С.С.РЕЗНИЧЕНКО
М.П.ПОДОЛЬСКИЙ
А.А.АШИХМИН

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИРОВАНИЕ В ПЛАНИРОВАНИИ И УПРАВЛЕНИИ ГОРНЫМ ПРОИЗВОДСТВОМ

101953

Допущено Государственным комитетом СССР по народному образованию в качестве учебника для студентов вузов, обучающихся по специальности "Экономика и управление в отраслях горной промышленности и геологоразведке"

Наукова бібліотека
Київського
ордена Леніна
політехнічного інституту

МОСКВА "НЕДРА" 1991

ББК 65.050.9(2)
Р 34
УДК 658.012.2.001.57:622.012(075.8)

Рецензенты:
кафедра управления производством и анализа хозяйственной деятельности промышленных предприятий Донецкого политехнического института
д-р техн. наук И.Б. Табакман

Р $\frac{2501010000-057}{043(01)-91}$ 82–91

©С.С. Резниченко, М.П. Подольский, А.А. Ашихмин, 1991

ISBN 5-247-01000-0

ПРЕДИСЛОВИЕ

Современные шахты, рудники, карьеры, обогатительные фабрики представляют собой сложные комплексные предприятия, оснащенные мощной горной техникой. Планирование и управление технологическими процессами, горными предприятиями и объединениями требуют от руководителя любого ранга умения быстро и правильно принимать различные решения. При этом его функции все более усложняются из-за роста объемов производства, ухудшения горно-геологических условий, дальнейшего развития техники, повышения требований к максимальному использованию недр и охране окружающей среды.

В этих условиях решения, принимаемые на основе личного опыта и инженерной интуиции, зачастую становятся малоэффективными, так как не учитывают целого ряда противодействующих факторов. Современное производство отличается не только размерами и сложностью, затрудняющими принятие решения, но и очень высокой капиталоемкостью, резко повышающей ущерб от ошибок в проектировании, планировании и управлении. Такая ситуация характерна для всей современной экономики, а также и науки, что создает предпосылки для разработки новой технологии принятия решений, основанной на количественных оценках вариантов, исключающей или уменьшающей значение субъективных факторов. Появились условия для зарождения новой дисциплины – исследования операций – науки о количественном обосновании решений. Как самостоятельное научное направление оно возникло в середине 30-х годов и начало бурно развиваться с появлением цифровых вычислительных машин в конце 40-х. Первоначально исследование операций было применено в военном деле, а в настоящее время эти методы в большей или меньшей степени используются почти во всех областях деятельности человека.

В основе исследования операций лежат следующие принципы: воспроизводимость результатов, доступность для проверки всех вычислений и допущений, независимость выводов от личности, репутации или частных интересов.

Под операцией понимается любое мероприятие (система действий), объединенное единым замыслом, возглавляемое одним центром и направленное на достижение поставленной цели. Данное понятие операции значительно шире, чем используемое в отдельных отраслях (технике, спорте, медицине, военном деле).

Примеры операций в горном деле следующие: разработка месторождения, составление проекта горного предприятия, выбор оборудования, организация технологических процессов, планирование горных работ,

организация перевозок грузов, мероприятия, направленные на повышение надежности функционирования технологической системы, обеспечение бесперебойности работы обогатительной фабрики и т.д.

Каждая операция планируется заблаговременно и предполагает определенные условия ее проведения. Однако эти условия не всегда точно известны заранее. Они могут меняться вследствие появления непредвиденных, неучтенных или случайных факторов. Например, нельзя заранее точно рассчитать производительность погрузочного оборудования, так как практически невозможно предусмотреть его поломки, простои, а также влияние климатических условий, работы транспорта и т.д.

Основная цель исследования операций – количественное обоснование оптимальных решений, т.е. выбор из возможных ее элементов тех, которые по каким-либо соображениям более выгодны.

Помимо основной цели исследование операций решает ряд вспомогательных задач, например, сравнительную оценку различных вариантов и анализ влияния отдельных факторов (условий проведения операций и элементов решения) на результат.

Принятие решения является компетенцией ответственного лица – мастера, диспетчера, начальника смены (цеха, отдела) и руководителя предприятия. Ответственные лица, основываясь на имеющихся у них данных (в том числе на математических расчетах и исследованиях), выбирают окончательное решение в пределах своей компетенции. Это правило остается в силе и при автоматизированных системах. В этом случае исследование операций не заменяет деятельности руководства, а является необходимым инструментом в условиях современного сложного производства. Применение методов исследования операций позволяет руководителям решать возникающие сложные задачи в короткое время и получать на этой основе эффективные результаты.

