МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО СПЕЦИАЛЬНОГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ УЗБЕКИСТАН

Т. ХОДЖАЕВ, И.АЗИЗОВ, С.ОТАКУЛОВ

ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

Рекомендовано Министерством Высшего и среднего специального образования Республики Узбекистан в качестве учебного пособия для студентов высших учебных заведений

ТАШКЕНТ – 2007

Ходжаев Т., Азизов И., Отакулов С. Исследование операций. (Учебное пособе) – Изд-во «Aloqachi», 2007, 176 стр.

Учебное пособие «Исследование операций» написано на основе программы предмета одноименного курса. В работе проработаны и изложены: общая характеристика исследования операций, основные этапы и принципы исследования операций, векторная оптимизация, элементы теории игр, игра двух лиц с нулевой суммой, игры с природой, основы сетевого планирования и управления, расчет временных параметров сетевого графика, вероятностные сети, оптимизация комплекса операций по времени и но стоимости, алгоритм решения задачи о максимальном потоке, задача о кратчайшем маршруте, задача о потоке минимальной стоимости, алгоритм Басакера-Гоуэна, задачи теории расписаний, задача о назначениях, марковские случайные процессы. Изложенные материалы оснащены практическими примерами и задачами. Для более глубокого изучения и проработки материалов приведены вопросы и задания для самостоятельной работы и проблемного творческого подхода к их решению.

Настоящее учебное пособие будет полезным магистрантам, обучающимся по специальностям 5А460106, 5А480105, 5А480108 и специалистам, занимающимся научными исследованиями.

Рецензенты: **Х. ТУРАЕВ** – д.т.н., профессор;
А. КРАСИНСКИЙ – д.ф. - м.н., профессор;
М. ТУХТАСИНОВ – доцент.

ISBN 978-9943-326-07-1

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Развитие и становление любого общества обусловлено целенаправленным исследованием и выполнением множества задач различных отраслей народного хозяйства. Решение каждой из этих задач есть, по сути, некая определенная операция, реализация которой должна обеспечить эффективность поставленной цели исследования, и, причем в разумные сроки.

Заметим, что независимо от того, какая операция, сложная или простая, исследуется, качественная ее реализация зависит, в первую очередь, от правильного формализованного ее представления и правильного выбора метода решения. Исследование операций, как раз, и есть та наука, которая обладает мощным комплексом, апробированных научных методов и подходов, применяющихся для решения задач по эффективному управлению любой разумной человеческой деятельности, в частности, по оперативному управлению отраслей народного хозяйства.

Развитие исследования операций, можно сказать, имеет глубокие исторические корни, уходящие в далекое прошлое. Однако, датой рождения исследования операций, как самостоятельного научного направления, принято считать начало второй четверти XX столетия. В появившихся, в то время, первых публикациях по исследованию операций были изложены, разработанные методы и подходы, примененные для решения военных задач, в частности, для анализа и исследования боевых операций. Позднее и в настоящее время принципы и методы исследования операций применялись и достаточно широко применяются в сфере промышленного, финансового управления народным хозяйством.

Во второй половине XX столетия, точнее в 1957 году, была создана, и, в настоящее время продолжает функционировать, Международная федерация обществ по исследованию операций – IFORS (International Federation of Operations Research Sociaties), в состав которой вошли национальные комитеты и общества по исследованию операций многих стран.

Большой вклад в формирование и развитие исследования операций, в создании современного математического аппарата и развития ряда ее направлений внесли ученые Р.Акоф, Р. Беллман, Г. Данциг, Г. Кун, Т. Саати, Р. Чермен (США), А. Кофман, Р. Фор (Франция), Л.В. Канторович, Б.В. Гнеденко, Н.П. Бусленко, Д.Б. Юдин, Н.П. Федоренко (Россия) и др.

Настоящее учебное пособие формализовано авторами на базе тех материалов предмета «Исследования операций», которые преподносились ими, в течение более чем 20 лет, студентам факультета прикладной математики и информатики Самаркандского государственного университета. В настоящее время этот курс, в качестве одной из основных дисциплин, читается студентам, обучающимся в бакалавриатуре факультета информатики и информационных технологий названного университета.

Авторы отмечают, что работа не исключена от возможности допущения некоторых субъективных мнений по части формализации и изложения материалов. Любые пожелания будут приняты ими с благодарностью.

ГЛАВА I. ОСНОВНЫЕ ПРИНЦИПЫ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

§ 1. ОБЩАЯ ХАРАКТЕРИСТИКА ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Операция, управление, система, стратегия, оптимальная стратегия, критерий

1.1 Предмет исследования операций

Исследование операций является, по сути, еще новой интенсивно развивающейся наукой. Точного и полного определения дать ей довольно-таки трудно, но, тем не менее, для понимания предмета исследования операций, ее содержания и целей, такое определение необходимо.

Исследование операций — это наука, занимающаяся разработкой и практическим применением методов наиболее рационального (оптимального) управления организационными системами.

Предмет исследования операций представляет собой систему организационного управления, которая состоит из большого числа взаимодействующих между собой подразделений, причем интересы подразделений не всегда согласуются между собой и могут быть противоположны.

Целью исследования операций является количественное обоснование, принимаемых решений по управлению организационными системами.

Основной задачей исследования операций является поиск наилучших путей, ведущих к достижению цели, и их оценка.

Как и любая другая наука, исследование операций обусловлено многими специфическими понятиями и терминами, среди которых основным является понятие операции. Под операцией понимается комплекс действий, направленных на достижение поставленной цели.

В проведении любой операции могут быть задействованы совокупность лиц, устройств, автоматов, которые стремятся достичь своей поставленной цели и несут ответственность за выполнение операции. Их будем называть оперирующей стороной. Оперирующая сторона всегда имеет в своем распоряжении некоторый запас ресурсов. Эти ресурсы используются и расходуются оперирующей стороной для достижения поставленной цели и называются они активными средствами. Способы их использования, расходования считаются стратегиями оперирующей стороны.

Для достижения поставленной цели оперирующая сторона всегда пользуется наиболее эффективными стратегиями, которые, несомненно, отражают суть цели. Такие стратегии называются оптимальными.

1.2. Критерий эффективности

При исследовании какой-либо операции, обязательно возникает необходимость выделения некоторого критерия оценки достижения поставленной цели, который, в первую очередь, должен обеспечить меру эффективности результата выполнения операции.

В общем, слово критерий происходит от греческого kritērion и означает средство для суждения, т.е. — признак, на основании которого производится оценка, определение или классификация чего-либо; мерило суждения. Следовательно, применительно к задачам исследования операций, критерий – есть главное средство для количественной оценки решений, сравнения их между собой и выбора наилучшего (оптимального).

Исследователь операции, который входит в состав оперирующей стороны, оценивает различные исходы операции, соответствующие различным стратегиям, пользуется именно таким критерием, вернее он подбирает этот критерий. Итак, критерий эффективности – это математический эквивалент цели операции, позволяющий количественно оценивать меру достижения этой цели.

В результате решения поставленной задачи критерий эффективности должен принимать наилучшее значение. В одних задачах этот критерий принимает максимальное значение, а в других – минимальное. Иначе говоря, в одних задачах критерий эффективности необходимо обратить в максимум, а в других – в минимум.

Задача, когда критерий эффективности необходимо минимизировать, легко сводится к задаче максимизации, и, наоборот (для этого достаточно изменить знак критерия эффективности на противоположный).

Критерий эффективности должен обладать некоторыми особенными свойствами:

– не содержать большого числа, затрудняющих его исследование, побочных связей и факторов, т.е. – простотой;

– отражать основную, а не побочные цели операции, т.е. – представительностью;

– в достаточной степени реагировать на изменения параметров, характеризующих выбор той или иной стратегии, т.е. – критичностью.

Еще одним особенным, но не всегда достигаемым, свойством является единственность критерия эффективности. Это свойство характеризуется тем, что каждой операции должен соответствовать единственный критерий эффективности.

Анализируя цели выполнения операции, можно выделить два вида целей и соответственно два вида критерия эффективности:

1. **Цели, сущность которых состоит в том, что они могут быть достигнуты или нет. Такие цели называются качественными.** Качественный критерий эффективности Φ, соответствующий качественной цели, может быть записан так:

$$\Phi = \begin{cases} 1, \text{если цель достигается}, \\ 0 \text{ или } -\infty, \text{если цель не достигается}. \end{cases}$$

2. **Количественные цели.** Сущностью количественных целей является стремление к увеличению или уменьшению цели, зависящей от стратегий. По сути дела-это показатель, являющийся количественной мерой степени достижения цели операции.

Рассмотрим пример планирования межотраслевых потоков выпуска продукции. Пусть имеется ряд отраслей народного хозяйства, каждая из которых занимается выпуском своей продукции. Эти отрасли структурно связаны некоторыми своими производственными подразделениями, исполняющими характерные им работы. Каждое подразделение, для достижения собственных целей, исходит из своих возможностей, которые порой оказываются не выполнимыми из-за ряда причин. Это связано со множеством факторов, которые в свою очередь являются зависимыми от действий других подразделений отрасли. В этом случае возникают ситуации, где сталкиваются противоречивые интересы подразделений, устранение которых приводит к некоторой задаче компромисса. Задачей планирования теперь является поиск такого пути, который приведет к реализации поставленной цели.

Каждая отрасль стремится выпускать как можно больше продукции при наименьших затратах. Поэтому она заинтересована в возможно более длительном и непрерывном производстве. Такой процесс, во-первых, позволяет выпускать продукцию в большем количестве, и, во-вторых, приводит к снижению ее себестоимости. Далее, выпуск изделий большими партиями требует создания больших объемов запасов материалов, комплектующих изделий и т.д. Это, в свою очередь, наталкивается на соответствующие трудности снабжения и хранения запасов. Отсюда возникают противоречивые обстоятельства внутри подразделений.

Руководство каждой отрасли, стремясь минимизировать расходы функционирования своего производственного комплекса, пытается усовершенствовать свои внешние связи. **Например,** завоз ресурсов как можно с ближних производств, организация разумной рекламы, изучение конъюнктуры рынка и т.д. Здесь приведены достаточно много различных целей производства, каждую из которых можно измерять соответствующими единицами измерения. Они являются примерами количественных целей.

Рассмотрим другой пример. Пусть имеем дело с некоторым технологическим процессом, реализация которого зависит от наличия энергетических ресурсов. Здесь осуществление технологического процесса полностью зависит от наличия в достаточном количестве энергетических ресурсов. Цель достигается при наличии этих ресурсов, иначе – нет.

Этот пример обусловлен наличием качественной цели.

1.3 Основные особенности исследования операций

Известно, что изучение объектов и явлений как систем вызвало формирование нового подхода в науке – системного подхода. Системный подход как общеметодический принцип используется в различных отраслях науки и деятельности человека. Системный подход – это подход к исследованию объекта (проблемы, явления, процессы), т.е. системы, в которой выделены элементы, внутренние и внешние связи, наиболее существенным образом влияющие на исследуемые результаты ее функционирования, а цели каждого из элементов, исходят из общего предназначения объекта.

Можно также сказать, что системный подход – это такое направление методологии научного познания и практической деятельности, в основе которого лежит исследование любого объекта как сложной целостной социально-экономической системы. Любая задача, какой бы частной она не казалась на первый взгляд, рассматривается с точки зрения ее влияния на критерий функционирования всей системы.

Характерной особенностью исследования операций является системный подход к анализу поставленной проблемы. Исследование операций обусловлено тем, что при решении каждой проблемы возникают все новые и новые задачи.

Одной из главных и существенных особенностей исследования операций является стремление найти оптимальное решение, поставленной задачи. Однако, часто, такое решение оказывается недостижимым из-за ограничений, накладываемых имеющимися в наличии ресурсами (денежные средства, машинное время) или же возможностями современной науки. Для многих задач, в которых применяются комбинаторные методы решения, например, в задачах календарного планирования, при числе станков $n<4$, найти оптимальное решение оказывается возможным при использовании метода перебора. Отметим, что и при небольших n, число возможных вариантов оказывается настолько большим, что перебор всех вариантов практически не возможен даже при использовании современных средств вычислительной техники.

Один из основателей исследования операций – Т. Саати дал следующее определение этой науке: «Исследование операций – это искусство давать плохие ответы на те практические вопросы, на которые даются еще худшие ответы другими методами».

Для проведения исследования операций создается некоторая группа специалистов. Это – управленцы, инженеры, математики, экономисты, социологи, психологи, программисты и другие. Задачей создания подобных групп является комплексное исследование всего множества факторов, влияющих на решение проблемы и использование идей и методов различных наук. Это является еще одной особенностью исследования операций, состоящая в том, что все мероприятия проводятся комплексно, по многим направлениям.

Проблемные задания

1. Сформулируйте какую-нибудь проблему, отражающую задачу организационной системы.
2. Определите характерные признаки задачи задания 1 и сформулируйте их в терминах исследования операций.
3. Для задачи задания 1 определите цели операции и сформулируйте соответствующий критерий эффективности для каждой цели.

Вопросы для самоподготовки

1. Определение исследования операций.
2. Основная цель исследования операций.
3. Определение операции.
4. Характерные особенности критерия эффективности.
5. Виды критерия эффективности.
6. Назовите основные особенности исследования операций.

§ 2. ОСНОВНЫЕ ЭТАПЫ И ПРИНЦИПЫ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Операция, задача, модель, математическая модель, метод решения, линейное программирование, динамическое программирование, стохастическое программирование.

2.1. Основные этапы исследования операций

Исследование операций охватывает широкий круг задач, которые характеризуются многообразием свойств и принципов. Этому многообразию свойственно то, что каждое исследование операции проходит последовательно через следующие основные этапы:

1) постановка задачи;

2) построение математической модели;

3) поиск и выбор метода решения;

4) проверка и (при необходимости) корректировка модели;

5) внедрение результатов решения.

Дадим краткую характеристику этим этапам.

Постановка задачи – наиболее сложный и ответственный этап исследования операции. Здесь, первоначально, задача формируется и представляется с позиций заказчика. Порою, такая постановка задачи не бывает окончательной. Анализируя исследуемую систему, приходится совершенствовать постановку задачи, т.е. – уточнять ее детали.

На втором этапе происходит **построение математической модели**, т.е. математическая формализация задачи. Этот процесс является творческим. Здесь каждый исследователь операции строит модель исходя из своих возможностей, своей интуиции. Поэтому, одна и та же постановка задачи может быть представлена моделями, отличающимися по разным признакам. Этот процесс так же является трудным. Здесь, от исследователя требуется достаточных знаний по математике и знаний законов той области науки, для которой будет применяться разрабатываемая модель.

Математическая модель задачи, полученная в результате проведения этого этапа, в некотором случае, может иметь вид:

найти экстремум (максимум или минимум) функции $z = f(x, y)$ при ограничениях $g_i(x, y) \le (\text{или} \ge, \text{или} =) b_i$, $i = \overline{1, m}$.

Здесь $f(x, y)$ – целевая функция (скалярная, показатель качества или эффективность системы); x – вектор управляемых переменных; y – вектор неуправляемых переменных; $g_i(x, y)$ – функция потребления i-го ресурса; b_i – величина i-го ресурса (например, плановый фонд машинного времени группы токарных автоматов в станко-часах).

Поиск и выбор метода решения. Определение оптимального решения (если такое существует) рассматриваемой задачи зависит, во-первых, от вида целевой функции $f(x, y)$, и, во-вторых, от функций $g_i(x, y)$, $i = \overline{1, m}$, с помощью которых описываются ограничения задачи. Это дает возможность отнести задачу к той или иной области математического программирования и применить соответствующие методы определения оптимального решения.

Этап **проверки модели на адекватность** необходим для установления оценки точности (в достаточной степени) разработанной модели. Здесь выясняются характеристические особенности модели и степень их соответствия реальным параметрам рассматриваемого объекта.

В сложных системах, к которым относятся системы организационного типа, модель лишь частично отражает реальный процесс. Поэтому здесь крайне необходима проверка степени соответствия или адекватности модели к реальному процессу.

Естественно, любая построенная модель не может однозначно отражать все реальности действительного объекта. Поэтому полученную модель нужно подвергать корректировке, путем проведения дополнительных исследований объекта и уточнения структуры его математической модели. Для этого необходимо возвратится к начальному этапу исследования.

Таким образом, четыре названные выше этапа повторяют многократно до тех пор, пока не будет достигнуто практическое соответствие между выходными параметрами объекта и составленной моделью.

Внедрение результатов решения. Этот этап является важнейшим этапом, завершающим исследование операций. Он также является сложным и требует решения множества задач, находящихся вне области исследования операций. Отметим, что процесс внедрения результатов исследований можно рассматривать как самостоятельную задачу, где необходимо применение методов системного подхода и системного анализа.

2.2. Типичные задачи исследования операций

Практически, любая целенаправленная деятельность человека связана с решением каких-то проблем, относящихся к задачам исследования операций. По каким-то признакам эти задачи можно подразделить на соответствующие классы. Из этих классов задач можно выделить следующие типичные задачи.

Задача об оптимальном плане выпуска продукции. Суть этой задачи заключается в следующем, пусть некоторое предприятие выпускает несколько видов продукции, для которых необходимо определенное количество видов сырья. От реализации каждого вида продукции предприятие имеет соответствующую прибыль. Требует-

2

ся составить такой план выпуска продукции, который был бы осуществим, удовлетворял бы поставленным ограничениям на выпуски каждого вида продукции и в то же время приносил бы общую прибыль предприятию.

Задача определения наилучшего состава смеси. Пусть задано содержание необходимого количества питательных веществ в различных кормах, применяемых для кормления животных. Кроме того, известна также цена единицы каждого вида корма. Требуется выбрать такой рацион, т.е. набор и количество кормов, чтобы каждое питательное вещество содержалось в нем в требуемом количестве, необходимом для кормления животного и также, чтобы суммарные расходы на этот рацион были минимальны.

Задачи управления запасами. Эти задачи составляют наиболее распространенный и изученный, в настоящее время, класс задач исследования операций. Суть этих задач заключается в следующем. Любое хозяйство, для нормального функционирования, стремится иметь определенный запас ресурсов. Увеличение запасов приводит к увеличению расходов на их хранение, но при этом уменьшаются потери из-за возможной их нехватки. Следовательно, одна из задач управления запасами заключается в определении такого уровня запасов, который минимизирует сумму ожидаемых затрат по хранению запасов, а также потерь из-за их дефицита.

В задачах управления запасами могут возникать следующие ситуации:

– запас пополняется мгновенно;

– запас пополняется в течение определенного момента времени;

– пополнение запаса производится на основе вероятностной оценки ситуации.

Транспортная задача. Пусть требуется доставить некоторый однородный груз из пунктов производителей в пункты потребителей. Количество этого груза, выпускаемого каждым производителем и потребное его количество каждым потребителем, считаются известными. Необходимо составить такой план перевозок, чтобы запросы каждого потребителя были удовлетворены, чтобы весь груз был вывезен из пунктов производителей и, причем, с наименьшими общими затратами.

В этой задаче могут возникнуть две ситуации:

– количество груза равно количеству запроса потребителей;

– количество груза не равно количеству запроса потребителей.

Задачи распределения ресурсов. Эти задачи возникают в тех ситуациях, когда существует некоторый набор, подлежащих исполнению, работ (операций) При этом, может оказаться, что наличие ресурсов для эффективного выполнения каждой работы зачастую не хватает.

В этих задачах, в зависимости от условий, могут возникнуть следующие ситуации:

– известны и работы, и ресурсы. Необходимо так распределить ресурсы между работами, чтобы прибыль была максимальной, или же затраты были минимальны;

– известны только наличные ресурсы. Требуется определить, такой состав работ, чтобы обеспечить максимум дохода с учетом этих ресурсов;

– известны только работы. Необходимо определить, какие ресурсы нужны для того, чтобы минимизировать суммарные издержки производства.

Задачи массового обслуживания. В этих задачах исследуются вопросы образования и функционирования очередей, определения различных параметров систем обслуживания на различной стадии их работы, с которыми приходится сталкиваться на практике. В зависимости от условий, здесь могут быть рассмотрены следующие ситуации:

– требования для обслуживания поступают из неограниченного источника;

– в системе нет мест для ожидания, т.е. требования не обслуживаются из-за занятости всех мест обслуживания;

– в системе имеются места для ожидания. В случае занятости всех мест обслуживания, поступившие требования могут подождать в очереди на обслуживание.

Задачи теории расписаний (задачи упорядочения). Эти задачи обусловлены следующими особенностями. Известны, подлежащие исполнению работы и устройства (машины люди и т.п.) для выполнения этих работ. Также известно время, отведенное на выполнение каждой работы на каждом устройстве. Необходимо составить такой план (такое расписание) выполнения этих работ на устройствах, чтобы максимизировать (или минимизировать) некоторый критерий эффективности, например, минимизировать суммарную продолжительность работ.

Задачи сетевого планирования и управления. Здесь рассматривается комплекс операций, где продолжительность времени выполнения каждой операции считается известной. Также может быть задано директивное время выполнения каждой операции и комплекса операций в целом. В этих задачах определяются временные параметры, устанавливающие зависимости операций друг от друга. Они могут быть рассмотрены в следующих постановках:

– известна продолжительность всего комплекса. Необходимо определить сроки начала каждой операции, при которой минимизируется (или максимизируется) какой-нибудь критерий;

– известны общие ресурсы. Нужно определить сроки начала каждой операции, которые позволяют минимизировать продолжительность выполнения всего комплекса работ.

Задача о назначениях. Имеется определенное количество различных работ и устройств (машин, людей и т.п.) для их выполнения. Каждое устройство может быть назначено на любую работу, причем известны ожидаемые эффекты от использования каждой работы на каждом устройстве. Суть этой задачи состоит в определении такого плана назначений, чтобы суммарные эффективности были максимальными.

Математические модели задач исследования операций

Любая задача в своей постановке и дальнейшем исследовании немыслима без формализации и представления ее в математическом виде. Этот процесс обусловлен использованием законов и символов математики, правильное формирование которых, приводит к, так называемой, математической модели.

Математическая модель любой задачи представляет собой некоторое приближенное описание рассматриваемых явлений, событий, операций, при помощи математической символики. Математическая модель позволяет познать и оценить структуру исследуемых объектов, прогнозировать поведение этих объектов на определенный интервал времени и принимать решение по эффективному их управлению.

Процесс математического моделирования, как было отмечено выше, является вторым основным этапом исследования операций. Этот этап будучи творческим процессом, требует от исследователя всесторонней научной эрудированности.

Процесс математического моделирования можно подразделить на 3 этапа.

Первый этап – формализация концепций, определяющих структурные связи исследуемых объектов модели. Здесь требуется качественный сбор информации, относящейся к изучаемым явлениям, что в последствии позволяет представить исследуемый объект в математических терминах.

Второй этап направлен на исследование и решение сформулированных задач, представленных математическими моделями. На этом этапе важную роль приобретают математический аппарат, необходимый для анализа математических моделей.

Третий этап – оценка адекватности сформированной модели практическим аналогам. Здесь устанавливается согласованность результатов опытов с теоретическими концепциями.

Как было отмечено ранее, для оценки различных исходов операции, соответствующие различным стратегиям, необходим некоторый критерий эффективности.

В достаточно простых ситуациях исследования операций, удается ограничиться единственным критерием оптимальности. Соответствующие задачи принятия решений называются одноцелевыми или однокритериальными. Можно так же отметить, что реально, многие задачи исследования операций характеризуются многокритериальностью (например, выпустить как можно больший объем продукции с наименьшими затратами и т.н.).

В начале рассмотрим однокритериальные задачи исследования операций. Пусть имеется некоторая операция, т.е. управляемое мероприятие, на исход которого, оперирующая сторона может влиять в какой-то мере. Эффективность этого управления характеризуется некоторым критерием эффективности z, допускающим количественное представление. Здесь критерий эффективности может быть задан либо в виде функции, либо в виде функционала, либо иметь лишь алгоритмическое значение.

Величина критерия эффективности z зависит от ряда факторов, который можно разбить на две группы:

1) контролируемые (управляемые) факторы, выбор которых находится в распоряжении оперирующей стороны. Каждый конкретный выбор значений контроли-

руемых факторов, составляет стратегию оперирующей стороны;

2) неконтролируемые (неуправляемые) факторы, на которые оперирующая сторона влиять не может. В состав неконтролируемых факторов входит и время, если в операции участвуют динамические объекты, изменяющие свои свойства и поведение во времени.

В свою очередь, неконтролируемые факторы, в зависимости от информированности о них исследователя операций, можно разбить на следующие три группы:

1) детерминированные факторы — неслучайные фиксированные факторы, т.е. известные факторы;

2) стохастические факторы — случайные фиксированные факторы-величины и процессы с известными оперирующей стороне законами распределения;

3) неопределенные факторы, для каждого из которых известна только область (определения) возможных знаний фактора или область, внутри которой находится закон распределения, если фактор случаен.

В соответствии с выделенными факторами, критерий эффективности z, можно представить в виде зависимости:

$$z = f(x, a, y, v, t) ,$$

где x — вектор контролируемых факторов, a — вектор неконтролируемых фиксируемых факторов, y — вектор неконтролируемых стохастических факторов, v — вектор неконтролируемых неопределенных факторов, t — время.

Все величины, влияющие на исход операции, обычно, ограничены рядом естественных причин, что математически можно выразить в виде следующих условий:

$$g_i = g_i(x, a, y, v, t) \le (\text{или} \ge, \text{или} =) b_i , \quad i = \overline{1, m} .$$

Эти условия определяют области допустимых значений, внутри которых расположены возможные значения всех факторов.

Поскольку, критерий эффективности есть количественная мера степени достижения цели операции, то математически, цель выражается в стремлении к возможному увеличению (или уменьшению) его значения.

Средством достижения этой цели, является соответствующий выбор оперирующей стороной стратегии x из области допустимых значений контролируемых факторов.

Таким образом, перед лицом, ответственным за принятие решения, стоит задача, которую можно сформулировать следующим образом:

при заданных значениях неконтролируемых факторов a и y, с учетом неконтролируемых неопределенных факторов v, найти оптимальные значения x^0 из области допустимых значений, которые обращали бы в максимум (минимум) критерий эффективности z.

Эту задачу формально можно представить в виде:

$$z \to \max \;\; (\min),$$

$$g_i = g_i(x, a, y, v, t) \leq (\text{или} \geq, \text{или} =) b_i, \;\; i = \overline{1, m}.$$

2.4. Классификация математических моделей задач исследования операций

В настоящее время, не существует общепринятой универсальной классификационной схемы математических моделей задач исследования операций. Можно выделить отдельные их классификационные признаки, а именно:

1. Количество целей операции, преследуемых одной оперирующей стороной и соответствующих этим целям критериев эффективности.

2. Наличие или отсутствие зависимости критериев эффективности и дисциплинирующих условий от времени.

3. Наличие случайных и неопределенных факторов, влияющих на исход операции.

По первому классификационному признаку задачи исследования операций делятся на два больших класса: однокритериальные (скалярные) и многокритериальные (векторные).

По второму классификационному признаку задачи исследования операций так же делятся на два класса: статические и динамические. В статических задачах исследования операций функция, с помощью которой определяется критерий (цель) операции и функция ограничений не зависят от времени. В динамических же задачах – имеется зависимость от времени. Эти задачи сложные и в настоящее время они еще не получили широкого применения на практике.

По третьему классификационному признаку – определенность-риск-неопределенность – задачи исследования операций делятся на три больших класса:

1. Детерминированные задачи исследования операций, т.е. исследование операций при определенности. Эти задачи характеризуются однозначностью и детерминированностью между принятым решением и исходом. Здесь, относительно каждой стратегии, оперирующей стороне заранее известно, что она приводит к некоторому конкретному результату. В этих задачах критерий эффективности и дисциплинирующие условия зависят только от стратегии оперирующей стороны и фиксированных неконтролируемых факторов.

2. Стохастические задачи исследования операций, т.е. исследование операций при риске. Здесь, каждая стратегия оперирующей стороны может привести к одному из множества возможных исходов, причем каждый исход имеет определенную вероятность появления.

3. Задачи исследования операций в условиях неопределенности. В этих задачах критерий эффективности зависит, кроме стратегий оперирующей стороны и фиксированных факторов, также от неопределенных факторов, неподвластных оперирующей стороне и неизвестных ей в момент принятия решений. В результате влияния неопределенных факторов, каждая стратегия оперирующей стороны оказывается связанной с множеством возможных исходов, вероятности которых либо не известны оперирующей стороне, либо вовсе не имеют смысла.

Еще одним классификационным признаком задач исследования операций может служить признак подразделения задач по содержательной постановке. Систематизация задач исследования операций позволяет выделить по указанному признаку типичные классы задач, часть из которых была приведена выше.

2.5. Место математического программирования в исследовании операций

При решении задач исследования операций основополагающая роль, можно сказать, принадлежит теории математического программирования. Математический аппарат этой теории позволяет исследовать экстремальные задачи и определять экстремумы функций в ограниченной области допус-тимых значений переменных, определяемых системой ограничений. Математическое программирование – раздел науки об исследовании операций, охватывающий широкий класс задач управления, математическими моделями которых являются конечномерные экстремальные задачи. Задачи математического программирования находят применение в различных областях человеческой деятельности, где необходим выбор одного из возможных образов действий.

При определении возможных методов решения задач исследовании операций, необходимо исходить из анализа математической модели этой задачи, позволяющий проникнуть в сущность изучаемых явлений. Этот процесс в дальнейшем естественно связывается с вычислительными аспектами.

Характерной особенностью вычислительной стороны решения задач исследования операций, методами математического программирования, является то, что применение этих методов неразрывно связано с использованием современных средств вычислительной техники.

Задачи математического программирования, подобно задачам исследования операций, можно классифицировать по разным признакам. Например, на задачи линейного и нелинейного программирования, выпуклого и невыпуклого программирования и т.д. Также существуют такие разделы математического программирования, к которым следует отнести задачи целочисленного, с булевыми переменными, квадратичного, сепарабельного, геометрического программирования. Более подробно эти задачи изложены и исследуются в курсе «Математическое программирование».

Проблемные задания

1. Сформулируйте конкретную задачу из какой-либо отрасли народного хозяйства.

2. Для задачи задания 1 исследуйте последующие этапы исследования операций.

3. Приведите и конкретизируйте возможные проблемы моделирования задачи задания 1.

4. Исследуйте возможные методы реализации этой задачи.

5. Приведите конкретные примеры задач, адекватно (либо не адекватно), отражающие характерные признаки этапов исследования операций.

6. Сформулируйте задачу планирования производства продукции с одноцелевым и многоцелевым критериями эффективности.

7. Сформулируйте задачу, в которой принятие решений осуществлялось бы при стохастических факторах.

8. Сформулируйте ситуационные задачи, в одной из которых участвует фактор времени, а в другой — нет.

9. Определите критерии эффективности для задач предыдущего задания.

10. Приведите пример, отражающий по характеру задачу управления запасами. Исследуйте задачу для случаев, когда в ней один критерий эффективности и когда их много.

Вопросы для самоподготовки

1. Основные этапы исследования операций.
2. Характеристики задач исследования операций.
3. Классификация задач исследования операций.
4. Методы решения задач исследования операций.
5. Факторы, влияющие на критерий эффективности.
6. Однокритериальная задача исследования операций.
7. Типичные задачи исследования операций.
8. Какими задачами можно было бы дополнить приведенный выше список типичных задач исследования операций?

ГЛАВА II.
§ 3. СУЩНОСТЬ И ПРОБЛЕМЫ ЗАДАЧ ВЕКТОРНОЙ ОПТИМИЗАЦИИ

Операция, критерий эффективности, оценка, правило компромисса, область согласия, область компромисса, нормализация критериев, приоритет критериев

3.1. Сущность задачи векторной оптимизации

Характерной особенностью, рассмотренных выше задач исследования операций (в основном) является наличие единственного критерия эффективности. Однако, на практике встречаются ситуации, когда операция не может быть оценена с помощью единственного критерия эффективности. Это вызвано тем, что большие по объему сложные операции трудно поддаются исследованию. Не все связи и зависимости между факторами операции удается установить, и она не может быть однозначно охарактеризована с помощью единственного критерия эффективности. Тогда операция описывается с помощью вектора $z=(z_1,z_2,...,z_n)$ критериев эффективности, причем, как правило, одни показатели необходимо обратить в максимум, а другие — в минимум.

