(1,6) = 3$; $c(1,7) = 10$; $c(2,3) = 7$; $c(2,7) = 14$; $c(2,8) = 7$; $c(2,9) = 12$; $c(3,4) = 15$; $c(3,9) = 8$; $c(4,12) = 16$; $c(5,7) = 13$; $c(5,11) = 7$; $c(5,13) = 12$; $c(6,5) = 9$; $c(6,7) = 10$; $c(7,8) = 8$; $c(7,10) = 17$; $c(7,11) = 4$; $c(7,13) = 11$; $c(8,4) = 8$; $c(8,10) = 8$; $c(8,12) = 9$; $c(9,4) = 6$; $c(9,8) = 7$; $c(10,12) = 18$; $c(10,13) = 12$; $c(11,13) = 19$; $c(12,13) = 6$.

№ 3-2

$c(1,2) = 7$; $c(1,6) = 2$; $c(1,7) = 17$; $c(2,3) = 4$; $c(2,7) = 2$; $c(2,8) = 12$; $c(2,9) = 7$; $c(3,4) = 11$; $c(3,9) = 10$; $c(4,12) = 4$; $c(5,7) = 6$; $c(5,11) = 14$; $c(5,13) = 6$; $c(6,5) = 12$; $c(6,7) = 14$; $c(7,8) = 10$; $c(7,10) = 7$; $c(7,11) = 12$; $c(7,13) = 14$; $c(8,4) = 3$; $c(8,10) = 15$; $c(8,12) = 12$; $c(9,4) = 12$; $c(9,8) = 10$; $c(10,12) = 6$; $c(10,13) = 11$; $c(11,13) = 4$; $c(12,13) = 12$.

№ 3-3

$c(1,2) = 2$; $c(1,6) = 2$; $c(1,7) = 15$; $c(2,3) = 4$; $c(2,7) = 4$; $c(2,8) = 11$; $c(2,9) = 9$; $c(3,4) = 7$; $c(3,9) = 7$; $c(4,12) = 8$; $c(5,7) = 6$; $c(5,11) = 13$; $c(5,13) = 5$; $c(6,5) = 7$; $c(6,7) = 11$; $c(7,8) = 19$; $c(7,10) = 7$; $c(7,11) = 15$; $c(7,13) = 11$; $c(8,4) = 3$; $c(8,10) = 4$; $c(8,12) = 12$; $c(9,4) = 12$; $c(9,8) = 7$; $c(10,12) = 6$; $c(10,13) = 14$; $c(11,13) = 11$; $c(12,13) = 12$.

№ 3-4

$c(1,2) = 4$; $c(1,6) = 5$; $c(1,7) = 5$; $c(2,3) = 7$; $c(2,7) = 2$; $c(2,8) = 13$; $c(2,9) = 8$; $c(3,4) = 6$; $c(3,9) = 11$; $c(4,12) = 4$; $c(5,7) = 14$; $c(5,11) = 6$; $c(5,13) = 8$; $c(6,5) = 12$; $c(6,7) = 9$; $c(7,8) = 11$; $c(7,10) = 3$; $c(7,11) = 4$; $c(7,13) = 9$; $c(8,4) = 11$; $c(8,10) = 12$; $c(8,12) = 7$; $c(9,4) = 6$; $c(9,8) = 8$; $c(10,12) = 2$; $c(10,13) = 4$; $c(11,13) = 18$; $c(12,13) = 11$.

№ 3-5

$c(1,2) = 4$; $c(1,6) = 11$; $c(1,7) = 3$; $c(2,3) = 17$; $c(2,7) = 12$; $c(2,8) = 5$; $c(2,9) = 8$; $c(3,4) = 2$; $c(3,9) = 10$; $c(4,12) = 13$; $c(5,7) = 14$; $c(5,11) = 9$; $c(5,13) = 7$; $c(6,5) = 6$; $c(6,7) = 10$; $c(7,8) = 12$; $c(7,10) = 7$; $c(7,11) = 11$; $c(7,13) = 11$; $c(8,4) = 7$; $c(8,10) = 3$; $c(8,12) = 4$; $c(9,4) = 8$; $c(9,8) = 10$; $c(10,12) = 6$; $c(10,13) = 2$; $c(11,13) = 12$; $c(12,13) = 4$.

№ 3-6

$c(1,2) = 17$; $c(1,6) = 8$; $c(1,7) = 12$; $c(2,3) = 8$; $c(2,7) = 2$; $c(2,8) = 14$; $c(2,9) = 16$; $c(3,4) = 15$; $c(3,9) = 10$; $c(4,12) = 6$; $c(5,7) = 6$; $c(5,11) = 13$; $c(5,13) = 7$; $c(6,5) = 19$; $c(6,7) = 12$; $c(7,8) = 10$; $c(7,10) = 3$; $c(7,11) = 11$; $c(7,13) = 9$; $c(8,4) = 8$; $c(8,10) = 2$; $c(8,12) = 6$; $c(9,4) = 11$; $c(9,8) = 7$; $c(10,12) = 1$; $c(10,13) = 5$; $c(11,13) = 7$; $c(12,13) = 12$.

№ 3-7

$c(1,2) = 7$; $c(1,6) = 5$; $c(1,7) = 14$; $c(2,3) = 15$; $c(2,7) = 10$; $c(2,8) = 12$; $c(2,9) = 4$; $c(3,4) = 7$; $c(3,9) = 1$; $c(4,12) = 4$; $c(5,7) = 13$; $c(5,11) = 7$; $c(5,13) = 13$; $c(6,5) = 8$; $c(6,7) = 6$; $c(7,8) = 4$; $c(7,10) = 19$; $c(7,11) = 4$; $c(7,13) = 8$; $c(8,4) = 7$; $c(8,10) = 8$; $c(8,12) = 11$; $c(9,4) = 5$; $c(9,8) = 9$; $c(10,12) = 12$; $c(10,13) = 6$; $c(11,13) = 6$; $c(12,13) = 19$.

№ 3-8

$c(1,2) = 10$; $c(1,6) = 19$; $c(1,7) = 7$; $c(2,3) = 14$; $c(2,7) = 5$; $c(2,8) = 11$; $c(2,9) = 6$; $c(3,4) = 17$; $c(3,9) = 8$; $c(4,12) = 12$; $c(5,7) = 13$; $c(5,11) = 11$; $c(5,13) = 9$; $c(6,5) = 9$; $c(6,7) = 10$; $c(7,8) = 17$; $c(7,10) = 4$; $c(7,11) = 2$; $c(7,13) = 11$; $c(8,4) = 6$; $c(8,10) = 18$; $c(8,12) = 8$; $c(9,4) = 8$; $c(9,8) = 7$; $c(10,12) = 9$; $c(10,13) = 19$; $c(11,13) = 7$; $c(12,13) = 2$.

№ 3-9

$c(1,2) = 11$; $c(1,6) = 10$; $c(1,7) = 3$; $c(2,3) = 4$; $c(2,7) = 7$; $c(2,8) = 15$; $c(2,9) = 8$; $c(3,4) = 15$; $c(3,9) = 11$; $c(4,12) = 2$; $c(5,7) = 6$; $c(5,11) = 2$; $c(5,13) = 2$; $c(6,5) = 9$; $c(6,7) = 8$; $c(7,8) = 17$; $c(7,10) = 6$; $c(7,11) = 4$; $c(7,13) = 11$; $c(8,4) = 16$; $c(8,10) = 13$; $c(8,12) = 15$; $c(9,4) = 2$; $c(9,8) = 4$; $c(10,12) = 18$; $c(10,13) = 12$; $c(11,13) = 6$; $c(12,13) = 5$.

№ 3-10

$c(1,2) = 2$; $c(1,6) = 5$; $c(1,7) = 4$; $c(2,3) = 6$; $c(2,7) = 3$; $c(2,8) = 2$; $c(2,9) = 1$; $c(3,4) = 10$; $c(3,9) = 15$; $c(4,12) = 8$; $c(5,7) = 6$; $c(5,11) = 14$; $c(5,13) = 6$; $c(6,5) = 12$; $c(6,7) = 2$; $c(7,8) = 16$; $c(7,10) = 8$; $c(7,11) = 10$; $c(7,13) = 5$; $c(8,4) = 4$; $c(8,10) = 15$; $c(8,12) = 11$; $c(9,4) = 3$; $c(9,8) = 6$; $c(10,12) = 9$; $c(10,13) = 10$; $c(11,13) = 3$; $c(12,13) = 15$.

№ 3-11

$c(1,2) = 9$; $c(1,6) = 5$; $c(1,7) = 10$; $c(2,3) = 3$; $c(2,7) = 12$; $c(2,8) = 13$; $c(2,9) = 7$; $c(3,4) = 11$; $c(3,9) = 9$; $c(4,12) = 8$; $c(5,7) = 5$; $c(5,11) = 14$; $c(5,13) = 7$; $c(6,5) = 12$; $c(6,7) = 13$; $c(7,8) = 19$; $c(7,10) = 8$; $c(7,11) = 12$; $c(7,13) = 14$; $c(8,4) = 13$; $c(8,10) = 17$; $c(8,12) = 11$; $c(9,4) = 11$; $c(9,8) = 10$; $c(10,12) = 6$; $c(10,13) = 10$; $c(11,13) = 4$; $c(12,13) = 3$.

№ 3-12

$c(1,2) = 10$; $c(1,6) = 11$; $c(1,7) = 12$; $c(2,3) = 5$; $c(2,7) = 7$; $c(2,8) = 6$; $c(2,9) = 4$; $c(3,4) = 1$; $c(3,9) = 9$; $c(4,12) = 7$; $c(5,7) = 6$; $c(5,11) = 14$; $c(5,13) = 6$; $c(6,5) = 12$; $c(6,7) = 11$; $c(7,8) = 7$; $c(7,10) = 10$; $c(7,11) = 9$; $c(7,13) = 12$; $c(8,4) = 7$; $c(8,10) = 15$; $c(8,12) = 19$; $c(9,4) = 5$; $c(9,8) = 13$; $c(10,12) = 6$; $c(10,13) = 11$; $c(11,13) = 4$; $c(12,13) = 12$.

№ 3-13

$c(1,2) =5$; $c(1,6) = 6$; $c(1,7) = 7$; $c(2,3) = 5$; $c(2,7) = 6$; $c(2,8) = 7$; $c(2,9) = 8$; $c(3,4) = 11$; $c(3,9) = 10$; $c(4,12) = 9$; $c(5,7) = 2$; $c(5,11) = 20$; $c(5,13) = 1$; $c(6,5) = 5$; $c(6,7) = 3$; $c(7,8) = 10$; $c(7,10) = 7$; $c(7,11) = 12$; $c(7,13) = 14$; $c(8,4) = 3$; $c(8,10) = 17$; $c(8,12) = 19$; $c(9,4) = 12$; $c(9,8) = 10$; $c(10,12) = 4$; $c(10,13) = 5$; $c(11,13) = 8$; $c(12,13) = 9$.

№ 3-14

$c(1,2) = 8$; $c(1,6) = 7$; $c(1,7) = 15$; $c(2,3) = 2$; $c(2,7) = 18$; $c(2,8) = 4$; $c(2,9) = 3$; $c(3,4) = 19$; $c(3,9) = 10$; $c(4,12) = 5$; $c(5,7) = 6$; $c(5,11) = 12$; $c(5,13) = 15$; $c(6,5) = 7$; $c(6,7) = 17$; $c(7,8) = 10$; $c(7,10) = 7$; $c(7,11) = 13$; $c(7,13) = 11$; $c(8,4) = 2$; $c(8,10) = 15$; $c(8,12) = 10$; $c(9,4) = 16$; $c(9,8) = 10$; $c(10,12) = 6$; $c(10,13) = 11$; $c(11,13) = 5$; $c(12,13) = 10$.

№ 3-15

$c(1,2)=4$; $c(1,6) = 8$; $c(1,7) = 18$; $c(2,3) = 11$; $c(2,7) = 4$; $c(2,8) = 17$; $c(2,9) = 12$; $c(3,4) = 15$; $c(3,9) = 8$; $c(4,12) = 7$; $c(5,7) = 13$; $c(5,11) = 11$; $c(5,13) = 17$; $c(6,5) = 4$; $c(6,7) = 2$; $c(7,8) = 18$; $c(7,10) = 7$; $c(7,11) = 5$; $c(7,13) = 14$; $c(8,4) = 3$; $c(8,10) = 9$; $c(8,12) = 3$; $c(9,4) = 16$; $c(9,8) = 5$; $c(10,12) = 4$; $c(10,13) = 14$; $c(11,13) = 9$; $c(12,13) = 3$.

№ 3-16

$c(1,2) = 2$; $c(1,6) = 3$; $c(1,7) = 4$; $c(2,3) = 5$; $c(2,7) = 6$; $c(2,8) = 7$; $c(2,9) = 8$; $c(3,4) = 9$; $c(3,9) = 10$; $c(4,12) = 11$; $c(5,7) = 11$; $c(5,11) = 7$; $c(5,13) = 19$; $c(6,5) = 12$; $c(6,7) = 5$; $c(7,8) = 11$; $c(7,10) = 11$; $c(7,11) = 5$; $c(7,13) = 6$; $c(8,4) = 3$; $c(8,10) = 4$; $c(8,12) = 10$; $c(9,4) = 12$; $c(9,8) = 9$; $c(10,12) = 8$; $c(10,13) = 11$; $c(11,13) = 5$; $c(12,13) = 4$.

№ 3-17

$c(1,2) = 4$; $c(1,6) = 3$; $c(1,7) = 2$; $c(2,3) = 5$; $c(2,7) = 7$;$c(2,8) = 6$; $c(2,9) = 8$; $c(3,4) = 10$; $c(3,9) = 11$; $c(4,12) = 9$;$c(5,7) = 12$; $c(5,11) = 13$; $c(5,13) = 16$; $c(6,5)= 14$; $c(6,7) = 15$; $c(7,8) = 2$; $c(7,10) = 17$; $c(7,11) = 1$; $c(7,13) = 8$; $c(8,4) = 4$; $c(8,10) = 6$; $c(8,12) = 7$; $c(9,4) = 4$; $c(9,8) = 5$; $c(10,12) = 8$; $c(10,13) = 8$; $c(11,13) = 10$; $c(12,13) = 11$.

№ 3-18

$c(1,2) = 10$; $c(1,6) = 6$; $c(1,7) = 11$; $c(2,3) = 5$; $c(2,7) = 9$; $c(2,8) = 3$; $c(2,9) = 8$; $c(3,4) = 10$; $c(3,9) = 4$; $c(4,12) = 8$; $c(5,7) = 15$; $c(5,11) = 4$; $c(5,13) = 3$; $c(6,5) = 4$; $c(6,7) = 2$; $c(7,8) = 2$; $c(7,10) = 15$; $c(7,11) = 7$; $c(7,13) = 18$; $c(8,4) = 3$; $c(8,10) = 12$; $c(8,12) = 8$; $c(9,4) = 11$; $c(9,8) = 9$; $c(10,12) = 4$; $c(10,13) = 7$; $c(11,13) = 7$; $c(12,13) = 13$.

№ 3-19

$c(1,2) = 1$; $c(1,6) = 9$; $c(1,7) = 10$; $c(2,3) = 9$; $c(2,7) = 11$; $c(2,8) = 2$; $c(2,9) = 10$; $c(3,4) = 2$; $c(3,9) = 7$; $c(4,12) = 4$; $c(5,7) = 10$; $c(5,11) = 8$; $c(5,13) = 7$; $c(6,5) = 5$; $c(6,7) = 20$; $c(7,8) = 10$; $c(7,10) = 8$; $c(7,11) = 6$; $c(7,13) = 2$; $c(8,4) = 11$; $c(8,10) = 15$; $c(8,12) = 4$; $c(9,4) = 3$; $c(9,8) = 10$; $c(10,12) = 11$; $c(10,13) = 15$; $c(11,13) = 8$; $c(12,13) = 17$.

№ 3-20

$c(1,2) = 10$; $c(1,6) = 11$; $c(1,7) = 12$; $c(2,3) = 5$; $c(2,7) = 6$; $c(2,8) = 2$; $c(2,9) = 8$; $c(3,4) = 9$; $c(3,9) = 11$; $c(4,12) = 13$; $c(5,7) = 5$; $c(5,11) = 9$; $c(5,13) = 10$; $c(6,5) = 17$; $c(6,7) = 4$; $c(7,8) = 3$; $c(7,10) = 5$; $c(7,11) = 11$; $c(7,13) = 16$; $c(8,4) = 4$; $c(8,10) = 9$; $c(8,12) = 10$; $c(9,4) = 12$; $c(9,8) = 17$; $c(10,12) = 4$; $c(10,13) = 7$; $c(11,13) = 9$; $c(12,13) = 10$.

№ 3-21

$c(1,2) = 1$; $c(1,6) = 7$; $c(1,7) = 2$; $c(2,3) = 4$; $c(2,7) = 7$; $c(2,8) = 13$; $c(2,9) = 2$; $c(3,4) = 4$; $c(3,9) = 8$; $c(4,12) = 5$; $c(5,7) = 2$; $c(5,11) = 14$; $c(5,13) = 6$; $c(6,5) = 8$; $c(6,7) = 10$; $c(7,8) = 1$; $c(7,10) = 6$; $c(7,11) = 5$; $c(7,13) = 12$; $c(8,4) = 4$; $c(8,10) = 9$; $c(8,12) = 2$; $c(9,4) = 2$; $c(9,8) = 7$; $c(10,12) = 12$; $c(10,13) = 9$; $c(11,13) = 1$; $c(12,13) = 5$.

№ 3-22

$c(1,2) = 7$; $c(1,6) = 4$; $c(1,7) = 4$; $c(2,3) = 6$; $c(2,7) = 3$; $c(2,8) = 9$; $c(2,9) = 2$; $c(3,4) = 10$; $c(3,9) = 7$; $c(4,12) = 16$; $c(5,7) = 11$; $c(5,11) = 8$; $c(5,13) = 15$; $c(6,5) = 6$; $c(6,7) = 3$; $c(7,8) = 15$; $c(7,10) = 8$; $c(7,11) = 4$; $c(7,13) = 2$; $c(8,4) = 6$; $c(8,10) = 9$; $c(8,12) = 1$; $c(9,4) = 12$; $c(9,8) = 4$; $c(10,12) = 13$; $c(10,13) = 6$; $c(11,13) = 9$; $c(12,13) = 8$.

№ 3-23

$c(1,2) = 5$; $c(1,6) = 9$; $c(1,7) = 4$; $c(2,3) = 3$; $c(2,7) = 6$; $c(2,8) = 4$; $c(2,9) = 3$; $c(3,4) = 15$; $c(3,9) = 5$; $c(4,12) = 6$; $c(5,7) = 19$; $c(5,11) = 15$; $c(5,13) = 5$; $c(6,5) = 2$; $c(6,7) = 4$; $c(7,8) = 7$; $c(7,10) = 6$; $c(7,11) = 3$; $c(7,13) = 11$; $c(8,4) = 8$; $c(8,10) = 8$; $c(8,12) = 16$; $c(9,4) = 2$; $c(9,8) = 7$; $c(10,12) = 3$; $c(10,13) = 9$; $c(11,13) = 5$; $c(12,13) = 9$.

№ 3-24

$c(1,2) =6$; $c(1,6) = 7$; $c(1,7) = 5$; $c(2,3) = 4$; $c(2,7) = 1$; $c(2,8) = 5$; $c(2,9) = 8$; $c(3,4) = 11$; $c(3,9) = 4$; $c(4,12) = 7$; $c(5,7) = 5$; $c(5,11) = 3$; $c(5,13) = 8$; $c(6,5) = 5$; $c(6,7) = 1$; $c(7,8) = 12$; $c(7,10) = 9$; $c(7,11) = 4$; $c(7,13) = 2$; $c(8,4) = 7$; $c(8,10) = 16$; $c(8,12) = 8$; $c(9,4) = 5$; $c(9,8) = 6$; $c(10,12) = 7$; $c(10,13) = 12$; $c(11,13) = 6$; $c(12,13) = 3$.

№ 3-25

$c(1,2) = 6$; $c(1,6) = 1$; $c(1,7) = 8$; $c(2,3) = 4$; $c(2,7) = 5$; $c(2,8) = 11$; $c(2,9) = 5$; $c(3,4) = 6$; $c(3,9) = 8$; $c(4,12) = 3$; $c(5,7) = 9$; $c(5,11) = 6$; $c(5,13) = 4$; $c(6,5) = 11$; $c(6,7) = 9$; $c(7,8) = 4$; $c(7,10) = 8$; $c(7,11) = 2$; $c(7,13) = 6$; $c(8,4) = 4$; $c(8,10) = 15$; $c(8,12) = 8$; $c(9,4) = 10$; $c(9,8) = 6$; $c(10,12) = 12$; $c(10,13) = 1$; $c(11,13) = 5$; $c(12,13) = 9$.

№ 3-26

$c(1,2) = 6$; $c(1,6) = 8$; $c(1,7) = 4$; $c(2,3) = 1$; $c(2,7) = 6$; $c(2,8) = 4$; $c(2,9) = 2$; $c(3,4) = 8$; $c(3,9) = 5$; $c(4,12) = 10$; $c(5,7) = 4$; $c(5,11) = 3$; $c(5,13) = 9$; $c(6,5) = 16$; $c(6,7) = 8$; $c(7,8) = 4$; $c(7,10) = 4$; $c(7,11) = 6$; $c(7,13) = 6$; $c(8,4) = 11$; $c(8,10) = 3$; $c(8,12) = 12$; $c(9,4) = 2$; $c(9,8) = 8$; $c(10,12) = 9$; $c(10,13) = 10$; $c(11,13) = 5$; $c(12,13) = 4$.

4. Задача о максимальном потоке в графе (сети)

Дана транспортная сеть, представляющая собой связный ориентированный граф, состоящий из вершин $x_1 ... x_6$ и дуг, характеризуемых матрицей пропускных способностей $c(i, j)$.

Необходимо найти такое распределение потоков по дугам графа, при котором величина потока сети P была бы наибольшей для заданной структуры графа и заданных пропускных способностей.

№ 4-1

	x1	x2	x3	x4	x5	x6
x1		50	20			
x2				15	30	
x3						10
x4			20			5
x5				10		15
x6						

№ 4-2

	x1	x2	x3	x4	x5	x6
x1		12	11	19		
x2				10	7	
x3						14
x4			9		13	
x5						22
x6						

№ 4-3

	x1	x2	x3	x4	x5	x6
x1		25	29	27		
x2				10	19	
x3						20
x4			14		16	
x5			8			14
x6						

№ 4-4

	x1	x2	x3	x4	x5	x6
x1		29	30	25		
x2				11	17	
x3						29
x4			15		14	
x5			21			16
x6						

№ 4-5

	x1	x2	x3	x4	x5	x6
x1		38	43			
x2					41	
x3				18		12
x4		7			8	12
x5						14
x6						

№ 4-6

	x1	x2	x3	x4	x5	x6
x1		48	22			
x2				15	27	
x3						11
x4			25			6
x5				12		16
x6						

№ 4-7

	x1	x2	x3	x4	x5	x6
x1		21	12	19		
x2				11	10	
x3						15
x4				5	10	
x5						20
x6						

№ 4-8

	x1	x2	x3	x4	x5	x6
x1		29	17	10	20	
x2					8	19
x3				16		11
x4					17	
x5						15
x6						

№ 4-9

	x1	x2	x3	x4	x5	x6
x1		45	33	22		
x2				11	15	
x3						36
x4			13		25	
x5			15			26
x6						

№ 4-10

	x1	x2	x3	x4	x5	x6
x1		51	24			
x2				17	29	
x3						12
x4			27			7
x5				11		14
x6						

№ 4-11

	x1	x2	x3	x4	x5	x6
x1		55	17			
x2					42	
x3				18		35
x4		10			5	14
x5						11
x6						

№ 4-12

	x1	x2	x3	x4	x5	x6
x1		41	49			
x2					45	
x3				25		18
x4		8			7	10
x5						27
x6						

№ 4-13

	x1	x2	x3	x4	x5	x6
x1		17	54	18	20	
x2					9	25
x3				20		14
x4					31	
x5						15
x6						

№ 4-14

	x1	x2	x3	x4	x5	x6
x1		50	19			
x2				27	18	
x3						14
x4			15			10
x5				20		25
x6						

№ 4-15

	x1	x2	x3	x4	x5	x6
x1		37	27	17		
x2				6	11	
x3						20
x4			15		23	
x5						28
x6						

№ 4-16

	x1	x2	x3	x4	x5	x6
x1		38	30	25	15	
x2					10	19
x3				14		11
x4					15	
x5						16
x6						

№ 4-17

	x1	x2	x3	x4	x5	x6
x1		44	26			
x2				16	31	
x3						13
x4			19			7
x5				11		16
x6						

№ 4-18

	x1	x2	x3	x4	x5	x6
x1		25	10	20		
x2				6	8	
x3						15
x4			7		15	
x5						20
x6						

№ 4-19

	x1	x2	x3	x4	x5	x6
x1		35	45			
x2					51	
x3				26		16
x4		13			17	11
x5						27
x6						

№ 4-20

	x1	x2	x3	x4	x5	x6
x1		41	32	23	14	
x2					14	20
x3				15		10
x4					17	
x5						15
x6						

№ 4-21

	x1	x2	x3	x4	x5	x6
x1		40	30	20		
x2				15	15	
x3						30
x4			10		20	
x5			10			20
x6						

№ 4-22

	x1	x2	x3	x4	x5	x6
x1		25	40			
x2					21	
x3				32		19
x4		13			25	12
x5						19
x6						

№ 4-23

	x1	x2	x3	x4	x5	x6
x1		45	25			
x2				10	35	
x3						15
x4			17			10
x5				10		20
x6						

№ 4-24

	x1	x2	x3	x4	x5	x6
x1		40	50			
x2					35	
x3				27		12
x4		12			10	11
x5						16
x6						

№ 4-25

	x1	x2	x3	x4	x5	x6
x1		40	21			
x2				16	31	
x3						14
x4			19			9
x5				10		14
x6						

№ 4-26

	x1	x2	x3	x4	x5	x6
x1		20	17			
x2					14	
x3				24		19
x4		30			10	22
x5						16
x6						

5. Задача об оптимальном распределении заданного потока в транспортной сети

Величину потока P_0, подлежащую распределению по сети, принять равной $P_{max} - \Delta P$, ΔP – величина, задаваемая преподавателем для каждой учебной группы.

Ниже приведены матрицы удельных стоимостей, которые следует взять для расчетов по соответствующему варианту.

№ 5-1

	x1	x2	x3	x4	x5	x6
x1		50	20			
x2				8	5	
x3						10
x4			20			5
x5				10		3
x6						

№ 5-2

	x1	x2	x3	x4	x5	x6
x1		12	11	7		
x2				10	7	
x3						2
x4			9		13	
x5						5
x6						

№ 5-3

	x1	x2	x3	x4	x5	x6
x1		1	7	10		
x2				10	19	
x3						4
x4			14		16	
x5			8			14
x6						

№ 5-4

	x1	x2	x3	x4	x5	x6
x1		9	3	4		
x2				11	17	
x3						2
x4			15		14	
x5			21			16
x6						

№ 5-5

	x1	x2	x3	x4	x5	x6
x1		4	7			
x2					5	
x3				18		12
x4		7			8	12
x5						14
x6						

№ 5-6

	x1	x2	x3	x4	x5	x6
x1		2	9			
x2				15	27	
x3						11
x4			25			6
x5				12		16
x6						

№ 5-7

	x1	x2	x3	x4	x5	x6
x1		5	12	19		
x2				11	10	
x3						15
x4				5	10	
x5						6
x6						

№ 5-8

	x1	x2	x3	x4	x5	x6
x1		7	17	10	11	
x2					8	19
x3				16		11
x4					17	
x5						2
x6						

№ 5-9

	x1	x2	x3	x4	x5	x6
x1		45	3	4		
x2				11	15	
x3						5
x4			13		25	
x5			15			6
x6						

№ 5-10

	x1	x2	x3	x4	x5	x6
x1		5	24			
x2				17	3	
x3						12
x4			1			7
x5				11		14
x6						

№ 5-11

	x1	x2	x3	x4	x5	x6
x1		5	17			
x2					42	
x3				18		2
x4		10			5	14
x5						11
x6						

№ 5-12

	x1	x2	x3	x4	x5	x6
x1		5	11			
x2					45	
x3				25		18
x4		8			7	10
x5						8
x6						

№ 5-13

	x1	x2	x3	x4	x5	x6
x1		17	5	18	20	
x2					9	3
x3				20		14
x4					31	
x5						8
x6						

№ 5-14

	x1	x2	x3	x4	x5	x6
x1		50	1			
x2				5	18	
x3						14
x4			15			10
x5				20		7
x6						

№ 5-15

	x1	x2	x3	x4	x5	x6
x1		3	9	17		
x2				6	11	
x3						20
x4			15		23	
x5						5
x6						

№ 5-16

	x1	x2	x3	x4	x5	x6
x1		5	30	3	15	
x2					10	19
x3				5		11
x4					15	
x5						16
x6						

№ 5-17

	x1	x2	x3	x4	x5	x6
x1		8	7			
x2				16	31	
x3						13
x4			19			7
x5				11		5
x6						

№ 5-18

	x1	x2	x3	x4	x5	x6
x1		7	10	5		
x2				6	8	
x3						15
x4			7		15	
x5						2
x6						

№ 5-19

	x1	x2	x3	x4	x5	x6
x1		35	6			
x2					1	
x3				26		16
x4		13			17	11
x5						4
x6						

№ 5-20

	x1	x2	x3	x4	x5	x6
x1		1	32	5	14	
x2					14	20
x3				15		10
x4					17	
x5						10
x6						

№ 5-21

	x1	x2	x3	x4	x5	x6
x1		5	30	20		
x2				15	15	
x3						10
x4			10		20	
x5			10			20
x6						

№ 5-22

	x1	x2	x3	x4	x5	x6
x1		2	40			
x2					11	
x3				32		19
x4		13			2	12
x5						19
x6						

№ 5-23

	x1	x2	x3	x4	x5	x6
x1		45	2			
x2				10	3	
x3						15
x4			7			10
x5				10		20
x6						

№ 5-24

	x1	x2	x3	x4	x5	x6
x1		4	7			
x2					35	
x3				27		12
x4		12			10	11
x5						10
x6						

№ 5-25

	x1	x2	x3	x4	x5	x6
x1		5	21			
x2				16	4	
x3						14
x4			8			9
x5				10		14
x6						

№ 5-26

	x1	x2	x3	x4	x5	x6
x1		20	17			
x2					14	
x3				6		19
x4		5			10	7
x5						16
x6						

Глава 5. **НЕЛИНЕЙНЫЕ ЗАДАЧИ СТАТИЧЕСКОЙ ОПТИМИЗАЦИИ**

§ 5.1. Аналитическое решение нелинейных задач статической оптимизации

Рассмотрим задачу оптимизации об отыскании значений $x_i = x_i^0$ ($i = \overline{1,n}$), образующих вектор $\bar{x} = \bar{x}^0$, при которых ЦФ $y = f(\bar{x})$ принимает наименьшее (либо наибольшее) значение при выполнении условий-ограничений: $h_j(\bar{x}) = 0, \quad g_l(\bar{x}) \ge 0; \quad j = \overline{1,m}, \quad l = \overline{1,k}.$

Как отмечалось в гл. 1, в частных случаях задача оптимизации может не содержать ограничений. При этом решение задачи, если оно существует, совпадает с одним из локальных экстремумов ЦФ, называемым глобальным экстремумом. Отыскание решения задачи безусловной оптимизации сводится к выявлению всех локальных экстремумов, вычислению в каждом из них ЦФ и выбору наименьшего из минимумов (либо наибольшего из максимумов).

При нахождении локальных экстремумов используется необходимое условие экстремума, в соответствии с которым в экстремальной точке $\bar{x} = \bar{x}^0$ все частные производные дифференцируемой функции равны нулю:

$$\frac{\partial y}{\partial x_i} = 0, \quad i = \overline{1,n} \quad \text{или} \quad \nabla f = \text{grad} f = 0, \quad \nabla f = \left[\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right]^T. \quad (5.1)$$

Точки, удовлетворяющие (5.1), называются стационарными. Следует подчеркнуть, что все экстремумы дифференцируемой функции находятся среди стационарных точек, но среди стационарных точек могут оказаться и не экстремальные. Заметим, что экстремум может оказаться в точке, где функция не дифференцируема.

Для того чтобы проверить, является ли стационарная точка $\bar{x}^0$ экстремальной и определить ее характер (min или max), можно использовать достаточное условие экстремума. Оно формулируется с помощью матрицы H вторых производных, обозначаемой $\nabla^2 f$ и называемой матрицей Гессе. Ее элементы $h_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

Если $f(\bar{x})$ имеет в некоторой окрестности стационарной точки непрерывные вторые частные производные, то в случае, когда матрица Гессе отрицательно определена, точка $\bar{x}^0$ является точкой максимума, а в случае, когда она положительно определена, $\bar{x}^0$ – точка минимума.

Матрица H положительно определена, когда все ее определители $\Delta_i (i=\overline{1,n})$, получаемые выделением первых i строк и i столбцов, положительны; если же $(-1)^i \Delta_i > 0$ для $i=\overline{1,n}$, то матрица отрицательно определена.

Рассмотренные условия позволяют в некоторых случаях получить решение поставленной оптимизационной задачи.

Пример 1. Исследовать на экстремум функцию

$$y = 30 + x_1^3 + 3x_1 x_2{}^2 - 15x_1 - 12x_2.$$

Решение. Используя условие (5.1), получаем:

$$\frac{\partial y}{\partial x_1} = 3x_1^2 + 3x_2^2 - 15 = 0; \quad x_1^2 + x_2^2 = 5;$$

$$\frac{\partial y}{\partial x_2} = 6x_1 x_2 - 12 = 0; \quad x_1 x_2 = 2.$$

Подставив $x_2 = 2/x_1$ в 1-е уравнение, полученную систему сводим к уравнению: $x_1^4 - 5x_2^2 + 4 = 0$. Решив его, получаем стационарные точки:

$$\bar{x}_A(1;\ 2), \quad \bar{x}_B(2;\ 1), \quad \bar{x}_C(-1;\ -2), \quad \bar{x}_Д(-2;\ -1).$$

Составляем матрицу Гессе

$$H = \begin{bmatrix} 6x_1 & 6x_2 \\ 6x_2 & 6x_1 \end{bmatrix}.$$

Подставляя в H координаты найденных точек, получаем:

$$H_A = \begin{bmatrix} 6 & 12 \\ 12 & 6 \end{bmatrix}, \quad \begin{matrix} \Delta_1 = 6 > 0 \\ \Delta_2 = -108 < 0 \end{matrix}; \quad H_B = \begin{bmatrix} 12 & 6 \\ 6 & 12 \end{bmatrix}, \quad \begin{matrix} \Delta_1 = 12 > 0 \\ \Delta_2 = 108 > 0 \end{matrix};$$

$$H_C = \begin{bmatrix} -6 & -12 \\ -12 & -6 \end{bmatrix}, \quad \begin{matrix} \Delta_1 = -6 < 0 \\ \Delta_2 = -108 < 0 \end{matrix}; \quad H_Д = \begin{bmatrix} -12 & -6 \\ -6 & -12 \end{bmatrix}, \quad \begin{matrix} \Delta_1 = -12 < 0 \\ \Delta_2 = 108 > 0 \end{matrix}.$$

Анализируя знаки определителей, заключаем, что $\bar{x}_B$ – точка минимума, в ней $y_{\min} = 2$; $\bar{x}_Д$ – точка максимума, в ней $y_{\max} = 58$; в $\bar{x}_A$ и $\bar{x}_C$ – экстремума нет.

Пример 2. В эксперименте регистрируется процесс $y^Э(t)$, представляющий собой сумму плавно изменяющейся так называемой полезной составляющей и быстро изменяющейся компоненты, порожденной действием случайных факторов (рис. 5.1).

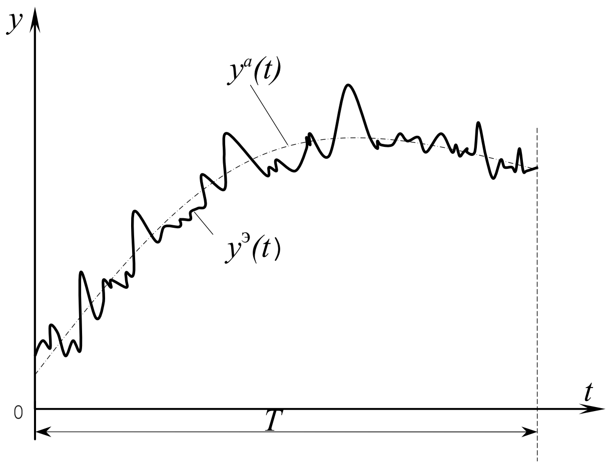


Рис. 5.1

Требуется по результатам наблюдений $y^{э}(t)$, выполненным на промежутке [0, *T*], найти аналитическое описание процесса в виде некоторой «плавной» функции $y^{a}(t)$, сглаживающей получаемый экспериментально процесс и приближенно воспроизводящей полезную составляющую процесса.

Решение. Примем в качестве сглаживающей функции

$$y^{a}(t) = A_0 + A_1\left(\frac{t}{T}\right) + A_2\left(\frac{t}{T}\right)^2, \quad 0 \le t \le T. \qquad (5.2)$$

Ее параметры A_0, A_1, A_2 подберем так, чтобы интегральное квадратическое отклонение

$$Q = \int_0^T \left[y^{э}(t) - y^{a}(t)\right]^2 dt = \int_0^T \left[y^{э}(t) - A_0 - A_1\left(\frac{t}{T}\right) - A_2\left(\frac{t}{T}\right)^2 \right]^2 dt$$

было минимальным.

Используя необходимое условие экстремума (5.1), получаем:

$$\frac{\partial Q}{\partial A_0} = \int_0^T 2\left[y^{э}(t) - A_0 - A_1\left(\frac{t}{T}\right) - A_2\left(\frac{t}{T}\right)^2 \right] (-1)dt = 0,$$

или

$$TA_0 + \frac{T}{2}A_1 + \frac{T}{3}A_2 = \int_0^T y^{э}(t)dt.$$

Аналогично рассматривая условие $\dfrac{\partial Q}{\partial A_1} = 0$ и $\dfrac{\partial Q}{\partial A_2} = 0$, получаем систему уравнений:

$$A_0 + \frac{1}{2}A_1 + \frac{1}{3}A_2 = s_0; \;\; \frac{1}{2}A_0 + \frac{1}{3}A_1 + \frac{1}{4}A_2 = s_1; \;\; \frac{1}{3}A_0 + \frac{1}{4}A_1 + \frac{1}{5}A_2 = s_2, \qquad (5.3)$$

где $s_0 = \frac{1}{T}\int_0^T y^э(t)dt; \quad s_1 = \frac{1}{T}\int_0^T y^э(t)\left(\frac{t}{T}\right)dt; \quad s_2 = \frac{1}{T}\int_0^T y^э(t)\left(\frac{t}{T}\right)^2 dt.$ (5.4)

Решив систему линейных алгебраических уравнений (5.3), получаем следующие выражения для неизвестных параметров аппроксимирующей параболы (5.2):

$$\begin{cases} A_0 = 9s_0 - 36s_1 + 30s_2, \\ A_1 = -36s_0 + 192s_1 - 180s_2 \\ A_2 = 30s_0 - 180s_1 + 180s_2. \end{cases} \tag{5.5}$$

Отметим, что в случае, когда процесс $y^э(t)$ регистрируется в дискретные моменты времени $t_0 = 0,\ t_1 = h, t_2 = 2h, ...,\ t_k = kh, ...,\ t_N = Nh = T$, разбивающие промежуток $[0,\ T]$ на достаточно мелкие отрезки длины h, для вычисления интегралов (5.4), входящих в (5.5), можно использовать численные методы, в частности формулу прямоугольников [17]. Тогда

$$s_0 \approx \frac{1}{N}\sum_{k=1}^{N} y_k^э; \quad s_1 \approx \frac{1}{N}\sum_{k=1}^{N} y_k^э\left(\frac{kh}{T}\right); \quad s_2 \approx \frac{1}{N}\sum_{k=1}^{N} y_k^э\left(\frac{kh}{T}\right)^2,$$

где $y_k^э = y^э(kh)$.

Рассмотрим теперь задачу оптимизации, в которой наряду с ЦФ задана система ограничений-равенств, т.е. задачу на условный экстремум. Ее можно свести к задаче на безусловный экстремум, используя один из следующих двух приемов.

1. Допустим, что систему m уравнений удается разрешить относительно каких-либо m переменных из общего списка n неизвестных $x_i (m < n)$, например:

$$\left.\begin{aligned} x_1 &= \varphi_1(x_{m+1}, ..., x_n), \\ x_2 &= \varphi_2(x_{m+1}, ..., x_n), \\ &\ldots \\ x_m &= \varphi_m(x_{m+1}, ..., x_n). \end{aligned}\right\} \tag{5.6}$$

Подставляя соотношения (5.6) в ЦФ, получаем задачу об отыскании $n-m$ переменных $x_{m+1}, ..., x_n$, при которых функция

$$y = f(\varphi_1(x_{m+1}, ..., x_n), ..., \varphi_m(x_{m+1}, ..., x_n)) = \tilde{f}(x_{m+1}, ..., x_n) \tag{5.7}$$

принимает наименьшее значение. Подставив полученные в результате решения задачи на безусловный экстремум для функции (5.7) значения $x_{m+1}^0, ..., x_n^0$ в (5.6), находим полный набор неизвестных $x_1^0, ..., x_n^0$.

2. Для задачи (1.7) – (1.8) введем вспомогательную функцию:

$$\tilde{y} = f(\bar{x}) + \sum_{k=1}^{m} \lambda_k h_k(\bar{x}),$$

называемую функцией Лагранжа. В нее входят: заданная ЦФ *f*, выражения h_k, стоящие в левых частях ограничений-равенств (1.8), и вспомогательные переменные λ_k, называемые множителями Лагранжа.

Для решения задачи в этом случае можно использовать необходимое условие $\dfrac{\partial \tilde{y}}{\partial x_i} = 0, \quad i = \overline{1,n}; \quad \dfrac{\partial \tilde{y}}{\partial \lambda_j} = 0, \quad j = \overline{1,m}.$

Пример. Для формализованной в примере 1 из § 1.1 задачи найдем параметры *R* и *H*, при которых ЦФ:

$$s = 2\pi R^2 + 2\pi RH \tag{5.8}$$

минимальна при выполнении условия

$$\pi R^2 H - V_0 = 0. \tag{5.9}$$

Используя метод исключения, из (5.9) выражаем $H = V_0 / \pi R^2$.

Подставляя это соотношение в (5.8), получаем ЦФ: $s = 2\pi R^2 + \dfrac{2V_0}{R}$, зависящую только от *R*. Из условия $ds/dR = 4\pi R - 2V_0/R^2 = 0$ находим

$$R^* = \sqrt[3]{V_0/2\pi}; \quad H^* = \frac{V_0}{\pi}\left(\frac{V_0}{2\pi}\right)^{-2/3} = 2\sqrt[3]{V_0/2\pi} = 2R^*.$$

Заметим, что при найденных значениях параметров $R = R^*$ и $H = H^*$:

$$\frac{d^2 s}{dR^2} = 4\pi + 2V_0 \frac{1}{R^3} = 4\pi + 4V_0 \frac{1}{V_0/2\pi} = 12\pi > 0,$$

следовательно, эти параметры действительно дают минимум *s*.

Пользуясь методом множителей Лагранжа, сформируем функцию: $\tilde{s} = 2\pi R^2 + 2\pi RH + \lambda(\pi R^2 H - V_0)$ и для нее запишем условия:

$$\begin{cases} \dfrac{\partial \tilde{s}}{\partial R} = 4\pi R + 2\pi H + 2\pi\lambda RH = 0; \\ \dfrac{\partial \tilde{s}}{\partial H} = 2\pi R + \pi\lambda R^2 = 0; \\ \dfrac{\partial \tilde{s}}{\partial \lambda} = \pi R^2 H - V_0 = 0. \end{cases}$$

Из второго уравнения получаем $\lambda^* = -2/R$. Подстановка этого соотношения в 1-е уравнение дает $4\pi R + 2\pi H - 4\pi H = 0; \quad H^* = 2R^*$.

Задача, в которой ЦФ (1.7) рассматривается при наличии ограничений-неравенств (1.9), относится к задачам математического (нелинейного) программирования. При этом обычно неравенства (1.9) задают в пространстве переменных $x_1,...,x_n$ некоторую замкнутую ограниченную область *D*. Следовательно, нужно найти такую точку $\overline{x} \in D$, для которой ЦФ (1.7) является наименьшей (или наибольшей). Ясно, что искомая точка может оказаться внутри *D*, тогда она является точкой экстремума, либо на границе *D*. Учитывая это, для решения рассматриваемой задачи можно поступить следующим образом:

1) найти все локальные экстремумы ЦФ (1.7), выбрать те из них, которые удовлетворяют (1.9), т.е. лежат в *D*, а остальные исключить из рассмотрения;

2) вычислить ЦФ в выделенных точках;

3) найти наибольшие (или наименьшие) значения ЦФ на линиях (поверхностях), образующих границу *D*;

4) из всех найденных значений (в п. 2 и 3) выбрать наибольшее (наименьшее).

Пример. Найти наибольшее и наименьшее значения ЦФ:

$$y = x_1^2 + x_2^2 - x_1x_2 + x_1 + x_2$$

при выполнении условий:

$$x_1 \le 0, \quad x_2 \le 0, \quad x_1 + x_2 \ge -3.$$

Эти условия задают область *D* в виде треугольника *OAB* (рис. 5.2).

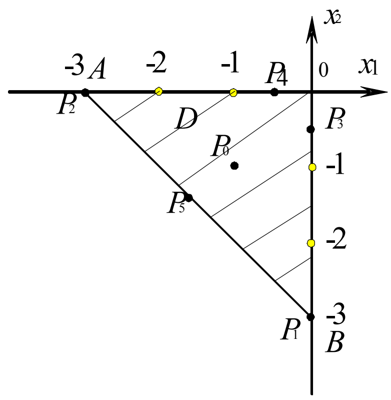


Рис. 5.2

Решение 1. Находим стационарные точки из условий:

$$\partial y / \partial x_1 = 2x_1 - x_2 + 1 = 0;$$

$$\partial y / \partial x_2 = 2x_2 - x_1 + 1 = 0.$$

В результате получаем $x_1 = -1$, $x_2 = -1$ (точка P_0 на рис. 5.2). В ней $y = -1$.

2. Рассматриваем функцию на границах области D (см. рис. 5.2):

OB: $x_1 = 0, \quad y = x_2^2 + x_2; \quad -3 \le x_2 \le 0;$

$$y'_{x2} = 2x_2 + 1 = 0, x_2 = -1/2, y(x_2 = -1/2) = -1/4,$$

кроме того, $y(x_2 = -3) = 6, \quad y(x_2 = 0) = 0.$

Вывод: на этой границе $y_{\sup} = 6$ при $(x_1 = 0; \quad x_2 = -3)$ – точка P_1;

$y_{\inf} = -1/4$ при $(x_1 = 0; \quad x_2 = -1/2)$ – точка P_3.

OA: $x_2 = 0; \quad y = x_1^2 + x_1$, аналогично получаем

$$y_{\sup} = 6 \text{ при } (x_1 = -3; \quad x_2 = 0) \text{ – точка } P_2;$$

$$y_{\inf} = -1/4 \text{ при } (x_1 = -1/2; \quad x_2 = 0) \text{ – точка } P_4.$$

AB: $x_2 = -3 - x_1$; $\quad y = 3x_1^2 + 9x_1 + 6, \quad -3 \le x_1 \le 0$, аналогично находим:

$$y_{\inf} = -3/4 \text{ при } (x_1 = -3/2; \quad x_2 = -3/2) \text{ – точка } P_5;$$

$$y_{\sup} = 6 \text{ в точках } P_1 \text{ и } P_2.$$

Сопоставляя все полученные значения ЦФ, заключаем, что наибольшего значения $y = 6$ эта функция достигает в угловых точках P_1 и P_2 области D, а наименьшее значение $y = -1$ получается в стационарной точке P_0.

Рассмотренный в данном параграфе материал показывает, что нелинейные задачи оптимизации (как при отсутствии, так и при наличии ограничений) в отдельных случаях допускают аналитическое решение.

Обратимся теперь к следующей задаче. Найти значение $x = x^*$, при котором $y = \frac{1}{2}x^2 + e^{-x}$ принимает наименьшее значение.

Наличие решения, совпадающего с локальным минимумом этой довольно простой функции, здесь очевидно. Его отыскание с помощью необходимого условия экстремума приводит к уравнению

$$y' = x - e^{-x} = 0, \quad \text{или} \quad e^{-x} = x,$$

которое аналитически невозможно разрешить.

К сожалению, при решении практических задач подобная ситуация, когда необходимое условие экстремума (5.1) приводит к неразрешимой аналитически системе уравнений, довольно типична. Кроме того, нередко встречаются задачи, в которых невозможно получить аналитическое выражение для частных производных от ЦФ. Например, в тех случаях, когда аналитическое выражение для ЦФ отсутствует, а значения ЦФ получаются в результате экспериментов (наблюдений). Во всех подобных случаях для решения оптимизационной задачи приходится использовать численные методы.

§ 5.2. Численные методы решения одномерных задач статической оптимизации

Рассмотрим задачу об отыскании значения переменной $x = x^*$ из заданного промежутка $[a, b]$, при котором ЦФ $y = f(x)$ принимает наименьшее значение. Эта задача минимизации функции одной переменной имеет самостоятельное практическое значение. Кроме того, она нередко возникает как вспомогательная задача при многомерной оптимизации.

Будем исходить из того, что в нашем распоряжении имеется ЦВМ, позволяющая вычислять значения функции f (а возможно, и ее производных) в любой выбранной пробной точке x_k. Решение задачи численным методом состоит в такой организации выбора пробных точек x_k, при которой в результате выполнения некоторого количества вычислений будет найдено приближенное значение искомой величины x^*.

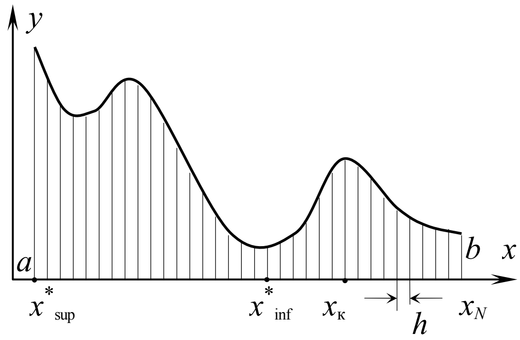


Рис. 5.3

Для рассматриваемой одномерной задачи этот подход можно реализовать путем разбиения отрезка $[a, b]$ на N достаточно мелких частей длины h; вычисления функции в полученных точках:
$x_0 = a;\ x_1 = a + h;...;\ x_k = a + kh;...;$
$x_N = a + Nh = b$; сравнения значений функции $y_0, y_1,...$ в этих точках и выбора точки, в которой значение этой функции оказалось наименьшим (рис. 5.3).

Достоинством этого метода, называемого методом равномерного пассивного поиска, является то, что при достаточно мелком шаге h он позволяет находить решение и в том случае, когда функция имеет несколько экстремумов. Его недостаток – большой объем вычислений, необходимых для отыскания решения с заданной точностью ε (нужно выбирать $N \geq \dfrac{b-a}{\varepsilon}$). Блок-схема алгоритма, реализующего этот метод, представлена на рис. 5.4.

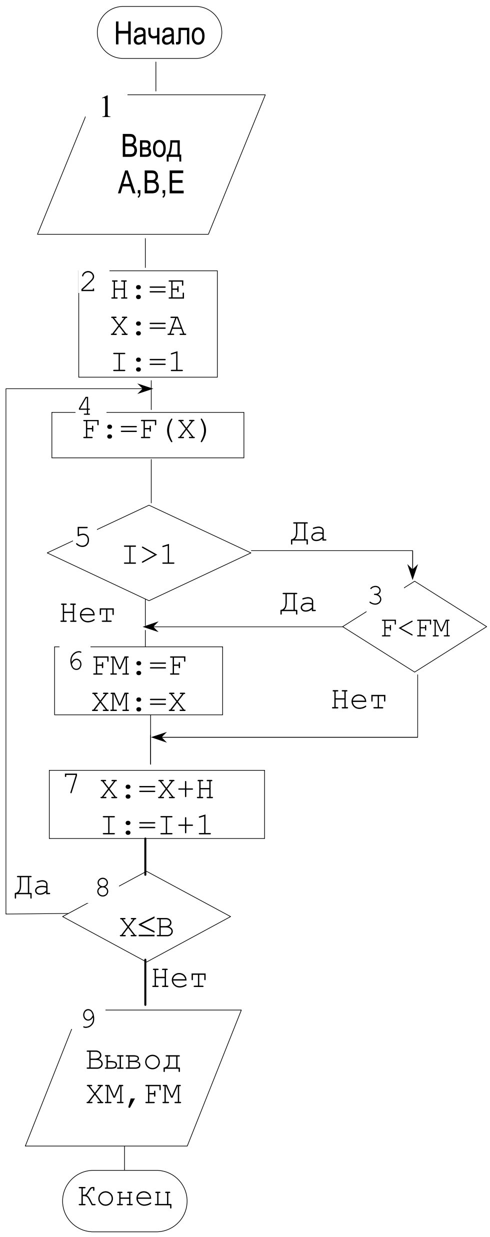


Рис. 5.4

Программирование конкретной функции в этой схеме предусмотрено в блоке 4.