Обоснование принимаемых решений с помощью методов исследования операций предполагает изучение и формальное, математическое описание разрешаемой ситуации, поиск параметров операции, обеспечивающих решения, экстремум выбранной системы критериев качества, исследование и оценку полученных результатов.

Каждый этап решения операционной задачи имеет свою методическую основу и составляет самостоятельный раздел исследования. Методы формального и количественного описания оптимизируемой ситуации составляют раздел математического моделирования, методы поиска оптимального решения по сформированной математической модели – математическое программирование. Объединяет эти разделы и позволяет оценить полученные результаты теория выбора и принятия решения.

Настоящий учебник написан на основе курса лекций, читаемого в Московском горном институте.

Цель учебника – научить студентов формализовать задачи планирования, организации и управления горным производством, строить экономико-математические модели, принимать решения на основе полученных оптимальных значений переменных.

Учитывая небольшой объем курса и его практическую направленность в данном учебнике, изложение методов математического программирования ведется в минимальном объеме, необходимом только для понимания сути процесса получения решения. Такой подход представляется оправданным, если учесть, что в практической деятельности горный инженер-экономист должен прежде всего уметь сформулировать задачу, правильно оценить и использовать результаты ее решения, а расчеты могут быть выполнены по стандартным программам или с привлечением специалистов-математиков, обслуживающих АСУ предприятия.

Предисловие, гл. IV–VI, VIII, IX, XI–XIV написаны С.С. Резниченко; гл. I–III – М.П. Подольским; VII, X – А.А. Ашихминым; гл. VIII § 2–4, IX § 8–11 – совместно с А.М. Валуевым; гл. XIV § 3, 5, 6 – совместно с М.А. Кушнир.

Глава I

ВВЕДЕНИЕ В МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

§ 1. НАЗНАЧЕНИЕ И ТИПЫ МОДЕЛЕЙ

В своей деятельности каждый инженер непрерывно сталкивается с необходимостью принятия решений по целому ряду производственных вопросов, затрагивающих технологию и организацию производства, организацию труда и заработную плату отдельных трудящихся, финансы предприятия и т.д. В основе решений может лежать инженерная интуиция, базирующаяся на прошлых аналогичных результатах, эксперимент или количественная оценка возможных вариантов. К сожалению, задачи проектирования, планирования и организации производства почти не допускают экспериментов из-за их продолжительности, высокой стоимости или тяжелых последствий в случае неудач. Интуитивные инженерные решения подвержены субъективным влияниям, что снижает их качество, а в отдельных случаях интуиция не подкрепляется опытом, т.е. сложившаяся ситуация не имеет аналогов в прошлом.

Все это создает необходимость выработать единую методику количественной оценки вариантов принимаемых решений.

С этой целью следует ввести ряд общих понятий, позволяющих описать все многообразие решаемых инженерных задач.

Каждое решение предполагает некоторый конкретный объект реализации, который в дальнейшем будет называться моделируемой системой.

Система – совокупность элементов, находящихся во взаимодействии и образующих единое целостное образование со свойствами, отличными от свойств отдельных элементов.

Структура системы – это описание ее организации, определяющее возможность разделения на отдельные элементы и их отношения.

Состояние системы – определяется значением ряда характеризующих параметров. Например, состояние очистного забоя, как системы, характеризуется его положением в пространстве, уровнем метана на исходящей воздушной струе, нахождением комбайна, крепи, наличием и мощностью породных пропластков и т.д. Чем сложнее система, тем больше параметров необходимо определить для описания ее состояния.

Операция – целенаправленное изменение состояния системы.

Проведение операции предполагает изменение одних характеристик системы – так называемых ее параметров x_1, x_2, x_3, x_4, при учете других – условий операции $a_1, a_2, a_3, \ldots, m$.

Во главе каждой операции стоит руководитель, который принимает

решение относительно ее параметров, обеспечивающих достижение цели при существующих условиях.