Например, при оценке деятельности промышленного предприятия в качестве частных (локальных) критериев эффективности могут быть приняты такие:

- z_1 — прибыль, z_2 — объем выпускаемой продукции, z_3 — себестоимость, z_4 — рентабельность и т.д. Следует отметить, что решение, доставляющее оптимальное значение одним критериям, как правило, является далеко не оптимальным по другим критериям. Тем не менее, задачи такого рода приходится решать, основываясь на некотором способе нахождения компромисса между несколькими критериями. Следовательно, для сравнения двух решений, необходимо определить правило сравнения этих векторов, на основе которого можно выбрать одно из решений.

Естественно, что все условия задач исследования операций однозначно применимы и при исследовании этих задач. Здесь речь идет о тех факторах, которые могут повлиять на результат решения задачи. Для простоты исследований рассмотрим задачу, которая характеризуется наличием только лишь контролируемого фактора.

Пусть задана некоторая операция S, эффективность которой будет оценена с помощью векторного критерия эффективности z. Эта операция выполняется в такой среде, которая обусловлена множеством допустимых стратегий X. Каждый выбор стратегии $x \in X$, обеспечивает векторному критерию эффективности $z = z(x)$ соответствующее значение.

Теперь задачу сформулируем следующим образом. Необходимо определить такую стратегию x по управлению операцией S, чтобы она удовлетворяла следующие условия:

– стратегия x должна быть осуществимой, т.е. принадлежать области допустимых значений X;

– стратегия x должна быть наилучшей в смысле принятого в операции правила сравнения двух векторов, которое в дальнейшем будем называть правилом компромисса.

В такой постановке задача исследования операций будет называться задачей векторной оптимизации. Ее иногда называют многокритериальной или многоцелевой задачей исследования операций.

В задаче векторной оптимизации, структура критерия эффективности z позволяет разбить его на отдельные критерии, которые принято называть локальными критериями эффективности.

Под термином «решить задачу векторной оптимизации» будем понимать следующее. Среди элементов множества стратегий X найти такой, на котором критерий эффективности z «принимает максимальное или минимальное значение». Известно, что для векторов не существует понятий больше и меньше. Поэтому, приведенная постановка задачи векторной оптимизации является не вполне определенной. В связи с этим, можно сказать, что задачи векторной оптимизации являются, в какой-то мере, задачами оптимизации в условиях неопределенности.

Несмотря на то, что для точного определения решения задачи векторной оптимизации, априорная информация является недостаточной, тем не менее, имеющаяся в наличие информация оказывается достаточной для описания свойств, которым должно обладать каждое из возможных решений поставленной задачи. Это свойство называется эффективностью.

Стратегия $x' \in X$ называется эффективной (оптимальной по Парето), если не существует другой стратегии $\bar{x} \in X$, такой, что $z(\bar{x}) \geq z(x')$ (или $z(\bar{x}) \leq z(x')$) и $z(\bar{x}) \neq z(x')$.

Необходимо отметить, что в задачах векторной оптимизации, могут быть даны и другие определения относительно эффективности (оптимальности) стратегии.

3.2. Основные проблемы векторной оптимизации

Исследование задач векторной оптимизации, так же как и множество других задач исследования операций, обусловлено рядом проблем. Эти проблемы характерны не только при формализации условий задачи, но и при самой ее постановке. Решение этих проблем является трудным, а иногда и просто невозможным. Из множества проблем, возникающих при исследовании и решении задач векторной оптимизации, рассмотрим наиболее важные.

Проблема 1 – определение области компромисса. При решении задач векторной оптимизации между некоторыми локальными критериями возникает противоречие, суть которого состоит в том, что при улучшении какого-либо локального критерия, другой, как правило, ухудшается (например, повышение надежности технического устройства и снижение его стоимости, часто является противоречивым).

Подмножество $X^c \subseteq X$ будем называть областью согласия, если в нем улучшение решения по любому локальному критерию может быть осуществлено без ухудшения решений по другим локальным критериям.

Подмножество $X^к \subseteq X$ будем называть областью компромисса, если здесь стремление к улучшению решения по одному из локальных критериев приведет обязательно к ухудшению по некоторому другому критерию.

Области X^c и $X^к$ обладают следующими свойствами:

$$X^c \cup X^к = X, \; X^c \cap X^к = \varnothing .$$

Ясно, что оптимальное решение будет принадлежать только $X^к$, т.е. решение в X^c всегда может быть улучшено по всем критериям.

Проблема 1, является не основной проблемой в процессе поиска оптимального решения. Решение этой проблемы заключается в выделении области компромисса $X^к$ из области допустимых решений X. Ее разрешение позволяет сузить область допустимых значений, в которой будут находиться оптимальные решения. Отметим, что в некоторых случаях исследование и решение задач векторной оптимизации заканчивается выделением области компромисса, обеспечивая, приемлемую для практических нужд, точность получаемых решений.

Проблема 2 — выбор схемы компромисса. Выше было отмечено, что задачам векторной оптимизации свойственны внутренние противоречия даже при начальной их постановке. В связи с этим, поиск оптимального решения может быть осуществлен лишь после того, как будет выбрана некоторая схема компромисса, т.е. указано правило сравнения двух векторов решений. Это правило является отправной точкой определения области компромисса.

Во многих случаях схема компромисса позволяет привести исходную векторную задачу к скалярной, где оптимизируется единственный критерий эффективности. Это, естественно, позволяет реализовать вычислительные схемы однокритериальных оптимизационных задач.

Проблема 3 — нормализация критериев. Данная проблема неразрывно связана с предыдущей и возникает только в тех задачах, в которых локальные критерии имеют различные единицы измерения. Для безошибочного учета единиц измерения локальных критериев, их обычно приводят к единому, безразмерному масштабу измерения. Именно в этом и состоит суть решения проблемы нормализации критериев. В зависимости от выбранной схемы компромисса проблема нормализации может и отсутствовать.

Проблема 4 — учет приоритета критериев. При исследовании и решении задач векторной оптимизации, довольно часто, локальные критерии имеют различную степень важности. Эта степень обычно задается в виде вектора приоритетов при постановке задачи. Ее необходимо учитывать при решении задачи.

Перечисленные выше проблемы задач векторной оптимизации являются характерными для широкого класса задач различных отраслей народного хозяйства. Во многих случаях их решение осуществляется с помощью различного рода эвристических процедур.

Проблемные задания

1. Сформулируйте задачу и составьте ее математическую модель, позволяющую выделить область компромисса из области допустимых решений.
2. Исследуйте возможность определения оптимального решения задачи, сформулированной в задании 1.
3. Исследуйте возможность сведения локальных критериев задачи, сформулированной в задании 1 (если таковые есть), к единому.

Вопросы для самоподготовки

1. Охарактеризуйте сущность задачи векторной оптимизации.
2. Что такое компромисс?
3. Назовите и охарактеризуйте основные проблемы векторной оптимизации.

§ 4. ПОДХОДЫ К РЕШЕНИЮ ПРОБЛЕМ ВЕКТОРНОЙ ОПТИМИЗАЦИИ

Операция, вектор, оптимизация, критерий, локальный критерий, вектор весовых коэффициентов, свертка, процедура, цель

4.1. Примеры реализации проблем векторной оптимизации

Рассмотрим некоторые подходы к решению приведенных проблем и проиллюстрируем их с помощью примеров. Иллюстрацию проблем векторной оптимизации удобнее всего представить геометрически. Здесь, для удобства восприятия понятий перейдем от области допустимых решений X к области Z возможных значений локальных критериев $(z_1, z_2, ..., z_n) = z$. Тогда, область Z также можно подразделить на области согласия $Z^c \subset Z$ и компромисса $Z^к \subseteq Z$, для которых $Z^c \cup Z^к = Z$ и $Z^c \cap Z^к = \varnothing$.

Предположим, что все локальные критерии необходимо обратить в максимум и для простоты иллюстрации проблем векторной оптимизации рассмотрим только два критерия эффективности – z_1 и z_2: $z = (z_1, z_2) \in Z$.

Сначала проиллюстрируем геометрически проблемы 1 и 2 (определение области компромисса и выбор схемы компромисса). Для этого примем декартовую систему координат Oz_1z_2, осям которой поставим в соответствие критерии эффективности z_1 и z_2. Проследим связь между областями $X^c \leftrightarrow Z^c$ и $X^к \leftrightarrow Z^к$.

Пусть множество Z представлено в виде заштрихованной области рис. 1. Отметим, что эта область включает и границу фигуры, т.е. замкнутую линию $OABCD$. Здесь дуга BC является частью границы области, с которой совпадает область компромисса $Z^к$. Оставшаяся часть фигуры образует область согласия $Z^c = Z \setminus Z^k$, и, ей соответствует область $Z^c = Z \setminus (BC)$.

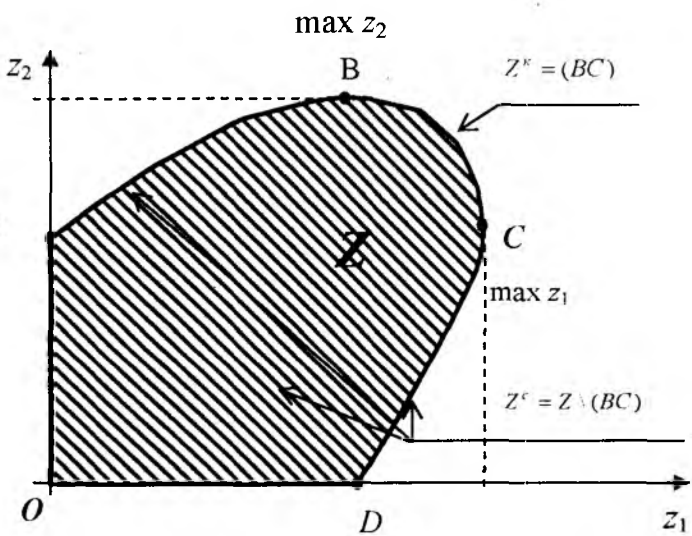


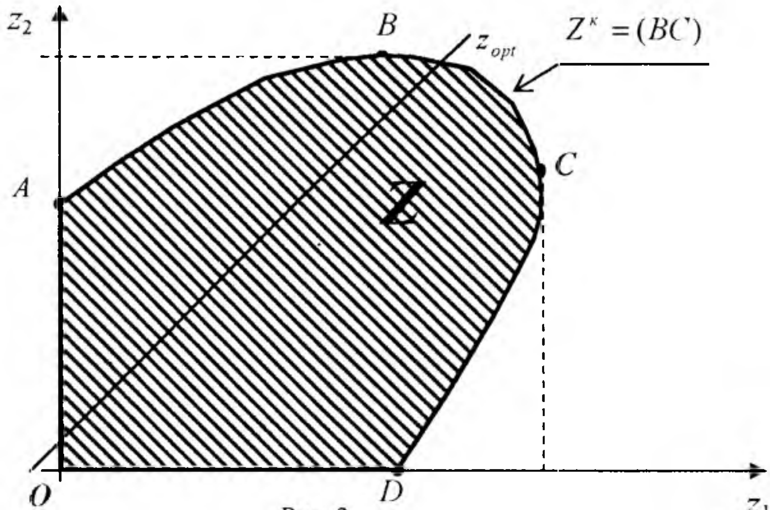
Рис. 1

Заметим, что на рис. 1 область допустимых значений Z изображена выпуклой областью. Если область не является выпуклой, то выделение области компромисса будет особенно сложным.

Решение проблемы выбора схем компромисса обусловлено рядом особенностей, одной из которых является практическая неограниченность возможных схем компромисса. Поэтому, обычно используются такие простые подходы, которые основаны на принципах равномерности и справедливой уступки.

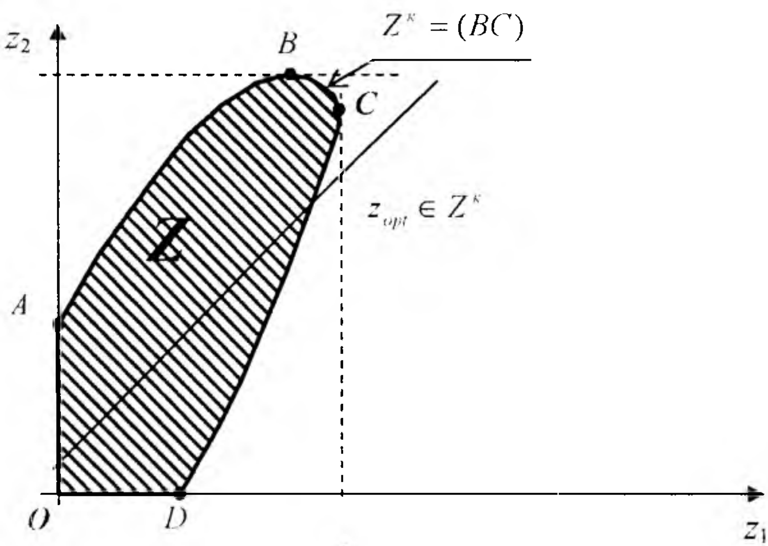
Коротко поясним идеи принципов равномерности и справедливой уступки. Принцип равномерности предполагает, что все локальные критерии должны в равной степени участвовать при решении задачи, а так же они должны равномерно и гармонично улучшатся. Одним из простых и удобных разновидностей принципа равномерности является принцип равенства. При его использовании оптимальным компромиссом считается такое, которое обеспечивает равенство всех локальных критериев (рис. 2).

Этот принцип является весьма «жестким» и можно привести примеры, в которых, его использование, может приводить к решению вне области компромисса Z^* (рис. 3).



z_2
B
z_{opt}
$Z^{\kappa} = (BC)$
C
A
Z
O
D
z_1

Рис. 2



$Z^{\kappa} = (BC)$
z_2
B
C
$z_{opt} \in Z^{\kappa}$
Z
A
O
D
z_1

Рис. 3

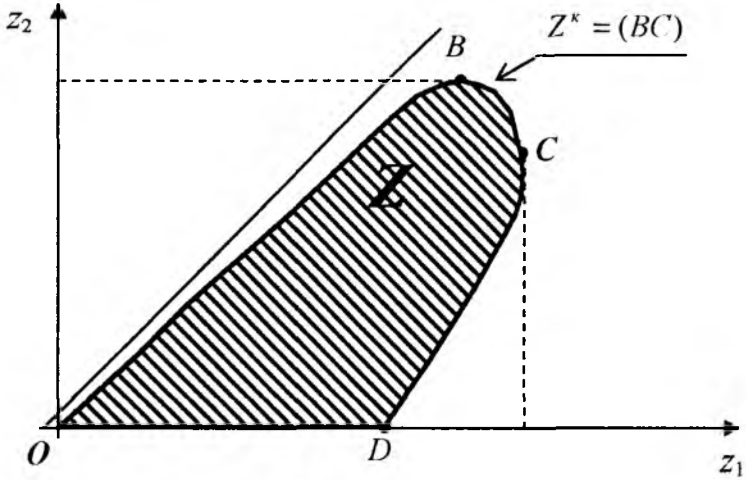


Рис. 4

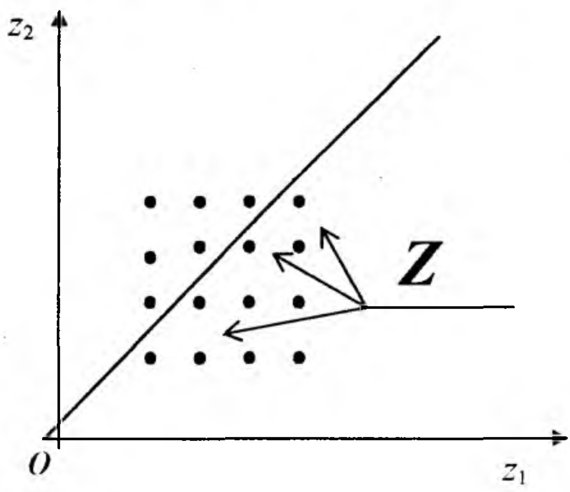


Рис. 5

В некоторых случаях, использование этого принципа может вообще не приводить к решению (рис. 4), особенно, когда рассматривается дискретная задача (рис. 5).

Математическая запись принципа равномерности, где используется оператор оптимизации opt и определяется выбор наилучшего решения в соответствии с выбранной схемой, может быть представлена следующим образом:

$$z \to \underset{z \in Z^K, z_1 = z_2 = \ldots = z_n}{opt}.$$

Выше были отмечены некоторые недостатки принципа равномерности, не позволявшие, в той или иной степени, получить решение задачи. Другим подходом, при помощи которого можно прийти к решению задачи, является следующий принцип- среди имеющегося множества критериев выбирается наихудший (т.е. принимающий наименьшее значение), который затем максимизируется. Этот принцип математически может быть записан так:

$$\min_{1 \le i \le n} z_i \to \max_{z \in Z^K}.$$

Такой подход при применении принципа равномерности принято называть принципом максимина.

Следующим принципом компромисса является принцип справедливой уступки. Его можно подразделить на два вида: принцип справедливой абсолютной уступки и принцип справедливой относительной уступки.

Идея принципа справедливой абсолютной уступки заключается в следующем. Здесь, справедливым считается такой компромисс, суммарный абсолютный уровень снижения одного или нескольких критериев при котором, не может быть больше суммарного абсолютного уровня повышения других критериев.

Пусть рассматривается векторный критерий $z=(z_1,z_2,...,z_n)$. Для сравнения значений векторов $z^1=(z_1^1,z_2^1,...,z_n^1)$ и $z^2=(z_1^2,z_2^2,...,z_n^2)$ среди локальных критериев вектора $z=(z_1,z_2,...,z_n)$ выберем $z_{i_1},z_{i_2},...,z_{i_k}$, $k=\overline{1,n-1}$. Согласно принципа справедливой абсолютной уступки $z^1=(z_1^1,z_2^1,...,z_n^1)$ считается лучше, чем $z^2=(z_1^2,z_2^2,...,z_n^2)$, если выполняется условие $\sum\limits_{i_s\in\{i_1,i_2,...,i_k\}}(z_s^2-z_s^1)\le\sum\limits_{i_s\notin\{i_1,i_2,...,i_k\}}(z_s^1-z_s^2)$.

Обозначим через $\Delta_{абс}$ величину суммарной абсолютной уступки при переходе от z^1 к z^2:

$$\Delta_{абс}=\sum_{i_s\in\{i_1,i_2,...,i_k\}}(z_s^2-z_s^1)-\sum_{i_s\notin\{i_1,i_2,...,i_k\}}(z_s^1-z_s^2)=\sum_{s=1}^{n}(z_s^2-z_s^1).$$

Тогда, согласно рассматриваемого принципа, z^1 будет лучше чем z^2, если $\Delta_{абс}\le 0$.

Проиллюстрируем принцип справедливой абсолютной уступки на примере двумерной векторной задачи. Пусть имеются два критерия эффективности z_1 и z_2, т.е. $z=(z_1,z_2)\in Z$. И пусть требуется сравнить два решения: $z^1=(9,\ 3)$ и $z^2=(5,\ 6)$, принадлежащие области Z (рис. 6). Вычислим величину суммарной абсолютной уступки

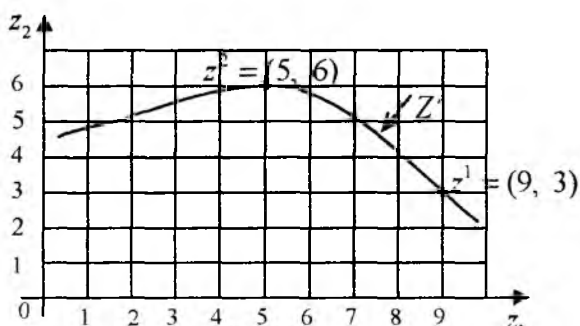


Рис. 6

$\Delta_{абс}$ при переходе от z^1 к z^2.

$$\Delta_{абс} = (z_1^2 - z_1^1) + (z_2^2 - z_2^1) = (5-9) + (6-3) = -1 < 0 .$$

Следовательно, на основании принципа абсолютной уступки решение z^1 лучше, чем z^2.

Очевидно, что если величина суммарной абсолютной уступки $\Delta_{абс}$ при переходе от некоторого решения z^0 к любому другому z будет отрицательной, то такое решение будет оптимальным, т.е. $\Delta_{абс} = \sum_{s=1}^{n}(z_s - z_s^0) = \sum_{s=1}^{n} z_s - \sum_{s=1}^{n} z_s^0 < 0$.

Последнее условие равносильно следующему $\sum_{s=1}^{n} z_s < \sum_{s=1}^{n} z_s^0$.

Таким образом, согласно принципа справедливой абсолютной уступки, оптимальным будет такое решение, которое обеспечивает оптимальность суммы локальных критериев, т.е.

$$\sum_{s=1}^{n} z_s \to \underset{z \in Z^*}{opt} .$$

Основным достоинством этого принципа является простота вычисления критерия оптимальности.

Среди недостатков этого принципа отметим то, что при различных единицах измерения локальных критериев операция суммирования их значений оказывается не выполнимой. Кроме того, еще одним недостатком этого принципа является то, что при разных уровнях значений локальных критериев, высокое значение суммарного критерия может быть достигнуто за счет высокого уровня некоторых локальных критериев.

Теперь рассмотрим принцип справедливой относительной уступки. Согласно этого принципа, справедливым считается такой компромисс, при котором суммарный относительный уровень снижения одного или нескольких критериев, не может быть больше суммарного относительного уровня повышения других критериев.

Рассмотрим векторный критерий $z=(z_1,z_2,\ldots,z_n)$. Для сравнения значений векторов $z^1=(z_1^1,z_2^1,\ldots,z_n^1)$ и $z^2=(z_1^2,z_2^2,\ldots,z_n^2)$ выберем локальные критерии $z_{i_1},z_{i_2},\ldots,z_{i_k}$, $k=\overline{1,n-1}$. Согласно принципа справедливой относительной уступки $z^1=(z_1^1,z_2^1,\ldots,z_n^1)$ считается лучше, чем $z^2=(z_1^2,z_2^2,\ldots,z_n^2)$, если выполняется условие $\sum\limits_{i_s\in\{i_1,i_2,\ldots,i_k\}}\frac{z_s^2-z_s^1}{z_s^1}\le\sum\limits_{i_s\notin\{i_1,i_2,\ldots,i_k\}}\frac{z_s^1-z_s^2}{z_s^1}$.

Обозначим через $\Delta_{отн}$ величину суммарной относительной уступки при переходе от z^1 к z^2:

$$\Delta_{отн}=\sum_{i_s\in\{i_1,i_2,\ldots,i_k\}}\frac{z_s^2-z_s^1}{z_s^1}-\sum_{i_s\notin\{i_1,i_2,\ldots,i_k\}}\frac{z_s^1-z_s^2}{z_s^1}=\sum_{s=1}^{n}\frac{z_s^2-z_s^1}{z_s^1}.$$

Тогда, согласно рассматриваемого принципа, z^1 будет лучше чем z^2, если $\Delta_{отн}=\sum\limits_{s=1}^{n}\frac{z_s^2-z_s^1}{z_s^1}\le 0$.

Можно доказать, что при некоторых условиях, согласно принципа справедливой относительной уступки, оптимальным будет такое решение, при котором произведение локальных критериев достигает максимального значения, т.е.

$$\prod_{s=1}^{n} z_s \to \max_{z\in Z}.$$

Проиллюстрируем принцип справедливой относительной уступки для выше приведенного примера двумерной векторной задачи и сравним два его решения, $z^1=(9,\ 3)$ и $z^2=(5,\ 6)$, принадлежащие области Z. Для этого вычислим величину суммарной оптимальной уступки при переходе от z^1 к z^2:

$$\Delta_{отн} = \sum_{s=1}^{n} \frac{z_s^2 - z_s^1}{z_s^1} = \frac{9-5}{5} + \frac{3-6}{6} = \frac{4}{5} + \frac{-3}{6} = 0{,}3 \cdot$$

Следовательно, согласно принципа справедливой относительной уступки, решение $z^2 = (5,\ 6)$ лучше, чем $z^1 = (9,\ 3)$.

Преимуществом этого принципа, перед принципом справедливой абсолютной уступки, является то, что здесь на процесс вычисления не мешают различные единицы измерения локальных критериев.

Заметим, что в процессе исследования проблем 1 и 2 (определение области компромисса и выбор схемы компромисса), так или иначе, осуществляется переход от векторного критерия эффективности к некоторому другому – скалярному критерию. Этот процесс, т.е. переход от векторного критерия эффективности к скалярному, является главным при упрощении задач векторной оптимизации. Другим способом упрощения задач векторной оптимизации является процедура получения обобщенного критерия эффективности, который называют сверткой векторного критерия эффективности.

Перейдем теперь к рассмотрению проблемы 3 (нормализация критериев). Здесь отметим, что решение этой проблемы, в основном, связано с введением вектора идеального качества операции. За вектор идеального качества операции можно принять вектор, установленный и оцененный на базе научных и экспериментальных исследований. Далее этот вектор обозначим через $z^u = (z_1^u, z_2^u, \ldots, z_n^u)$.

В процессе решения проблемы 3 вектор идеального качества операции z^u позволяет оптимизируемый вектор z привести к безразмерному виду. В этом случае имеем дело с вектором $z^н = \left(\frac{z_1}{z_1^u}, \frac{z_2}{z_2^u}, \ldots, \frac{z_n}{z_n^u} \right)$, который назовем нормализованным вектором. Теперь, вместо задачи оптимизации вектора z, решается задача оптимизации нормализованного вектора $z^н$.

В отдельных случаях, для нормализованного вектора, можно привести некоторые условия его ограниченности, это, например, условия, когда компоненты вектора $z^н$ могут находиться в интервале $[0,1]$, если $z_i^н > 0$, $i = \overline{1,n}$.

Выбор идеального вектора осуществляется различными способами. Это, когда идеальный вектор качества задается заранее или, когда в качестве идеального вектора принимается такой, компонентами которого являются максимальные значения локальных критериев.

Для исследования проблемы 4 (проблема учета приоритета критериев) существует некоторые подходы. Приведем два из этих подходов:

- введение ряда приоритета, который позволяет упорядочить множество индексов вектора критерия эффективности согласно приоритета каждого локального критерия;
- введение вектора весовых коэффициентов, с помощью которого определяется относительная важность каждого критерия эффективности.

Первый подход основывается на субъективных мнениях экспертов или же лиц, принимающих решение.

Второй подход так же основывается на экспертных оценках лиц, ответственных за принятие решения. В то же время, решающее значение имеет, так называемый вектор весовых коэффициентов $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, который определяет относительную важность каждого критерия. Компоненты вектора α обычно обладают следующими свойствами: $0 \le \alpha_i \le 1$, $i = \overline{1,n}$, $\sum_{i=1}^{n} \alpha_i = 1$.

При использовании первого подхода, для решения задачи с векторным критерием эффективности, часто применяется принцип жесткого приоритета, который требует некоторого упорядочения локальных критериев.

Решение задачи оптимизации, в этом случае, практически сводится к последовательной оптимизации локальных критериев, начиная с критерия, обладающего высшим приоритетом. Преимущество этого подхода состоит в том, что он не требует задания весовых коэффициентов, которые обязательно должны присутствовать при практической реализации второго подхода. Недостатком же этого подхода является то, что процесс решения задачи может практически заканчиваться после оптимизации первого, самого важного критерия.

4.2. Способы свертки критериев

Как было отмечено выше, сверткой векторного критерия эффективности называлась процедура получения критерия объединенной операции. При этом обобщенный критерий эффективности z_Σ получается как функция частных критериев эффективности $z_1, z_2, \ldots, z_n$, т.е. $z_\Sigma = F(z_1, z_2, \ldots, z_n)$. Рассмотрим некоторые элементарные способы свертки, предполагая, что требуется максимизировать все частные критерии $z_1, z_2, \ldots, z_n$, а также обобщенный критерий z_Σ.

1. Суммирование, или «экономический» способ свертки. Предположим, что заданы некоторые параметры $\alpha_i, i = \overline{1, n}$, которые могут иметь самый различный смысл, а именно определять относительную важность каждого из частных критериев и удовлетворять условия: $0 \le \alpha_i \le 1$, $i = \overline{1, n}$, $\sum_{i=1}^{n} \alpha_i = 1$, где частные критерии эффективности имеют различный масштаб измерения. Эти параметры $\alpha_i, i = \overline{1, n}$ назовем нормирующими множителями.

Идея этого способа заключается в том, что вместо задачи максимизации векторного критерия, рассматривается задача максимизации критерия объединенной операции

$$z_\Sigma = \sum_{i=1}^{n} \alpha_i z_i ,$$

который получается в результате суммирования всех частных критериев с учетом нормирующих множителей.

Суммированию частных критериев эффективности, т. е. экономическому способу свертки критериев, соответствует один из принципов компромисса, а именно – принцип справедливой абсолютной уступки.

2. Способ свертки, основанный на представлении обобщенного критерия в виде качественного. Предположим, что задан некоторый вектор $z^0 = (z_1^0, z_2^0, \ldots, z_n^0)$, компоненты которого определяют необходимые уровни достижения частных критериев. Этот способ свертки критериев предполагает переход от векторного критерия z к критерию z_Σ согласно

$$z_\Sigma = \begin{cases} 1, \text{ если } z_j \geq z_j^0, \ j = \overline{1,n}, \\ 0, \text{ или } -\infty \text{ в противном случае.} \end{cases}$$

Этот способ свертки имеет право на существование. Отметим, что способ свертки, основанный на представлении обобщенного критерия в виде качественного, полностью зависит от объективности вектора z^0 и не может быть применен в случае его неизвестности.

3. Способ свертки, основанный на последовательном достижении частных целей. Каждая последующая операция, при этом способе свертки, учитывается лишь тогда, когда достигнуты абсолютные максимальные значения критериев предыдущих операций.

4. Логическое свертывание критериев. Этот способ свертки критериев применяется, только тогда, когда все частные критерии являются качественными, т.е. прини-

мают значения 0 или 1. В этом случае обобщенные критерии можно получить следующими способами:

а) введением противоположной цели, такой, которая приводит к невыполнению какой-то цели;

б) способом логического умножения всех частных критериев;

в) способом логического сложения всех частных критериев.

5. Обобщенное логическое свертывание. Этот способ свертки является прямым обобщением действий предыдущего способа, позволяющим проводить объединение качественных целей.

6. Случайное и неопределенное свертывание. Здесь, обобщенным критерием эффективности может быть объявлен любой из частных критериев в зависимости от состояния неконтролируемых факторов.

Проблемные задания

1. Сформулируйте задачу, обусловленную проблемами векторной оптимизации, в которой все локальные критерии следовало бы обратить в максимум.

2. Для определения локальных критериев задачи задания 1 исследуйте принцип равномерной справедливой уступки

3. Для решения задачи задания 1 исследуйте принцип максимина.

4. Для решения задачи задания 1 исследуйте принципы относительной и абсолютной уступки.

5. Для решения задачи задания 1 рассмотрите способы свертки критериев.

Вопросы для самоподготовки

1. В чем заключается суть принципа равномерности и справедливости уступки?

2. Разностью какого принципа является принцип максимина?

3. В чем состоит суть принципа максимина?

4. Охарактеризуйте принцип справедливой относительной уступки.

5. Назовите основные способы свертки критериев.

6. При каком способе свертки критериев все частные критерии являются качественными?

7. В чем заключается цель логического способа сложения критериев?

8. Определите суть обобщенного логического свертывания критериев.

ГЛАВА III.