В тех случаях, когда в рассматриваемом промежутке заданная функция имеет один экстремум, она называется унимодальной. Для минимизации такой функции существует ряд эффективных алгоритмов (метод Фибоначчи, методы золотого сечения, дихотомии, квадратичной аппроксимации и др.) [2, 15, 26, 37]. Рассмотрим один из них – простой и весьма эффективный метод золотого сечения (МЗС). Он основывается на очевидном свойстве унимодальной функции (рис. 5.5, *а*), в соответствии с которым наличие значений функции в двух точках, выбранных на интервале L, содержащем экстремум и называемом интервалом неопределенности, позволяет отбросить часть этого интервала как заведомо не содержащую экстремального значения, и тем самым сократить интервал неопределенности (рис. 5.5, *б*, *в*).

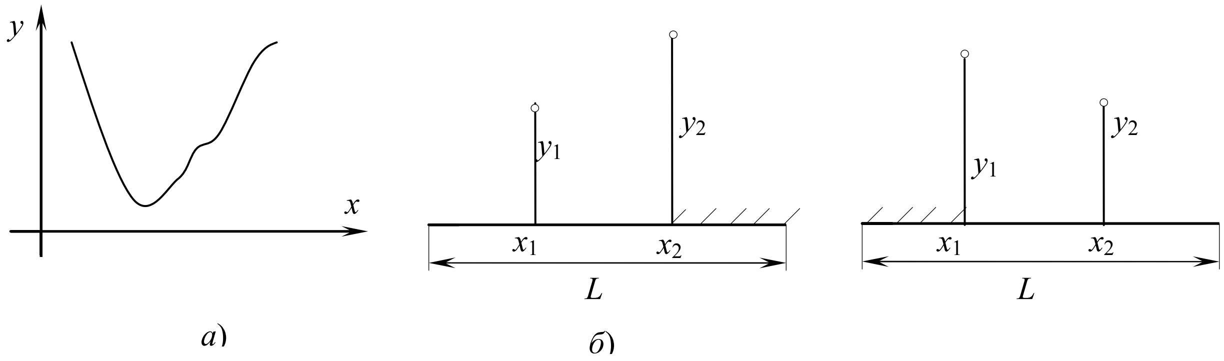


Рис. 5.5

Эту процедуру нужно повторять до тех пор, пока длина интервала не станет меньше допустимой погрешности ε.

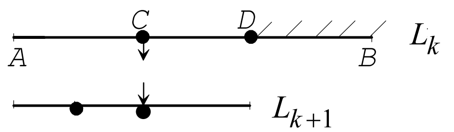


Рис. 5.6

Особенностью МЗС является то, что при выборе пробных точек используется правило, при котором одна из точек в L_{k+1} совпадает с одной из точек в L_k (рис. 5.6).

Благодаря этому на каждом интервале (кроме начального) нужно вычислять функцию не в двух, а в одной точке. Можно показать, что такое сечение получается в том случае, когда на отрезке AB выбираются точки C и D, удовлетворяющие соотношению

$$\frac{AB}{AD} = \frac{AD}{DB},$$

называемому золотым сечением. Такое сечение дают точки:

$$x_C = x_A + (1-\tau)(x_B - x_A), \quad x_D = x_A + \tau(x_B - x_A),$$

где $\tau = (-1+\sqrt{5})/2 = 0{,}618034$.

Блок-схема алгоритма, реализующего МЗС, представлена на рис. 5.7. В блоках 4, 10, 13 предусмотрено обращение к подпрограмме *F* вычисления заданной целевой функции; величина *D* представляет собой длину текущего интервала неопределенности.

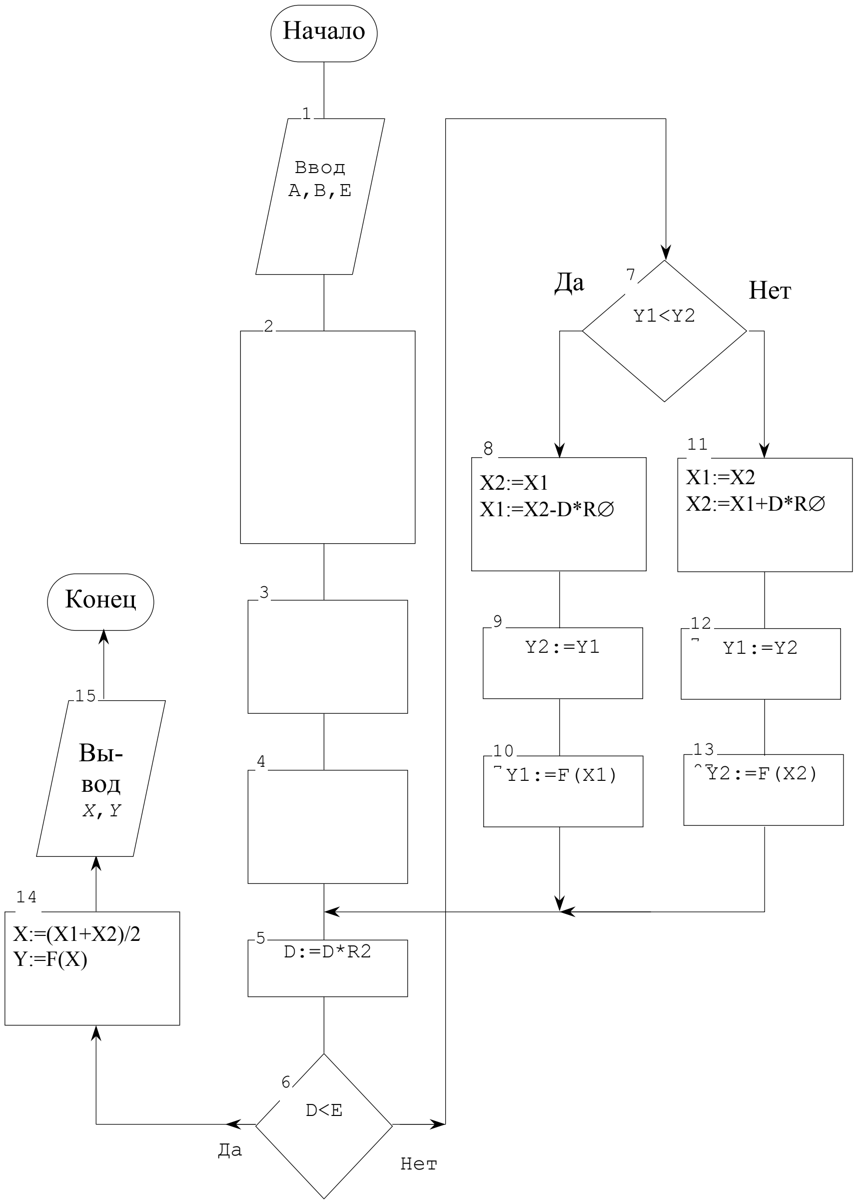


Рис. 5.7

В тех случаях, когда возможно вычисление производных от ЦФ, для отыскания оптимального решения можно использовать метод Ньютона – Рафсона. В соответствии с ним для отыскания точки x^* минимума $f(x)$

строится числовая последовательность $x_0, x_1, x_2, ...$ по рекуррентному соотношению

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \tag{5.10}$$

где x_0 – некоторое начальное значение.

Как только для некоторого k будет получено x_k, для которого $f'(x_k)$ – мало, процесс вычислений прекращают, полагая $x^* \approx x_k$.

Если вычисление f'' невозможно или затруднено, то поиск можно вести по соотношению

$$x_{k+1} = x_k - \gamma f'(x_k), \tag{5.11}$$

в котором γ – некоторое число, подлежащее выбору.

§ 5.3. Численные методы многомерной безусловной оптимизации с использованием производных

Для поиска значений $x_i = x_i^* \ (i = \overline{1,n})$, при которых ЦФ $y = f(x_1,...,x_n) = f(\bar{x})$ принимает наименьшее значение, можно использовать распространение метода Ньютона-Рафсона (5.10) на многомерный случай

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \Gamma_k \nabla f(\bar{x}^{(k)}), \tag{5.12}$$

где $\Gamma_k = H_k^{-1}; H_k = \nabla^2 f(\bar{x}^{(k)})$ – матрица Гессе; $\bar{x}^{(0)}$ – произвольная начальная точка.

При выполнении определенных условий, в частности, для дважды дифференцируемой функции *f*, имеющей положительно определенную матрицу Гессе, получаемая по (5.12) последовательность точек $\bar{x}^{(0)}$, $\bar{x}^{(1)}$,... сходится к искомой экстремальной точке, причем в некоторых случаях скорость сходимости оказывается очень высокой. Так, для квадратичной ЦФ решение при произвольной $\bar{x}^{(0)}$ получается за один шаг [26].

При практическом использовании (5.12) могут возникнуть сложности с вычислением вторых производных, а главное – с обращением матрицы *H*, так как она часто оказывается близкой к вырожденной.

Возможны различные модификации алгоритма (5.12), не использующие вторых производных. При этом в качестве матрицы Г в (5.12) берут, например,

$$\Gamma = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & \gamma_n \end{bmatrix},$$

где γ_i – некоторые числа. Тогда в развернутом виде (5.12) принимает следующий вид:

$$\begin{aligned} x_1^{(k+1)} &= x_1^{(k)} - \gamma_1 \left.\frac{\partial f}{\partial x_1}\right|_{\bar{x}=\bar{x}^{(k)}}; \\ . \quad . \quad & . \quad . \quad . \\ x_n^{(k+1)} &= x_n^{(k)} - \gamma_n \left.\frac{\partial f}{\partial x_n}\right|_{\bar{x}=\bar{x}^{(k)}}. \end{aligned} \tag{5.13}$$

Характерно, что при $\gamma_1 = ... = \gamma_n$ в этом случае переход от $\bar{x}^{(k)}$ к $\bar{x}^{(k+1)}$ осуществляется в направлении антиградиента, т.е. направлении наибольшего убывания ЦФ. Поэтому рассматриваемую процедуру поиска минимума называют методом градиентного спуска. Возможны различные модификации этого метода, отличающиеся способом выбора длины шага в указанном направлении. Так, при градиентном методе с дроблением шага [37] числа $\gamma_1,...,\gamma_k$ принимаются равными некоторому числу $\gamma = \gamma_0$ и проводятся вычисления по (5.13), контролируя на каждом шаге выполнение условия монотонного убывания функции:

$$f(\bar{x}^{(k+1)}) < f(\bar{x}^{(k)}).$$

Если оно выполняется, то γ_0 остается неизменным; при его нарушении это число уменьшается до тех пор, пока монотонность не восстановится.

Для окончания вычислений можно использовать различные критерии, в частности,

$$\left\|\nabla f(\bar{x}^{(k)})\right\| = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\right)^2} < \varepsilon.$$

На рис. 5.8. представлена блок-схема этого алгоритма.

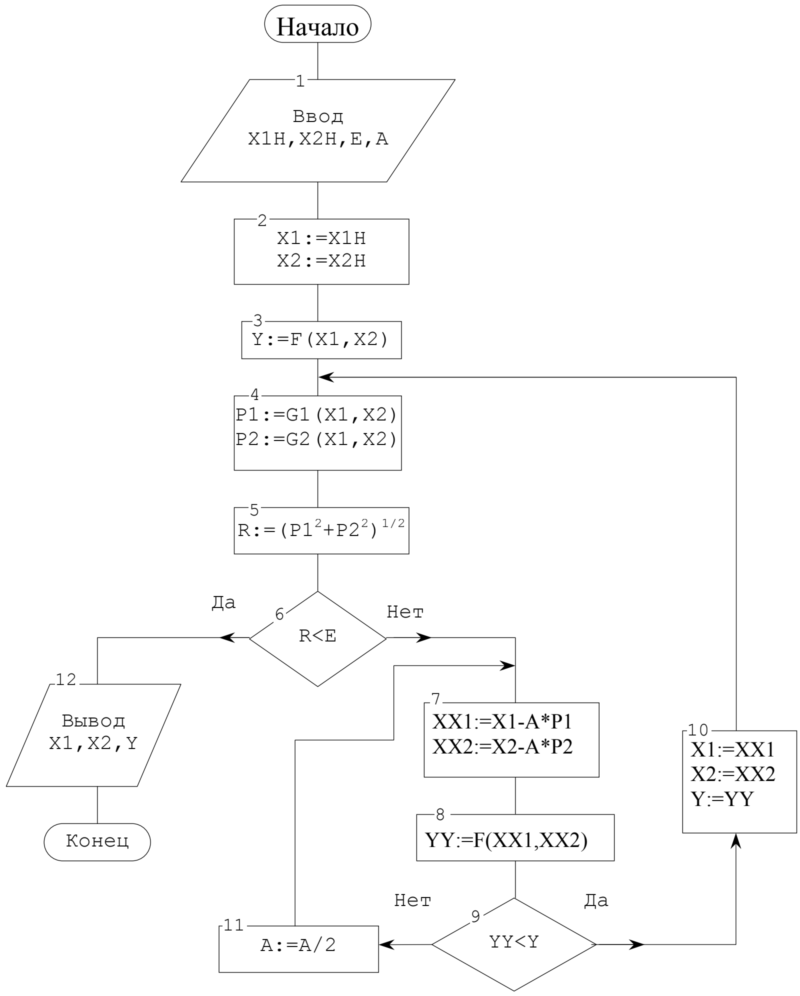


Рис. 5.8

Для простоты рассмотрен случай ЦФ двух переменных. В программе предусмотрен ввод координат начальной точки *X1H*, *X2H*, начального значения параметра $\gamma_0 = A$ и малого числа $\varepsilon = E$.

Для рассмотренных модификаций градиентных методов характерна медленная сходимость (большое количество необходимых для достижения экстремума итераций) при обработке функций, у которых одни переменные на них сильно влияют, а другие – незначительно. В этом случае график ЦФ двух переменных представляет собой поверхность в виде длинного узкого оврага, при этом линии равного уровня функции представляют собой вытянутые замкнутые линии (рис. 5.9).

При использовании градиентных методов (5.13) для таких функций обычно происходит быстрый спуск со склона на дно оврага, а затем медленное зигзагообразное движение к точке минимума. В таких случаях гораздо эффективнее с точки зрения ускорения сходимости поиска оказывается модификация градиентного метода, называемая методом сопряженных градиентов, в частности, его вариант, предложенный Флетчером и Ривсом [37]. Он состоит в выполнении следующих операций:

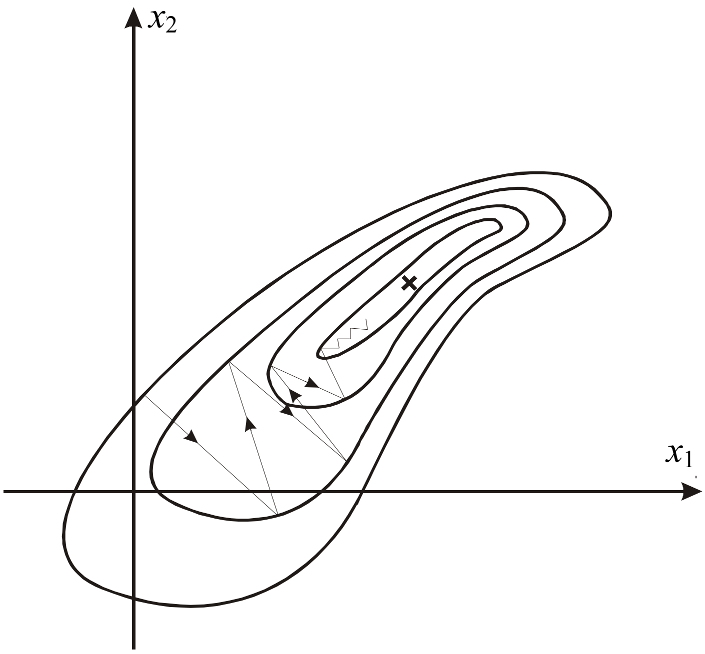


Рис. 5.9

1. Для выбранной начальной точки $\bar{x}^{(0)}$ вычисляется антиградиент:

$$\bar{s}^{(0)} = -\nabla f(\bar{x}^{(0)}),$$

2. На k-м шаге (k = 1, 2,…) рассматривается функция

$$f(\bar{x}^{(k)} + \alpha\, \bar{s}^{(k)})$$

и, используя какую-либо процедуру одномерного поиска, находится число $\alpha = \alpha_k^*$, при котором эта функция принимает минимальное значение; фиксируется получающаяся при этом точка минимума

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \alpha_k^* \bar{s}^{(k)},.$$

3. Вычисляются

$$f(\bar{x}^{(k+1)}) \text{ и } \nabla f(\bar{x}^{(k+1)}).$$

4. Проверяется условие прекращения поиска, например

$$\left\| \nabla f(\bar{x}^{(k+1)}) \right\| = \sqrt{\sum \left(\frac{\partial f}{\partial x_i} \right)^2} < \varepsilon$$

или $$\left\| \bar{x}^{(k+1)} - \bar{x}^{k}) \right\| = \sqrt{\sum \left[x_i^{(k+1)} - x_i^{(k)} \right]^2} < \varepsilon,$$

при его выполнении решение прекращается, принимая $\bar{x}^{(k+1)} \approx \bar{x}^*$ за искомое решение, иначе осуществляется переход к следующей операции.

5. Находится величина

$$\beta_k = \begin{cases} c_k / d_k, & \text{если} \quad k+1 \notin I, \\ 0, & \text{если} \quad k+1 \in I, \end{cases}$$

где числа c_k и d_k определяются как скалярные произведения

$$d_k = \left[\nabla f(\bar{x}^{(k)})\right]^T \left[\nabla f(\bar{x}^{(k)})\right];$$

$$c_k = \left[\nabla f(\bar{x}^{(k+1)})\right]^T \left[\nabla f(\bar{x}^{(k+1)} - \nabla f(\bar{x}^{(k)})\right];$$

$I = \{0, n, 2n, ...\}$ – множество индексов.

6. Вычисляется $s^{(k+1)} = -\nabla f(\bar{x}^{(k+1)}) + \beta_k s^{(k)}$, а затем все процедура повторяется.

Реализующая эти операции укрупненная блок-схема алгоритма представлена на рис. 5.10 [37].

В ней предусмотрены: ввод начальной точки $\overline{XH}$ и допустимой погрешности $E = \varepsilon$, вычисление ЦФ и ее частотных производных в подпрограммах F и G, определение величин:

$$sp = c_k; \quad pN = d_k; \quad BET = \beta_k.$$

Пример. Найти наименьшее значение функции Розенброка [6]:

$$y = 100\,(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Заметим, что искомое решение здесь, очевидно, $y_{\min} = 0$ при $x_1^* = 1$, $x_2^* = 1$. Однако отыскание его численным методом в данном случае затруднено из-за того, что эта функция плохо масштабирована (1-е слагаемое в отличие от 2-го входит с большим коэффициентом), поэтому она относится к функциям овражного типа. Ее линии равного уровня показаны на рис. 5.11. Указанный особый характер этой функции при простоте аналитического выражения обусловил широкое использование ее в качестве тестовой функции при исследовании эффективности различных алгоритмов оптимизации [2, 37].

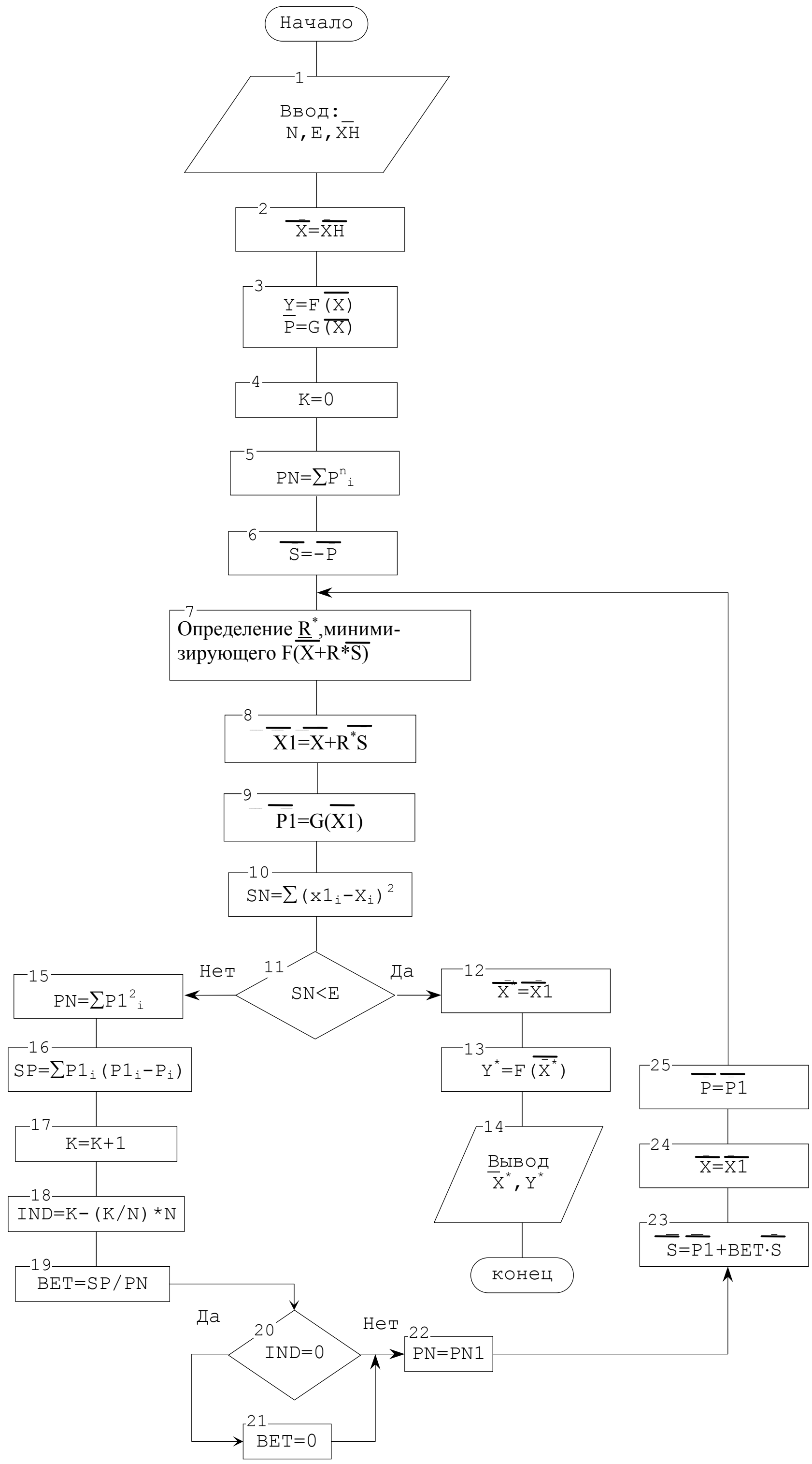

Начало
Ввод:
N,E,XH
X=XH
Y=F(X)
P=G(X)
K=0
PN=∑Pⁿi
S=-P
Определение R*,миними-
зирующего F(X+R*S)
X1=X+R*S
P1=G(X1)
SN=∑(x1i-Xi)2
Нет
SN<E
Да
PN=∑P12i
SP=∑P1i(P1i-Pi)
K=K+1
IND=K-(K/N)*N
BET=SP/PN
X*=X1
Y*=F(X*)
Вывод
X*,Y*
конец
Да
IND=0
Нет
PN=PN1
BET=0
S=P1+BET·S
X=X1
P=P1

Рис. 5.10

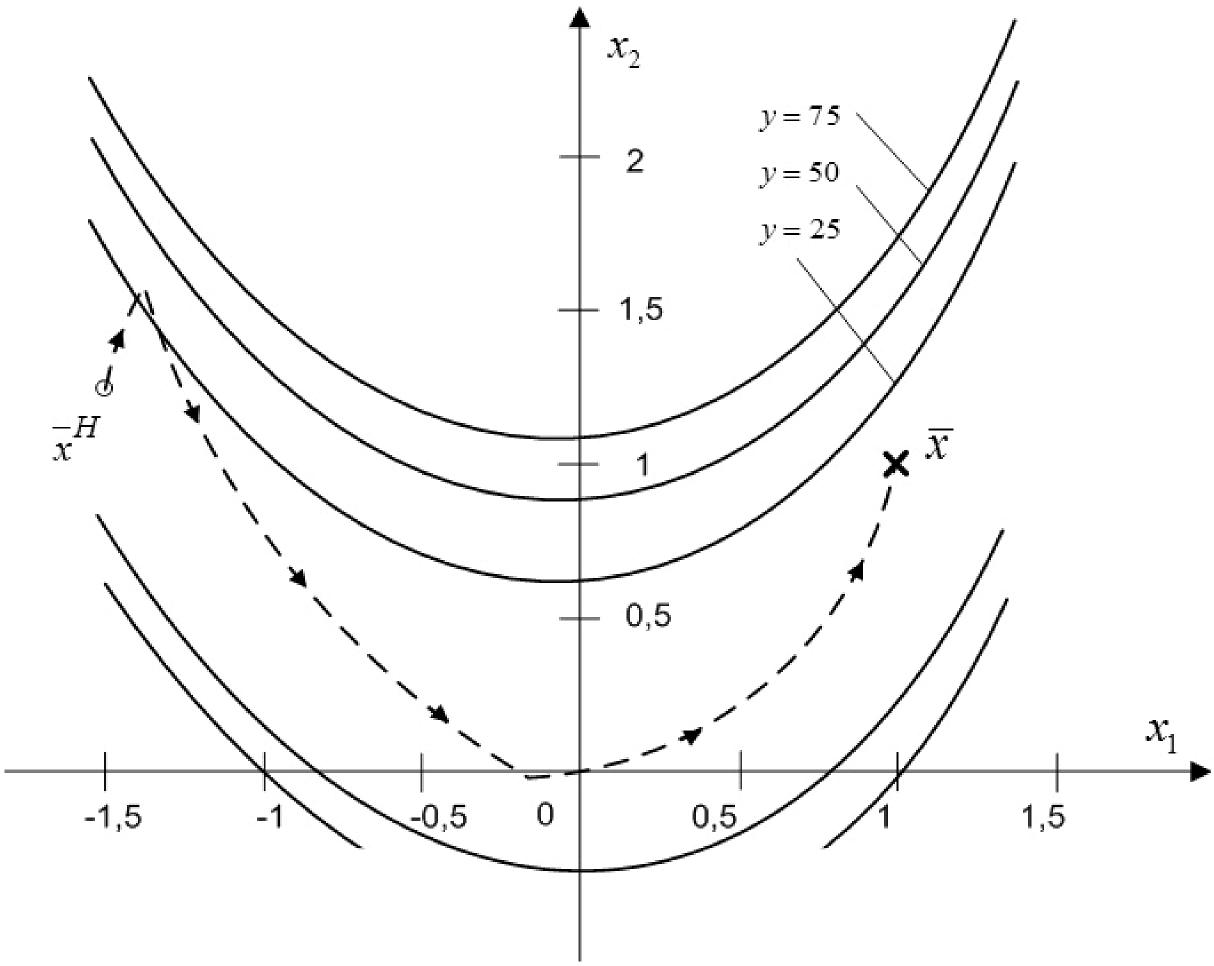


Рис. 5.11

На рис. 5.11 пунктирной линией показана последовательность точек, получающихся в процессе вычислений по методу сопряженных градиентов для случая, когда в качестве начальной точки была принята точка $\left(x_1^H = -1{,}5,\ x_2^H = 1{,}5\right)$. Поиск завершился при $x_1 = 0{,}991;\ x_2 = 0{,}992.$

§ 5.4. Численные методы многомерной оптимизации, без использования производных

Методы этой группы применяются в тех случаях, когда в процессе поиска получать информацию о производных от ЦФ по оптимизируемым параметрам трудно или невозможно. В частности, их используют в задачах оптимизационного планирования, связанных с экспериментальным определением оптимального сочетания варьируемых факторов, при которых качество ее функционирования, оцениваемое заданным показателем, будет наилучшим.

Простейшим методом такого рода является метод покоординатной оптимизации (метод поочередного изменения переменных), сущность которого состоит в следующем.

1. Выбираются начальные значения оптимизируемых переменных $x_1 = x_{10},..., x_n = x_{n0}$ заданной ЦФ $y = f(x_1,..., x_n)$.

2. Решается задача одномерной оптимизации по переменной x_1 при неизменных остальных переменных: $x_2 = x_{20},..., x_n = x_{n0}$. В результате получается некоторое значение $x_1{}^*$.

3. При $x_1 = x_1^*$, $x_3 = x_{30},..., x_n = x_{n0}$ осуществляется оптимизация ЦФ по x_2.

4. Используя $x_1 = x_1^*, x_2 = x^*{}_2,..., x_n = x_{n0}$, находим x_3^*, оптимизирующее ЦФ и т.д., пока все переменные поочередно не будут рассмотрены.

5. Затем весь процесс повторяется с п. 1 при использовании в качестве x_{i0} значения x_i^*.

Вычисления заканчиваются при попадании в точку $\bar{x}^{\text{опт}}$, для которой смещения по каждой из координат x_i не дают улучшения ЦФ.

Поиск рассмотренным методом для $n = 2$ проиллюстрирован на рис. 5.12.

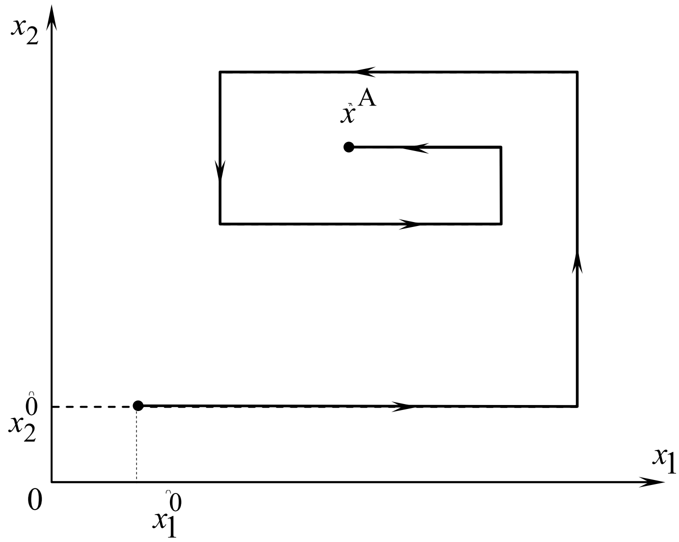


Рис. 5.12

Ясно, что количество вычислений, необходимых для поиска оптимума этим методом, быстро увеличивается с ростом размерности задачи.

К числу более эффективных методов из рассматриваемой группы относится так называемый симплекс-метод [37]. Поясним его идею, рассматривая для наглядности минимизацию ЦФ двух переменных.

1. Выберем на плоскости оптимизируемых переменных x_1, x_2 три точки (A_0, B_0, C_0), являющиеся вершинами некоторого равностороннего треугольника (рис. 5.13).

2. Находим значения ЦФ в каждой из этих точек, сравниваем их между собой и выбираем ту, в которой ЦФ оказалась наибольшей, например, т. A_0.

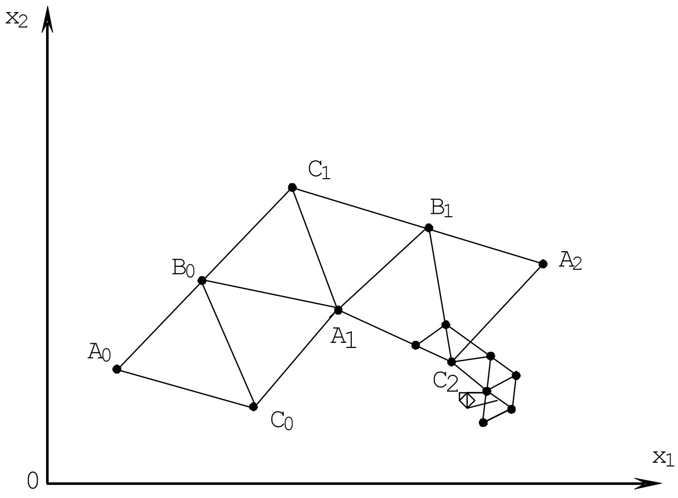


Рис. 5.13

3. Осуществляя зеркальное отображение выделенной точки A_0 относительно противолежащей стороны (B_0, C_0), получаем новую точку A_1. Вычисляем ЦФ в этой точке A_1, а затем сравниваем полученное значение $f(A_1)$ с имеющимися значениями $f(B_0)$ и $f(C_0)$, выбираем «наихудшую» точку и осуществляем ее зеркальное отображение.

Этот процесс продолжается до тех пор, пока на каком-то шаге не произойдет «зацикливание» – в результате отображения возвращаемся к уже пройденной точке $(A_2 \leftrightarrow A_1)$. В таком случае следует уменьшить размер треугольника, например в 2 раза, и продолжить поиск по тем же правилам, выполняя отображения и сжатия треугольников. Процесс вычислений прекращается, когда размеры полученного в ходе поиска треугольника станут меньше заданной малой величины ε.

Аналогично можно решать оптимизационную задачу с ЦФ трех переменных, рассматривая при этом в пространстве $(x_1 x_2 x_3)$ четыре точки, образующие тетраэдр, а также ЦФ n переменных, используя $n+1$ точку в n–мерном пространстве. Используемые геометрические фигуры: треугольник, тетраэдр и т.д. называют симплексами.

Важным достоинством метода является то, что при его использовании на каждом шаге (кроме 1-го) вычисляется значение ЦФ только в одной точке, в то же время смещение происходит в направлении уменьшения ЦФ сразу по всем переменным.

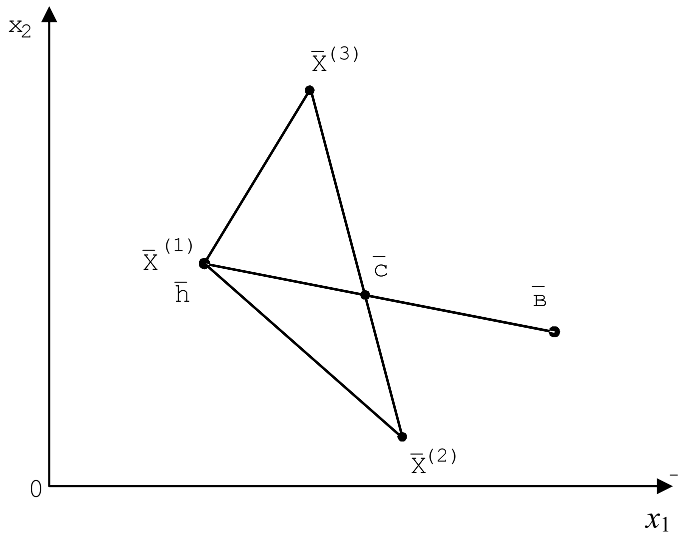


Рис. 5.14

На основе изложенной идеи разработаны различные варианты метода. Одним из них является метод деформируемого многогранника (метод Нелдера – Мида) [2, 37]. Он состоит в следующем.

1. В n-мерном пространстве переменных $x_1,...,x_n$ выбираются точки $\overline{x}^{(1)},...,\overline{x}^{(n)},\overline{x}^{(n+1)}$, образующие многогранник $s^{(0)}$ (рис. 5.14).

2. Вычисляются значения ЦФ f в выбранных точках $\overline{x}^{(k)}$, $k=\overline{1,n+1}$.

3. Выделяется точка $\overline{x}^{(j)}=\overline{h}$, в которой ЦФ принимает наихудшее (наибольшее при минимизации) значение, и точка $\overline{x}^{(i)}=\overline{l}$ – с наилучшим значением ЦФ.

4. Определяются координаты $c_j=\dfrac{1}{n}\left[\left(\sum\limits_{i=1}^{n+1}x_j^{(i)}\right)-h_j\right]$ точки $\overline{c}$, называемой центром тяжести.

5. Находится новая точка $\overline{x}=\overline{b}$ проектированием точки $\overline{h}$ через центр тяжести $\overline{c}$. Эта операция называется отражением и определяется соотношением $\overline{b}=\overline{c}+\alpha(\overline{c}-\overline{h})$, где $\alpha>0$ – число, называемое коэффициентом отражения (обычно принимают $\alpha=1$).

6. Вычисляются значения ЦФ в точках $\overline{b}$ и $\overline{l}$, а затем $f(\overline{b})$ и $f(\overline{l})$ сравниваются; если $f(\overline{b})<f(\overline{l})$, то выделенное направление (от $\overline{c}$ к $\overline{b}$) признается перспективным и вектор ($\overline{b}$ – $\overline{c}$) растягивается в соответствии с соотношением

$$\overline{q}=\overline{c}+\gamma(\overline{b}-\overline{c}),$$

где $\gamma > 1$ – число, называемое коэффициентом растяжения (обычно принимают $\gamma = 2$).

7. Вычисляется значение $f(\overline{q})$.

8. Формируется новый многогранник $s^{(1)}$, отличающийся от $s^{(0)}$ тем, что вершина $\overline{h}$ исключается, а к остальным вершинам добавляется одна новая. Если $f(\overline{g}) < f(\overline{l})$, то добавляемой вершиной является $\overline{x} = \overline{h}$, иначе $\overline{x} = \overline{b}$.

9. Для полученного многогранника повторяются все перечисленные выше операции.

10. Если при сравнении ЦФ в точках $\overline{b}$ и $\overline{l}$ (см. п. 6) оказывается, что условие $f(\overline{b}) \le f(\overline{l})$ не выполняется, то $f(\overline{b})$ сравнивается со значением ЦФ в остальных вершинах *s*. Если $f(\overline{b}) > f(\overline{x}^{(i)})$ для всех вершин кроме $\overline{h}$, то вектор ($\overline{h} - \overline{c}$) сжимается в соответствии с формулой $\overline{r} = \overline{c} + \beta(\overline{h} - \overline{c})$, где $0 < \beta < 1$ – коэффициент сжатия, принимаемый обычно равным 1/2. Затем в $s^{(0)}$ вершина $\overline{h}$ заменяется на точку $\overline{r}$, и для полученного в результате многогранника $s^{(1)}$ повторяются все операции с п. 2. Если же $f(\overline{b}) > f(\overline{h})$, то во избежание зацикливания осуществляется так называемая редукция: все векторы $(\overline{x}^{(i)} - \overline{l})$ уменьшаются в два раза, т.е. $\overline{x}^{(i)} \to \overline{l} + \frac{1}{2}\left(\overline{x}^{(i)} - \overline{l}\right), i = \overline{1, n+1}$.

Для полученного многогранника повторяются все операции с п. 2.

Вычислительный процесс продолжается до тех пор, пока в результате поиска размеры многогранника не станут достаточно малы либо окажется выполненным условие: $\sqrt{\frac{1}{n+1}\sum_{i=1}^{n}\left[f\left(\overline{x}^{(i)}\right) - f\left(\overline{l}\right)\right]^2} < \varepsilon$.

Достоинством этого алгоритма является то, что деформируемый многогранник в отличие от регулярного симплекса гораздо лучше адаптируется к топографии ЦФ, вытягиваясь вдоль длинных наклонных поверхностей, изменяя направление в изогнутых впадинах и сжимаясь в окрестностях экстремума. Блок-схему программы, реализующей описанный выше алгоритм, см. в [37].

В [37] приведен листинг программы *SIMPLEX*, реализующей рассмотренный алгоритм. Чтобы воспользоваться этой программой для решения конкретной задачи безусловной оптимизации, необходимо в специальной подпрограмме *SUMR* предусмотреть вычисление соответствующей ЦФ, обозначенной *SUM(IN)*, от переменных *X*(1), *X*(2),... . Кроме того, нужно ввести число *NS* оптимизируемых переменных в ЦФ $1 \le NX \le 50$, начальные (предполагаемые) значения этих переменных, а также параметр *STEP*, определяющий размеры начального многогранника. Его можно выбрать, например, по правилу [37]:

$$STEP = \min\left\{ d_1, d_2, ..., d_n, \frac{0,2\sum_{i=1}^{n} d_i}{n} \right\},$$

где d_i – длина интервала $(x_i^a,\ x_i^b)$, внутри которого можно ожидать оптимальное значение x_i^*.

§ 5.5. Численные методы оптимизации при наличии ограничений

Как отмечалось в § 5.1, задачу оптимизации, в которой наряду с ЦФ $y = f(\bar{x})$ задана система ограничений-равенств $h_j(\bar{x}) = 0$ $(j = \overline{1, m})$, можно свести к задаче безусловной оптимизации. В частности, для этого можно использовать метод множителей Лагранжа. В этом случае исходная задача заменяется на задачу об отыскании значений $x_i = x_i^*, \lambda_j = \lambda_j^*$ $(i = \overline{1, n} : j = \overline{1, m})$, доставляющих минимум функции Лагранжа:

$$\tilde{y} = f(\bar{x}) + \bar{\lambda}^T \bar{h}(\bar{x}) = f(\bar{x}) + \lambda_1 h_1(\bar{x}) + ... + \lambda_m h_m(\bar{x}). \qquad (5.18)$$

Для ее решения можно использовать алгоритмы и программы безусловной оптимизации, рассмотренные в § 5.4.

Пример. Найти x_1, x_2, при которых ЦФ $y = x_1^2 + x_2^2$ принимает наименьшее значение при выполнении условия $x_1 + x_2 = 4$. Заметим, что решение этой задачи легко находится после подстановки $x_2 = 4 - x_1$ в выражение для ЦФ. В результате получается решение $x_1^* = x_2^* = 2$.

Для применения программы безусловной оптимизации с использованием метода множителей Лагранжа нужно рассматривать ЦФ

$$\tilde{y} = x_1^2 + x_2^2 + x_3(x_1 + x_2 - 4),$$

где x_3 играет роль λ, а затем находить ее минимум по x_1, x_2, x_3.

Возможен и другой подход для численного решения задачи условной оптимизации. Введем вспомогательную ЦФ

$$\tilde{y} = f(\bar{x}) + r\sum_{j=1}^{m} h_j^2(\bar{x}), \tag{5.19}$$

в которой второе слагаемое формируется по заданным ограничениям.

Ясно, что для тех $\bar{x}$, которые удовлетворяют всем ограничениям, второе слагаемое в (5.19) равно нулю и $\tilde{y} = y$. При достаточно большом значении параметра r для $\bar{x}$, не удовлетворяющих хотя бы одному из ограничений, это слагаемое будет принимать большие значения и, следовательно, минимального значения $\bar{x}$ для них получиться не может. Таким образом, благодаря второму слагаемому в (5.19) значения $\bar{x}$, не входящие в область допустимых решений, определяемую ограничениями, получают своеобразный штраф, поэтому функция (5.19) называется штрафной.

Для рассматриваемого выше примера она принимает вид

$$\tilde{y} = x_1^2 + x_2^2 + r(x_1 + x_2 - 4)^2.$$

Из условия $y'_{x_1} = 0, y'_{x_2} = 0$ нетрудно получить

$$x_2^* = 4/(2 + 1/r); \quad x_1^* = 4 - (1 + 1/r)x_2^*.$$

Как видим, при достаточно большом $r: x_2^* \to 2, \quad x_1^* \to 2$. При реализации этого подхода с использованием какой-либо программы безусловной оптимизации целесообразно осуществлять решение итерационно. При этом вначале параметр r принимается равным не очень большому числу r_1, полученное в результате поиска минимизирующее значение $\bar{x}^{*(1)}$ затем используется в качестве начальной точки для нового поиска с параметром $r = r_2 > r_1$ и т.д. до тех пор, пока число $r = r_k$ не станет достаточно большим. При выборе последовательности значений параметра штрафа можно использовать, например, $r_0 = 1$ и $r_k = k^2$.

Метод штрафных функций с успехом используется и для решения задач статической оптимизации с ограничениями-неравенствами, т.е. задач нелинейного программирования:

$$y = f(\bar{x}) \to inf, \quad g_j(\bar{x}) \ge 0, \quad j = \overline{1,m}.$$

Необходимую для этого штрафную функцию можно сконструировать, например, следующим образом [26, 37] :

$$\tilde{y}(\bar{x}) = f(\bar{x}) - r_k \sum_{j=1}^{m} ln\, g_j(\bar{x}) \tag{5.20}$$

либо

$$\tilde{y}(\bar{x}) = f(\bar{x}) + r_k \sum_{j=1}^{m} [1/g_j(\bar{x})]. \tag{5.21}$$

Их особенностью является то, что входящее в них второе слагаемое резко возрастает при приближении $\bar{x}$ к границе ОДР, соответствующей условиям $g_j(\bar{x}) = 0$, тем самым препятствуя выходу процесса поиска за пределы ОДР.

Поиск начинается с выбранной начальной точки $\bar{x}^{(0)}$, заведомо удовлетворяющей всем ограничениям, т.е. $\bar{g}(\bar{x}^{(0)}) > 0$ при некотором $r = r_0 > 0$. Полученный результат $\bar{x}^*$ используется в качестве начальной точки для нового поиска минимума при $r = r_1 < r_0$ и т.д. Последовательность r_k должна удовлетворять условию $r_k \to 0$ при $k \to \infty$. Можно, например, принять $r_0 = 1$ и $r_{k+1} = r_k / C$, где $C > 1$ – некоторая константа. В [37] рекомендуется использовать $C = 4$.

Следует подчеркнуть, что при минимизации вспомогательной (присоединенной) функции (5.20) или (5.21) численным методом не исключен выход $\bar{x}^{(k)}$ в процессе поиска из области допустимых значений. Поэтому процедуру безусловной минимизации функции $\tilde{y}(\bar{x})$ на каждой итерации необходимо подкреплять процедурой проверки допустимости текущего решения [26]. Отметим также, что в тех случаях, когда при выборе начальной точки $\bar{x}^{(0)}$, удовлетворяющей всем ограничениям $g_j(\bar{x}) > 0$, возникают затруднения, можно использовать поисковый алгоритм, изложенный в [37].

Указанных выше проблем с выбором $\bar{x}^{(0)}$ и проверкой допустимости $\bar{x}^{(k)}$ не возникает в тех случаях, когда для ограничений

$$\varphi_j(\bar{x}) \le 0, \quad j = \overline{1,m};$$

$$h_i(\bar{x}) = 0, \quad i = \overline{1,k}$$

используется вспомогательная функция [26, 37]

$$\tilde{y} = f(\bar{x}) + \rho_k \Phi(\bar{x}),$$

где функция штрафа φ (x) определяется выражением

$$\Phi(\bar{x}) = \sum_{j=1}^{m} [\varphi_j^+(\bar{x})]^2 + \sum_{i=1}^{k} h_i^2(\bar{x}), \tag{5.22}$$

в котором $\varphi_j^+(\bar{x})$ представляет собой так называемую «срезку» функции φ_j:

$$\varphi_j^+(\bar{x}) = \begin{cases} 0, & \text{если} \quad \varphi_j(\bar{x}) \le 0; \\ \varphi_j(\bar{x}), & \text{если} \quad \varphi_j(\bar{x}) > 0, \end{cases}$$

т.е. $\varphi_j^+(\bar{x}) = \max\{0, \varphi(\bar{x})\}$.

Примером программы, реализующей метод штрафных функций в форме (5.22), является программа, приведенная в [2]. Более подробно вопросы, связанные с методом штрафных функций, рассмотрены в [2, 26, 37].

Наряду с рассмотренным методом штрафных функций, сводящим задачу оптимизации при наличии ограничений к последовательности задач безусловной оптимизации вспомогательной функции, широкое применение находят алгоритмы, предусматривающие продвижение к оптимуму исходной целевой функции по последовательности допустимых или «почти» допустимых точек с монотонно убывающими значениями ЦФ с непосредственным контролем соблюдения ограничений [26, 37]. К числу таких алгоритмов относятся метод проекции градиента (метод Розена), метод возможных направлений [1] (метод допустимых направлений Зойтендейка) и другие [37].

Одним из универсальных, эффективных и надежных алгоритмов этой группы является метод скользящего допуска (МСД) [37]. Он позволяет решать задачи нелинейного программирования общего вида при наличии линейных и нелинейных ограничений равенств и неравенств. В отличие от многих других методов нелинейного программирования, при реализации которых на ЦВМ значительная доля машинного времени тратится на то, чтобы обеспечить строгое выполнение заданных ограничений, МСД позволяет улучшать значения ЦФ в процессе поиска как за счет информации, получаемой в допустимых точках пространства варьируемых переменных, так и за счет информации, которую удается получить в некоторых точках, лежащих вне допустимой области, но являющихся близкими к допустимым (почти допустимым). При этом интервалы, в пределах которых точки можно считать почти допустимыми в процессе поиска постепенно сужаются. Поэтому в пределе (по мере приближения к искомому решению задачи) рассматриваются только допустимые точки.

При реализации этой стратегии осуществляется процесс минимизации данной ЦФ $y = f(\bar{x})$ изложенным выше методом Нелдера – Мида и одновременно обеспечивается выполнение условия

$$T(\bar{x}) \le s^{(k)},$$

где $T(\bar{x}) = \left[\sum_{i=1}^{k} h_i(\bar{x}) + \sum_{j=1}^{m} u_j g_j^2(\bar{x}) \right]^{1/2}$; $u_j = \begin{cases} 0 & \text{при} \quad g_j(\bar{x}) \geq 0; \\ 1 & \text{при} \quad g_i(\bar{x}) < 0. \end{cases}$

В этих расчетных соотношениях величина $T(\bar{x})$ оценивает степень удовлетворения всех заданных ограничений, $s^{(k)}$ – представляет собой уменьшающееся в процессе поиска значение критерия, по которому текущие значения вектора $\bar{x}$ признаются допустимыми (или почти допустимыми).

Детальное изложение метода скользящего допуска приведено в [37]. Там же представлена реализующая этот метод программа *FLEX*.

Контрольные вопросы

1. Как формулируется необходимое условие экстремума?
2. Какие точки называются стационарными, какие критическими?
3. Как формулируется достаточное условие экстремума?
4. Как формируется функция Лагранжа, что называется множителями Лагранжа?
5. Как осуществляется отыскание решения задачи условной оптимизации методом множителей Лагранжа?
6. Какие препятствия возникают при практическом использовании необходимого условия экстремума для решения оптимизационных задач?
7. В чем состоит сущность метода равномерного пассивного поиска экстремума?
8. Можно ли использовать методы равномерного пассивного поиска для поиска экстремума функций нескольких переменных; возможно ли его применение для целевых функций, имеющих несколько экстремумов?
9. Какая функция называется унимодальной?
10. Что называется золотым сечением?
11. В чем состоит принципиальная особенность метода золотого сечения при отыскании экстремума?
12. Запишите расчетное соотношение для отыскания экстремума методом Ньютона-Рафсона в краткой векторно-матричной форме.

13. Запишите расчетное соотношение для отыскания экстремума градиентным методом в векторно-матричной форме и развернутом виде (для целевой функции двух переменных).
14. Какое условие для прекращения вычислений (критерий останова) используется при отыскании экстремума градиентным методом?
15. Какая целевая функция называется функцией «овражного типа», каким образом «овражный характер» этой функции отражается на поиске экстремума градиентным методом?
16. В чем состоит сущность поиска экстремума нелинейной целевой функции симплекс-методом (методом деформируемого многогранника)? Что такое «отражение» и «сжатие» в этом алгоритме, какой критерий останова используется?
17. В чем состоит сущность метода штрафных функций при решении задачи условной оптимизации?
18. Как формируется штрафная функция при поиске экстремума с ограничениями-неравенствами?

Задачи для самостоятельного решения

Используя аналитические методы, найти наименьшее или наибольшее значение заданных функций.

1. $y = x_1^2 + x_1 x_2 + x_2^2 - 3x_1 - 6x_2$.
2. $y = \dfrac{1}{2} x_1 x_2 + (47 - x_1 x_2)\left(\dfrac{x_1}{3} + \dfrac{x_2}{4}\right)$.
3. $y = x_1 x_2^2 (1 - x_1 - x_2)$.
4. $y = x_1^3 + 3x_1 x_2^2 - 15x_1 - 12x_2$.
5. $y = (x_1 - 5)^2 - 7x_2^2$.
6. $y = \dfrac{1 + x_1 - x_2}{\sqrt{1 + x_1^2 + x_2^2}}$.
7. $y = e^{(x_1 - x_2)} (x_1^2 - 2x_2^2)$.
8. $y = \sqrt{(9 - x_1)(9 - x_2)(x_1 + x_2 - 9)}$.
9. $y = 6 - 4x_1 - 3x_2$ при условии $x_1^2 + x_2^2 = 1$.
10. $y = x_1^2 + x_2^2$ при $\dfrac{x_1}{2} + \dfrac{x_2}{3} = 1$.
11. $y = x_1 x_2$ при условии $2x_1 + 3x_2 = 5$.