Что же представляет собой цель операции? По определению цель – это предвосхищаемый результат, на который направлено действие. Поэтому в цели должны содержаться значения параметров системы, соответствующие моменту ее достижения. Условия достижения цели могут задаваться в виде предельных уровней или конкретных значений параметров системы. Но эти условия могут обеспечиваться различными сочетаниями уровней параметров операции, в результате чего возникает необходимость выбора одного варианта из множества допустимых путей достижения цели. Следовательно, цель операции должна содержать не только желаемые значения состояния системы, но и условия выбора средств достижения конечного результата. Эти условия реализуются введением критерия эффективности операции K, который позволяет свести задачу выбора варианта к следующей схеме – из двух вариантов операции $(x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \ldots, x_n^{(1)})$ и $(x_1^{(2)}, x_2^{(2)}, \ldots, x_n^{(2)})$ выбирается тот, для которого значение критерия эффективности больше (меньше).

Распространяя это правило попарного сравнения на множестве допустимых вариантов проведения операции, можно прийти к следующему условию – достижение цели должно обеспечивать максимум (минимум) критерия эффективности при выполнении наложенных ограничений.

Таким образом цель операции имеет вид

$$Z = \left\{ a_1^{(z)} = A_1, \ldots, a_r^{(z)} = A_r; \right.$$
$$\left. a_{r+1}^{(z)} = A_{r+1}, \ldots, a_m^{(z)} \leqslant A_m; \quad K \to \max \right\}, \qquad \text{(I.1)}$$

где $a_1^{(z)}, a_2^{(z)}, \ldots, a_m^{(z)}$ – контролируемые параметры системы на момент достижения цели; $A_1, A_2, \ldots, A_m$ – условия проведения операции; K – критерий эффективности операции.

Например, цель работы угольной (сланцевой) шахты, как производственной системы, – добыча полезного ископаемого в соответствии с установленным для нее объемом при наилучших показателях производительности труда, себестоимости, рентабельности и обеспечении безопасных и комфортных условий для трудящихся, полноты выемки полезного ископаемого, сохранения и рекультивации поверхности после горных работ, чистоты воздушного пространства и водоемов.

Цель операции формулируется руководителем, который принимает решение о пути ее достижения. Формально принятие решения можно представить в виде задачи – найти оптимальное решение $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$, обеспечивающее max (min) критерия эффективности K на множестве допустимых решений X, т.е.

$$K(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}) = \max K$$

$$(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}) \in X. \qquad \text{(I.2)}$$

Напомним, что множество задается свойством, которым обладают все его элементы и не обладает ни один элемент, ему не принадлежащий.

Множество допустимых решений образуется всеми значениями параметров $(x_1, x_2, \ldots, x_n)$, которые не противоречат свойствам системы и обеспечивают выполнение граничных условий, заданных целью операции. Если установить зависимости между контролируемыми параметрами системы $a_1, a_2, \ldots, a_m$ и параметрами операции $x_1, x_2, \ldots, x_n$, играющими роль элементов решения, то множество допустимых решений будет задаваться решениями некоторой системы уравнений и неравенств

$$X = \begin{cases} f_1(x_1, x_2, \ldots x_n) = A_1, \\ f_r(x_1, x_2, \ldots, x_n) = A_r, \\ f_{r+1}(x_1, x_2, \ldots, x_n) \leqslant A_{r+1}, \\ f_m(x_1, x_2, \ldots, x_n) \leqslant A_m. \end{cases} \qquad \text{(I.3)}$$

Процесс изучения и количественного описания системы с целью решения операционной задачи составляет математическое моделирование.

Модель – специально созданная материальная или формальнологическая система, находящаяся в объективном соответствии с исследуемым объектом, отражающая наиболее существенные его свойства и дающая в процессе изучения информацию о нем. Процесс построения и изучения модели называется **моделированием**.

Практическое значение модель приобретает при условии, что ее изучение имеющимися средствами более доступно, чем изучение самого объекта.

В зависимости от выбора средств моделирования различают физические и абстрактные модели.

При физическом моделировании создают подобие изучаемого объекта из подобных материалов или подбирают аналогично протекающий процесс другой физической природы.

Например, для изучения разубоживания руды при различных режимах выпуска создают модель, в определенных пропорциях повторяющую участок месторождения с очистными выработками, разрушенным рудным телом и вмещающими породами. Роль руды и вмещающих пород играют сыпучие среды с подобными плотностью и гранулометрическими составами.

При решении операционных задач используют абстрактные модели,

к которым относятся словесные описания, графические и математические модели.

Словесные описания используются на начальных стадиях формулирования операционной задачи. Хотя словесные описания не содержат количественных мер и соотношений, они вводят понятийный аппарат и определяют смысл решаемой задачи.