§ 5. ЭЛЕМЕНТЫ ТЕОРИИ ИГР. ИГРА ДВУХ ЛИЦ С НУЛЕВОЙ СУММОЙ

Игра, конфликт, конфликтная ситуация. правила игры, парная игра, множественная игра, коалиция, игрок, стратегия, ход, платежная функция, выигрыш, проигрыш, нулевая сумма, функция полезности, платежная матрица, конечная игра

5.1. Понятие игры и способы ее описания

С появлением в 1944 году монографии Дж. фон Неймана и О. Моргенштерна «Теория игр и экономическое поведение» возникла теория игр. Теория игр является математической теорией конфликтных ситуаций и занимается выработкой рекомендаций по рациональному образу действий участников многократно повторяющегося конфликта. Ситуации, в которых эффективность принимаемого одной стороной решения зависит от действий другой стороны, называются конфликтными. Конфликт всегда связан с определенного рода разногласиями.

Игра представляет собой математическую модель реальной конфликтной ситуации, анализ которой ведется по определенным правилам.

Стороны, участвующие в игре называются игроками. Игроками могут быть отдельные лица или команды, воюющие стороны, предприятия, фирмы и, наконец, природа. формирующая условия, в которых необходимо принимать решения.

Стратегией игрока называется совокупность правил, определяющих выбор варианта действий каждым игроком в зависимости от ситуации, сложившейся в процессе игры.

Выбор одной из предусмотренных правилами игры стратегий и ее осуществление называется ходом. Ходы бывают личные и случайные.

В общем случае правилами игры устанавливаются последовательность ходов, объем информации каждой стороны о поведении другой и результат (исход) игры. Правила определяют также конец игры, когда некоторая возможная последовательность выборов уже сделана, и, больше ходов делать не разрешается.

В зависимости от числа участников, игры подразделяются на парные и множественные. В парной игре число участников равно двум, а во множественной – более двух. Участники множественной игры могут образовывать коалиции и игры в этом случае называются коалиционными.

Игра называется конечной, если число стратегий игроков конечно, и бесконечной, если, хотя бы у одного из игроков, число стратегий является бесконечным.

Стратегия игрока называется оптимальной, если, независимо от поведения противника, при многократном повторении игры, она обеспечивает данному игроку максимально возможный средний выигрыш или минимально возможный средний проигрыш. Отметим, что здесь могут быть использованы и другие показатели оптимальности.

Существуют два способа описания игр: позиционный и нормальный. Позиционный способ связан с развернутой формой игры и сводится к графику последовательных шагов (дереву игры). Нормальный способ заключается в явном представлении совокупности стратегий игроков и платежной функции. Для каждой совокупности выбранных игроками стратегий, платежная функция определяет выигрыш каждой из сторон.

Если в парной игре выигрыш одного игрока равен проигрышу второго, то такую игру принято называть игрой с нулевой суммой.

5.2. Игра двух лиц с нулевой суммой

Рассмотрим конечную игру двух лиц (I и II) с нулевой суммой. Предположим, что игрок I имеет m стратегий (обозначим их $A_1, A_2, \ldots, A_m$), а игрок II (противник игрока I) – n стратегий ($B_1, B_2, \ldots, B_n$). Такая игра называется игрой размерности $m \times n$.

Игра состоит из двух ходов: игрок I выбирает стратегию A_i, $i = \overline{1,m}$, а игрок II выбирает стратегию B_j, $j = \overline{1,n}$, (каждый выбор производится при полном незнании выбора другого игрока), после чего игроки получают соответственно выигрыши $w_1(A_i, B_j)$ и $w_2(A_i, B_j)$.

Так как рассматривается игра с нулевой суммой, то имеем $w_1(A_i, B_j) + w_2(A_i, B_j) = 0$. Выразим это равенство следующим образом $w_1(A_i, B_j) = w(A_i, B_j)$, $w_2(A_i, B_j) = -w(A_i, B_j)$.

Пусть, $w(A_i, B_j) = a_{ij}$, где значения a_{ij} известны при каждой паре стратегий A_i и B_j. Матрица $A = (a_{ij})$, $i = \overline{1,m}$, $j = \overline{1,n}$, называется платежной матрицей. Запишем эти значения в виде таблицы 1.

Строки этой таблицы соответствуют стратегиям игрока I, а столбцы – стратегиям игрока II. Каждый положительный элемент a_{ij} матрицы определяет величину выигрыша игрока I и проигрыша игрока II при применении ими соответствующих стратегий. Целью каждого игрока является максимизировать свой выигрыш или, что то же самое, – минимизировать свой проигрыш.

Таблица 1

II / I	B_1	B_2	...	B_n
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
...	...	...	...	...
A_m	a_{m1}	a_{m2}	...	a_{mn}

Проблемные задания

1. Оставьте задачу и составьте ее математическую модель, качественно отражающую игровую ситуацию.
2. Для задачи задания 1 исследуйте и определите:

а) число участников;

б) характер игры;

в) стратегии игроков.

3. Поставьте такую задачу, чтобы ее математическая модель отражала парную игровую ситуацию.

Вопросы для самоподготовки

1. Что такое игра?
2. Что называется стратегией?
3. Что называется оптимальной стратегией?
4. Какие вы знаете способы описания игр?
5. Что определяет в игре платежная функция?
6. Дайте определение игры с нулевой суммой.

§ 6. РЕШЕНИЕ МАТРИЧНЫХ ИГР ДВУХ ЛИЦ С НУЛЕВОЙ СУММОЙ. ПРИНЦИП МИНИМАКСА

Игра, платежная матрица, стратегия, максиминная стратегия, минимаксная стратегия, нижняя цена игры, верхняя цена игры, седловая точка, чистая стратегия, смешанная стратегия

6.1. Игра $m \times n$: основные свойства игры

Рассмотрим игру $m \times n$ с платежной матрицей, представленной в следующей таблице:

II / I	B_1	B_2	...	B_n
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{12}	a_{22}	...	a_{2n}
...	...	...	...	...
A_m	a_{m1}	a_{m2}	...	a_{mn}

Следует определить: а) наилучшую стратегию игрока I среди стратегий $A_1, A_2, \ldots, A_m$; б) наилучшую стратегию игрока II среди стратегий $B_1, B_2, \ldots, B_n$.

При определении наилучших стратегий игроков будем считать, что противники, участвующие в игре, одинаково разумны, и каждый из них делает все для того, чтобы добиться своей цели.

Используя этот принцип, найдем наилучшую стратегию игрока I, для чего проанализируем последовательно все его стратегии. Выбирая стратегию A_i, игрок I должен рассчитывать, что противник ответит на нее той из своих стратегий B_j, для которой выигрыш игрока I будет минимальным.

Вычислим $\alpha_i = \min\limits_{1 \le j \le n} a_{ij}$, $i = \overline{1, m}$, и запишем эти числа рядом с платежной матрицей в добавочный столбец.

Зная числа α_i, игрок I должен предпочесть другим стратегиям ту, для которой α_i максимально. Обозначим $\alpha = \max\limits_{1 \le i \le m} \alpha_i$, тогда $\alpha = \max\limits_{1 \le i \le m} \min\limits_{1 \le j \le n} a_{ij}$.

Величина α – гарантированный выигрыш, который может обеспечить себе игрок I, называется нижней ценой игры (максимином).

Стратегия, обеспечивающая получение нижней цены игры α, называется максиминной стратегией.

Если игрок I будет придерживаться своей максиминной (перестраховочной) стратегии, то ему гарантирован выигрыш, не меньший α при любом поведении игрока II.

Игрок II заинтересован уменьшить свой проигрыш или, что, то же самое, выигрыш игрока I обратить в минимум. Поэтому для выбора своей наилучшей стратегии он должен найти максимальное значение выигрыша в каждом из столбцов и среди этих значений выбрать наименьшее.

Обозначим $\beta_j = \max\limits_{1 \le i \le m} a_{ij}$ и $\beta = \min\limits_{1 \le j \le n} \beta_j$. Тогда, $\beta = \min\limits_{1 \le j \le n} \max\limits_{1 \le i \le m} a_{ij}$.

Стратегия игрока II, обеспечивающая «выигрыш» β является его минимаксной стратегией.

Если игрок II будет придерживаться своей минимаксной стратегии, то в любом случае проиграет не больше β.

Существуют игры, для которых $\alpha = \beta$. Такие игры называются играми с седловой точкой.

Общее значение нижней и верхней цены в играх с седловой точкой называется чистой ценой игры, а стратегии A_i^* и B_j^*, позволяющие достичь этого значения – оптимальными.

Оптимальные стратегии и чистая цена являются решением игры с седловой точкой.

Чистую цену игры $\gamma = \alpha = \beta$ в игре с седловой точкой, при условии одинаковой разумности партнеров, игрок I не может увеличить, а игрок II – уменьшить.

Если игра имеет седловую точку, то говорят, что она решается в чистых стратегиях. Под чистой стратегией понимается такая стратегия, которая выбрана игроком сознательно, без использования механизма случайного выбора.

Следует отметить, что платежная матрица игры может иметь более одной седловой точки.

Таким образом, если матрица игры содержит седловую точку, то ее решение находится по принципу минимакса.

В основной теореме игр утверждается, что любая конечная игра двух лиц с нулевой суммой имеет, по крайней мере, одно решение, возможно, в смешанных стратегиях, т.е., каждая конечная игра имеет цену.

Цена игры γ – средний выигрыш, приходящийся на одну партию – всегда удовлетворяет условию $\alpha \le \gamma \le \beta$. Следовательно, каждый игрок при многократном повторении игры, придерживаясь смешанных стратегий, получает более выгодный для себя результат.

6.2. Решение игры $m \times n$ без седловой точки

Теперь рассмотрим игру $m \times n$, для которой $\alpha < \beta$. Здесь, применение минимаксных стратегий каждым из игроков обеспечивает первому выигрыш не меньше нижней цены игры α, а второму – проигрыш не больше верхней ценой игры β. Учитывая, что $\alpha < \beta$, естественным желанием игрока I является увеличить свой выигрыш, а игрока II – уменьшить свой проигрыш. Поиск такого решения приводит к применению сложной стратегии, состоящей в случайном применении двух и более чистых стратегий с определенными частотами.

Такая сложная стратегия в теории игр называется смешанной.

Смешанные стратегии игроков I и II обозначим соответственно через $p_A = (p_1, p_2, ..., p_m)$ и $q_B = (q_1, q_2, ..., q_n)$, где $p_i \geq 0$, $q_j \geq 0$ – вероятности применения чистых стратегий A_i и B_j, $i = \overline{1,m}$, $j = \overline{1,n}$, при этом $\sum_{i=1}^{m} p_i = 1$, $\sum_{j=1}^{n} q_j = 1$.

Стратегии игроков, входящие в их оптимальные смешанные стратегии, называются активными.

Теорема. Применение оптимальной смешанной стратегии обеспечивает игроку максимальный средний выигрыш (или минимальный средний проигрыш), равный цене игры γ, независимо от того, какие действия предпринимает другой игрок, если только он не выходит за пределы своих активных стратегий.

Доказательство. Предположим, что найдено оптимальное решение игры $m \times n$ в смешанных стратегиях, в котором первые r стратегий ($r \leq m$) игрока I и первые s стратегий ($s \leq n$) игрока II являются активными (это не нарушает общности, т.к. стратегии всегда можно перенумеровать таким образом, чтобы первыми были активные), т.е.

$$p_A^* = (p_1, p_2, ..., p_r, 0, ..., 0), \quad \sum_{i=1}^{r} p_i = 1,$$

$$q_B^* = (q_1, q_2, ..., q_s, 0, ..., 0), \quad \sum_{j=1}^{s} q_j = 1.$$

Выигрыш, полученный в результате применения этих стратегий, равен цене игры γ.

Выигрыш игрока I, если он пользуется оптимальной смешанной стратегией p_A^*, а игрока II – чистыми стратегиями $B_1,B_2,\ldots,B_s$, обозначим через $\gamma_1, \gamma_2,\ldots, \gamma_s$. Из свойства оптимального решения игры следует, что отклонение игрока II от оптимальной стратегии q_B^* может лишь увеличить его проигрыш. Следовательно, $\gamma_j \geq \gamma, j=\overline{1,s}$.

Выразим теперь цену игры γ, при оптимальных смешанных стратегиях игроков p_A^* и q_B^*, через γ_1, $\gamma_2,\ldots, \gamma_s$. Поскольку, в оптимальной смешанной стратегии q_B^*, чистые стратегии $B_1,B_2,\ldots,B_s$ применяются с вероятностями $q_1,q_2,\ldots,q_s$, то $\gamma=\gamma_1 q_1+\gamma_2 q_2+\ldots+\gamma_s q_s$, при этом $\sum_{j=1}^{s} q_j = 1$.

Сумма $\gamma_1 q_1+\gamma_2 q_2+\ldots+\gamma_s q_s$ есть средневзвешенное значение, которое было бы больше γ, если хотя бы один из выигрышей γ_j был больше γ. Следовательно $\gamma_1=\gamma_2=\ldots=\gamma_s=\gamma$.

Рассмотрим наиболее простую игру 2×2:

	B_1	B_2
A_1	a_{11}	a_{12}
A_2	a_{21}	a_{22}

В этой игре каждый из игроков обладает только двумя стратегиями. Если игра 2×2 имеет седловую точку, то ее решение очевидно.

Пусть $\alpha \neq \beta$. Требуется найти оптимальные смешанные стратегии игроков $p_A^*=(p_1,p_2)$, $q_B^*=(q_1,q_2)$ и цену игры γ.

Очевидно, что в игре 2×2, не имеющей седловую точку, обе стратегии игроков являются активными. Поэтому, согласно теореме об активных стратегиях, если игрок I будет применять свою оптимальную смешанную стратегию, то независимо от действий игрока II, выигрыш его будет равен цене игры γ.

Поскольку, игрок I для получения оптимального выигрыша применяет стратегию A_1 с вероятностью p_1 и стратегию A_2 с вероятностью p_2, то, если игрок II применит стратегию B_1, тогда значение выигрыша игрока I определится из уравнения $a_{11}p_1 + a_{21}p_2 = \gamma$.

Если же игрок II применит стратегию B_2, то выигрыш игрока I не изменится и определится равенством $a_{12}p_1 + a_{22}p_2 = \gamma$.

Принимая во внимание условие $p_1 + p_2 = 1$, будем иметь систему трех линейных уравнений с тремя неизвестными p_1, p_2 и γ. Решив эту систему уравнений находим $p_A^* = (p_1, p_2)$ и γ.

Оптимальная стратегия $q_B^* = (q_1, q_2)$ игрока II, аналогично определится из системы уравнений:

$$\begin{aligned} &a_{11}q_1 + a_{12}q_2 = \gamma, \\ &a_{21}q_1 + a_{22}q_2 = \gamma, \\ &q_1 + q_2 = 1. \end{aligned}$$

Проблемные задания

1. Составьте игровую задачу двух лиц, решение которой возможно только в смешанных стратегиях.
2. Исследуйте возможность построения оптимальных смешанных стратегий для игры, составленной в задании 1.

Вопросы для самоподготовки

1. Что называется нижней ценой игры?
2. Что такое максиминная стратегия?

3. Что такое минимаксная стратегия?
4. Что называется игрой с седловой точкой?
5. Что такое чистая цена игры?
6. Что называется смешанной стратегией?
7. Что называется активной стратегией?

§ 7. УПРОЩЕНИЕ ИГР. СВЕДЕНИЕ МАТРИЧНОЙ ИГРЫ К ЗАДАЧЕ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Игра, платежная матрица, седловая точка, дублирующие стратегии, заведомо невыгодные стратегии, конечная игра, линейное программирование

7.1. Основные определения упрощения игр

Если платежная матрица игры не содержит седловой точки, то определение оптимальных смешанных стратегий тем сложнее, чем больше размерность матрицы. Поэтому, перед решением игры, целесообразно уменьшить (насколько это возможно) размерность ее платежной матрицы. Это осуществляется путем вычеркивания дублирующих и заведомо невыгодных стратегий, а также замены некоторых групп чистых стратегий – смешанными.

Определение 1. Если в платежной матрице игры все элементы строки (столбца) равны соответствующим элементам другой строки (столбца), то соответствующие строкам (столбцам) стратегии называются дублирующими.

Определение 2. Если в платежной матрице $\{a_{ij}\}$ игры все элементы некоторой строки (столбца), определяющей стратегию A_i (B_j) игрока I (игрока II) не больше (не меньше) соответствующих элементов другой строки (столбца), то стратегия A_i (B_j) называется заведомо невыгодной.

7.2. Решение игр $2 \times n$ и $m \times 2$

Любая конечная игра $m \times n$ имеет решение, в котором число активных стратегий каждого игрока не превосходит $\min(m,n)$. Следовательно, у игр $2 \times n$ и $m \times 2$ всегда имеется решение, содержащее не более двух активных стратегий у каждого из игроков. Если эти активные стратегии игроков будут найдены, то игра $2 \times n$ или $m \times 2$ превращается в игру 2×2, решение которой осуществляется элементарно.

Практически решение игры $2 \times n$ осуществляется следующим образом:

-строится графическое изображение игры;

-выделяется нижняя граница выигрыша и находится наибольшая ордината нижней границы, которая равна цене игры γ;

-определяется пара стратегий, пересекающихся в точке оптимума.

Решение игры $m \times 2$ осуществляется аналогично.

Рассмотрим игру $2 \times n$, представленную в следующей таблице:

II I	B_1	B_2	$\cdots$	B_n
A_1	a_{11}	a_{12}	$\cdots$	a_{1n}
A_2	a_{21}	a_{22}	$\cdots$	a_{2n}

Обозначим через $p_A^* = (p_1, p_2)$ оптимальную смешанную стратегию игрока I. Поскольку, игрок I имеет только две стратегии и игра не имеет седловой точки, то обе стратегии будут активными и $p_2 = 1 - p_1$, $p_1 \geq 0$, $p_2 \geq 0$.

Ожидаемый выигрыш γ_j игрока I, соответствующий стратегии B_j игрока II, равен

$$\gamma_j = a_{1j}p_1 + a_{2j}p_2 = a_{1j}p_1 + a_{2j}(1 - p_1) = (a_{1j} - a_{2j})p_1 + a_{2j}.$$

Отсюда видно, что ожидаемый выигрыш γ_j игрока I линейно зависит от p_1.

В соответствии с критерием минимакса, для игр в смешанных стратегиях игрок I должен выбирать p_1 так, чтобы максимизировать свой минимальный ожидаемый выигрыш.

Эта задача может быть решена графически построением прямых линий, соответствующих линейным функциям от p_1.

Пример. Найти решение и дать геометрическую интерпретацию игры, платежная матрица которой представлена в следующей таблице:

II I	B_1	B_2	B_3
A_1	5	−1	3
A_2	2	4	−2

Эта игра не имеет седловой точки, поскольку нижняя цена $\alpha = -1$, а верхняя цена $\beta = 3$, т.е. $\alpha < \beta$. Ожидаемые выигрыши игрока I, соответствующие чистым стратегиям B_j, $j = 1,2,3$, будут равны: $\gamma_1(p_1) = 3p_1 + 2$ при $j = 1$; $\gamma_2(p_1) = -5p_1 + 4$ при $j = 2$ и $\gamma_3(p_1) = 5p_1 - 2$ при $j = 3$.

Изобразим три прямые, являющиеся графиками функций $\gamma_1(p_1)$, $\gamma_2(p_1)$, $\gamma_3(p_1)$ (рис. 1). В соответствии с критерием минимакса игрок I должен выбирать p_1 так, чтобы максимизировать свой минимальный ожидаемый выигрыш. Минимальный ожидаемый выигрыш игрока I на рис.1 изображен в виде ломанной ABC. Максимум минимального ожидаемого выигрыша игрока I достигается при $p_1^* = \frac{3}{5}$.

Следовательно, $p_2^* = 1 - p_1^* = 1 - \frac{3}{5} = \frac{2}{5}$ и $\gamma^* = -5 \cdot \frac{3}{5} + 4 = 1$. Таким образом, найдены оптимальная стратегия игрока I $p_I^* = (p_1, p_2)$ и цена игры γ^*: $p_I^* = \left(\frac{3}{5}, \frac{2}{5}\right)$, $\gamma^* = 1$.

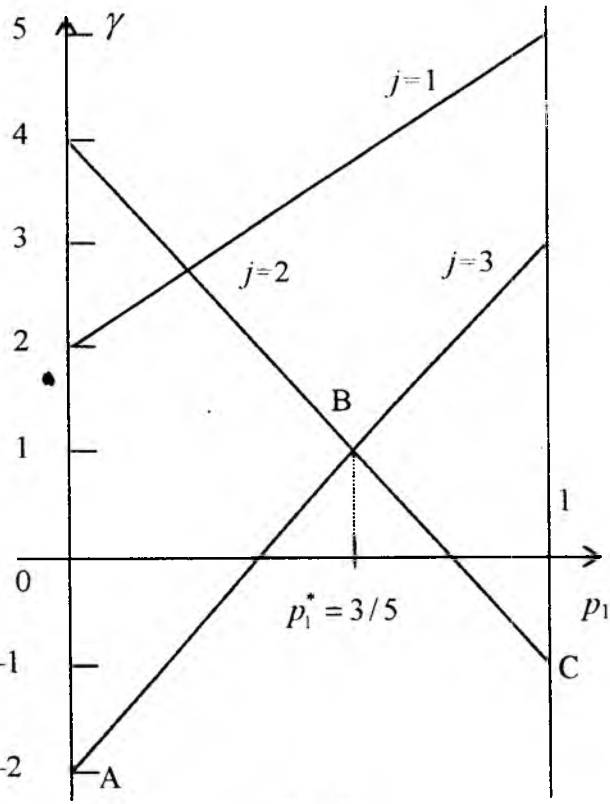


Рис. 1

7.3. Сведение матричной игры к задаче линейного программирования

Рассмотрим игру $m \times n$. Будем считать, что все элементы платежной матрицы неотрицательны (если это не так, то можно ко всем элементам матрицы добавить некоторое достаточно большое число L, переводящее платежи в область неотрицательных значений, при этом цена игры увеличится на L, а решение задачи не изменится). Следовательно, можно принять, что $\gamma > 0$.

Пусть платежная матрица игры не имеет седловой точки. Следовательно, игра решается в смешанных стратегиях. Оптимальные смешанные стратегии игроков I и II обозначим соответственно через $p_A^* = (p_1, p_2, ..., p_m)$ и $q_B^* = (q_1, q_2, ..., q_n)$.

Применяя оптимальную смешанную стратегию p_A^* игрок I гарантирует себе, независимо от поведения игрока II, выигрыш, не меньший цены игры γ.

Допустим, что игрок II применяет свою чистую стратегию B_j, а игрок I – свою оптимальную стратегию p_A^*. Тогда средний выигрыш игрока I будет равен

$$\gamma_j = a_{1j} p_1 + a_{2j} p_2 + ... + a_{mj} p_m, \quad j = \overline{1,n}.$$

Учитывая, что γ_j, $j = \overline{1,n}$, не может быть меньше γ, можем записать условия:

$$\sum_{i=1}^{m} a_{ij} p_i \geq \gamma, \; j = \overline{1,n}.$$

Введя обозначение $x_i = \frac{p_i}{\gamma}$, $i = \overline{1,m}$, получим

$$\sum_{i=1}^{m} a_{ij} x_i \geq 1, \; j = \overline{1,n}.$$

Поскольку $p_i \geq 0$ и $\gamma > 0$, то имеем $x_i \geq 0$.

Из равенства $\sum_{i=1}^{m} p_i = 1$ следует, что $\sum_{i=1}^{m} x_i = \frac{1}{\gamma}$. Учитывая, что игрок I стремится максимизировать γ, приходим к задаче минимизации линейной функции $\sum_{i=1}^{m} x_i$.

Следовательно, задача решения игры свелась к следующей задаче линейного программирования:

– найти такой вектор $x = (x_1, x_2, ..., x_m)$, который при условиях $\sum_{i=1}^{m} a_{ij} x_i \geq 1, j = \overline{1,n}$, $x_i \geq 0$, $i = \overline{1,m}$, даст минимальное значение функции $\sum_{i=1}^{m} x_i$, т.е.: $\sum_{i=1}^{m} x_i \to \min$, $\sum_{i=1}^{m} a_{ij} x_i \geq 1, j = \overline{1,n}$, $x_i \geq 0$, $i = \overline{1,m}$.

Решая задачу линейного программирования, определяем ее оптимальный план $x^* = (x_1^*, x_2^*, ..., x_m^*)$. Затем, используя x^*, по формулам: $\gamma = \frac{1}{\sum_{i=1}^{m} x_i^*}$, $p_i = \frac{x_i}{\sum_{i=1}^{m} x_i^*}$, $i = \overline{1,m}$ находим цену игры γ и оптимальную стратегию p_A^* игрока I.

Рассуждая аналогично, можно утверждать, что оптимальную стратегию $q_B^* = (q_1, q_2, ..., q_n)$ игрока II можно найти по формуле $q_j = \frac{u_j^*}{\sum_{j=1}^{n} u_j}$, $j = \overline{1,n}$ где $u^* = (u_1^*, u_2^*, ..., u_n^*)$, есть оптимальный план задачи линейного программирования

$$\sum_{j=1}^{n} u_j \to \max, \quad \sum_{j=1}^{n} a_{ij} u_j \leq 1, \; i = \overline{1,m}, \quad u_j \geq 0, \; j = \overline{1,n}.$$

Эта и предыдущая задачи линейного программирования являются взаимно двойственными.

Таким образом, игра $m \times n$ свелась к паре взаимно двойственных задач линейного программирования.

Проблемные задания

1. Составьте платежную матрицу игры, которую:

а) возможно свести к игре 2×2;

б) невозможно свести к игре 2×2.

2. Решите задачу задания 1 в обоих случаях.

3. Приведите задачу задания 1 к задаче линейного программирования.

4. Составьте и решите игровые задачи $2\times n$ и $m\times2$ для различных $n>2$ и $m>2$.

Вопросы для самоподготовки

1. Какие стратегии называются дублирующими?

2. Какие стратегии называются заведомо невыгодными?

3. Опишите последовательность графического решения игры $2\times n$.

4. Напишите математическую модель матричной игры, сведенной к задаче линейного программирования.

§ 8. ИГРЫ С ПРИРОДОЙ

Игра, игра с природой, операция, риск, вероятность, оптимальная стратегия, матрица рисков, средние значения, оптимальное решение, эксперимент, неопределенность, планирование

8.1. Определение игры с природой. Игровая постановка задачи принятия решения в условиях неопределенности

Участники вышерассмотренных игр принимают решение, исходя из своих разумных действий, и, являются антагонистическими противниками. Каждый из участников, в этих играх, предпринимает именно те действия, которые наиболее выгодны ему и менее выгодны противнику. Однако, встречаются и такие ситуации, которые сопровождаются неконтролируемыми, случайными возмущениями. Эти факторы находятся вне сознательных действий участников игры. Зачастую, на них нельзя оказать практического воздействия. Они зависят от некой, неизвестной участникам игры объективной действительности – природы. Примечательно то, что в этом случае «природу» также можно считать участником игры, т.е. игроком. Такого рода ситуации принято называть играми с природой или же статистическими играми.

В игре с природой игрок II (природа) не является разумным игроком, так как рассматривается как некая незаинтересованная сторона, которая не выбирает для себя оптимальных стратегий. Здесь возможные состояния природы (ее стратегии) реализуются случайным образом.

Рассмотрим игровую постановку задачи принятия решения в условиях неопределенности.

Пусть оперирующей стороне необходимо выполнить операцию в недостаточно известной обстановке, относительно состояний которой, можно сделать n предположений. Эти предположения обозначим через $П_1, П_2, \ldots, П_n$ и они будут рассматриваться в качестве стратегий природы. Предположим, что оперирующая сторона в своем распоряжении имеет m возможных стратегий — $A_1, A_2, \ldots, A_m$.

Эффект, получаемый игроком I от использования им своей стратегии A_i, $i=\overline{1,m}$, при состоянии (стратегии) природы $П_j$, $j=\overline{1,n}$, обозначим через a_{ij}, $i=\overline{1,m}$, $j=\overline{1,n}$, и назовем выигрышем игрока I. Этот выигрыш, при каждой паре стратегий A_i и $П_j$, предполагается известным и задается платежной матрицей $A=\left(a_{ij}\right)_{i=\overline{1,m},\, j=\overline{1,n}}$.

Задача заключается в определении такой стратегии (чистой или смешанной), которая, при ее применении, обеспечила бы игроку I наибольший выигрыш.

Для решения этой задачи необходимо сначала анализировать матрицу выигрышей, т.е. выявлять и отбрасывать дублирующие и заведомо невыгодные стратегии игрока I. Но ни одну из стратегий природы отбрасывать нельзя.

После упрощения платежной матрицы игры с природой целесообразно не только оценить выигрыш при той или иной игровой ситуации, но и определить разность между максимально возможным выигрышем при данном состоянии природы и выигрышем, который будет получен при применении стратегии A_i в тех же условиях. Эта разность называется риском.

Пусть $\beta_j = \max\limits_{1 \le i \le m} a_{ij}$, $j=\overline{1,n}$. Риск игрока I, при применении им стратегии A_i в условиях $П_j$, обозначим через r_{ij}, $i=\overline{1,m}$, $j=\overline{1,n}$. Тогда, $r_{ij} = \beta_j - a_{ij} \ge 0$.

Во многих случаях матрица рисков $R = (r_{ij})_{i=\overline{1,m},\, j=\overline{1,n}}$ позволяет более глубоко понять неопределенную ситуацию, чем матрица выигрышей A.

В некоторых случаях удается снизить степень неопределенности ситуации, которая достигается нахождением вероятностей состояний природы на основе данных статистических наблюдений. Предположим, что вероятности состояний природы известны:

$$p(\Pi_j) = Q_j,\quad j=\overline{1,n},\quad \sum_{j=1}^{n} Q_j = 1.$$

Обозначим через $\overline{\sigma}_i$, $i=\overline{1,m}$, среднее значение (математическое ожидание) выигрыша, которое можно определить по формуле:

$$\overline{\sigma}_i = \sum_{j=1}^{n} a_{ij} Q_j,\ i=\overline{1,m}.$$

Игрок I стремится максимизировать свой выигрыш, т.е. ему необходимо найти такую стратегию A_i, которая обеспечила бы ему максимальное значение $\overline{\sigma}$ среднего выигрыша:

$$\overline{\sigma} = \max_{1\le i\le m} \overline{\sigma}_i = \max_{1\le i\le m}\left\{\sum_{j=1}^{n} a_{ij} Q_j\right\}.$$

Оптимальную стратегию игрока I, при известных вероятностных состояний природы, можно найти, используя понятие риска. Для этого, сначала необходимо найти среднее значение риска $\overline{r}_i$, $i=\overline{1,m}$ при применении игроком I стратегии A_i:

$$\overline{r}_i = \sum_{j=1}^{n} r_{ij} Q_j,\ i=\overline{1,m}.$$

В качестве оптимальной стратегии выбирается та, которая обеспечивает минимальное среднее значение риска $\overline{r}$:

$$\bar{r} = \min_{1 \le i \le m} \bar{r}_i = \min_{1 \le i \le m} \left\{ \sum_{j=1}^{n} r_{ij} Q_j \right\}.$$

Теперь рассмотрим ситуацию, когда объективные оценки состояний получить невозможно. В этом случае, вероятности состояний природы могут быть оценены субъективно на основе следующих принципов:

1) Принцип недостаточного основания Лапласа, который применяется тогда, когда ни одно состояние природы нельзя предпочесть другому: $Q_1 = Q_2 = \ldots = Q_n = \frac{1}{n}$.

2) Принцип убывающей арифметической прогрессии: $Q_1 : Q_2 : \ldots : Q_n = n : (n-1) : \ldots : 1$. Можно доказать, что при этом принципе $Q_j = \frac{2(n-j+1)}{n(n+1)}$, $j = \overline{1,n}$.