12. $y = x_1 x_2 + x_2 x_3$ при условиях $x_1 + x_2 = 4$, $x_2 + x_3 = 4$.

13. $y = 3x_1^2 + 2x_1 + 2x_2^2 + 4x_2 x_3$ при условиях $x_1^2 + 2x_2^2 = 19$, $x_1 + 2x_2 x_3 = 11$.

14. $y = x_1 x_2 x_2$ при условиях $2x_1 x_2 + x_2 x_3 = 12$, $2x_1 - x_2 = 8$.

15. $y = x_1 - 2x_2 + 2x_3$ при условии $x_1^2 + x_2^2 + x_3^3 = 9$.

16. $y = x_1 x_2$ при условии $x_1^2 + x_2^2 \le 1$.

17. $y = 1 + x_1 + 2x_2$ при условии $x_1 \ge 0,\ x_2 \ge 0,\ x_1 + x_2 \le 1$.

18. $y = x_1^2 x_2^2 - x_1 x_2 + 4x_1$ при $x_1 \ge 0,\ x_2 \ge 0,\ 2x_1 + 3x_2 \le 12$.

19. $y = x_1 x_2 + x_1 + x_2$ при $1 \le x_1 \le 2,\ 2 \le x_2 \le 3$.

20. $y = x_1^3 + x_2^3 - 3x_1 x_2$ при $0 \le x_1 \le 2,\ -1 \le x_2 \le 2$.

Составить программу и провести вычисления, для отыскания наименьшего значения заданной функции одной переменной с помощью одного из следующих методов:

- равномерного пассивного поиска;
- метода дихотомии;
- метода золотого сечения;
- метода Ньютона;
- квадратичной аппроксимации.

21. $y = 1 - xe^{-x}$.

22. $y = \dfrac{1}{2}x^2 + e^{-x}$.

23. $y = (x-1)e^{-x^2}$.

24. $y = \dfrac{x-2}{1+(x-2)^2}$.

25. $y = 4 - 10x + 2x^2 - 0{,}1x^3$.

26. $y = 1 - e^{-x}\ln x$.

27. $y = 2 - 3x + x^3$.

28. $y = -1 + e^{-2x}\cos x$.

29. $y = \dfrac{x^3}{(x-2)^2}$.

30. $y = e^{-2x} - e^{-3x}$

Указанное выше задание выполнить для следующих функций двух переменных, используя методы [37]:

- Ньютона – Рафсона;
- градиентного спуска с дроблением шага;
- Флетчера – Ривса;
- покоординатной оптимизации (Гаусса – Зайделя);
- Нелдера – Мида;
- Хука – Дживса;
- Давидона – Флетчера – Пауэлла.

Используя результаты вычислений на различных итерациях, построить на плоскости варьируемых переменных (с нанесенным на нее семейством линий равного уровня) траектории изменения положения точки $\bar{x}$ в процессе поиска для различных начальных точек.

31. $y=(x_1-4)^2+\frac{1}{4}(x_2-7)^2$.

32. $y=4x_1^2+3x_2^2-4x_1x_2+x_1$.

33. $y=8x_1^2+4x_1x_2+5x_2^2$.

34. $y=(x_2-x_1^2)^2+(1-x_1)^2$.

35. $y=100(x_2-x_1^2)^2+(1-x_1)^2$.

36. $y=2x_1^3+4x_1x_2^3-10x_1x_2+x_2^2$.

37. $y=(x_1^2+x_2-11)^2+(x_1+x_2^2-7)^2$.

38. $y=\frac{1}{10}\left[12+x_1^2+\frac{1+x_2^2}{x_1^2}+\frac{x_1^2x_2^2+100}{(x_1x_2)^4}\right]$.

39. $y=2x_1^2+4x_1x_2^3-10x_1x_2+x_2^2$.

40. $y=u_1^2+u_2^2+u_3^2$, где $u_1=1{,}5-x_1(1-x_2)$; $u_2=2{,}25-x_1(1-x_2^2)$; $u_3=2{,}625-x_1(1-x_2)^3$.

Глава 6. **ЗАДАЧИ ОПТИМАЛЬНОГО ДИНАМИЧЕСКОГО УПРАВЛЕНИЯ И ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ**

§ 6.1. Понятие об управляемых динамических системах

Рассмотрим некоторую ***систему*** любой природы. Это может быть техническая система, например самолет, экономическая система, например производственное предприятие и т.п. Будем считать, что ее поведение можно характеризовать некоторым набором величин и обозначим их $s_1, \dots, s_m$. Они описывают ***состояние*** системы, поэтому будем называть их переменными состояния.

Система является ***управляемой***, если в ней имеется некоторый набор факторов (величин, параметров), выбирая которые соответствующим образом можно целенаправленно влиять на переменные состояния. Обозначим их $x_1, \dots, x_k$. Это своеобразные рули, распоряжаясь которыми можно добиться нужного поведения системы. Назовем их управляющими воздействиями, или управляющими переменными, или короче – управлениями.

Рассмотренную систему можно представить графически, если управляющие воздействия интерпретировать как входы, а переменные состояния – как выходы системы (рис. 6.1).

Рис. 6.1

Удобно использовать векторные обозначения, полагая

$$\bar{x} = \begin{bmatrix} x_1 \\ \dots \\ x_k \end{bmatrix}; \qquad \bar{s} = \begin{bmatrix} s_1 \\ \dots \\ s_m \end{bmatrix}.$$

Система называется ***статической***, если x_i, s_i – постоянные величины (числа). Если же они изменяются во времени и описываются функциями времени $x_i = x_i(t)$, $s_i = s_i(t)$, то система называется ***динамической***.

Обычно поведение системы рассматривается на некотором промежутке времени $[t_н, t_к]$. Если при этом время непрерывно изменяется от начального момента до конечного, пробегая все возможные значения, то процессы $\bar{x}(t)$, $\bar{s}(t)$ и система называются ***непрерывными***. Существует немало систем, в том числе экономических, в которых процессы рассматриваются, контролируются или изменяются в отдельные изолированные моменты времени $\{t_0 = t_{\text{н}},\ t_1,\ t_2, ..., t_{\text{к}}\}$. Такие процессы и системы называют ***дискретными***. Обычно эти моменты следуют друг за другом через одинаковый промежуток времени T, называемый периодом дискретизации, или шагом по времени. Тогда $t_i = t_{\text{н}} + iT$, при этом i называют номером такта, шага или этапа.

Будем рассматривать ***детерминированные*** системы, для которых, задав некоторое конкретное управление $\bar{x}(t)$ и начальное состояние $\bar{s}(t_{\text{н}}) = \bar{s}^{(0)}$, мы получим однозначно определенный закон изменения вектора переменных состояния $\bar{s} = \bar{s}(t)$.

При этом следует иметь в виду, что не все реальные системы, с которыми приходится иметь дело на практике, в том числе в экономике, можно считать детерминированными. Часто на поведение системы помимо целенаправленно формируемых управлений $\bar{x}(t)$ действуют неподвластные нам скрытые и другие факторы, которые можно назвать возмущениями. К ним относятся, например, факторы спроса на тот или иной вид продукции, действия конкурентов, природные факторы и т.д. Подобные системы называют системами ***с неопределенностями,*** или ***стохастическими*** системами.

Задачу управления можно представить как поиск и формирование вектора управляющих воздействий $\bar{x}$, при которых система ведет себя в соответствии с поставленной целью. Следует подчеркнуть, что не любое, а только целенаправленное воздействие на систему называется управляющим. Для различных конкретных систем характер управляющих воздействий, переменных состояния и поставленной цели управления может быть разным. Ниже будут приведены соответствующие примеры для экономических систем. Характерно, что для таких систем управление часто состоит в планировании объема выпускаемых изделий, закупаемых товаров, вложенных средств и т.д.

§ 6.2. Формулировка классической задачи об оптимальном динамическом управлении

В общей теории динамического управления обычно изучается следующая базовая задача: найти закон изменения управляющих переменных $x_i = x_i(t)$, при которых рассматриваемая система переходит из заданного начального состояния в заданное конечное состояние:

$$\bar{s}(t_{\text{н}}) = \bar{s}^{(\text{н})} \xrightarrow[t_{\text{н}} \le t \le t_{\text{к}}]{} \bar{s}(t_{\text{к}}) = \bar{s}^{(\text{к})}. \tag{6.1}$$

Оказывается, эта задача, как правило, имеет не единственное решение – можно сформировать множество различных законов управления, обеспечивающих выполнение условия (6.1). В подобной ситуации вводят некий показатель качества управления

$$Y = F[\bar{x}(t), \bar{s}(t)], \tag{6.2}$$

где Y – число, оценивающее эффективность или качество процесса управления, F – правило, по которому оно определяется.

В связи с этим дополнительно к задаче управления (6.1), которую можно считать основной, ставится новая задача, которую можно рассматривать как «сверхзадачу»: из множества возможных управлений $\bar{x} = \bar{x}(t)$, обеспечивающих выполнение условия (6.1), найти такое управление, при котором получается наибольшее (или наименьшее) значение показателя качества (6.2):

$$\bar{x} = \bar{x}^*(t), \;\; Y = F[\bar{x}(t), \bar{s}(t)] \to \sup(\inf). \tag{6.3}$$

Следует отметить, что в структуре рассматриваемой задачи могут присутствовать дополнительные условия, чаще всего в виде равенств или неравенств, например

$$a_j \le x_j \le b_j.$$

Обычно они задают ограничения на ресурсы. В общем случае эти ограничения можно представить следующим образом:

$$\left.\begin{array}{l} \bar{x}(t) \in \Omega_x, \\ \bar{s}(t) \in \Omega_s, \end{array}\right\} \tag{6.4}$$

где Ω_x, Ω_s – область допустимых значений в пространстве управлений и пространстве состояний.

Управление заданной системой $\bar{x}(t) = \bar{x}^*(t)$, обеспечивающее выполнение условий (6.1) – (6.4) называется ***оптимальным***, а показатель качества (6.2) называется ***критерием оптимальности***. Заметим, что соотношение (6.2) функциям $\bar{x}(t), \bar{s}(t)$ ставит в соответствие число Y и представляет собой функционал. Это соотношение является аналогом целевой функции $y = f(\bar{x})$ в задачах статической оптимизации.

Изложенная ситуация имеет следующую наглядную интерпретацию. Пусть на плоскости (рис. 6.2) заданы две точки А и В. Требуется сформировать траекторию (линию), двигаясь по которой можно из точки А попасть в точку В.

Ясно, что таких линий существует бесконечное множество. Дополнительно можно потребовать, чтобы сформированная траектория позволяла бы попасть из А в В по кратчайшему пути или за минимальное время или с минимальной затратой средств (например топлива). Длина пути, время или затраченные средства – это примеры показателей качества Y, а формулы, позволяющие рассчитывать соответствующий показатель Y по траектории $\bar{x}(t)$, будут определять критерий качества (оптимальности).

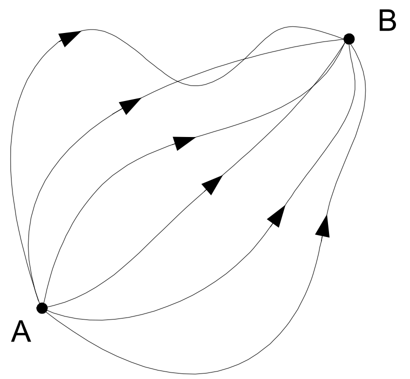


Рис. 6.2

Для решения задачи об отыскании оптимального управления необходимо задать или найти математическую модель рассматриваемой системы, т.е. некоторый математический преобразователь G, позволяющий для выбранного управления $\bar{x}(t)$ и заданного начального условия $\bar{s}(0)$ находить переменные состояния:

$$\text{для } t \geq 0 \qquad \bar{s}(t) = G(\bar{s}(0), \bar{x}(t)). \tag{6.5}$$

Например, для непрерывной управляемой системы такой моделью обычно являются дифференциальные уравнения

$$\frac{d\bar{s}}{dt} = \bar{g}(\bar{s}, \bar{x}, t). \tag{6.5.1}$$

Таким образом, задача оптимального динамического управления (задача динамической оптимизации) состоит в том, чтобы для заданной системы, описываемой математической моделью (6.5), найти закон изменения управляющих факторов $\bar{x}(t) = \bar{x}^*(t)$, при котором система переходит из заданного начального состояния в заданное конечное состояние (6.1), а показатель качества (6.2) принимает наибольшее (наименьшее) значение, при этом соблюдаются ограничения (6.4).

Следует отметить, что задача об отыскании функций $\bar{x}(t)$, при которых функционал (6.2) принимает экстремальное значение, давно известна в математике. Такие задачи, правда без какой-либо связи с прикладными задачами управления, рассматривались Эйлером, Лагранжем и др. В их трудах были заложены основы специального раздела математики, который называется вариационным исчислением [4, 17, 30, 34].

Специфика задач управления оказалась довольно существенной и потребовала использования специального математического аппарата, который был разработан во второй половине XX в. в СССР (принцип максимума Л.С. Понтрягина [4]) и в США (динамическое программирование Р. Беллмана [6, 30, 32]). Впоследствии оказалось, что они во многом близки. Однако детальное их изучение и опыт их использования показали, что принцип максимума (ПМ) лучше приспособлен для решения задач непрерывного динамического управления, а динамическим программированием (ДП) целесообразнее пользоваться для решения задач управления дискретными системами.

Характерно, что в экономических системах переменные, показатели, разного рода факторы контролируются и изменяются, как правило, по дням, месяцам, годам, т.е. являются дискретными процессами. Поэтому в математической экономике в большей мере востребовано динамическое программирование.

§ 6.3. Формулировка классической задачи динамического программирования (ДП)

Рассмотрим задачу (6.1) – (6.5) об оптимальном управлении некоторой, например экономической, системой, предполагая, что она является детерминированной, динамической и дискретной, т.е. процессы в ней изменяются во времени и эти изменения происходят по шагам.

Кроме того, сделаем ряд принципиально важных, характерных для ДП предположений, касающихся свойств управляемой системы.

1. Система описывается m переменными состояния ($s_1,...,s_m$), и ее функционирование осуществляется по этапам. Будем считать, что управление состоит из n шагов, причем на i-м шаге система под действием управления $\bar{u}(i)$ переходит из состояния $\bar{s}(i-1)$ в следующее состояние $\bar{s}(i)$ (рис. 6.3).

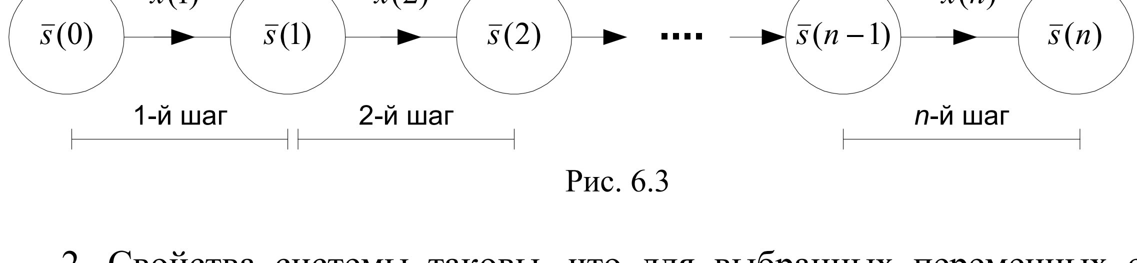


Рис. 6.3

2. Свойства системы таковы, что для выбранных переменных состояния ее математическая модель (6.5) может быть представлена в виде

$$\bar{s}(i) = G_i\left[\bar{s}(i-1), \bar{x}(i)\right],\ i = \overline{1,n}, \qquad (6.6)$$
$$\bar{s}(0) = \bar{s}^{\text{н}},$$

где G_i – некоторый оператор.

Это соотношение отражает принципиально важное для ДП свойство системы – ее новое состояние $\bar{s}(i)$ зависит от текущего управления, реализуемого на текущем этапе, и предшествующего состояния $\bar{s}(i-1)$ и не зависит от того, каким образом система пришла в это состояние. Это свойство называется отсутствием последействия [13].

3. Показатель качества управления Y, например доход предприятия, полученный на всем промежутке времени за все n этапов, складывается из оценок качества управления y_i на отдельных этапах:

$$\left.\begin{aligned} Y &= \sum_{i=1}^{n} y_i, \\ y_i &= f_i(\bar{s}(i-1), \bar{x}(i)), \end{aligned}\right\} \qquad (6.7)$$

где f_i – функция, определяющая показатель качества, например доход, на i-м этапе по состоянию системы $\bar{s}(i-1)$ в начале этапа и реализованному на этом этапе управлению $\bar{x}(i)$. Это свойство называется аддитивностью критерия качества.

Задача ДП состоит в том, чтобы для каждого этапа (шага) найти значения управляющих факторов $\bar{x}(1),\ldots,\bar{x}(n)$, при которых система из заданного начального состояния переходит в заданное конечное состояние, а критерий (6.7) принимает наибольшее значение, если он определяет выигрыш (доход, прибыль и т.д.), или принимает наименьшее значение, если он выражает потери. При этом в процессе решения нужно учитывать модель системы в виде уравнений состояния (6.6), а также ограничения на управления и состояния, если они заданы.

Эта задача – классическая (стандартная) задача динамического программирования, хотя лучше было бы назвать ее задачей многоэтапного планирования, поскольку, как уже отмечалось ранее, термин «программирование» известен в ином более употребительном значении.

§ 6.4. Принцип оптимальности Р. Беллмана

Существуют различные формулировки принципа оптимальности, лежащего в основе ДП. Все они эквивалентны и не принципиально отличаются друг от друга. Они исходят из простой, почти очевидной идеи.

Планируя многошаговую операцию, надо выбирать управляющее воздействие на каждом шаге с учетом всех его будущих последствий на еще предстоящих шагах [6].

Эту же идею можно сформулировать несколько иначе – каково бы ни было состояние $\bar{s}$ системы в результате какого-то числа шагов, необходимо выбирать управление на ближайшем шаге так, чтобы оно в совокупности с оптимальным управлением на всех последующих шагах приводило к оптимальному эффекту на всех оставшихся шагах, включая данный [6].

Этот принцип может показаться довольно простым и даже очевидным. Хотя это вовсе не так. Проиллюстрируем эту ситуацию на примере [13]. Дана транспортная сеть (рис. 6.4) с начальным пунктом А, конечным пунктом В и промежуточными пунктами A_1, A_2, …, которые соединены транспортными магистралями, длины которых заданы и отмечены. Требуется выбрать оптимальный (наикратчайший) путь из А в В. Эту задачу можно представить как задачу поэтапного выбора 1-го, 2-го, … участков искомого маршрута. Эта задача рассматривалась в § 4.2, там же был изложен метод отыскания кратчайшего пути.

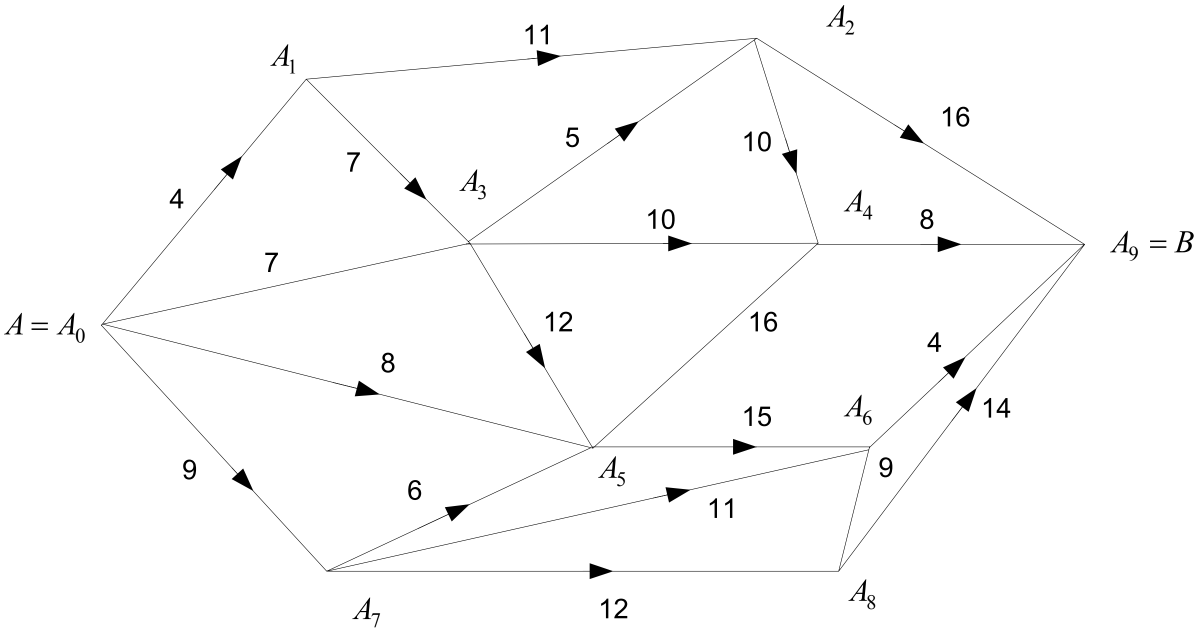


Рис. 6.4

Тот, кто не знаком с соответствующим разделом теории оптимизации, возможно попытается для решения этой задачи использовать следующий принцип «рационального» поведения: для того чтобы получить оптимальное значение критерия качества многошаговой операции, нужно на каждом шаге выбирать наилучший вариант из имеющихся. Если исходить из этого принципа в решении указанной задачи и на каждом этапе выбирать кратчайший путь из рассматриваемой промежуточной точки в следующую соседнюю (смежную) точку, не заглядывая вперед, то получится маршрут $A_0, A_1, A_3, A_2, A_4, B$, имеющий длину 34 ед. Этот маршрут не является оптимальным. Нетрудно найти путь с меньшей длиной, например A_0, A_3, A_4, B с длиной 25, хотя и он не является кратчайшим.

Таким образом, с точки зрения оптимальности всего процесса управления в многошаговой операции нельзя каждый шаг оптимизировать, не заглядывая вперед. Принцип «жить сегодняшним днем», безусловно, ошибочен как в житейском, так и в научном плане. Потратив все имеющиеся деньги на максимально роскошную жизнь сегодня, можно остаться без средств к существованию завтра, послезавтра и т.д. и в итоге оказаться в бедственном положении.

Может показаться, что изложенный выше принцип оптимальности носит слишком общий, быть может, философский характер, и исходя из него вряд ли удастся сформулировать вычислительный алгоритм. Однако это не так. Как отмечается в [12], заслуга Р. Беллмана состоит не только в том, что он в своих основополагающих работах сформулировал этот принцип. Главное состоит в том, что Р. Беллман разработал концепцию практического использования принципа оптимальности для организации вычислительного процесса и «с удивительной изобретательностью стал применять его буквально к сотням оптимизационных задач, возникающих в математике, технике, экономике, исследовании операций и других областях знаний» [12].

На базе этих работ сформировалась техника ДП, которая имеет определенную достаточно универсальную общую концепцию и предполагает наполнение ее для каждого класса решаемых задач своим содержанием, определяемым спецификой этой задачи.

§ 6.5. Сущность метода ДП

Суть общей концепции ДП состоит прежде всего в том, что решение осуществляется в две стадии.

На первой стадии управляемый процесс рассматривается «вспять», начиная с последнего шага и поэтапно переходя к начальному. Характерно, что указанный последний шаг можно анализировать на эффективность без оглядки на будущее, т.к. его в этом случае нет. Важно, что для анализа управления на каждом шаге, включая последний, нужно знать состояние, с которым завершился предыдущий шаг и который будет основой для нового шага. Поскольку это состояние пока неизвестно, то приходится рассматривать все возможные состояния и для каждого из них находить оптимальное управление. Аналогично поступают на предпоследнем и других шагах, вплоть до начального шага.

Эту стадию называют условной оптимизацией.

Вторая стадия, называемая безусловной оптимизацией, является более простой и представляет собой движение от заданного начального состояния к конечному состоянию с использованием результатов условной оптимизации, выполненной на первой стадии.

Это описание концепции ДП покажется довольно туманным и непонятным тому, кто впервые знакомится с ДП. Опыт обучения студентов показывает, что этот туман рассеивается и изложенная концепция воспринимается достаточно ясно после рассмотрения следующего примера.

Задача об отыскании оптимального пути, соединяющего два заданных узла на прямоугольной сетке.

Рассмотрим на плоскости прямоугольник ABCD. Стороны этого прямоугольника разбиты на M частей по горизонтали и на N частей по вертикали, в результате получается сетка (рис. 6.5).

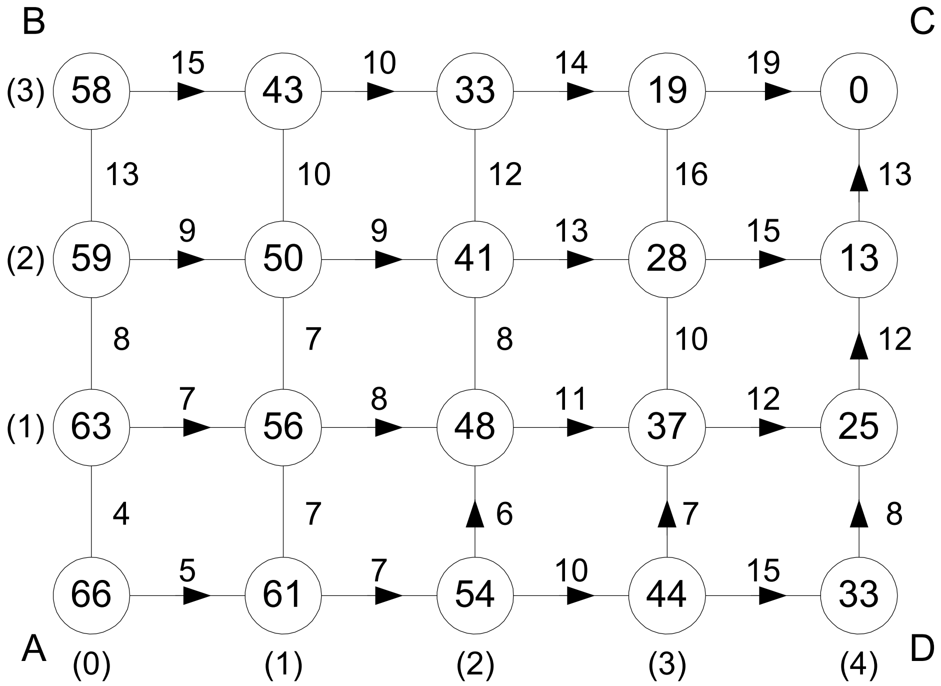


Рис. 6.5

Из точки A (начала) нужно проложить маршрут в точку C (конец), двигаясь по шагам только по сетке и перемещаясь на каждом шаге либо вправо (на восток), либо вверх (на север). Ясно, что этот маршрут будет состоять из $M + N$ шагов и таких маршрутов можно сформировать большое количество.

Пусть каждому элементарному участку сетки приписано неотрицательное число, которое можно рассматривать как затраты, связанные с преодолением этого участка. На рис. 6.5 эти числа проставлены рядом с соответствующими ячейками сетки.

Необходимо выбрать маршрут, для которого продвижение из начальной точки А в конечную точку С было бы связано с наименьшими суммарными затратами.

Решение. Будем рассматривать процесс формирования маршрута как процесс управления системой (точкой), перемещающейся по координатной сетке под влиянием управления (принимаемого решения). Состояние системы меняется дискретно, переходя с одного узла сетки в другой соседний узел. Это состояние можно характеризовать координатами (x, y), где $x \in \{0,1,2,3,4\}$, $y \in \{0,1,2,3\}$. Управление (команда) на каждом шаге может принимать одно из двух значений: «вправо» или «вверх». Обозначим их условными символами: П, В. Процесс будет состоять из 7 шагов и может быть задан набором указанных символов, например (П, В, В, П, П, В, П).

В каждый узел координатной сетки поместим кружок и в нем запишем вспомогательное число – оптимальную величину затрат, с которыми можно из этого узла переместиться в конечную точку С.

Стадия условной оптимизации. Ее начинаем с конечной точки С. В нее в результате последнего 7-го шага можно попасть двумя способами: из точки (4; 2) с управлением (В) и затратами в 13 ед. и из точки (3; 3) с управлением (П) и затратами в 19 ед.

Запишем соответствующие величины затрат в кружки и отметим стрелками направление движения.

Рассмотрим теперь предпоследний 6-й шаг. Он может начаться в одной из трех точек с координатами (2; 3), (3; 2) и (4; 1). Для точки (2; 3) управление может быть одно (П), общие затраты 14 + 19 = 33 проставляем в кружке. Аналогично для точки (4; 1) управление возможно только (В), суммарные затраты 12 + 13 = 25. В точке (3; 2) ситуация интереснее: из нее можно идти вправо с суммарными затратами 15 + 13 = 28 ед. либо вверх с суммарными затратами 16 + 19 = 35 ед. Выбираем оптимальный вариант смещения право, отмечаем его стрелкой и проставляем величину 28 в кружок.

Аналогично рассматриваем 5-й шаг и перебираем четыре точки, с которых он может начаться, затем 4-й, 3-й, 2-й, 1-й шаги. В результате заполняем кружки во всех узлах и проставляем стрелки в соответствующих участках сетки.

Числа в кружках для каждого узла равны оптимальной величине затрат, с которыми можно из этого узла добраться до конечной точки С, при этом стрелки указывают направление оптимального смещения.

Стадия безусловной оптимизации. Она начинается с исходной точки А. Приписанное ей число 66 означает искомую оптимальную величину затрат для маршрута, соединяющего точку А и точку С. Сам оптимальный маршрут легко получается, если из точки А последовательно продвигаться в направлении стрелок. В результате получаем оптимальное управление (П, П, В, П, П, В, В).

Изложенная процедура применении ДП для решения рассмотренной задачи наглядно иллюстрирует и поясняет черты этого метода, в том числе суть стадий условной и безусловной оптимизации. Кроме того, это решение позволяет сделать ряд выводов, которые носят общий характер.

1. В рассматриваемой задаче фактор времени в явном виде не присутствует. Однако псевдодинамику с дискретным временем в этой задаче удалось ввести искусственно, рассматривая управление (принятие решения о формировании маршрута) как поэтапную, многошаговою процедуру, что вполне соответствует концепции ДП. Эта ситуация довольно типична для многих приложений ДП.

2. В ходе условной оптимизации может возникнуть ситуация, когда на некотором этапе возможно не одно, а два или более управлений (в данном примере – направление «П» или «В») с одинаковыми суммарными затратами. В таком случае можно выбирать любое из них. В результате задача будет иметь не единственное решение (оптимальный маршрут), но все решения будут иметь одну и ту же оптимальную величину критерия качества, в частности суммарную величину затрат.

3. Выполнив первую стадию решения методом ДП (условную оптимизацию), мы получили результат гораздо более содержательный, чем предполагалось поставленной задачей. Найдены оптимальные величины затрат и ориентиры (стрелки) для формирования оптимального маршрута для любого состояния системы (для каждого узла сетки), а не только заданного начального состояния (стартовой точки А).

4. В рассматриваемой задаче и многих других подобных задачах нужно выбрать решение из конечного множества возможных, легко формируемых решений. Возникает вопрос – не проще ли организовать и осуществить прямой перебор этих вариантов с оценкой каждого их них и выбором наилучшего? Решение, реализуемое с помощью ДП, также предполагает некоторый перебор вариантов, но при ДП этот перебор осуществляется целенаправленно. Это позволяет значительно сократить количество

рассматриваемых (обрабатываемых) вариантов по сравнению с «примитивным» перебором всех возможных вариантов. Подробнее об оценке эффективности ДП по сравнению с полным перебором вариантов см. например [11].

5. Сравнивая методику и результаты решения задач на основе ДП с представленными в гл. 4 методами решения оптимизационных задач на графах (задача о кратчайших путях, поиск критического пути, задачи на транспортных сетях), можно увидеть много общего – решение в две стадии с движением от конца к началу и наоборот; получение меток, содержание которых аналогично информации, проставленной в кружках на рис. 6.5 и др.

§ 6.6. Основное функциональное уравнение ДП

При решении задач с помощью ДП бывает целесообразно использовать символьную (формализованную) запись изложенного выше принципа оптимальности Р. Беллмана. Для ее получения необходимо помимо обозначений, введенных в § 6.2, и формул (6.1) – (6.5) ввести дополнительные обозначения.

Будем наряду с общей оценкой эффективности всего процесса управления, выражаемой показателем

$$Y = \sum_{i=1}^{n} y_i ,$$

где y_i – эффективность отдельного i-го шага, рассматривать величину

$$Y_k = \sum_{i=k}^{n} y_i . \tag{6.8}$$

Она определяет эффективность участка рассматриваемого многошагового процесса от k-го шага до последнего включительно.

Условимся снабжать символом (*) переменные, соответствующие оптимальному решению, в частности:

$(\bar{x}^*(1), \bar{x}^*(2), ..., \bar{x}^*(n))$ – оптимальное управление на всех шагах рассматриваемого процесса;

$(\bar{s}^*(1), \bar{s}^*(2), ..., \bar{s}^*(n))$ – последовательность оптимальных состояний системы;

Y^* – оптимальное значение критерия эффективности на всем промежутке управления;

Y_k^* – оптимальное значение критерия качества на участке от k-го шага до последнего.

Тогда принцип оптимальности можно записать следующим образом:

$$Y_k^*(\overline{s}(k-1)) = \max_{\overline{x}(k)}\left[f_k(\overline{s}(k-1), \overline{x}(k)) + Y_{k+1}^*(\overline{s}(k)) \right], \qquad (6.9)$$

где $k = 1, 2, ..., n-1$.

Для последнего шага $k = n$ справедливо соотношение

$$Y_k^*(\overline{s}(k-1)) = \max_{\overline{x}(k)}\left[f_k(\overline{s}(k-1), \overline{x}(k)) \right]. \qquad (6.10)$$

Выражение в правой части (6.9) означает, что управления $\overline{x}(k)$ на k-м шаге выбираются исходя из максимальности функции, стоящей в квадратных скобках. Оно определяется значением критерия эффективности на этом шаге в совокупности с оптимальным значением суммарного эффекта на всех последующих шагах с $(k+1)$-го до n-го включительно. Поскольку за последним шагом нет никаких шагов, формула (6.9) справедлива только до предпоследнего шага включительно, а на последнем шаге справедлива формула (6.10).

Таким образом, принцип оптимальности и его формализованная запись (6.9), (6.10) имеют довольно простой и ясный смысл. Однако их практическое применение требует от пользователя усилий, прежде всего, для описания поставленной задачи как задачи ДП и соответствующей организации вычислительного процесса. Многое здесь определяется спецификой задачи.

Проиллюстрируем методологию ДП на нескольких примерах (§ 6.7 – 6.11).

§ 6.7. Задача об оптимальном единовременном распределении выделенных средств между предприятиями

Рассматривается группа предприятий $\Pi_1, \ldots, \Pi_n$, входящих в некий холдинг. Его руководство выделяет некоторый объем средств s_0 для развития этих предприятий. Известны функции, определяющие получаемый доход в зависимости от величины вложенных средств $f_i(\overline{x}(i))$ для каждого предприятия $\varPi_i$. Требуется найти такое распределение величины s_0 между предприятиями, которое даст максимальный общий доход.

Если обозначить через x_i величину средств, выделенных i-му предприятию, то задача сводится к определению чисел $x_i = x_i^*$, при которых целевая функция принимает наибольшее значение:

$$Y = \sum_{i=1}^{n} f_i(x_i) \to \sup, \tag{6.11}$$

при этом выполняется условие

$$x_1 + x_2 + ... + x_n = s_0. \tag{6.12}$$

В этой задаче фактор времени отсутствует. По сути она является статической задачей и относится к задачам условной оптимизации. Такие задачи рассматривались в § 5.1 и § 5.5. Однако подобного рода задачи можно искусственно свести к динамическим, если принятие решения о распределении средств представить как пошаговую процедуру, при которой на первом этапе выделяются средства предприятию $\varPi_1$, на втором – $\varPi_2$ и т.д. В результате получается дискретный процесс, состоящий из n этапов.

Искомые n чисел $x_1,...,x_n$ в таком случае будут рассматриваться как n значений одного изменяющегося поэтапно фактора x, который можно интерпретировать как управление. Введем еще одну переменную $s = s_i$ – остаток средств после их распределения на предшествующих этапах: s_0 – выделенная для распределения начальная величина средств: $s_1 = s_0 - x_1$ – величина средств, остающихся для распределения после первого шага, когда x_1 средств выделено $\varPi_1$ и т.д. Эту переменную можно интерпретировать как состояние. При этом соотношение

$$s_i = s_{i-1} - x_i \tag{6.13}$$

является уравнением состояний, аналогичным (6.6).

Важно, что показатель качества

$$Y = \sum_{i=1}^{n} f_i(x_i)$$

соответствует условию аддитивности (6.7).

Введенная интерпретация и полученные соотношения (6.11) – (6.13) дают возможность рассматривать и решать поставленную задачу как задачу ДП.

Для любого k-го шага ($k = 1,...,(n-1)$) этого процесса принятия решений о выделении средств можно записать уравнение Беллмана

$$Y_k^*(s_{k-1}) = \max_{0 \le x_k \le s_{k-1}} \left[f_k(x_k) + Y_{k+1}^*(s_k) \right]. \tag{6.14}$$

Значение $x_k \in \left[0, s_{k-1}\right]$ является допустимым управлением для k-го шага.

Для последнего шага $k = n$ и справедливо соотношение (6.10)

$$Y_k^*(s_{k-1}) = \max_{0 \le x_k \le s_{k-1}} \left[f_k(x_k) \right]. \tag{6.15}$$

Пусть $n = 4$. Рассмотрим уравнение Беллмана (6.14) – (6.15) в соответствии с процедурой условной оптимизации, принятой в ДП, последовательно от последнего шага к первому. Получим следующие расчетные соотношения:

$$k = 4 \qquad Y_4^*(s_3) = \max_{0 \le x_4 \le s_3} \left[f_4(x_4) \right],$$

$$k = 3 \qquad Y_3^*(s_2) = \max_{0 \le x_3 \le s_2} \left[f_3(x_3) + Y_4^*(s_3) \right],$$

$$k = 2 \qquad Y_2^*(s_1) = \max_{0 \le x_2 \le s_1} \left[f_2(x_2) + Y_3^*(s_2) \right],$$

$$k = 1 \qquad Y_1^*(s_1) = \max_{0 \le x_1 \le s_0} \left[f_1(x_1) + Y_2^*(s_1) \right],$$

при этом $s_k = s_{k-1} - x_k$.

Получили последовательность расчетных соотношений, в которых результат, полученный на одном шаге, используется в расчетах на другом шаге. Такие соотношения называются рекуррентными. Поэтому условие оптимальности (6.9) – (6.10) часто называют рекуррентным уравнением Беллмана.

Далее для решения задачи нужно конкретизировать исходные условия. Будем, следуя [13], считать, что распределению подлежит величина $s_0 = 200$ у.е., например 200 млн руб. или 200 тыс. долларов. Средства выделяются в размерах, кратных 40 у.е., а функции дохода заданы в табл. 6.1.

Таблица 6.1

x \ f	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
40	8	6	3	4
80	10	9	4	6
120	11	11	7	8
160	12	13	11	13
200	18	15	18	16

Оптимизация на каждом этапе в соответствии с приведенными выше условиями будет связана с перебором вариантов, который удобно осуществить с помощью формы, представленной в табл. 6.2.

Таблица 6.2

s_{k-1}	x_k $(0 \le x_k \le s_{k-1})$	$s_{k-1} - x_k$	$f_k(x_k)$	Y_{k+1}^*	$\left[f_k + Y_{k+1}^*\right]$	$\max\left[f_k + Y_{k+1}^*\right] = Y_k^*$
1	2	3	4	5	6	7

В результате получаем следующую последовательность таблиц (табл. 6.3 – 6.6).

Таблица 6.3

s_3	x_4 $(0 \le x_4 \le s_3)$	Для $k = 4$ не заполняется	$f_4(x_4)$	Для $k = 4$ не заполняется	$[f_4]$	$\max[...] = Y_4^*$
1	2	3	4	5	6	7
40	0 40		0 4		0 4	4
80	0 40 80		0 4 6		0 4 6	6
120	0 40 80 120		0 4 6 8		0 4 6 8	8
160	0 40 80 120 160		0 4 6 8 13		0 4 6 8 13	13
200	0 40 80 120 160 200		0 4 6 8 13 16		0 4 6 8 13 16	16

Примечание: $k = 4$

Таблица 6.4

s_2	x_3 $(0 \le x_3 \le s_2)$	$s_2 - x_3 = s_3$	$f_3(x_3)$	$Y_4^*(s_3)$	$\left[f_3 + Y_4^*\right]$	$\max[...] = Y_3^*$
1	2	3	4	5	6	7
40	0	40	0	4	4	4
	40	0	3	0	3	
80	0	80	0	6	6	
	40	40	3	4	7	7
	80	0	4	0	4	
120	0	120	0	8	8	
	40	80	3	6	9	9
	80	40	4	4	8	
	120	0	7	0	7	
160	0	160	0	13	13	13
	40	120	3	8	11	
	80	80	4	6	10	
	120	40	7	4	11	
	160	0	11	0	11	
200	0	200	0	16	16	
	40	160	3	13	16	
	80	120	4	8	12	
	120	80	7	6	13	
	160	40	11	4	15	
	200	0	18	0	18	18

Примечание. $k = 3$.

Таблица 6.5

s_1	x_2 $(0 \le x_2 \le s_1)$	$s_1 - x_2 = s_2$	$f_2(x_2)$	$Y_3^*(s_2)$	$\left[f_2 + Y_3^*\right]$	$\max[...] = Y_2^*$
1	2	3	4	5	6	7
40	0	40	0	4	4	
	40	0	6	0	6	6
80	0	80	0	7	7	
	40	40	6	4	10	10
	80	0	9	0	9	
120	0	120	0	9	9	
	40	80	6	7	13	
	80	40	9	4	13	13
	120	0	11	0	11	
160	0	160	0	13	13	
	40	120	6	9	15	
	80	80	9	7	16	16
	120	40	11	4	15	
	160	0	13	0	13	
200	0	200	0	18	18	
	40	160	6	13	19	19
	80	120	9	9	18	
	120	80	11	7	18	
	160	40	13	4	17	
	200	0	15	0	15	

Примечание. $k = 2$.

Т а б л и ц а 6.6

s_0	x_1 $(0 \le x_1 \le s_0)$	$s_0 - x_1 = s_1$	$f_1(x_1)$	$Y_2^*(s_1)$	$\left[f_1 + Y_2^*\right]$	$\max[\ ] = Y_1^*$
1	2	3	4	5	6	7
40	0	40	0	6	6	
	40	0	8	0	8	8
80	0	80	0	10	10	
	40	40	8	6	14	14
	80	0	10	0	10	
120	0	120	0	13	13	
	40	80	8	10	18	18
	80	40	10	6	16	
	120	0	11	0	11	
160	0	160	0	16	16	
	40	120	8	13	21	21
	80	80	10	10	20	
	120	40	11	6	17	
	160	0	12	0	12	
200	0	200	0	19	19	
	40	160	8	16	24	24
	80	120	10	13	23	
	120	80	11	10	21	
	160	40	12	6	18	
	200	0	18	0	18	

Примечание. $k = 1$.

Пояснения. Табл. 6.3, соответствующая последнему этапу, определяется упрощенным соотношением (6.10), и она могла бы быть составлена проще, но для единообразия представлена в том же виде, что и остальные табл. 6.4 – 6.6. В каждой из них колонки 1, 2, 3 заполняются очевидным способом. Колонка 4 заполняется из соответствующей колонки исходной табл. 6.1. Колонка 5 заполняется данными, полученными в последней колонке 7 предыдущей таблицы. Числа в колонке 6 получаются как сумма чисел из 4-й и 5-й колонок. В 7-й колонке выбирается наибольшее число из чисел, полученных в 6-й колонке.

Формирование этих таблиц (табл. 6.3 – 6.6) соответствует стадии условной оптимизации.

Перейдем теперь ко второй стадии – безусловной оптимизации. В последней из полученных выше таблиц (табл. 6.6.), соответствующей первому шагу, получено значение $Y_1^*(200) = 24$. Это означает, что максимальный доход, который можно получить при распределении средств в размере 200 у.е., составляет 24 ед. Здесь же получается оптимальное значение $x_1 = x_1^* = 40$. Из уравнения состояния находим $s_1^* = 200 - 40 = 160$. Из табл. 6.5 для этого состояния оптимальным значением является $x_2^* = 80$. Вычисляем $s_2^* = 160 - 80 = 80$. Из табл. 6.4 находим $x_3^* = 40$, тогда $s_3^* = s_2^* - x_3^* = 80 - 40 = 40$. Из табл. 6.3 в строке, соответствующей $S_3^* = 40$, получаем $x_4^* = 40$.

Таким образом, максимальный доход, равный 24 у.е., будет получен при следующем распределении выделенных средств в размере 200 у.е.:

$П_1 \to 40$ у.е., $П_2 \to 80$ у.е., $П_3 \to 40$ у.е., $П_4 \to 40$ у.е.

§ 6.8. Задача об оптимальном поэтапном распределении выделенных средств между предприятиями в течение планового периода

Рассмотрим более сложный вариант задачи о распределении средств между предприятиями, чем в § 6.7. Необходимо разработать план финансирования предприятий $П_1, ..., П_m$ на период в n лет. Начальный общий объем выделенных средств для всех предприятий задан и составляет s_0 у.е. Следуя [13], допустим, что средства x, вложенные в $П_j$, позволяют получить в конце текущего года доход $f_j(x)$ и выделить средства в размере $\varphi_j(x)$ для формирования общего фонда, который используется для нового распределения средств между предприятиями при финансировании их в следующем году. При этом никаких иных средств для финансирования не поступает.

Требуется найти оптимальный план распределения средств между предприятиями на каждый год планируемого периода.

Решение. Будем рассматривать процесс распределения средств как многошаговый, полагая, что номер шага k соответствует номеру года. Управляемую систему предприятий $(П_1, ..., П_m)$ можно характеризовать одной переменной состояния $s = s(k-1)$ $(k = 1, 2, ..., n)$ – количеством средств, подлежащих перераспределению в начале k-го года. Переменных управления будет m: $x_1(k), ..., x_m(k)$ – количество средств, выделенных $П_1, П_2, ..., П_m$ на k-й год. Так как средства ежегодно перераспределяются полностью, то

$$x_1(k) + x_2(k) + ... + x_m(k) = s(k-1). \qquad (6.16)$$

Следуя [13], будем далее решать эту задачу при следующих условиях:

$$m = 2;\ n = 4;\ s_0 = 10000;$$

$$\left.\begin{array}{l} f_1(x) = 0,4x;\ f_2(x) = 0,3x; \\ \varphi_1(x) = 0,5x;\ \varphi_2(x) = 0,8x. \end{array}\right\} \qquad (6.17)$$

В этом случае уравнение (6.16) примет вид

$$x_1(k) + x_2(k) = s(k-1) \tag{6.18}$$

или

$$x_2(k) = s(k-1) - x_1(k). \tag{6.19}$$

Уравнение (6.18) означает, что фактически в процессе решения на каждом шаге нужно находить значение только одной переменной $x_1 = x_1^*(k)$, а вторая величина x_2 однозначно определяется из этого уравнения (6.18). Это позволяет рассматривать управляемый процесс с двумя переменными (x_1, x_2) как одномерный. Будем переменную $x_1(k)$ для краткости обозначать x_k.

В этой задаче показатель эффективности – доход, полученный от обоих предприятий за период в n лет – определяется соотношением

$$Y = \sum_{i=1}^{n} \left[f_1(x_i) + f_2(s_{i-1} - x_i) \right]. \tag{6.20}$$

Уравнение состояния, выражающее средства, формируемые по окончании i-го шага для последующего распределения между предприятиями, имеет вид

$$s_i = \varphi_1(x_i) + \varphi_2(s_{i-1} - x_i). \tag{6.21}$$

Подставляя в это уравнение соотношения (6.17), получаем

$$s_i = 0{,}5x_i + 0{,}8(s_{i-1} - x_i) = 0{,}8s_{i-1} - 0{,}3x_i. \tag{6.22}$$

Доход, полученный в k-м году от обоих предприятий, в сумме составит величину

$$f_k = f_1(x_k) + f_2(s_{k-1} - x_k) = 0{,}4x_k + 0{,}3(s_{k-1} - x_k) = 0{,}1x_k + 0{,}3s_{k-1}. \tag{6.23}$$

Запишем для рассматриваемого случая рекуррентное соотношение (6.14), выражающее принцип оптимальности

$$Y_k^*(s_{k-1}) = \max_{0 \le x_k \le s_{k-1}} \left[f_1(x_k) + f_2(s_{k-1} - x_k) + Y_{k+1}^*(s_k) \right],$$

где $k = 1, 2, \ldots, n-1$. Для $k = n$

$$Y_n^*(s_{n-1}) = \max_{0 \le x_n \le s_{n-1}} \left[f_1(x_n) + f_2(s_{n-1} - x_n) \right].$$

Подставляя в эти уравнения соотношения (6.22), (6.23), получим

$$k = 4: Y_4^*(s_{k-1}) = \max_{0 \le x_4 \le s_3} (0{,}1x_k + 0{,}3s_{k-1}),$$

$$k = 1, 2, 3: Y_k^*(s_{k-1}) = \max_{0 \le x_k \le s_{k-1}} (0{,}1x_k + 0{,}3s_{k-1} + Y_{k+1}^*(0{,}8s_{k-1} - 0{,}3x_k)).$$

Реализуем стадию условной оптимизации, рассматривая процесс от последнего года к первому. Для $k = 4$ справедливо соотношение

$$Y_4^*(s_3) = \max_{0 \le x_4 \le s_3} (0{,}1x_4 + 0{,}3s_3) = 0{,}4s_3,$$

т.к. выражение в скобках является линейной функцией, а она достигает максимума в конце промежутка, т.е. при $x_4^* = s_3$.

Для $k = 3$: $Y_3^*(s_2) = \max\limits_{0 \le x_3 \le s_2} (0{,}1x_3 + 0{,}3s_2 + 0{,}4s_3) =$

$$= \max_{0 \le x_3 \le s_2} \left[0{,}1x_3 + 0{,}3s_2 + 0{,}4(0{,}8s_2 - 0{,}3x_3)\right] = \max_{0 \le x_3 \le s_2} \left[-0{,}02x_3 + 0{,}62s_2\right].$$

Выражение в скобках здесь, как и в предыдущем случае, является линейным, коэффициент при x_3 отрицателен, поэтому наибольшее значение это выражение получит в начале промежутка, т.е. при $x_3 = 0$. При этом

$$Y_3^*(s_2) = 0{,}62s_2,$$

$$k = 2:\ Y_2^*(s_1) = \max_{0 \le x_2 \le s_1} (0{,}1x_2 + 0{,}3s_1 + 0{,}62(0{,}8s_1 - 0{,}3x_2)) =$$

$$= \max_{0 \le x_2 \le s_1} \left[0{,}796s_1 - 0{,}086x_2\right] = 0{,}796s_1 \ \ (\text{при } x_2^* = 0),$$

$$k = 1:\ Y_1^*(s_0) = \max_{0 \le x_1 \le s_0} (0{,}1x_1 + 0{,}3s_0 + 0{,}796(0{,}8s_0 - 0{,}3x_1)) =$$

$$= \max_{0 \le x_1 \le s_0} \left[-0{,}1388x_1 + 0{,}9368s_0\right] = 0{,}9368s_0 \ \ (\text{при } x_1^* = 0).$$

Таким образом, на стадии условной оптимизации получены следующие результаты:

$Y_1^*(s_0) = 0{,}9368s_0;\ x_1^* = 0;$

$Y_2^*(s_1) = 0{,}796s_1;\ x_2^* = 0;$

$Y_3^*(s_2) = 0{,}62s_2;\ x_3^* = 0;$

$Y_4^*(s_3) = 0{,}4s_3;\ x_4^* = s_3.$

Переходим к стадии безусловной оптимизации.