Дальнейшая формализация задачи может потребовать изучения и описания организации системы, для чего используются графические и графоаналитические структурные модели. Эти модели отражают неколичественные отношения между элементами системы.

Количественные отношения между параметрами системы отражают аналитические и эмпирические зависимости. Эти зависимости получаются путем аналитических выводов, экспериментальных наблюдений или в результате имитационного моделирования. Суть его заключается в том, что создается алгоритмическая модель, описывающая поведение системы или ее части и на ЭВМ имитируется функционирование системы в течение некоторого времени. Полученные таким образом результаты используются как экспериментальные наблюдения для построения эмпирических зависимостей. В дальнейшем из полученных аналитических и эмпирических зависимостей составляется оптимизационная модель, используемая для принятия решения.

Приведенная здесь очень укрупненная классификация отражает лишь роль отдельных классов моделей в процессе решения операционной задачи. В дальнейшем эти классы моделей будут разделяться на более мелкие группы по методам моделирования, типам решаемых задач и т.д.

§ 2. СВОЙСТВА ПРОИЗВОДСТВЕННЫХ СИСТЕМ И ВИДЫ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Математическая модель – это система математических выражений, описывающих взаимосвязь характеристик объекта моделирования, анализ которых позволяет получить дополнительную информацию о системе, принять инженерное решение. Вид моделей, математический аппарат их построения и исследования определяются свойствами моделируемой системы и в первую очередь – типом ее параметров.

При моделировании производственных систем выделяют три типа параметров: детерминированные, стохастические и качественные.

Детерминированные параметры имеют количественную меру и сохраняют свое значение при неизменных определяющих условиях. Эти параметры могут быть получены путем измерений и расчетов с любой доступной точностью.

Например, длина выработки, наличие денежных средств в кассе предприятия в данный момент времени, план выпуска продукции на период и т.д. В зависимости от своей природы детерминированные па-

раметры могут иметь любые действительные значения, например, уклон выработки, неотрицательные значения – объем добычи за истекший период, или принимать только положительные целые значения – численность трудящихся.

Стохастические параметры представляют собой случайные величины, изменяющиеся в процессе работы предприятия.

Случайной величиной называется величина, которая в результате опыта может принять то или иное неизвестное заранее значение, при сохранении исходных условий.

В отличие от детерминированной, постоянной величины случайная величина не имеет единственного, истинного значения. Оно может случайным образом меняться от опыта к опыту независимо от точности наблюдения при прочих равных условиях.

По своей природе стохастические параметры производственных систем могут быть дискретными или непрерывными случайными величинами.

Дискретная случайная величина принимает в результате испытаний ряд значений, которые можно перенумеровать. Например, число электровозов, безотказно проработавших смену.

Значения непрерывной случайной величины, полученные в результате бесчисленного множества испытаний, непрерывно заполняют некоторый числовой интервал. Например, фактический объем суточной добычи из очистного забоя.

Полученное в результате измерения или расчета значение детерминированного параметра может служить оценкой его истинной величины и использовано в дальнейшем. В отличие от этого разовое измерение стохастического параметра не может служить его оценкой, так как в силу своей природы при следующем испытании он может принять иное значение. Поэтому стохастические параметры производственных систем задаются законами распределения или отдельными числовыми характеристиками распределений.

Закон распределения случайной величины – это соотношение, устанавливающее связь между возможными значениями случайной величины и вероятностями их появления. Закон распределения может иметь разные формы, но наиболее употребительные следующие.

Ряд распределения дискретной случайной величины x представляет собой таблицу, в которой перечислены возможные значения случайной величины $x_1, x_2, \ldots, x_n$ и соответствующие им вероятности $P_1, P_2, \ldots, P_n$.

Значения величины	x_i	x_1	x_2	...	x_n
Вероятности значения	P_i	P_1	P_2	...	P_n

где $P_i = P(X = x_i)$; (I.4)

$$\sum_{i=1}^{n} P_i = 1. \quad \text{(I.5)}$$

Выражение (I.4) читается следующим образом: P_i – это вероятность того, что величина X примет значение x_i.