3) Принцип получения средних значений вероятностей $\bar{Q}_1, \bar{Q}_2, \ldots, \bar{Q}_n$ состояний природы $П_1, П_2, \ldots, П_n$, используя оценки группы экспертов.

8.2. Критерии Вальда, Сэвиджа и Гурвица

Существуют и другие подходы нахождения оптимального решения в условиях полной неопределенности, основанные на применении других критериев.

Максиминный критерий Вальда. Этот критерий называют еще и критерием крайнего пессимизма. В соответствии с этим критерием, в качестве оптимальной стратегии рекомендуется выбрать ту, которая гарантирует игроку I, в наихудших условиях, максимальный выигрыш, т.е,

$$\sigma = \max_{1 \le i \le m} \min_{1 \le j \le n} a_{ij}.$$

Название «максиминный» критерия Вальда исходит из этой расчетной формулы.

Критерий (минимаксного риска) Сэвиджа. Этот критерий так же является критерием крайнего пессимизма. В качестве оптимальной стратегии рекомендуется выбирать ту, при которой, в наихудших условиях, величина риска r принимает наилучшее значение:

$$r = \min_{1\le i\le m}\max_{1\le j\le n} r_{ij}.$$

Критерий Гурвица. Этот критерий называют критерием обобщенного максимума или пессимизма-оптимизма. Критерий Гурвица рекомендует выбрать в качестве оптимальной стратегии ту, которая находится как бы между стратегиями, являющимися «крайним пессимизмом» и «легкомысленным оптимизмом». Формула, с помощью которой определяется выигрыш σ игрока I, использующего критерий Гурвица имеет вид

$$\sigma = \max_{1\le i\le m}\{\lambda \min_{1\le j\le n} a_{ij} + (1-\lambda)\max_{1\le j\le n} a_{ij}\},$$

где λ $(0 \le \lambda \le 1)$ – некий коэффициент.

Очевидно, что при $\lambda = 1$ критерий Гурвица превращается в пессимистический критерий Вальда, а при $\lambda = 0$ – в критерий крайнего оптимизма.

Коэффициент λ выбирается на основании субъективных соображений (опыта, здравого смысла и т.д.) и зависит от настроя исследователя и его оценки ситуации – чем опаснее ситуация, чем больше хотелось бы в этой ситуации «подстраховаться», тем ближе к 1 должно быть значение λ.

Пример. Руководство универмага заказывает неделимый товар вида А. Известно, что спрос на данный вид товара лежит в пределах от 6 до 9 ед. Если заказанного товара окажется недостаточно для удовлетворения спроса, то руководство может срочно заказать и завести недостающее количество. Если же спрос будет меньше наличного количества товара, то нереализованный товар хранится на складе универмага. Условия транспортировки таковы, что этого товара нужно доставлять за один рейс по 2.

Требуется определить такой объем заказа на товар, при котором дополнительные затраты, связанные с хранением и срочным завозом были бы минимальными, если расходы на хранение 1 единицы товара составляют 1 денежная единица, а по срочному заказу и завозу – 2 денежные единицы.

В этом примере покупательский спрос выступает в качестве игрока II, т.е. природы, состояния которой определяются данными спроса: 6 ед. – $П_1$, 7 ед. – $П_2$, 8 ед. – $П_3$,9 ед. – $П_4$.

Игроком I является руководство универмага, стратегиями которого являются: A_1 – завоз 6 ед. товара, A_2 – завоз 8 ед. товара, A_3 – завоз 10 ед. товара.

Представленную ситуацию можно рассмотреть как игру с природой. Для составления ее платежной матрицы заметим, что если товара завозится 6 единиц и реализуется столько же, то затраты, связанные с хранением и срочным завозом равны 0. Если товара завозится 6 единиц, а спрос составляет 7 единиц, то необходимо срочно заказать и совершить 1 рейс (завести еще две единицы товара), на что необходимо затратить 4 денежных единиц. Это будет относится в пассив универмага, потому пишется «–4». В случае, спрос составляет 8 единиц, то все равно нужно совершить 1 рейс – затраты те же что и в случае, когда спрос составляет 7 единиц. Наконец, если спрос составляет 9 единиц, то затраты будут равны 6 денежным единицам.

Если товара завезено 8 единиц, а реализовано всего 6 единиц, то для хранения на складе требуется затратить за каждую единицу товара по 1 денежной единице, значит затраты равны 2 денежным единицам. Продолжая таким образом получим платежную 3×4-матрицу, которая содержится в таблице 1.

Таблица 1

	$П_1$ 6 ед.	$П_2$ 7 ед.	$П_3$ 8 ед.	$П_4$ 9 ед.
A_1 (6 ед.)	0	−4	−4	−6
A_2 (8 ед.)	−2	−1	0	−2
A_3 (10 ед.)	−4	−3	−2	−1

Найдем решение этой игры по критериям Вальда, Сэвиджа и Гурвица при λ=0,2.

1) Критерий Вальда. Сначала определим $\sigma_i = \min\limits_{1\le j\le 4} a_{ij}$, $i=\overline{1,3}$: $\sigma_1 = -6$, $\sigma_2 = -2$, $\sigma_3 = -4$. Теперь можно найти $\sigma = \max\limits_{1\le i\le 3} \sigma_i = \sigma_2 = -2$, т.е. оптимальной является стратегия A_2. Необходимо заказывать 8 ед. товара.

2) Критерий Сэвиджа. Составим матрицу риска $R=(r_{ij})$, $i=\overline{1,3}$, $j=\overline{1,4}$:

$$R = \begin{pmatrix} 0 & 3 & 4 & 5 \\ 2 & 0 & 0 & 1 \\ 4 & 2 & 2 & 0 \end{pmatrix}.$$

найдем $r_i = \max\limits_{1\le j\le 4} r_{ij}$, $i=\overline{1,3}$, и $r = \min\limits_{1\le i\le 3} r_i$: $r_1 = 5$, $r_2 = 2$, $r_3 = 4$, $r = \min\limits_{1\le i\le 3} r_i = r_2 = 2$. Здесь также оптимальной является стратегия A_2.

3) Критерий Гурвица при λ=0,2. Значения $\min\limits_{1\le j\le 4} a_{ij}$, $i=\overline{1,3}$, известны: $\min\limits_{1\le j\le 4} a_{1j} = -6$, $\min\limits_{1\le j\le 4} a_{2j} = -2$, $\min\limits_{1\le j\le 4} a_{3j} = -4$. Находим $w_i = \max\limits_{1\le j\le 4} a_{ij}$, $i=\overline{1,3}$: $w_1 = 0$, $w_2 = 0$,

$w_3 = -1$. Теперь найдем для $i = \overline{1,3}$ значения $\sigma_i = \lambda\alpha_i + (1-\lambda)w_i = 0{,}2\alpha_i + 0{,}8w_i$: $\sigma_1 = -1{,}2$, $\sigma_2 = -0{,}4$, $\sigma_3 = -1{,}6$. Определим $\sigma = \max\limits_{1 \le i \le 3} \sigma_i = \sigma_2 = -0{,}4$. Значит и в этом случае оптимальной будет стратегия A_2.

8.3. Принятие решений в условиях неопределенности

В качестве еще одного примера игры с природой рассмотрим задачу принятия решений в условиях неопределенности. Предположим, что необходимо выполнить некоторую операцию. Условия, при которых будут выполняться действия операции, недостаточно выяснены. Для их выполнения можно провести эксперимент, однако, это требует затрат средств. Следует определить: нужно ли проводить эксперимент или же лучше от него воздержаться?

С экономической точки зрения эксперимент целесообразно проводить в том случае, если затраты на его проведение не превышают выигрыша, который можно получить при более точном знании обстановки. Эту задачу можно рассмотреть как игру с природой.

Пусть известна матрица выигрышей $A = (a_{ij})$, $i = \overline{1,m}$, $j = \overline{1,n}$, и вероятности $Q_1, Q_2, \ldots, Q_n$ различных состояний природы $\Pi_1, \Pi_2, \ldots, \Pi_n$. Известны также затраты на проведение эксперимента, которые составляют C денежных единиц.

Рассмотрим случай идеального эксперимента, проведение которого позволяет точно определить состояние природы Π_j, при котором будет осуществляться операция.

Если эксперимент не проводится, то средний выигрыш игрока I будет равен $\overline{\sigma} = \max\limits_i \{\sum\limits_{j=1}^{n} a_{ij} Q_1\}$.

Полагаем теперь, что эксперимент проведен и выяснено действительное состояние природы, при котором будет осуществляться операция.

Если этим состоянием оказалось $П_j$, то выигрыш игрока I будет равен $\beta_j = \max\limits_{1 \le i \le m} a_{ij}$. Однако, на самом деле, истинное состояние природы неизвестно. Поэтому, интуитивно, средний выигрыш $\overline{\beta}$ игрока I определится формулой:

$$\overline{\beta} = \sum_{j=1}^{n} \beta_j Q_j .$$

Эксперимент нужно проводить, если

$$C < \sum_{j=1}^{n} \beta_j Q_j - \max_{1 \le i \le m} \left\{ \sum_{j=1}^{n} a_{ij} Q_j \right\}$$

или

$$C < \min_{1 \le i \le m} \overline{r}_i .$$

Таким образом, если затраты на осуществление эксперимента меньше минимального среднего риска, то его следует проводить. В качестве оптимальной стратегии в этом случае следует выбирать ту, для которой средний риск минимален.

Проблемные задания

1. Имеются ли общие характерные признаки и особенности между обычной игрой и игрой с природой?
2. Сформулируйте и смоделируйте игровую задачу принятия решений в условиях неопределенности.
3. Исследуйте возможность установления максимального количества рисков при решении конкретной игровой задачи принятия решений в условиях неопределенности.
4. Исследуйте пути выявления наиболее характерных особенностей процесса планирования эксперимента в условиях неопределенности.

5. Сформулируйте задачу, характеризующую суть процесса планирования эксперимента, проводимого в условиях неопределенности.

6. Составьте математическую модель задачи задания 5 и исследуйте пути ее реализации.

7. Дайте оценку полученному решению.

Вопросы для самоподготовки

1. Дайте определение игры с природой.
2. Что такое риск? Напишите формулу определения риска.
3. Напишите формулу максимального критерия Вальда.
4. Формула критерия Сэвиджа.
5. Формула критерия Гурвица.

ГЛАВА IV.
§ 9. ОСНОВЫ СЕТЕВОГО ПЛАНИРОВАНИЯ. СЕТЕВОЙ ГРАФИК КОМПЛЕКСА ОПЕРАЦИЙ И ПРАВИЛА ЕГО ПОСТРОЕНИЯ

Сетевое планирование, управление, сетевой график, событие, комплекс операций, операция-ожидание, контур, транзитивность

9.1. Сетевой график комплекса операций

Метод сетевого планирования и управления является одним из математических методов современной теории управления большими системами. Основой метода сетевого планирования и управления является сетевой график, отражающий логическую взаимосвязь и взаимообусловленность входящих в него операций.

В системах сетевого планирования и управления используются следующие, наиболее распространенные способы построения сетевых графиков:

1) Сетевые графики в терминах **«дуги–операции»**. В таких графиках вершины, называемые событиями, соответствуют моментам времени начала или окончания одной или нескольких операций, а дуги – операциям.

2) Сетевые графики в терминах **«дуги–связи»**, в которых операции изображаются вершинами сети, а дуги показывают порядок выполнения (взаимосвязь) отдельных операций.

В сетевом графике различают три вида событий: исходное, завершающее и промежуточное.

Исходное – это такое событие, с которого начинается выполнение комплекса операций. Завершающее событие соответствует достижению конечной цели, т.е.

завершению комплекса операций. Сетевые графики с несколькими завершающими событиями называются многоцелевыми. Все остальные события относятся к промежуточным.

На сетевых графиках события, обычно, обозначаются кружками, точками, а операции – линиями с указанием направления. Предполагается, что события не имеют продолжительности и наступают как бы мгновенно.

Моментом свершения события, считается момент окончания выполнения всех входящих в это событие операций. Пока не выполнены все входящие операции, не может свершится само событие, а следовательно, не может быть начата ни одна из непосредственно следующих за ним операций.

Различают три вида операций:

– **действительная операция** (→) – процесс, требующий затрат времени и ресурсов (разработка проекта, подвоз материалов, выполнение монтажных работ и т.д.);

– **операция-ожидание** (–·–·–>) – процесс, требующий только затраты времени (затвердение бетона, естественная сушка штукатурки перед началом малярных работ, рост растений и т.д.);

– **фиктивная операция** (– – – – >), или логическая зависимость, отражает технологическую или ресурсную зависимость в выполнении некоторых операций.

9.2. Правила построения сетевого графика

При построении сетевых графиков необходимо соблюдать определенные правила:

1) в сети не должно быть событий (кроме исходного), в которые не входит ни одна дуга;

2) не должно быть событий (кроме завершающего), из которых не выходит ни одной дуги;

3) сеть не должна содержать контуров (контур – замкнутый маршрут, у которого все вершины различны, кроме первой и последней);

4) любая пара событий сетевого графика может быть соединена не более чем одной дугой. Если изобразить одновременно (параллельно) выполняемые три различные операции *b, c, d* с общим начальным и конечным событиями (рис. 1), возникает путаница из-за того, что различные операции имеют одно и то же обозначение (2,5). В этом случае рекомендуется ввести дополнительные события и соединить их с последующими фиктивными операциями (рис. 2);

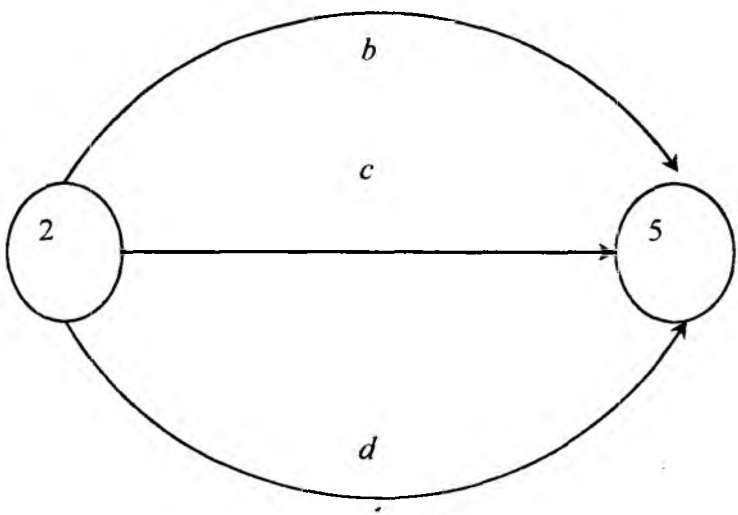


Рис. 1

5) если, какие-либо операции могут быть начаты до полного окончания непосредственно предшествующей им операции, то последнюю целесообразно представить

как ряд последовательно выполняемых операций, завершающихся определенными событиями. Например, если операции c и d могут быть начаты до полного окончания операции b, то операцию b рекомендуется разбить на элементарные операции b_1, b_1, b_3, и представить выполнение всех операций в виде графика, изображенного на рис. 3.

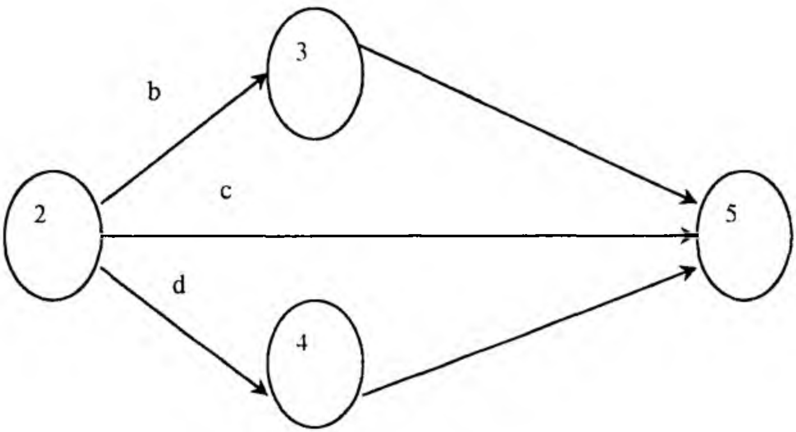


Рис.2

Для отражения технической или ресурсной зависимости, при выполнении операций применяют фиктивные операции. Предположим, что операция c может выполняться после завершения операций a и b, а операция d – только после завершения операции b. Эта зависимость представлена на рис. 4, из которого видно, что операция c следует за операцией a и фиктивной операцией (2,3). В свою очередь операция (2,3) следует за операцией b. Тогда в силу транзитивности (транзитивность – свойство бинарного отношения R, состоящее в том, что из aRb и

bRc следует aRc; примеры транзитивных бинарных отношений: =, ≥, >, ≤, <) выполнение операции b предшествует выполнению операции c.

Построение сетевого графика начинается с составления списка операций (работ), подлежащих выполнению (эта последовательность в списке операций может быть произвольной). Порядок нумерации операций осуществляется в соответствии с последовательностью их записи в списке. Перечень операций тщательно продумывается и детализируется в зависимости от конкретных условий. Включенные в список операции шифруются путем указания пары событий, которые считаются начальным и конечным событиями данной операции. В данный список обычно не включаются фиктивные операции. Операции, включенные в список, характеризуются определенной продолжительностью, которая устанавливается на основе действующих нормативов или по аналогии с ранее выполнявшимися операциями. Такие временные оценки называются детерминированными. Если же нормативные данные временных оценок операций отсутствуют, то определяются вероятностные оценки.

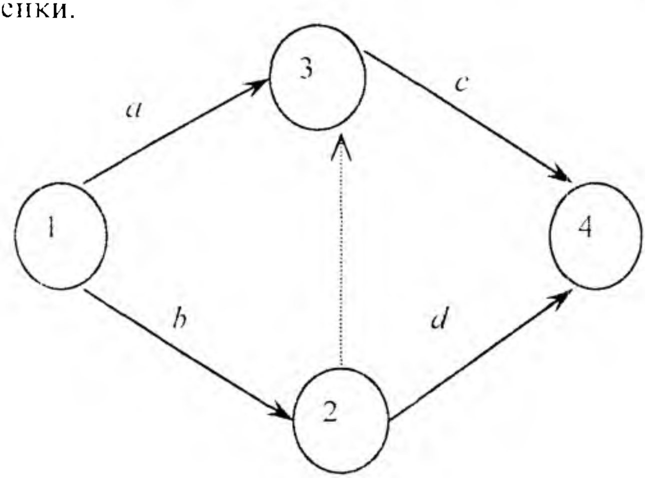


Рис. 4

После составления списка операций, приступают к процедуре построения сети. Фиктивные операции включаются в сетевой график в виде дуги графика с продолжительностью, равной нулю.

Пример. Необходимо построить укрупненный сетевой график выполнения комплекса операций по реконструкции цеха. Список операций представлен в таблице.1.

Таблица 1

Операция	Шифр операции	Наименование операции	Опирается на операции	Продолжительность (дни)
a_1	(1,2)	Подготовительные работы	–	5
a_2	(1,3)	Демонтаж старого оборудования	–	3
a_3	(2,6)	Ремонтные строительно-монтажные работы	a_1	30
a_4	(3,4)	Подготовка фундамента под новое оборудование	a_1, a_2	16
a_5	(2,4)	Подготовка к монтажу нового оборудования	a_1	10
a_6	(2,5)	Электротехнические работы	a_1	12
a_7	(4,5)	Монтаж нового оборудования	a_4, a_5	8
a_8	(5,7)	Подключение оборудования к электросети	a_6, a_7	2
a_9	(7,8)	Наладка и технологические испытания оборудования	a_8	6
a_{10}	(6,8)	Отделочные работы	a_3, a_6 a_7	8
a_{11}	(8,9)	Приемка цеха в эксплуатацию	a_3, a_{10}	1

Сетевой график комплекса операций изображен на рис.5.

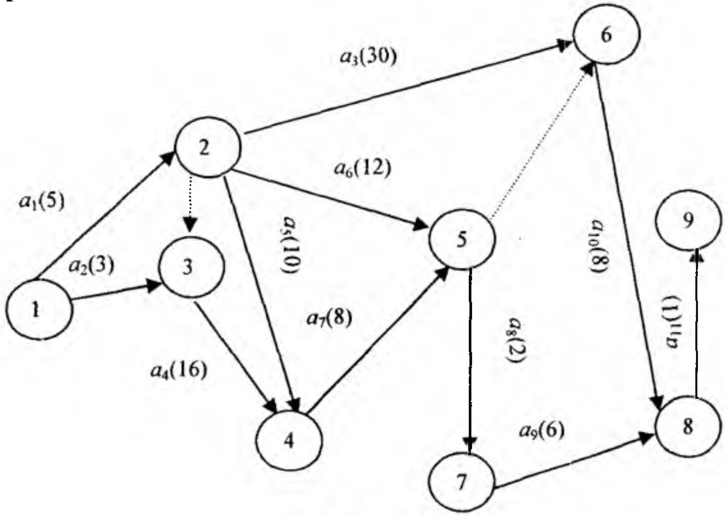


Рис. 5

Все операции графика, за исключением фиктивных операций (2,3) и (5,6) являются действительными. Числа в скобках, приписанные дугам, означают продолжительность выполнения соответствующих операций. Операции a_1 и a_2 не опираются ни на какие операции, следовательно, на графике они изображаются дугами, выходящими из исходного события (1), означающего момент начала выполнения комплекса операций.

Операции a_3, a_5 и a_6 опираются на операцию a_1, поэтому на графике дуги a_3, a_5 и a_6 непосредственно следуют за дугой a_1. Событие (2) означает момент окончания операции a_1 и начала операций, представленных дугами, выходящими из этого события. Операция a_4 опирается на операции a_1 и a_2. Графически это условие отражено посредством последовательного изображения операций (1,3) и (3,4) и введения фиктивной операции (2,3). Событие (3) инцидентно операциям (1,3) и (2,3), следовательно, моментом свершения события (3) будет такой момент, к которому будут выполнены все входящие в это событие операции и может быть начата операция, отраженная дугой, выходящей из него. Аналогично, с учетом технологии выполнения, изображены на графике остальные операции. Завершающее событие (9) означает момент окончания выполнения всего комплекса операций по реконструкции цеха. Шифры операций состоят из номеров начального и конечного событий и практически в список заносятся после составления графика.

Проблемные задания

1. Составьте пример, сетевой график которого будет иметь несколько завершающих событий.
2. Исследуйте и определите, всегда ли противопоказано содержание контуров в сети.
3. Исследуйте и определите практическую реальность использования фиктивных операций.

Вопросы для самоподготовки

1. Дайте определение исходного и завершающего событий.
2. Как называются сетевые графики с несколькими завершающими событиями?
3. Можно ли свести сетевые графики с несколькими завершающими событиями к сетевому графику с одним завершающим событием?
4. Что такое момент свершения события?

5. Чем отличаются друг от друга действительная операция, операция-ожидание и фиктивная операция?

6. Перечислите правила, которые необходимо соблюдать при построении сетевых графиков.

7. Какие временные оценки называются детерминированными, а какие вероятностными?

§ 10. РАСЧЕТ ВРЕМЕННЫХ ПАРАМЕТРОВ СЕТЕВОГО ГРАФИКА

Оперирующая сторона, сетевая модель, событие, полный путь, критический путь, предельный срок, ранний срок, полный резерв времени, свободный и частный резервы времени

10.1. Понятия о временных параметрах сетевого графика

Для управления ходом выполнения комплекса операций, представленного сетевой моделью, оперирующая сторона должна располагать количественными параметрами элементов сети. К таким параметрам относятся: продолжительность выполнения всего комплекса операций, сроки выполнения отдельных операций и их резервы времени. Важнейшим параметром сетевого графика является также критический путь.

Различают следующие виды путей: полный, предшествующий событию, следующий за событием.

Путь сетевого графика называется полным, если его начальная вершина совпадает с исходным событием, а конечная – с завершающим.

Предшествующий событию путь – это путь от исходного события до данного.

Следующий за событием путь – это путь от данного события до завершающего.

Критическим называется путь (полный), имеющий наибольшую продолжительность во времени.

Операция и события, принадлежащие критическому пути, называются соответственно критическими операциями и критическими событиями.

Суммарная продолжительность операций, принадлежащих критическому пути, равна критическому времени $t_{кр}$ выполнения комплекса операций в целом.

10.2. Расчет временных параметров сетевого графика

Расчет временных параметров сетевого графика может осуществляться различными методами. Рассмотрим один из них на примере.

Пример. Предположим, что продолжительности выполнения операций t_{ij} известны и приписаны у соответствующих дуг сетевого графика, представленного на рис. 1.

Определим, прежде всего, ожидаемые (ранние) сроки свершения событий t_i сетевого графика. Исходное событие означает момент начала выполнения комплекса операций, следовательно, $t_i = 0$. Событие (2) свершится, очевидно, спустя 2 ед. времени после свершения события (1), т.к. время выполнения операции (1,2) равно 2.

Следовательно, $t_2 = t_1 + t_{12} = 0 + 2 = 2$. Событию (3) предшествуют два пути: $\mu_1 = (1-3)$ и $\mu_2 = (1-2-3)$. Продолжительность первого пути равна 1 ед. времени, а второго 2 ед. времени, т.к. $t_{13} = 1$, $t_{12} + t_{23} = 2 + 0 = 2$.

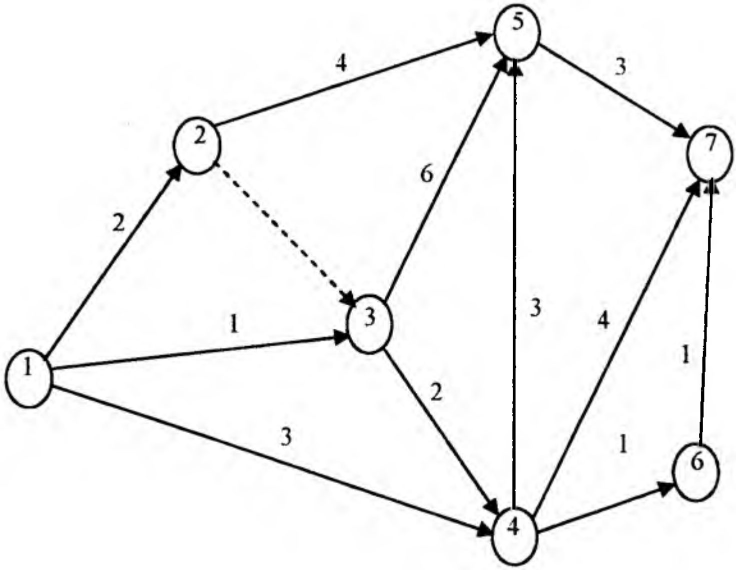


Рис. 1

Продолжительность второго пути можно найти добавлением к ожидаемому сроку свершения события (2) времени выполнения операции (2,3), т.е. $t_2 + t_{23} = 2 + 0 = 2$.

Поскольку событие (3) может совершиться не раньше момента окончания всех входящих в него операций, то $t_3 = \max(t_1 + t_{13};\ t_2 + t_{23}) = \max(0 + 1;\ 2 + 0) = 2$.

В событие (4) входят две пути, исходящие из событий (1) и (3), для которых ожидаемые сроки свершения найдены. Следовательно, ожидаемый срок свершения события (4) $t_4 = \max(t_1 + t_{14};\ t_3 + t_{34}) = \max(0 + 3;\ 2 + 2) = 4$.

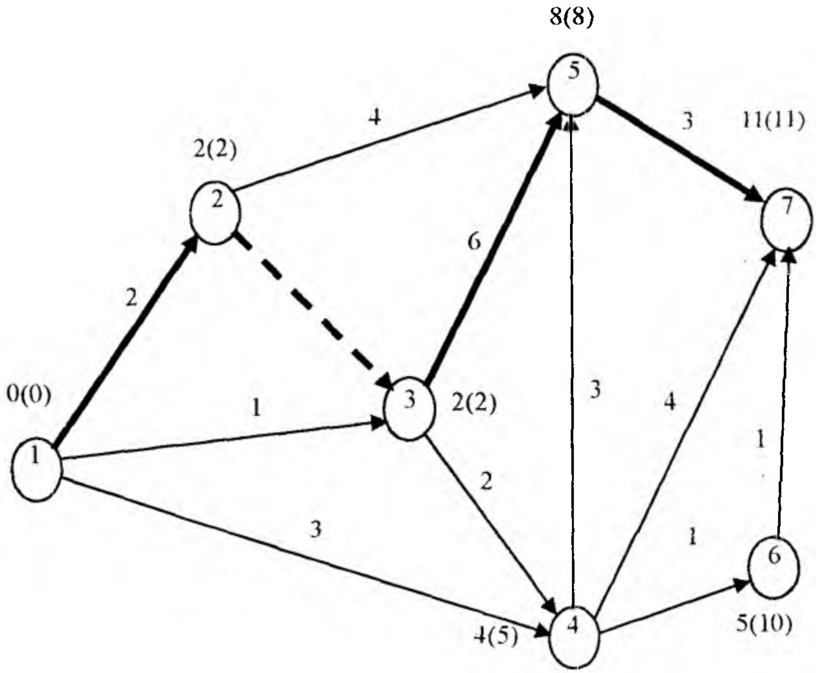


Рис. 2

Аналогично находятся ожидаемые сроки свершения событий (5), (6) и (7). Значения t_i, $i = \overline{1,7}$, приписаны соответствующим событиям на рис. 2.

Общую формулу для нахождения ожидаемых сроков совершения событий можно записать так:

$$t_1 = 0, \quad t_j = \max_{\{(\overrightarrow{i,j})\}} (t_i + t_{ij}), \quad j = 2,3,\dots, n,$$

где $\{(\overrightarrow{i,j})\}$ - подмножество дуг сети, входящих в событие (j).

Ожидаемый срок свершения события (7) $t_7 = 11$ совпадет с критическим временем (суммарной продолжительностью операций, принадлежащих критическому пути). Возвращаясь теперь от завершающего события к исходному, выделим операции, принадлежащие критическому пути. Из трех операций, входящих в событие (7), $t_{кр}=11$ определила операция (5,7), выполнение которой начинается после свершения события (5) и продолжается 3 ед. времени: $t_5+t_{57}=8+3=11$.

Момент свершения события (5) определила операция (3,5), т.к. $t_3+t_{35}=2+6=8$. В свою очередь, момент свершения события (3) определила операция (2,3), а события (2) – операция (1,2). Эти операции на графике (рис. 2) выделены жирной линией.

Таким образом, критическим является путь $\mu_{кр}=(1-2-3-5-7)$. Увеличение времени выполнения любой операции, принадлежащий критическому пути, ведет к увеличению времени выполнения комплекса операций в целом. Увеличение же времени выполнения или задержка с выполнением некритических операций, может не отразится на сроке свершения завершающего события. Например, время выполнения операции (4,5) может быть увеличено или начало ее выполнения может быть отсрочено на 1 ед. времени, и это не отразится на сроке свершения события (5), а следовательно, и всего комплекса операций.

Начало выполнения операции (4,7) может быть отсрочено на 3 ед. времени.

Отсюда следует, что для рассмотренных событий, не лежащих на критическом пути, существуют предельные (поздние) сроки свершения событий.

Обозначим через $t_i^*, i=\overline{1,n}$ предельный срок свершения события (i) сетевого графика. Отметим, что ожидаемый и предельный сроки свершения завершающего события (n) совпадают: $t_n=t_n^*$.

Предельный срок свершения любого события сетевого графика равен минимальной разности между предельными сроками окончания операций, исходящих из данного события, и временем выполнения соответствующих операций. Нахождение предельного срока осуществляется по формуле:

$$t^* = t_n, \quad t_i^* = \min_{\overset{<}{\{(i,j)\}}}(t_j^* - t_{ij}), \quad i = \overline{1, n-1},$$

где $\overset{<}{\{(i,j)\}}$ - подмножества дуг сети, которые исходят из события (i).