Полагая $s_0 = 10000$, находим $Y_{\max} = 9368$, $x_1^* = 0$. Используя соотношение (6.22), получаем

$$s_1^* = 0{,}8 \cdot 10000 - 0{,}3 \cdot 0 = 8000.$$

Для второго года планирования при этом значении s_1^* и $x_2^* = 0$ находим

$$s_2^* = 0{,}8s_1^* - 0{,}3x_2^* = 0{,}8 \cdot 8000 = 6400.$$

Аналогично для $x_3^* = 0$ находим

$$s_3^* = 0{,}8s_2^* - 0{,}3x_3^* = 0{,}8 \cdot 6400 = 5120.$$

И, наконец, определяем

$$x_4^* = s_3 = 5120.$$

В исходной постановке задачи фигурировали две искомые переменные $x_1(k)$ и $x_2(k)$ для $П_1$ и $П_2$ соответственно. В полученном решении найдены значения только одной переменной $x_1(k)$. Вторая определяется из соотношения (6.19)

$$x_2(k) = s(k-1) - x_1(k).$$

В результате получаем план финансирования двух предприятий, который представлен в табл. 6.7.

Таблица 6.7

Год / $П_j$	1-й	2-й	3-й	4-й
$П_1$	0	0	0	5120
$П_2$	10000	8000	6400	0

При таком использовании первоначально выделенных средств $s_0 = 10000$ через четыре года будет получен максимально возможный доход $Y_{\max} = 9368$.

Заметим, что в этой задаче в отличие от предыдущей задачи переменные x и s могут изменяться непрерывно, а зависимости $f(x)$ и $\varphi(x)$ заданы аналитически, а не таблично. В [13] рассмотрен дискретный вариант этой задачи, когда переменные x изменяются с некоторым шагом Δx, а зависимости $f(x)$ и $\varphi(x)$ задаются таблично.

§ 6.9. Задача об оптимальном плане замены оборудования

Одной из распространенных и важных задач оптимального планирования в экономике является задача определения наиболее эффективной стратегии замены действующего оборудования на новое. В качестве оборудования могут рассматриваться станки, автомобили, агрегаты, технологические линии, производственные или торговые здания и помещения и т.д.

С течением времени в процессе эксплуатации любое оборудование стареет – происходит его физический и моральный износ. При этом возрастают финансовые и временные затраты на его ремонт и обслуживание, а вместе с этим снижается производительность оборудования и так называемая ликвидная стоимость. Настает момент, когда эксплуатационные затраты для действующего оборудования вырастают настолько, что его выгоднее продать и заменить новым оборудованием, хотя, конечно, приобретение его связано с некоторыми, как правило, значительными затратами.

В связи с изложенными обстоятельствами возникает задача об определении оптимального с точки зрения выбранного (или заданного) критерия срока замены оборудования. В качестве такого критерия обычно принимают либо суммарную прибыль, полученную при реализации производственного или иного процесса на старом и новом оборудовании (с учетом затрат и на эксплуатацию старого оборудования и на приобретение нового), либо суммарные затраты на эксплуатацию оборудования в течение определенного промежутка времени. Заметим, что замена оборудования может состоять в приобретении оборудования того же вида, что и старое (естественно, без износа или с меньшим износом), либо нового более совершенного.

Рассмотрим далее изложенную задачу в более узкой и конкретной формулировке [10, 13].

Задача 1. Рассматривается плановый промежуток времени, состоящий из n лет. Вопрос о замене оборудования на новое или сохранении действующего оборудования рассматривается в начале каждого года. Известны следующие исходные условия:

- зависимость производительности оборудования от времени эксплуатации (со временем она уменьшается);
- зависимость затрат на содержание, эксплуатацию и ремонт этого оборудования от времени (они со временем увеличиваются);
- величина затрат на приобретение, установку (монтаж) и запуск в эксплуатацию нового оборудования;
- заменяемое старое оборудование списывается (не продается).

Требуется составить план замены оборудования, при котором общая прибыль, получаемая при эксплуатации оборудования за n лет, была бы максимальна.

Решение. Для получения математической модели этой задачи введем следующие обозначения:

t – текущее время в плановом периоде; решение о замене или сохранении оборудования принимается в моменты времени $t = 0;\ 1;\ 2;....;\ n-1$;

τ – возраст оборудования; $\tau \in \{0,\ 1,\ 2,...\}$; $\tau = 0$ соответствует использованию нового оборудования, $\tau = 1$ соответствует эксплуатации оборудования с возрастом 1 год и т.д.;

$r = r(\tau)$ – производительность оборудования – годовой выпуск продукции (объем выпускаемой продукции или объем проданных товаров, оказанных услуг и т.д.) в стоимостном выражении для оборудования возраста τ;

$z = z(\tau)$ – эксплуатационные затраты, т.е. затраты на обслуживание и ремонт, которые приходится осуществлять в течение одного года для оборудования, имеющего возраст τ;

$p = p_0$ – стоимость нового оборудования.

В рассматриваемом случае управляемой ***системой*** является оборудование, ***управлением*** – решение о сохранении или замене оборудования. Если это управление обозначить переменной x, то она будет принимать два значения

$$x = \begin{cases} x_1 = x^c - \text{сохранить действующее оборудование;} \\ x_2 = x^з - \text{заменить оборудование на новое.} \end{cases}$$

В качестве ***состояния*** этой системы будем рассматривать возраст оборудования: $s = \tau$.

Для получения уравнения состояний (6.6) рассмотрим i-й год планового периода.

Пусть к началу этого года оборудование подошло имея возраст τ, тогда в случае принятия решения о его сохранении к следующему моменту рассмотрения это оборудование станет на год старше $s(i) = \tau + 1$.

Если же оборудование будет заменено на новое, то по истечении рассматриваемого промежутка времени (одного года) его возраст будет равен $s(i) = 1$. Таким образом, получаем следующее уравнение состояния:

$$s(i-1) = \tau \xrightarrow[x=x(i)]{} s(i) = \begin{cases} \tau + 1 \text{ при } x = x^c, \\ 1 \ \text{ при } \ x = x^з. \end{cases} \tag{6.24}$$

Рассмотрим теперь величину критерия эффективности использования оборудования, т.е. принятия решений о сохранении действующего оборудования или его замене на новое в течение пятилетки. Следуя [1, 10], примем в качестве этого критерия следующий показатель:

$$Y = \sum_{i=1}^{5} y_i\,, \tag{6.25}$$

где $y(i) = \begin{cases} r(\tau_i) - z(\tau_i), \text{ если } x_i = x^c - \text{оборудование сохраняется;} \\ r(0) - z(0) - p, \text{ если } x_i = x^з - \text{оборудование заменяется.} \end{cases}$

Заметим, что этот показатель интерпретируют как общую прибыль от эксплуатации оборудования в течение пятилетки [1] либо суммарный доход [10], хотя в действительности это не соответствует общепринятым в экономике понятиям прибыли и дохода.

Функциональное рекуррентное уравнение Беллмана (6.9) в данном случае принимает вид

$$Y_k^*(\tau_k, x_k) = \max_{x_k} \begin{cases} r(\tau_k) - z(\tau_k) + Y_{k+1}^*(\tau_k + 1), \\ r(0) - z(0) + Y_{k+1}^*(1). \end{cases}$$

В соответствии с [1, 10] будем считать, что функции $r(\tau)$ и $z(\tau)$, определяющие производительность оборудования и эксплуатационные затраты, заданы таблично (табл. 6.8), а затраты, связанные с приобретением и установкой нового оборудования, $p = 40$ тыс. руб., причем заменяемое оборудование списывается.

Т а б л и ц а 6.8

Функция	Возраст τ, лет					
	-	1	2	3	4	5
$r(\tau)$, тыс. руб.	80	75	65	60	60	55
$z(\tau)$, тыс. руб.	20	25	30	35	45	55

Рассмотрим стадию условной оптимизации, начиная с последнего (5-го) года планового периода и предполагая, что вначале оборудование было новым, т.е. имело возраст $\tau = 0$.

Полагаем $k = 5$. В этом случае множество допустимых состояний s к началу пятого года (т.е. множество возможных значений возраста оборудования) будет состоять из значений: $\tau_5 \in \{1,\ 2,\ 3,\ 4\}$.

В соответствии с (6.10) для последнего этапа Y_{k+1}^* в уравнении Беллмана равно 0. Поэтому

$$Y_5^*(\tau_5) = \max_{x_5} \begin{cases} r(\tau_5) - z(\tau_5), \\ r(0) - z(0) - p. \end{cases}$$

Используя значения функций $r(\tau)$ и $z(\tau)$ из табл. 6.8, получаем следующие результаты:

$$\tau_5 = 1,\ Y_5^*(\tau_5) = \max \begin{cases} r(1) - z(1) \\ r(0) - z(0) - p \end{cases} = \max \begin{Bmatrix} 75 - 25 \\ 80 - 20 - 40 \end{Bmatrix} = \max \begin{Bmatrix} 50 \\ 20 \end{Bmatrix};$$

$$Y_5^*(\tau_5) = 50 \text{ при } x = x^c;$$

$$\tau_5 = 2,\ Y_5^*(\tau_5) = \max \begin{cases} r(2) - z(2) \\ r(0) - z(0) - p \end{cases} = \max \begin{Bmatrix} 65 - 30 \\ 80 - 20 - 40 \end{Bmatrix} = \max \begin{Bmatrix} 35 \\ 20 \end{Bmatrix};$$

$$Y_5^*(\tau_5) = 35 \text{ при } x = x^c;$$

$$\tau_5 = 3,\ Y_5^*(\tau_5) = \max \begin{cases} r(3) - z(3) \\ r(0) - z(0) - p \end{cases} = \max \begin{Bmatrix} 60 - 35 \\ 80 - 20 - 40 \end{Bmatrix} = \max \begin{Bmatrix} 25 \\ 20 \end{Bmatrix};$$

$$Y_5^*(\tau_5) = 25 \text{ при } x = x^c;$$

$$\tau_5 = 4,\ Y_5^*(\tau_5) = \max\begin{cases} r(4) - z(4) \\ r(0) - z(0) - p \end{cases} = \max\begin{Bmatrix} 60 - 45 \\ 80 - 20 - 40 \end{Bmatrix} = \max\begin{Bmatrix} 15 \\ 20 \end{Bmatrix};$$

$$Y_5^*(\tau_5) = 20 \text{ при } x = x^з.$$

Полученные результаты сведем в табл. 6.9.

Таблица 6.9

τ_5	$Y_5^*(\tau_5)$	x_5^*
1	50	x^c
2	35	x^c
3	25	x^c
4	20	$x^з$

Полагаем $k = 4$. В этом случае $\tau_4 \in \{1,\ 2,\ 3\}$,

$$Y_4^*(\tau_4) = \max_{x_4}\begin{cases} r(\tau_4) - z(\tau_4) + Y_5^*(\tau_4 + 1), \\ r(0) - z(0) - p + Y_5^*(1). \end{cases}$$

$$\tau_4 = 1,\ Y_4^*(\tau_4) = \max\begin{cases} r(1) - z(1) + Y_5^*(2) \\ r(0) - z(0) - p + Y_5^*(1) \end{cases} = \max\begin{Bmatrix} 75 - 25 + 35 \\ 80 - 20 - 40 + 50 \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} 85 \\ 70 \end{Bmatrix} = 85 \text{ при } x_4 = x^c;$$

$$\tau_4 = 2,\ Y_4^*(\tau_4) = \max\begin{cases} r(2) - z(2) + Y_5^*(3) \\ r(0) - z(0) - p + Y_5^*(1) \end{cases} = \max\begin{Bmatrix} 65 - 30 + 25 \\ 80 - 20 - 40 + 50 \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} 60 \\ 70 \end{Bmatrix} = 70 \text{ при } x_4 = x^з;$$

$$\tau_4 = 3,\ Y_4^*(\tau_4) = \max\begin{cases} r(3) - z(3) + Y_5^*(4) \\ r(0) - z(0) - p + Y_5^*(1) \end{cases} = \max\begin{Bmatrix} 60 - 35 + 20 \\ 80 - 20 - 40 + 50 \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} 45 \\ 70 \end{Bmatrix} = 70 \text{ при } x_4 = x^з.$$

Полученные результаты представим в табл. 6.10.

Таблица 6.10

τ_4	$Y_4^*(\tau_4)$	x_4^*
1	85	x^c
2	70	$x^з$
3	70	$x^з$

Полагаем $k = 3$. В этом случае $\tau_3 \in \{1, 2\}$,

$$Y_3^*(\tau_3) = \max_{x_3}\begin{cases} r(\tau_3) - z(\tau_3) + Y_4^*(\tau_3 + 1), \\ r(0) - z(0) - p + Y_4^*(1); \end{cases}$$

$$\tau_3 = 1,\; Y_3^*(\tau_3) = \max\begin{Bmatrix} r(1) - z(1) + Y_4^*(2) \\ r(0) - z(0) - p + Y_4^*(1) \end{Bmatrix} = \max\begin{Bmatrix} 75 - 25 + 70 \\ 80 - 20 - 40 + 85 \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} 120 \\ 105 \end{Bmatrix} = 120 \text{ при } x_3 = x^c;$$

$$\tau_3 = 2,\; Y_3^*(\tau_3) = \max\begin{Bmatrix} r(2) - z(2) + Y_4^*(3) \\ r(0) - z(0) - p + Y_4^*(1) \end{Bmatrix} = \max\begin{Bmatrix} 65 - 30 + 70 \\ 80 - 20 - 40 + 85 \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} 105 \\ 105 \end{Bmatrix} = 105 \text{ при } x_3 = x^c \text{ или } x_3 = x^{\text{з}}.$$

Как видим, в том случае, когда к началу 3-го года пятилетки возраст оборудования составляет два года, можно в качестве условно оптимального принять любое решение: сохранить оборудование или заменить, эффективность использования оборудования в каждом из этих случаев будет одинаковой. Примем для определенности $x_3 = x^3$.

Результат для k = 3 сведем в табл. 6.11.

Т а б л и ц а 6.11

τ_3	$Y_3^*(\tau_3)$	x_3^*
1	120	x^c
2	105	$x^{\text{з}}$

Полагаем k = 2. К началу второго года возраст оборудования может быть равен только одному году, т.е. $\tau_2 = 1$.

$$Y_2^*(\tau_2) = \max\begin{Bmatrix} r(1) - z(1) + Y_3^*(2) \\ r(0) - z(0) - p + Y_3^*(1) \end{Bmatrix} = \max\begin{Bmatrix} 75 - 25 + 105 \\ 80 - 20 - 40 + 120 \end{Bmatrix} =$$

$$= \max\begin{Bmatrix} 155 \\ 140 \end{Bmatrix} = 155 \text{ при } x_2 = x^c.$$

Этот результат представим в виде табл. 6.12.

Т а б л и ц а 6.12

τ_2	$Y_2^*(\tau_2)$	x_2^*
1	155	x^c

Полагаем, наконец, k = 1. В соответствии с исходными условиями в начале пятилетки, т.е. в начале 1-го года, установлено новое оборудование, значит, его возраст $\tau_1 = 0$. Поэтому выбора на 1-м этапе фактически нет – оборудование на 1-м году сохраняется. Значит, условно оптимальным решением является $x_1 = x^c$. При этом

$$Y_1^*(\tau_1) = r(0) - z(0) + Y_2^*(1) = 80 - 20 + 155 = 215.$$

Это означает, что максимально возможное значение оптимизируемого критерия (6.25) равно 215 у.е. Ему соответствует оптимальный план замены оборудования, который получается на второй стадии вычислительного процесса – стадии безусловной оптимизации. Она состоит в последовательном использовании табл. 6.9 – 6.12 для 1-го, 2-го, …, 5-го года пятилетки.

Для первого года решение единственное – оборудование следует сохранить ($x_1 = x^c$). Значит его возраст к началу второго года будет равен $\tau = 1$. Тогда в соответствии с табл. 6.12 оптимальным для второго года будет решение о сохранении оборудования. Значит, к началу третьего года его возраст будет равен $\tau = 2$. В соответствии с табл. 6.11 на третьем году оборудование следует заменить. В таком случае к началу четвертого года его возраст составит $\tau = 1$. В соответствии с табл. 6.10 при таком возрасте на четвертом году пятилетки оборудование менять не следует. При этом к началу пятого года его возраст станет равен $\tau = 2$ и в соответствии с табл. 6.9 оборудование менять не нужно.

Таким образом, мы получили следующий оптимальный план принятия решений о судьбе рассматриваемого оборудования:

1-й год – сохранить; 2-й год – сохранить; 3-й год – заменить;

4-й год – сохранить; 5-й год – сохранить.

Рассмотрим далее еще одну задачу о замене оборудования. Ее особенностью будет то, что затраты на эксплуатацию, начальная и ликвидная стоимость оборудования зависят не только от возраста оборудования, но и от времени, прошедшего с начала анализируемого процесса.

Формулировка задачи [13]. Анализируется ситуация, связанная с эксплуатацией автомобиля. В начале анализируемого промежутка времени машина является новой, она приобретена за $p_0 = 5000$ у.е. Из-за инфляции и других причин цены на новые машины со временем постоянно растут.

По прогнозам этот рост составляет 10 % в год, поэтому через k лет автомобиль этой же модели будет стоить (у. е.)

$$p_k = 5000 + 500(k-1),\ k = 1, 2, 3, ... \tag{6.26}$$

Ежегодные затраты на эксплуатацию автомашины, ее ремонт и обслуживание со временем растут пропорционально возрасту машин на 10 % в год. Эту зависимость можно описать следующим образом:

$$r_k(\tau) = 0{,}1 p_k(\tau+1),\ \tau = 0,\ 1,\ 2, ... \tag{6.27}$$

Ликвидная стоимость, по которой автомашину можно продать, со временем падает. Анализ рынка показывает, что после τ лет эксплуатации машину, приобретенную в k-м году, можно продать по цене

$$z_k(\tau) = \frac{p_k}{2^\tau} \text{ у.е.} \tag{6.28}$$

В начале каждого года из рассматриваемого промежутка времени нужно принимать решение – продолжить дальнейшую эксплуатацию машины или заменить ее на новую.

Требуется сформировать оптимальный план действий, т.е. определить, когда автомашину следует продать и купить новую так, чтобы суммарные затраты были минимальны.

Анализируемый промежуток времени ограничен семью годами, причем решение о продолжении эксплуатации машины или ее замене принимается по истечении 1-го, 2-го, …, 6-го, а после 7-го года эксплуатации она в любом случае продается, а новая не покупается.

Решение. Сначала рассмотрим для этой задачи сущность ключевых понятий: управления, состояния и критерия эффективности, а также математические соотношения, определяющие их. Управление в этой задаче носит тот же характер, что и в предыдущем случае. Оно может принимать два значения:

$$x = \begin{cases} x^c - \text{если машина сохраняется,} \\ x^з - \text{если она заменяется.} \end{cases}$$

Состояние системы s_{k-1} в начале k-го шага (года) характеризуется возрастом машины τ. Уравнение состояния, как и в предыдущем случае, определяется соотношением (6.24), т.е.

$$s_k = \begin{cases} s_{k-1} + 1, \text{ если } x_k = x^c, \\ 1, \text{ если } x_k = x^з. \end{cases}$$

Если в предыдущем случае критерием эффективности был доход от эксплуатации оборудования, который нужно было сделать максимальным, то в данном случае эффективность определяется затратами и их, естественно, нужно минимизировать.

Если в начале *k*-го года принято решение не продавать машину, то затраты на эксплуатацию будут определяться величиной $r_k(\tau)$ в соответствии с (6.27). Если же машина в начале *k*-го года продана, то затраты будут определяться расходами за год эксплуатации нового автомобиля, т.е.

$$r_k(0) = 0{,}1 p_k .$$

Кроме того, в затраты нужно включить расходы на приобретение нового автомобиля в *k*-м году p_k и вычесть выручку от продажи старой машины $z_k(\tau)$, определяемую (6.28).

Таким образом, мы получаем

$$y_k(\tau) = \begin{cases} 0{,}1 p_k(\tau + 1), \text{ если } x_k = x^c, \\ p_k + 0{,}1 p_k - p_k 2^{-\tau} = p_k(1{,}1 - 2^{-\tau}), \text{ если } x_k = x^з. \end{cases} \quad (6.29)$$

Используя это соотношение, запишем функциональное уравнение Беллмана (6.9) для рассматриваемого случая:

$$Y_k^*(\tau) = \min \begin{cases} 0{,}1 p_k(\tau + 1) + Y_{k+1}^*(\tau + 1), \text{ если } x_k = x^c, \\ p_k(1{,}1 - 2^{-\tau}) + Y_{k+1}^*(\tau + 1), \text{ если } x_k = x^з \end{cases} \quad (6.30)$$

для $k = 1, 2, ..., 5$;

При $k = 6$ уравнение принимает следующий вид:

$$Y_6^*(\tau) = \min \begin{cases} 0{,}1 p_6(\tau + 1) - p_7 2^{-(\tau+1)}, \text{ если } x_6 = x^c, \\ p_6(1{,}1 - 2^{-\tau}) - p_7 2^{-1}, \text{ если } x_6 = x^з. \end{cases} \quad (6.31)$$

В этом уравнении вторая компонента учитывает выручку от продажи автомобиля в конце последнего 7-го года. Если подставить в (6.31) значения $p_6 = 5000 + 500 \cdot 5 = 7500$, $p_7 = 5000 + 500 \cdot 6 = 8000$, определенные в соответствии с (6.26), то вместо (6.31) получим:

$$Y_6^*(\tau) = \min \begin{cases} 750(\tau + 1) - 8000 \cdot 2^{-(\tau+1)}, \text{ если } x_6 = x^c, \\ 4250 - 7500 \cdot 2^{-\tau}, \text{ если } x_6 = x^з, \end{cases} \quad (6.32)$$

где $\tau \in \{0, 1, 2, 3, 4, 5\}$.

Рассмотрим стадию условной оптимизации, начиная с 6-го шага (года). Для него найдем компоненты, входящие в (6.32) для различных τ. Результаты вычислений сведем в табл. 6.13.

Т а б л и ц а 6.13

τ	$750(\tau+1)-8000\cdot 2^{-(\tau+1)}$	$4250-7500\cdot 2^{-\tau}$	$Y_6^*(\tau)$	$x_6(\tau)$
0	750 – 4000 = –3250	4250 – 7500 = – 3250	– 3250	x^c или $x^з$
1	1500 – 2000 = –500	4250 – 3750 = 500	– 500	x^c
2	2250 – 1000 = 1250	4250 – 1875=2375	1250	x^c
3	3000 – 500 = 2500	4250 – 937,5=3312,5	2500	x^c
4	3750 – 250 = 3500	4250 – 468,75=3781,25	3500	x^c
5	4500 – 125 = 4375	4250 – 234,375=4015,625	4015, 625	$x^з$

Условная оптимизация на 5-м, 4-м, …, 1-м шагах будет определяться соотношением (6.30), которое после подстановки вместо p_k выражения (6.26) и некоторых преобразований примет следующий вид:

$$Y_k^*(\tau)=\min\begin{cases}50(k+9)(\tau+1)-Y_{k+1}^*(\tau+1), \text{ если } x_k=x^c,\\ 500(k+9)(1{,}1-2^{-\tau})+Y_{k+1}^*(1), \text{ если } x_k=x^з.\end{cases} \qquad (6.33)$$

Вычисления, проведенные по (6.33) для различных $k=5,4,3,2,1$ и различных $\tau<k$, приведены в табл. 6.14. В каждой строке этой таблицы $Y_k^*(\tau)$ выбирается как минимальное значение из $Y_k(\tau,x^c)$ и $Y_k(\tau,x^з)$, в таблице оно подчеркнуто. Соответственно выбирается и x_k^*.

Т а б л и ц а 6.14

k	τ	50(τ+1)* *(k+9)	$Y_{k+1}^*(\tau+1)$	$Y_k(\tau,x^c)$	500(k+9)* *(1,1-2$^{-\tau}$)	$Y_{k+1}^*(1)$	$Y_k(\tau,x^з)$	$Y_k^*(\tau)$	x_k^*
5	0	700	-500	<u>200</u>	7700-7000=700	-500	<u>200</u>	200	$x^c(x^з)$
	1	1400	1250	<u>2650</u>	7700-3500=4200	-500	3700	2650	x^c
	2	2100	2500	<u>4600</u>	7700-1750=5950	-500	5450	4600	x^c
	3	2800	3500	<u>6300</u>	7700-875=6825	-500	6325	6300	x^c
	4	3500	4015,6	7515,6	7700-437,5=7262,5	-500	<u>6762,5</u>	6792,5	$x^з$
4	0	650	2650	<u>3300</u>	7150-6500=650	2650	<u>3300</u>	3300	$x^c(x^з)$
	1	1300	4600	<u>5900</u>	7150-3250=3900	2650	6550	5900	x^c
	2	1950	6300	8250	7150-1625=5525	2650	<u>8175</u>	8175	x^c
	3	2600	6762,5	9362,5	7150-812,5=6337,5	2650	<u>8987,5</u>	8987,5	x^c

k	τ	50(τ+1)* *(k+9)	$Y^*_{k+1}(\tau+1)$	$\overline{Y_k}(\tau, x^c)$	500(k+9)* *(1,1-2^{-τ})	$Y^*_{k+1}(1)$	$\overline{Y_k}(\tau, x^з)$	$Y^*_k(\tau)$	x^*_k
3	0	600	5900	<u>6500</u>	6600-6000=600	5900	<u>6500</u>	6500	$x^c(x^з)$
	1	1200	8187	<u>9375</u>	6600-3000=3600	5900	9500	9375	x^c
	2	1800	8987,5	<u>10787,5</u>	6600-1500=5100	5900	11000	10787,5	x^c
2	0	550	9375	<u>9925</u>	6050-5500= 550	9375	<u>9925</u>	9925	$x^c(x^з)$
	1	1100	10787,5	<u>11887,5</u>	6050-2750=3300	9375	12675	11887,5	x^c
1	0	5500	11887,5	<u>17387,5</u>				17387,5	x^c

На стадии безусловной оптимизации, двигаясь по табл. 6.14 снизу вверх, получаем последовательность действий

$$x_1^* = x^c; x_2^* = x^c; x_3^* = x^c; x_4^* = x^з; x_5^* = x^c; x_6^* = x^c.$$

При этом минимальная величина суммарных затрат на эксплуатацию автомашины, включая ее перепродажу на 4-м году, составляет

$$Y^*_{\min} = 17\,387{,}5 \text{ у.е.}$$

§ 6.10. Задача календарного планирования трудовых ресурсов

Рассмотрим задачу об оптимальном регулировании численности работников в следующей формулировке [32]. Предпринимателю требуется составить план приема и увольнения рабочих определенной профессии на ближайшие 5 месяцев. Исходя из имеющихся договоров, на поставку продукции предприятия составлен производственный план на указанный период, а по нему определена потребность a_i в рабочих на каждый месяц. Пусть в i-м месяце предусматривается работа x_i человек. При излишней численности работников, когда $x_i > a_i$, возникают убытки, связанные с простоями $(x_i - a_i)$ человек. Их величина υ_i пропорциональна избытку рабочей силы

$$\upsilon_i = b(x_i - a_i). \tag{6.34}$$

При найме новых рабочих возникают затраты, связанные с их оформлением, страхованием, обучением и т.д. Величину этих затрат w_i определим соотношением:

$$w_i = \begin{cases} c + k(x_i - x_{i-1}), \text{ если } x_i > x_{i-1}, \\ 0, \text{ если } x_i \le x_{i-1}. \end{cases} \tag{6.35}$$

В первом случае при переходе с (i-1)-го месяца в i-й месяц приняты новые рабочие, во втором случае – новых рабочих не принимали, их численность сохранилась или произведены увольнения. Предполагается, что при увольнении накладных расходов не возникает.

Требуется составить оптимальный план регулирования численности рабочих на рассматриваемый 5-месячный период, при котором суммарные затраты, связанные с приемом новых рабочих и простоями при избытке работников, были бы минимальны. Кроме того, будем предполагать, что исходное количество имеющихся работников задано: $x_0 = a_0$. Будем также исходить из того, что на любом *i*-м этапе $x_i \geq a_i$, т.е. недостатка рабочих, когда $x_i < a_i$, допускать нельзя.

Для конкретности примем

$$a_0 = a_1 = 5;\ a_2 = 7;\ a_3 = 8;\ a_4 = 4;\ a_5 = 6;$$

$$b = 3;\ c = 4;\ k = 2.$$

Для этих данных график потребности в рабочих представлен на рис 6.6 пунктирной линией.

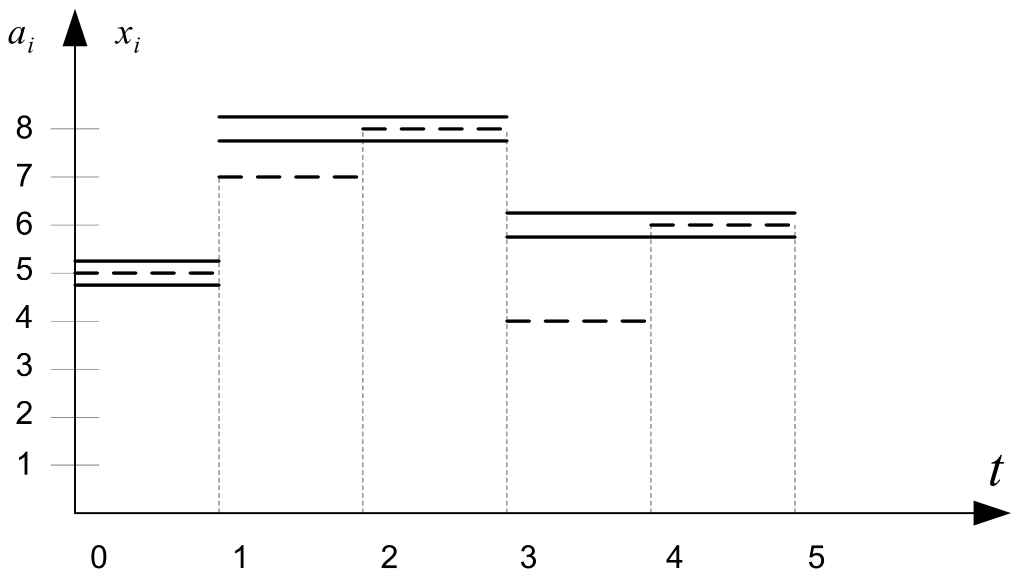


Рис. 6.6

Решение. Определим сначала ключевые для ДП понятия – этап, состояние, управление, критерий оптимальности. В этой задаче этапы определяются естественным образом – каждый месяц является соответствующим этапом. Состояние s_{i-1} в начале *i*-го этапа определим количеством рабочих x_{i-1}, имеющихся на предприятии к концу (*i*-1)-го месяца. Управление на *i*-м этапе состоит в определении численности работников x_i.

Критерий эффективности – суммарные затраты, связанные с наймом новых рабочих и простоями при их избытке, определяется соотношением

$$Y = \sum_{i=1}^{5} y_i , \tag{6.36}$$

где $y_i = \upsilon_i + w_i$.

Требуется найти значения неизвестных величин $x_i = x_i^*$ ($i = \overline{1,5}$), при которых $Y \to \inf$.

Из сущности сформулированной задачи вытекают дополнительные условия, которым должно удовлетворять решение x_i:

$x_0 = x^{\text{н}} = 5$;

$x_5 = a_5 = 6$ – на этом этапе нет смысла содержать лишнюю рабочую силу, которая могла бы пригодиться в будущем, т.к. этот этап последний;

$a_i \le x_i \le a_{\text{макс}} = \sup\limits_{1 \le i \le 5} \{a_i\} = 8$; $i = \overline{1,4}$;

$a_i \le s_i \le a_{\text{макс}}$;

$X_i = S_{i-1}$, где X и S – множества возможных значений x_i и s_i на соответствующих этапах.

Уравнение Беллмана (6.9) в рассматриваемом случае примет следующий вид:

$$\left.\begin{aligned} Y_i^*(s_{i-1}) &= \min_{x_i \ge a_i} \left\{ \upsilon_i + w_i + Y_{i+1}^*(s_i) \right\}, \text{ для } i = \overline{1,4}; \\ Y_i^*(s_{i-1}) &= \min_{x_i = a_i} \left\{ \upsilon_i + w_i \right\}, \text{ для } i = 5, \end{aligned}\right\} \tag{6.37}$$

где υ_i и w_i определяются соотношениями (6.34) и (6.35).

Стадия условной оптимизации

<u>Этап № 5</u>. В этом случае, как отмечалось, $x_5 = a_5$ и $(a_4 = 4) \le s_4 \le (a_{\max} = 8)$, т.е. $s_4 \in \{4,\ 5,\ 6,\ 7,\ 8\}$. Рассмотрим каждое возможное начальное состояние и проведем вычисления по второй формуле из (6.37).

Так как $x_5 = a_5 = 6$, то простоев на этом этапе нет, поэтому затраты $\upsilon_i = 0$. В зависимости от состояния s_4 переход на 5-й этап будет связан либо с наймом рабочих (при $x_5 > s_4$), либо с их увольнением (при $x_5 < s_4$).

Следовательно, получим:

для $s_4 = 4$ нужно нанять двух работников, поэтому

$$Y_5^*(s_4) = \upsilon_i + w_i = w_i = 4 + 2 \cdot 2 = 8;$$

для $s_4 = 5$ нужно нанять одного работника, поэтому

$$Y_5^*(s_4) = w_i = 4 + 2 \cdot 1 = 6;$$

для $s_4 = 6$ не требуется ни нанимать, ни увольнять рабочих, поэтому

$$Y_5^*(s_4) = w_i = 0;$$

для $s_4 = 7$ и $s_4 = 8$ нужно уволить одного или двух человек – по условиям задачи это не требует каких-либо затрат, поэтому в том и в другом случае:

$$Y_5^*(s_4) = 0.$$

Для всех начальных состояний $x_5^* = 6$.

Этап № 4. В этом случае $X_4 = S_4$, т.е. $x_4 \in \{4, 5, 6, 7, 8\}$ и $(a_3 = 8) \le s_3 \le (a_{\max} = 8)$, т.е. $s_3 = 8$. Так как $x_4 \le s_3 = 8$, то в любом случае найма работников не требуется, а необходимы будут увольнения, поэтому $w_4 = 0$.

По первой формуле из (6.37) получаем следующие результаты:

для $x_4 = 4$ простоев нет ($\upsilon_4 = 0$) и требуется уволить $(s_4 - x_4) = 4$ чел., ($w_4 = 0$), поэтому

$$Y_4(s_3) = 0 + 0 + Y_5^*(s_4) = 8;$$

для $x_4 = 5$ получаются убытки из-за простоя одного работника $\upsilon_4 = 3 \cdot 1$ и требуется уволить 3 чел., ($w_4 = 0$), поэтому

$$Y_4(s_3) = 3 \cdot 1 + 0 + Y_5^*(s_4) = 3 + 6 = 9;$$

для $x_4 = 6$ получаются убытки из-за простоя двух человек, поэтому

$$Y_4(s_3) = 3 \cdot 2 + 0 + Y_5^*(s_4) = 6 + 0 = 6;$$

для $x_4 = 7$ получаются убытки из-за простоя трех человек, поэтому

$$Y_4(s_3) = 3 \cdot 3 + 0 + 0 = 9;$$

для $x_4 = 8$ излишняя численность равна 4, поэтому

$$Y_4(s_3) = 3 \cdot 4 + 0 + 0 = 12.$$

Наименьшее значение Y_5^* получается равным $Y_5^*(s_3 = 8) = 6$ при $x_4^* = 6$.

Этап № 3. На этом этапе $a_3 = 8$ и $x_3 = 8$; $(a_2 = 7) \le s_2 \le (a_{\max} = 8)$, т.е. $s_2 \in \{7, 8\}$. Простоев нет, поэтому $\upsilon_3 = 0$;

для $s_2 = 7$ при $x_3 = 8$ следует нанять одного работника: $w_3 = 4 + 2 \cdot 1 = 6$, в результате получаем $Y_3^*(s_2) = 0 + 6 + 6 = 12$; для $s_2 = 8$ при $x_3 = 8$ не требуется ни нанимать, ни увольнять работников. В результате получаем

$$Y_3^*(s_2) = 0 + 0 + 6 = 6.$$

Отметим, что $x_3^* = 8$.

Этап № 2. В этом случае $a_2 = 7$; $X_2 = S_2 = \{7, 8\}$; $(a_1 = 5) \le s_1 \le (a_{\max} = 8)$, т.е. $s_1 \in \{5, 6, 7, 8\}$.

Для $s_1 = 5$ при $x_2 = 7$ простоев нет, следовательно, $\upsilon_2 = 0$; требуется нанять двух человек:

$$Y_2(s_1) = 0 + 4 + 2 \cdot 2 + Y_3^*(s_2) = 0 + 8 + 12 = 20;$$

при $x_2 = 8$ будут приняты три человека, но один из них будет простаивать:

$$Y_2(s_1) = 3 \cdot 1 + 4 + 2 \cdot 3 + 6 = 19.$$

Выбираем наименьшее значение из полученных: $Y_2^*(s_1 = 5) = 19$ при $x_2^* = 8$.

Для $s_1 = 6$ при $x_2 = 7$ аналогично получим

$$Y_2(s_1) = 0 + 4 + 2 \cdot 1 + 12 = 18;$$

при $x_2 = 8$

$$Y_2(s_1) = 3 \cdot 1 + 4 + 2 \cdot 2 + Y_3^*(s_2) = 3 + 8 + 6 = 17.$$

Выбираем наименьшее значение

$$Y_2^*(s_1 = 6) = 17 \text{ при } x_2^* = 8.$$

Для $s_1 = 7$ при $x_2 = 7$ находим

$$Y_2(s_1) = 0 + 0 + 12 = 12;$$

при $x_2 = 8$

$$Y_2(s_1) = 3 \cdot 1 + 4 + 2 \cdot 1 + 6 = 15.$$

Выбираем $Y_2^*(s_1 = 7) = 12$ при $x_2^* = 7$.

Для $s_1 = 8$ при $x_2 = 7$ получаем

$$Y_2(s_1) = 0 + 0 + 12 = 12;$$

при $x_2 = 8$

$$Y_2(s_1) = 3 \cdot 1 + 0 + 6 = 9.$$

Выбираем

$$Y_2^*(s_1 = 8) = 9 \text{ при } x_2^* = 8.$$

Этап № 1. $a_1 = 5$; $X_1 = S_1 = \{5, 6, 7, 8\}$; $s_0 = s_{1\text{н}} = 5$.

Для $s_0 = 5$ при $x_1 = 5$ имеем

$$Y_1(s_0) = 0 + 0 + 19 = 19;$$

при $x_1 = 6$

$$Y_1\ (s_0) = 3 \cdot 1 + 4 + 2 \cdot 1 + 17 = 26;$$

при $x_1 = 7$

$$Y_1\ (s_0) = 3 \cdot 2 + 4 + 2 \cdot 2 + 12 = 26;$$

при $x_1 = 8$

$$Y_1\ (s_0) = 3 \cdot 3 + 4 + 2 \cdot 3 + 9 = 28.$$

Выбираем наименьшее значение

$$Y_1^*(s_0 = 5) = 19 \text{ при } x_1^* = 5.$$

В краткой форме результаты проделанных вычислений приведены в табл. 6.15 – 6.19.

Этап № 5 Таблица 6.15

s_4	$y_i = \upsilon_i + w_i$	$Y_5^*(s_4)$	x_5^*
	$x_5 = 6$		
4	$3 \cdot 0 + 4 + 2 \cdot 2 = 8$	8	6
5	$3 \cdot 0 + 4 + 2 \cdot 1 = 6$	6	6
6	$3 \cdot 0 + 0 + 2 \cdot 0 = 0$	0	6
7	$3 \cdot 0 + 0 + 2 \cdot 0 = 0$	0	6
8	$3 \cdot 0 + 0 + 2 \cdot 0 = 0$	0	6

Этап № 4 Таблица 6.16

s_3	$\upsilon_i + w_i + Y_5^*(s_4)$					$Y_4^*(s_3)$	x_4^*
	$x_4 = 4$	$x_4 = 5$	$x_4 = 6$	$x_4 = 7$	$x_4 = 8$		
8	3·0+0+8=8	3·1+0+ 0+6=9	3·2+0+0=6	3·3+0+0=9	3·4+0+0=12	6	6

Этап № 3 Таблица 6.17

s_2	$\upsilon_i + w_i + Y_4^*(s_3)$	$Y_3^*(s_2)$	x_3^*
	$x_3 = 8$		
7	0 + 4 + 2 · 1 + 6 = 12	12	8
8	0 + 0 + 0 + 6 = 6	6	8

Этап № 2 Т а б л и ц а 6.18

s_1	$\upsilon_i + w_i + Y_3^*(s_2)$		$Y_2^*(s_1)$	x_2^*
	$x_2 = 7$	$x_2 = 8$		
5	3·0+4+2·2+12=20	3·1+4+2·3+6=$\underline{19}$	19	8
6	3·0+4+2·1+12=18	3·1+4+2·2+6=$\underline{17}$	17	8
7	3·0+0+12=$\underline{12}$	3·1+4+2·1+6=15	12	7
8	3·0+0+12=12	3·1+0·1+6=$\underline{9}$	9	8

Этап № 1 Т а б л и ц а 6.19

s_0	$\upsilon_i + w_i + Y_2^*(s_1)$				$Y_1^*(s_0)$	x_1^*
	$x_1 = 5$	$x_1 = 6$	$x_1 = 7$	$x_1 = 8$		
5	0+0+19=19	3·1+4+2·1+17 = 26	3·2+4+2·2+12 = 26	3·3+4+2·3+9 =28	19	5

Стадия безусловной оптимизации.

Двигаясь от начального этапа к финальному пятому этапу, последовательно находим

$$x_0 = 5 = x_1^* \to x_2^* = 8 \to x_3^* = 8 \to x_4^* = 6 \to x_5^* = 6 .$$

Оптимальный план регулирования численности рабочих представлен в табл. 6. 20.

Т а б л и ц а 6. 20

Месяц	Минимальная потребность, чел	Оптимальное число работников, чел	Необходимое решение
1-й	5	5	Никого не принимать, никого не увольнять
2-й	7	8	Нанять трех рабочих
3-й	8	8	Никого не принимать, никого не увольнять
4-й	4	6	Уволить двух рабочих
5-й	6	6	Никого не принимать, никого не увольнять

Контрольные вопросы

1. Какая система называется управляемой?
2. Какая система называется статической, а какая – динамической?
3. Какая система является непрерывной, а какая – дискретной? Что называется периодом (шагом) дискретности?
4. Какая система считается детерминированной?
5. Какие системы называются системами с неопределенностями?
6. Как формулируется классическая задача об оптимальном динамическом управлении?
7. Что представляет собой критерий оптимизации в задаче динамического управления?
8. Сформулировать классическую задачу динамического программирования, в чем ее принципиальные особенности?
9. Как формулируется принцип оптимальности Р. Беллмана?
10. Записать функциональное уравнение ДП, реализующее принцип оптимальности Р. Беллмана.
11. На какие стадии разделяется процесс решения задач динамического программирования методом Р. Беллмана?
12. Как формулируется задача об оптимальном единовременном распределении выделенных средств между предприятиями? Каким образом определяются в ней понятия «управление», «состояние» и «критерий оптимальности»?
13. Как формулируется задача об оптимальном поэтапном распределении выделенных средств между предприятиями в течение планового периода?
14. Как формируется задача об оптимальном плане замены оборудования? Как определяются понятия «управление», «состояние» и «критерий оптимальности» в этой задаче?
15. Как формулируется задача календарного планирования трудовых ресурсов? Как определяются «управление», «состояние» и «критерий оптимальности» в этом случае?

Задачи для самостоятельного решения

Найти решение сформулированной в § 6.7 задачи об оптимальном единовременном распределении средств между предприятиями для $s_0 = 250$; $\Delta s = \Delta x = 50$; $k = 4$ и функций доходности, заданных приведенными ниже таблицами.

№ 1-1

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	5	7	4	8
100	7	10	7	13
150	8	12	10	17
200	9	13	12	18
250	10	14	14	19

№ 1-2

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	4	3	7	5
100	5	5	9	8
150	7	6	11	11
200	10	8	12	13
250	15	9	12	15

№ 1-3

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	2	4	5	7
100	6	8	8	10
150	9	11	12	11
200	11	13	15	13
250	12	14	16	15

№ 1-4

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	7	5	6	4
100	9	7	9	8
150	11	8	12	10
200	12	9	14	12
250	13	11	15	13

№ 1-5

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	6	7	5	4
100	10	11	8	8
150	13	13	11	12
200	15	14	13	15
250	16	15	14	17

№ 1-6

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	5	4	6	5
100	9	9	9	7
150	13	14	12	9
200	15	16	15	12
250	16	18	17	15

№ 1-7

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	7	6	6	5
100	10	9	7	8
150	13	11	8	11
200	15	12	10	14
250	17	13	12	16

№ 1-8

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	5	7	6	8
100	9	10	9	9
150	13	13	12	10
200	15	15	14	11
250	17	16	15	11

№ 1-9

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	6	8	5	7
100	10	11	9	10
150	14	13	12	13
200	16	15	14	15
250	17	16	15	17

№ 1-10

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	7	6	8	5
100	9	8	11	9
150	11	9	14	13
200	12	10	16	15
250	13	10	17	16

№ 1-11

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	5	7	8	6
100	7	10	10	9
150	8	12	12	11
200	9	14	13	12
250	9	15	14	14

№ 1-12

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	4	6	5	5
100	8	8	8	9
150	12	10	10	13
200	14	11	13	16
250	15	11	16	18

№ 1-13

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	7	6	4	5
100	10	10	7	9
150	13	12	10	12
200	15	14	12	14
250	16	15	14	15

№ 1-14

$x \backslash f$	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
50	6	4	5	7
100	10	9	8	9
150	13	13	11	12
200	15	15	12	15
250	17	16	12	18

Найти решение сформулированной в § 6.8 задачи о поэтапном распределении средств между предприятиями в течение планового периода при $m = 2$; $n = 4$; $s_0 = 100000$ у.е. для функций f_i и φ_i, заданных следующими формулами:

№ 2-1

$f_1 = 0,3x$; $f_2 = 0,5x$;
$\varphi_1 = 0,4x$; $\varphi_2 = 0,7x$.

№ 2-2

$f_1 = 0,7x$; $f_2 = 0,6x$;
$\varphi_1 = 0,8x$; $\varphi_2 = 0,5x$.

№ 2-3

$f_1 = 0,4x$; $f_2 = 0,6x$;
$\varphi_1 = 0,7x$; $\varphi_2 = 0,6x$.

№ 2-4

$f_1 = 0,4x$; $f_2 = 0,3x$;
$\varphi_1 = 0,5x$; $\varphi_2 = 0,4x$.

№ 2-5

$f_1 = 0,35x$; $f_2 = 0,5x$;
$\varphi_1 = 0,6x$; $\varphi_2 = 0,8x$.

№ 2-6

$f_1 = 0,55x$; $f_2 = 0,8x$;
$\varphi_1 = 0,9x$; $\varphi_2 = 0,65x$.

№ 2-7

$f_1 = 0,9x$; $f_2 = 0,7x$;
$\varphi_1 = 0,65x$; $\varphi_2 = 0,75x$.

№ 2-8

$f_1 = 0,7x$; $f_2 = 0,5x$;
$\varphi_1 = 0,6x$; $\varphi_2 = 0,8x$.

№ 2-9

$f_1 = 0,2x$; $f_2 = 0,3x$;
$\varphi_1 = 0,4x$; $\varphi_2 = 0,3x$.

№ 2-10

$f_1 = 0,8x$; $f_2 = 0,6x$;
$\varphi_1 = 0,4x$; $\varphi_2 = 0,5x$.

№ 2-11

$f_1 = 0,35x$; $f_2 = 0,45x$;
$\varphi_1 = 0,7x$; $\varphi_2 = 0,3x$.

№ 2-12

$f_1 = 0,6x$; $f_2 = 0,45x$;
$\varphi_1 = 0,4x$; $\varphi_2 = 0,8x$.

№ 2-13

$f_1 = 0,75x$; $f_2 = 0,5x$;
$\varphi_1 = 0,4x$; $\varphi_2 = 0,65x$.

№ 2-14

$f_1 = 0,4x$; $f_2 = 0,25x$;
$\varphi_1 = 0,5x$; $\varphi_2 = 0,9x$.

Решить сформулированную в § 6.9 задачу об определении оптимального плана замены оборудования при условии, что затраты по приобретению и установке нового оборудования равны $p = 50000$ у.е., а функции, определяющие производительность оборудования и эксплуатационные затраты, заданы приведенными ниже таблицами:

№ 3-1

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	100	80	75	70	66	60
$z(\tau)$, тыс. руб.	15	18	23	26	28	30

№ 3-2

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	90	80	75	72	70	64
$z(\tau)$, тыс. руб.	22	27	31	37	45	50

№ 3-3

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	70	60	46	40	37	33
$z(\tau)$, тыс. руб.	25	29	33	35	36	36

№ 3-4

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	80	65	55	48	41	38
$z(\tau)$, тыс. руб.	20	30	37	42	45	47

№ 3-5

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	75	74	73	70	65	60
$z(\tau)$, тыс. руб.	10	11	11	12	14	18

№ 3-6

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	100	98	95	91	86	80
$z(\tau)$, тыс руб.	15	16	17	19	23	28

№ 3-7

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс руб.	60	60	58	55	50	43
$z(\tau)$, тыс руб.	12	12	13	15	18	24

№ 3-8

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс руб.	50	50	48	46	42	36
$z(\tau)$, тыс. руб.	18	19	20	22	25	29

№ 3-9

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	65	64	64	63	61	56
$z(\tau)$, тыс. руб.	25	26	28	31	35	40

№ 3-10

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	85	84	82	79	75	69
$z(\tau)$, тыс. руб.	30	30	31	31	34	38

№ 3-11

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб	95	95	93	90	86	81
$z(\tau)$, тыс. руб.	11	12	13	15	18	22

№ 3-12

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс руб.	120	119	117	112	106	100
$z(\tau)$, тыс руб.	22	23	25	28	32	35

№ 3-13

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб	150	148	145	140	130	110
$z(\tau)$, тыс. руб	25	26	28	32	37	45

№ 3-14

	τ, лет					
	0	1	2	3	4	5
$r(\tau)$, тыс. руб.	130	128	125	120	110	95
$z(\tau)$, тыс. руб.	30	32	34	40	46	55

Решить сформулированную в § 6.10 задачу календарного планирования трудовых ресурсов для следующих исходных данных:

№ 4-1

$a_0 = a_1 = 7;\ a_2 = 5;\ a_3 = 8;$

$a_4 = 6;\ a_5 = 4;$

$\upsilon_i = 2(x_i - a_i);$

$$w_i = \begin{cases} 1 + 3(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-2

$a_0 = a_1 = 6;\ a_2 = 8;\ a_3 = 6;$

$a_4 = 5;\ a_5 = 4;$

$\upsilon_i = 1{,}5(x_i - a_i);$

$$w_i = \begin{cases} 2 + (x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-3

$a_0 = a_1 = 8;\ a_2 = 6;\ a_3 = 5;$

$a_4 = 7;\ a_5 = 5;$

$\upsilon_i = 2{,}5(x_i - a_i);$

$$w_i = \begin{cases} 2 + 1{,}5(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-4

$a_0 = a_1 = 4;\ a_2 = 7;\ a_3 = 6;$

$a_4 = 8;\ a_5 = 7;$

$\upsilon_i = 3(x_i - a_i);$

$$w_i = \begin{cases} 1 + 2(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-5

$a_0 = a_1 = 6;\ a_2 = 8;\ a_3 = 10;$

$a_4 = 7;\ a_5 = 6;$

$\upsilon_i = (x_i - a_i);$

$$w_i = \begin{cases} 2{,}5 + 1{,}5(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-6

$a_0 = a_1 = 7;\ a_2 = 10;\ a_3 = 9;$

$a_4 = 8;\ a_5 = 5;$

$\upsilon_i = 2(x_i - a_i);$

$$w_i = \begin{cases} 2 + (x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-7

$a_0 = a_1 = 8;\ a_2 = 7;\ a_3 = 10;$

$a_4 = 6;\ a_5 = 5;$

$\upsilon_i = 4(x_i - a_i);$

$$w_i = \begin{cases} 0,5 + 1,5(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-8

$a_0 = a_1 = 7;\ a_2 = 9;\ a_3 = 8;$

$a_4 = 10;\ a_5 = 6;$

$\upsilon_i = 1,5(x_i - a_i);$

$$w_i = \begin{cases} 1 + 0,75(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0 \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-9

$a_0 = a_1 = 4;\ a_2 = 7;\ a_3 = 8;$

$a_4 = 10;\ a_5 = 6;$

$\upsilon_i = 3,5(x_i - a_i);$

$$w_i = \begin{cases} 4 + 0,5(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0,75(x_i - x_{i-1}) \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-10

$a_0 = a_1 = 6;\ a_2 = 5;\ a_3 = 9;$

$a_4 = 7;\ a_5 = 10;$

$\upsilon_i = 2,5(x_i - a_i);$

$$w_i = \begin{cases} 3 + 4(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 0,5(x_i - x_{i-1}) \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-11

$a_0 = a_1 = 3;\ a_2 = 8;\ a_3 = 7;$

$a_4 = 9;\ a_5 = 5;$

$\upsilon_i = 4(x_i - a_i);$

$$w_i = \begin{cases} 1 + 2(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 1,25(x_i - x_{i-1}) \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-12

$a_0 = a_1 = 4;\ a_2 = 9;\ a_3 = 10;$

$a_4 = 7;\ a_5 = 6;$

$\upsilon_i = 3(x_i - a_i);$

$$w_i = \begin{cases} 1,5 + 3,5(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 2,5(x_i - x_{i-1}) \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-13

$a_0 = a_1 = 2;\ a_2 = 4;\ a_3 = 8;$

$a_4 = 7;\ a_5 = 5;$

$\upsilon_i = 2(x_i - a_i);$

$$w_i = \begin{cases} 0,5 + 2(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 1,5(x_i - x_{i-1}) \text{ при } x_i \le x_{i-1}. \end{cases}$$

№ 4-14

$a_0 = a_1 = 10;\ a_2 = 8;\ a_3 = 6;$

$a_4 = 5;\ a_5 = 7;$

$\upsilon_i = 1,5(x_i - a_i);$

$$w_i = \begin{cases} 1 + 4(x_i - x_{i-1}) \text{ при } x_i > x_{i-1}; \\ 1,5(x_i - x_{i-1}) \text{ при } x_i \le x_{i-1}. \end{cases}$$

Глава 7. **ОСНОВЫ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ И ЕГО ПРИМЕНЕНИЕ ДЛЯ РЕШЕНИЯ ЗАДАЧ ДИНАМИЧЕСКОЙ ОПТИМИЗАЦИИ**

§ 7.1. Основные понятия вариационного исчисления

Вариационное исчисление (ВИ) является областью математики, рассматривающей задачи об отыскании функций, при которых заданный функционал в виде определенного интеграла принимает максимальное или минимальное значение. Оно является развитием дифференциального исчисления (ДИ), позволяющего находить максимумы и минимумы (т.е. экстремумы) функции. Как отмечалось в § 5.1, необходимое условие экстремума дифференцируемой функции $x = x(t)$ состоит в требовании

$$x'(t) = 0,$$

позволяющем определить точки t_i, «подозрительные» на экстремум.