Функция распределения случайной величины X

$$F(x) = P, \quad (X < x), \quad \text{(I.6)}$$

выраженная вероятность того, что X примет значение меньшее, чем x. $F(x)$ – неубывающая функция непрерывная слева такая, что

$$F(+\infty) = 1, \quad F(-\infty) = 0.$$

Вероятность попадания величины X в интервал от x_1 до x_2 (включая x_1) определяется из выражения

$$P(x_1 \leqslant X < x_2) = F(x_2) - F(x_1). \quad \text{(I.7)}$$

Отсюда следует, что вероятность конкретного значения x_0 непрерывной случайной величины X

$$P(X = x_0) = F(x_0) - F(x_0) = 0. \quad \text{(I.8)}$$

Это означает, что если заменить непрерывную случайную величину X ее значением x_0 в какой-либо модели, то такая модель с вероятностью равной единице не будет точно соответствовать системе, так как вероятность того, что X примет значение x_0, равна нулю.

Из этого можно сделать вывод, что стохастические параметры системы имеет смысл задавать в виде некоторых интервалов значений, учитывая вероятность попадания величины в интервал. Обычно этот интервал включает математическое ожидание случайной величины, представляющее собой ее теоретическое среднее значение.

Определение закона распределения случайного параметра, оценка математического ожидания и других числовых характеристик выполняются методами математической статистики.

Но основная трудность моделирования систем со стохастическими параметрами заключается в том, что случайное изменение величин может нарушить соответствие модели и системы и требует принципиально иного подхода к принятию решения на ее основе.

Пусть моделируется система с одним детерминированным параметром a и стохастическим параметром ϵ. Модель $Y = \varphi(a, \epsilon)$ соответствует системе и дает возможность принять правильное решение, если $\epsilon_1 \leqslant \leqslant \epsilon < \epsilon_2$.

Допустим, что вероятность попадания ϵ в указанный интервал $P = 0{,}9$. Если обратиться к определению, то вероятность – это предел,

к которому стремится отношение числа благоприятных исходов к их общему числу при бесконечном количестве испытаний. То есть вероятность указывает на закономерность в частоте появления события с ростом числа опытов.

В данном примере вероятность $P = 0{,}9$ говорит о том, что примерно в 90 случаях из 100 модель будет давать правильные результаты, а в остальных – ложные. Сравнительно редкая возможность ошибки моделирования не гарантирует от неудачи при одноразовом использовании модели и делает ее почти неизбежной при многократном применении. Это заставляет в обоих случаях страховаться от неудач в принятии решений, применяя для этого специальную процедуру.

Качественные параметры системы не имеют количественной меры. Производственные системы отличаются большим числом таких показателей. Например, вид технологии, тип оборудования, применяемая форма организации производства, труда, заработной платы, вид выпускаемой продукции и т.д.

Измерение или расчет количественного показателя системы позволяет соотнести его величину с эталонной мерой. Для определения качественных показателей также устанавливают эталон, который представляет собой минимально необходимый набор качественных признаков (призначный минимум). Если система обладает призначным минимумом, то считают, что она имеет качество, определяемое данным качественным показателем, и наоборот.

Прямое включение в модель качественных параметров не представляется возможным. Поэтому множество качественных признаков заменяют множеством бинарных переменных так, что каждому признаку A соответствует переменная α, принимающая значения $\alpha = 1$, если признак A имеется у системы; $\alpha = 0$, если признак A отсутствует.

Важно понять следующее: математическая модель системы – это не просто формальная запись некоторой абстрактной функции, существующей на множестве действительных чисел, а описание взаимосвязи параметров реального объекта, лежащих в определенных пределах, характерных для данной системы в строго определенных условиях. То есть модель всегда предполагает выполнение следующего гипотетического условия – если X и Y – параметры системы и она находится в допустимом состоянии, то реальные значения параметров X и Y будут связаны соотношением, описанным моделью. В случае детерминированных параметров предполагается, что каждому состоянию системы соответствует единственное значение детерминированного параметра, которое сохраняется, если состояние системы остается неизменным. Стохастические параметры могут менять свое значение при неизменном уровне других показателей.

Лежащий в основе моделирования принцип воспроизводимости результатов предполагает, что состояние системы может быть повторено любое число раз и соответственно определены значения ее параметров.

Именно воспроизводимость результатов позволяет планировать развитие производства в будущем, но этот принцип заставляет строго разделить стохастические и детерминированные параметры.

Тип параметров системы определяет возможные взаимосвязи между ними, а следовательно и модель, описывающую эту взаимосвязь.

Какие же взаимосвязи определены для параметров систем разных типов и как это влияет на метод моделирования?