В данном примере $t_7^* = t_7 = 11$. Определим этот показатель для оставшихся событий. Из события (5) исходит одна операция, следовательно, $t_5^* = t_7^* - t_{17} = 11 - 3 = 8$. Аналогично $t_6^* = t_7^* - t_{67} = 11 - 1 = 10$. Из события (4) исходят три операции, поэтому

$$t_4^* = \min(t_5^* - t_{45};\, t_6^* - t_{46};\, t_7^* - t_{47}) = \min(8-3; 10-1; 11-4) = 5\,.$$

Аналогично находим, что $t_3^* = 2$, $t_2^* = 2$ и $t_1^* = 0$. На рис. 2 предельные сроки свершения событий указаны в скобках. Для критических событий эти сроки совпадают с ожидаемыми.

Некритические события имеют резервы времени, которые показывают, на какой предельно допустимый срок может задержаться свершение событий без изменения срока свершения завершающего события. Резерв времени R_i события (i) равен разности между предельным и ожидаемым сроками его свершения: $R_i = t_i^* - t_i$.

Ожидаемые и предельные сроки свершения событий находятся в диалектическом единстве со сроком начала и окончания операций:

- ранний срок начала выполнения операции (i, j) равен ожидаемому сроку свершения события (i): $t_{ij}^{рн} = t_i$;

- поздний срок окончания операции совпадает с поздним сроком свершения ее конечного события: $t_{ij}^{п.о.} = t_j^*$;

- поздний срок начала выполнения операции равен разности между предельным сроком свершения ее конечного события и продолжительностью: $t_{ij}^{п.н.} = t_j^* - t_{ij}$;

- ранний срок окончания операции равен сумме ожидаемого срока свершения ее начального события и продолжительности: $t_{ij}^{р.о} = t_i + t_{ij}$.

Сроки выполнения операций находятся в границах, определяемых параметрами: $t_{ij}^{р.н.}$, $t_{ij}^{п.н.}$, $t_{ij}^{р.о.}$, $t_{ij}^{п.о.}$. Следовательно, операции, как и события, могут иметь некоторый резерв времени.

Различают четыре разновидности резервов времени операций: **полный, свободный, частный первого вида и частный второго вида.**

1. Полный резерв времени операции $R_{ij}^{п}$ показывает, насколько можно сдвинуть начало выполнения операции или увеличить ее продолжительность, не изменяя ожидаемого срока свершения начального события, при условии, что конечное, для данной операции, событие свершится не позднее своего предельного срока. Величина полного резерва времени вычисляется по формуле:

$$R_{ij}^{п} = t_j^* - (t_i + t_{ij}) = t_j^* - t_{ij}^{р.о.}.$$

2. Свободный резерв времени операции $R_{ij}^{с}$ показывает, насколько можно увеличить продолжительность или отсрочить начало выполнения операции (i,j) при условии, что начальное и конечное ее события свершаются в ожидаемое время:

$$R_{ij}^{c} = t_j - (t_i + t_{ij}) = t_j - t_{ij}^{p.o.}.$$

3. Частный резерв времени первого вида R'_{ij} – это запас времени, которым можно располагать при выполнении операции (i,j) в предположении, что начальное и конечное ее события свершаются в предельные сроки:

$$R'_{ij} = t_j^* - (t_i^* + t_{ij}) = t_{ij}^{n.н.} - t_i^*.$$

4. Частный резерв времени второго вида R''_{ij} – это запас времени, которым можно располагать при выполнении операции (i,j) в предположении, что ее начальное событие свершится в предельное, а конечное - в ожидаемое время. Для некоторых операций интервал времени между предельным сроком свершения начального срока события может быть меньше их продолжительности. В этом случае R''_{ij} принимается равным нулю. Определяется частный резерв времени второго вида по формуле:

$$R''_{ij} = \max(t_j - t_i^* - t_{ij};0).$$

Найдем резервы времени операции (4,6) сетевого графика (рис. 2), для чего заметим, что $t_6^* = 10$, $t_4 = 4$, $t_{46} = 1$. Имеем:

$$R_{46}^{n} = t_6^* - (t_4 + t_{46}) = 10 - (4 + 1) = 5,$$

$$R_{46}^{c} = t_6 - (t_4 + t_{46}) = 5 - (4 + 1) = 0,$$

$$R'_{46} = t_6^* - (t_4^* + t_{46}) = 10 - (5 + 1) = 4,$$

$$R''_{46} = \max(t_6 - t_4^* - t_{46};0) = \max(5 - 5 - 1;0) = 0.$$

Проблемные задания

1. Сформулируйте задачу управления ходом выполнения комплекса операций, какого либо технического процесса.

2. Составьте сетевую модель задачи задания 1.

3. Исследуйте и определите подходы, и возможности установления резервов времени в этой задаче.

4. Определите возможные и предельные сроки свершения событий задачи задания 1.

5. Для одного из некритических и одного критического событий задачи задания 1 определите:

а) ранний срок начала выполнения операций;

б) поздний срок окончания операций;

в) поздний срок начала выполнения операций;

г) ранний срок окончания операции;

д) полный резерв времени операций;

е) свободный резерв времени;

ж) частный резерв времени 1-го вида;

з) частный резерв времени 2-го вида.

Вопросы для самоподготовки

1. Что такое критический путь?

2. Что называется критической операцией и критическим событием?

3. Дайте определения позднего и раннего сроков окончания операций.

4. Дайте определения полного и свободного резервов времени операций.

5. Дайте определения частных резервов времени I и II видов.

§ 11. ВЕРОЯТНОСТНЫЕ СЕТИ

Сетевой график, детерминированная и стохастическая структуры, вероятностные сети, функция распределения, дисперсия, математическое ожидание, критическое время, критический путь

11.1. Структуры вероятностных сетей

Сетевые графики комплекса операций могут иметь детерминированную или стохастическую структуру. Если все операции комплекса и их взаимосвязи точно определены, то такая структура графика называется детерминированной. Стохастическая структура означает, что все операции включаются в сеть с некоторой вероятностью. **Например,** в научно-исследовательских и опытно-конструкторских разработках заранее неизвестны не только продолжительность отдельных операций, но и их перечень, а так же структура сети.

На практике обычно применяются детерминированные сети со случайными временными оценками операций. Такие сети называются вероятностными.

При исследовании вероятностных сетей могут встретиться два случая:

1) операции не являются новыми, и мы приближенно знаем для каждой из них функцию распределения продолжительности выполнения;

2) операции являются новыми, малоизученными, и для них функции распределения продолжительностей неизвестны.

В первом случае по неизвестной функции распределения нетрудно определить среднее значение продолжительности выполнения каждой операции (математическое ожидание) и дисперсию.

Во втором случае применяется метод усреднения. Исходными данными для метода усреднения являются вероятностные оценки продолжительности каждой операции:

a — минимальная продолжительность (оптимистическая оценка) операции,

b — максимальная продолжительность (пессимистическая оценка) операции,

m — наиболее вероятная продолжительность (мода) операции.

Исследования позволили обосновать возможность использования β -распределения в качестве типового распределения продолжительности операций с оценками a, b и m.

Функция плотности β-распределения (рис. 1) имеет вид

$$f(t)=\begin{cases} c(t-a)^p(b-t)^q & \text{для} \quad a \le t \le b; \\ 0 & \text{для} \quad -\infty < t < a, b < t < \infty, \end{cases}$$

где p, q — параметры распределения, зависящие от вида операций; c - нормирующий множитель, определяемый из условия

$$c\int_a^b (t-a)^p(b-t)^q dt = 1 \cdot$$

По известной функции распределения $f(t)$ находятся числовые характеристики операций:

— среднее значение (математическое ожидание) продолжительности операции t

$$M[t]=\bar{t}=\int_a^b tf(t)dt=\frac{(p+q)m+(a+b)}{p+q+2};$$

– дисперсия $D[t] = \sigma_t^2 = \int_a^b t^2 f(t)dt - \bar{t}^2$.

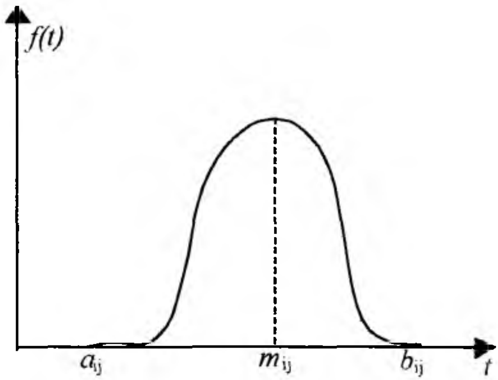


Рис.1

Статистический анализ, проведенный эмпирико-экспериментальным путем, позволили установить, что $p+q \approx 4$. Следовательно,

$$\bar{t} = \frac{a + 4m + b}{6}; \quad D[t] \approx \left(\frac{b-a}{6}\right)^2.$$

После определения математических ожиданий продолжительностей операций проводится расчет временных параметров сети, как и в детерминированном случае. Длительность критического пути рассматривают как математическое ожидание случайной величины $t_{кр}$:

$$M[t_{kp}] = \bar{t}_{kp} = \sum_{(i,j)_{kp}} \bar{t}_{i,j}.$$

Дисперсию продолжительности пути считают равной сумме дисперсий продолжительностей операций, находящихся на критическом пути:

$$D[t_{kp}] = \sum_{(i,j)\in \mu_{kp}} D_{ij}[t].$$

Практически, расчет временных параметров сети по средним значениям продолжительностей операций не позволяет строго определить срок завершения комплекса операций. Фактически отклонение случайных величин t_{ij} от их средних значений $\bar{t}_{ij}$ может быть как в большую, так и в меньшую сторону. Поэтому, фактическая продолжительность выполнения комплекса операций $t_ф$ может быть больше или меньше $\bar{t}_{кр}$. В связи с этим представляет большой интерес оценка вероятности завершения комплекса операций к определенному сроку, которая зависит от дисперсии $D[t_{кр}]$ продолжительности критического пути. При одних значениях величин t_{ij} может быть один критический путь, при других – другой. Однако, если продолжительности работ отклоняются от своих средних значений на такую малую величину, что критический путь не изменяется, и если на критическом пути лежит значительное число операций (5, 6 или более), то на основании центральной предельной теоремы можно приближенно считать, что его продолжительность подчиняется нормальному закону распределения с параметрами $\bar{t}_{кр}$, $D[t_{кр}]$. Тогда вычисление вероятности того, что фактическая продолжительность выполнения комплекса операций $t_ф$ меньше планового директивного срока $T_{пл}$, производится по формуле:

$$p(t_ф - T_{пл}) = Ф(u) + 0{,}5,$$

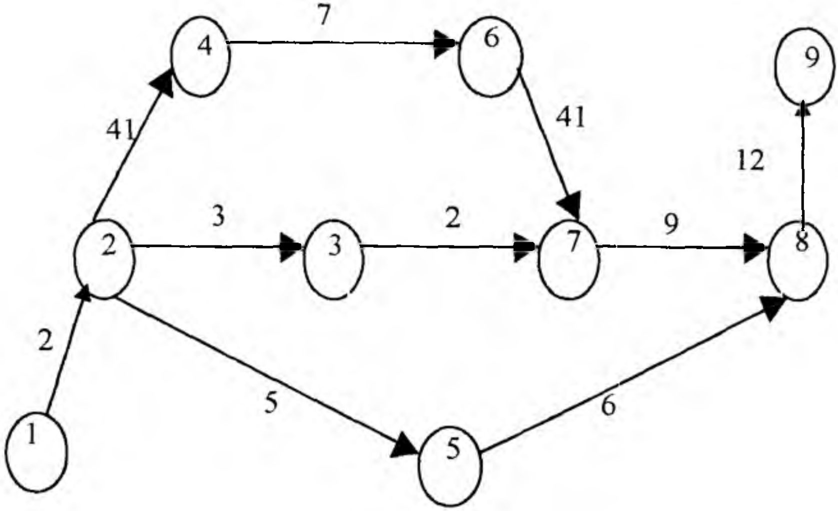


Рис 2

где $\Phi(u)=\frac{1}{\sqrt{2\pi}}\int\limits_{o}^{u}e^{-\frac{t^2}{2}}dt$ функция Лапласа, $u=\frac{T_{пл}-\bar{t}_{kp}}{\sigma_{kp}}$, $\sigma_{kp}=\sqrt{D[t_{kp}]}$ – среднеквадратичное отклонение.

По формуле $p(t_{ф}-T_{пл})=\Phi(u)+0.5$ можно вычислить вероятность выполнения любой операции в заданный срок.

Рассмотрим подход к определению математического ожидания $\bar{t}_{ij}$ и дисперсии, $D_{ij}[t]$ операции (i,j) сетевого проекта на основе двух оценок: оптимистической a и пессимистической b.

На основе многочисленных, эмперико-экспериментальных исследований двухоценочной методики выявлено, что в β-распределении величины p и q, определенные для большого количества сетевых моделей, близки к постоянным значениям $p=1$, $q=2$. Выбрав их в качестве стандартных показателей степени, получают функцию, которая относится к классу β - распределений и имеет следующие параметры:

математическое ожидание: $\bar{t}_{ij} = \int_a^b tf(t)dt = \dfrac{3a+2b}{5}$;

дисперсия: $D_{ij}[t] = \int_a^b t^2 f(t)dt - \bar{t}_{ij}^2 = \left(\dfrac{b-a}{5}\right)^2$.

Применение двух временных оценок существенно уменьшает объем информации, который требуется от ответственного исполнителя, так как он освобождается от задания наиболее вероятной оценки.

11.2. Пример

Найти критическое время $t_{кр}$ выполнения комплекса операций, представленного на рис. 2, используя средние оценки продолжительности и дисперсию, а также определить:

1) вероятность выполнения комплекса операций за:

а) $T_{пл}$=35 дней; б) $T_{пл}$=42 дня;

2) время, за которое комплекс операций будет выполнен с вероятностью, не меньшей:

а) p=0,75; б) p=0,35;

3) вероятность завершения операции (2,5) в восьмой день.

Оптимистическая оценка b_{ij} для каждой операции задана в таблице 1. Случайные отклонения времени выполнения операций от математических ожиданий не меняют критического пути.

Вычисляя математическое ожидание $\bar{t}_{ij}$ и дисперсию $D_{ij}[t]$ и дополняя ими таблицу 1, получим таблицу 2.

Продолжительность критического пути, найденного по средним оценкам времени (средние оценки принисаны дугам графа), t_{kp}=38 дней, по оптимистическим оценкам $t_{\text{кр.опт}} = 27.5$ дня, по пессимистическим $t_{\text{кр.пес}} = 53.75$ дня. Практически, комплекс операций может быть выполнен с некоторой вероятностью в любой срок интервала [27,5; 53,75].

Таблица 1

Исходные параметры		
(i,j)	a_{ij}	b_{ij}
(1,2)	1	3,5
(2,3)	2	4,5
(2,4)	2,5	6,25
(2,5)	4	6,5
(3,7)	1,5	2,75
(4,6)	5	10
(5,8)	4,5	8,25
(6,7)	3	5,5
(7,8)	8	10,5
(8,9)	8	18

Так как критический путь включает 6 операций и, согласно центральной предельной теоремы теории вероятностей, его длина подчиняется нормальному закону распределения, то он характеризуется параметрами:

$$\bar{t}_{kp} = 38; \qquad D[t_{kp}] = \sum_{(i,j)kp} D_{ij} = 6.31; \qquad \delta_{kp} = \sqrt{D[t_{kp}]} = 2.51.$$

Вычислим вероятность выполнения комплекса операций за $T_{n,l}$=*35* дней:

$$u = \frac{35-38}{2.51} = -\frac{3}{2.51} = -1.19,$$

$$P(t_\phi < T_{\text{пт}}) = \Phi\left(\frac{35-28}{2.51}\right) + 0.5 =$$

$$= \Phi(1.55) + 0.5 = 0.439 + 0.5 = 0.939.$$

Таблица 2

Исходные параметры			Расчетные параметры	
(i,j)	a_{ij}	b_{ij}	t_{ij}	D_{ij}
(1,2)	1	3,5	2	0,25
(2,3)	2	4,5	3	0,25
(2,4)	2,5	6,25	4	0,56
(2,5)	4	6,5	5	0,25
(3,7)	1,5	2,75	2	0,063
(4,6)	5	10	7	1
(5,8)	4,5	8,25	6	0,56
(6,7)	3	5,5	4	0,25
(7,8)	8	10,5	9	0,25
(8,9)	8	18	12	4

Определим время, за которое комплекс операций будет выполнен с вероятностью не меньшей чем p=0,75. Значению $\Phi(u) = p - 0{,}5 = 0{,}75 - 0{,}5 = 0{,}25$ функции Лапласа соответствует значение аргумента u=0,68. Следовательно, искомое время будет равно $T_{пл} = \bar{t}_{kp} + \sigma_{kp}u = 38 + 2{,}51 \cdot 0{,}68 \approx 40$ дней.

Для p=0,35 функция Лапласа $\Phi(u) = 0{,}35 - 0{,}5 = -0{,}15$ и $u = -0{,}39$.Таким образом, в этом случае $T_{пл} = 38 + 0.51(-0.39) \approx 37$ дней.

Ожидаемый срок свершения пятого события t_5=7. Сумма дисперсий операций, принадлежащих пути (1-2-5), ведущему к пятому событию, будет равна $D_{12}(t) + D_{25}(t) = 0.25 + 0.25 = 0.5$.

Таким образом,

$$P_{(2,5)}(t_{\phi} < T_{пл}) = \Phi\left(\frac{8-7}{\sqrt{0{,}5}}\right) + 0{,}5 = \Phi(1{,}41) + 0{,}5 =$$

$$= 0.4207 + 0.5 = 0.9207 \approx 0.921 .$$

Следовательно, операция (2,5) с вероятностью 0,921 будет завершена в плановый срок.

Проблемные задания

1. Сформулируйте задачу выполнения комплекса операций для какого-либо технологического процесса (например, выпуск продукции).

2. Представьте детерминированную сетевую модель задачи пункта 1, которая будет характеризоваться случайными временными оценками операций.

3. В задаче сетевой модели пунктов 1 и 2 найдите критическое время $t_{кр}$ выполнения комплекса операций и определите:

а) вероятность выполнения комплекса операций за $T_{пл}$(плановое время установите сами);

б) время, за которое комплекс операций будет выполнено с вероятностью не меньшей: p_1 =0,75; p_2 =0,4.

4. Оптимистическую оценку a_{ij} и пессимистическую оценку b_{ij} для каждой операции задайте сами.

5.

Вопросы для самоподготовки

1. Дайте определение вероятностных сетей.

2. Определите числовые характеристики операций по известным функциям распределения.

3. Напишите формулу вычисления вероятности того, что фактическая продолжительность выполнения комплекса операций $t_ф$ меньше планового директивного срока $T_{пл}$.

§ 12. ОПТИМИЗАЦИЯ КОМПЛЕКСА ОПЕРАЦИЙ ПО ВРЕМЕНИ И ПО СТОИМОСТИ

Операция, сетевой график, фиктивная операция, критический путь, затраты, фиксированный срок

12.1. Оптимизация комплекса операций по времени

Оптимизация комплекса операций по времени обусловлена сокращением продолжительности критического пути. Такая задача возникает тогда, когда критическое время выполнения комплекса операции превосходит срок T_0, заданный оперирующей стороной.

Задачи оптимизации комплекса операций по времени решаются с привлечением дополнительных средств и с использованием внутренних резервов.

Рассмотрим одну из постановок задачи оптимизации с использованием дополнительных средств. Пусть задан сетевой график $G(E,e)$ выполнения комплекса операций. Время выполнения каждой операции равно t_{ij}. Вложение x_{ij} дополнительных средств в операцию (i,j) сокращает время выполнения с t_{ij} до $t'_{ij} = f_{ij}(x_{ij}) < t_{ij}$.

Отметим, однако, что насыщение критических операций ресурсами не может быть бесконечным, так как для каждой операции задается минимально возможное время ее выполнения, равное d_{ij}. Необходимо определить время начала $T_{ij}^{\text{н}}$ и окончания T_{ij}^{o} выполнения операций, а также, сколько дополнительных средств x_{ij} необходимо вложить в каждую из операций (i,j), чтобы:

– общее время выполнения комплекса операций было минимальным;

– сумма вложенных дополнительных средств не превышала заданной величины B;

Математическая модель этой задачи может быть представлена следующим образом:

$$t_{кр} = T^{о}_{n-1,n} \to \min ,$$

$$\sum_{(i,j)\in e} x_{ij} \le B ,$$

$$T^{о}_{ij} - T^{н}_{ij} \ge d_{ij} \text{ для всех } (i,j)\in e ,$$

$$t_{ij}(x_{ij}) = T^{о}_{ij} - T^{н}_{ij} \text{ для всех } (i,j)\in e ,$$

$$T^{н}_{jr} \ge T^{о}_{ij} \text{ для всех } i,j,r \in E ,$$

$$T^{н}_{ij} \ge 0 ,\ T^{о}_{ij} \ge 0 ,\ x_{ij} \ge 0 , \text{ для всех } (i,j)\in e .$$

Добавив, при необходимости, фиктивную операцию, выходящую из последнего события, целевую функцию любого графика можно записать в виде целевой функции задачи.

Ограничение-равенство задачи показывает зависимость продолжительности выполнения операций от вложенных средств. Последующие неравенства при ограничениях задачи обеспечивают выполнение условий предшествования операций в соответствии с топологией сети (время начала выполнения каждой операции должно быть не меньше времени окончания непосредственно предшествующей ей операции).

Критический путь $\mu_{кр}$ в данной задаче является функцией от объемов дополнительно вкладываемых средств x_{ij}.

Сформулированная задача относится к классу задач математического программирования и может быть решена методами линейного или нелинейного программирования в зависимости от вида функций $f_{ij}(x_{ij})$.

Пример: Комплекс операций представлен сетевым графиком на рис. 1. Цифры, приписанные дугам графика, означают соответственно продолжительность t_{ij} и минимально возможное время d_{ij} выполнения операций.

Продолжительность выполнения операций линейно зависит от дополнительно вложенных средств и выражается соотношением $t'_{ij}=t_{ij}(1-k_{ij}x_{ij})$, где $k_{12}=0,2$; $k_{13}=0,1$; $k_{14}=0,3$; $k_{23}=0,2$; $k_{24}=0,5$; $k_{45}=0,3$. Требуется построить развернутую математическую модель для определения времени начала и окончания выполнения операций и количества средств, вкладываемых в каждую операцию, чтобы время выполнения комплекса операций было минимальным, а сумма вложенных средств не превышала 10 ед.

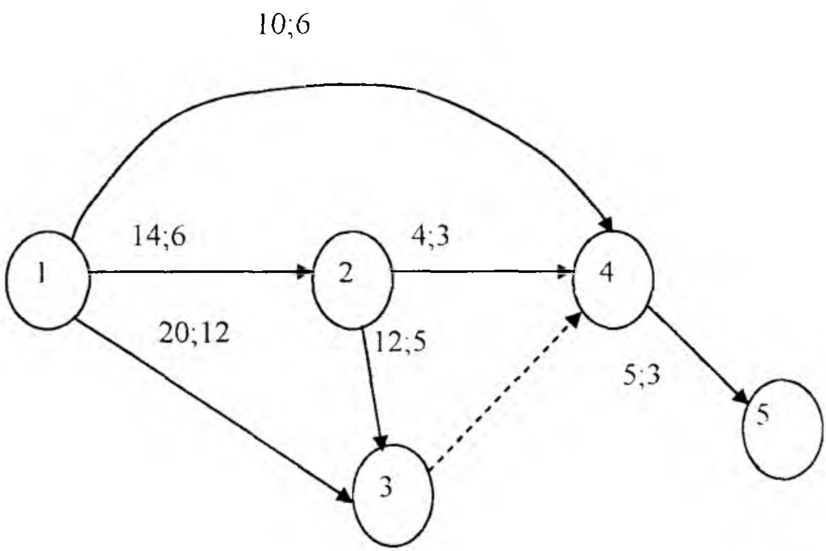


Рис. 1

В силу особенностей сетевого графика целевая функция имеет вид $t_{кр} = T^{o}_{15} \to \min$.

Запишем ограничения задачи.

Сумма вложенных средств не должна быть больше наличного их количества:

$$x_{12}+x_{13}+x_{14}+x_{23}+x_{24}+x_{45}\le 10,$$

время выполнения каждой операции должно быть не меньше минимального времени:

$$T_{12}^{0}-T_{12}^{н}\ge 6\,;\; T_{13}^{0}-T_{13}^{н}\ge 12\,;\; T_{14}^{0}-T_{14}^{н}\ge 6\,;$$
$$T_{23}^{0}-T_{23}^{н}\ge 5\,;\; T_{24}^{0}-T_{24}^{н}\ge 3\,;\; T_{34}^{0}-T_{34}^{н}\ge 0\,;$$
$$T_{45}^{0}-T_{45}^{н}\ge 3\,.$$

Зависимость продолжительностей операций от вложенных средств дают ограничения-равенства:

$$T_{12}^{0}-T_{12}^{н}=14\cdot(1-0{,}2x_{12})\,;\; T_{13}^{0}-T_{13}^{н}=20\cdot(1-0{,}1x_{13})\,;$$
$$T_{14}^{0}-T_{14}^{н}=10\cdot(1-0{,}3x_{14})\,;\; T_{23}^{0}-T_{23}^{н}=12\cdot(1-0{,}2x_{23})\,;$$
$$T_{24}^{0}-T_{24}^{н}=4\cdot(1-0{,}5x_{24})\,;T_{45}^{0}-T_{45}^{н}=5\cdot(1-0{,}3x_{45});.$$

Время начала выполнения каждой операции должно быть не меньше времени окончания непосредственно предшествующей ей операции (моменты времени $T_{12}^{н}=T_{13}^{н}=T_{14}^{н}=0$): $T_{23}^{н}\ge T_{12}^{0}$; $T_{24}^{н}\ge T_{12}^{0}$; $T_{34}^{н}\ge T_{23}^{0}$; $T_{34}^{н}\ge T_{13}^{0}$; $T_{45}^{н}\ge T_{14}^{0}$; $T_{45}^{н}\ge T_{34}^{0}$; $T_{45}^{н}\ge T_{24}^{0}$.

Условия неотрицательности неизвестных:

$T_{ij}^{н}\ge 0$, $T_{ij}^{o}\ge 0$, $x_{ij}\ge 0$, для всех $(i,j)\in e$.

Сформулированная математическая модель задачи, содержащая 17 неизвестных и 20 ограничений, может быть решена, например, симплекс-методом.

12.2. Оптимизация комплекса операций по стоимости

Рассмотрим частный случай оптимизации комплекса операций – оптимизацию по затратам. Предположим, что затраты на выполнение отдельных операций находятся в обратной зависимости от продолжительности их выполнения. Коэффициент дополнительных затрат k_{ij} операции (i,j) вычисляется по формуле $k_{ij} = \frac{c'_{ij} - c''_{ij}}{t''_{ij} - t'_{ij}}$, где t'_{ij} – срочный режим выполнения операции (наименьшая продолжительность), которому соответствуют наибольшие затраты c'_{ij}; t''_{ij} – нормальный режим выполнения операции (наибольшая продолжительность), которому соответствуют минимальные затраты c''_{ij}.

Следовательно, продолжительность критического пути будет наибольшей, а стоимость выполнения комплекса операций наименьшей (минимальной). Необходимо сократить критический путь до некоторого возможного минимального значения при минимальном возрастании стоимости выполнения комплекса операций.

В общем случае сетевой график может содержать несколько критических путей, взаимосвязь между операциями которых может быть довольно сложной.

Рассмотрим более простой случай: будем полагать, что если график содержит несколько критических путей, то, или они не имеют общих операций, или же имеется одна либо несколько общих операций для всех критических путей.

При этом предположении алгоритм оптимизации комплекса операций по стоимости (затратам) сводится к следующему:

Предварительный шаг. Определяется коэффициент дополнительных затрат. Используя продолжительность операций t''_{ij}, находят: критический путь, длину критического пути $t_{кр}$, полные резервы времени операций R''_{ij} сетевого графика и затраты на реализацию комплекса операций C.

Общий шаг: 1. Среди критических операций находим ту, для которой коэффициент дополнительных затрат наименьший. Если найденная операция является общей для всех критических путей или если критический путь один, то она и подлежит сокращению. Если же найденная операция не является общей для критических путей, (пути могут иметь одну или несколько общих операций), то на каждом из них находим операцию с наименьшим коэффициентом дополнительных затрат. Просуммируем коэффициенты дополнительных затрат этих операций и сравним с коэффициентом дополнительных затрат той из общих операций, для которой он наименьший. Если сумма коэффициента дополнительных затрат операций больше коэффициента дополнительных затрат общей операции, то сокращению подлежит общая для критических путей операция. Если критические пути не имеют общих операций, то на каждом из них находится операция с наименьшим коэффициентом дополнительных затрат.

2. Производим сокращение продолжительности этой операции (операций) до тех пор, пока она (они) не достигнет (достигнут) минимальной продолжительности t_{ij} или не образуется новый критический путь (если полный резерв одной из критических операций будет равен 0).

3. Для данного варианта сетевого графика определим продолжительность критический путь $t_{кр}$, $R^{п}_{ij}$ и C.

4. Проверим, все ли операции критического пути достигли минимальной продолжительности. Если «да», то действия алгоритма считаются законченными, ибо сокращение продолжительности некритических операций увеличивает стоимость выполнения всего комплекса операций, не влияя на длину критического пути. Если же «нет», то перейдем к п. 1.

Пример. Сетевой график комплекса операций изображен на рис.1. Цифры, приписанные дугам графика, означают продолжительности выполнения операций в пормальном и срочном режимах соответственно. Прямые затраты на выполнение операций следующие:

$$C''_{12}=150; C'_{12}=190; C''_{13}=111; C'_{13}=175; C''_{14}=30; C'_{14}=90;$$

$$C''_{23}=66; C'_{23}=87; C''_{24}=72; C'_{24}=112; C''_{45}=89; C'_{45}=123;.$$

Требуется сократить критический путь при минимальном возрастании стоимости выполнения всех операций.

Предварительный шаг: Определяем коэффициент дополнительных затрат:

$k_{12}=5;\ k_{13}=8;\ k_{14}=15;\ k_{23}=3;\ k_{24}=40;\ k_{45}=17$.

Находим, что при наибольшей продолжительности операций t''_{ij} критический путь $\mu_{kp}=(1-2-3-4-5)$, $t_{kp}=31$ резервы времени некритических операций $R''_{13}=6$, $R''_{14}=16$, $R''_{24}=8$ и стоимость выполнения комплекса операций *C=518*. Результаты заносим в таблицу 1.

Первый шаг:

1) Среди критических операций наименьший коэффициент дополнительных затрат имеет операция (2,3): $k_{23}=3$.

2) Сокращаем время выполнения операции (2,3) на величину $(t''_{23}-t'_{23}; R''_{13}; R''_{14}; R''_{24})=\min(7;6;16;8)=6$.

3) В результате сокращения операции (2,3) образовались два критических пути: $\mu'=(1-2-3-4-5)$ и $\mu''=(1-3-4-5)$ с общими операциями (3,4) и (4,5). Продолжительность критического пути уменьшилась на 6 ед.: $t'_{kp}=25$.

Таблица 1

Параметры			Операции (i,j) (1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)	(4,5)	Продолжительность критического пути	Стоимость С
t''_{ij}			14	20	10	12	4	0	5		
t'_{ij}			6	12	6	5	3	0	3		
k_{ij}			5	8	15	3	40	∞	17		
t_{ij}	Шаг оптимизации	Предв.	14	20	10	12	4	0	5	31	518
		1	14	20	10	6	4	0	5	25	536
		2	14	19	10	5	4	0	5	24	547
		3	7	12	10	5	4	0	5	17	638
		4	7	12	10	5	4	0	5	15	670
R^{H}_{ij}		Предв.	0	6	16	0	8	0	0		
		1	0	0	10	0	2	0	0		
		2	0	0	9	0	1	0	0		
		3	0	0	2	0	1	0	0		
		4	0	0	2	0	1	0	0		
Сокращение времени Δt			7	8	0	7	0	0	2		
Приращение Стоимости Δc			35	64	0	21	0	0	34		

Резервы времени некритических операций составляют: $R^{п}_{14}=10$, $R^{п}_{24}=2$, а критических равны 0. Стоимость выполнения операций C=536 (данные занесены в строки первого шага оптимизации таблицы 1).