Если в основе ДИ лежит понятие приращения и дифференциала, а также связанного с ним понятия производной, то в основе ВИ аналогичную роль играет понятие вариации функции.

Пусть задана некоторая функция $x = x^0(t)$, рассматриваемая на промежутке $t_1 \le t \le t_2$. Выберем новую функцию $x = x(t)$ и рассмотрим разность

$$\delta_x(t) = x(t) - x^0(t),$$

которая характеризует отклонения между указанными функциями $x^0(t)$ и $x(t)$ для различных $t \in [t_1, t_2]$. Ее и называют вариацией.

Рассмотрим некоторый функционал

$$Y = F[\overline{x}(t)].$$

Следует отметить, что правило F функциям $\overline{x}(t) = \left[x_1(t),\, x_2(t) ... \, x_n(t)\right]$ ставит в соответствие число Y. Такое правило (отображение) называется функционалом.

Типичным примером функционала является определенный интеграл

$$Y = \int_0^{t_к} \varphi\left[\overline{x}(t),\, t\right] dt,$$

где φ – некоторое выражение.

В [24] в качестве показателя, оценивающего качество функционирования экономической системы, рассматривается, например, суммарная дисконтированная величина непроизводственного потребления, определенная по промежутку времени $t_{\text{н}} = 0 \le t \le t_{\text{к}}$:

$$Y = \int\limits_{t_{\text{н}}}^{t_{\text{к}}} \gamma(t) c(t) dt \to \sup,$$

где $c(t)$ – непроизводственное потребление,

$\gamma(t)$ – функция дисконтирования, которая отражает меру предпочтения потребления в различные моменты времени.

Обычно предполагают [24], что $\gamma(t) \ge 0$; $\max\limits_{0 \le t \le t_k} \gamma(t) = \gamma(0)$ и $d\gamma/dt < 0$.

Другим примером интегрального показателя качества может служить выражение

$$Y = \int\limits_{0}^{t_{\text{к}}} e^2(t) dt \to \inf,$$

где $e = s^0(t) - s(t)$,

$s^0(t)$ – желаемый закон изменения фактора s,

$s(t)$ – реальный процесс, наблюдаемый в системе,

e – мера несоответствия (ошибка или рассогласование).

В дискретных системах аналогом интегральных показателей являются суммарные показатели, например

$$Y = \sum_{i=1}^{N} \varepsilon^2(t_i) \to \inf,$$

где i – номер шага, такта или этапа.

Заметим, что в динамических системах показатель качества (функционал (6.2)) необязательно выражается интегралом. Приведем два примера показателей иного рода:

$$Y = \sup\limits_{0 \le t \le t_{\text{к}}} \left| e(t) \right| \to \inf,$$

$$Y = x_i(t_{\text{к}}) \to \sup.$$

Будем считать, что рассматриваемый функционал F на функции $x = x^0(t)$ имеет локальный минимум, если существует такое число $\varepsilon > 0$, что $F[x^0(t)] < F[x^0(t)+\delta_x(t)]$ для всех вариаций $\delta_x(t)$, удовлетворяющих условию

$$0 < \left|\delta_x(t)\right| < \varepsilon .$$

Аналогично определяются максимумы функционала. Функционал, также как и функция, может иметь несколько локальных экстремумов. В таком случае глобальный максимум (минимум), который обычно необходимо найти в прикладных задачах, следует искать среди локальных максимумов (минимумов), определив в каждом из них значение показателя Y и выбрав наилучший результат.

§ 7.2. Классические задачи ВИ и соотношения для их решения

Задача 1. Необходимо отыскать функцию $x = x(t)$ для $t \in [t_\text{н}, t_\text{к}]$, удовлетворяющую условиям:

$$x(t_\text{н}) = x^\text{н}; \quad x(t_\text{к}) = x^\text{к} \tag{7.1}$$

и для которой функционал в виде интеграла

$$J = \int_{t_\text{н}}^{t_\text{к}} \varphi[x(t), x'(t), t]dt \tag{7.2}$$

принимает максимальное (или минимальное) значение.

Ключевой результат, относящийся к ВИ, представляет собой необходимое условие экстремума этого функционала. В соответствии с ним функция $x = x(t)$, доставляющая минимум или максимум функционалу (7.2), должна удовлетворять дифференциальному уравнению

$$\frac{\partial\varphi}{\partial x} - \frac{d}{dt}\left(\frac{\partial\varphi}{\partial x'}\right) = 0 . \tag{7.3}$$

Это условие называют уравнением Эйлера [17, 22, 34], или уравнением Эйлера – Лагранжа [30, 31].

Помимо приведенного уравнения Эйлера существуют и другие необходимые условия экстремальности функционала: условия Лежандра, Вейерштрасса, Якоби [22, 31].

Пример. Найти функцию $x = x(t)$, для которой

$x(0) = 0; \quad x(1) = 1$, а также

$$J = \int_0^1 (x^2(t) + x'^2(t))dt \to \min .$$

Решение. В этом случае $\varphi = x^2 + x'^2$;

$$\frac{\partial\varphi}{\partial x} = 2x; \quad \frac{\partial\varphi}{\partial x'} = 2x'.$$

Условие (7.3) дает следующее уравнение:

$$2x - (2x')' = 0 \text{ или } x'' - x = 0.$$

Рассмотрим характеристическое уравнение для этого ДУ:

$$\lambda^2 - 1 = 0.$$

По его корням $\lambda_1 = 1$; $\lambda_2 = -1$ запишем общее решение

$$x = c_1 e^t + c_2 e^{-t}.$$

Подставляя граничные условия, получаем уравнения для отыскания неизвестных констант c_1, c_2:

$$\left.\begin{aligned} x(0) = c_1 + c_2 = 0, \\ x(1) = c_1 e^1 + c_2 e^{-1} = 1, \end{aligned}\right\},$$

откуда находим

$$c_1 = -c_2; \; c_1 = \frac{1}{e^1 - e^{-1}}.$$

Таким образом, искомая оптимальная траектория, соединяющая точки (0;0) и (1;1), определяется следующим образом:

$$x = x^*(t) = \frac{1}{e^1 - e^{-1}}\left(e^t - e^{-t}\right); \quad t \in [0,1].$$

Задача 2. Она является обобщением задачи 1 и состоит в отыскании нескольких функций $x_1(t), \dots, x_n(t)$, удовлетворяющих граничным условиям

$$x_i(t_\text{н}) = x_i^\text{н}; \quad x_i(t_\text{к}) = x_i^\text{к}, \; i = \overline{1,n} \tag{7.4}$$

и обеспечивающих максимальное или минимальное значение функционала

$$J = \int\limits_{t_\text{н}}^{t_\text{к}} \varphi[x_1(t),\dots,x_n(t); x_1'(t),\dots,x_n'(t); t]\,dt. \tag{7.5}$$

Необходимое условие экстремальности этого показателя представляет собой следующую систему ДУ [17, 22, 31]:

$$\frac{\partial\varphi}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial\varphi}{\partial x_i'}\right) = 0, \; i = 1, 2, \dots, n. \tag{7.6}$$

Задача 3. Требуется найти систему функций $\{x_1(t),...,x_n(t)\}$, при которых показатель качества в виде функционала (7.5) принимает максимальное (или минимальное) значение и, кроме того, искомые функции должны удовлетворять заданным граничным условиям (7.4) и m дополнительным условиям, которые называют уравнениями связи:

$$\gamma_k\left[x_1(t),...,x_n(t);x_1'(t),...,x_n'(t)\right]=0, \tag{7.7}$$

$$k=1,...,m;\ \ m<n.$$

Типичным и наиболее важным примером дополнительных уравнений связи, возникающих в задачах динамической оптимизации, являются дифференциальные уравнения, описывающие управляемую систему:

$$\left.\begin{aligned} x_1' &= f_1(x_1,...,x_n),\\ &\dots\\ x_m' &= f_m(x_1,...,x_n).\end{aligned}\right\} \tag{7.8}$$

Они возникают при математическом моделировании реальной системы, когда в описании ее свойств участвуют производные от тех или иных величин. Это относится и к экономическим системам, т.к. при их анализе в динамике, как правило, рассматриваются такие факторы, как темпы развития или скорости изменения величин.

В качестве примера можно привести так называемую однопродуктовую модель динамики фондов [24]. Если при рассмотрении производственных процессов предположить, что валовые инвестиции (валовые капитальные вложения) I расходуются на прирост основных производственных фондов и на амортизационные отчисления A, то получим следующее соотношение [24]:

$$I=q\frac{dK}{dt}+A,$$

где K – величина основных производственных фондов;

q – коэффициент пропорциональности (параметр модели).

Амортизационные отчисления обычно определяются пропорционально капитальным вложениям

$$A=\mu K,$$

где μ – коэффициент амортизации. В таком случае полученное выражение принимает следующий вид:

$$\frac{dK}{dt}=\frac{1}{q}(-\mu K+I).$$

Оно является линейным дифференциальным уравнением первого порядка.

Следует отметить, что при моделировании динамики реальных систем могут получаться уравнения различного вида, в том числе дифференциальные. Для использования соответствующих математических методов анализа, оперирующих с моделями, как правило, необходимо полученные уравнения привести к стандартному виду.

Что касается дифференциальных уравнений, то обычно рассматривают две стандартные формы их представления:

– дифференциальное уравнение *n*-го порядка, описывающее поведение фактора $s(t)$, которое определяется действием внешней (входной) величины $x(t)$:

$$\frac{d^n s}{dt^n} = F\left[s, s', ..., s^{(n-1)}; x, x', ..., x^{(k-1)}; t\right],$$

в частности, линейное дифференциальное уравнение:

$$a_n s^{(n)} + a_{n-1} s^{(n-1)} + ... + a_1 s' + a_0 s = b_k x^{(k)} + ... + b_1 x' + b_0 x;$$

– система дифференциальных уравнений в нормальной форме:

$$\left.\begin{array}{l} s_1' = f_1(s_1, ..., s_n; x_1, ..., x_m; t), \\ \dots \\ s_n' = f_n(s_1, ..., s_n; x_1, ..., x_m; t). \end{array}\right\}$$

Если ввести в рассмотрение векторы $\bar{s}$ и $\bar{x}$, а также вектор (столбец) правых частей этих уравнений

$$\bar{f} = \begin{bmatrix} f_1 \\ ... \\ f_n \end{bmatrix},$$

то получим краткую векторно-матричную запись указанной системы:

$$\frac{d\bar{s}}{dt} = \bar{f}(\bar{s}, \bar{x}, t).$$

Заметим, что от модели в виде одного дифференциального уравнения *n*-го порядка можно перейти к системе из *n* дифференциальных уравнений первого порядка. Например, дано уравнение 2-го порядка

$$s'' + a_1 s' + a_0 s = b_0 x.$$

Введем новые переменные, полагая

$$s = s_1; \quad s' = s_2,$$

тогда вместо исходного ДУ 2-го порядка получим следующую систему из двух уравнений первого порядка:

$$\begin{cases} s_1' = s_2, \\ s_2' = -a_0 s_1 - a_1 s_2 + b_0 x. \end{cases}$$

В дискретных системах аналогом ДУ являются разностные уравнения. В них участвуют значения анализируемых процессов в различные тактовые моменты времени, например

$$a_0 s(t_i) + a_1 s(t_{i-1}) + a_2 s(t_{i-2}) = b_0 x(t_i),$$

где i-й текущий номер временного такта (этапа или шага).

ДУ (7.8) сводятся к стандартному виду (7.7), если принять

$$\gamma_k = x_k' - f_k.$$

Установлено [17, 31], что эту задачу с дополнительными ограничениями-равенствами можно свести к предыдущей задаче 2, если рассматривать новый вспомогательный функционал

$$\tilde{J} = \int_{t_{\text{н}}}^{t_{\text{к}}} \left[\varphi + \sum_{k=1}^{m} \lambda_k \gamma_k \right] dt = \int_{t_{\text{н}}}^{t_k} \tilde{\varphi} dt, \tag{7.9}$$

где $\lambda_k = \lambda_k(t)$ – вспомогательные функции, называемые множителями Лагранжа; γ_k – выражения, стоящие в левых частях ограничений-равенств (7.7).

Для отыскания искомого решения следует применить уравнение Эйлера (7.6) к функционалу (7.9), рассматривая получаемые при этом ДУ с граничными условиями (7.4) и уравнениями связи (7.7).

Задача 4. Иногда возникают задачи, в которых одно или несколько граничных значений $x_i(t_k)$, фигурирующих в задачах 1 – 3, не заданы и не связаны какими-либо условиями. Известно, что в таком случае для решения можно использовать изложенные выше условия и приемы, но каждое отсутствующее условие заменить так называемым условием трансверсальности [17, 31]

$$\left. \frac{\partial \varphi}{\partial x_i'} \right|_{t=t_k} = 0. \tag{7.10}$$

Задача 5. Рассмотрим задачи 1 – 2 в ситуации, когда граничные значения искомой функции $x_i(t)$ при $t = t_{\text{н}}$ и $t = t_{\text{к}}$ заданы, но один из концов промежутка $[t_{\text{н}}, t_{\text{к}}]$ обычно t_k не задан. В этом случае соотношения, определяющие искомое решение, дополняются следующим условием трансверсальности [17]:

$$\sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i'} x_i' - \varphi = 0 \quad (t = t_k). \tag{7.11}$$

Задача 6. В рассмотренных выше задачах необходимо найти функции, минимизирующие интегральный показатель качества (7.5) или (7.9). Эти задачи называют задачами Лагранжа. На практике возникают задачи и иного рода – когда требуется минимизировать некоторую функцию конечного состояния

$$J = \beta[\bar{x}(t_{к})]. \tag{7.12}$$

Такая задача называется задачей Майера [30, 31].

Задача 7. Обобщением задачи Лагранжа и Майера является задача Больца, в которой минимизируется показатель

$$J = \beta[\bar{x}(t_{к})] + \int_{t_{н}}^{t_{к}} \varphi[\bar{x}(t), \bar{x}'(t), t]\, dt. \tag{7.13}$$

Заметим, что в задачах Больца и Майера правый конец траектории $t = t_к$; $\bar{x} = \bar{x}(t_к)$ предполагаются незафиксированными.

В соответствии с [31] для решения такой задачи следует использовать соотношения (7.6), дополнив их следующими условиями трансверсальности:

$$\left.\frac{\partial\varphi}{\partial x_i'}\right|_{t=t_к} = -\frac{\partial\beta}{\partial x_i(t_к)},\ i = \overline{1,n}. \tag{7.14}$$

§ 7.3. Специфика задач оптимального динамического управления и использование ВИ для их решения

Рассмотрим изложенную в § 6.2 классическую задачу об оптимальном управлении динамической системой. В ней требуется найти законы изменения управляющих факторов (управлений) $u_i = u_i(t)$, при которых система из заданного начального состояния переходит в заданное конечное состояние

$$\bar{x}(t_н) = \bar{x}^н \to \bar{x}(t_к) = \bar{x}^к;\ t_н \le t \le t_к$$

и при этом показатель качества, выражаемый функционалом

$$Y = F(\bar{x}(t), \bar{u}(t)),$$

например

$$Y = \int_{t_н}^{t_к} f_0(x_1(t),...,x_n(t); u_1(t),...,u_m(t))\, dt \to \inf, \tag{7.15}$$

принимает наименьшее значение.

При этом следует учитывать наличие связи между переменными состояния $\bar{x}(t)$ и управлениями $\bar{u}(t)$, которая задается математической моделью системы в виде ДУ:

$$\left.\begin{aligned} &x_1' = f_1(x_1,...,x_n;u_1,...,u_m),\\ &\dots,\\ &x_n' = f_n(x_1,...,x_n;u_1,...,u_m) \end{aligned}\right\}$$

или $\bar{x}' = \bar{f}(\bar{x},\bar{u})$.

Сравнивая эту задачу с классическими задачами ВИ, рассмотренными в § 7.2, можно прийти к выводу, что эта задача близка к задаче 3 об отыскании системы функций при наличии ограничений.

В ней также рассматриваются функции $x_1(t),...,x_n(t)$, удовлетворяющие граничным условиям (7.4), и интегральный показатель качества (7.5). Участвующую в формулировке задачи модель легко свести к ограничениям-равенствам (7.7), фигурирующим в задаче ВИ, если ДУ переписать в виде

$$x_i' - f_i = 0$$

и принять $\gamma_i = x_i' - f_i$.

Однако в задаче управления есть существенная специфика: в функционале (7.5), рассматриваемом в ВИ, в подынтегральном выражении участвуют $(x_1,...,x_n;x_1',...,x_n')$, а в показателе качества (7.15) из задачи управления стоят $(x_1,...,x_n;u_1,...,u_m)$.

Для стыковки этих двух задач можно переменные управления $u_1,u_2,...,u_m$ рассматривать как дополнительные переменные x_{n+1},x_{n+2} и т.д. Если учесть, что в показателе качества производные от этих переменных отсутствуют, то можно получить следующие уравнения Эйлера –Лагранжа для рассматриваемой задачи [34].

$$\left.\begin{aligned} &\frac{\partial L}{\partial x_i} - \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial x_i'}\right) = 0,\ i=\overline{1,n},\\ &\frac{\partial L}{\partial u_j} = 0,\ j=\overline{1,m},\\ &L = f_0 + \sum_{i=1}^{n}\lambda_i(f_i - x_i'), \end{aligned}\right\} \tag{7.16}$$

где λ_i – множители Лагранжа.

Как отмечалось в § 6.2, в задачах управления наряду с моделью (7.8) могут задаваться ограничения-неравенства на переменные состояния и управления, например

$$a_j \le u_j \le b_j .$$

Такие ограничения средства ВИ также позволяют учесть. Для этого имеются специальные приемы, например метод штрафов и метод Валентайна [31], хотя при этом возникают дополнительные трудности в решении.

Пример 1 [31]. Рассмотрим систему, поведение которой характеризуется выходной переменной y. Система имеет одно управляющее воздействие u. Известна модель системы

$$y'' + y = u .$$

Требуется найти закон формирования $u = u(t)$, $t \ge 0$.

$$y(0) = y_0; \quad y'(0) = 0;$$

$$y(\infty) = y'(\infty) = 0 .$$

Кроме того, необходимо минимизировать показатель качества:

$$J = \int_0^\infty (y^2 + y'^2 + \alpha u^2)\, dt \to \min ,$$

где α – постоянный параметр.

Решение

Перейдем от исходной модели к системе уравнений первого порядка, полагая

$$y = x_1; \;\; y' = x_2 .$$

Тогда

$$\left.\begin{array}{l} x_1' = x_2, \\ x_2' = -x_1 + u \end{array}\right\} \text{ или } \left.\begin{array}{l} x_1' - x_2 = \gamma_1 = 0, \\ x_2' + x_1 - u = \gamma_2 = 0, \end{array}\right\}$$

$$J = \int_0^\infty (x_1^2 + x_2^2 + \alpha u^2)\, dt .$$

Запишем подынтегральное выражение в форме (7.9), учитывающей ограничения-равенства:

$$\tilde{L} = x_1^2 + x_2^2 + \alpha u^2 + \lambda_1 (x_1' - x_2) + \lambda_2 (x_2' + x_1 - u) + \lambda_3 (x_3' - u) .$$

Рассмотрим производные от $\tilde{L}$, которые войдут в уравнение Эйлера – Лагранжа (7.16):

$$\frac{\partial \tilde{L}}{\partial x_1} = 2x_1 + \lambda_2; \quad \frac{\partial \tilde{L}}{\partial x_2} = 2x_2 - \lambda_1;$$

$$\frac{\partial \tilde{L}}{\partial x_1'} = \lambda_1; \quad \frac{\partial \tilde{L}}{\partial x_2'} = \lambda_2; \quad \frac{\partial \tilde{L}}{\partial u} = 2\alpha u - \lambda_2 .$$

Запишем уравнения (7.16)

$$\frac{\partial \tilde{L}}{\partial x_i} - \left(\frac{\partial \tilde{L}}{\partial x_i'} \right)' = 0 \text{, и } \frac{\partial \tilde{L}}{\partial u} = 0;$$

$$(2x_1 + \lambda_2) - \lambda_1' = 0;$$

$$(2x_2 - \lambda_1) - \lambda_2' = 0;$$

$$2\alpha u - \lambda_2 = 0 .$$

В результате получаем следующую систему уравнений:

$$\left. \begin{aligned} & x_1' = x_2, \\ & x_2' = -x_1 + u, \\ & \lambda_1' = 2x_1 + \lambda_2, \\ & \lambda_2' = 2x_2 - \lambda_1, \\ & u = \frac{1}{2\alpha}\lambda_2. \end{aligned} \right\}$$

В матричном виде ее можно записать следующим образом:

$$\begin{bmatrix} x_1' \\ x_2' \\ \lambda_1' \\ \lambda_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & \frac{1}{2\alpha} \\ 2 & 0 & 0 & 1 \\ 0 & 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \lambda_1 \\ \lambda_2 \end{bmatrix}.$$

Для записи решения этой системы рассмотрим характеристическое уравнение

$$\det(\lambda I - A) = 0 ,$$

$$\begin{vmatrix} \lambda & -1 & 0 & 0 \\ 1 & \lambda & 0 & -\frac{1}{2\alpha} \\ -2 & 0 & \lambda & -1 \\ 0 & -2 & 1 & \lambda \end{vmatrix} = 0 .$$

Раскрывая этот определитель, получим

$$\lambda^4 + \left(2 - \frac{1}{\alpha}\right)\lambda^2 + \left(1 + \frac{1}{\alpha}\right) = 0.$$

Положим для определенности $\alpha = 1/9$, тогда уравнение примет вид

$$\lambda^4 - 7\lambda^2 + 10 = 0.$$

Обозначая $\lambda^2 = y$ и решая квадратное уравнение

$$y^2 - 7y + 10 = 0,$$

получаем

$$y_1 = 5; \; y_2 = 2.$$

$$\lambda_{1,2} = \pm\sqrt{5}; \; \lambda_{3,4} = \pm\sqrt{2}.$$

Общее решение для x_1 можно записать следующим образом:

$$x_1(t) = c_1 e^{\sqrt{5}t} + c_2 e^{-\sqrt{5}t} + c_3 e^{\sqrt{2}t} + c_4 e^{-\sqrt{2}t}.$$

Константы c_i необходимо найти из граничных условий.

Заметим, что условие $\bar{x}(\infty) = 0$ может выполняться лишь при $c_1 = c_3 = 0$.

Соответственно

$$x_1 = c_2 e^{-\sqrt{5}t} + c_4 e^{-\sqrt{2}t}.$$

При $t = 0$ $x_1 = x^0$; $x_2 = 0$, поэтому

$$\left.\begin{array}{l} c_2 + c_4 = x^0, \\ -\sqrt{5}c_2 - \sqrt{2}c_4 = 0, \end{array}\right\}$$

Отсюда получаем

$$c_2 = \frac{\sqrt{2}}{\sqrt{2} - \sqrt{5}} x^0; \; c_4 = \frac{\sqrt{5}}{\sqrt{5} - \sqrt{2}} x^0.$$

Таким образом, оптимальный управляемый процесс будет определяться следующим образом:

$$y^* = x_1^* = \frac{x^0}{\sqrt{5} - \sqrt{2}} \left(\sqrt{5}e^{-\sqrt{2}t} - \sqrt{2}e^{-\sqrt{5}t}\right).$$

Оптимальное управление можно найти из соотношения

$$x_2' = -x_1 + u,$$

т.е.

$$u^* = x_1 + x_2' = x_1 + x_1''.$$

Дифференцируя найденное выражение для $x_1 = x_1^*$, получаем

$$u^* = \frac{3x^0}{\sqrt{5}-\sqrt{2}}\left(\sqrt{5}e^{-\sqrt{2}t} - 2\sqrt{2}e^{-\sqrt{5}t}\right).$$

Пример 2. Обобщением рассмотренной задачи является задача об управлении системой n-го порядка, модель которой представлена следующим образом:

$$\left.\begin{array}{l} x_1' = a_{11}x_1 + \ldots + a_{1n}x_n + b_1 u \\ \ldots \\ x_n' = a_{n1}x_1 + \ldots + a_{nn}x_n + b_n u \end{array}\right\} \qquad (7.17)$$

или

$$\bar{x}' = A\bar{x} + \bar{b}u,$$

где $\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$; $\bar{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$; $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$. $\qquad (7.17')$

Необходимо найти управление $u = u^*(t)$, при котором система переходила бы из начального состояния $\bar{x}(t_{\text{Н}}) = \bar{x}^{\text{Н}}$ в заданное конечное состояние $\bar{x}(t_{\text{К}}) = \bar{x}^{\text{К}}$ $(t_н \le t \le t_к)$, а обобщенный интегральный квадратичный показатель качества

$$J = \int_{t_{\text{Н}}}^{t_{\text{К}}} (c_1 x_1^2 + \ldots + c_n x_n^2 + \alpha u^2)\, dt \to \min$$

принимал минимальное значение.

Решение этой задачи с использованием средств ВИ см., например, в [31].

§ 7.4. Приближенные методы решения задач динамической оптимизации средствами ВИ

При использовании необходимых условий оптимальности в виде уравнений Эйлера – Лагранжа для решения реальных практических задач редко удается получить искомый закон формирования управлений в аналитическом виде так, как это удалось сделать выше для простейших примеров.

В подобных случаях приходится использовать приближенные численные методы. При этом возникает ситуация, сходная с той, которая рассматривалась в гл. 5 в связи с отысканием экстремума функции: можно использовать необходимое условие экстремума, решая систему уравнений $\frac{\partial f}{\partial x_i} = 0$ каким-либо приближенным методом, а можно использовать процедуру поиска экстремума, рассматривая непосредственно (напрямую) заданную функцию *f(x)*. На практике чаще прибегают к использованию прямых методов решения.

Рассмотрим сущность наиболее известных подходов подобного рода на примере простейшей задачи 1, рассмотренной в § 7.2.

Требуется найти функцию $x = x(t),\ t \in [t_\text{н}, t_\text{к}]$, для которой

$$x(t_\text{н}) = x^\text{н}; \quad x(t_\text{к}) = x^\text{к} \text{ – заданы,}$$

а функционал

$$J = \int_{t_\text{н}}^{t_\text{к}} \varphi[x(t), x'(t), t]dt \to \min$$

принимает минимальное значение.

Метод Ритца

Введем в рассмотрение некоторую функцию $\alpha_0(t)$, удовлетворяющую заданным граничным условиям:

$$\alpha_0(t_\text{н}) = x^\text{н}; \qquad \alpha_0(t_\text{к}) = x^\text{к},$$

а также систему линейно-независимых функций $\{\alpha_1(t), \alpha_2(t), ... \alpha_m(t)\}$, удовлетворяющих условиям

$$\alpha_i(t_\text{н}) = \alpha_i(t_\text{к}) = 0.$$

Искомое решение будем искать в виде

$$\tilde{x}(t) = \alpha_0(t) + \sum_{i=1}^{m} c_i \alpha_i(t), \tag{7.18}$$

где c_i – подлежащие определению коэффициенты.

Подставив эту функцию в функционал (7.2) и выполнив интегрирование, получим выражение, зависящее от неизвестных величин c_i:

$$J = J(c_1, ..., c_m). \tag{7.19}$$

Теперь задача свелась к отысканию численных значений $c_i = c_i^*$, при которых функция (7.19) будет минимальна.

Заметим, что представление решения в форме (7.18) дало возможность свести задачу о минимизации функционала к задаче о минимизации функции.

Для ее решения можно использовать необходимое условие:

$$\frac{\partial J}{\partial c_i} = 0, \quad i = \overline{1, m}.$$

Если решение этой системы уравнений вызывает затруднения, то можно использовать рассмотренные в § 5.3 – 5.4 численные методы поиска экстремума.

Пример. Рассмотрим задачу об отыскании функции $x = x(t)$, для которой

$$0 \le t \le 1; \qquad x(0) = 0; \qquad x(1) = 1;$$

$$J = \int_0^1 \left(x^2(t) + x'^2(t)\right) dt \to \min.$$

В § 7.2 было найдено решение этой задачи с помощью уравнения Эйлера – Лагранжа.

Выберем в качестве базисных функций $\alpha_i(t)$ многочлены

$$\alpha_0(t) = t, \qquad \text{при этом } \alpha_0(0) = 0; \quad \alpha_0(1) = 1,$$

$$\alpha_1(t) = t(t-1), \qquad \alpha_2(t) = t^2(t-1), \ldots, \ \alpha_n(t) = t^n(t-1).$$

Ограничимся аппроксимацией 2-го порядка, полагая n=1. Тогда

$$x \approx \tilde{x}(t) = t + c_1 t(t-1) = ct^2 + (1-c)t.$$

Заметим, что для любого c граничные условия выполняются.

Рассмотрим компоненты, входящие в функционал:

$$x^2 = c^2 t^4 - 2c(c-1)t^3 + t^2(1-c)^2;$$

$$x' = 2ct + (1-c);$$

$$x'^2 = 4c^2 t^2 + (1-c)^2 + 4c(1-c)t.$$

Тогда

$$J \approx \tilde{J} = \int_0^1 \left[c^2 t^4 - 2c(c-1)t^3 + t^2\left((1-c)^2 + 4c^2\right) + t4c(1-c) + (1-c)^2 \right] dt.$$

После интегрирования и подстановки пределов получаем

$$J =| \tilde{J} = \frac{11}{30}c^2 - \frac{1}{6}c + \frac{4}{3}.$$

Используя условие

$$\frac{d\tilde{J}}{dc} = \frac{11}{15}c - \frac{1}{6} = 0,$$

находим

$$c = c^* = \frac{5}{22}.$$

В результате получаем искомое приближенное решение

$$\tilde{x}^* = t + \frac{5}{22}t(t-1).$$

Заметим, что точное решение, найденное в § 7.2, определяется выражением

$$x^* = \frac{1}{e - e^{-1}}(e^t - e^{-t}).$$

Для оценки полученного результата в табл. 7.1 приведены значения x^* и $\tilde{x}^*$ для нескольких значений t_i.

Таблица 7.1

t	Точные значения, x^*	Приближенные значения, $\tilde{x}^*$	
		Метод Ритца	Метод Эйлера
0	0	0	0
0,2	0,171	0,164	0,171
0,4	0,350	0,345	0,350
0,6	0,542	0,545	0,542
0,8	0,756	0,764	0,766
1	1	1	1

Метод Эйлера

Он основан на разбиении рассматриваемого промежутка времени на небольшие части $\Delta t = h$ и представлении искомой функции $x = x(t)$ отдельными дискретами:

$$x(t_0) = x^{\text{н}}, x_1 = x(t_0 + h), \ldots, x_i = (t_0 + ih), \ldots, x_N = x(t_0 + Nh) = x^k,$$

где $N = (t_{\text{к}} - t_{\text{н}})/h$.

При этом производные приближенно заменяются конечными разностями

$$x'(t_i) \cong \frac{\Delta x}{\Delta t} = \frac{x_{i+1} - x_i}{h},$$

а функционал (7.2) приближенно вычисляется по формуле прямоугольников

$$J = \int_{t_н}^{t_к} \varphi\, dt \approx \Delta t\left[\varphi_0 + \varphi_1 + \ldots + \varphi_{N-1}\right] \tag{7.20}$$

В результате получается выражение, зависящее от неизвестных значений искомой функции $x_1, x_2, \ldots x_{N-1}$. Задача сводится к отысканию значений $x_i = x_i^*$, при которых функция

$$J = J(x_1, \ldots, x_{N-1}) \to \min$$

принимает наименьшее значение.

Эту задачу, как и в предыдущем случае, можно решить аналитически, используя необходимые условия экстремума, или с помощью численных методов поиска экстремума.

Пример. Рассмотрим ту же, что и в предыдущем случае, задачу. Промежуток $0 \le t \le 1$ разобьем на 5 частей с шагом $h = \Delta t = 0{,}2$. В результате искомая функция будет представлена дискретными отсчетами

$$\{x_0, x_1, x_2, x_3, x_4, x_5\},$$

причем $x_0 = x(0) = 0$ и $x_5 = x(1) = 1$, а числа $x_1, \ldots, x_4$ необходимо определить так, чтобы функционал (7.20) достиг минимума.

Для рассматриваемого случая запишем следующие соотношения:

$$J \approx h\left[\varphi_0 + \varphi_1 + \varphi_2 + \varphi_3 + \varphi_4\right], \qquad \varphi_i = x_i^2 + x'^2_i = x_i^2 + \frac{(x_{i+1} - x_i)^2}{h^2};$$

$$J \approx h(x_0^2 + x_1^2 + x_2^2 + x_3^2 + x_4^2) + \frac{1}{h}\Big[(x_1 - x_0)^2 + (x_2 - x_1)^2 + (x_3 - x_2)^2 +$$

$$+ (x_4 - x_3)^2 + (x_5 - x_4)^2\Big] = h(x_0^2 + x_1^2 + x_2^2 + x_3^2 + x_4^2) +$$

$$+ \frac{1}{h}\left(x_0^2 + 2(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2)\right) - \frac{2}{h}\left(x_1 x_0 + x_2 x_1 + x_3 x_2 + x_4 x_3 + x_5 x_4\right).$$

После приведения подобных членов с учетом $x_0 = 0$ и $x_5 = 1$ получаем

$$J = \frac{1}{h} + \left(h + \frac{2}{h}\right)\left(x_1^2 + x_2^2 + x_3^2 + x_4^2\right) - \frac{2}{h}\left(x_2 x_1 + x_3 x_2 + x_4 x_3 + 1 * x_4\right). \tag{7.21}$$

Рассмотрим условия экстремальности J по x_i:

$$\left.\begin{aligned}
\frac{\partial J}{\partial x_1} &= \left(h+\frac{2}{h}\right)2x_1 - \frac{2}{h}x_2 = 0,\\
\frac{\partial J}{\partial x_2} &= \left(h+\frac{2}{h}\right)2x_2 - \frac{2}{h}(x_1+x_3) = 0,\\
\frac{\partial J}{\partial x_3} &= \left(h+\frac{2}{h}\right)2x_3 - \frac{2}{h}(x_2+x_4) = 0,\\
\frac{\partial J}{\partial x_4} &= \left(h+\frac{2}{h}\right)2x_4 - \frac{2}{h}(x_3+1) = 0.
\end{aligned}\right\}$$

После преобразований эта система принимает следующий вид:

$$\left.\begin{aligned}
ax_1 &= x_2,\\
ax_2 &= x_1 + x_3,\\
ax_3 &= x_2 + x_4,\\
ax_4 &= x_3 + x_1,
\end{aligned}\right\},$$

где $a = h^2 + 2 = 2{,}04$.

Последовательно исключая переменные, можно получить следующие выражения, дающие значения оптимального решения в выбранных точках t_i:

$$x_1^* = 1/\Delta; \qquad x_2^* = a/\Delta; \qquad x_3^* = (a^2-1)/\Delta; \qquad x_4^* = a(a^2-2)/\Delta,$$

где $\Delta = a^4 - 3a^2 + 1$.

После подстановки в эти формулы значения $a = 2{,}04$ получаются отсчеты приближенного решения x_i, приведенные в табл. 7.1.

Заметим, что в данном случае минимизацию функционала (7.21) удалось осуществить аналитически. В более сложных случаях для этого можно использовать численные алгоритмы поиска экстремума, изложенные в § 5.3 – 5.4.

Рассмотренные прямые методы Ритца и Эйлера можно применить и для решения более сложных задач ВИ и динамической оптимизации, используя специальные приемы, позволяющие учесть ограничения, фигурирующие в условиях задачи.

Проиллюстрируем это на примере задачи управления с минимизацией затрат энергии (более подробно эта задача будет рассмотрена в § 8.3). В этом случае необходимо найти управляющее воздействие $u=u(t)$, $0 \le t \le t_к$, при котором показатель качества

$$J = \int_0^{t_к} u^2(t)dt \to \min$$

принимает минимальное значение и выполняются условия связи

$$\left.\begin{array}{l} y' = v, \\ v' = ku, \end{array}\right\}$$

а также граничные условия

$$y(0) = y_0;\ \ y'(0) = \text{v}_0;\ \ y(t_к) = y^k;\ y'(t_к) = \text{v}^к \text{ – заданы.}$$

Для использования метода Эйлера осуществляем следующие действия:

- разобьем промежуток времени $[0;t_к]$ на N частей с шагом $\Delta t = h = t_к/N$;
- будем рассматривать переменные u,y,v в полученные дискретные моменты времени: $y_i=y(t_i)$; $v_i=v(t_i)$; $u_i=u(t_i)$; $i=0,1,..N$;
- введем вспомогательный функционал, используя идею метода штрафов для учета уравнений модели

$$\tilde{J} = \int_0^{t_k} \left\{ u^2(t) + M\left[(y'-v)^2 + (v'-ku)^2 \right] \right\} dt, \qquad (7.22)$$

где M – некоторое достаточно большое число;

- заменим этот функционал (7.22) его дискретным аналогом, используя метод прямоугольников для интегрирования и конечные разности для дифференцирования:

$$\tilde{J} \approx h \sum_{i=0}^{N-1} \left\{ u_i^2 + M\left[\frac{y_{i+1} - y_i}{h} - v_i \right]^2 + M\left[\frac{v_{i-1} - v_i}{h} - ku_i \right]^2 \right\}, \qquad (7.23)$$

где y_0, v_0 и y_N, v_N – заданы, а дискреты $\{u_0, u_1..u_n\}$, $\{y_1,..y_{N-1}\}$, $\{v_2,..v_{N-1}\}$ подлежат определению;

- для полученного выражения (7.23) применим одну из процедур поиска минимума по переменным u_i, y_i, v_i.

Следует отметить, что хотя все изложенные выше действия являются вполне обоснованными, тем не менее, как показывает опыт их применения, при их практической реализации возникают серьезные трудности [35]:

медленная сходимость, ненадежность результатов и др. Эти трудности связаны, в частности, с тем, что минимизируемые функции, например (7.23), обычно имеют «овражный» характер, обладают несколькими локальными экстремумами, в каждом из которых процедура поиска может остановиться и др.

Контрольные вопросы

1. Какой класс задач рассматривается в вариационном исчислении?
2. Как определяются понятия «вариация», «функционал» и «экстремум функционала»?
3. Приведите примеры функционалов.
4. Как формулируется классическая задача ВИ и как записывается необходимое условие экстремума функционала (уравнение Эйлера-Лагранжа):
   – для одной искомой функции,
   – для нескольких искомых функций.
5. Как формулируется классическая задача ВИ при наличии дополнительных уравнений связи и каким образом можно их учесть при использовании уравнений Эйлера-Лагранжа?
6. В каких случаях используется условие трасверсальности и как оно записывается?
7. Как формулируются задачи Лагранжа, Майера и Больца, в чем их сходство и отличие?
8. В чем состоит специфика задач динамического оптимального управления и каким образом эти задачи можно свести к классическим задачам вариационного исчисления?
9. В каких случаях приходится использовать приближенные методы при решении задач ВИ?
10. В чем состоит сущность метода Ритца для решения задач ВИ?
11. В чем состоит сущность метода Эйлера при решении задач ВИ?

Примечание. Задачи для самостоятельного решения по данной теме приведены после 9-й главы.

Глава 8. **ПРИНЦИП МАКСИМУМА И ЕГО ПРИМЕНЕНИЕ ДЛЯ СИНТЕЗА ОПТИМАЛЬНЫХ УПРАВЛЕНИЙ В НЕПРЕРЫВНЫХ СИСТЕМАХ**

§ 8.1. Формулировка принципа максимума для непрерывных систем

Рассмотрим задачу о динамическом управлении непрерывной системой, математическая модель которой сведена к стандартной форме

$$\left.\begin{array}{l} x_1' = f_1(x_1,...,x_n;\, u_1,...,u_m;t), \\ x_2' = f_2(x_1,...,x_n;\, u_1,...,u_m;t), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n' = f_n(x_1,...,x_n;u_1,...,u_m;t). \end{array}\right\} \tag{8.1}$$

или

$$\bar{x}' = \bar{f}(\bar{x},\bar{u},t), \tag{8.1 а}$$

где $\bar{f} = \left[f_1, f_2,..., f_n\right]^{\mathrm{T}}$ – столбец правых частей уравнений; n – порядок системы; $x_i = x_i(t)$ – переменные состояния; $u_j = u_j(t)$ – управляющие воздействия.

Необходимо найти управление $\bar{u} = \bar{u}(t)$ на промежутке времени $t_{\text{н}} \le t \le t_{\text{к}}$ так, чтобы выполнялись условия:

1) система была переведена из заданного начального состояния $\bar{x}(t_{\text{н}}) = \bar{x}_{\text{н}}$ в такое состояние в конечный момент времени $t = t_{\text{к}}$, чтобы первые l переменных x_i совпали с заданными значениями $x_i(t_{\text{к}}) = x_i^{\text{к}}$, $i = \overline{1,l}$; $l \le n$;
2) заданный показатель качества управления в виде функционала

$$J = \int\limits_{t_{\text{н}}}^{t_{\text{к}}} f_0(\bar{x},\bar{u},t)dt \to \min \tag{8.2}$$

достигал минимума;

3) выполнялись все ограничения на управление $\bar{u}$ (и возможно на переменные состояния $\bar{x}$), например

$$a_i \le u_i \le b_i. \tag{8.3}$$

Заметим, что для части переменных состояния ($x_{l+1},..., x_n$) их значения в конечный момент времени не заданы и могут быть любыми. Если принять $l = n$, то переменных со свободными значениями в момент времени $t = t_к$ нет.

Будем предполагать, что функции $f_0, f_1, ..., f_n$ являются непрерывными по *u*, непрерывно дифференцируемыми по *x* и кусочно-непрерывными по *t* [17].

Введем вспомогательные функции $\psi_i(t), i = 0,1,...,n$. С их помощью сформируем следующее выражение:

$$H = H(\bar{x}, \bar{u}, \bar{\psi}, t) = \psi_0 f_0 + \psi_1 f_1 + ... + \psi_n f_n. \tag{8.4}$$

Оно определяет так называемую функцию Гамильтона.

Теорема (принцип максимума Понтрягина) [31]. Если $\bar{u} = \bar{u}^*(t)$ – оптимальное управление, а $\bar{x} = \bar{x}^*(t)$ – соответствующие оптимальные переменные состояния, определяемые (8.1), то существуют ненулевые непрерывные функции $\psi_i(t)$, при которых:

1) функция Гамильтона достигает максимума, т.е.

$$H(\bar{x}^*, \bar{u}^*, \bar{\psi}, t) \geq H(\bar{x}, \bar{u}, \bar{\psi}, t); \tag{8.5}$$

2) функции ψ_i удовлетворяют уравнениям

$$\frac{d\psi_i}{dt} = -\frac{\partial H}{\partial x_i}, \; i = \overline{1,n}; \tag{8.6}$$

3) $\psi_0 = -1$; $\psi_i(t_к) = 0$ при $i = l+1,...,n$. (8.7)

Эта теорема определяет необходимые, но в общем случае недостаточные условия оптимальности управления.

Следует отметить, что принцип максимума (ПМ) обладает важными отличительными особенностями по сравнению с базовыми результатами вариационного исчисления:

1) он сформулирован непосредственно для задач управления, в его структуре фигурируют переменные состояния $\bar{x}$, управления $\bar{u}$, связывающая их модель (8.1) и ограничения на них, например (8.3);
2) формулировка имеет довольно удобный и универсальный характер;
3) он охватывает известные в ВИ необходимые условия экстремальности функционала;

4) ограничения на свойства управляемой системы (правые части модели $f_1,...,f_n$) и критерий качества f_0, предполагаемые в ПМ, имеют гораздо менее жесткий характер, чем в ВИ, в частности, оптимальное управление может иметь разрывный, скачкообразный характер, типичный для многих прикладных задач;
5) для важного частного случая, когда управляемая система описывается линейной моделью, ПМ определяет необходимые и достаточные условия оптимальности управления [31].

Если в конечный момент времени $t = t_{\text{к}}$ состояние системы не зафиксировано или задано частично только для нескольких переменных:

$$x_i(t_{\text{к}}) = x_i^{\text{к}}, \quad i = \overline{1,l}, \quad l < n, \tag{8.8}$$

то можно рассматривать более общую форму показателя качества, чем (8.2):

$$J = F(\overline{x}(t_{\text{к}})) + \int_{t_{\text{н}}}^{t_{\text{к}}} f_0(\overline{x},\overline{u},t)\,dt \to \min \qquad \text{(задача Больца).} \tag{8.9}$$

Можно показать (см., например, [24, 30]), что в этом случае определяемые принципом максимума условия остаются теми же, но дополняются условиями трансверсальности:

$$\psi_i(t_{\text{к}}) = -\left.\frac{\partial F(\overline{x})}{\partial x_i}\right|_{\overline{x}=\overline{x}^*}, \quad i = l+1, l+2, ..., n, \tag{8.10}$$

которые записываются для индексов i, соответствующих незафиксированным переменным.

В том случае, когда конечный момент времени $t_{\text{к}}$ заранее не определен (задача с нефиксированным временем), для решения необходимо ввести дополнительное условие [31]:

$$\left. H(\overline{x}^*, \overline{u}^*, \overline{\psi}, t)\right|_{t=t_{\text{к}}} = 0, \tag{8.11}$$

т.е. функция Гамильтона, соответствующая оптимальному управлению, должна равняться нулю в конечный момент времени $t_{\text{к}}$.

Важным примером задачи с нефиксированным временем является задача об оптимальном быстродействии (см. § 8.4).

Проиллюстрируем методику практического использования ПМ на следующих примерах (см. § 8.2 – 8.6).

§ 8.2. Классическая задача Эйлера

Рассмотрим одну из задач, которые изучались основоположниками вариационного исчисления [30].

Задача. Требуется найти функции $u = u(t)$, $x = x(t)$, $0 \le t \le T$, для которых выполняются условия

$$x' = u, \tag{8.12}$$

$$x(0) = 0,\ x(T) = a \tag{8.13}$$

и минимизируется функционал

$$J = \int_0^T \left(x^2 + u^2 \right) dt \to \min . \tag{8.14}$$

Продемонстрируем на примере этой задачи технику применения принципа максимума.

Решение. Для заданного ДУ первого порядка (8.12) и критерия качества (8.14) запишем функцию Гамильтона (8.4):

$$H = \psi_0 \left(x^2 + u^2 \right) + \psi_1 u .$$

Поскольку ограничений на значения u нет, то наибольшее значение H может быть только экстремумом (максимумом). Используя необходимое условие экстремума, получаем:

$$\frac{\partial H}{\partial u} = 2\psi_0 u + \psi_1 = 0 .$$

Учитывая, что $\psi_0 = -1$, получаем:

$$u^* = \frac{1}{2}\psi_1(t) . \tag{8.15}$$

Рассмотрим уравнение (8.6) для вспомогательной функции ψ_1, а также уравнение (8.12), в котором u определяется полученным выражением (8.15):

$$\left. \begin{aligned} \frac{d\psi_1}{dt} &= -\frac{\partial H}{\partial x} = 2x, \\ \frac{dx}{dt} &= \frac{1}{2}\psi_1. \end{aligned} \right\}$$

Получили систему взаимосвязанных (сопряженных) уравнений. Продифференцируем второе уравнение, выразим $\psi_1' = 2x''$ и подставим этот результат в первое уравнение:

$$2x'' = 2x \text{ или } x'' - x = 0 .$$

Для полученного однородного ДУ следует записать характеристическое уравнение $\lambda^2 - 1 = 0$. По его корням $\lambda_{1,2} = \pm 1$ можно сформировать общее решение этого ДУ:

$$x(t) = C_1 e^t + C_2 e^{-t} . \tag{8.16}$$

Произвольные константы в (8.16) следует найти из граничных условий (8.13):

$$\left. \begin{aligned} x(0) &= C_1 + C_2 = 0, \\ x(T) &= C_1 e^T + C_2 e^{-T} = a. \end{aligned} \right\}$$

Отсюда получаем

$$C_1 = \frac{a}{e^T - e^{-T}} ; \; C_2 = -C_1 .$$

Тогда оптимальный процесс и оптимальное управление будут определяться выражениями

$$x^* = \frac{a}{e^T - e^{-T}} \left(e^t - e^{-t} \right), \qquad u^* = x' = \frac{a}{e^T - e^{-T}} \left(e^t + e^{-t} \right).$$

§ 8.3. Задача оптимального управления с минимизацией затрат энергии на управление

Рассмотрим систему, поведение которой характеризуется фактором y. На него можно влиять с помощью воздействия (управления) u. Модель, определяющая связь между ними, задана дифференциальным уравнением

$$y'' = k u, \quad u = u(t), \; y = y(t),$$

где k – параметр модели.

В начальный момент времени $y(0) = y^0 ; \; y'(0) = y_1^0 .$

Требуется сформировать управление $u = u(t)$ $(0 \le t \le t_к)$ так, чтобы:

1) по окончании процесса управления, т.е. при $t = t_к$, система оказалась в заданном состоянии: $y(t_к) = 0; \;\; y'(t_к) = 0$;

2) показатель качества в виде функционала $J = \int_0^{t_к} f_0(y, y', u)\, dt$ принял минимальное значение.

Эта задача имеет простую и наглядную физическую интерпретацию (рис. 8.1).

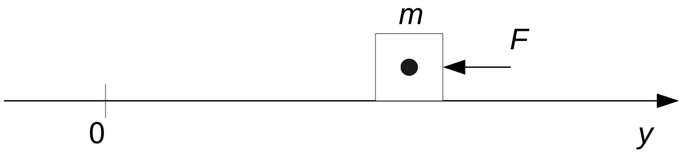


Рис. 8.1

Материальный объект, имеющий массу m, в начальный момент времени находится в точке $y = y(0)$ и имеет начальную скорость $v(0) = y_0'$. Необходимо найти закон изменения силы $F = F(t)$, прикладываемой к этому объекту, при котором он оказался бы в точке 0 и оставался в ней, а для этого он должен прийти в эту точку с нулевой скоростью.

При некоторых упрощающих предположениях эта система описывается уравнением, соответствующим второму закону Ньютона в механике:

$$ma = F\,.$$

Учитывая, что $y' = v$ – скорость, а $y'' = v' = a$ – ускорение, получаем:

$$y'' = k\,u\,, \tag{8.17}$$

где $k = 1/m$; $u = F$.

Приведем модель системы (8.17) к стандартному виду (8.1), полагая $y = x_1$, $y' = x_2$:

$$\left.\begin{aligned} x_1' &= x_2, \\ x_2' &= k\,u. \end{aligned}\right\} \tag{8.18}$$

Рассмотрим эту задачу, оценивая качество процессов управления величиной затрат энергии. Можно показать [31, 34], что эти затраты определяются функционалом

$$J = \int_0^{t_к} u^2(t)dt \to \min\,. \tag{8.19}$$

Нетрудно убедиться в том, что решение задачи об оптимальном управлении не изменится, если перед интегралом (8.19) поставить какой-либо числовой множитель. Будем считать, что в качестве такого множителя принят коэффициент, равный $\frac{1}{2}$.