Между детерминированными параметрами, с учетом их физической сущности и единиц измерения, возможны следующие взаимосвязи:

равенство $x = y$;

неравенство $x < y$;

функциональная связь в виде числовой функции действительных переменных $y = F(x_1, x_2, \ldots, x_n)$.

В этом случае модели будут иметь вид систем уравнений и неравенств, составленных на основе теоретических расчетных формул, физических законов, аналитических зависимостей, полученных в процессе анализа системы. Конкретные числовые значения коэффициентов функций получаются путем измерений или расчетов. При получении, исследовании и преобразовании этих моделей используются элементарная алгебра, геометрия, тригонометрия, дифференциальное и интегральное исчисления.

Например, модель, связывающая плановую производительность труда горнорабочего очистного забоя с плановой нагрузкой на участки шахты, может быть получена из двух выражений

$$П = A_{ш}/N; \tag{I.9}$$

$$A_{ш} = A_1 + A_2 + \ldots + A_m, \tag{I.10}$$

где $П$ – производительность труда горнорабочего очистного забоя, т/смену; N – планируемое количество выходов рабочих на всех участках шахты в смену; $A_{ш}$ – суммарная плановая добыча угля из очистных забоев, т; A_i – плановая нагрузка на i-й очистной забой, т; m – количество очистных забоев.

Выражение (I.9) вытекает из определения показателя производительности труда, в (I.10) – выводится из физической сущности $A_{ш}$.

В случае стохастических параметров возможные взаимосвязи и методы моделирования резко меняются.

Для случайных величин равенство $\alpha = \beta$ и неравенство $\alpha \leqslant \beta$ теряют смысл, так как в каждой реализации случайные величины α и β будут принимать новые значения и отношение $\alpha_i = \beta_i$ или $\alpha_i \leqslant \beta_i$ будет выполняться непостоянно при $i = 1, 2, \ldots, m$. Отсюда можно сделать вывод, что равенство и неравенство должны определяться только для постоянных характеристик случайных величин, которыми являются законы распределения и их числовые параметры. Так, например, воз-

можно отношение $F(\alpha) = F(\beta)$ или $M_\alpha \leqslant M_\beta$, где M_α и M_β – математические ожидания случайных величин α и β.

Для дискретных случайных величин, имеющих определенную вероятность отдельных значений, возможно введение равенства по вероятности $P(\alpha_i = \beta_i) = P_i$, означающее, что отношение $\alpha_i = \beta_i$ выполняется с вероятностью P_i.

Для непрерывных случайных величин, имеющих нулевую вероятность отдельных значений, введение отношения равенства по вероятности теряет смысл. Неравенство по вероятности можно вводить как для дискретных, так и для непрерывных случайных величин выражением $P(\alpha \leqslant \beta) = P$, где P – вероятность выполнения отношения.

Изменчивость значений случайных величин в различных испытаниях позволяет в одних случаях говорить о функциональной связи между ними, а в других – о статистической. Так, фактическая производительность горнорабочего очистного забоя на выход $П_ф = A_{ш.ф}/N_ф$ – является неслучайной функцией двух случайных аргументов: $A_{ш.ф}$ – фактическая добыча угля из очистных забоев, т; $N_ф$ – количество выходов горнорабочих в смену.

При этом $П_ф$ также становится случайной величиной. В этом случае имеет место так называемая функция случайного аргумента.

Если рассмотреть отношение $A_{ш.ф}$ и $N_ф$, то из их физической сущности можно предположить, что уровень добычи прямо пропорционально зависит от количества выходов, но практика показывает, что их связь не имеет функционального характера. Это происходит потому, что на $A_{ш.ф}$ трудоемкость работ по выемке угля влияет целый ряд случайных факторов. В результате между $A_{ш.ф}$ и $N_ф$ прослеживается так называемая статистическая зависимость.

В отличие от функциональной связи статистическая зависимость не устанавливает взаимооднозначного соответствия между значениями аргумента и функции, а определяют линию, вокруг которой группируются их значения в различных испытаниях. Статистическая связь случайных величин α и β описывается уравнением вида

$$\alpha = f(\beta) + \epsilon, \qquad (I.11)$$

где $f(\beta)$ – функция регрессии, связывающая математические ожидания случайных величин; ϵ – случайная величина, представляющая собой отклонение фактических значений α от линии регрессии.

Следует отметить, что функция регрессии не является взаимно обратной, поэтому из функции регрессии α от β нельзя получить функцию регрессии β от α путем обычных алгебраических преобразований.