Поскольку, критические операции выполняются за время большее, чем t'_{ij}, то переходим ко второму шагу оптимизации.

Второй шаг:

1) Критической операцией с наименьшим коэффициентом дополнительных затрат является операция (2,3), для которой k_{23}=3. Но эта операция принадлежит только пути μ'', и уменьшение ее продолжительности не дает желаемого результата. Поэтому, на пути μ'' находим операцию с наименьшим коэффициентом дополнительных затрат, которая выполняется параллельно операции (2,3). Такой операцией является единственная операция (1,3), для которой k_{13}=8. Сумма коэффициентов дополнительных затрат $k_{13}+k_{23}$=11, что меньше $k_{34}=\infty$ и k_{45}=17, следовательно, сокращению подлежат операции (1,3) и (2,3).

2) Операции (1,3) и (2,3) сокращаем на 1 ед., т.к. наименьшая продолжительность операции (1,3) равна 5 и дальнейшее сокращение ее не возможно.

3) Рассчитав параметры сетевого графика, занесем их в строки второго шага оптимизации (таблица 1). Продолжительность операции (2,3): $t_{23}=t'_{23}=5$ выделяем жирными цифрами. Критические пути после сокращения операций сохранились: $\mu'=(1-2-3-4-5)$ и $\mu''=(1-3-4-5)$.

4) Учитывая, что не все критические операции выполняются в срочном режиме, переходим к выполнению третьего шага.

Третий шаг.

1) Из оставшихся критических операций наименьший коэффициент дополнительных затрат имеет операция (1,2), принадлежащая пути μ', для которой k_{12}=5. Из пути μ'' выбираем операцию (1,3), которая выполняется параллельно операции (1,2). Сумма $k_{12}+k_{13}$=13, что меньше $k_{34}=\infty$ и k_{45}=17. Следовательно, сокращению подлежат операции (1,2) и (1,3).

2) Сокращаем продолжительности операций (1,2) и (1,3) на 7 ед., т.к.

$$\min(t''_{12} - t'_{12}; t''_{13} - t'_{13}) = \min(14 - 6; 19 - 12) = 7$$

и эта величина меньше полного резерва некритических операций (1,4). Занесем продолжительности операций в строку третьего шага оптимизации. Дальнейшее сокращение продолжительности критической операции (1,3) невозможно, поэтому t_{13}=12 отмечаем жирной цифрой.

3) Рассматриваем параметры сетевого графика и заносим в соответствующие строки третьего шага оптимизации. Критические пути остались прежними: $\mu' = (1-2-3-4-5)$ и $\mu'' = (1-3-4-5)$.

4) Переходим к четвертому шагу оптимизации.

Четвертый шаг:

1) Из оставшихся критических операций наименьший коэффициент дополнительных затрат имеет операция (1,2). Однако, сокращать ее не имеет смысла, потому что уменьшение ее продолжительности не повлияет на длину критического пути, а лишь увеличит стоимость выполнения комплекса операций. Поэтому сокращению подлежит операция (4,5) для которой k_{45}=17.

2) Операцию (4,5), принадлежащую обоим критическим путям, сокращаем на 2 ед. времени.

3) Рассчитываем параметры сети и заносим их в таблицу. Дальнейшее сокращение операции (4,5) невозможно, поэтому $t_{45} = t'_{45} = 3$ отмечаем жирной цифрой.

4) Все операции критического пути (1-3-4-5) уменьшены до минимальных продолжительностей. Следовательно, действия алгоритма закончены.

Проблемные задания

1. Сформулируйте задачу оптимизации комплекса операций по времени.
2. Представьте задачу задания 1 сетевым графиком.
3. Исследуйте задачу с привлечением дополнительных средств и оцените количество неизвестных и ограничений задачи.
4. Решите задачу, предварительно установив метод ее решения.

Вопросы для самоподготовки

1. Напишите условия задачи оптимизации комплекса операций по времени.
2. Изобразите графически задачу комплекса операций по времени.
3. Напишите формулу вычисления коэффициента дополнительных затрат.
4. Напишите формулу вычисления минимальной стоимости выполнения комплекса операций.

§ 13. ЗАДАЧА О МАКСИМАЛЬНОМ ПОТОКЕ

Сеть, вершина, дуга, граф, симметричный граф, интенсивность, пропускная способность, поток, максимальный поток, разрез

13.1. Постановка задачи о максимальном потоке

Пусть задана сеть, состоящая из множества вершин E и множества дуг, соединяющих некоторые пары вершин, взятых из E. Предположим, что она является симметрическим графом, т.е., если дуга (E_i,E_j) входит в сеть, то в нее входит и симметричная дуга (E_i,E_j), хотя реально такой дуги может и не быть. Для определенности присвоим вершинам сети следующие обозначения: $E_0,E_1,E_2,\ldots,E_n$. Каждая вершина E_i характеризуется интенсивностью $d(E_i)$. Вершины, для которых $d(E_i)>0$, назовем источниками, вершины, для которых $d(E_i)<0$, – стоками, а остальные – промежуточными.

Из источников в стоки, по путям сети, направляется однородное вещество (газ, жидкость, транспорт и т.д.). Каждой дуге (E_i,E_j) сети поставлено в соответствие число b_{ij}, называемое пропускной способностью дуги, под которой понимается максимальное количество вещества, которое она может пропустить за ед. времени.

Пусть $d(E_0)>0$ и $d(E_n)>0$, тогда E_0 – единственный источник, E_n – единственный сток, а $E_1,E_2,\ldots,E_{n-1}$ промежуточные вершины сети.

Под потоком в сети (из источника в сток) будем понимать совокупность потоков $\{x_{ij}\}$ по всем дугам сети, где x_{ij} – поток по дуге (E_i,E_j), $i,j=\overline{0,n}, i \neq j$, равный количеству вещества, перемещаемого по ней в единицу времени. Отметим, что если пропускные способности симметричных дуг равны между собой, то эти дуги могут быть заменены ребрами. В силу сказанного, следующая постановка задачи справедлива для смешанных и неориентированных графов.

Ставится задача: определить для заданной сети максимальную величину потока из источника E_0 в сток E_n.

Обозначим сеть $G(E,e)$, где e – множество дуг или ребер или дуг, и ребер. Математически задача о максимальном потоке формулируется так:

найти неотрицательные значения x_{ij} (для всех $(E_i,E_j)\in e$), максимизирующие $v=\sum_{j=1}^{n}x_{0j}=\sum_{i=0}^{n-1}x_{in}$

при ограничениях $0\le x_{ij}\le b_{ij},\ i,j=\overline{0,n},\ i\ne j$,

$$\sum_{i=0}^{n-1}x_{ik}-\sum_{j=1}^{n}x_{kj}=0,\ k=\overline{1,n-1}$$

Условие $v=\sum_{j=1}^{n}x_{0j}=\sum_{i=0}^{n-1}x_{in}$ отражает величину максимального потока, который равен количеству вещества, вытекающего из источника, или количеству вещества, притекающего в сток. Условия $0\le x_{ij}\le b_{ij},\ i,j=\overline{0,n},\ i\ne j$ означают, что поток по каждой дуге должен быть неотрицательным и не превышать ее пропускной способности; из условия $\sum_{i=0}^{n-1}x_{ik}-\sum_{j=1}^{n}x_{kj}=0,\ k=\overline{1,n-1}$, следует, что количество вещества, протекающего в любую промежуточную вершину, равно количеству вещества, вытекающего из нее.

13.2. Понятие разреза

Важная роль в обосновании алгоритма решения задачи о максимальном потоке играет понятие разреза.

Разобьем множество всех вершин сети на два непересекающихся подмножества R и $\overline{R}$, $R \cup \overline{R} = E$, так, чтобы $E_0 \in R$, а $E_n \in \overline{R}$. Если выделить все дуги, начальные вершины которых принадлежат подмножеству R, а конечные - $\overline{R}$, то эти дуги будут образовывать разрез $(R, \overline{R})$, отделяющий источник E_0 от стока E_n. Таким образом, разрезом, отделяющим E_0 от E_n, называется совокупность всех дуг (E_i, E_j), которые исходят из вершин $E_i \in R$ и заканчивается в вершинах $E_j \in \overline{R}$.

Каждый разрез характеризуется пропускной способностью в $(R, \overline{R})$, которая численно равна сумме пропускных способнос-тей дуг, его образующих:

$$b(R, \overline{R}) = \sum_{E_i \in R,\ E_j \in \overline{R}} b_{ij}.$$

Любой путь из источника в сток содержит хотя бы одну дугу разреза $(R, \overline{R})$. Поэтому, если удалить все дуги какого ни будь разреза, то все пути из источника в сток будут блокированы. Поскольку, пропуская способность пути равна наименьшей из пропускных способностей дуг, входящих в этот путь, то величина потока v, перемещаемого по всем возможным путям, соединяющим E_0 с E_n не может превышать пропускной способности любого разреза сети, т.е. $v \le b(R, \overline{R})$. Следовательно, если удастся построить такой поток, то его величина v^* окажется равной пропускной способности некоторого разреза $(R^*, \overline{R}^*)$, т.е., $v^* = b(R^*, \overline{R}^*)$, то этот поток будет максимальным, а $(R^*, \overline{R}^*)$ – разрезом с минимальной пропускной способностью.

13.3. Теорема Форда – Фалкерсона

Прежде чем перейти к формулировке и доказательству теоремы, введем понятие прямой и обратной дуги цепи. Под цепью в данном случае, будем понимать последовательность сцепленных дуг сети без учета их ориентации.

Дугу, принадлежащую некоторой цепи, называют прямой, если ее направление совпадает с направлением обхода вершин этой цепи, обратной – в противном случае. Например, цепь $\mu=(E_3-E_5-E_4-E_7-E_6-E_8)$ (рис. 1), связывающая вершину E_3 с вершиной E_8, содержит три прямые дуги – (E_3,E_5), (E_4,E_7), (E_6,E_8) – и две обратные – (E_4,E_5), (E_6,E_7).

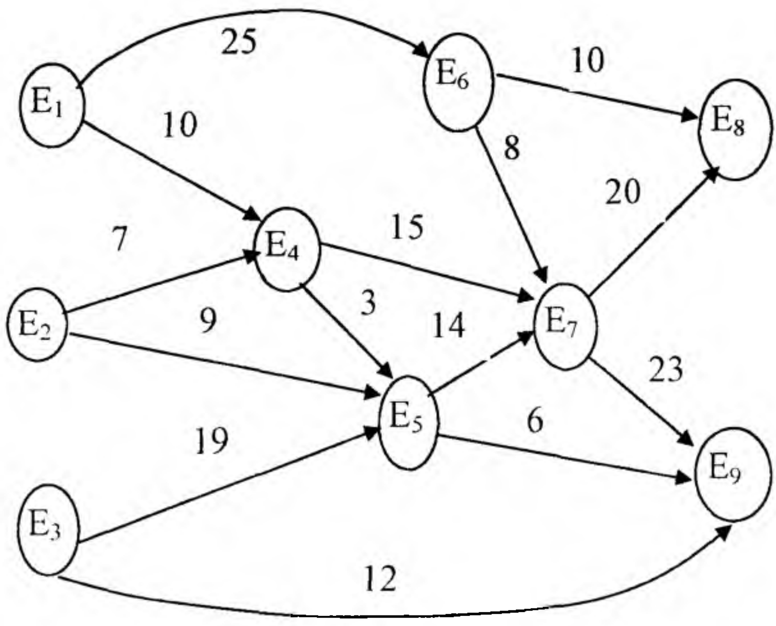


Рис. 1

Теорема. В любой сети максимальная величина потока из источника E_0 в сток E_n равна минимальной пропускной способности разреза, отделяющего E_0 от E_n.

Доказательство. Пусть имеем максимальный поток в сети. Если $x_{ij}=b_{ij}$, то будем говорить, что дуга (E_i,E_j) насыщена потоком.

Предположим, что v - величина максимального потока в сети $G(E,e)$. Необходимо доказать, что найдется такой разрез $(R,\bar{R})$ на этой сети, пропускная способность которого равна v, т.е. в $(R,\bar{R})= v$.

Такой разрез можно построить, если подмножество R включая все вершины, которые достигаются по некоторой цепи из E_0 , а подмножество $\bar{R}$ - все остальные вершины.

Вершины составляющие подмножество R, определяются последовательно, начиная с источника E_0 по следующему правилу:

$$E_0 \in R,$$
$$\text{если } E_i \in R \text{ и } x_{ij} < b_{ij}, \text{ то } E_j \in R;$$
$$\text{если } E_i \in R \text{ и } x_{ij} > 0, \text{ то } E_j \in R.$$

Покажем, что применение этих правил приводит к построению разреза. Очевидно, разрез будет построен в том случае, если сток $E_n \in R$.

Предположим, что $E_n \in R$. Тогда из вышеприведенных правил следует, что существует цепь из E_0 в E_n с пропускной способностью больше нуля, $\mu=(E_{i1}-E_{i2}-E_{i3}-\ldots -E_{im})$, где $E_{i1}=E_0$; $E_{im}=E_n$, так как все прямые дуги, входящие в цепь, ненасыщенные $(x_{i_s i_{s+1}} < b_{i_s i_{s+1}})$, а для всех обратных дуг $(E_{i_{s+1}}, E_{i_s})$ величина потока больше нуля $(x_{i_{s+1} i_s} > 0)$.

Обозначим через δ_1 наименьшую разность между пропускной способностью и величиной потока, взятую по всем прямым дугам цепи. Определим величину $\delta = \min(\delta_1, \delta_2) > 0$. Увеличим поток на δ по всем прямым дугам цепи и уменьшим на эту же величину по всем ее обратным дугам. Таким образом, величина нового потока равна $\nu+\delta$, что противоречит предположению о максимальности ν. Следовательно, $E_n \in \overline{R}$, и множество дуг $(R,\overline{R})$ есть разрез, отделяющий источник от стока.

Докажем теперь, что пропускная способность построенного разреза равна максимальному потоку ($b(R,\overline{R})=\nu$). Из правил построения разреза $(R,\overline{R})$, точнее нахождения подмножества вершин R следует, что если $E_i \in R$, $E_j \in \overline{R}$, то $x_{ij}=b_{ij}$, в противном случае, вершина E_j входила бы в R. Таким образом:

$$\sum_{i\in R,\, j\in \overline{R}} x_{ij} = \sum_{i\in R,\, j\in \overline{R}} b_{ij} = b(R,\overline{R}) = \nu.$$

Теорема доказана.

Проблемные задания

1. Сформулируйте задачу о максимальном потоке на примере функционирования газораспределительной сети.
2. Исследуйте возможности практического использования в сети симметричной дуги.
3. Исследуйте возможность сведения задачи о максимальном потоке задания 1, к задаче с единственными источником и стоком.
4. Исследуйте возможности определения всех разрезов для задачи задания 1.

Вопросы для самоподготовки

1. Охарактеризуйте цель задачи о максимальном потоке.

2. Что называется пропускной способностью?

3. Сформулируйте математически задачу о максимальном потоке.

4. Чем характеризуется каждый разрез?

§ 14. АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ О МАКСИМАЛЬНОМ ПОТОКЕ

Сеть, вершина, дуга, граф, симметричный граф, интенсивность, пропускная способность, поток, максимальный поток, алгоритм Форда, задача со многими источниками и стоками

14.1. Алгоритм Форда решения задачи о максимальном потоке

Приведем табличный алгоритм Форда нахождения максимального потока в сети, состоящий из ряда шагов.

Предварительный шаг. Запишем пропускные способности дуг сети в таблицу размерности $(n+1)\times(n+1)$, где n+1 – количество вершин сети $G(E,e)$. Если пропускная способность дуги $(E_i, E_j)>0$, а симметричной ей дуги (E_j, E_i) равна нулю, то в клетку (i,j) заносим элементы b_{ij}, а в клетку (j,i) – нуль; если же $b_{ij}=b_{ji}=0$, то клетки (i,j) и (j,i) не заполняем.

Общий шаг. Общий повторяющийся шаг состоит из трех действий:

1. Находим по таблице путь из E_0 в E_n с пропускной способностью больше нуля. Для этого столбец, соответствующий вершине E_0 пометим знаком *. Отыщем в E_0 – й строке элементы $b_{oj}>0$ и столбцы, в которых они находятся, пометим сверху номером просматриваемой строки (цифрой 0). В результате окажутся выделенными все дуги (E_0,E_j) с положительной пропускной способностью. Эти дуги могут служить первыми дугами пути из E_0 в E_n.

Далее просмотрим те строки номера, которых совпадают с номерами помеченных столбцов. В каждой такой строке, например, E_i -й, отыщем элементы $b_{ij}>0$, находящиеся в непомеченных столбцах, и пометим эти столбцы номером просматриваемой строки. Таким образом, окажутся выделенными дуги, которые могут служить вторыми дугами путей из E_0 в E_n. Продолжим аналогичный просмотр строк, номера которых совпадают с номерами помеченных столбцов, до тех пор, пока:

а) либо не будет помечен столбец E_n (т.е. сток). Это означает, что удалось выделить путь с пропускной способностью больше нуля из E_0 в E_n;

б) либо не убедимся в том, что нельзя пометить новые столбцы (в просматриваемых строках не окажется положительных элементов, расположенных в непомеченных столбцах). В этом случае отсутствует путь из E_0 в E_n, проходящий по дугам с положительной пропускной способностью.

В случае а) искомый путь μ из E_0 в E_n находим, используя пометки столбцов.

Пусть последняя вершина пути E_n помечена номером k. По этой пометке находим предшествующую вершину E_k (при просмотре строки E_k были помечены столбцы E_n и дуга (E_k,E_n) – последняя в искомом пути). Элемент $b_{kn}>0$, стоящий на пересечении E_k -й строки и E_n -го столбца, отметим знаком минус (получим b_{kn}^-), а симметричный ему элемент b_{nk}, находящийся на пересечении E_k-й строки и E_n-го столбца, – знаком плюс (получим b_{nk}^+)). Т.к просматривалась E_k-ая строка, то перед этим был помечен, например, номером r E_k-й столбец. Поэтому двигаясь от элемента b_{nk} по E_k-му столбцу до r-ой строки (дуга (E_r,E_k) предшествует дуге (E_k,E_r)) и отмечаем b_{rk} знаком минус, а b_{kr} – знаком плюс. Продолжаем этот процесс до тех пор, пока не придем к E_0-й строке и не отметим знаком минус элемент этой строки и знаком плюс симметричной ему элемент. Переходим к действию 2.

В случае б) столбец E_n, соответствующий стоку, пометить не возможно. Следовательно, не существует больше пути с пропускной способностью больше 0 из в E_0 в E_n, и общий повторяющийся шаг закончен. Вершины, находящихся в помеченных столбцах таблицы, образуют подмножество R (эти вершины достижимы по некоторому пути из источника E_0), остальные вершины входят в подмножество $\overline{R}$. Дуги, исходящие из вершин $E_i \in R$ и входящие в вершины $E_j \in \overline{R}$, образуют разрез с минимальной пропускной способностью $b(R,\overline{R}) = v = \sum\limits_{E_i \in R,\, E_j \in \overline{R}} b_{ij}$, где b_{ij} - пропускные способности дуг исходной сети. Переходим к заключительному шагу алгоритма.

2. Находим пропускную способность θ найденного пути μ. Под пропускной способностью понимается максимальное количество вещества, которое можно переместить из E_0 в E_n по пути μ в единицу времени. Пропускная способность пути равна наименьшей из пропускных способностей дуг, входящих в этот путь, т.е. $\theta = \min\limits_{(i,j) \in \mu} \{b_{ij}\}$.

3. Определяем остаточные пропускные способности дуг, найденного пути и симметричных к ним. Для этого из элементов таблицы b_{ij}^- вычитаем θ, а к элементам b_{ij}^+ прибавляем θ. В результате выполнения этого действия получим новую таблицу, аналогичную исходной, но с измененными пропускными способностями. После получения новой таблицы возвращаемся к действию 1 общего шага, который применяем, до тех пор, пока не получим окончательную таблицу, в которой нет ни одного пути из E_0 в E_n с пропускной способностью больше нуля.

Заключительный шаг. Из элементов исходной таблицы вычитаем соответствующие элементы таблицы, полученной на последнем шаге. В результате получим таблицу, положительные элементы которой равны величинам потоков x_{ij} по соответетствующим дугам (E_i,E_j), и максимальный поток в сети – сумме элементов E_0-й строки или E_n-го столбца, т.е. $v = \sum_{j=1}^{n} x_{0j} = \sum_{i=0}^{n-1} x_{in}$.

14.2. Пример применения алгоритма Форда

Для сети, изображенной на рис. 1, найти максимальный поток из вершины E_0 в вершину E_4.

Легко заметить, что сеть, изображенная на рис. 1, содержит, наравне, с дугами (они нарисованы в виде направленной линии – стрелки) и ребра. Ребра на рисунке изображены с помощью линий без направлений. Для них пропускная способность, написанная рядом с линией, считается в обоих направлениях, а для дуг – только в направлении дуги.

Применим алгоритм Форда для решения этого примера.

Предварительный шаг. Формируем матрицу пропускных способностей дуг сети (таблица 1).

Таблица 1

	(*)	(0)	(0)	(1)	(2)
E_j / E_i	E_0	E_1	E_2	E_3	E_4
E_0		17	19^-		
E_1	0		4	12	
E_2	0^+	4		8	9^-
E_3		12	6		24
E_4			0^+	0	

Первый шаг. 1) Находим по таблице 1 какой-либо путь с положительной пропускной способностью из вершины E_0 в вершину E_4.

Для этого столбец E_0 помогаем знаком *. В строке E_0 положительные элементы расположены в столбцах E_1 и E_2. Следовательно, эти столбцы помечаем сверху цифрой 0 (номером рассматриваем строки). Далее просматриваем элементы строк, номера которых совпадают с номерами помеченных столбцов. Просматривая строку E_1, помечаем номером 1 столбец E_3, и просматривая очередную строку E_2, помечаем номером 2 столбец E_4. Т.к. вершина E_4 – сток, то процесс пометок закончен и искомый путь из источника в сток существует. Последней дугой этого пути является дуга (E_2,E_4). Отмечаем элемент $b_{24}=9$ знаком минус, симметричный ему элемент $b_{42}=0$ – знаком плюс. Поскольку столбец E_2 помечен номером 0, то элемент $b_{02}=19$ отмечаем знаком минус, а элемент $b_{20}=0$ – знаком плюс. В результате получим путь $\mu_1=(E_0-E_2-E_4)$.

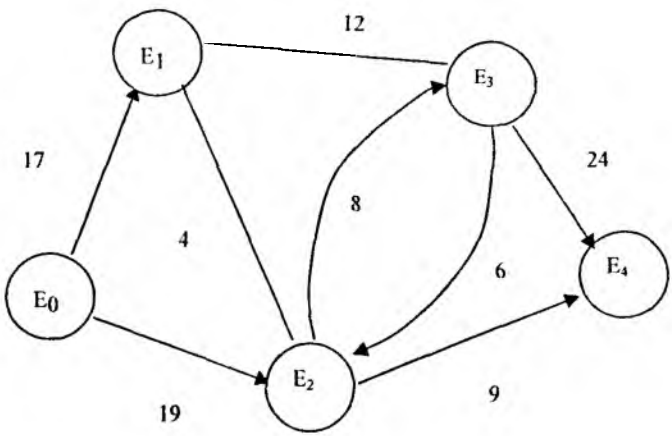


Рис. 1

2) Определяем пропускную способность найденного пути $\theta_1 = \min\{b_{02}^-, b_{24}^-\} = \min\{19,9\} = 9$.

3) Пропускные способности дуг найденного пути уменьшим на $\theta_1=9$, а симметричных к ним дуг увеличим на ту же величину. Получим таблицу 2.

Таблица 2

	(*)	(0)	(0)	(1)	(3)
E_j / E_i	E_0	E_1	E_2	E_3	E_4
E_0		17^-	10		
E_1	0^+		4	12^-	
E_2	9	4		8	0
E_3		12^+	6		24^-
E_4			9	0^+	

Второй шаг. 1) Пометив столбцы таблицы 2 и расставив знаки, находим путь $\mu_2=(E_0-E_1-E_3-E_4)$. При этом элементы b_{01}, b_{13}, b_{34} окажутся помеченными знаком минус, а симметричные им элементы b_{10}, b_{31}, b_{43} – знаком плюс.

2) Пропускная способность пути μ_2 равна $\theta_2=\min\{b_{01}^-, b_{13}^-, b_{34}^-\}=\min\{17,12,24\}=12$.

3) Изменив пропускные способности дуг на θ_2, получим таблицу 3.

Таблица 3

	(*)	(0)	(0)	(2)	(3)
E_j / E_i	E_0	E_1	E_2	E_3	E_4
E_0		5	10^-		
E_1	12		4	0	
E_2	9^+	4		8^-	0
E_3		24	6^+		12
E_4			9	12^+	

Третий шаг: 1) Пометив столбцы, находим путь из источника E_0 в сток E_4: $\mu_3=(E_0-E_2-E_3-E_4)$.

2) Величина потока по пути μ_3 равна $\theta_3 = \min\{10,8,12\} = 8$.

3) Вычислив новые пропускные способности дуг, приходим к таблице 4.

Таблица 4

E_j / E_i	(*) E_0	(0) E_1	(0) E_2	E_3	E_4
E_0		5	2		
E_1	12		4	0	
E_2	17	4		0	0
E_3		24	14		4
E_4			9	20	

Четвертый шаг: 1) Столбец E_0 помечаем знаком *. Просматривая E_0–ю строку, помечаем номером 0 столбцы E_1 и E_2. Продолжая просмотр строк, убеждаемся, что столбцы E_3 и E_4 пометить не возможно. Следовательно, больше не существует ни одного пути с положительной пропускной способностью из вершины E_0 в вершину E_4 Подмножество R образуют помеченные вершины E_0, E_1, E_2 (см таблицу 4), подмножество $\overline{R}$ – непомеченные вершины E_3 и E_4. Разрез с минимальной пропускной способностью образуют дуги, начальные вершины которых принадлежат подмножеству R, а конечные – $\overline{R}$. Таким образом, разрез с минимальной пропускной способностью $(R^*, \overline{R}^*) = \{(E_1,E_3),(E_2,E_3),(E_2,E_4)\}$. Действительно, удалив дуги разреза, мы блокируем все пути из источника в сток. Пропускная способность разреза $b((R^*,\overline{R}^*))=b_{13}+b_{23}+b_{24}=12+8+9=29$.

Заключительный шаг. Вычитая из элементов таблицы 1, соответствующие элементы таблицы 4 получим таблицу 5.

Положительные элементы таблицы 5 характеризуют величины дуговых потоков. Следовательно, $x_{01}=12$, $x_{02}=17$, $x_{13}=12$, $x_{23}=8$, $x_{24}=9$, $x_{34}=20$, а по всем остальным дугам потоки равны нулю.

Таблица 5

E_j / E_i	E_0	E_1	E_2	E_3	E_4
E_0		12	17		
E_1	-12		0	12	
E_2	-17	0		8	9
E_3		-12	8		20
E_4			-9	-20	

Величина максимального потока равна сумме элементов E_0–ой строки таблицы 5 или сумме элементов E_4–го столбца: $v=\sum_{j=1}^{4} x_{ij}=12+17=\sum_{i=0}^{3} x_{ij}=9+20=29$.

Как видно, $v^*=b(R^*,\overline{R}^*)$. Дуги разреза $(R^*,\overline{R}^*)$ насыщены потоком ($x_{13}=b_{13}=12$; $x_{23}=b_{23}=8$; $x_{24}=b_{24}=9$).

14.3. Сведение задачи с несколькими источниками и стоками к задаче с одним источником и одним стоком

Рассмотрим сеть $G=(E,e)$, состоящую из множества источников $S=\{E_1,\ldots,E_q\}$, множества промежуточных вершин (узлов) $N=\{E_{q+1},\dots,E_m\}$ и множества стоков $T=\{E_{m+1},\dots,E_n\}$. Под потоком в сети из S в T будем понимать действительную функцию v, определенную на множестве дуг e и удовлетворяющую ограничения вида:

$$0 \le x_{ij} \le b_{ij}, \quad i,j=\overline{0,n}; \quad i \ne j; \quad \sum_{i=0}^{n-1} x_{ik} - \sum_{j=1}^{n} x_{kj} = 0; \quad k=\overline{1,n-1}.$$

Величина потока находится из выражения

$$v = \sum_{i=1}^{l}\sum_{j=l+1}^{m} x_{ij} + \sum_{i=1}^{l}\sum_{j=m+1}^{n} x_{ij} = \sum_{i=1}^{l}\sum_{j=m+1}^{n} x_{ij} + \sum_{i=l+1}^{m}\sum_{j=m+1}^{n} x_{ij}.$$

Расширив сеть $G(E,e)$ до сети $G'=(E',e')$, добавив две фиктивные вершины E_0 и E_{n+1} (E_0 – источник; E_{n+1} – сток) и фиктивные дуги (E_0,E_j), $j=\overline{1,q}$ и (E_i,E_{n+1}), $i=\overline{m+1,n}$. Пропускные способности фиктивных дуг $b_{0j}=d(E_j)$, $j=\overline{1,q}$ и $b_{i,n+1}=d(E_i)$, $i=\overline{m+1,n}$.

Величину потока как в исходной сети $G(E,e)$, так и в расширенной сети $G'=(E',e')$ определяют пропускные способности дуг исходной сети. Таким образом, задача о максимальном потоке из множества источников S во множество стоков T сети $G(E,e)$ равносильна задаче о максимальном потоке из единственного источника E_0 в единственный сток E_{n+1} сети $G'=(E',e')$.

Проблемные задания

1. Поставьте реальную задачу нахождения максимального потока, составьте ее математическую модель и установите суть предварительного шага алгоритма Форда для рассматриваемой задачи.

2. Выполняйте все три действия общего шага алгоритма для задачи пункта 1 (при необходимости повторите эти действия нужное количество раз).

3. Исследуйте заключительный шаг алгоритма для задачи пункта 1 и дайте оценку полученному результату.

Вопросы для самоподготовки

1. Проанализируйте общий шаг алгоритма Форда.
2. Как определяется пропускная способность найденного пути?
3. Как определяются остаточные пропускные способности дуг найденного пути?
4. Напишите формулу определения максимального потока сети.
5. Сформулируйте способ сведения задачи со многими источниками и стоками к задаче с одним источником и одним стоком.

§ 15. ЗАДАЧА О КРАТЧАЙШЕМ МАРШРУТЕ

Задача, сеть, путь, вершина сети, кратчайший путь, оптимальный поток, подмножество, дуга.

15.1. Математическая постановка задачи

Рассмотрим задачу определения кратчайшего пути между двумя вершинами сети. Пусть задана сеть $G(E,e)$, каждой дуге (ребру) которого соответствует некоторое расстояние l_{ij}. Требуется найти кратчайший маршрут из вершины E_0 в вершину E_n.