Тогда показатель качества будет соответствовать общей форме (8.2) при $f_0 = \frac{1}{2}u^2(t)$. Учитывая то, что управляемая система имеет второй порядок и описывается уравнениями (8.18), вводим вспомогательные функции ψ_0, ψ_1, ψ_2 и формируем функцию Гамильтона (8.4):

$$H = \psi_0 f_0 + \psi_1 f_1 + \psi_2 f_2 = \frac{1}{2}\psi_0 u^2 + \psi_1 x_2 + \psi_2 k u . \tag{8.20}$$

Записываем уравнения (8.6) для функций ψ_i:

$$\left.\begin{aligned} \frac{d\psi_1}{dt} &= -\frac{\partial H}{\partial x_1} = 0 \\ \frac{d\psi_2}{dt} &= -\frac{\partial H}{\partial x_2} = -\psi_1. \end{aligned}\right\} \tag{8.21}$$

Нетрудно найти решения этой системы:

$$\left.\begin{aligned} \psi_1 &= C_1, \\ \psi_2 &= -C_1 t + C_2, \end{aligned}\right\} \tag{8.22}$$

где $C_{1,2} = \text{const}$.

В соответствии с ПМ функция H должна быть максимальна при оптимальном управлении, поэтому $dH/du = 0$.

Дифференцируя (8.20) и учитывая, что $\psi_0 = -1$, получаем:

$$-u + k\psi_2 = 0, \quad u = u^* = k\psi_2(t).$$

Используя (8.22), находим форму оптимального управления:

$$u^* = k(C_2 - C_1 t). \tag{8.23}$$

Таким образом, с помощью ПМ нам удалось установить важный результат: для того чтобы осуществить оптимальное по затратам энергии управление, нужно изменять управляющее воздействие на заданном промежутке времени по линейному закону (рис. 8.2).

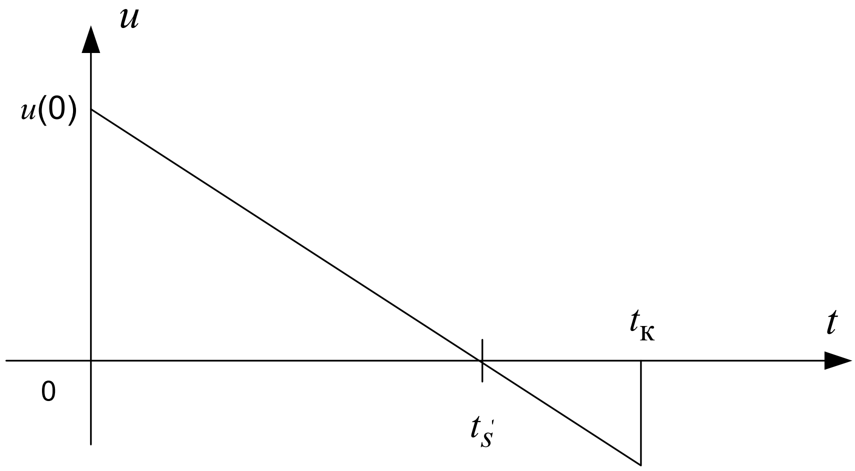


Рис. 8.2

На промежутке времени $[0,\ t_s]$ осуществляется «разгон» объекта, хотя интенсивность воздействия постепенно уменьшается, а на промежутке $\left[t_s,\ t_к\right]$ – «торможение»; при $t \ge t_к$ должно быть $u = 0$. Для того чтобы полностью определить управление, необходимо найти пока неизвестные константы C_1, C_2.

Из уравнения модели (8.18) для синтезированного управления (8.23) получаем:

$$x_2' = k^2(C_2 - C_1 t), \qquad x_2 = k^2\left(C_2 t - C_1\frac{t^2}{2}\right) + C_3,$$

$$x_1' = C_3 + k^2 C_2 t - \frac{k^2 C_1}{2}t^2, \quad x_1 = C_3 t + \frac{k^2 C_2}{2}t^2 - \frac{k^2 C_1}{6}t^3 + C_4.$$

В полученные выражения подставляем граничные условия:
при $t = 0$

$$\left.\begin{aligned} x_1 &= C_4 = y_0, \\ x_2 &= C_3 = y_0'; \end{aligned}\right\}$$

при $t = t_к = T$

$$\left.\begin{aligned} x_1 &= 0 = y_0' T + \frac{k^2 T^2}{2}C_2 - \frac{k^2 T^3}{6}C_1 + y_0, \\ x_2 &= 0 = y_0' + k^2 T C_2 - \frac{k^2 T^2}{2}C_1. \end{aligned}\right\} \tag{8.24}$$

Получили систему двух линейных алгебраических уравнений относительно двух неизвестных C_1 и C_2. Решив ее и определив соответствующие значения C_1^* и C_2^*, полностью определяем программу изменения оптимального управления во времени.

Рассмотрим далее решение задачи при следующих условиях:

$$\left.\begin{aligned} x_1(0) &= x_1^{\text{н}}, \\ x_2(0) &= 0; \end{aligned}\right\} \left.\begin{aligned} x_1(T) &= 0, \\ x_2(T) &= 0; \end{aligned}\right\} k = 1.$$

Тогда вместо (8.24) получим следующие уравнения:

$$\left.\begin{aligned} &C_2\frac{T^2}{2} - C_1\frac{T^3}{6} + x_1^{\text{н}} = 0, \\ &C_2 T - C_1\frac{T^2}{2} = 0. \end{aligned}\right\}$$

Второе уравнение дает соотношение $C_2 = \frac{T}{2}C_1$, а из первого уравнения находим $C_1 = -\frac{12}{T^3}x_1^{\text{н}}$.

В результате получаем следующее выражение для оптимального управления:

$$u^*(t) = \frac{6}{T^2}x_1^{\text{н}}\left(-1 + \frac{2}{T}t\right).$$

Ему соответствует график, представленный на рис. 8.3.

Заметим, что при положительном начальном отклонении, когда $x_1^{\text{н}} > 0$, на начальном этапе при $0 \le t \le \frac{T}{2}$ управление должно быть отрицательным, а на втором этапе $\left(\frac{T}{2} \le t \le T\right)$ – положительным.

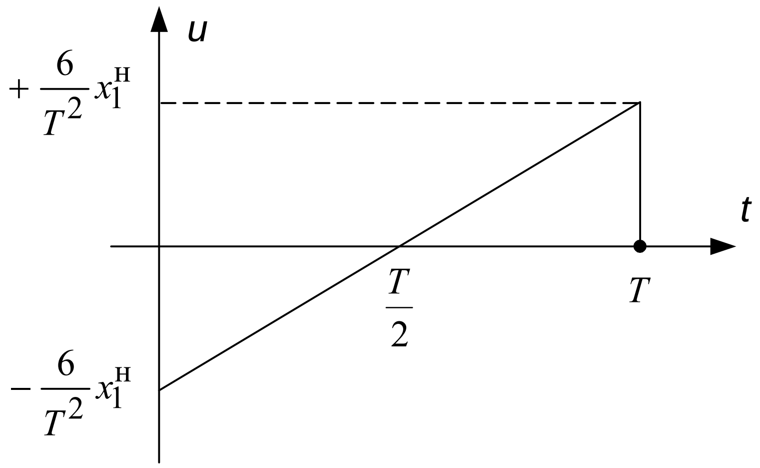


Рис. 8.3

Для физического примера, представленного на рис. 8.1, это означает, что сначала сила, приложенная к объекту, должна быть направлена противоположно положительному направлению оси y, а затем – наоборот, т.е. на финальном этапе воздействие должно иметь тормозящий характер.

Кроме того, обратим внимание на следующее свойство полученного оптимального управления: чем меньше выделено времени на переходный процесс (T), тем больше должны быть начальная и конечная величины управляющего воздействия $\left(\frac{6}{T^2}x_1^{\text{н}}\right)$ и круче прямая (см. рис. 8.3), определяющая программу изменения этого воздействия.

§ 8.4. Задача об оптимальном по быстродействию управлении

Задача 1. Рассмотрим задачу управления тем же объектом (8.18), но с двумя изменениями:

1) вместо критерия (8.19) будем рассматривать показатель
$$J = t_{\text{к}} - t_{\text{н}} = t_{\text{к}} \to \min\,, \tag{8.25}$$
который выражает затраты времени на управление;
2) примем условие $|u| \le U_0$, учитывающее ограниченность управляющего ресурса.

Решение. На первый взгляд может показаться, что критерий (8.25) не соответствует общему критерию (8.2), фигурирующему в ПМ. Однако в действительности выражение (8.25) легко получается из (8.2), если принять $f_0 \equiv 1$. Тогда
$$J = \int_0^{t_{\text{к}}} f_0 dt = \int_0^{t_{\text{к}}} dt = t_{\text{к}}\,.$$

Запишем выражение для функции Гамильтона в этом случае:
$$H = \psi_0 \cdot 1 + \psi_1 x_2 + \psi_2 ku\,, \tag{8.26}$$
$$-U^0 \le u \le U^0\,. \tag{8.27}$$

Как видим из (8.26), в данном случае H линейно зависит от u. Значит, наибольшее значение H может принимать только на концах промежутка (8.27):
$$u^* = \begin{cases} +U^0, \text{если } \psi_2 \ge 0, \\ -U^0, \text{если } \psi_2 < 0. \end{cases} \tag{8.28}$$

Уравнения для ψ_1 и ψ_2 остаются теми же, что и в предыдущем случае, значит
$$\psi_2 = C_2 - C_1 t\,.$$

Следовательно, оптимальное по быстродействию управление будет иметь вид, представленный на рис. 8.4.

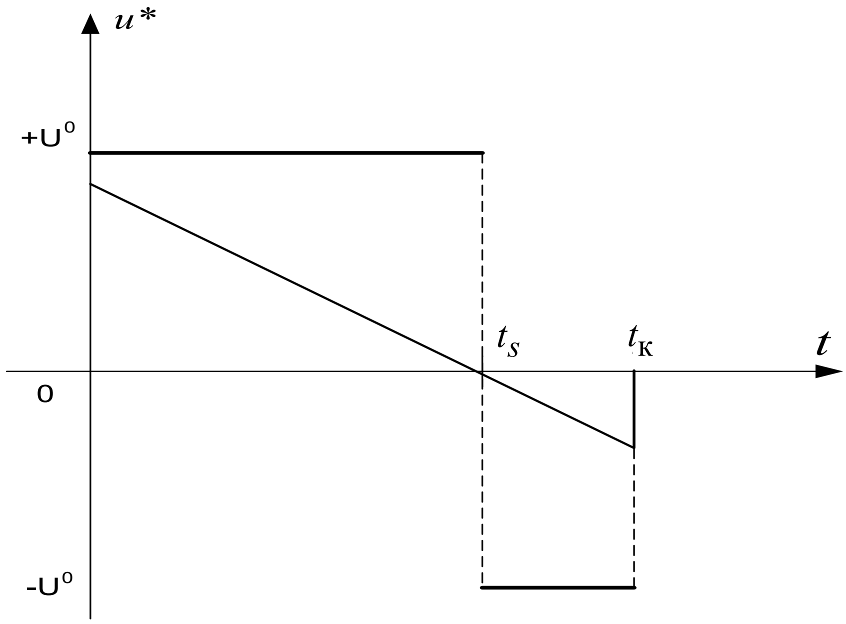


Рис. 8.4

Таким образом, в рассмотренном случае получается кусочно-постоянная программа формирования оптимального управления:

- на промежутке $[0,\ t_s]$ подается максимально возможное воздействие, «разгоняющее» объект;
- на промежутке $[t_s,\ t_к]$ – максимально возможное «тормозящее» воздействие;
- при $t \geq t_к$ $u \equiv 0$.

Для окончательной конкретизации этого оптимального управления нужно найти два момента времени t_s и $t_к$, а для этого следует определить константы $C_{1,2}$, входящие в ψ_2.

Характер действий, необходимых для этого, остается тем же, что и в предыдущем случае: для получения управления u^* (8.28), зависящего от $C_{1,2}$, нужно найти решения уравнения состояния (8.18) $x_1(t)$ и $x_2(t)$, которые будут зависеть от дополнительных констант C_3 и C_4.

Используя граничные условия – два начальных и два конечных условия, можно получить четыре соотношения для отыскания четырех неизвестных C_i.

Интегрируя уравнения (8.18) при постоянном управлении u, можно получить следующие выражения:

для $u = +U^0$

$$\left.\begin{aligned} x_1(t) &= x_1(0) + x_2(0)t + A\frac{t^2}{2}, \\ x_2(t) &= x_2(0) + At; \end{aligned}\right\} \tag{8.29}$$

для $u = -U^0$

$$\left.\begin{aligned} x_1(t) &= x_1(0) + x_2(0)t - A\frac{t^2}{2}, \\ x_2(t) &= x_2(0) - At, \end{aligned}\right\} \tag{8.30}$$

где $A = kU^0$.

Будем далее рассматривать следующие краевые условия:

$$\left.\begin{aligned} x_1(0) &= x_1^{\text{н}}, \\ x_2(0) &= 0; \end{aligned}\right\} \quad \left.\begin{aligned} x_1(t_{\text{к}}) &= 0, \\ x_2(t_{\text{к}}) &= 0. \end{aligned}\right\} \tag{8.31}$$

В этом случае при $x_1^{\text{н}} > 0$ нетрудно установить, что на первом этапе управления воздействие должно быть отрицательным, а на втором – положительным. Из соотношений (8.30) получаем:

1-й этап ($u = -U^0$):

$$\left.\begin{aligned} x_1(t) &= x_1^{\text{н}} - \frac{A}{2}t^2, \\ x_2(t) &= -At. \end{aligned}\right\}$$

Обозначим $t = T_1$ продолжительность первого этапа. Для него в конечный момент будут получаться следующие значения переменных состояния:

$$\left.\begin{aligned} x_1(T_1) &= x_1^{\text{н}} - \frac{A}{2}T_1^2 = x_1^s, \\ x_2(T_1) &= -AT_1 = x_2^s. \end{aligned}\right\} \tag{8.32}$$

Их следует рассматривать в качестве начальных условий для второго этапа.

2-й этап ($u = +U^0$):

$$\left.\begin{aligned} x_1(\tilde{t}) &= x_1^s + x_2^s\,\tilde{t} + \frac{A}{2}\tilde{t}^2, \\ x_2(\tilde{t}) &= x_2^s + A\tilde{t}, \end{aligned}\right\} \tag{8.33}$$

где $\tilde{t} = t - T_1$ – время, отсчитываемое от момента переключения.

В конце второго этапа при $\tilde{t} = T_2$ должно выполняться условие (8.31), т.е. $x_1(T_2) = 0$; $x_2(T_2) = 0$.

Используя соотношения (8.33) совместно с (8.32), получаем:

$$\left.\begin{aligned} &\left(x_1^{\text{н}} - \frac{A}{2}T_1^2\right) - (AT_1)T_2 + \frac{A}{2}T_2^2 = 0, \\ &-(AT_1) + AT_2 = 0. \end{aligned}\right\}$$

Из второго уравнения следует, что $T_1 = T_2$. В таком случае первое уравнение примет следующий вид: $x_1^{\text{н}} - \frac{A}{2}T_1^2 - AT_1^2 + \frac{A}{2}T_1^2 = 0$ или $x_1^{\text{н}} = AT_1^2$.

Таким образом, для заданных краевых условий (8.31) продолжительность обоих этапов оказывается одинаковой:

$$T_1 = T_2 = \sqrt{\frac{x_1^{\text{н}}}{kU^0}}.$$

График оптимального управления представлен на рис 8.5.

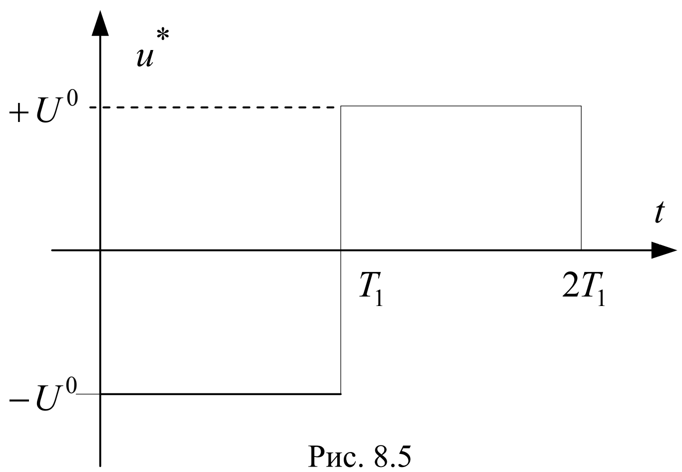


Рис. 8.5

Данное оптимальное по быстродействию управление существенно отличается от управления (см. рис. 8.3), полученного для этой же системы, но при минимизации затрат энергии на управление.

Примем для этого случая начальное условие $x_1^{\text{н}} = 1$, а время переходного процесса $T_Э^* = 1$ с. Тогда управление, имея линейный характер (см. рис. 8.3), будет лежать в диапазоне $-6 \le u \le 6$.

Если для оптимального по быстродействию управления принять $x_1^{\text{н}} = 1$, $U^0 = 6$, то время переходного процесса будет равно:

$$T_Б^* = 2T_1 = 2\sqrt{\frac{1}{6}} = \sqrt{\frac{2}{3}} < T_Э^*.$$

Следует отметить, что принцип максимума позволяет доказать, что оптимальное по быстродействию управление не только объектом (8.18), но и любой системой, имеющей линейную модель (7.17) при ограничении $|u| \le U^0$, будет кусочно-постоянным и должна принимать только предельные значения из заданного промежутка допустимых значений (8.27). Кроме того, установлено, что в тех случаях, когда все собственные значения системной матрицы A в модели (7.17) являются вещественными, управление будет иметь не более n интервалов постоянства, где n – порядок модели. Этот результат называют теоремой об n интервалах [34].

Задача 2 [34]. Рассмотрим задачу об управлении тем же объектом (8.18), т.е.

$$\left.\begin{aligned} x_1' &= x_2, \\ x_2' &= u, \end{aligned}\right\}$$

с краевыми условиями:

$$x_1(0) = x_2(0) = 0;\; x_1(t_\text{к}) = x_1^\text{к} = a;\; x_2(t_\text{к}) = 0.$$

Требуется найти оптимальное по быстродействию управление, при котором $J = t_\text{к} \to \min$.

Рассмотренное выше в задаче 1 ограничение на величину управления $|u| \le U_0$ заменим интегральным ограничением

$$\int_0^{t_\text{к}} u^2(t)dt = b. \tag{8.34}$$

Решение. Введем новую переменную $x_3 = \int_0^t u^2(t)dt$. Для нее можно записать следующие соотношения: $x_3' = u^2$; $x_3(0) = 0$; $x_3(t_\text{к}) = b$.

Сформируем функцию Гамильтона: $H = \psi_0 \cdot 1 + \psi_1 x_2 + \psi_2 u + \psi_3 u^2$.

Условие ее максимальности дает уравнение $\dfrac{dH}{du} = \psi_2 + 2\psi_3 u = 0$.

Отсюда получаем

$$u^* = -\frac{1}{2}\left(\frac{\psi_2}{\psi_3}\right).$$

Запишем уравнения для сопряженных переменных:

$$\left.\begin{aligned}\psi_1' &= -\frac{\partial H}{\partial x_1} = 0,\\ \psi_2' &= -\frac{\partial H}{\partial x_2} = -\psi_1,\\ \psi_3' &= -\frac{\partial H}{\partial x_3} = 0.\end{aligned}\right\}$$

Из этих уравнений получаем: $\psi_1 = C_1$; $\psi_2 = -C_1 t + C_2$; $\psi_3 = C_3$.

Используя эти результаты, можно записать выражение, определяющее форму оптимального управления:

$$u^* = -\frac{\psi_2}{2\psi_3} = \frac{C_1 t - C_2}{2C_3} = C_4 t - C_5,$$

где $C_4 = \dfrac{C_1}{2C_3}$; $C_5 = \dfrac{C_2}{2C_3}$.

Подставляем полученное выражение для управляющего воздействия в уравнения (8.18):

$$\left.\begin{aligned}x_1' &= x_2,\\ x_2' &= C_4 t - C_5.\end{aligned}\right\}$$

В результате интегрирования получаем:

$$\left.\begin{aligned}x_2 &= C_4\frac{t^2}{2} - C_5 t + C_6,\\ x_1 &= C_4\frac{t^3}{6} - C_5\frac{t^2}{2} + C_6 t + C_7.\end{aligned}\right\}$$

При $t = 0$ $x_1 = x_2 = 0$; поэтому $C_6 = 0$ и $C_7 = 0$.

При $t = t_к$ $x_1 = a$; $x_2 = 0$, поэтому:

$$\left.\begin{aligned}C_4\frac{t_к^2}{2} - C_5 t_к &= 0,\\ C_4\frac{t_к^3}{6} - C_5\frac{t_к^2}{2} &= a.\end{aligned}\right\} \tag{8.35}$$

Кроме этого нужно рассмотреть условие (8.34):

$$x_3 = \int_0^{t_к} (C_4^2 t^2 - 2C_4 C_5 t + C_5^2)\,dt = b$$

или

$$\frac{C_4^2}{3} t_к^3 - C_4 C_5 t_к^2 + C_5^2 t_к = b. \tag{8.36}$$

Из системы уравнений (8.35) получаем:

$$C_5 = \frac{1}{2} t_к C_4; \quad C_4 = -(12a)/t_к^3; \quad C_5 = -(6a)/t_к^2.$$

Из условия (8.36) находим:

$$t_к = \sqrt[3]{\frac{12a^2}{b}}.$$

При этом оптимальное по быстродействию управление оказалось линейным. Оно определяется выражением

$$u^* = C\left[1 - 2\left(\frac{t}{t_к}\right)\right], \text{ где } C = \frac{6a}{t_к^2}.$$

График формирования оптимального управления показан на рис.8.6.

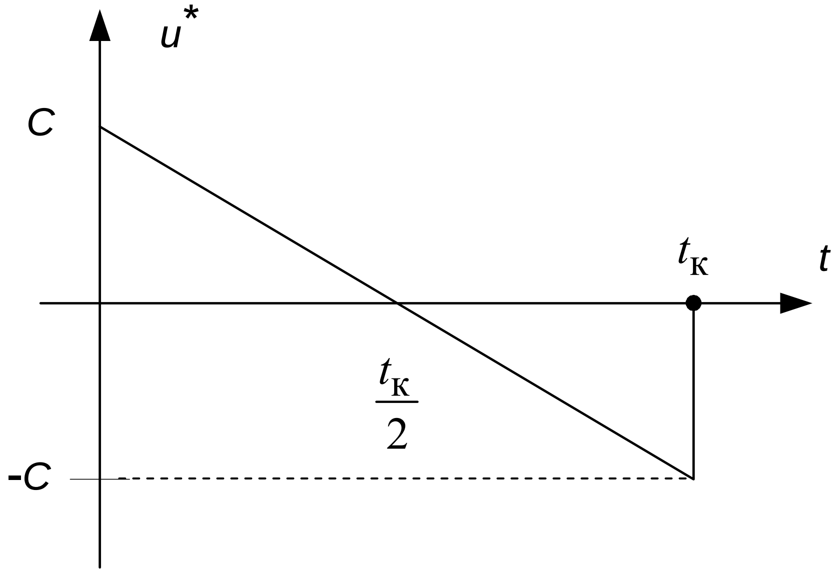


Рис. 8.6

§ 8.5. Задачи об управлении линейной динамической системой со свободным правым концом

Рассмотрим примеры применения ПМ для решения задач об оптимальном управлении, сформулированных в форме задач Больца [31, 34]. Как было показано в § 8.1, решение таких задач можно осуществить с помощью ПМ, дополнив его условием трансверсальности (8.10).

Задача 1 [24]. Рассматривается управляемая линейная система второго порядка, модель которой имеет следующий вид:

$$\left.\begin{aligned} \frac{dx_1}{dt} &= x_1 + x_2, \\ \frac{dx_2}{dt} &= u. \end{aligned}\right\} \tag{8.37}$$

Допустимые значения управляющего воздействия ограничены диапазоном $0 \le u \le 1$. Управление осуществляется на промежутке времени $0 \le t \le 3$.

Начальное состояние управляемой системы задано:

$$x_1(0) = 1;\ \ x_2(0) = 0, \tag{8.38}$$

а конечное состояние $\bar{x}(3)$ не зафиксировано.

Требуется найти управление $u = u^*(t)$, $0 \le t \le 3$, которое минимизировало бы заданный функционал качества:

$$J = x_1(3) + \int_0^3 (x_1 + 2x_2 - 3u)dt. \tag{8.39}$$

Решение

Сформируем функцию Гамильтона:

$$H = \psi_0(x_1 + 2x_2 - 3u) + \psi_1(x_1 + x_2) + \psi_2 u,$$

где $\psi_0 = -1$.

Это выражение можно записать в следующем виде:

$$H = u(\psi_2 + 3) + \psi_1(x_1 + x_2) - (x_1 + 2x_2).$$

Учитывая линейность полученного выражения по u и заданные ограничения на управление, из условия максимальности H можно получить следующие соотношения, определяющие вид оптимального управления:

$$u^* = \begin{cases} 1, \text{ если } \left(\psi_2 + 3\right) > 0, \\ 0, \text{ если } \left(\psi_2 + 3\right) < 0, \\ \forall u \in [0;1], \text{ если} (\psi_2 + 3) = 0. \end{cases} \tag{8.40}$$

Уравнения (8.6) для вспомогательных переменных ψ_i в данном случае принимают следующий вид:

$$\left.\begin{aligned} \psi_1' &= -\frac{\partial H}{\partial x_1} = -\left(\psi_1 - 1\right) = 1 - \psi_1, \\ \psi_2' &= -\frac{\partial H}{\partial x_2} = -\left(\psi_1 - 2\right) = 2 - \psi_1. \end{aligned}\right\} \tag{8.41}$$

Первое уравнение $\psi_1' + \psi_1 = 1$ является линейным неоднородным дифференциальным уравнением первого порядка. Используя стандартные приемы, несложно найти его решение:

$$\psi_1(t) = Ce^{-t} + (1 - e^{-t}), \tag{8.42}$$

где C – подлежащая определению константа. Для ее отыскания воспользуемся условием трансверсальности:

$$\psi_1(t_к) = -\frac{\partial F}{\partial x_1} = -\frac{\partial}{\partial x_1}(x_1) = -1. \tag{8.43}$$

При $t_к = 3$ из (8.42) – (8.43) получаем $Ce^{-3} + (1 - e^{-3}) = -1$. Из этого соотношения находим искомую константу: $C = 1 - 2e^3$. Подставив ее в (8.42), можно получить следующее выражение: $\psi_1 = 1 - k\,e^{-t}$, где $k = 2e^3 = 40{,}171$.

Рассмотрим теперь уравнение, определяющее вторую переменную из (8.41): $\psi_2' = 2 - (1 - ke^{-t}) = 1 + ke^{-t}$.

Интегрируя это выражение, получаем:

$$\psi_2 = t - ke^{-t} + C. \tag{8.44}$$

Для определения C воспользуемся вторым условием трансверсальности:

$$\psi_2(t_к) = -\frac{\partial F}{\partial x_2}. \tag{8.45}$$

Правая часть в (8.45) будет равна нулю, так как в рассматриваемой задаче $F = x_1(3)$ от x_2 не зависит. Получаем условие $3 - ke^{-3} + C = 0$ или $C = -3 + 2e^3\,e^{-3} = -1$. Таким образом, $\psi_2 = t - ke^{-t} - 1$.

В соответствии с (8.40) величина оптимального управления определяется знаком переменной: $\sigma = \psi_2 + 3 = t - ke^{-t} + 2$.

Заметим, что функция $\sigma(t)$ обладает следующими свойствами:

$$\sigma(0) = 2 - k \approx -37{,}4; \quad \sigma(3) = 3 - 2e^3 e^{-3} + 2 = 3;$$

$$\sigma' = 1 + ke^{-t} > 0.$$

Эта функция монотонно вырастает от отрицательного значения на левом конце промежутка $0 \le t \le 3$ до положительного значения на правом конце. Учитывая непрерывность этой функции, можно сделать вывод о том, что она имеет единственную точку $t = \tau$, в которой величина σ меняет знак. Вид этой функции показан на рис. 8.7, *а*.

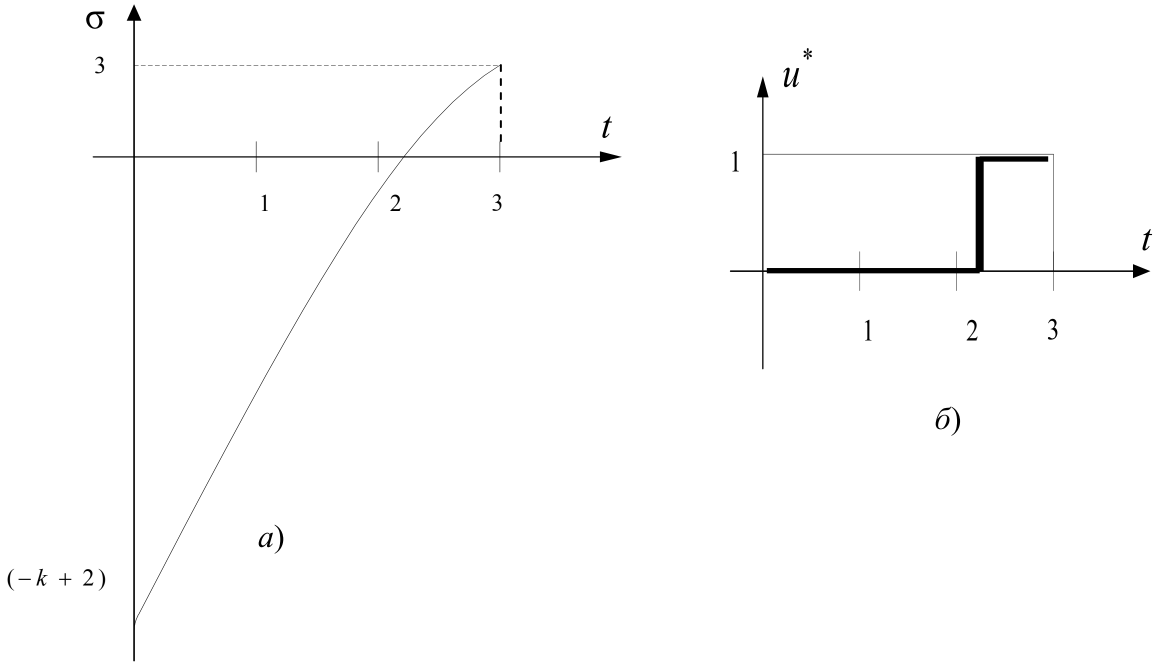


Рис. 8.7

Уравнение $\tau - ke^{-\tau} + 2 = 0$ является трансцендентным. Его приближенное решение можно найти каким-либо численным методом, например методом деления пополам [17]. В результате получается $\tau = 2{,}247$.

В соответствии с (8.40) оптимальное управление определяется следующим образом:

$$u^* = \begin{cases} 0 & \text{при } 0 \le t \le \tau, \\ 1 & \text{при } \tau < t \le 3. \end{cases} \tag{8.46}$$

Его вид показан на рис. 8.7, *б*.

Для отыскания оптимальных переменных состояния рассмотрим (8.37) с начальным условием (8.38) для функции $u = u(t)$, определяемой (8.46).

На первом этапе $0 \le t \le \tau$, когда $u = 0$, получаем уравнение

$$\left.\begin{aligned} x_1' &= x_1 + x_2, \\ x_2' &= 0, \end{aligned}\right\}$$

$$x_1(0) = 1;\ x_2(0) = 0.$$

Из второго уравнения получаем $x_2 = C = \text{const}$, а учитывая начальное условие $x_2(0) = 0$, следует принять $x_2 = \text{const} = 0$.

Из первого уравнения $x_1' - x_1 = 0$ можно получить следующее решение: $x_1(t) = x_1(0)e^t$. Для заданного начального условия получается $x_1(t) = e^t$.

На границе первого этапа процесса управления переменные будут равны следующим значениям: $x_1(\tau) = e^{\tau}$; $x_2(\tau) = 0$.

Их следует принять в качестве начальных условий для второго этапа, на котором система будет описываться уравнениями:

$$\left.\begin{aligned} x_1' &= x_1 + x_2, \\ x_2' &= 1. \end{aligned}\right\}$$

Для второго этапа введем новый отсчет времени, полагая $\tilde{t} = t - \tau$.

Из уравнения $x_2' = 1$ при $x_2(\tilde{t} = 0) = 0$ находим $x_2(\tilde{t}) = \tilde{t}$.

Тогда первое уравнение примет следующий вид:

$$x_1'(\tilde{t}) - x_1(\tilde{t}) = \tilde{t},$$

$$x_1(\tilde{t} = 0) = e^{\tau}.$$

Интегрируя это уравнение, получим: $x_1(\tilde{t}) = e^{\tau}e^{\tilde{t}} + \left[e^{\tilde{t}} - 1 - \tilde{t}\right]$.

Если вернуться к времени t, то решение, определяющее оптимальный процесс, примет вид:

$$x_1^* = \begin{cases} e^t, & 0 \le t \le \tau, \\ e^t + e^{t-\tau} - 1 - t + \tau, & \tau \le t \le 3, \end{cases}$$

$$x_2^* = \begin{cases} 0, & 0 \le t \le \tau, \\ t - \tau, & \tau < t \le 3. \end{cases}$$

Задача 2 [24]. Найти оптимальное управление линейной системой первого порядка:

$$x' = x + u,\ x(0) = x_0,\ 0 \le t \le t_{\text{к}} = T, \tag{8.47}$$

при котором

$$J = \frac{1}{2}x^2(t_{\text{к}}) + \int_0^{t_{\text{к}}} (x^2 + u^2)dt \to \min. \tag{8.48}$$

Решение. Запишем функцию Гамильтона:

$$H=\psi_0(x^2+u^2)+\psi_1(x+u).$$

Для вспомогательных переменных в соответствии с ПМ получаем соотношения

$$\left.\begin{aligned}\psi_1' &= -\frac{\partial H}{\partial x}=-(2\psi_0 x+\psi_1),\\ \psi_0 &= -1\end{aligned}\right\}$$

или

$$\psi_1'+\psi_1=2x. \tag{8.49}$$

Рассмотрим условие максимальности H. Учитывая то, что ограничений на управление в данной задаче нет, максимум H определяется из необходимого условия экстремума:

$$\frac{\partial H}{\partial u}=-2u+\psi_1=0,$$

откуда получаем

$$u^*=\frac{1}{2}\psi_1. \tag{8.50}$$

Таким образом, задача свелась к рассмотрению системы из двух дифференциальных уравнений (8.47) и (8.49), в которой u определяется (8.50), т.е.

$$\left.\begin{aligned}x' &= x+\frac{1}{2}\psi,\\ \psi' &= 2x-\psi.\end{aligned}\right\} \tag{8.51}$$

Для решения этой системы выполним следующие преобразования:

– из второго уравнения выразим x:

$$x=\frac{1}{2}\left(\psi'+\psi\right); \tag{8.52}$$

– дифференцируем это выражение:

$$x'=\frac{1}{2}\left(\psi''+\psi'\right);$$

– затем оба соотношения подставим в первое уравнение из системы (8.51):

$$\frac{1}{2}\left(\psi''+\psi'\right)=\frac{1}{2}\left(\psi'+\psi\right)+\frac{1}{2}\psi;$$

– после упрощений получаем однородное уравнение второго порядка:

$$\psi'' - 2\psi = 0.$$

Его характеристическое уравнение $\lambda^2 - 2 = 0$ имеет корни: $\lambda_1 = \sqrt{2}$; $\lambda_2 = -\sqrt{2}$. Общее решение будет иметь вид:

$$\psi(t) = C_1 e^{\sqrt{2}t} + C_2 e^{-\sqrt{2}t}. \tag{8.53}$$

Подставив это выражение, а также $\psi' = \sqrt{2}C_1 e^{\sqrt{2}t} - \sqrt{2}e^{-\sqrt{2}t}$ в уравнение (8.52), получаем соотношение, определяющее оптимальный переходный процесс:

$$x = C_1\left(\frac{1+\sqrt{2}}{2}\right)e^{\sqrt{2}t} + C_2\left(\frac{1-\sqrt{2}}{2}\right)e^{-\sqrt{2}t}. \tag{8.54}$$

В найденные выражения (8.53) – (8.54) для $\psi(t)$ и $x(t)$ входят две неизвестные константы C_1 и C_2. Для их определения нужно воспользоваться начальным условием $x(0) = x_0$ и условием трансверсальности (8.10):

$$\psi(T) = -\left.\frac{\partial F}{\partial x}\right|_{t=T} = -\frac{\partial}{\partial x}\left(\frac{1}{2}x^2\right) = -x(T).$$

Первое условие дает соотношение

$$C_1\left(\frac{1+\sqrt{2}}{2}\right) + C_2\left(\frac{1-\sqrt{2}}{2}\right) = x_0. \tag{8.55}$$

Второе условие приводит к уравнению

$$C_1 e^{\sqrt{2}T} + C_2 e^{-\sqrt{2}T} = -\left[C_1\left(\frac{1+\sqrt{2}}{2}\right)e^{\sqrt{2}T} + C_2\left(\frac{1-\sqrt{2}}{2}\right)e^{-\sqrt{2}T}\right],$$

которое преобразуется к виду

$$C_1 e^{\sqrt{2}T}\left(1 + \frac{1+\sqrt{2}}{2}\right) + C_2 e^{-\sqrt{2}T}\left(1 + \frac{1-\sqrt{2}}{2}\right) = 0$$

или

$$C_1 e^{\sqrt{2}T}\left(\frac{3+\sqrt{2}}{2}\right) + C_2 e^{-\sqrt{2}T}\left(\frac{3-\sqrt{2}}{2}\right) = 0,$$

или

$$C_2 = -C_1 e^{2\sqrt{2}T}\,\frac{3+\sqrt{2}}{3-\sqrt{2}}. \tag{8.56}$$

Подставив это соотношение в (8.55), находим:

$$C_1 = \frac{2x_0}{(1+\sqrt{2})-(1-\sqrt{2})e^{2\sqrt{2}T}\,\dfrac{(3+\sqrt{2})}{(3-\sqrt{2})}}.$$

Полученное значение C_1 совместно C_2, вычисленным по (8.56), полностью определяет оптимальное управление:

$$u^* = \frac{1}{2}\left(C_1 e^{\sqrt{2}t} + C_2 e^{-\sqrt{2}t}\right),\ 0 \le t \le T,$$

и оптимальный процесс $x = x^*(t)$ из (8.54).

§ 8.6. Задача об управлении линейной динамической системой с минимизацией обобщенного квадратичного интегрального показателя

Рассмотрим задачу об управлении системой, модель которой можно представить в виде линейных дифференциальных уравнений, т.е.

$$\frac{d\overline{x}}{dt} = A\overline{x} + B\overline{u}, \tag{8.57}$$

где A и B – матрицы $(n \times n)$ и $(n \times m)$ коэффициентов a_{ij} и b_{ij}.

Требуется найти закон формирования управляющих воздействий $\overline{u} = \overline{u}^*(t)$, при которых система переходит из начального состояния $\overline{x}(0) = \overline{x}_{\text{н}}$ в конечное состояние $\overline{x}(t_{\text{к}}) = 0$ и при этом минимизируется показатель качества в виде следующего функционала:

$$J = \frac{1}{2}\int_0^{t_{\text{к}}} (\overline{x}^{\text{T}} Q\overline{x} + \overline{u}^{\text{T}} R\overline{u})\, dt \to \min, \tag{8.58}$$

где R и Q – квадратные симметрические матрицы заданных или выбранных весовых коэффициентов, причем Q – неотрицательно-определенная матрица, а R – положительно-определенная матрица [17]; в этом случае $\overline{x}^{\text{T}} Q\,\overline{x} \ge 0$ для любых $\overline{x} \ne 0$, $\overline{u}^{\text{T}} R\,\overline{u} > 0$ для любых $\overline{u}$.

Входящие в подынтегральное выражение компоненты являются квадратичными формами, например:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} q_{11}x_1^2 + 2\,q_{12}x_1x_2 + q_{22}x_2^2 \end{bmatrix}.$$

Будем считать все элементы матриц *A, B, Q, R* постоянными числами, а $t_к \to \infty$.

Для отыскания оптимального управления воспользуемся принципом максимума. Сформируем функцию Гамильтона (8.4) для уравнения (8.57) и функционала (8.58):

$$H = \Psi_0 \frac{1}{2}\left(\overline{x}^{\mathrm{T}} Q \overline{x} + \overline{u}^{\mathrm{T}} R \overline{u}\right) + \overline{\Psi}^{\mathrm{T}}\left(A\overline{x} + B\overline{u}\right), \tag{8.59}$$

где $\Psi_0 = -1$, а функции $\Psi_i(t)$, $i = \overline{1, n}$, входящие в $\overline{\Psi}$, удовлетворяют системе уравнений

$$\frac{d\overline{\Psi}}{dt} = -\frac{\partial H}{\partial \overline{x}} = -\left(-Q\overline{x} + A^{\mathrm{T}}\overline{\Psi}\right). \tag{8.60}$$

При этом использовались правила дифференцирования матричных выражений, приведенные в приложении.

Учитывая, что для оптимального управления функция Гамильтона должна иметь максимум, запишем необходимое условие экстремума:

$$\frac{\partial H}{\partial \overline{u}} = 0.$$

Дифференцируя выражение (8.59) по $\overline{u}$, получаем $-R\overline{u} + B^{\mathrm{T}}\overline{\Psi} = 0$, откуда находим следующее выражение для оптимального управления:

$$\overline{u} = \overline{u}^* = R^{-1}B^{\mathrm{T}}\overline{\Psi}. \tag{8.61}$$

Сделаем важное предположение о том, что вектор вспомогательных переменных $\overline{\Psi}$ можно представить в виде

$$\overline{\Psi} = P\overline{x}, \tag{8.62}$$

где *P* – неизвестная пока матрица коэффициентов p_{ij}, связывающих $\overline{\Psi}$ и $\overline{x}$.

Если принять это предположение, то алгоритм оптимального управления (8.61) будет иметь следующий вид:

$$\overline{u} = R^{-1}B^{\mathrm{T}}P\,\overline{x}$$

или

$$\overline{u} = K\overline{x}, \tag{8.63}$$

где

$$K = R^{-1}B^{\mathrm{T}}P. \tag{8.64}$$

Для определения неизвестной матрицы P подставим (8.62) в (8.60), предполагая матрицу P не зависящей от t:

$$P\frac{d\overline{x}}{dt} = Q\overline{x} - A^{\mathrm{T}}P\overline{x}. \qquad (8.65)$$

После подстановки выражения (8.63) для $\overline{u}$ в уравнение модели (8.57) получаем $\dfrac{d\overline{x}}{dt} = A\overline{x} + BR^{-1}B^{\mathrm{T}}P\overline{x}$ или $\left(PA + PBR^{-1}B^{\mathrm{T}}P - Q + A^{\mathrm{T}}P\right)\overline{x} = 0$.

Это равенство должно выполняться для любых $\overline{x} = \overline{x}(t)$, а для этого необходимо, чтобы выражение в скобках равнялось нулю. Отсюда получаем уравнение для определения неизвестной матрицы P:

$$PA + PBR^{-1}B^{\mathrm{T}}P - Q + A^{\mathrm{T}}P = 0. \qquad (8.66)$$

Таким образом, в случае, когда уравнение (8.66) имеет решение $P = P^{*}$, предположение (8.62) можно считать оправданным и при этом оптимальное управление довольно просто выражается через переменные состояния системы и реализуется в виде обратных связей:

$$\left.\begin{aligned} u_1^{*} &= k_{11}x_1 + \ldots + k_{1n}x_n,\\ u_2^{*} &= k_{21}x_1 + \ldots + k_{2n}x_n,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ u_m^{*} &= k_{m1}x_1 + \ldots + k_{mn}x_n. \end{aligned}\right\} \qquad (8.67)$$

Настроечные коэффициенты k_{ij} алгоритма управления (8.63) или (8.67) определяются из соотношений (8.64) и (8.66).

Пример. Пусть управляемая система имеет один выход y и один вход u, а модель системы представлена уравнением $T_0\,y'' + y' = k_0\,u$.

Обозначим $y = x_1$; $y' = x_2$, тогда

$$\left.\begin{aligned} x_1' &= x_2,\\ x_2' &= -\frac{1}{T_0}x_2 + \frac{k_0}{T_0}u. \end{aligned}\right\}$$

Эта модель соответствует (8.57) при

$$A = \begin{bmatrix} 0 & 1\\ 0 & -a \end{bmatrix},\quad B = \overline{b} = \begin{bmatrix} 0\\ b \end{bmatrix},$$

где $a = 1/T_0$; $b = k_0/T_0$.

Качество управления оценивается по критерию (8.58) при

$$Q=\begin{bmatrix} q_1 & 1 \\ 0 & q_2 \end{bmatrix} \text{ и } R=[r].$$

Рассмотрим уравнение (8.66) для данного случая:

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & -a \end{bmatrix}+\frac{1}{r}\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & b^2 \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}+$$

$$+\begin{bmatrix} 0 & 0 \\ 1 & -a \end{bmatrix}\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}=\begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix},$$

где использовано соотношение

$$B R^{-1} B^{\mathrm{T}}=\begin{bmatrix} 0 \\ b \end{bmatrix}\left[r^{-1}\right]\begin{bmatrix} 0 & b \end{bmatrix}=\frac{1}{r}\begin{bmatrix} 0 & 0 \\ 0 & b^2 \end{bmatrix}.$$

Выполнив перемножение матриц, находим:

$$\begin{bmatrix} 0 & p_{11}-a\,p_{12} \\ 0 & p_{12}-ap_{22} \end{bmatrix}+\frac{b^2}{r}\begin{bmatrix} p_{12}p_{21} & p_{12}p_{22} \\ p_{22}p_{21} & p_{22}^2 \end{bmatrix}+\begin{bmatrix} 0 & 0 \\ p_{11}-ap_{21} & p_{12}-ap_{22} \end{bmatrix}=\begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}.$$

Приравняв соответствующие элементы матриц в левой и правой частях этого соотношения, запишем четыре уравнения для отыскания четырех неизвестных элементов матрицы P:

$$\left.\begin{aligned}
&\frac{b^2}{r}p_{12}\,p_{21}=q_1,\\
&\left(p_{11}-a\,p_{12}\right)+\frac{b^2}{r}p_{12}\,p_{22}=0,\\
&\left(p_{11}-a\,p_{21}\right)+\frac{b^2}{r}p_{22}\,p_{21}=0,\\
&\left(p_{21}-a\,p_{22}\right)+\left(p_{12}-a\,p_{22}\right)+\frac{b^2}{r}p_{22}^2=q_2.
\end{aligned}\right\} \tag{8.68}$$

Разрешим второе и третье уравнения относительно p_{12} и p_{21}:

$$\left.\begin{aligned}
p_{12}&=-\frac{p_{11}}{-a+\dfrac{b^2}{r}p_{22}},\\
p_{21}&=-\frac{p_{11}}{-a+\dfrac{b^2}{r}p_{22}}.
\end{aligned}\right\}$$

Сравнивая полученные выражения, заключаем, что $p_{12} = p_{21}$, значит матрица P – симметрическая. Обозначив $p_{12} = p_{21} = p$, из 1-го уравнения системы (8.68) находим:

$$p^2 = \frac{q_1 r}{b^2} \text{ или } p = p_{12} = p_{21} = \pm\frac{\sqrt{q_1 r}}{b}. \qquad (8.69)$$

Последнее уравнение (8.68) с учетом полученного соотношения (8.69) можно записать в виде

$$\frac{b^2}{r} p_{22}^2 - 2a\, p_{22} - q_2 + 2p = 0\,.$$

Корни уравнения

$$p_{22}^2 - \frac{2a\,r}{b^2} p_{22} - \frac{r}{b^2}\left(q_2 \mp \frac{2\sqrt{q_1 r}}{b} \right) = 0$$

будем определять выражением

$$p_{22} = \frac{a\,r}{b^2} \pm \sqrt{\frac{a^2 r^2}{b^4} + \frac{r}{b^2}\left(q_2 \mp \frac{2\sqrt{q_1 r}}{b} \right)}. \qquad (8.70)$$

Рассмотрим алгоритм оптимального управления для исследуемой системы

$$u = \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} k_1 x_1 + k_2 x_2 \end{bmatrix}, \qquad (8.71)$$

где $K = \begin{bmatrix} k_1 & k_2 \end{bmatrix} = \frac{1}{r}\begin{bmatrix} 0 & b \end{bmatrix} \begin{bmatrix} p_{11} & p \\ p & p_{22} \end{bmatrix} = \frac{b}{r}\begin{bmatrix} p & p_{22} \end{bmatrix}$.

Можно показать, что для получения устойчивой системы необходимо принять концепцию отрицательной обратной связи, а для этого в формулах (8.69) и (8.70), содержащих знак «±», взять знак «–».

Таким образом, оптимальный по критерию (8.58) алгоритм управления определяется следующим соотношением:

$$u = -(k_1 x_1 + k_2 x_2)\,, \qquad (8.72)$$

где

$$\left.\begin{aligned} k_1 &= \left(\frac{b}{r}\right)\frac{\sqrt{q_1 r}}{b} = \sqrt{\frac{q_1}{r}}, \\ k_2 &= -\frac{1}{k_0} + \sqrt{\frac{1}{k_0^2} + \frac{q_2}{r} + \frac{2T_0}{k_0}\sqrt{\frac{q_1}{r}}}. \end{aligned}\right\} \qquad (8.73)$$

Ясно, что разрешить матричное уравнение (8.66) относительно p_{ij} и получить в явном виде формулы, определяющие параметры k_i алгоритма управления через параметры модели и критерия оптимальности, удалось благодаря простоте и в первую очередь низкому порядку рассмотренной системы.

Для более сложных систем это невозможно. В этих случаях для решения уравнения (8.66) можно использовать следующую итерационную процедуру для отыскания матрицы P:

$$P_{k+1} = P_k + h\overline{f}(P_k), \tag{8.74}$$

где k – номер итерации; P_k – матрица P, получаемая на k-й итерации; P_0 – некоторое начальное приближение, например $P_0 = I$; $\overline{f}$ – матричное выражение, стоящее в левой части решаемого уравнения (8.66); h – шаг решения.

При достаточно малом шаге h и большом количестве выполненных шагов матрицы P_k будут приближаться к искомой матрице P, удовлетворяющей уравнению $f(P) = 0$. Разумеется, что реализацию вычислительного процесса по формуле (8.74) следует осуществлять с помощью компьютера, подготовив соответствующую программу для организации вычислений по (8.74) и (8.66). При этом можно использовать стандартные подпрограммы для выполнения матричных операций.

Изложенную выше методику решения можно применять для синтеза оптимального по критерию (8.58) управления в более сложных условиях, когда управляемая система нестационарна из-за того, что элементы $a_{ij}(t)$; $b_{ij}(t)$ матриц, входящих в модель (8.57), не являются постоянными числами. Кроме того, в условиях задачи могут рассматриваться процессы управления на конечном промежутке времени $0 \le t \le t_\text{к} < \infty$.

Можно показать (см., например, [30]), что и в этом случае алгоритм управления будет иметь вид (8.63), но матрица K будет зависеть от времени. Ее определение потребует решения более сложного дифференциального уравнения вместо уравнения (8.66).

В заключение отметим, что оптимальное по критерию (8.58) управление зависит не только от параметров управляемой системы (элементов матриц A и B), но и от параметров q_{ij} и r_{ij}, входящих в критерий оптимальности.

Заметим, в частности, что второе слагаемое в критерии (8.58), как отмечалось ранее, оценивает затраты энергии на управление. Если в обобщенном показателе J им придать малый вес ($r \to 0$), то определяющие по (8.73) оптимальное управление настроечные коэффициенты $k_{1,2}$, а вслед за ними и само управление u будет стремиться к бесконечности, а это не соответствует реальным условиям, поскольку имеющиеся для управления ресурсы всегда ограничены.

При решении реальных практических задач параметры критерия $q_{1,2}$ и r обычно не заданы и их приходится выбирать. Обоснованный выбор конкретных значений указанных параметров обычно осуществить довольно трудно. Для этого можно, например, воспользоваться следующими рекомендациями [34]:

1) матрицы Q и R брать диагональными;
2) исходя из специфики конкретной рассматриваемой управляемой системы (технологических, экономических и других требований), оценить желательную длительность процесса управления $T_н = t_к - t_н$, а также максимально допустимые значения $|x_i(t)| = x_{iМ}$, $|u_i(t)| = u_{iМ}$.