При взаимодействии стохастических и детерминированных параметров возможна только статистическая связь, в которой детерминированная величина играет роль аргумента.

Таким образом, статистическая связь между параметрами системы принципиально отличается от функциональной связи и требует использования иных видов моделей и методов моделирования. При наличии стохастических параметров выбор вида моделей и их интерпретация выполняется с помощью теории вероятностей, а их оценка, сравнение, построение регрессионных зависимостей и т.д. проводится с использованием теории планирования эксперимента и математической статистики.

Наличие качественных параметров в системе требует использования специфических моделей и методов моделирования, так как это порождает принципиально отличные типы взаимосвязей параметров, требующих своих методов описания. Взаимосвязь может иметь место между двумя качественными или качественным и количественным параметрами. Отсутствие количественной меры делает неопределенными равенство, неравенство и функциональную связь между качественными параметрами, но это не исключает эквивалентности и связи вообще.

Равенство является простейшим случаем эквивалентности, основанном на следующем логическом правиле сравнения: если в результате измерения двух величин A и B их меры совпадают, то выполняется отношение $A = B$. Вместо количественной меры качественный показатель характеризует наличие некоторого множества качественных признаков у системы, поэтому логическое правило проверки эквивалентности качественных признаков можно сформулировать следующим образом: если в результате определения двух качественных показателей устанавливается наличие одних и тех же признаков, то они эквивалентны.

Определение качественного показателя приводит к двум взаимоисключающим исходам – наличию или отсутствию описываемого качества. Отсюда взаимосвязь качественных показателей проявляется в логике и вероятности наступления событий. Таким образом, взаимосвязь качественных параметров имеет тип логических отношений, носящих детерминированный или стохастический характер.

Для описания такого рода взаимосвязей используются булева алгебра, теория графов, теория вероятности.

Проведенный предварительный анализ свойств показателей производственных систем позволяет сделать следующие выводы. Во-первых, характер параметров системы определяет вид модели, метод моделирования и способ принятия решения.

Во-вторых, ошибка в определении типа параметра может привести к методическим ошибкам в моделировании и ложным результатам в изучении системы.

В дальнейшем, если это будет оговорено особо, изложение методов математического программирования будет вестись для случая детерминированных параметров с известными функциональными связями.

§ 3. КАЧЕСТВЕННЫЕ ХАРАКТЕРИСТИКИ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Цель построения математической модели – получение формального описания объекта для его дальнейшего изучения или управления им. Поэтому ее основные качественные характеристики должны оценивать адекватность модели системе и решаемой задаче.

Термин "адекватность" вводится в теории познания и означает соответствие образа знания своему оригиналу. Математическое моделирование представляет один из способов отображения реальной действительности, поэтому на него можно распространить это понятие. Адекватность предполагает комплексную оценку отображения по полноте и точности описания. Что же представляют собой эти характеристики и как их соотнести с показателями математической модели?

Всякая модель предполагает описание состояний и свойств системы. Под свойством здесь понимается присущий системе комплекс параметров с отношениями между ними. Свойства системы порождают множество ее состояний, поэтому знание свойств и настоящего состояния позволяют полностью определить ее развитие. Количество описанных свойств определяет полноту модели. Чем полнее модель, тем выше адекватность и больше количество состояний, в которых модель объективно описывает свойства системы.

Часть свойств представляют собой отношения между количественными параметрами системы, что вызывает необходимость в оценке точности моделирования. Точность модели можно определить как оценку разности между расчетными и истинными значениями моделируемых величин. Необходимая точность моделирования должна обеспечивать однозначное описание реальной взаимосвязи величин.

Например, выражения $a = b$ и $a < b$ теряют однозначность, если интервальные оценки этих величин, построенные с учетом точности моделирования $\pm \epsilon_a$ и $\pm \epsilon_b$, перекрываются.

Введенные показатели полноты и точности не позволяют прямо оценить адекватность модели по следующим причинам.

Большая часть свойств производственных систем и истинные значения параметров не известны и устанавливаются в процессе моделирования. Поэтому оценка адекватности путем сравнения множеств описанных и реально существующих свойств является диалектически неразрешимой задачей, как прямое сравнение абсолютной и относительной истины. В реальном моделировании имеет смысл говорить об относительной адекватности различных моделей.