Математическая постановка задачи имеет вид:

минимизировать $\sum\limits_{(i,j)\in e} l_{ij} x_{ij}$

при ограничениях

$$\sum_{(k,j)\in e} x_{kj} - \sum_{(i,k)\in e} x_{ik} = \begin{cases} 1, \text{если } E_k = E_0; \\ 0, \text{если } E_k \neq E_0, E_n, k = \overline{1, n-1}; \\ -1, \text{если } E_k = E_n; \end{cases}$$

$$x_{ij} = \begin{cases} 1, \text{если дуга } (E_i, E_j) \text{ входит в путь;} \\ 0 - \text{в противном случае.} \end{cases}$$

Двойственная задача сводится к максимизации выражения $\varepsilon_0 - \varepsilon_n$

при ограничениях: $\varepsilon_j - \varepsilon_i \leq l_{ij}$ для всех $(E_i,E_j)\in e$,

где все двойственные переменные не ограничены по знаку. Значение $\varepsilon_k, k=0,n$ равно наименьшему расстоянию из узла E_0 в узел E_k.

Рассмотрим алгоритм отыскания значений двойственных значений переменных. Обозначим через R подмножество помеченных вершин сети, а через $\overline{R}$ – подмножество непомеченных вершин, и $RO\overline{R} = E$..

15.2. Алгоритм решения задачи

Решение задачи с учетом обозначений сводится к следующему алгоритму:

Предварительный шаг: Помечаем источник E_0 числом $\varepsilon_0 = 0$ и переходим к общему шагу (после выполнения этого шага вершина $E_0 \in R$).

Общий шаг: 1. Находим дуги, начальные вершину (E_i) которых принадлежат подмножеству R, а конечные (E_j) – в $\overline{R}$. Для каждой из этих дуг $(E_i, E_j) \in (R, \overline{R})$ определяем величину $h_{ij} = \varepsilon_i + l_{ij}$, где ε_i – числа (пометки), приписанные вершинам $E_i \in R$.

Находим значение $\varepsilon_j = \min\limits_{(E_i, E_j) \in (R, \overline{R})} h_{ij}$ и выделяем дуги, на которых достигается этот минимум. Вершинам $E_j \in \overline{R}$, являющимися конечными вершинами выделенных дуг, припишем значение ε_j.

2. Проверяем, выполняется ли условие $\varepsilon_i + l_{ij} \ge \varepsilon_j$ для всех дуг сети, оба конца которых принадлежат R. Если это условие для какой-то дуги не выполняется, т.е. $\varepsilon_j > \varepsilon_i + l_{ij}$, то соответствующее значение ε_j заменим на $\varepsilon_i + l_{ij}$. Выделим дугу (E_i, E_j) и переходим к пункту 1. Пометку вершин продолжим до тех пор, пока не будет помечен сток E_n. Длину кратчайшего маршрута от E_0 до E_n указывает значение ε_n.

Заключительный шаг. Определим оптимальный маршрут или оптимальные маршруты (если их несколько), двигаясь по выделенным дугам от стока E_n к источнику E_0 в направлении, обратном их ориентации. При этом в маршрут включаются те дуги (E_i,E_j) для которых $\varepsilon_j - l_{ij} = \varepsilon_i$.

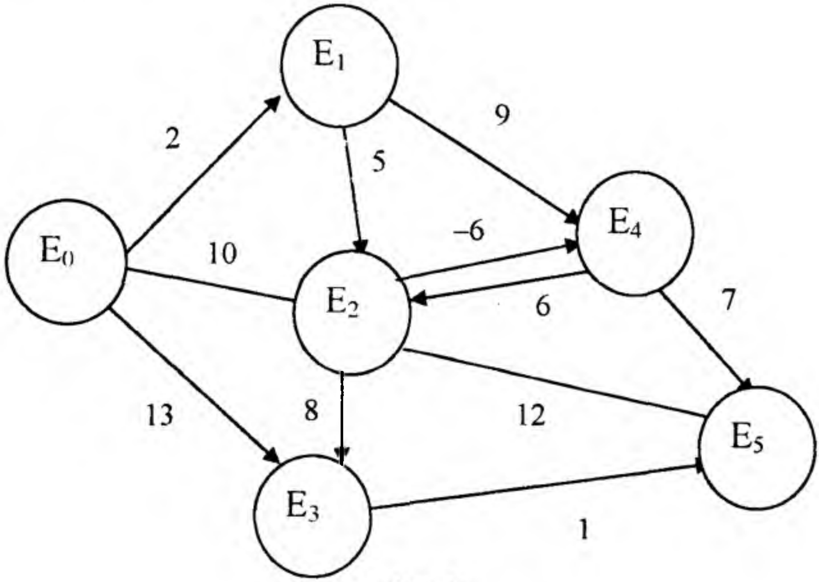


Рис. 1

Алгоритм сходится за конечное число шагов при условии, что сумма длин дуг любого контура, содержащегося в сети, неотрицательна.

Пример: Найти кратчайший маршрут из E_0 в E_5 в сети, показанной на рис. 1. Значения l_{ij} приписаны дугам и ребрам.

Предварительный шаг: Источнику приписываем пометку $\varepsilon_0 = 0$ ($E_0 \in R$).

Первый шаг: а) $R = \{E_0\}$, $\overline{R} = \{E_1, E_2, E_3, E_4, E_5\}$, $(R,\overline{R}) = \{(E_0,E_1),(E_0,E_2),(E_0,E_3)\}$.

Находим $h_{01} = \varepsilon_0 + l_{01} = 0 + 2 = 2$, $h_{02} = \varepsilon_0 + l_{02} = 0 + 10 = 10$ и $h_{03} = \varepsilon_0 + l_{03} = 0 + 13 = 13$. Среди величин h_{ij} минимальной является h_{01}, поэтому вершине E_1 приписываем пометку $\varepsilon_1 = 2$ и выделяем дугу (E_0, E_1) жирной линией ($E_1 \in R$).

б) Для (E_0, E_1) условие $\varepsilon_i + l_{ij} \geq \varepsilon_j$ выполняется.

Второй шаг: а) $R = \{E_0, E_1\}$, $\overline{R} = \{E_2, E_3, E_4, E_5\}$ $(R, \overline{R}) = \{(E_0, E_2), (E_0, E_3), (E_1, E_2), (E_1, E_n)\}$. $h_{02}=10$, $h_{03}=13$, $h_{12}=7$, $h_{14}=11$, $\min\{h_{02}, h_{03}, h_{12}, h_{14}\}=h_{12}=7$. Минимум достигается по дуге (E_1, E_2).

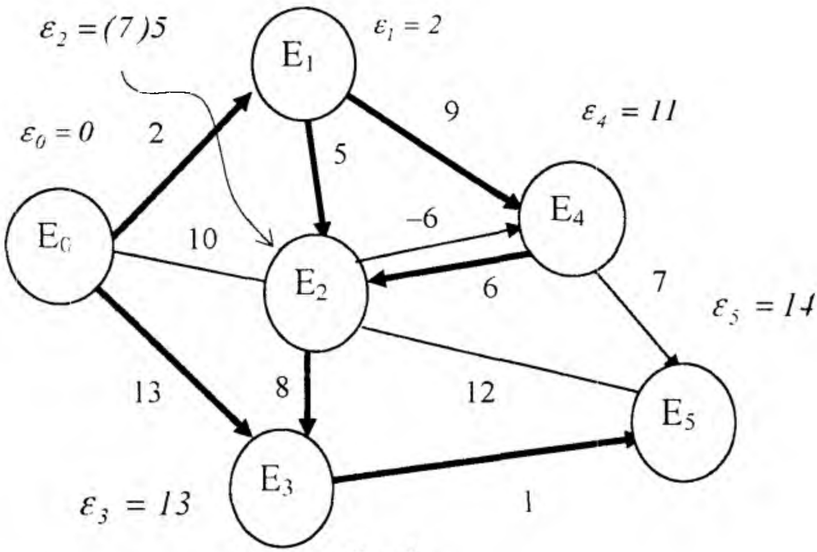


Рис. 2

Вершине E_2 приписываем $\varepsilon_2 = 7$, ($E_2 \in R$).

б) Для (E_0, E_1) имеем $\varepsilon_0 + l_{01} = \varepsilon_1$, т.е. $0+2 \geq 2$, условие выполняется и т.д. для всех остальных рассматриваемых дуг условия так же выполняются.

Третий шаг: а) $R = \{E_0, E_1, E_2\}$ $\bar{R} = \{E_3, E_4\}$ $(R, \bar{R}) = \{(E_0, E_3), (E_1, E_4), (E_2, E_3), (E_2, E_4)\}$, $h_{03}=13$, $h_{14}=11$, $h_{23}=15$, $h_{24}=13$, $\min\{h_{03}, h_{14}, h_{23}, h_{24}\}=h_{14}=11$, $\varepsilon_4 = 11$.

б) (E_0, E_1) – условие $\varepsilon_0 + l_{01} \geq \varepsilon_1$ $(0+2 \geq 2)$ выполняется;

(E_1, E_2) – условие $\varepsilon_1 + l_{12} \geq \varepsilon_2$ $(2+5 \geq 7)$ выполняется;

(E_0, E_2) – условие $\varepsilon_0 + l_{02} \geq \varepsilon_2$ $(0+10 \geq 10)$ выполняется;

(E_1, E_4) – условие $\varepsilon_1 + l_{14} \geq \varepsilon_4$ $(2+9 \geq 11)$ выполняется;

(E_2, E_4) – условие $\varepsilon_2 + l_{24} \geq \varepsilon_4$ $(7+6 \geq 11)$ выполняется;

(E_4, E_2) – условие $\varepsilon_4 + l_{42} \geq \varepsilon_2$ $(11+(-6) < 7)$ не выполняется.

Поэтому производим замену $\varepsilon_2 = 7$ на $\varepsilon_4 + l_{42} = 5$ и выделяем дугу (E_4, E_2).

Четвертый шаг: а) $R = \{E_0, E_1, E_2, E_4\}$ $\bar{R} = \{E_3, E_5\}$, $(R, \bar{R}) = \{(E_2, E_5), (E_3, E_5), (E_4, E_5)\}$, $h_{25}=17$, $h_{35}=14$, $h_{45}=18$, $\min\{h_{25}, h_{35}, h_{45}\}=h_{35}=14$, $\varepsilon_5 = 14$ – минимум достигается по дуге (E_3, E_5).

Заключительный шаг. Находим оптимальный маршрут.

Так как $\varepsilon_5 - l_{35} = 13 = \varepsilon_3$, $\varepsilon_3 - l_{23} = 5 = \varepsilon_2$, $\varepsilon_2 - l_{42} = 11 = \varepsilon_4$, $\varepsilon_4 - l_{14} = 2 = \varepsilon_4$, $\varepsilon_1 - l_{01} = 0 = \varepsilon_0$, то кратчайшим маршрутом является путь (рис. 2) $\mu_1 = (E_0 - E_1 - E_4 - E_2 - E_3 - E_5)$, и он для данной сети не единственный. Другим кратчайшим маршрутом будет путь $\mu_2 = (E_0 - E_3 - E_5)$, поскольку, $\varepsilon_3 - l_{03} = 0 = \varepsilon_0$.

Проблемные задания

1. Сформулируйте задачу определения кратчайшего пути и постройте ее математическую модель.
2. Исследуйте алгоритм отыскания значений двойственных переменных для задачи задания 1.
3. Решите задачу задания 1 и дайте оценку полученному результату.

Вопросы для самоподготовки

1. Сформулируйте и напишите математическую постановку задачи о кратчайшем маршруте.
2. Какие операции выполняются при предварительном шаге алгоритма нахождения кратчайшего маршрута?
3. Напишите формулу определения числа, приписываемого вершинам сети.

§ 16. ЗАДАЧА О ПОТОКЕ МИНИМАЛЬНОЙ СТОИМОСТИ. АЛГОРИТМ БАСАКЕРА – ГОУЭНА

Задача, сеть, путь, вершина сети, кратчайший путь, оптимальный поток, подмножество, дуга

16.1. Постановка задачи о потоке минимальной стоимости

Пусть задана сеть $G(E,e)$ с одним источником E_0, одним стоком E_n и промежуточными вершинами E_1, $E_2, \ldots, E_{n-1}$. Каждой дуге (E_i, E_j) (ребру $(\overline{E_i, E_j})$) сети поставлены в соответствие две величины: пропускная способность дуги (ребра) b_{ij}; дуговая стоимость c_{ij} (стоимость доставки единицы потока по дуге (E_i, E_j) или ребру $(\overline{E_i, E_j})$), одинаковая в обоих направлениях. Необходимо найти поток. Под стоимостью будем понимать стоимость доставки того или иного количества вещества из источника в сток. При этом предполагается, что заданная величина потока не превышает величины максимального потока из E_0 в E_n

Формальная запись задачи имеет вид:

$$f = \sum_{(i,j)\in e} c_{ij} x_{ij} \to min \cdot$$

$$0 \le x_{ij} \le b_{ij}, \quad i,j = \overline{0,n}, \quad i \ne j, \sum_{i=1}^{n-1} x_{ij} - \sum_{n=1}^{n} x_{kj} = 0, \quad k = \overline{1,n-1},$$

$$\sum_{j=1}^{n} x_{0j} = \sum_{i=0}^{n-1} x_{in} = B.$$

Если бы не было ограничений на пропускные способности дуг (ребер), то для решения задачи достаточно было бы найти самый экономичный путь (путь минимальной стоимости) из E_0 в E_n и пропустить по нему весь поток. Путь минимальной стоимости – это путь, сумма стоимостей которого, приписанных дугам, является минимальной. При наличии ограничений на пропускные способности дуг (ребер) можно последовательно находить различные пути минимальной стоимости и пропускать потоки по ним до тех пор, пока суммарная величина потока по всем путям не будет равна заданной величине потока. Ниже предлагается рассмотреть алгоритм нахождения потока минимальной стоимости, основанный на этом подходе.

16.2. Алгоритм Басакера – Гоуэна

Первый шаг. В исходной сети все дуговые потоки из E_0 в E_n полагаем равными нулю.

Второй шаг. Находим путь μ минимальной стоимости из E_0 в E_n, используя стоимости, приписанные дугам на первой итерации, и модифицированные стоимости на последующих итерациях.

Третий шаг. Определим пропускную способность пути μ по формуле $\theta = \min\limits_{(E_i,E_j)\in\mu} b_{ij}$, добавим к величине старого потока $v_{стар}$ величину $\theta' = \min(\theta, B - v_{стар})$ и сравним с заданной величиной потока B. Если величина суммарного потока равна B, то задача решена и переходим к шестому шагу. В противном случае переходим к четвертому шагу.

Четвертый шаг. Находим величину потока по каждой дуге, принадлежащей пути μ, для чего к старым величинам дуговых потоков пути μ добавляем величину θ. Пропускные способности дуг, симметричных дугам пути

μ, полагаем равными величинам соответствующих дуговых потоков, т.е. $b_j = x_{ij}$.

Пятый шаг. Определим модифицированные дуговые стоимости c^*_{ij} по формуле:

$$c^*_{ij} = \begin{cases} c_{ij}, & \text{если} \quad 0 \le x_{ij} \le b_{ij}, \\ \infty, & \text{если} \quad x_{ij} = b_{ij}. \end{cases}$$

Если для какой-то дуги $x_{ij} < 0$, то модифицированная стоимость симметричной ей дуги (j,i) равна величине c_{ij}, взятой с обратным знаком. Переходим к выполнению второго шага.

Шестой шаг. Минимальную стоимость потока заданной величины определим по формуле:

$$f = \sum_{(i,j)\in c} c_{ij} x_{ij} \cdot$$

16.3. Пример

Найти поток из E_0 в E_5 величины $B=3$, обладающей минимальной стоимостью, для сети, представленной на рис. 1. Первое число, приписанное каждому ребру, означает пропускную способность, второе – стоимость доставки единицы потока по ребру из одной вершины в другую.

Доставка потока по ребру может осуществляться в любом направлении.

Итерация 1. Первый шаг. Полагаем величину потока по каждому ребру сети равной нулю $(x_{ij}=0)$.

Второй шаг. Применяя алгоритм нахождения кратчайшего маршрута, определяем, что минимальная стоимость доставки единицы потока из E_0 в E_5, равная $\varepsilon_3=4$, достигается по путям $\mu_1=(E_0\text{-}E_1\text{-}E_3\text{-}E_5)$ и $\mu_2=(E_0\text{-}E_2\text{-}E_3\text{-}E_5)$, изображенным на рис. 2 жирными линиями.

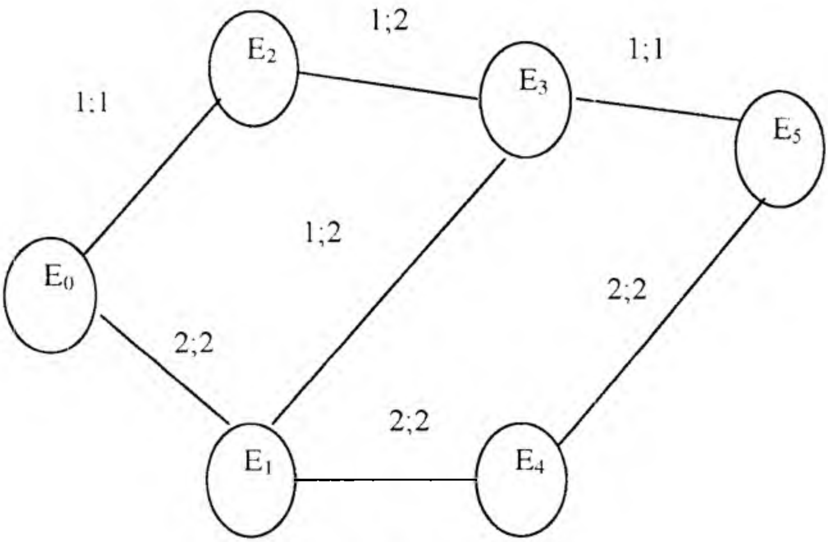


Рис. 1

Третий шаг. Определяем пропускные способности выделенных путей μ_1 и μ_2:

$$\theta_1 = \min_{(E_i,E_j)\in\mu_1} b_{ij} = \min(2,\ 1,\ 1) = 1\,,$$

$$\theta_2 = \min_{(E_i,E_j)\in\mu_2} b_{ij} = \min(1,\ 1,\ 1) = 1\,.$$

Так как дуга (E_3,E_5) с пропускной способностью $b_{35}=1$ не позволяет использовать два пути, то выберем первый путь. Учитывая, что $\mu_2=1<B=3$, переходим к четвертому шагу.

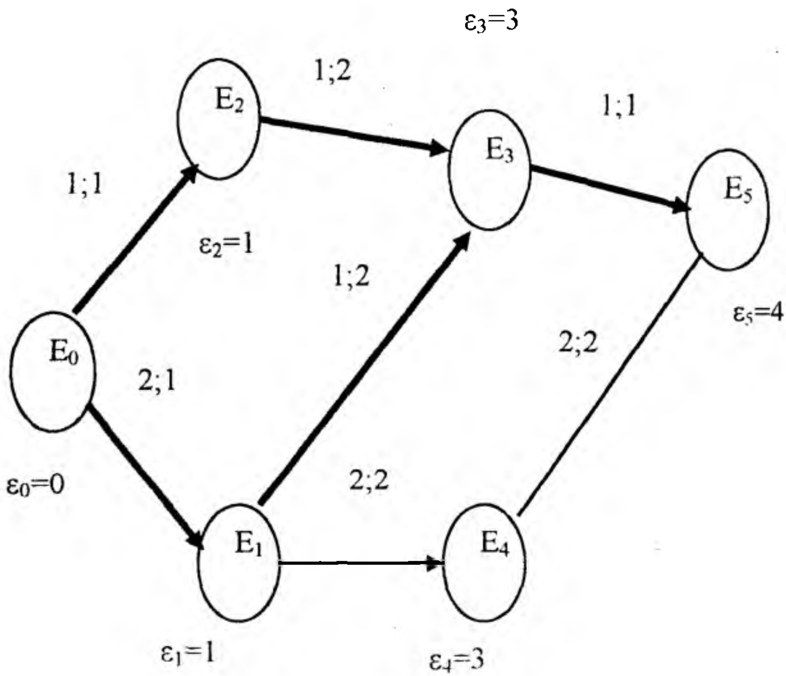


Рис. 2

Четвертый шаг. Величины потоков по дугам пути μ_1 равны: $x_{01}=x_{13}=x_{35}=0+1=1$. Эти значения на рис. 3 приписаны дугам сети в скобках.

Пропускная способность дуг, симметричных дугам пути, равны: $b_{10}=x_{01}=1$; $b_{31}=x_{13}=1$; $b_{53}=x_{35}=1$.

Пятый шаг. Определяем модифицированные стоимости:

$c^*_{01}=1$, так как $x_{01}=1<b_{01}=2$ и $c^*_{10}=-1$;

$c^*_{13} = \infty$, так как $x_{13}=b_{13}=1$ и $c^*_{31} = -2$;

$c^*_{35} = \infty$, так как $x_{35}=b_{35}=1$ $c^*_{53} = -1$;

все остальные $c^*_{ij} = c_{ij}$, так как для них $x_{ij} = 0 < b_{ij} \neq 0$.

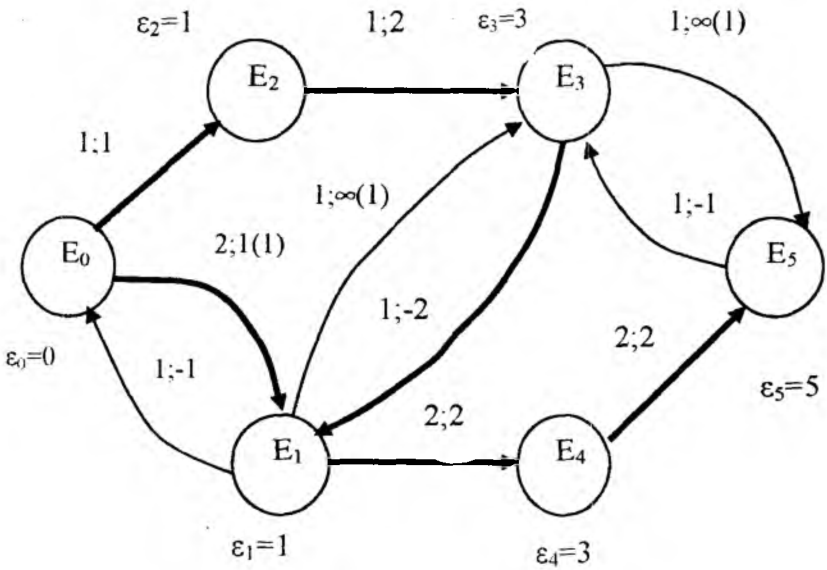


Рис. 3

Итерация 2. Второй шаг. Находим пути минимальной стоимости $\mu_3=(E_0-E_1-E_4-E_5)$ и $\mu_4=(E_0-E_2-E_3-E_5-E_1-E_4-E_5)$.

Третий шаг. Пропускные способности третьего пути $\theta_3=\min(1,\ 2,\ 2)=1$, четвертого пути $\theta_4=\min(1,\ 1,\ 1,\ 2,\ 2)=1$. В этом случае поток пропускаем сразу по обоим путям, так как пропускные способности общих дуг

(E_1, E_4) и (E_4, E_5) равны 2 ед. Сумма $\theta_1+\theta_3+\theta_3+\theta_4=3$ и равна заданной величине потока. Следовательно, задача решена.

Шестой шаг. Оптимальная стоимость потока

$$f = \sum_{(i,j)\in e} c_{ij}x_{ij} = c_{01}x_{01} + c_{02}x_{02} + c_{14}x_{14} +$$

$$+ c_{23}x_{23} + c_{35}x_{35} + c_{45}x_{45} = 14.$$

Фактическое движение потока изображено на рис. 4. На этом рисунке числа у дуг равны дуговым потокам ($x_{01}=2$, $x_{02}=1$, $x_{14}=2$, $x_{23}=1$, $x_{35}=1$, $x_{45}=2$, все остальные неизвестные равны нулю). Поток по дугам (E_1, E_3) и (E_3, E_1) «погасился».

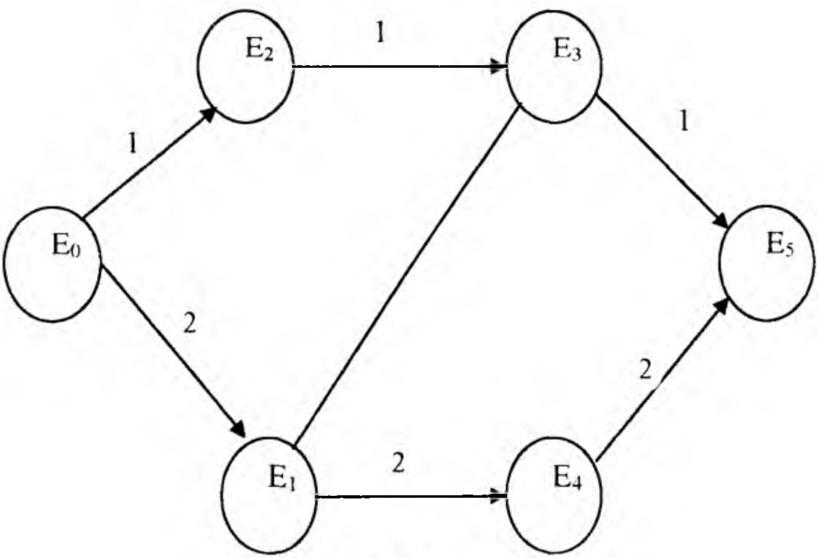


Рис. 4

Примечание: Пропускная способность пути $\mu_3=(E_0-E_1-E_4-E_5)$, для которого $\theta_3=\min(1, 2, 2)=1$. (1) – пропускная способность дуги (E_0, E_1) получилась вследствие того, что ранее по этому пути ($E_0-E_1-E_3-E_5$) был пропущен поток, равный 1, что указано в скобках на этих дугах. Т.к. пропускная способность дуги (E_0, E_1) ранее была равна 2, то теперь, следовательно она уменьшается на 1. Пропускные способности по остальным дугам этого пути будут равны 0. Поэтому путь μ_3 изменяется.

Проблемные задания

1. Сформулируйте конкретную задачу о потоке минимальной стоимости и дайте ее математическое описание с учетом ограничений на пропускные способности.
2. Исследуйте алгоритм Басакера – Гоуэна для задачи задания 1.
3. Решите задачу задания 1 и оцените ее результат.

Вопросы для самоподготовки

1. Напишите математическую модель задачи о потоке минимальной стоимости.
2. Что такое минимальная стоимость?
3. Напишите формулу определения модифицированных дуговых стоимостей.
4. Напишите формулу определения минимальной стоимости потока заданной величины.
5. Напишите формулу определения пропускной способности пути μ.

ГЛАВА V.

§ 17. ЗАДАЧИ ТЕОРИИ РАСПИСАНИЙ

Задачи планирования и управления, расписание,система ресурсов, календарное планирование производства, материальные ре-сурсы, задачи упорядочения, директивное время, штраф за ожидание, стоимость выполнения, задержка работы, критерии эффективности

17.1. Общие сведения о задачах теории расписаний

Многие задачи планирования и управления требуют упорядочения во времени фиксированной системы ресурсов для выполнения определённой совокупности работ. В теории сетевого планирования основное внимание уделялось распределению времени материальных ресурсов при выполнении заданного комплекса работ. Развитие календарного планирования производства вызвало к жизни теорию расписаний.

Теория расписаний является разделом исследования операций и занимается изучением специфических моделей. Эти модели широко распространены в оперативно-календарном планировании, в производстве, в учебном процессе и т.п. В теории расписаний исследуются вопросы, связанные с разработкой математических методов построения наилучших (в смысле заданного критерия эффективности) календарных планов.

Среди задач теории расписаний особое место занимают задачи упорядочения. В качестве примера задачи упорядочения рассмотрим одну из задач оперативно-календарного планирования работы производственного участка, обеспечивающего выпуск некоторого количества деталей различных типов. Для каждого типа деталей предполагаются известными технологическая последовательность обработки деталей на станках и время обработки каждой детали на каждом из станков.

Требуется принять решения, направленные на эффективную организацию работы участка, т.е. определить такой порядок запуска деталей в производство, при котором общее время пребывания их на обработке было-бы минимальным.

В теории расписаний есть ряд задач, которые имеют большое практическое значение и сравнительно просто решаются. К их числу может быть отнесена задача с одним обслуживающим устройством.

Пусть для выполнения на одной машине одновременно поступает множество $N=\{1, 2, \ldots, n\}$ работ. Также предположим, что предположительность выполнения каждой работы на машине известна. Обозначим ее через t_i, $i \in N$.

Задача, связанная с построением расписания для выполнения множества работ, при котором некоторый заданный критерий эффективности будет принимать оптимальное значение, является конечной целью.

Метод решения задачи существенно зависит от критерия эффективности.

Введём следующие обозначения:

$\underline{t}_i$ – время начало выполнения работы $i \in N$;

$\overline{t}_i$ – время окончания выполнения работы $i \in N$;

d_i – директивное время, в течение которого должно быть завершено выполнение работы $i \in N$;

α_i – штраф за ожидание работы $i \in N$ в единицу времени до начала её обработки;

β_i - стоимость выполнения работы $i \in N$.

Время T, необходимое для выполнения всех работ множества N, не зависит от порядка выполнения работ и равно сумме времён выполнения всех работ: $T = \sum_{i\in N} t_i$.

Учитывая принятые обозначения, можно записать что задержка z_i, работы i составит $z_i = \max(0, \overline{t_i} - d_i)$, $i \in N$.

Критерий эффективности, определяющий величину суммарных издержек, связанных с опозданием, будет иметь вид $\Phi_1 = \sum_{i\in N} \alpha_i z_i$.

Критерий, позволяющий вычислять максимальный штраф, связанный также с опозданием выполнения работ, имеет вид $\Phi_2 = \max_{i\in N} \alpha_i z_i$.

Часто, встречается критерий $\Phi_3 = \sum_{i\in N} \alpha_i \underline{t_i}$, представляющий сумму штрафов, связанных с ожиданием работ.

Приведём алгоритм построения оптимального расписания по критерию Φ_2.

Предварительный шаг. Вычислим T и перейдем к первому шагу алгоритма.

Первый шаг. Среди всех неупорядоченных работ находим такую (s), для которой $\alpha_s z_s = \min \alpha_i z_i$, где $z_i = \max(0, T - d_s)$.

Переходим ко второму шагу.

Второй шаг. Работу с номером s выполняем среди рассматриваемого множества. Исключаем работу s из рассмотрения. Если множество работ пусто, то задача решена. Иначе, заменяем T на $T - t_s$, и переходим к первому шагу.

17.2. Задача теории расписании с двумя последовательными обслуживающими устройствами

Одной из простых задач упорядочения является задача Джонсона. Имеется множество $N=\{1,2,...,n\}$ работ, которые должны быть выполнены на m машинах. Время выполнения работы i на машине j обозначим через t_{ij}, предполагая его заранее известным. Задача построения расписания состоит в указании порядка, в котором должны выполняться работы, чтобы суммарное время простая всех машин было минимальным.

Рассмотрим случай, когда число машин равно двум. Каждая работа состоит из двух операций, которая выполняется сначала на первой машине, затем на второй. Время выполнения работы j на первой, затем на второй машине равно t_{1j} и t_{2j} соответственно, где $j \in N$.

Приведём алгоритм построения оптимального расписания, который называется алгоритмом Джонсона.

Предварительный шаг. Записываем матрицу $T=\begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \end{pmatrix}$ времен выполнения операций. Переходим к первому шагу.

Первый шаг. Выбираем в матрице T минимальный элемент. Если он находится в первой строке (соответствующей первой машине), то данную работу выполняем первой, если во второй строке – то последней. Переходим ко второму шагу.

Второй шаг. Исключаем из рассмотрения время выполнения операций, относящихся к упорядоченной работе. Если множество элементов матрицы T уже пусто, то задача решена. Если нет, то переходим к первому шагу.

17.3. Пример

Имеется пять работ. Каждая работа состоит из двух операций, которые выполняются сначала на первой, затем на второй машинах. Время выполнения операций задано в таблице 1:

Таблица 1

M \ P	P_1	P_2	P_3	P_4	P_5
M_1	5	3	4	1	2
M_2	2	2	3	2	3

Проведя вычисления в соответствии с алгоритмом Джонсона, находим, что оптимальный порядок выполнения работ следующий: P_4, P_5, P_3, P_1, P_2.