Тогда параметры обобщенного критерия оптимальности можно определить следующим образом [34]:

$$q_{ii} = \frac{1}{T_н |x_{iМ}|^2}, \qquad r_{ii} = \frac{1}{T_н |u_{iМ}|^2}.$$

§ 8.7. Два подхода к реализации оптимального управления

Выше при рассмотрении задачи оптимального управления в § 8.1 – 8.5 предполагалось, что управляющее воздействие должно быть найдено, а затем реализовано в виде искомой функции времени $\bar{u} = \bar{u}^*(t)$. Этот закон изменения управляющих факторов нередко называют программой управления.

Однако в § 8.6 был рассмотрен иной подход к формированию управляющего воздействия в виде $\bar{u} = \bar{u}^*(\bar{x})$, т.е. функции от текущих значений переменных состояния.

Обе концепции формирования управления можно проиллюстрировать схемами, представленными на рис. 8.8, где слева показана программная реализация управления, а справа – управление по принципу обратной связи.

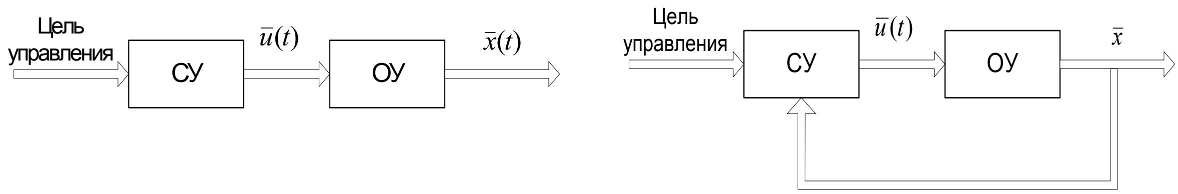


Рис. 8.8

Концепция обратной связи оказалась плодотворной при решении задачи об управлении оптимальном по обобщенному квадратичному интегральному показателю качества; рассмотренный в § 8.6 закон управления получается довольно простым и для общего случая (8.63), и в частном случае (8.72).

Синтез управления в виде функции от переменных состояния оказывается возможным и целесообразным и в других случаях.

Рассмотрим в качестве примера задачу об оптимальном быстродействии, анализ которой был проведен в § 8.4. Вернемся к модели системы (8.18)

$$\left.\begin{aligned} \frac{dx_1}{dt} &= x_2, \\ \frac{dx_2}{dt} &= ku, \end{aligned}\right\}$$

$$u \in \{-U^0; +U^0\}.$$

Разделив второе уравнение на первое, получим:

$$\frac{dx_2}{dx_1} = \frac{ku}{x_2}$$

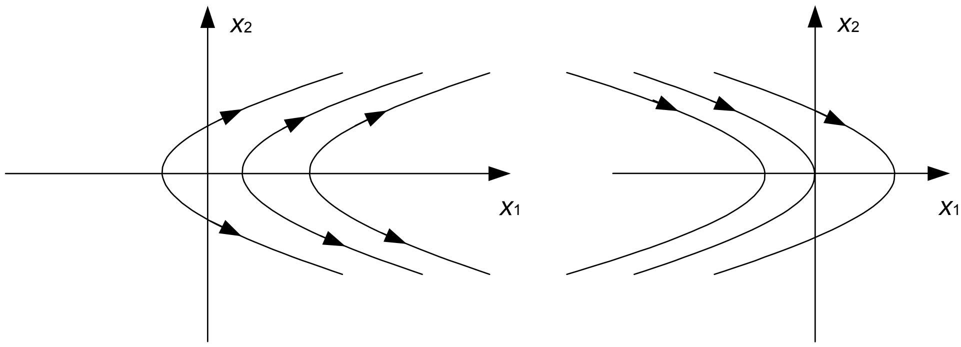
или

$$x_2 dx_2 = k \cdot u dx_1 . \tag{8.75}$$

Рассмотрим стадию процесса управления, когда $u = +U^0 = \text{const}$. Для нее, интегрируя (8.75), находим

$$\frac{x_2^2}{2} = kU^0 x_1 + C . \tag{8.76}$$

Введем в рассмотрение плоскость переменных состояния (x_1, x_2) (рис. 8.9, *а*). На ней уравнению (8.76) будет соответствовать семейство парабол, получающихся при различных значениях *С*.



а) *б*)

Рис. 8.9

Переходный процесс в системе будет связан с изменением переменных состояния $x_1 = x_1(t)$ и $x_2 = x_2(t)$, при этом точка на указанной плоскости будет скользить по параболе, проходящей через точку, соответствующую начальному состоянию системы.

Аналогичное уравнение и семейство парабол получается для стадии процесса, на которой $u = -U^0$. Они показаны на рис. 8.9, *б*.

Рассмотрим теперь поведение системы «в целом», объединяя отдельные стадии. При этом будем учитывать, что переходный процесс для любого начального состояния должен завершиться попаданием в начало координат, а количество тактов управления должно быть не более двух.

Это означает, что процесс на плоскости переменных состояния (рис. 8.10) в общем случае будет состоять из двух участков:

– завершающего участка, ведущего в т. *O* по полупараболе *MO* или *NO*;

– предшествующего участка, ведущего из начальной точки на линию *MON*.

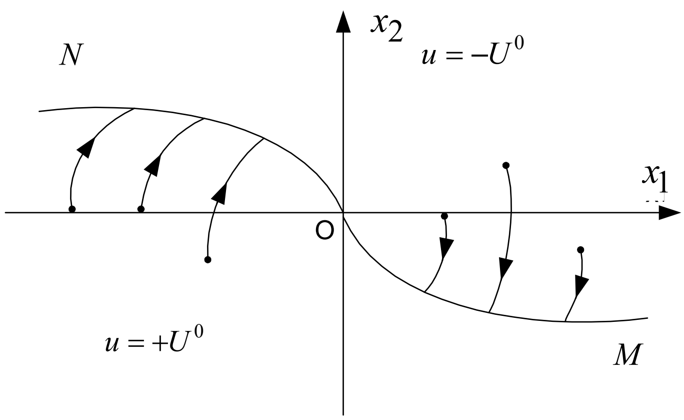


Рис. 8.10

Как видим, линия *MON*, составленная из двух симметричных полупарабол, в формировании управления играет особую роль. При попадании точки, изображающей состояние системы, на эту линию должно происходить скачкообразное переключение значения управления с $+U^0$ на $-U^0$ или наоборот. Поэтому линию *MON* называют линией оптимальных переключений. Она определяется двумя уравнениями:

$$x_1 = \frac{x_2^2}{kU^0} \text{ при } x_2 < 0 \text{ и } x_1 = -\frac{x_2^2}{kU^0} \text{ при } x_2 > 0.$$

Эти условия можно объединить в одно уравнение

$$z = x_1 + \frac{1}{kU^0} x_2 |x_2| = 0.$$

Используя вспомогательную переменную z, оптимальный алгоритм управления можно записать следующим образом:

$$u = \begin{cases} +U^0, \text{ если } z < 0, \\ -U^0, \text{ если } z > 0, \end{cases}$$

$$u = \begin{cases} +U^0, \text{ если } z = 0 \text{ и } x_2 < 0 \\ -U^0, \text{ если } z = 0 \text{ и } x_2 > 0, \end{cases}$$

$$u = 0, \text{ если } x_1 = x_2 = 0.$$

Рассмотренный подход дает достаточно простой алгоритм управления, а главное позволяет обойти довольно сложную проблему отыскания констант C_i, необходимых для формирования управления в программном режиме.

§ 8.8. Приближенные методы отыскания решения задач оптимального управления

Примеры, рассмотренные в гл. 8, показывают, что даже в простейших случаях, когда управляемая система описывается линейным ДУ второго порядка, решение задачи об оптимальном управлении с помощью принципа максимума, так же как и по уравнениям Эйлера – Лагранжа, редко удается довести до результата в явном (аналитическом) виде, определяющем программу формирования оптимального управления. Ситуация значительно усложняется с ростом порядка модели или появлением в ней нелинейных эффектов.

В этих случаях приходится использовать приближенные, итерационные процедуры. Сущность таких процедур, анализ трудностей, которые могут возникнуть при их реализации, изложены в специальной литературе, например в [35].

Рассмотрим один из таких подходов к приближенному отысканию оптимальных управлений. Допустим, что с помощью ПМ удалось установить вид, форму или характер искомого управления и для конкретизации остается определить численные значения нескольких параметров.

Типичным примером подобного рода может служить решение задачи об оптимальном быстродействии. В § 8.4 было показано, что оптимальное управление будет кусочно-постоянным и примет предельные возможные значения.

Это довольно важный результат, позволяющий существенно сузить класс функций, из которых должно быть выбрано искомое управляющее воздействие, и свести задачу к определению длительностей интервалов постоянства $\{t_1, t_2, ..., t_n\}$.

Пусть управляемая система имеет одно управляющее воздействие $u(t)$ и один «выход» $y(t)$, связь между которыми задается моделью в виде дифференциального уравнения

$$y^{(n)} = f(y, y', ..., y^{(n-1)}, u). \tag{8.77}$$

Предполагается, что начальное состояние и конечное состояние заданы, а управление ограничено: $-U^0 \le u \le U^0$. Требуется найти управление, которое переводит систему из начального состояния в заданное конечное состояние за минимальное время.

Перейдем от уравнения (8.77) к системе уравнений первого порядка, полагая $y = x_1$; $y' = x_2, ...$; $y^{(n-1)} = x_n$. Тогда

$$\left.\begin{aligned} & x_1' = x_2, \\ & x_2' = x_3, \\ & \ldots\ldots\ldots\ldots \\ & x_n' = f(x_1, x_2, ... x_n; u). \end{aligned}\right\} \tag{8.78}$$

Допустим, что искомое управление является кусочно-постоянным и состоит из n тактов. Его форма показана на рис. 8.11 для $n = 3$.

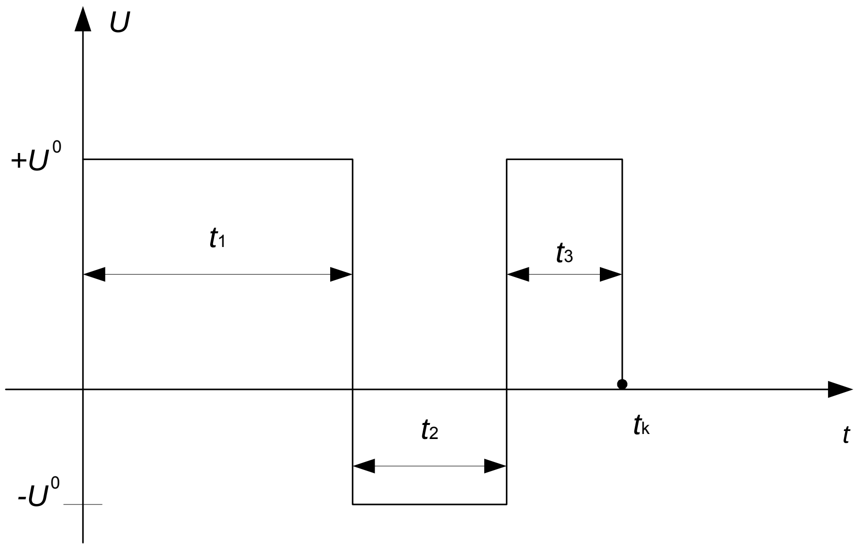


Рис. 8.11

Как отмечалось в § 8.4, в соответствии с теоремой о n интервалах таким будет оптимальное по быстродействию управление, если управляемая система линейна и ее собственные значения являются вещественными.

Будем характеризовать выбранное управление упорядоченным набором из n чисел $\{T_1, T_2, \ldots, T_n\}$, в котором $|T_i| = t_i$, а знак T_i определяется знаком управляющего воздействия на соответствующем i-м такте. Например, набор {+2,3; –1,7; +3,6} определяет релейное управление, в котором на первом такте $u = +U^0$ и он продолжается $t_1 = 2{,}3$ с; на 2-м такте $u = -U^0$ и он длится $t_2 = 1{,}7$ с; а на третьем такте $u = +U^0$ и его продолжительность равна $t_3 = 3{,}6$ с.

Для отыскания искомого оптимального управления осуществим следующие действия.

1. Зададимся некоторым начальным набором из n чисел $\left\{T_1^{(0)}, T_2^{(0)}, \ldots, T_n^{(0)}\right\}$, определяющим некоторое кусочно-постоянное управление $u = u^{(0)}(t)$ указанного выше типа.

2. Для него (управления $u^{(j)}(t)$) найдем процессы в рассматриваемой системе; для этого нужно решить систему уравнений (8.78) при заданных начальных условиях и выбранном управлении. Решение можно осуществить каким-либо численным методом, например методом Эйлера [17]. В

таком случае отрезок времени [0, T], где $T = \sum_{j=1}^{n} |t_j|$, разбивается на мелкие части Δt и дискреты искомых переменных $x_i(k\Delta t)$ находятся по рекуррентным соотношениям:

$$\left.\begin{aligned}
&x_1(k+1) = x_1(k) + \Delta t x_2(k),\\
&x_2(k+1) = x_2(k) + \Delta t x_3(k),\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&x_n(k+1) = x_n(k) + \Delta t f\left(x_1(k), x_2(k), \dots, x_n(k); u(k)\right),
\end{aligned}\right\}$$

где $x_1(0), \dots, x_n(0)$ – заданные начальные условия; $k = 0, 1, 2, \dots, N-1$, $N = T/\Delta t$.

3. Фиксируем значения полученных переменных состояния в конечный момент времени и сравниваем их с заданными граничными условиями:

$$\left.\begin{aligned}
&\varepsilon_1 = x_1^0(T) - x_1(N\Delta t),\\
&\varepsilon_2 = x_2^0(T) - x_2(N\Delta t),\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\varepsilon_n = x_n^0(T) - x_n(N\Delta t).
\end{aligned}\right\}$$

4. Находим обобщенную меру несоответствия заданных и полученных значений переменных состояния, например

$$\varepsilon = \sum_{m=1}^{n} |\varepsilon_m|.$$

Сравниваем эту величину с заданной величиной допустимой погрешности $\varepsilon_{доп}$. Если $\varepsilon < \varepsilon_{доп}$, то искомое кусочно-постоянное управление, переводящее систему из заданного начального состояния в заданное конечное состояние, можно считать найденным. В противном случае переходим к следующему действию.

5. Корректируем управляющее воздействие по формуле

$$T_i^{(j+1)} = T_i^{(j)} + \left(\Delta T_i\right)^{(j)}, \quad j = 1, 2, \dots,$$

где i – номер такта в управлении; j – номер итерации в процессе поиска решения; $\left(\Delta T_i\right)^{(j)}$ – величина, на которую корректируется i-й такт после j-й итерации.

Затем переходим к действию 2.

Изложенный порядок действий реализуется до тех пор, пока на некоторой итерации условие останова $\varepsilon < \varepsilon_{доп}$ не окажется выполненным.

Для реализации конкретного алгоритма формирования величины $(\Delta T_i)^{(j)}$ можно использовать различные подходы.

1. Использование универсальных алгоритмов минимизации функций.

Вычисления, осуществляемые в соответствии с изложенными выше пп. 1 – 4, позволяют каждому набору чисел $\{T_1, T_2, ..., T_n\}$ поставить в соответствие число ε и тем самым определить некоторую функцию $\varepsilon = F(T_1, T_2, ..., T_n)$, где F – правило соответствия, определяемое через действия пп. 1 – 4.

Корректировку ΔT_i можно осуществлять, используя один из алгоритмов поиска минимума функции $\varepsilon = F(\overline{T})$, рассмотренных в гл 5.

2. Использование специального алгоритма поиска интервалов постоянства.

В [39] предложен оригинальный, довольно простой и эффективный алгоритм корректировки интервалов кусочно-постоянного воздействия:

$$(\Delta T_i)^{(j)} = b_i \, \varepsilon_i^{(j)}, \tag{8.79}$$

где b_i – некоторые весовые коэффициенты, которые рекомендуется выбирать равными $b_i \in \{0{,}2; 0{,}4; 0{,}8; ...\}$.

Как видим из (8.79), в этом алгоритме длина первого интервала корректируется по величине невязки (отклонения) по первой координате, длина второго интервала – по величине ошибки по второй координате, т.е. скорости изменения ошибки, третьего интервала – по ускорению ошибки и т.д.

Как показывают эксперименты, проведенные для различных тестовых систем, даже при жестких требованиях к точности и наугад выбранном начальном управлении $\{T_i^{(0)}\}$ рассмотренный алгоритм позволяет найти искомое управление за сравнительно небольшое количество итераций, по крайней мере, гораздо быстрее, чем при использовании универсальных алгоритмов поиска экстремума.

Контрольные вопросы

1. Какая задача об оптимальном управлении непрерывной системой рассматривается при формулировке принципа максимума?
2. Как формулируется принцип максимума Л.С. Понтрягина для непрерывной системы?

3. В каком случае используется и как записывается условие трансверсальности в дополнение к принципу максимума?
4. Каким образом используется принцип максимума для решения классической задачи Эйлера об отыскании функций $u(t), x(t), 0 \le t \le T$, для которых $x' = u$; $x(0) = 0$; $x(T) = a$; $J = \int_0^T (x^2 + u^2)dt \to \min$?
5. Каким образом используется принцип максимума для решения задачи об отыскании функций $u(t)$, $y(t)$, $0 \le t \le t_к$, для которых $y'' = ku$; $y(0) = y^0$; $y'(0) = y_1^0$; $y(t_к) = y'(t_к) = 0$; $J = \int_0^{t_к} u^2(t)dt \to \min$?
6. Каким образом используется принцип максимума для решения задачи об оптимальном по быстродействию управлении, при котором $y'' = ku$; $0 \le t \le t_к$; $y(0) = y^0$; $y'(0) = y_1^0$; $y(t_к) = y'(t_к) = 0$; $|u| \le U^0$; $J = t_к \to \min$?
7. Какой вид имеет оптимальное по быстродействию управление, как формулируется теорема об n интервалах?
8. Как формулируется задача об управлении линейной динамической системой с минимизацией обобщенного квадратичного интегрального критерия качества? Какой вид имеет оптимальное управление в этом случае?
9. Пояснить с помощью структурной схемы два возможных подхода к реализации оптимального управления по разомкнутому и замкнутому принципам.
10. Чем обусловлена необходимость в использовании приближенных методов для отыскания оптимального управления динамической системой?
11. Какие принципы и какие расчетные соотношения используются при отыскании оптимального управления на основе приближенных подходов?

Примечание. Задачи для самостоятельного решения по данной теме приведены в конце гл. 9.

Глава 9. ИСПОЛЬЗОВАНИЕ ПРИНЦИПА МАКСИМУМА ДЛЯ ОТЫСКАНИЯ ОПТИМАЛЬНЫХ УПРАВЛЕНИЙ В ДИСКРЕТНЫХ СИСТЕМАХ

§ 9.1. Условия оптимальности для дискретных динамических систем

Как отмечалось в § 6.1, существует немало систем, в том числе экономических, в которых процессы имеют дискретный характер, т.е. изменяются или фиксируются в отдельные изолированные моменты времени t_i, обычно следующие друг за другом с одинаковым шагом T, называемым периодом дискретности.

Рассмотрим классическую задачу об управлении такой системой, полагая, что ее математическая модель представляет собой систему разностных уравнений, которую запишем в следующем виде:

$$\left.\begin{array}{l} x_1(t+1) = f_1(x_1(t),...,x_n(t);u_1(t),...,u_m(t);t), \\ x_2(t+1) = f_2(x_1(t),...,x_n(t);u_1(t),...,u_m(t);t), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_n(t+1) = f_n(x_1(t),...,x_n(t);u_1(t),...,u_m(t);t), \end{array}\right\} \tag{9.1}$$

где t – дискретное время (номер шага или этапа), $t = 0,\ 1,\ 2,...,K-1$; K – количество шагов; n – порядок системы; m – количество управляющих воздействий.

Будем, как и прежде, пользоваться матрицами-столбцами (векторами), тогда модель (9.1) можно записать в кратком виде:

$$\bar{x}(t+1) = \bar{f}[\bar{x}(t),\bar{u}(t),t]. \tag{9.1 a}$$

Требуется найти (сформировать) вектор управлений $\bar{u} = \bar{u}^*(t)$, при котором система (9.1) была бы переведена из заданного начального состояния $\bar{x}(0) = \bar{x}_{\text{н}}$ в конечное состояние $\bar{x}(k) = \bar{x}_{\text{К}}$ и при этом минимизировался показатель качества в виде функционала

$$J = \sum_{t=0}^{k-1} f_0(\bar{x}(t),\bar{u}(t),t) + F(\bar{x}(k)) \to \min. \tag{9.2}$$

Можно показать, что необходимые условия оптимальности процессов для этой задачи во многом аналогичны условиям принципа максимума для непрерывных систем и их использование сводится к следующим действиям [24, 30, 34]:

1. Формируем функцию Гамильтона:

$$H(x,u,\psi,t) = \sum_{i=0}^{n} \psi_i f_i \qquad (9.3)$$

где f_0 – функция, фигурирующая в критерии качества, а $f_1,...,f_n$ – выражения, стоящие в правых частях системы разностных уравнений (9.1); $\psi_i = \psi_i(t)$ – вспомогательные функции, причем $\psi_0 = -1$.

2. Составляем так называемую сопряженную систему для вспомогательных функций ψ_i $(i = \overline{1,n})$:

$$\psi_i(t) = \frac{\partial H(\overline{x}(t), \overline{u}^*(t), \overline{\psi}(t+1), t)}{\partial x_i}, \qquad (9.4)$$

и рассматриваем ее совместно с исходной моделью управляемого процесса (9.1).

3. Анализируем условие экстремальности функции Гамильтона на оптимальном решении. При отсутствии прямых ограничений на управления это условие приводит к уравнениям:

$$\left.\frac{\partial H(\overline{x}^*(t), \overline{u}^*(t), \overline{\psi}(t+1), t)}{\partial u_i}\right|_{\overline{u}=\overline{u}^*(t)} = 0 \quad (i = \overline{1,m}). \qquad (9.5)$$

4. Учитываем краевые условия: $\overline{x}(0) = \overline{x}_{\text{н}}$; $\overline{x}(k) = \overline{x}_{\text{к}}$, или, если правый конец свободен, условия трансверсальности:

$$\psi_i(k) = -\left.\frac{\partial F(\overline{x})}{\partial x_i}\right|_{\overline{x}^*(k)}. \qquad (9.6)$$

Поясним сущность изложенных операций по отысканию оптимального дискретного управления на следующих примерах.

Задача 1. Рассмотрим дискретную систему первого порядка, которая описывается уравнением

$$x(t+1) = -x(t) + u(t)\ . \qquad (9.7)$$

Ее поведение рассматривается на пяти этапах. Задано нулевое начальное условие: $x(0) = 0$, а конечное значение $x(5)$ не зафиксировано.

Требуется найти дискретные значения управления $u^* = u(t)$, $t = 0,1,2,3,4$, при которых минимизируется показатель качества:

$$J = x(5) + \sum_{t=0}^{4} [x^2(t) + u^2(t)]. \qquad (9.8)$$

Решение

Запишем функцию Гамильтона:

$$H = \psi_0[x^2(t) + u^2(t)] + \psi_1(-x(t) + u(t)), \tag{9.9}$$

где $\psi_0 = -1$.

Рассмотрим условие (9.5) экстремальности этой функции

$$\frac{\partial H}{\partial u} = -2u(t) + \psi_1(t+1) = 0,$$

из которого получаем соотношение, определяющее оптимальное управление:

$$u^*(t) = \frac{1}{2}\psi_1(t+1). \tag{9.10}$$

Запишем сопряженное уравнение (9.4) для функции (9.9):

$$\psi(t) = \frac{\partial}{\partial x}[-x^2(t) - u^2(t) + \psi_1(t+1)(-x(t) + u(t))]$$

или

$$\psi(t) = -2x(t) - \psi_1(t+1). \tag{9.11}$$

Так как правый конец управляемого процесса свободен, рассматриваем условие трансверсальности (9.6):

$$\psi_1(5) = -\left.\frac{\partial F}{\partial x}\right|_{x=x(5)} = -1. \tag{9.12}$$

Полученные соотношения (9.10) и (9.11) с учетом (9.7) образуют следующую взаимосвязанную систему разностных уравнений:

$$\left.\begin{aligned} x(t+1) &= -x(t) + \frac{1}{2}\psi(t+1), \\ \psi(t+1) &= -2x(t) - \psi(t), \end{aligned}\right\}$$

для которой задано начальное условие по переменной x и конечное условие по переменной ψ:

$$x(0) = 0;\ \psi(5) = -1. \tag{9.13}$$

Подставив выражение для $\psi(t+1)$ из второго уравнения в правую часть первого уравнения, получаем систему разностных уравнений стандартного типа:

$$\left.\begin{aligned} x(t+1) &= -2x(t) - \frac{1}{2}\psi(t), \\ \psi(t+1) &= -2x(t) - \psi(t). \end{aligned}\right\} \tag{9.14}$$

Рассмотрим дискреты анализируемых процессов x и ψ, определяемых этой системой (9.14) поэтапно, учитывая при этом условия (9.13):
$t = 0$:

$$\left.\begin{aligned} x(1) &= -2x(0) - \frac{1}{2}\psi(0) = -\frac{1}{2}\psi(0), \\ \psi(1) &= -2x(0) - \psi(0) = -\psi(0); \end{aligned}\right\}$$

$t = 1$:

$$\left.\begin{aligned} x(2) &= -2x(1) - \frac{1}{2}\psi(1) = \psi(0) + \frac{1}{2}\psi(0) = \frac{3}{2}\psi(0), \\ \psi(2) &= -2x(1) - \psi(1) = \psi(0) + \psi(0) = 2\psi(0); \end{aligned}\right\}$$

$t = 2$:

$$\left.\begin{aligned} x(3) &= -2x(2) - \frac{1}{2}\psi(2) = -3\psi(0) - \psi(0) = -4\psi(0), \\ \psi(3) &= -2x(2) - \psi(2) = -3\psi(0) - 2\psi(0) = -5\psi(0); \end{aligned}\right\}$$

$t = 3$:

$$\left.\begin{aligned} x(4) &= -2x(3) - \frac{1}{2}\psi(3) = 8\psi(0) + \frac{5}{2}\psi(0) = \frac{21}{2}\psi(0), \\ \psi(4) &= -2x(3) - \psi(3) = 8\psi(0) + 5\psi(0) = 13\psi(0); \end{aligned}\right\}$$

$t = 4$:

$$\left.\begin{aligned} x(5) &= -2x(4) - \frac{1}{2}\psi(4) = -21\psi(0) - \frac{13}{2}\psi(0) = -\frac{55}{2}\psi(0), \\ \psi(5) &= -2x(4) - \psi(4) = -21\psi(0) - 13\psi(0) = -34\psi(0). \end{aligned}\right\}$$

Воспользуемся условием (9.13): $\psi(5) = -34\psi(0) = -1$, откуда находим неизвестный параметр, вошедший во все предыдущие соотношения: $\psi(0) = \frac{1}{34}$.

Теперь можно легко найти дискреты оптимального управляемого процесса:

$$x(0) = 0;\ x(1) = -\frac{1}{68};\ x(2) = \frac{3}{68};\ x(3) = -\frac{2}{17};\ x(4) = \frac{21}{68};\ x(5) = -\frac{55}{68},$$

а также дискреты оптимального управления из соотношения (9.10):

$$u(0) = \frac{1}{2}\psi(1) = -\frac{1}{68};\ u(1) = \frac{1}{2}\psi(2) = \frac{3}{17};\ u(2) = \frac{1}{2}\psi(3) = -\frac{5}{68};$$

$$u(3) = \frac{1}{2}\psi(41) = \frac{13}{34};\ u(4) = \frac{1}{2}\psi(5) = -\frac{1}{2}.$$

Задача 2 [30]. Дана дискретная система первого порядка:

$$x(t+1) = x(t) + \alpha u(t), \tag{9.15}$$

поведение которой рассматривается на десяти этапах.

Требуется сформировать управление $u = u(t)$, которое переводит систему из единичного начального состояния в нулевое конечное состояние, т.е.

$$x(0) = 1;\ x(10) = 0. \tag{9.16}$$

Кроме того, необходимо минимизировать функционал качества:

$$J = \frac{1}{2}\sum_{t=0}^{9} u^2(t). \tag{9.17}$$

Решение. Формируем функцию Гамильтона:

$$H = -\frac{1}{2}u^2(t) + \psi(t+1)[x(t) + \alpha u(t)], \tag{9.18}$$

и рассматриваем для нее необходимое условие экстремума:

$$\frac{\partial H}{\partial u} = -u(t) + \alpha\psi(t+1) = 0,$$

откуда получаем:

$$u^*(t) = \alpha\psi(t+1). \tag{9.19}$$

Записываем уравнение (9.4) для функции (9.18)

$$\psi(t) = \frac{\partial H}{\partial x} = \psi(t+1). \tag{9.20}$$

Рассматриваем систему уравнений (9.15), (9.19) и (9.20):

$$\left.\begin{array}{l} x(t+1) = x(t) + \alpha^2\psi(t+1), \\ \psi(t+1) = \psi(t). \end{array}\right\}$$

Второе уравнение показывает, что в рассматриваемой задаче все значения ψ будут одинаковыми, обозначим $\psi(t) = c = \mathrm{const}$. Тогда $x(t+1) = x(t) + \alpha^2 c$. Для различных t получим:

$$x(1) = x(0) + \alpha^2 c;$$

$$x(2) = x(1) + \alpha^2 c = x(0) + 2\alpha^2 c;$$

$$x(3) = x(2) + \alpha^2 c = x(0) + 3\alpha^2 c;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$x(10) = x(0) + 10\alpha^2 c.$$

По условию $x(10)=0$. Из него находим неизвестную величину $\psi(t)=c=-\dfrac{x(0)}{10\alpha^2}$.

Таким образом, для выполнения поставленных условий управление должно быть постоянным и иметь значение: $u^*=\alpha\psi=-\dfrac{x_0}{10\alpha}$. При этом дискреты $x(t)$ управляемого процесса будут определяться выражением $x^*(t)=x_0+t\alpha^2c=x_0-t\dfrac{x_0}{10}=x_0(1-\dfrac{t}{10})$.

Заметим, что на границах промежутка управления $0\le t\le 10$ это соотношение дает $x^*(0)=x_0$; $x^*(10)=0$, что соответствует исходным условиям.

§ 9.2. Управление линейной дискретной системой произвольного порядка с оптимизацией суммарного обобщенного квадратичного критерия

Рассмотрим дискретный аналог задачи об оптимальном управлении линейной непрерывной системой, изложенной в § 8.6.

Задача. Пусть дискретная система описывается разностными уравнениями (9.1), причем все выражения f_i, определяющие правые части этих уравнений, являются линейными. В матричном виде такая модель описывается следующим образом:

$$\bar{x}(t+1)=A\bar{x}(t)+B\bar{u}(t), \tag{9.21}$$

где $\bar{x}(0)=\bar{x}^{\text{н}}$; A и B – матрицы размерами $n\times n$, $n\times m$; $\bar{x}$ и $\bar{u}$ – столбцы; $t=0,\ 1,\ 2,...,k$. В общем случае элементы матриц A и B могут зависеть от t (нестационарная система), в частном случае их элементы – постоянные числа (стационарная система).

Пусть функции F и f_0, входящие в функционал качества (9.2), представляют собой квадратичные формы (см. приложение), т.е.

$$J=\frac{1}{2}\bar{x}^T(k)\,S\,\bar{x}(k)+\frac{1}{2}\sum_{t=0}^{k-1}\Big(\bar{x}^T(t)\,Q\,\bar{x}(t)+\bar{u}^T(t)\,R\,\bar{u}(t)\Big), \tag{9.22}$$

где Q, R, S – заданные положительно определенные матрицы весовых коэффициентов, причем Q и R в принципе могут зависеть от t.

Требуется найти функции, описывающие управления $\bar{u}^* = \bar{u}(t)$ и состояния $\bar{x}^* = \bar{x}(t)$, при которых бы минимизировался показатель качества (9.22).

Решение. Составим гамильтониан:

$$H = \psi_0 \frac{1}{2}\left(\bar{x}^T(t) Q \bar{x}(t) + \bar{u}^T(t) R \bar{u}(t)\right) + \bar{\psi}^T(t+1)\left(A\bar{x}(t) + B\bar{u}(t)\right). \quad (9.23)$$

По нему сформируем сопряженную систему:

$$\bar{\psi}(t) = \frac{\partial H}{\partial \bar{x}} = -Q\bar{x}(t) + A^T \bar{\psi}(t+1). \quad (9.24)$$

Рассматриваем условие экстремальности H:

$$\frac{\partial H}{\partial \bar{u}} = -R\bar{u}(t) + B^T \bar{\psi}(t+1) = 0,$$

откуда получаем:

$$\bar{u}^*(t) = R^{-1} B^T \bar{\psi}(t+1). \quad (9.25)$$

Поскольку правый конец не зафиксирован, рассматриваем условие трансверсальности (9.6):

$$\bar{\psi}(k) = -\frac{\partial F}{\partial \bar{x}(k)} = -S\bar{x}(k). \quad (9.26)$$

Таким образом, как и в предыдущих примерах, мы получаем две взаимосвязанные системы уравнений, определяющих процессы $\bar{x}(t)$ и $\bar{\psi}(t)$:

$$\bar{x}(t+1) = A\bar{x}(t) + BR^{-1}B^T\bar{\psi}(t+1), \quad (9.27)$$

$$\bar{\psi}(t) = -Q\bar{x}(t) + A^T\bar{\psi}(t+1), \quad (9.28)$$

с начальным условием $\bar{x}(0) = \bar{x}^{\text{н}}$ и конечным условием $\bar{\psi}(k) = -S\,\bar{x}(k)$.

Предположим, как и в § 8.6 для непрерывной системы, что процессы $\bar{x}(t)$ и $\bar{\psi}(t)$ связаны соотношением

$$\bar{\psi}(t) = P\bar{x}(t), \quad (9.29)$$

где P – пока неизвестная матрица ($n \times n$).

Выясним теперь, оправдано ли такое предположение и какой должна быть матрица в случае, если это так.

Подставив (9.29) в уравнения (9.27), (9.28), получаем:

$$\left.\begin{aligned} &\bar{x}(t+1) = A\bar{x}(t) + BR^{-1}B^T P(t+1)\bar{x}(t+1),\\ &P(t)\bar{x}(t) = -Q\bar{x}(t) + A^T P(t+1)\bar{x}(t+1),\\ &P\bar{x}(k) = -S\bar{x}(k). \end{aligned}\right\} \quad (9.30)$$

Из первого уравнения выражаем $\bar{x}(t+1)$:

$$\bar{x}(t+1) = \left(I - B\,R^{-1}\,B^T\,P(t+1)\right)^{-1} A\,\bar{x}(t),$$

и подставляем полученный результат во второе уравнение:

$$P(t)\bar{x}(t) = -Q\,\bar{x}(t) + A^T\,P(t+1)\left(I - B\,R^{-1}\,B^T\,P(t+1)\right)^{-1} A\,\bar{x}(t)$$

или

$$\left[P(t) + Q - A^T\,P(t+1)\left(I - B\,R^{-1}\,B^T\,P(t+1)\right)^{-1} A\right]\bar{x}(t) = 0.$$

Чтобы это соотношение было справедливо для произвольного $\bar{x}(t)$, необходимо выполнение условия

$$\left.\begin{array}{l} P(t) + Q - A^T\,P(t+1)\left(I - B\,R^{-1}\,B^T\,P(t+1)\right)^{-1} A = 0, \\ P(k) = -S. \end{array}\right\} \qquad (9.31)$$

Получили так называемое разностное уравнение Риккати. Решая это матричное уравнение в обратном по времени порядке от $t = k$ к $t = 0$, можно получить матрицу $P(t)$. С ее использованием получается следующий алгоритм управления:

Сначала из (9.30) находим

$$A^T\,\bar{\psi}(t+1) = P\,\bar{x}(t) + Q\,\bar{x}(t) = (P+Q)\,\bar{x}(t),$$

затем

$$\bar{\psi}(t+1) = \left(A^T\right)^{-1}(P+Q)\,\bar{x}(t).$$

Тогда из (9.25) получаем:

$$\bar{u}^*(t) = R^{-1}\,B^T\left(A^T\right)^{-1}(P+Q)\,\bar{x}(t)$$

или

$$\bar{u}^*(t) = K\,\bar{x}(t), \qquad (9.32)$$

где

$$K = R^{-1}B^T\left(A^T\right)^{-1}(P+Q). \qquad (9.33)$$

Таким образом, алгоритм оптимального управления дискретной системой (9.21) для критерия (9.22), так же как и непрерывной системой (8.57) для критерия (8.58), оказывается пропорциональным.

В нем «матрица усиления» *K* определяется выражением (9.33), в которое входит матрица *P*, определенная из (9.31).

§ 9.3. Отыскание оптимального управления для дискретного прототипа непрерывной динамической системы

Существуют способы, позволяющие для изучаемой непрерывной динамической системы подобрать дискретную систему, процессы в которой близки к процессам в исходной непрерывной системе, если фиксировать их в одни и те же дискретные моменты времени $t = i\,h$, где $h = \Delta t$ – шаг дискретизации по времени (рис. 9.1).

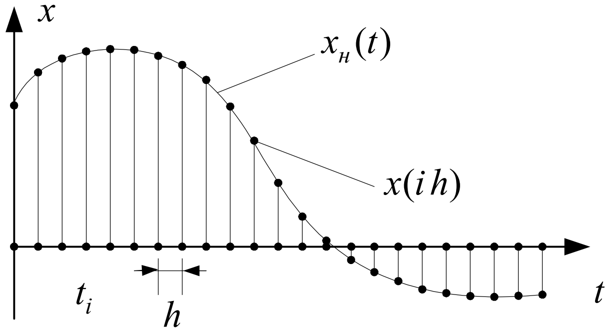


Рис. 9.1

Рассмотрим один из наиболее простых и часто используемых приемов получения такого дискретного прототипа на простом примере. В § 8.3 было изложено решение задачи об оптимальном управлении непрерывной системой второго порядка:

$$\left.\begin{aligned} x_1'(t) &= x_2(t),\\ x_2'(t) &= u(t); \end{aligned}\right\} \tag{9.34}$$

$$\left.\begin{aligned} x_1(0) &= x_1^{\text{н}},\\ x_2(0) &= 0, \end{aligned}\right\} \quad \left.\begin{aligned} x_1(t_{\text{к}}) &= 0,\\ x_2(t_{\text{к}}) &= 0; \end{aligned}\right\} \tag{9.35}$$

$$J = \frac{1}{2}\int_0^{t_{\text{к}}} u^2(t)\,dt \to \min. \tag{9.36}$$

Разобьем промежуток времени $[0,\ t_{\text{к}}]$ на k достаточно малых частей $\Delta t = h$. Будем рассматривать процессы $x_1(t), x_2(t)$ и $u(t)$ в полученные дискретные моменты времени: $t = t_i = i\,h$ $(i = 0,\ 1,\ 2,\ldots,k)$.

Заменим производные в (9.34) отношением приращений, учитывая малость этих приращений:

$$x'(i\,h) \approx \frac{x((i+1)\,h) - x(i\,h)}{h}.$$

Кроме того, заменим интеграл в критерии (9.36) суммой по методу прямоугольников:

$$J \approx \frac{1}{2} h \sum_{i=0}^{k-1} u^2(i\,h).$$

Для краткости аргумент $(i\,h)$ у переменных x и u заменим на i, считая, что здесь будет указываться вместо дискретного момента времени номер соответствующего отсчета.

Для системы (9.34) получаем:

$$\left.\begin{aligned} \frac{x_1(i+1)-x_1(i)}{h} &= x_2(i), \\ \frac{x_2(i+1)-x_2(i)}{h} &= u(i) \end{aligned}\right\}$$

или

$$\left.\begin{aligned} x_1(i+1) &= x_1(i)+h\,x_2(i), \\ x_2(i+1) &= x_2(i)+hu(i), \end{aligned}\right\} \tag{9.37}$$

$$J=\frac{h}{2}\sum_{i=0}^{k-1}u^2(i)\to\inf . \tag{9.38}$$

Отметим, что в основе перехода от непрерывной системы (9.34) к дискретной (9.37) лежит замена производной отношением приращений, т.е. разностей, поэтому получаемые соотношения (9.37) и вообще уравнения вида (9.1) называют разностными.

Для решения полученной задачи (9.37) – (9.38) воспользуемся дискретным принципом максимума (см. § 9.1).

Формируем функцию Гамильтона (9.3):

$$H=-\frac{h}{2}u^2(i)+\psi_1(i+1)\big[x_1(i)+h\,x_2(i)\big]+\psi_2(i+1)\big[x_2(i)+h\,u(i)\big].$$

Рассматриваем условие ее максимальности:

$$\frac{\partial H}{\partial u}=-h\,u(i)+\psi_2(i+1)\,h=0 .$$

В результате получаем соотношение, определяющее оптимальное управление:

$$u^*(i)=\psi_2(i+1). \tag{9.39}$$

Записываем уравнения (9.4), определяющие вспомогательные переменные ψ_i:

$$\left.\begin{aligned} \psi_1(i) &= \frac{\partial H}{\partial x_1}=\psi_1(i+1), \\ \psi_2(i) &= \frac{\partial H}{\partial x_2}=h\psi_1(i+1)+\psi_2(i+1). \end{aligned}\right\}$$

Первое соотношение показывает, что все значения переменной ψ_1 одинаковы. Обозначим их $\psi_1(i)=C=\text{const}$.

Второе уравнение запишем в виде $\psi_2(i+1)=\psi_2(i)-b$, где $b=hC$.

Полагая $i=0,\ 1,\ 2,...,$ получаем:

$$\psi_2(1) = \psi_2^0 - b; \; \psi_2(2) = \left(\psi_2^0 - b\right) - b = \psi_2^0 - 2b;$$

$$\psi_2(3) = \left(\psi_2^0 - 2b\right) - b = \psi_2^0 - 3b \text{ и т.д.}$$

Таким образом,

$$\psi_2(i) = \psi_2^0 - ib. \tag{9.40}$$

Рассмотрим теперь уравнения (9.37) для переменных состояния, используя полученные соотношения (9.39) и (9.40):

$$\left.\begin{aligned} &x_1(i+1) = x_1(i) + hx_2(i), \\ &x_2(i+1) = x_2(i) + h\left[\psi_2^0 - (i+1)b\right] = x_2(i) + a - ki, \end{aligned}\right\} \tag{9.41}$$

где

$$a = h\psi_2^0 - hb; \; k = hb. \tag{9.42}$$

Пусть $t_к = 1$, $h = 0{,}1$; $k = 10$. В таком случае получим результаты, которые приведены в табл. 9.1

Т а б л и ц а 9.1

i	$x_2(i+1) = x_2(i) + a - k\,i$	$x_1(i+1) = x_1(i) + h\,x_2(i)$
0	$x_2(1) = a$	$x_1(1) = x_1^н$
1	$x_2(2) = 2a - k$	$x_1(2) = x_1^н + h\,a$
2	$x_2(3) = 3a - 3k$	$x_1(3) = x_1^н + 3\,a\,h - kh$
3	$x_2(4) = 4a - 6k$	$x_1(4) = x_1^н + 6\,a\,h - 4kh$
4	$x_2(5) = 5a - 10k$	$x_1(5) = x_1^н + 10a\,h - 10kh$
5	$x_2(6) = 6a - 15k$	$x_1(6) = x_1^н + 15a\,h - 20kh$
6	$x_2(7) = 7a - 21k$	$x_1(7) = x_1^н + 21a\,h - 35kh$
7	$x_2(8) = 8a - 28k$	$x_1(8) = x_1^н + 28\,a\,h - 56kh$
8	$x_2(9) = 9a - 36k$	$x_1(9) = x_1^н + 36\,a\,h - 84kh$
9	$x_2(10) = 10a - 45k$	$x_1(10) = x_1^н + 45a\,h - 120kh$

Учитывая заданное конечное условие (9.35) и используя результаты, представленные в последней строке табл. 9.1, можно записать следующие соотношения:

$$\left.\begin{aligned} x_1(10) &= x_1^{\text{н}} + 45ah - 120kh = 0, \\ x_2(10) &= 10a - 45k = 0. \end{aligned}\right\}$$

Подставляя в эти соотношения выражение (9.42), после несложных преобразований получаем:

$$10h(\psi_2^0 - h\psi_1^0) - 45h^2\psi_1^0 = 0$$

или

$$\psi_2^0 = 5{,}5h\psi_1^0; \tag{9.43}$$

$$x_1^{\text{н}} + 45h^2\psi_2^0 - 165h^3\psi_1^0 = 0.$$

Используя полученный результат (9.43), находим:

$$\psi_2^0 = -\frac{5{,}5}{82{,}5}\frac{x_1^{\text{н}}}{h^2}. \tag{9.44}$$

Рассматривая (9.39), (9.40), (9.43) и (9.44), получаем следующее выражение для оптимального управления:

$$u^*(i) = \psi_2(i+1) = \left[\psi_2^0 - (i+1)h\psi_1^0\right] = x_1^{\text{н}}\left[-\frac{4{,}5}{82{,}5h^2} + i\frac{1}{82{,}5h^2}\right]. \tag{9.45}$$

Если учесть, что в наших расчетах $h = 0{,}1$, то выражение (9.45) примет вид

$$u^*(i) = \frac{x_1^{\text{н}}}{0{,}825}(-4{,}5 + i); \quad i = 0, 1, 2, \ldots, 9$$

или

$$u^*(t_i) = 5{,}454x_1^{\text{н}}(-1 + 2{,}222t_i); \; t_i \in \{0;\ 0{,}1;\ 0{,}2; \ldots;\ 0{,}9\}.$$

На рис. 9.2 этот результат представлен в графической форме. Приближенно он соответствует функции $u^*(t)$ и графику (см. рис. 8.3) для оптимального управления в исходной непрерывной системе, рассмотренной в § 8.3.

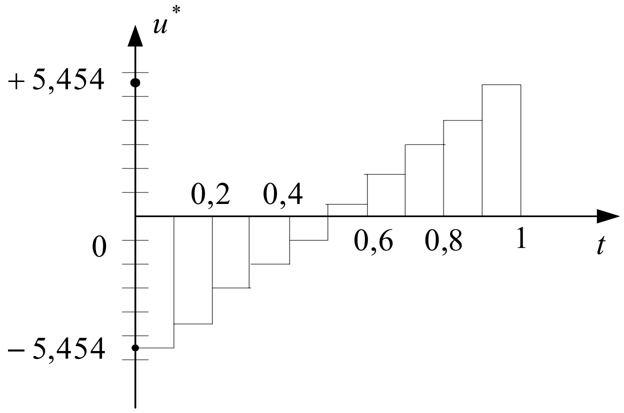


Рис. 9.2

В заключение отметим, что рассмотренный прием, позволяющий перейти от заданной непрерывной системы к дискретному прототипу на основе разностных схем, часто используется для приближенного решения практических задач об оптимальном управлении, так как решение исходных уравнений состояния совместно с сопряженной системой уравнений при заданных краевых условиях легче осуществить для дискретных соотношений, чем для непрерывных.

§ 9.4. Задача календарного планирования производства и поставки продукции

Рассмотрим задачу об оптимальном дискретном динамическом управлении экономического содержания, изложенную в [24]. Некое предприятие выпускает определенный вид продукции. Отдел маркетинга этого предприятия провел анализ имеющихся заявок на эту продукцию и другой информации, касающейся возможного ее сбыта на предстоящий плановый период, разбитый, например, на месяцы. В результате получен план прогнозируемого (возможного) спроса на продукцию в виде таблицы, например табл. 9.2, где t – номер планируемого этапа, например месяца; y – объем продукции, который можно продать в соответствующий месяц. При этом указанная таблица задает функцию спроса $y = y(t)$.

Таблица 9.2

t	0	1	2	3	4	5
y	1	2	1	5	4	8

В силу многих причин этот график может оказаться весьма неравномерным, и это создает для предприятия серьезные проблемы, поскольку для производства наиболее удобным и рациональным был бы равномерный график выпуска продукции. При этом удается избежать простоя рабочих или сверхурочной работы, возможного найма временных работников, длительного хранения продукции, подлежащей продаже в следующие месяцы, повышенного спроса на продукцию и т.д. Ясно, что перепроизводство, даже с условием последующей продажи, связано с дополнительными издержками, но и неудовлетворение спроса влечет за собой проблемы из-за уменьшения прибыли, возможной потери потребителей и т.п.

В этих условиях возникает задача об определении наиболее рационального календарного плана производства, при котором были бы учтены и минимизированы различные издержки и потери на планируемый период.

Для решения этой задачи, изложенной пока в словесной формулировке, необходимо, прежде всего, перейти к ее математической модели. Осуществим это, следуя [24]. Будем рассматривать процесс производства и сбыта продукции в дискретные моменты времени $t = 0,\ 1,\ 2,...,k$, где k – номер последнего этапа (месяца), определяющий продолжительность планового периода.

Введем обозначения:

$x = x(t)$ – количество продукции (в штуках, тоннах и т.д.), планируемой к выпуску в t-м месяце;

$y = y(t)$ – количество продукции, которое можно продать в соответствующем месяце;

$z = x(t) - y(t)$ – величина дефицита при $z < 0$ или перепроизводства при $z > 0$.

Как отмечалось выше, при $z \neq 0$ возникают дополнительные издержки: при $z > 0$ – из-за необходимости хранения продукции, поиска новых потребителей и т.д.; при $z < 0$ – из-за недополучения прибыли, возможной потери потребителей и т.д. Ясно, что масштабы этих потерь разные. Обозначим через $f_1 = f_1(z)$ функцию, описывающую указанные потери. Ее характер показан на рис. 9.3.

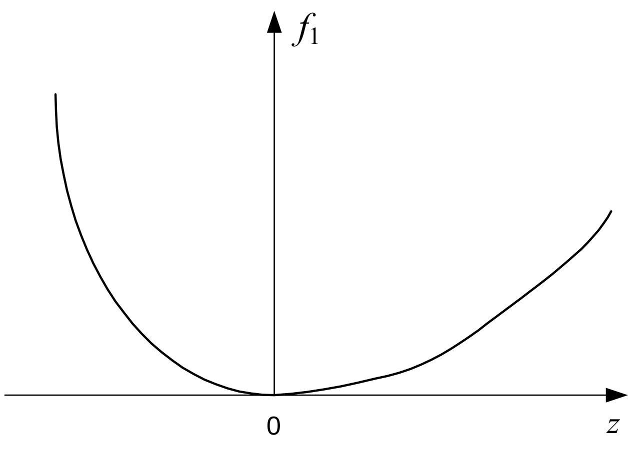


Рис. 9.3

При этом учтено, что рассматриваемая функция является несимметричной и крутизна положительной и отрицательной ветвей этого графика разная. Есть основания полагать [24], что масштабы потерь при $z < 0$ больше, чем при $z > 0$.

Будем аппроксимировать эту зависимость следующим образом:

$$f_1(z) = \begin{cases} a_1\, z^2 & \text{при } z > 0, \\ b_1\, z^2 & \text{при } z < 0,\ b_1 > a_1 > 0. \end{cases} \tag{9.46}$$

Как отмечалось, с точки зрения производства наиболее предпочтительным был бы его постоянный уровень интенсивности, когда $x(t+1) = x(t)$ для всех t.

Введем величину

$$u(t) = x(t+1) - x(t), \tag{9.47}$$

которая выражает увеличение (при $u > 0$) или снижение (при $u < 0$) объема выпускаемой продукции при переходе от одного планового этапа (месяца) к другому. Выше пояснялось, что при $u > 0$ и $u < 0$ возникают издержки. Их величину обозначим f_2. В результате получаем функцию потерь из-за неравномерности производства $f_2 = f_2(u)$. Ее характер показан на рис. 9.4. Как и функцию f_1, ее можно описать выражением:

$$f_2(z) = \begin{cases} a_2\, u^2 & \text{при } u > 0, \\ b_2\, u^2 & \text{при } u < 0, \end{cases} \tag{9.48}$$

хотя заранее нельзя сказать, какая из ветвей этих парабол будет круче.

Теперь введем показатель, учитывающий общие потери от обеих составляющих f_1 и f_2 на всем планируемом периоде:

$$J = \sum_t \left[f_1(z(t)) + f_2(u(t)) \right]. \tag{9.49}$$

В этом выражении суммирование должно осуществляться по всем этапам (месяцам), т.е. $t = 0, 1, \ldots, k$. Однако, как отмечается в [24], при $t = k$ значение $u(k) = x(k+1) - x(k)$, а значит, и $f_2(u(k))$ не определено, так как величина $x(k+1)$ выходит за пределы рассматриваемого периода. Поэтому показатель качества (9.49) запишем в следующем виде:

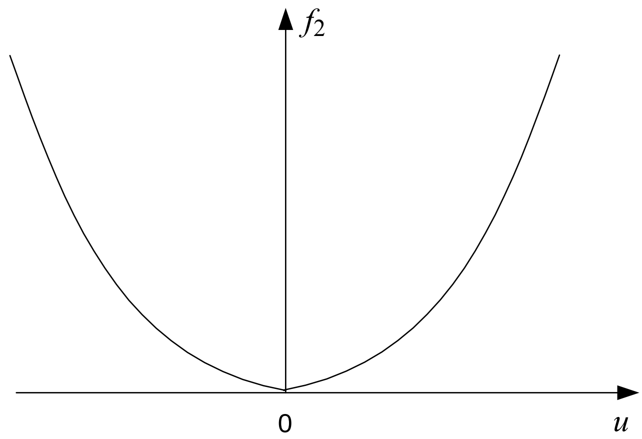


Рис. 9.4

$$J = \sum_{t=0}^{k-1} \left[f_1(x(t) - y(t)) + f_2(u(t)) \right] + f_1(x(k) - y(k)) \to \min. \tag{9.50}$$

При этом следует учитывать соотношение (9.47), т.е.

$$x(t+1) = x(t) + u(t). \tag{9.51}$$

Таким образом, в результате формализации рассматриваемая задача календарного планирования свелась к следующему: для заданного планового периода, задаваемой функции спроса $y = y(t)$, а также принятых функций потерь f_1 и f_2, определяемых (9.46) и (9.48), требуется найти функцию управления $u = u(t)$ и функцию состояния $x = x(t)$, при которых обобщенные потери, выражаемые функционалом (9.50), были бы минимальны. При этом процесс описывается уравнением (9.51) с начальным условием $x(0) = x_0$, а значение x в конце планового периода не задано (правый конец не зафиксирован).

Заметим, что в формулировке задачи было бы оправдано ввести условие $x(t) \geq 0$, так как поставка продукции не может быть отрицательной. Будем иметь в виду это ограничение, хотя специальных мер, исключающих возникновение таких недопустимых значений, принимать не будем. Это можно было бы сделать на основе идеи метода штрафных функций, но это значительно усложнит задачу.

Решение

1. Запишем функцию Гамильтона, используя модель (9.51) и критерий (9.50):

$$H = (-1)\left[f_1\left(x(t) - y(t)\right) + f_2\left(u(t)\right)\right] + \psi(t+1)\left[x(t) + u(t)\right]. \qquad (9.52)$$

2. Из условия $\dfrac{\partial H}{\partial u} = \psi(t+1) - \dfrac{\partial f_2(u(t))}{\partial u} = 0$, используя выражение (9.48) для f_2, находим:

$$\psi(t+1) = \begin{cases} 2a_2 u^*(t), \text{ если } u^*(t) \geq 0, \\ 2b_2 u^*(t), \text{ если } u^*(t) < 0. \end{cases} \qquad (9.53)$$

Заметим, что $a_2 > 0$ и $b_2 > 0$, поэтому из полученного соотношения (9.53) следует, что $u^*(t)$ и $\psi(t+1)$ будут иметь одинаковый знак. Поэтому из (9.53) можно получить следующее выражение, определяющее оптимальное управление:

$$u^*(t) = \begin{cases} \dfrac{1}{2a_2}\psi(t+1), \text{ если } \psi(t+1) \geq 0, \\ \dfrac{1}{2b_2}\psi(t+1), \text{ если } \psi(t+1) < 0. \end{cases} \qquad (9.54)$$

3. Уравнение (9.4) для сопряженной переменной

$$\psi(t) = \left.\frac{\partial H}{\partial x}\right|_{x^*(t)}$$

для функции (9.52) дает следующее выражение:

$$\psi(t) = \psi(t+1) - \frac{\partial f_1(x(t) - y(t))}{\partial x(t)},$$

которое с учетом функции (9.46), определяющей f_1, можно записать в виде:

$$\psi(t)=\psi(t+1)-2\begin{cases} a_1\left(x^*(t)-y(t)\right), \text{ если } x^*(t)\ge y(t), \\ b_1\left(x^*(t)-y(t)\right), \text{ если } x^*(t)< y(t). \end{cases} \tag{9.55}$$

4. Подставляем выражение (9.54) для u^* в уравнение модели управляемого процесса (9.51). В результате получаем:

$$x(t+1)=x(t)+\begin{cases} \dfrac{1}{2a_2}\psi(t+1), \text{ если } \psi(t+1)\ge 0, \\ \dfrac{1}{2b_2}\psi(t+1), \text{ если } \psi(t+1)< 0. \end{cases}$$

или

$$x(t)=x(t+1)-\begin{cases} \dfrac{1}{2a_2}\psi(t+1), \text{ если } \psi(t+1)\ge 0 \\ \dfrac{1}{2b_2}\psi(t+1), \text{ если } \psi(t+1)< 0. \end{cases} \tag{9.56}$$

5. Поскольку значение x на правом конце промежутка управления $0\le t\le k$ не задано, следует обратиться к условию трансверсальности (9.6):

$$\psi(k)=-\frac{\partial f_1}{\partial x}\bigg|_{x^*(k)}=-\begin{cases} 2a_1\left[x^*(k)-y(k)\right], \text{ если } x^*\ge y, \\ 2b_1\left[x^*(k)-y(k)\right], \text{ если } x^*< y. \end{cases} \tag{9.57}$$

Рассмотрим далее технику использования полученных расчетных соотношений на примере с конкретными числовыми данными: $k=5;\ x_0=1;\ a_1=1;\ b_1=2;\ a_2=b_2=3$.

Функция спроса задана в табл. 9.2. Расчеты необходимо провести, обеспечив погрешность в определении x не более $\varepsilon=0{,}1$.

Как уже отмечалось выше, особенностью предстоящих расчетов, выполняемых по соотношениям, полученным из принципа максимума, является то, что в ходе расчетов необходимо найти два процесса $x(t)$ и $\psi(t)$, по одному из них задано начальное условие $x(0)=x_0$, а по другому – конечное. Это серьезно затрудняет расчеты и приходится их осуществлять итерационно. При этом возможны два варианта проведения расчетов: можно отталкиваться от начального условия $x(0)=x_0$, последовательно увеличивая t, определить значения переменных при $t=k$ и добиваться, чтобы $\psi(k)$ получило нужное значение, а можно, отталкиваясь от значения переменных при $t=k$, двигаться вспять и добиваться нужного значения $x(0)$. Воспользуемся второй стратегией.

Итерация 1. По смыслу задачи ясно, что значения $x(t)$ не должны сильно отличаться от $y(t)$. Примем для начала, что при $t=5$ $x^*(5)=y(5)=8$. В этом случае по формуле (9.57) получаем $\psi(5)=0$. Полагая $t+1=5$, $t=4$, из формул (9.56) и (9.54) находим: $x^*(4)=8$, $u^*(4)=0$, $\psi(4)=-8$.

Аналогично, приняв $t+1=4$, $t=3$, вычисляем $x^*(3)$, $u^*(3)$, $\psi(3)$ и т.д. Выполнив все подобные действия, получаем $x^*(0)=30{,}58$. Сравнивая это значение с заданным значением $x(0)=1$, видим их несоответствие. Это делает необходимым выполнение следующей итерации с изменением $x^*(5)$.

Итерация 2. Уменьшим выбираемое значение $x^*(5)$, приняв $x^*(5)=6{,}125$.

Выполнив по изложенной выше схеме последовательно расчеты по всем этапам от последнего до начального, получим: $x^*(0)=3{,}923$.

Как видим, результат стал ближе к заданному значению $x(0)=1$, но еще существенно отличается от него.

Итерация 3. Примем $x^*(5)=6{,}02$, тогда в результате выполнения расчетов находим $x^*(0)=2{,}253$. Как видим, корректировка значений $x^*(5)$ привела к изменениям $x^*(0)$ в правильном направлении, но результат пока остается неприемлемым.

Итерация 4. Возьмем $x^*(5)=5{,}95$. Тогда, осуществляя расчеты, получаем $x^*(0)=1{,}053$. Этот результат можно принять, так как его отклонение от заданного значения укладывается в оговоренный размер погрешности: $|1-1{,}053|=0{,}053<0{,}1$.

Результаты расчетов величин x, ψ и u, полученные на этой итерации, приведены в табл. 9.3 [24]. Их можно принять как искомые оптимальные процессы.

Таблица 9.3

t	0	1	2	3	4	5
$y(t)$	1	2	1	5	4	8
$x^*(t)$	1,053	1,783	2,365	3,405	4,58	5,95
$\psi(t)$	4,35	4,37	3,5	6,23	7,04	8,2
$u^*(t)$	0,73	0,582	1,04	1,175	1,37	1,37

Контрольные вопросы

1. Запишите систему разностных уравнений в нормальной форме в развернутом и кратком векторно-матричных видах.
2. Запишите выражение, определяющее функционал качества, для дискретной динамической системы.
3. Как формулируется принцип максимума для дискретных систем?
4. Какой характер имеет принцип максимума: он определяет необходимые, достаточные или необходимые и достаточные условия оптимальности?
5. В каком случае используется и каким образом записывается условие трансверсальности для отыскания оптимального управления дискретной системой?
6. Запишите обобщенный суммарный квадратичный критерий качества для дискретной управляемой системы. Каким условиям удовлетворяют входящие в него матрицы?
7. Как определяется дискретный прототип непрерывной динамической системы?
8. Как формулируется задача календарного планирования производства и поставки продукции, каким образом записывается уравнение состояний и функционал качества для этой задачи?

Задачи для самостоятельного решения к главам 7 – 9

Используя изложенные в гл. 7 – 9 средства вариационного исчисления и принципа максимума, включая приближенные методы, необходимо решить нижеперечисленные непрерывные и дискретные задачи динамической оптимизации [24, 30, 31, 43]. Конкретный метод решения задачи следует выбрать самостоятельно или получить у преподавателя соответствующие указания.

1. Задачи управления непрерывными системами

1.1. $x'(t) = u(t)$; $x(0) = 1$; $x(4) = 0$; $|u| \le 1$;

$$J = \frac{1}{2}\int_0^4 x^2(t)dt \to \min.$$

1.2. $x'(t) = u(t)$; $x(0) = 0$; $x(1) = 1$; $|u| \le 1$;

$$J = -\int_0^1 \sqrt{1-u^2}\,dt \to \min .$$

1.3. $x'(t) = u(t)$; $x(0) = x^0$; $t_к$ – фиксировано;

$$J = \frac{1}{2}x^2(t_к) + \frac{1}{2}\int_0^{t_к} u^2 dt \to \min .$$

Найти $u = u(t)$ и $u = kx$.

1.4. $x' = -x + u$; $x(0) = 1$;

$$J = \frac{1}{2}\int_0^2 (x^2 + u^2)dt \to \min .$$

Найти управление в виде $u = kx$ и определить k.

1.5. $x' = -x + u$; $x(0) = 1$; $x(t_к) = 0$; $t_к$ – не зафиксировано;

$$J = \int_0^{t_к} \left[\alpha + u^2 + x^2\right] dt \to \min .$$

1.6. $x' = -x + u$; $x(0) = 10$; $x(1) = 0$;

$$J = \frac{1}{2}\int_0^1 (u')^2 dt \to \min .$$

1.7. $x' = \frac{1}{2}x + u$; $x(0) = x^0$;

$$J = \frac{1}{2}x^2(t_к) + \frac{1}{2}\int_0^{t_к} \left(2x^2 + u^2\right) dt \to \min .$$

1.8. $x' = -ax + bu$; $a > 0$; $b > 0$; $x(0) = x^0$; $x(\infty) = 0$;

$$J = \int_0^{\infty} \left(x^2 + u^2\right) dt \to \min .$$

1.9. $x' = ax + bu$; $a > 0$; $b > 0$; $x(0) = x^0$; $x(t_к) = 0$;

$$J = \int_0^{t_к} u^2 dt \to \min .$$

1.10. $x' = -x + 2u$; $0 \le u \le 1$; $x(0) = 1$; $x(10) = 0$;

$$J = \int_0^{10} (x + u) dt \to \min .$$

1.11. $x' = x + u$; $|u| \le 4$; $x(0) = 1$;

$$J = \int_0^{10} (u^2 + x)dt \to \min.$$

1.12. $x' = x + u$; $x(0) = 1$;

$$J = \int_0^{4} (u + u^2 + 2x^2)dt \to \min.$$

1.13. $x' = 2x + u$; $|u| \le 1$; $x(0) = 1$;

$$J = -2x(4) + \int_0^{4} (x + 5u)dt \to \min.$$

1.14. $x' = x + 2u$; $|u| \le 1$; $x(0) = 1$;

$$J = 2x(3) + \int_0^{3} (x + 6u)dt \to \min.$$

1.15. $x' = x + u$; $|u| \le 2$; $x(0) = 1$;

$$J = -3x(10) + \int_0^{10} (x + 2u + u^2)dt \to \min.$$

1.16. $x' = -2x + u$; $|u| \le 1$; $x(0) = 1$; $x(10) = 0$;

$$J = \int_0^{10} (2x + u)dt \to \min.$$

1.17. $x' = -\ 2x + u$; $|u| \le 1$; $x(0) = 1$;

$$J = 2x^2(10) + \int_0^{10} (2x + u)dt \to \min.$$

1.18. $x' = -x + u$; $|u| \le 1$; $x(0) = 2$;

$$J = x^2(5) + \int_0^{5} (u - x)dt \to \min.$$

1.19. $\left.\begin{aligned} x_1' &= x_2, \\ x_2' &= u; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= 0, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(1) &= a, \\ x_2(1) &= 0; \end{aligned}\right\}$

$$J = \int_0^{1} u^2 dt \to \min.$$

1.20. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = 0, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(1) = a, \\ x_2(1) - \text{свободно}; \end{array}\right\}$

$$J = \int_0^1 u^2 dt \to \min .$$

1.21. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = 0, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(t_\text{к}) - \text{свободно}, \\ x_2(t_\text{к}) = 0; \end{array}\right\}$

$$J = t_\text{к} \to \max .$$

1.22. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $|u| \le U_0$; $\left.\begin{array}{l} x_1(0) = 0, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(1) = a, \\ x_2(1) = 0; \end{array}\right\}$

$$J = t_\text{к} \to \min .$$

1.23. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = 0, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(t_\text{к}) = a, \\ x_2(t_\text{к}) = 0; \end{array}\right\}$ $|u| \le U_0$;

$$\int_0^{t_\text{к}} u^2 dt \le b; \quad J = t_\text{к} \to \min$$

1.24. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = 1, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(t_\text{к}) = 0, \\ x_2(t_\text{к}) = 0; \end{array}\right\}$ $|u| \le 1$;

$$J = \int_0^{t_\text{к}} \left(x_1^2 + x_2^2\right) dt \to \min ; \quad t_\text{к} = 5 .$$

1.25. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = 1, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(t_\text{к}) = 0, \\ x_2(t_\text{к}) = 0; \end{array}\right\}$ $|u| \le 1$;

$$J = \int_0^{t_\text{к}} \left(x_1^2 + x_2^2\right) dt \to \min ; \quad t_\text{к} - \text{не зафиксировано.}$$

1.26. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = a_1, \\ x_2(0) = a_2; \end{array}\right\}$ $x_1^2(t_\text{к}) + x_2^2(t_\text{к}) = 1$;

$$J = t_\text{к} \to \min ; \quad t_\text{к} - \text{не зафиксировано.}$$

1.27. $y'' = k_1 u, \quad y(0) = y_0, \quad y'(0) = u_0;$

$$\int_0^\infty u^2(t) dt = k_2 ; \quad J = \int_0^\infty [1 - y(t)]^2 \, dt \to \min.$$

1.28. $\left.\begin{array}{l} x_1' = x_2, \\ x_2' = -x_1 + u; \end{array}\right\}$ $\left.\begin{array}{l} x_1(0) = x_1^\text{н}, \\ x_2(0) = 0; \end{array}\right\}$ $\left.\begin{array}{l} x_1(t_\text{к}) = 0, \\ x_2(t_\text{к}) = 0; \end{array}\right\}$ $|u| \le 1$;

$$J = t_\text{к} \to \min .$$

1.29. $\left.\begin{aligned} x_1' &= x_2, \\ x_2' &= -2x_1 - 3x_2 + 2u; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= a_1, \\ x_2(0) &= a_2; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(t_к) &= 0, \\ x_2(t_к) &= 0; \end{aligned}\right\}$ $|u| \le 1$;

$$J = t_к \to \min .$$

Условия переключения управления интерпретировать графически.

1.30. $\left.\begin{aligned} x_1' &= -2x_1 + 2x_2, \\ x_2' &= -x_2 + u; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= a_1, \\ x_2(0) &= a_2; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(t_к) &= 0, \\ x_2(t_к) &= 0; \end{aligned}\right\}$ $|u| \le 1$;

$$J = t_к \to \min .$$

Условия переключения управления интерпретировать графически.

1.31. $\left.\begin{aligned} x_1' &= x_2 + u_1, \\ x_2' &= u_2; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= 1, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} &x_1(2) = 0, \\ &x_2(2) - \text{свободно}; \end{aligned}\right\}$

$$J = \frac{1}{2}\int_0^2 \left(u_1^2 + u_2^2\right) dt \to \min .$$

1.32. $\left.\begin{aligned} x_1' &= x_2 + u_1, \\ x_2' &= -x_1 + u_2; \end{aligned}\right\}$ $\left.\begin{aligned} |u_1| &\le 1, \\ |u_2| &\le 2; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= 1, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(t_к) &= 0, \\ x_2(t_к) &= 0; \end{aligned}\right\}$

$$J = t_к \to \min .$$

Найти кривые переключений и указать полярность u_1, u_2 для различных x_1, x_2.

1.33. $\left.\begin{aligned} x_1' &= x_2 - u, \\ x_2' &= x_1 + u; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= 2, \\ x_2(0) &= 0; \end{aligned}\right\}$ $|u| \le 2$;

$$J = -x_2(3) + \int_0^3 (x_1 + x_2 + 2u)dt \to \min .$$

1.34. $\left.\begin{aligned} x_1' &= x_2, \\ x_2' &= -2x_1 + 10u; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= x_1^н, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(t_к) &= 0, \\ x_2(t_к) &= 0; \end{aligned}\right\}$ $|u| \le 1$;

$$J = t_к \to \min .$$

1.35. $\left.\begin{aligned} x_1' &= x_2, \\ x_2' &= -2x_1 + 10u; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= x_1^н, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(\infty) &= 0, \\ x_2(\infty) &= 0; \end{aligned}\right\}$ $|u| \le 2$;

$$J = \int_0^\infty (x_1^2 + qx_2^2 + ru^2)dt \to \min .$$

2. Задачи управления дискретными системами

2.1. $x(i+1) = -x(i) + 2u(i)$; $x(0) = 1$; $x(5) = 0$;

$$J = \sum_{i=0}^{4}\left[x(i) + 2u(i) + u^2(i)\right] \to \min .$$

2.2. $x(i+1) = x(i) + 0{,}1u(i)$; $x(0) = 1$; $x(10) = 0$; $|u| \le 1$;

$$J = \sum_{i=0}^{9}\left[x^2(i) + u^2(i)\right] \to \min .$$

2.3. $x(i+1) = 0{,}9\,x(i) + 0{,}2\,u(i)$; $x(0) = 1$; $x(10) = 0$; $0 \le u \le 1$;

$$J = \sum_{i=0}^{9}\left[x(i) + u(i)\right] \to \min .$$

2.4. $x(i+1) = -x(i) + u(i)$; $x(0) = 0$;

$$J = x(5) + \sum_{i=0}^{4}\left[x^2(i) + u^2(i)\right] \to \min .$$

2.5. $x(i+1) = 0{,}5x(i) + u(i)$; $x(0) = 2$;

$$J = -x^2(4) + \sum_{i=0}^{3}\left[4x^2(i) - x(i)\right] \to \min .$$

2.6. $x(i+1) = -0{,}5\,x(i) + 2\,u(i)$; $x(0) = 1$;

$$J = 3x^2(4) + \sum_{i=0}^{3}\left[x(i) + u(i)\right] \to \min .$$

2.7. $x(i+1) = 0{,}2\,x(i) + u(i)$; $x(0) = 1$;

$$J = \sum_{i=0}^{4} u^2(i) + 5x(5) \to \min .$$

2.8. $\left.\begin{aligned} x_1(i+1) &= x_1(i) + 0{,}1x_2(i), \\ x_2(i+1) &= x_2(i) + 0{,}1u(i); \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= 0, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} &x_1(10) = 1, \\ &x_2(10) - \text{свободно}; \end{aligned}\right\}$

$$J = \sum_{i=0}^{9} u^2(i) \to \min .$$

2.9. $\left.\begin{aligned} x_1(i+1) &= x_1(i) + 0{,}1x_2(i), \\ x_2(i+1) &= x_2(i) + 0{,}1u(i); \end{aligned}\right\}$ $\left.\begin{aligned} x_1(0) &= 1, \\ x_2(0) &= 0; \end{aligned}\right\}$ $\left.\begin{aligned} x_1(k) &= 0, \\ x_2(k) &= 0; \end{aligned}\right\}$ $|u| \le 1$;

$$J = k \to \min .$$

ЗАКЛЮЧЕНИЕ

На основе изложенного выше теоретического и практического материала можно сделать следующие **обобщения** и **выводы**:

1. Решение прикладных, в том числе экономических, задач оптимизационного характера осуществляется по следующей общей схеме:
   - излагается сущность решаемой задачи в словесной форме:
     - с описанием анализируемой системы или процесса;
     - с выделением факторов и параметров, которые можно считать известными (заданными);
     - с перечислением переменных, на которые накладываются ограничения;
     - с указанием показателя, который нужно минимизировать или сделать как можно больше.

     Этот этап во многом предопределяет выбор метода решения и получаемый результат. Например, для одной и той же системы, рассмотренной в § 8.3, 8.4, при разных показателях, принятых в качестве критерия оптимальности, получены совершенно разные оптимальные управления;
   - выполняется формализация поставленной задачи, т.е. получение математической модели;
   - проводится анализ этой модели и отнесение ее к соответствующему типу классических задач;
   - осуществляется выбор соответствующего этому классу задач метода решения и получение последовательности (блок-схемы) расчетных соотношений;
   - составляется программа, реализующая вычисления по этой блок-схеме и расчетным соотношениям;
   - проводятся вычисления для заданных исходных условий;
   - осуществляется интерпретация полученных результатов.

     Внимательный анализ этих результатов иногда показывает на необходимость пересмотра или уточнения исходных условий либо на целесообразность использования полученной вычислительной схемы или программы для исследования чувствительности решения к изменению тех или иных параметров и т.д.
2. Не существует универсального алгоритма, позволяющего решать любые оптимизационные задачи, поэтому для каждого класса задач (линейного программирования, транспортных задач, сетевого планирования, динамического программирования и т.д.) разработан свой специфический алгоритм решения.

3. Существуют приемы, позволяющие задачу одного типа сводить к задаче другого типа, например, транспортные или сетевые задачи рассматривать как задачи линейного программирования и как задачи динамического программирования; задачи математического программирования, используя метод штрафных функций, рассматривать как задачи безусловной оптимизации; от задачи отыскания непрерывного динамического управления перейти к анализу дискретного прототипа и т.д. Это дает возможность для выбора наиболее удобного и эффективного подхода к решению в тех или иных условиях, например, в зависимости от имеющегося в наличии программного обеспечения.
4. Практически все изложенные в книге алгоритмы решения оптимизационных задач предполагают многократное выполнение одних и тех же математических операций с изменяющимися каждый раз данными. Такие вычислительные процессы называют итерационными. При правильной организации и настройке алгоритма от итерации к итерации решение будет все больше приближаться к искомому результату. Характерно, что некоторые алгоритмы, например, симплекс-метод для задач ЛП, алгоритм Дейкстры для задачи о кратчайших путях в графе, алгоритм Форда – Фалкерсона для задачи о максимальном потоке в транспортной сети и др., дают искомое оптимальное решение за конечное число итераций. Существуют задачи и алгоритмы, для которых точное решение задачи можно получить только при реализации бесконечного количества итераций, например, при отыскании экстремума нелинейной целевой функции. Правда и в этом случае для определенного допустимого уровня погрешности $\varepsilon > 0$ количество необходимых для получения решения итераций будет конечным.
5. Количество итераций, необходимых для получения искомого решения, существенно зависит от исходного (стартового) решения, выбранного в качестве начального. Часто его берут нулевым, см., например, алгоритм Форда – Фалкерсона (§ 4.5), но это, конечно, необязательно.
6. Опыт решения практических задач, а также изучение специальной литературы, в которой такой опыт аккумулирован, показывает, что, осуществляя вычисления по выбранному алгоритму для конкретной практической задачи, не всегда удается получить искомое решение, несмотря на то, что алгоритм представляется достаточно ясным, а в его реализации не допущено ошибок. Причиной такой неудачи могут быть многочисленные «подводные камни», которые не просматривались при изучении алгоритма решения и о сущест-

вовании которых трудно догадаться. Речь идет о таких факторах, как многоэкстремальность, вырожденность, очень высокая чувствительность к исходным данным, плохая обусловленность [15, 32, 37]. Часто такие факторы рассматривают как разного рода проявления так называемой некорректности задач [15]. Тем, кто знает о существовании таких «подводных камней», но не имеет еще достаточного практического опыта в решении задач, обычно кажется, что они являются какой-то экзотикой и встреча с ними представляется маловероятной. Однако, когда им приходится исследовать свою конкретную задачу, часто возникают указанные препятствия.

Автор имел возможность много раз убедиться в этом. Например, при отыскании наименьшего значения целевой функции, возникшей в одной прикладной задаче, одна и та же программа, реализованная на различных вычислительных машинах, при одних и тех же исходных данных, дала существенно различные результаты. Тщательный анализ ситуации позволил понять, что решаемая задача является плохо обусловленной и разницы в длине машинных слов, с которыми работали эти машины, оказалось достаточно для того, чтобы спровоцировать значительную разницу в получаемом решении.

7. Изложенные обстоятельства дают основания для того, чтобы согласиться с мнением [32], в соответствии с которым оптимизацию следует рассматривать и как науку, и как искусство. Безусловно, в основе решения любой оптимизационной задачи лежит научная база, связанная с построением математической модели, выбором и реализацией математического метода, использующего эту модель. Однако успех в использовании математического аппарата от формулировки задачи до получения решения во многом зависит от творческих способностей, интуиции, опыта специалиста, решающего задачу, в том числе и умения «обходить подводные камни», а это уже искусство.

РЕКОМЕНДУЕМЫЕ РАЗДЕЛЫ ДЛЯ САМОСТОЯТЕЛЬНОГО ИЗУЧЕНИЯ

К главе 1

[1]: § 1.1 – задачи об оптимальном планировании производства, задача об оптимальном раскрое ткани, задача о перевозках, задача о размещении предприятий, задача о распределении удобрений по посевным площадям, задача об оптимальном составе смеси.

[14]: гл. II – балансовые модели; гл. VI – транспортные задачи; гл. VII – задачи об оптимальном распределении взаимозаменяемых ресурсов, задачи об оптимальных назначениях; гл. VIII – задача об оптимальном планировании производства, задачи об оптимальном рационе и задачи о смесях, задача об оптимальном ассортименте, задача о раскрое, задача об оптимальном регулировании запасов; гл. X – целочисленные задачи оптимального планирования; гл. XI – нелинейные задачи оптимального планирования.

[20]: гл. 20, 21 – задачи оптимального планирования производства; гл. 23 – транспортные задачи; гл. 26 – задача о назначениях.

К главе 2

[1]: гл. 1 – общая формулировка и примеры задач линейного программирования, методика их решения, в том числе симплекс-метод, двойственные задачи и двойственный симплекс-метод, большой набор задач для самостоятельного решения и примеры решения; гл. 2 – специальные задачи ЛП, включая транспортные и целочисленные задачи, задачи параметрического, дробно-рационального и блочного программирования.

[15]: гл. 4 – 5 – общая теория линейного программирования, особо рассмотрены различные аспекты реализации вычислений по симплекс-методу, в том числе проблемы врожденности, рассмотрена методика декомпозиции задач ЛП.

[11]: гл. 2 – задачи ЛП, симплекс-метод, метод обратной матрицы и декомпозиционные методы решения ЗЛП большой размерности, многокритериальные ЗЛП.

[32]: гл. 2 – 7 – формулировка и примеры задач ЛП, градиентное решение и симплекс-метод, особые случаи в применении симплекс-метода, использование симплекс-таблиц для исследования модели ЗЛП на чувствительность к изменениям запасов ресурсов и вариации коэффициентов целевой функции, экономическая интерпретация двойственности, связь ЛП с транспортными и сетевыми задачами, матричное представление ЗЛП, модифицированный симплекс-метод, декомпозиция в ЛП, параметрическое ЛП, целочисленное ЛП, в том числе метод ветвей и границ, метод отсекающих плоскостей и булево программирование.

[3]: гл. 1 – 5 – общая теория ЛП и симплекс-метод, алгоритмы решения ЗЛП на базе симплекс-метода и листинги программ, реализующих эти алгоритмы на Бэйсике.
[14]: гл. II – V – применение матричной алгебры в экономических расчетах, балансовые модели, симплекс-метод, теория двойственности и экономическая интерпретация двойственных задач; большой набор задач для самостоятельного решения с примерами решений и ответами.

К главе 3

[6]: гл. 2 – различные формулировки транспортных задач и их решение на основе метода потенциалов.
[14]: гл. VI – решение транспортных задач методом потенциалов и распределительным методом, открытые модели ТЗ.
[23]: гл. 1 – решение ТЗ методом потенциалов и дельта-методом, приложения ТЗ к решению различных экономических задач – об увеличении производительности автомобильного транспорта за счет минимизации порожнего пробега, об оптимальных назначениях, об оптимизации суммарных расходов предприятия на производство и транспортировку продукции, об оптимальном размещении предприятий, при котором суммарные затраты на доставку сырья, производство продукции и ее транспортировку были минимальными.

К главе 4

[18]: гл. 2 – основные определения теории графов (кратко), задачи о кратчайшем пути в графе и графе наименьшей длины, а также задачи на транспортных сетях.
[21]: гл. 1 – 7 – основные понятия теории графов (подробно); гл. 8 – задачи о кратчайших путях; гл. 10 – задача о кратчайшем остове, а также задача коммивояжера и задача о назначениях; гл. 11 – потоки в сетях.
[22]: гл. 2 – основные определения теории графов, задача о кратчайших путях в графе, задачи на транспортных сетях.

[25]: гл. 1 – основные понятия; гл. 2 – алгоритмы построения деревьев; гл. 3 – алгоритмы поиска путей; гл. 4 – потоковые алгоритмы; гл. 5 – 8 – задачи о паросочетаниях, задача почтальона, задача коммивояжера и задачи размещения; гл. 9 – сетевые графики.
[29]: гл. 1 – 10 – подробное изложение основных понятий теории графов; гл. 15 – оптимизационные алгоритмы, включая алгоритмы отыскания кратчайших путей, отыскания оптимальных деревьев; решение задач об оптимальных назначениях и составлении расписаний, решения оптимизационных задач в транспортных сетях.

[36]: гл. 2 – задачи, интерпретируемые на графах, их связь с линейным программированием, алгоритмы решения, в частности, для задачи о кратчайших путях – алгоритмы Дейкстры, Флойда и метод двойного поиска, приведены оценки вычислительной сложности алгоритмов; гл. 4 – 5 – использование сетевых графиков для управления проектами; в приложении – описание и листинги программ, входящих в пакет сетевой оптимизации.

К главе 5

[15]: гл. 7 – 11 – общая теория нелинейных оптимизационных задач, различные алгоритмы одномерной и многомерной оптимизации, проблемы сходимости процессов поиска экстремума, понятие о корректных и некорректных задачах, методы регуляризации.

[26]: гл. I, II, IV, V – алгоритмы поиска экстремума при отсутствии и наличии ограничений.

[8]: гл. 3 – теоретические основы нелинейного программирования, включая метод множителя Лагранжа и теорему Куна – Таккера, квадратичное программирование и метод Вольфа, задачи с сепарабельными целевыми функциями.

[1]: гл. 3 – экономическая и графическая интерпретация задач нелинейного программирования, градиентные методы, решение задач с сепарабельными ЦФ, поиск оптимального решения при наличии ограничений методом Франка – Вульфа, методом штрафных функций, методом Эроу – Гурвица, большой набор задач для самостоятельного решения.

[2]: методы и алгоритмы поиска оптимума ЦФ, листинги программ, решающих эти алгоритмы на Бэйсике.

[37]: подробное описание алгоритмов поиска экстремума при отсутствии и наличии ограничений, листинги оптимизационных программ, результаты исследований эффективности различных алгоритмов.

[11]: гл. 5 – алгоритмы поиска экстремума функции без ограничений и при наличии ограничений, задачи и методы квадратичного и геометрического программирования.

К главе 6

[11]: гл. 6 – основы динамического программирования и специальные задачи ДП, включая задачи управления запасами, бесконечные задачи, сетевые задачи ДП, задачи ДП на марковских цепях.

[12]: ч. III, разд. 9 – методология и основные принципы ДП, модель ДП в условиях неопределенности.

[13]: гл. I – модель и вычислительная схема ДП; гл. II – решение задач ДП о распределении ресурсов; гл. III – решение задач ДП об управлении запасами; гл. IV – решение задач о замене; гл. V – анализ специальных задач

ДП (задач с мультипликативным критерием, целочисленные задачи ДП, задач о маршрутизации, стохастических задач).
[5]: гл. 8 – 12 – общая методология ДП, решение различных задач ДП (об управлении запасами, распределении усилий и др.), анализ вычислительных возможностей и области эффективного применения метода ДП.
[32]: гл. 9 – принципы построения и примеры моделей ДП, проблемы практического применения ДП, в том числе проблема размерности, применение методологии ДП для решения задач линейного программирования.

К главе 7

[22]: гл. 11 – задачи Лагранжа, Майера и Больца, уравнение Эйлера, условие Лежандра, условие трансверсальности, задачи с вырожденными функционалами и их решение с использованием метода Кротова; приближенные методы решения задачи вариационного исчисления (метод Ритца и метод Эйлера).
[30]: гл. 3 – классические задачи Лагранжа, Больца и Майера, необходимые условия в форме уравнения Эйлера – Лагранжа и условия трансверсальности, решение вариационных задач с дополнительными ограничениями в форме равенств и неравенств; примеры решений, задачи для самостоятельного решения.
[31]: гл. 13 – основные понятия вариационного исчисления, классические задачи и необходимые условия в форме уравнения Эйлера – Лагранжа, условие Лежандра, условия Вейерштрасса; задачи с ограничениями типа равенств и неравенств, применение аппарата ВИ для решения задач оптимального управления; примеры решений и задачи для самостоятельного решения.
[34]: гл. 10, §10.2 – вариационные задачи с закрепленными концами и фиксированным временем; задачи с подвижными концами и нефиксированным временем, уравнение Эйлера, использование множителей Лагранжа для решения вариационных задач с ограничениями.

К главе 8

[4]: гл. II – формулировка и доказательство принципа максимума, его связь с динамическим программированием; гл. III – линейные задачи об оптимальном управлении и принцип максимума как необходимое и достаточное условие оптимальности, решение различных задач об оптимальных быстродействиях, вычислительные методы решения этих задач.
[22]: гл. 12 – принцип максимума как обобщение и расширение классического вариационного исчисления, вывод уравнений принципа максимума и уравнений Гамильтона – Эйлера, применение принципа максимума для решения некоторых частных задач динамической оптимизации.

[24]: гл. 6 – формулировка принципа максимума, ПМ как необходимое и достаточное условие оптимальности, примеры применений ПМ для решения различных задач, задачи для самостоятельного решения; гл. 8 – применение ПМ для решения задачи о календарном планировании поставки продукции и задачи об оптимальном потреблении в однопродуктовой макроэкономической модели.

[30]: гл. 4 – связь принципа максимума с вариационным исчислением и с непрерывным динамическим программированием.

[31]: гл. 14 – формулировка и доказательство принципа максимума, его применение для решения задач об оптимальном быстродействии, управлении линейными системами, оптимальном управлении расходом топлива, задач Лагранжа, Майера, управлении нестационарными системами, оптимальном управлении при ограничениях на переменные состояния, 7 примеров решения задач на применение ПМ и 16 задач для самостоятельного решения.

[3]: § 10.3 – формулировка ПМ, его связь с ВИ, использование ПМ для отыскания оптимального управления линейной системой, решение задачи об оптимальном быстродействии, теорема об n интервалах, вырожденные и особые задачи динамической оптимизации.

К главе 9

[22]: §12 – 14 – формулировка дискретного варианта принципа максимума, задача об управлении линейной дискретной системой и принцип максимума как необходимое и достаточное условие оптимального управления, применение дискретного ПМ для решения транспортных задач.

[24]: гл. 7 – формулировка задачи об оптимальном многошаговом управлении, условия оптимальности при отсутствии и при наличии ограничений на управление.

[30] § 6.2 – 6.3 – формулировка дискретного ПМ, его применение для системы управления, оптимального по обобщенному суммарному квадратичному показателю качества, сравнение дискретного и непрерывного ПМ.

К задачам для самостоятельного решения к гл. 7 – 9

[24]: с. 83 – 85, с. 128 – непрерывные задачи оптимального управления (19 задач); с. 95 – 96, с. 128 – дискретные задачи оптимального управления (11 задач).

[30]: с. 47 – 49, с. 77 – 80 и 108 – 111 – непрерывные задачи оптимального управления (45 задач); с. 127 – 128 – дискретные задачи оптимального управления (12 задач).

[31]: § 12.8, 13.7, 14.9, 15.9, 16.8 – непрерывные задачи об оптимальном управлении (66 задач).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. **Акулич, И. Л.** Математическое программирование в примерах и задачах / И. Л. Акулич. – М. : Высш. шк., 1986. – 319 с.
2. **Банди, Б.** Методы оптимизации. Вводный курс / Б. Банди. – М. : Радио и связь, 1988. – 126 с.
3. **Банди, Б.** Основы линейного программирования / Б. Банди. – М. : Радио и связь, 1989. – 176 с.
4. **Болтянский, В. Г**. Математические методы оптимального управления / В.Г. Болтянский. – М. : Наука, 1966. – 307 с.
5. **Вагнер, Г.** Основы исследования операций : в 2 т. / Г. Вагнер. – М. : Мир, 1973. – Т. 2. – 488 с.
6. **Вентцель, Е. С.** Исследование операций / Е. С. Вентцель. – М. : Сов. радио. – 1972. – 552 с.
7. **Гилл, Ф.** Практическая оптимизация / Ф. Гилл, У. Мюррей, М. Райт. – М. : Мир, 1985. – 510 с.
8. **Дегтярев, Ю. И.** Исследование операций / Ю. И. Дегтярев. – М. : Высш. шк., 1986. – 320 с.
9. **Дегтярев, Ю. И.** Методы оптимизации / Ю. И. Дегтярев. – М. : Сов. радио, 1980. – 416 с.
10. **Деордица, Ю. С.** Исследование операций в планировании и управлении / Ю. С. Деордица, Ю. М. Нефедов. – Киев : Выща шк., 1991. – 270 с. – ISBN 5-11-002581-9.
11. **Зайченко, Ю. П.** Исследование операций / Ю. П. Зайченко. – Киев : Выща шк., 1988. – 552 с. – ISBN 5-11-002226-6.
12. Исследование операций : в 2 т. / под. ред. Дж. Моудера, С. Элмаграби. – М. : Мир, 1981. – Т. 1. – 712 с.
13. **Калихман, И. Л.** Динамическое программирование в примерах и задачах / И. Л. Калихман, М .А. Войтенко. – М. : Высш. шк., 1979. – 125 с.
14. **Калихман, И. Л.** Сборник задач по математическому программированию / И. Л. Калихман. – М. : Высш. шк., 1975. – 270 с.
15. **Карманов, В. Г.** Математическое программирование / В. Г. Карманов. – М. : Наука, 1986. – 288 с.
16. **Конюховский, В. П.** Математические методы исследования операций в экономике / В. П. Конюховский. – СПб. : Питер, 2000. – 208 с. – ISBN 5-8046-0190-3.
17. **Корн, Г.** Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М. : Наука, 1973. – 831 с.
18. **Коршунов, Ю. М.** Математические основы кибернетики / Ю. М. Коршунов. – М. : Энергия, 1980. – 424 с.
19. **Красс, М. С.** Математика для экономистов / М. С. Красс, Б. П. Чупрынов. – СПб. : Питер, 2004. – 464 с. – ISBN 5-94723-672-9.

20. **Красс, М. С.** Основы математики и ее приложения в экономическом образовании / М. С. Красс, Б. П. Чупрынов. – М. : Дело, 2002. – 688 с. – ISBN 5-7749-0186-6.

21. **Кристофидес, Н.** Теория графов. Алгоритмический подход / Н. Кристофидес. – М. : Мир, 1978. – 432 с.

22. **Кузин, Л. Т.** Основы кибернетики / Л. Т. Кузин. – М. : Энергия, 1973. – 504 с.

23. **Кузнецов, Ю. Н.** Математическое программирование / Ю. Н. Кузнецов, В. И. Кузубов, А. Б. Волощенко. – М. : Высш. шк., 1980. – 300 с.

24. **Лагоша, Б. А.** Оптимальное управление в экономике / Б. А. Лагоша. – М. : МГУЭСИ, 2004. – 128 с. – ISBN 5-7764-0392-8.

25. **Майника, Э.** Алгоритмы оптимизации на сетях и графах / Э. Майника. – М. : Мир, 1981. – 322 с.

26. **Моисеев, Н. Н.** Методы оптимизации / Н. Н. Моисеев, Ю. П. Иванилов, Е. М. Столярова. – М. : Наука, 1978. – 352 с.

27. **Ногин, В. Д.** Основы теории оптимизации / В. Д. Ногин, И. О. Протодьяконов, И. И. Евлампиев. – М. : Высш. шк., 1986. – 384 с.

28. Основы кибернетики. Математические основы кибернетики / под. ред. К. А. Пупкова. – М. : Высш. шк., 1974. – 413 с.

29. **Свами, М.** Графы, сети и алгоритмы / М. Свами, К. Тхуласираман. – М. : Мир, 1984. – 454 с.

30. **Сейдж, Э. П.** Оптимальное управление системами / Э. П. Сейдж, Ч. С. Уайт. – М. : Радио и связь, 1982. – 392 с.

31. **Сю, Д.** Современная теория автоматического управления и ее применение / Д. Сю, А. Мейер. – М. : Машиностроение, 1972. – 544 с.

32. **Таха, Х.** Введение в исследование операций : в 2 кн. / Х. Таха. – М. : Мир, 1985. – Кн. 1. – 479 с.

33. **Таха, Х.** Введение в исследование операций : в 2 кн. / Х. Таха. – М. : Мир, 1985. – Кн. 2. – 496 с.

34. Теория автоматического управления. В 2 ч. Ч. 2. Теория нелинейных и специальных систем автоматического управления / под. ред. А. А. Воронова. – М. : Высш. шк., 1986. – 504 с.

35. **Федоренко, Р. П.** Приближенное решение задач оптимального управления / Р. П. Федоренко. – М. : Наука, 1978. – 488 с.

36. **Филлипс, Д.** Методы анализа сетей / Д. Филлипс, А. Гарсиа-Диас. – М. : Мир, 1984. – 496 с.

37. **Химмельблау, Д.** Прикладное нелинейное программирование / Д. Химмельблау. – М. : Мир, 1975. – 534 с.

38. **Охорзин, В.А.** Оптимизация экономических систем. Примеры и алгоритмы в среде Mathcad : учеб. пособие / В.А. Охорзин. – М. : Финансы и статистика, 2005. – 144 с. – ISBN 5-279-02918-1.

39. **Привальский, В. Б.** О построении управления, приводящего объект в заданную точку / В. Б. Привальский [и др.] // X Всесоюз. совещ. по проблемам управления : тез. докл. В 2 кн. Кн. 1 / Акад. наук СССР. – М., 1986. – 564 с.

ПРИЛОЖЕНИЕ

Основные сведения об операциях над матрицами

1. Матрица. Совокупность mn чисел, расположенных в виде прямоугольной таблицы из n столбцов и m строк, называется прямоугольной матрицей с размерами $m \times n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Если $m = n$, то матрица называется квадртной. Числа a_{ij} – элементы матрицы. Аналогично можно рассматривать функциональные матрицы, если в таблице записываются не числа, а функции, например, $A(t)$ с элементами $a_{ij}(t)$.

2. Матрица-столбец. Это матрица с размерами $m \times 1$:

$$\overline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{bmatrix}.$$

3. Матрица-строка. Это матрица с размерами $1 \times n$:

$$\overline{x}^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}.$$

4. Скаляр. Это матрицы с размерами 1×1.

5. Транспонирование матриц. Это операция образования по матрице A новой матрицы A^{T}, строки которой являются столбцами матрицы A:

$$A^{\mathrm{T}} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

6. Симметрическая матрица. Это квадратная матрица A, для которой $A^{\mathrm{T}} = A$. Элементы такой матрицы располагаются симметрично относительно главной диагонали, например:

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & -2 & 5 \\ 3 & 5 & 4 \end{bmatrix}.$$

7. Сложение матриц. Складывать можно матрицы A и B одинакового размера. Сумма $A + B$ – это матрица, элементы которой получают сложением одинаково расположенных элементов матриц A и B.

8. Произведение матрицы и скаляра. Это матрица, у которой каждый элемент получен умножением соответствующего элемента исходной матрицы A на данное число α.

9. Произведение матрицы и столбца. Это матрица-столбец, определяемая следующим образом:

$$A\overline{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \end{bmatrix}.$$

10. Произведение строки и столбца. Это скаляр, получаемый по следующему правилу:

$$\overline{x}^{\mathrm{T}}\overline{y} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 + x_2y_2 + ... + x_ny_n \end{bmatrix}.$$

11. Умножение прямоугольных матриц. Матрицу A с элементами a_{ij} с размерами $m \times p$ можно умножить на матрицу B с размерами $p \times n$. В результате получим матрицу C с размерами $m \times n$ с элементами c_{ij}, причем

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} .$$

12. Квадратичная форма. Это многочлен n переменных $x_1, x_2, ..., x_n$, который можно ввести с помощью равенства:

$$y = \overline{x}^{\mathrm{T}} Q\, \overline{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} x_i x_j ,$$

где Q – квадратная симметрическая матрица. Например, для $n = 2$:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} q_{11}x_1 & q_{12}x_2 \\ q_{12}x_1 & q_{22}x_2 \end{bmatrix} =$$

$$= \left[x_1(q_{11}x_1 + q_{12}x_2) + x_2(q_{12}x_1 + q_{22}x_2)\right] = \left[q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2\right].$$

13. Дифференцирование матриц по скалярному параметру (например времени). Оно определяется следующим образом:

$$\text{если} \quad \overline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \text{ то } \frac{d\overline{x}(t)}{dt} = \begin{bmatrix} \dfrac{dx_1(t)}{dt} \\ \vdots \\ \dfrac{dx_n(t)}{dt} \end{bmatrix};$$

$$\text{если } A(t)=\begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \dots & \dots & \dots \\ a_{m1}(t) & \dots & a_{mn}(t) \end{bmatrix}, \text{ то } \frac{dA(t)}{dt}=\begin{bmatrix} \frac{da_{11}(t)}{dt} & \dots & \frac{da_{1n}(t)}{dt} \\ \dots & \dots & \dots \\ \frac{da_{m1}(t)}{dt} & \dots & \frac{da_{mn}(t)}{dt} \end{bmatrix}.$$

14. Дифференцирование по векторному параметру. Пусть $f(\bar{x})$ – скалярная функция от векторного аргумента $\bar{x}$, тогда

$$\frac{df}{d\bar{x}}=\begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix}^{\mathrm{T}}.$$

В результате получается вектор, называемый градиентом и обозначаемый

$$\frac{df(\bar{x})}{d\bar{x}}=\mathrm{grad} f(\bar{x})=\nabla_x f.$$

15. Частные случаи:

а) $y=f(\bar{x})=\bar{x}^{\mathrm{T}}Q\bar{x}$ – квадратичная форма. Можно доказать, что $\frac{d(\bar{x}^{\mathrm{T}}Q\bar{x})}{d\bar{x}}=2Q\bar{x}$. Проиллюстрируем это соотношение на простейшем примере:

$$y=\bar{x}^{\mathrm{T}}Q\bar{x}=\left[q_{11}x_1^2+2q_{12}x_1x_2+q_{22}x_2^2\right];$$

$$\frac{dy}{d\bar{x}}=\begin{bmatrix} \frac{dy}{dx_1} \\ \frac{dy}{dx_2} \end{bmatrix}=\begin{bmatrix} 2q_{11}x_1+2q_{12}x_2 \\ 2q_{12}x_1+2q_{22}x_2 \end{bmatrix}=2\begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=2Q\bar{x}.$$

б) $y=\overline{\Psi}^{\mathrm{T}}A\bar{x}$, где строка $\overline{\Psi}^{\mathrm{T}}$ и квадратная матрица A не зависят от $\bar{x}$. Можно доказать, что $\frac{d\overline{\Psi}^{\mathrm{T}}A\bar{x}}{d\bar{x}}=A^{\mathrm{T}}\overline{\Psi}$. Проиллюстрируем это соотношение примером:

$$y=\begin{bmatrix}\Psi_1 & \Psi_2\end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=[\Psi_1(a_{11}x_1+a_{12}x_2)+\Psi_2(a_{21}x_1+a_{22}x_2)];$$

$$\frac{dy}{d\bar{x}}=\begin{bmatrix} \frac{dy}{dx_1} \\ \frac{dy}{dx_2} \end{bmatrix}=\begin{bmatrix} \Psi_1a_{11}+\Psi_2a_{21} \\ \Psi_1a_{12}+\Psi_2a_{22} \end{bmatrix}=\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} \Psi_1 \\ \Psi_2 \end{bmatrix}=A^{\mathrm{T}}\overline{\Psi}.$$

УКАЗАТЕЛЬ ОСНОВНЫХ ОБОЗНАЧЕНИЙ

$x_i \quad (i=\overline{1,n})$; $x_{ij} \quad (i=\overline{1,n};\ j=\overline{1,m})$ – переменные, подлежащие определению в задачах статической оптимизации (линейного и нелинейного программирования); $\bar{x}=\left[x_1,...,x_n\right]^{\mathrm{T}}$ – вектор указанных переменных;

$x_1(i);\ x_2(i),...$ – значения определяемых переменных $x_1, x_2,...$ на i-м этапе (шаге) в задачах динамического программирования;

$x_1,...,x_n$ – обозначения вершин в графах;

$x_i = x_i(t)$ – функции, описывающие переменные состояния в задачах динамической оптимизации, в вариационном исчислении и принципе максимума;

y – показатель, оценивающий качество решения (показатель качества) в задачах статической оптимизации;

Y – показатель качества в задачах динамической оптимизации;

f – функция, определяющая зависимость показателя качества от варьируемых параметров x_i в задачах статической оптимизации (целевая функция);

f_0 – подынтегральная функция, входящая в интегральный показатель качества в задачах динамической оптимизации;

$f_1,...,f_n$ – правые части дифференциальных уравнений, описывающих управляемую систему в задачах динамической оптимизации;

t – время в задачах динамической оптимизации;

$i, j,...$ – номера переменных, этапов, шагов;

a_{ij} – элемент матрицы A; b_i – элементы столбца $\bar{b}$, фигурирующие в задачах линейного программирования;

c_i – коэффициенты при переменных в целевой функции в задачах линейного программирования;

c_{ij} – удельные стоимости в транспортных задачах;

c_{ij} – пропускные способности в транспортных сетях;

p_{ij} – потоки в транспортных сетях;

$u_i = u_i(t)$ – подлежащие определению функции управления в задачах динамической оптимизации;

$s_j = s_j(i)$ – переменные состояния на j-м этапе в задачах динамического программирования;

$\nabla f = \text{grad}\, f = \left[\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right]^{\text{T}}$ – градиент функции f;

$H = \nabla^2 f$ – матрица Гессе, используемая в достаточном условии экстремума;

H – функция Гамильтона в принципе максимума;

λ_i – вспомогательные величины (функции) – множители Лагранжа;

$\psi_i = \psi_i(t)$ – вспомогательные (сопряженные) функции в принципе максимума.

Учебное издание

ГАЛКИН Анатолий Александрович

МАТЕМАТИЧЕСКАЯ ЭКОНОМИКА

Учебник

Редакторы Е.А. Амирсейидова
Е.В. Невская
И.А. Арефьева
Технический редактор Н.В. Тупицына
Корректоры Е.В. Афанасьева, В.В. Гурова
Компьютерная верстка С.В. Павлухиной

Подписано в печать 24.11.06.
Формат 60x84/16. Бумага для множит. техники. Гарнитура Таймс.
Печать на ризографе. Усл. печ. л. 17,67. Уч.-изд. л. 17,97. Тираж 495 экз.
Заказ
Издательство
Владимирского государственного университета.
600000, Владимир, ул. Горького, 87.