Проблемные задания

1. Сформулируйте задачу упорядочения работ производственного участка, осуществляющего выпуск каких-либо деталей.
2. Дайте математическое описание сформулированной задачи.
3. Исследуйте возможные критерии эффективности принятия правильного решения.
4. Составьте упорядочение конечного множества работ для двух обслуживающих механизмов.

5. Используя алгоритм Джонсона, решите задачу пункта 4.

6. Используя график Ганта, оцените результат упорядочения работ для задачи пункта 5.

Вопросы для самоподготовки

1. Сформулируйте задачу теории расписаний.
2. Напишите формулу определения величины суммарных издержек, связанных с опозданием времени множества работ.
3. Напишите формулу определения задержки работ.
4. Напишите формулу определения суммарных штрафов.
5. Приведите алгоритм построения оптимального расписания для задачи Джонсона.

§ 18. ЗАДАЧА О НАЗНАЧЕНИЯХ

Работа, механизм, распределение механизмов по работам, производительность механизма, матрица эффективностей, эквивалентные матрицы, предварительные преобразования, венгерский алгоритм решения задачи

18.1. Постановка задачи о назначениях

Задача о назначениях является одной из разновидностей задач целочисленного программирования.

Пусть требуется выполнить *n* различных работ, для выполнения которых имеются *n* механизмов (машин). Здесь каждый механизм может быть использован для любой работы. Производительность каждого механизма на различных работах различна. Обозначим через c_{ij} производительность *i*-го механизма на *j*-ой работе. Задача заключается в таком распределении механизмов по работам, при котором суммарная производительность будет максимальной.

Построим математическую модель этой задачи. Сопоставим каждому из возможных вариантов распределения машин по работам набор значений неизвестных x_{ij}, относительно которых условимся, что $x_{ij} = 1$, если в данном варианте *i*-ый механизм назначается на *j*-ую работу, и $x_{ij} = 0$, если *i*-ый механизм не назначается на *j*-ую работу.

Для любого варианта среди чисел *i*, *j* должно быть точно *n* единиц, причем должны выполняться условия:

$\sum_{j=1}^{n} x_{ij} = 1,\ \ i = \overline{1,n}$, – (каждый механизм назначается на одну работу).

$\sum_{i=1}^{n} x_{ij} = 1, \quad j = \overline{1,n}$, (на каждую работу назначен один механизм).

Суммарная производительность при данном варианте назначения машин на работы выразится суммой $\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$.

Таким образом, математическая модель задачи будет представлена в следующем виде:

максимизировать $\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$

при ограничениях $\sum_{j=1}^{n} x_{ij} = 1, \quad i = \overline{1,n}$, $\sum_{i=1}^{n} x_{ij} = 1, \quad j = \overline{1,n}$,

$$x_{ij} = \begin{cases} 0 \\ 1 \end{cases} \quad i, j = \overline{1,n}.$$

Последние условия выводят задачу о назначениях из класса задач линейного программирования, поскольку они нелинейны, и позволяют отнести задачу о назначениях к задачам целочисленного программирования, точнее, к задачам с булевыми переменными. Практически, задачу о назначениях можно рассматривать как частный случай транспортной задачи. В самом деле, если, отбросить последние условия, заменив их условиями неотрицательности переменных, то задача превращается в транспортную задачу.

Пусть $C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{2n} & \dots & c_{nn} \end{pmatrix}$ матрица эффективностей задачи о назначениях.

В соответствии с постановкой этой задачи, решить ее означает следующее. В матрице эффективностей необходимо выбрать n элементов так, чтобы:

– во-первых, эти элементы выбирались по одному из каждой строки и каждого столбца (строки соответствуют механизмам, а столбцы – работам);

– во-вторых, сумма выбранных элементов, равная общей эффективности, соответствующей данному выбору, была наибольшей по сравнению с ее значениями при всех таких выборах.

В задаче о назначениях две матрицы называются эквивалентными, если одна из них, получается из другой, прибавлением к элементам каждой строки (и/или каждого столбца) одного и того же числа (для разных строк и/или столбцов матрицы эти числа могут быть различными). Это означает, что если для двух матриц эффективностей $C = (c_{ij})$, $i, j = 1,2,...,n$, и $D = (d_{ij})$, $i, j = 1,2,...,n$, выполняется условие $d_{ij} = c_{ij} + e_i + f_j$, $i, j = 1,2,...,n$, где e_i и f_j соответствующие числа, то эти матрицы эквивалентны.

Можно доказать, что множества оптимальных назначений двух задач выбора с эквивалентными матрицами совпадают. На основании этого факта, используя два последовательных преобразования матрицы C, осуществим переход от данной задачи с матрицей C к задаче с матрицей D:

$$C = (c_{ij}) \rightarrow C' = (c'_{ij}) = (\max_k c_{kj} - c_{ij}) \rightarrow D = (d_{ij}) = (c'_{ij} - \min_s c'_{is}).$$

Назовем эти преобразования предварительными.

18.2. Венгерский алгоритм решения задачи

Любая задача о назначениях может быть решена с использованием методов линейного программирования или алгоритма решения транспортной задачи. Заметим, что если сумма нескольких целочисленных неотрицательных переменных равна единице, то каждая из переменных может быть равна нулю, или единице. Поэтому, если решить транспортную задачу методом потенциалов или любым другим методом, который при целых правых частях ограничений приводит к целочисленному оптимальному решению, то полученное решение автоматически будет удовлетворять не учтенному условию булевости переменных.

Хотя для транспортной задачи есть методы, которые проще методов решения общей задачи линейного программирования, особенности рассматриваемой задачи позволяют решить ее с помощью еще более простых приемов. Поэтому, ввиду особой структуры данной задачи был разработан специальный алгоритм, получивший название венгерского метода. Венгерский метод решения этой задачи был впервые предложен американским математиком Х. Куном. Кун использовал теорию паросочетаний, известную ему по работам венгерских ученых Д. Кенига и Э. Эгервари, и поэтому назвал свой алгоритм венгерским методом. Этот алгоритм состоит из трех основных шагов.

Приведем венгерский алгоритм решения задачи о назначениях.

Пусть уже проделаны предварительные преобразования матрицы эффективности C данной задачи и получена неотрицательная матрица D, содержащая хотя бы по одному нулевому элементу в каждой строке и в каждом столбце:

1. а) Пометим звездочкой первый нуль, найденный при просмотре первого столбца матрицы D: (0^*); затем пометим звездочкой нуль во втором столбце, не лежащий в той строке, в которой находится 0^* из первого

столбца (если такой нуль во втором столбце найдется); далее пометим звездочкой нуль третьего столбца, лежащий в строке, где нет еще нулей со звездочкой (если такой нуль в третьем столбце найдется); и так далее пока не просмотрим все столбцы матрицы.

б) Возможны два случая:

1) число помеченных звездочкой нулей равно n;

2) число нулей со звездочкой меньше n.

В случае 1) процесс окончен: места, занимаемые нулями со звездочкой, соответствуют n переменным x_{ij}, равным 1 в оптимальном решении исходной задачи.

В случае 2) переходим к п.2 алгоритма.

2. а) Пометим знаком «+» столбцы матрицы, в которых есть 0* и будем считать эти столбцы занятыми.

В ходе процесса будут появляться и занятые строки. Элементы, расположенные на пересечении не занятого столбца и не занятой строки будем считать не занятыми, а остальные элементы – занятыми.

б) Если в матрице нет незанятых нулей, то переходим к п.5 алгоритма.

в) Если не занятые нули есть, то, просматривая поочередно элементы строки матрицы слева направо, выбираем первый из них. Отмечаем его штрихом (0'). Если в его строке нет нуля со звездочкой, то переходим к п.4 алгоритма; если в его строке 0* есть, то переходим к п.3 алгоритма.

3. Освободим, т.е. снимем знак «+» и будем считать снова не занятым столбец, в котором находится 0*, лежащий в той же строке, что и отмеченный только что штрихом нуль.

Пометим знаком «+» строку, в которой находится 0', и будем считать ее занятой. Переходим к подпункту б) пункта 2 алгоритма.

4. Начиная с только что отмеченного 0', построим цепочку из нулей. Для этого начнем движение от этого 0' по столбцу к 0* и от него по строке к 0' и т.д., пока это возможно. Цепочка оборвется (возможно, на первом

же 0′) на некотором 0′. Снимем звездочки у нулей из цепочки и заменим звездочками штрихи у нулей в цепочке. Новый набор нулей со звездочками содержит на один больше чем предыдущий. Снимем все метки, кроме звездочек, и переходим к подпункту б) пункта 1 алгоритма.

5. Отыщем минимальный элемент среди незанятых элементов матрицы и обозначим его через h. Вычтем h из всех элементов незанятых строк и прибавим ко всем элементам занятых столбцов. Никакие метки при этом не снимаются. В результате получим матрицу, эквивалентную предыдущей матрице, и, содержащую незанятые нули. Переходим к подпункту в) пункта 2.

Конец алгоритма.

Ясно, что при каждом выполнении пункта 4 алгоритма, количество 0^* увеличивается на единицу, и за конечное число шагов обязательно придем к оптимальному назначению. Если оптимальных назначений много, то каким будет это назначение, зависит от порядка действий при использовании пунктов алгоритма, где осуществляется поочередный просмотр элементов матрицы.

18.3. Пример

Пусть задана матрица эффективностей

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

задачи о назначениях. Для нахождения оптимального варианта назначений сначала выполним предварительные преобразования. В каждом из столбцов исходной матрицы C, найдем максимальный элемент, из него вычтем каждый элемент этого столбца, результаты запишем на соответствующих местах и получим матрицу:

$$C' = \begin{pmatrix} 2 & 3 & 1 & 3 & 0 \\ 2 & 0 & 5 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 7 & 2 & 4 & 3 & 3 \\ 0 & 3 & 1 & 0 & 2 \end{pmatrix}.$$

Далее, в каждой строке матрицы C' находим минимальный элемент. Вычитая его из каждого элемента этой строки, результаты запишем на соответствующих местах. Теперь получим матрицу D, эквивалентную исходной матрице C:

$$D = \begin{pmatrix} 2 & 3 & 1 & 3 & 0 \\ 2 & 0 & 5 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 5 & 0 & 2 & 1 & 1 \\ 0 & 3 & 1 & 0 & 2 \end{pmatrix}.$$

Используя венгерский алгоритм, решим задачу о назначениях. Результаты приведем в виде цепочки матриц, где проставлены соответствующие метки. Снятие какой-то метки отмечено заключением ее в прямоугольник.

$$
\begin{array}{ccccc}
\boxed{\cdot} & + & + & & + \\
\end{array}
\qquad\qquad\qquad
\begin{array}{ccccc}
 & + & \boxed{+} & & + \\
\end{array}
$$

$$
\begin{pmatrix}
2 & 3 & 1 & 3 & 0^* \\
2 & 0^* & 5 & 2 & 1 \\
2 & 1 & 0^* & 1 & 2 \\
5 & 0 & 2 & 1 & 1 \\
0^{\boxed{*}} & 3 & 1 & 0' & 2
\end{pmatrix}
\begin{array}{c} \\ \\ \\ \\ + \end{array}
\xrightarrow{h=1}
\begin{pmatrix}
1 & 3 & 1 & 2 & 0^* \\
1 & 0^* & 5 & 1 & 1 \\
1 & 1 & 0^* & 0 & 2 \\
4 & 0 & 2 & 0^{\boxed{\prime}*} & 1 \\
0^* & 4 & 2 & 0' & 3
\end{pmatrix}
\begin{array}{c} \\ \\ + \\ \\ + \end{array}
$$

Окончательно получим матрицу, где число нулей со звездочками (0^*) равно 5:

$$
\begin{pmatrix}
1 & 3 & 1 & 2 & 0^* \\
1 & 0^* & 5 & 1 & 1 \\
1 & 1 & 0^* & 0 & 2 \\
4 & 0 & 2 & 0^* & 1 \\
0^* & 4 & 2 & 0 & 3
\end{pmatrix}
$$

Значит, оптимальным назначением будет следующий план: $x_{15} = x_{22} = x_{33} = x_{44} = x_{51} = 1$, остальные $x_{ij} = 0$. Таким образом, если первый механизм назначается на пятую, второй – на вторую, третий – на третью, четвертый – на четвертую и пятый – на первую работы, то полученная суммарная эффективность будет наибольшей и равной 32 единицам.

Проблемные задания

Для следующих задач составьте математическую модель, матрицу эффективностей и выполните предварительные преобразования:

1. Предположим, что некоторое авиапредприятие выполняет авиарейсы по четырем маршрутам. Для обслуживания этих рейсов предприятие располагает четырьмя типами самолетов: ТУ-154, ИЛ-86, АН-24, ЯК-40. Каждый самолет может обслуживать любой из этих маршрутов. Прибыли, получаемые авиапредприятием, от обслуживания этих маршрутов (в условных денежных единицах), приведены в таблице 1. Требуется запланировать и так назначить каждый самолет на каждый рейс, чтобы суммарная прибыль авиапредприятия была максимальной.

Таблица 1

	Авиарейсы			
	1	2	3	4
ТУ-154	45	32	10	23
ИЛ-86	90	23	39	38
АН-24	20	12	10	6
ЯК-40	9	11	10	20

2. Пусть имеется 6 работников и 6 должностей. Каждому назначению i-го работника на j-ю должность сопоставим число c_{ij}, $i,j=1,2,...,6$, которое определяется согласно формуле $c_{ij}=ij+(7-i)i-j(j-8)$ и определяет эффективность данного назначения. Требуется так назначить каждого i-го работника на каждую j-ю должность, чтобы эффективность была максимальной.

3. Сформулируйте какую-нибудь задачу о назначениях с матрицей эффективностей 6×6.

4. Исследуйте все особенности венгерского алгоритма (если это возможно) для задачи задания 3.

5. Решите задачу задания 3 и оцените полученный результат.

Вопросы для самоподготовки

1. Сформулируйте задачу о назначениях.
2. Напишите математическую модель задачи о назначениях.
3. Какие матрицы называются эквивалентными в задаче о назначениях?
4. Приведите формулу предварительных преобразований.
5. Приведите последовательности венгерского алгоритма решения задачи о назначениях.

§ 19. МАРКОВСКИЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

Операция, случайные неконтролируемые факторы, случайный процесс, состояние системы, вероятности переходов, однородные и неоднородные марковские цепи, стохастическая матрица

19.1. Основные понятия марковского случайного процесса

При исследовании различных операций, с точки зрения выбора оптимального решения, приходится сталкиваться с ситуациями, когда обстановка поведения операции характеризуется случайными неконтролируемыми факторами. В этом случае можно говорить, что операция развивается по схеме случайных процессов, ход протекания которых и исход зависят от сопровождающих операций, случайных факторов.

Пусть имеется некоторая операция S, которую в дальнейшем будем называть системой, развивающуюся во времени и изменяющую свои состояния.

Состояние системы обозначим через $E_1, E_2, ..., E_i, ...$ Если число состояний системы S конечно, то она называется конечной системой.

Рассмотрим так называемый марковский случайный процесс, или «процесс без последствия» (или марковская цепь). Основные понятия и определения, связанные с марковскими цепями, были сформулированы и изучены русским математиком А.А.Марковым (1856–1922).

Характерной особенностью марковского процесса является следующее: для каждого момента времени t_0, вероятность пребывания системы в любом из состояний в момент t_1 ($t_1 > t_0$) зависит от того, в каком состоянии система находилась в момент t_0, и не зависит от того, когда и каким образом система пришла в это состояние (т.е. не зависит от развития процесса до момента t_0).

Марковский процесс является конечным, если число состояний системы конечно, в противном случае – бесконечным. Рассмотрим марковский процесс с конечным числом состояний $E_1, E_2, \ldots, E_n$.

Обозначим через $t_0, t_1, \ldots, t_m, \ldots$ ($t_0 < t_1 < \ldots < t_m < \ldots$) моменты перехода.

Для марковского процесса вероятность состояния системы S в момент t_m зависит лишь от состояния системы в момент t_{m-1}, т.е. имеем дело с условной вероятностью $p[j, t_m / i, t_{m-1}]$. Такие вероятности назовем переходными.

Если вероятности переходов зависят не от времени, когда осуществляется переход, а лишь от состояний E_i и E_j, то соответствующая марковская цепь называется однородной, в противном случае – неоднородной.

В дальнейшем рассмотрим однородные цепи Маркова. В случае однородных марковских цепей вероятность переходов из E_i, в E_j за один переход обозначим через p_{ij}.

Матрицу $p = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$ называют матрицей

вероятностей переходов. Для этой матрицы $0 \le p_{ij} \le 1$, $i, j = 1, 2, \ldots, n$; $\sum^{n} p_{ij} = 1$, $i = 1, 2, \ldots, n$. Матрица, обладающая этими свойствами, называется стохастической.

19.2. Вероятности состояний

Обозначим через $p_i(m)$ безусловную вероятность того, что система в момент времени m находится в состоянии E_i, $i=1,2,...,n$. Тогда, совокупность вероятностей $p_i(m)$, $i=1,2,...,n$ будет образовывать стохастический вектор $\overline{p}(m)$ состояний системы S:

$$\overline{p}(m) = (p_1(m), p_2(m),..., p_n(m)),$$

$$0 \le p_i(m) \le 1, \ i = 1,2,...,n, \ \sum_{i=1}^{n} p_i(m) = 1.$$

Рассмотрим задачу вычисления вероятности перехода из E_i в E_j не за один переход, а за m переходов. Обозначим эту вероятность через $p_{ij}^{(m)}$, а матрицу этих вероятностей — $P^{(m)}$. Тогда процесс перехода за m шагов может быть представлен в виде следующих двух этапов:

- сначала переход за k шагов ($1 \le k \le m$),
- затем за оставшиеся $m-k$ шагов (рис. 1).

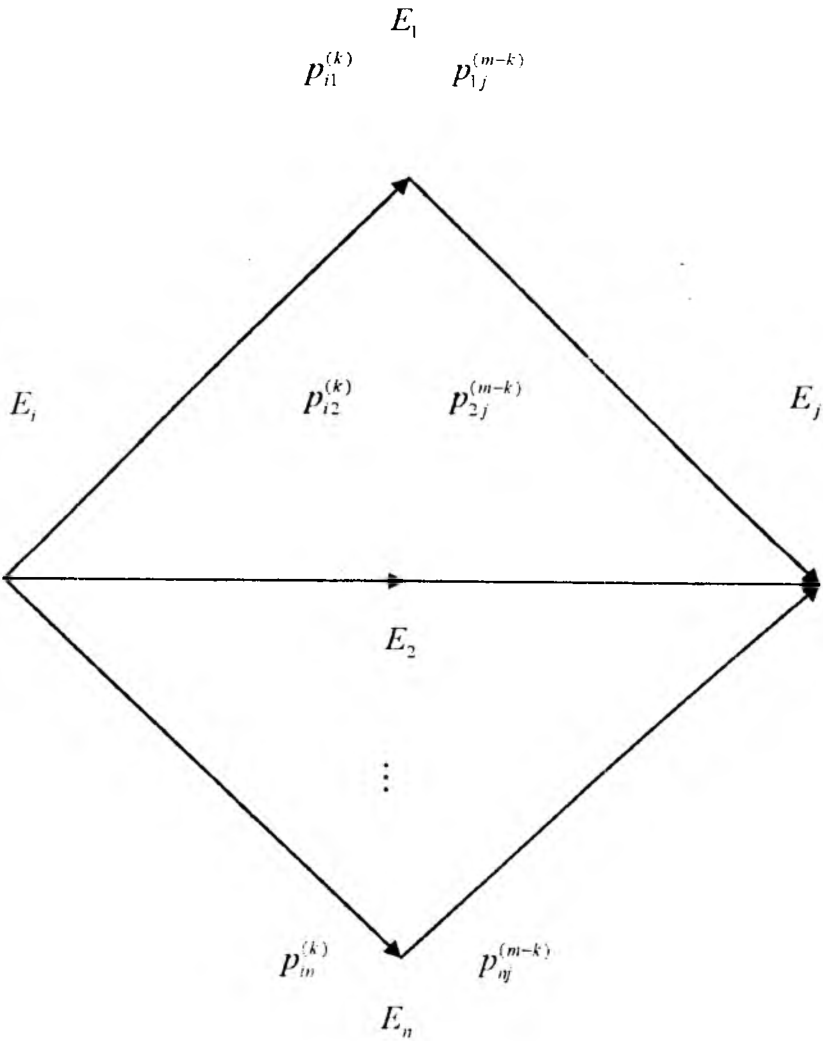


Рис. 1

Вероятность перехода по первому пути $E_i \to E_1 \to E_j$ будет равна $p_{i1}^{(k)} p_{1j}^{(m-k)}$, вероятность перехода по второму пути $E_i \to E_2 \to E_j$ будет равна $p_{i2}^{(k)} p_{2j}^{(m-k)}$, и т.д.

Таким образом,

$$p_{ij}^{(m)} = p_{i1}^{(k)} p_{1j}^{(m-k)} + p_{i2}^{(k)} p_{2j}^{(m-k)} + \ldots + p_{in}^{(k)} p_{nj}^{(m-k)} + \ldots$$

Используя матричное обозначение $P^{(m)} = P^{(k)} P^{(m-k)}$, заметим, что $P^{(1)} = P$. Поэтому, $P^{(m)} = P^m$. Для получения безусловных вероятностей $p_i^{(m)}$ при любом m, необходимо знать начальный вектор вероятностей состояний:

$$\overline{p}(0) = (p_1(0), p_2(0), \ldots, p_n(0)).$$

Теперь, используя формулу полной вероятности, можем записать

$$p_j(1) = \sum_{i=1}^{n} p_i(0) p_{ij}, \quad j = 1,2,\ldots,n, \text{ или } \overline{p}(1) = \overline{p}(0)P.$$

Аналогично для любого m:

$$\overline{p}(m) = \overline{p}(m-1)P = \overline{p}(0)P^m.$$

Рассмотрим эргодическую марковскую цепь. Эргодической называется такая цепь Маркова, для которой любое состояние E_i может быть достигнуто из любого состояния E_j за конечное число шагов.

Известно, что если система S обладает эргодическим свойством, то она является стохастически устойчивой.

Это означает, что предельные вероятности состояний $p_i(m)$ при $m \to \infty$, $i = 1,2,...,n$, не зависят от вектора начальных состояний. Если обозначить через p_i предельные вероятности состояний $p_i(m)$ при $m \to \infty$, $i = 1,2,...,n$, то их можно выразить как $\lim_{m \to \infty} p_i(m) = p_i(\infty) = p_i$.

Используя равенство $\overline{p}(m) = \overline{p}(m-1)P$ при $m \to \infty$, получим $\overline{p} = \overline{p}P$ или $\overline{p}(P - E) = 0$, где E – единичная матрица, а 0 – вектор соответствующих размерностей. Последнее равенство и $\sum_{i=1}^{n} p_i = 1$ образуют систему линейных уравнений относительно p_i, $i = 1,2,...,n$. Решая эту систему, найдем вектор $\overline{p}$.

19.3. Пример

Предприятие, в зависимости от потребности населения в изготавливаемой продукции, в конце работы, может оказаться в одном из двух состояний: E_1 – есть потребность, E_2 – нет потребности. Пусть $P = \begin{pmatrix} 4/7 & 3/7 \\ 5/7 & 2/7 \end{pmatrix}$ является матрицей вероятностей переходов для состояний предприятия (рис. 2).

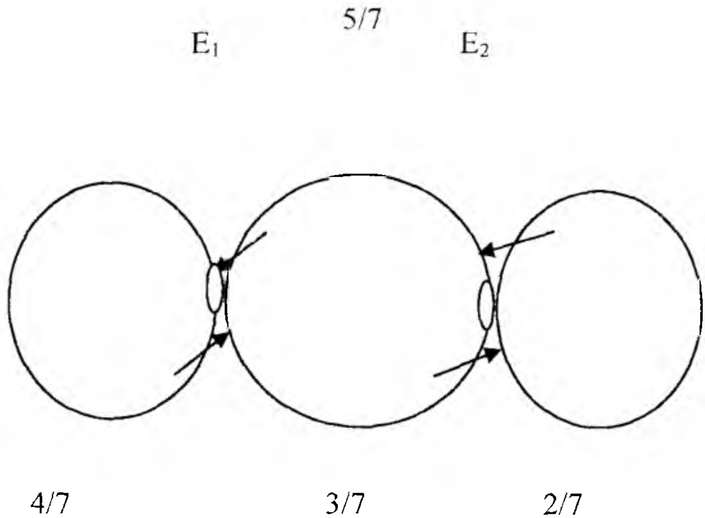


Рис. 2

Если в начальный момент времени предприятие находится в состоянии E_1, то $p(0)=(1,0)$. Вычислим вероятности пребывания предприятия в каждом из состояний в конце первого года:

$$p_1(1)=p_1(0)p_{11}+p_2(0)p_{21}=1\cdot\frac{4}{7}+0\cdot\frac{5}{7}=\frac{4}{7},$$

$$p_2(1)=p_1(0)p_{12}+p_2(0)p_{22}=1\cdot\frac{3}{7}+0\cdot\frac{2}{7}=\frac{3}{7}.$$

Итак, $p(1)=\left(\frac{4}{7},\frac{3}{7}\right)$. В конце второго года:

$$p_1(2)=p_1(1)p_{11}+p_2(1)p_{21}=\frac{4}{7}\cdot\frac{4}{7}+\frac{3}{7}\cdot\frac{5}{7}=\frac{31}{49},$$
$$p_2(2)=p_1(1)p_{12}+p_2(1)p_{22}=\frac{4}{7}\cdot\frac{3}{7}+\frac{3}{7}\cdot\frac{2}{7}=\frac{18}{49}.$$

Следовательно, $p(2)=\left(\frac{31}{49},\frac{18}{49}\right)$. Продолжая вычисления таким образом, определим вероятности пребывания предприятия в каждом из состояний в конце любого года. Для определения вероятностей предельного состояния предприятия $p=(p_1,p_2)$ решим следующую систему

$$p_1=p_1\cdot\frac{4}{7}+p_2\cdot\frac{5}{7},$$
$$p_2=p_1\cdot\frac{3}{7}+p_2\cdot\frac{2}{7},$$
$$p_1+p_2=1.$$

В результате получим $p_1=\frac{5}{8}$, $p_2=\frac{3}{8}$. Итак, вероятности предельного состояния предприятия приближаются, соответственно, к $\frac{5}{8}$ и $\frac{3}{8}$.

Проблемные задания

1. Сформулируйте задачу, в которой обстановка проведения операции характеризуется случайными неконтролируемыми факторами.

2. Исследуйте для задачи задания 1 характерные особенности марковского процесса.

3. Вычислите вероятности состояний системы задачи задания 1 и оцените полученный результат.

4. Фирма "Кожизделия" открывает новое производство дамских сумок. Все множество состояний фирмы можно разделить условно на 2 основных состояния:

– модели, которые она выпускает, находят спрос у покупателя;

– эти модели не находят спроса.

Матрица переходных вероятностей системы имеет вид

$$P = \begin{pmatrix} 0{,}25 & 0{,}75 \\ 0{,}6 & 0{,}4 \end{pmatrix}.$$

Определите вероятности состояний фирмы для двух начальных условий:

а) фирма начинает производство с удачной модели сумок;

б) фирма начинает производство с неудачной модели сумок.

Вопросы для самоподготовки

1. Дайте основные понятия и определения марковского процесса.

2. В чем заключается характерная особенность марковского процесса?

3. Какие вероятности называются переходными?

4. Какая матрица называется стохастической?

5. Дайте определение эргодической марковской цепи.

6. Напишите систему линейных уравнений, определяющих вероятности состояний.

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

1. Акулич И.Л. Математическое программирование в примерах и задачах. М.: «Высш. шк.»,1993. 336-с.

2. Афанасьев М. Ю., Суворов Б. П. Исследование операций в экономике: модели, задачи, решения. М.: «Инфра – М.»- 2003.

3. Бадалов Ф. Б. Оптималлаш назарияси ва математик программалаштириш. Т.: «Ўқитувчи», 1989.

4. Вагнер Г. Основы исследования операций.: В 3-х томах. – М.: «Мир», 1972-1973.

5. Волков И. К., Загоруйко Е. А. Исследование операций. М.:Изд. МГТУ им. Н. Э. Баумана.- 2000.

6. Габасов Р., Кириллова Ф.М. Методы оптимизации. Изд. БГУ. 1981.

7. Гольштейн Е.Г., Третьяков И.В. Модифицированные функции Лагранжа: Теория и методы оптимизации. М.: «Наука», 1989. 400-с.

8. Горелик В.А., Ушаков И.А. Исследование операций. М.: «Машиностроение», 1986. 286-с.

9. Давыдов Э.Г. Исследование операций. – М.: «Высш. шк.» 1990.

10. Евтушенко Ю.Г., Мазурик В.П. Программное обеспечение систем оптимизации. М.: «Знание», 1989. 48-с.

11. Зайченко Ю.П. Исследование операций. – Киев. «Вища школа», 1988.

12. Исследование операций. В 2-х томах. Под редакцией Дж.Моудера и С.Эльмаграби. – М.: «Мир», 1981.

13. Калихман В.Г. Сборник задач по математическому программированию. –М.: «Высш. шк.», 1975.

14. Карманов В.Г. Математическое программирование. М.: «Наука», 1986.

15. Конюховский П. В. Математические методы исследования операций. Ст. Петербург. Изд. «Питер».- 2001.

16. Косоруков О.А., Мищенко А.В. Исследование операций. М.: Издательство «Экзамен», 2003.

17. Костевич Л.С. Математическое программирование. Мн.: «Новое знание», 2003.

18. Костевич Л.С., Лапко А.А. Теория игр. Исследование опреаций. – Мн.: «Вышэйшая школа», 1982.

19. Лященко И.Н. и др. Линейное и нелинейное программирование, Киев, «Вища школа», 1975.

20. Мину М. Математическое программирование: Теория и алгоритмы. М.: «Наука», 1990. 488-с.

21. Нурминский Е.А. Численные методы выпуклой оптимизации. М.: «Наука», 1991. 167-с.

22. Сафаева Š., Бекназарова Н. Операцияларни текширишнинг математик усуллари. I қисм, Т.: «Ўқитувчи», 1984. II қисм, Т.: «Ўқитувчи», 1990.

23. Таха Х. Введение в исследование операций. – М.: «Мир», 1985.(в 2-х книгах).

24. Фролькис В.А. Введение в теорию и методы оптимизации для экономистов. 2-е издание С.-Пт.: Издательство: «Питер», 2002.

25. Шмырев В.И. Введение в математическое программирование. М.: Издательство: «Институт компьютерных исследований», 2002.

26. http://iasa.org.ua/iso.php?lang=rus

27. http://ask.cemi.rssi.ru/

28. http://fmi.asf.ru/vavilov/

29. www.xodjayev-t-azizov-i.narod.ru

ХОДЖАЕВ Т., АЗИЗОВ И., ОТАКУЛОВ С.

Исследование операции
(Учебное пособие)

Ташкент - Изд-во «Aloqachi» - 2007

Редактор	**М. МИРКОМИЛОВ**
Тех. редактор	**А. МОЙДИНОВ**
Корректор	**К. АВЕСБОЕВ**
Компьютерная верстка	**Ш. МИРКОСИМОВА**

Разрешено в печать 10. 07.2007. Формат 60x84 $^1/_{16}$.
Гарнитура «Times New Romans». Печать офсетная.
Условн. печ. лист 11,5. Издат. печ. лист 11,0.
Тираж 1000. Заказ № 2.

Отпечатано в типографии «Aloqachi matbaa markazi».
700000, г. Ташкент, ул. А.Тимура, 108.