Всякая производственная система обладает очень большим числом существенных и несущественных свойств, описать которые одной моделью можно только чисто теоретически. Реальные модели всегда описывают часть свойств, наиболее существенных с точки зрения исследователя. Поэтому сравнительная оценка адекватности двух моделей по числу описанных свойств не имеет реального смысла.

Моделирование взаимосвязи количественных параметров неизбежно связано с погрешностями расчетов, измерений. Чем полнее модель, тем больше последовательных расчетов и тем ниже точность их результатов, что вытекает из закона накопления ошибки последовательных вычислений. Следовательно, точность моделирования отдельных величин необходимо соотносить с полнотой модели.

В реальном моделировании не всегда удается точно сформулировать описываемое свойство, поэтому полноту модели оценивают по таким показателям, как число включенных в нее параметров и переменных, рассматриваемых соотношений. Точность моделирования определяется сравнением расчетных и известных значений параметров системы в прошлых периодах или специальных экспериментах, а также путем сравнения значений, полученных по разным моделям. Для оценки результатов моделирования с учетом уровня погрешности значений параметров определяют чувствительность модели, которая представляет собой относительное изменение результата моделирования при единичном изменении параметров.

Окончательный вывод об адекватности модели и системы выносят после ее экспериментальных испытаний.

Кроме адекватности системе модель должна соответствовать решаемой задаче. В определенной задаче математическая модель связывает управляемую систему и лицо, проводящее операцию, через цель операции. Поэтому проверка адекватности поставленной задачи сводится к анализу отражения цели в модели.

Из приведенных рассуждений можно сделать вывод о том, что адекватность модели следует всемерно повышать за счет введения большего числа параметров системы, связанных с целью операции, и увеличения точности описания зависимостей. Но, к сожалению, действуя подобным образом, не всегда удается успешно решить задачу из-за чрезмерной сложности полученной модели.

Сложность – понятие относительное и определяется через принятую меру сложности, которую целесообразно выбирать с точки зрения возможности получения конечного результата. В этом плане можно выделить две основные проблемы: выбор метода и составление алгоритма поиска оптимального решения, необходимый объем вычислений.

Сложность модели с точки зрения возможных способов поиска оптимального решения зависит от типа входящих в нее величин и вида зависимостей. Так как в настоящее время в математическом программировании разработаны стандартные методы решения лишь ограниченного круга задач, то практически оценка сложности модели сводится к проверке наличия стандартного метода решения. Грубо оценить сложность модели на начальном этапе ее разработки можно исходя из того, что, как правило, стохастическая модель сложнее модели с детерминированными величинами; линейная модель проще модели с любыми другими зависимостями в целевой функции или ограничениях; ограни-

чение целочисленности хотя бы одной переменной резко усложняет любую модель.

Практика показывает, что объем вычислений при решении задач одного класса и одним методом растет со скоростью показательной функции при увеличении порядка системы ограничений. Поэтому для сравнительной оценки вычислительной сложности двух моделей одного типа достаточно сравнить число переменных и ограничений.

Моделирование представляет собой лишь один из этапов решения задачи, поэтому при составлении модели необходимо не только стремиться к ее адекватности, но и ограничивать ее сложность.

Введение дополнительных переменных, усложнение зависимостей в погоне за точностью описания приводит к росту размерности задачи, повышают трудоемкость сбора исходной информации и усложняют методы поиска оптимального решения. В результате чрезмерно сложная модель может сделать неразрешимой поставленную задачу. С другой стороны, предметные упрощения могут настолько снизить точность результатов и ограничить область применения модели, что ее использование в операционной задаче потеряет всякий смысл. Поэтому при моделировании приходится искать "золотую" середину между чрезмерной сложностью и недопустимыми упрощениями.

§ 4. ОПТИМИЗАЦИОННЫЕ МОДЕЛИ, ИХ СТРУКТУРА

Процессы принятия решений при всем их многообразии имеют две характерные черты.

1. Условия задачи допускают большое количество возможных вариантов, из которых надо выбрать оптимальный. Притом с увеличением количества рассматриваемых вариантов растет объем информации, необходимой для решения задачи и ее описание становится более громоздким.

2. Принятие решения осуществляется для определенной цели, т.е. выбранное решение должно наилучшим образом обеспечивать достижение поставленной цели.

Для сравнения возможных вариантов и оценки их соответствия поставленной цели необходимо уметь количественно оценивать степень достижения цели при различных вариантах или эффективность операции. Количественный критерий, с помощью которого определяется эффективность операции, называется показателем эффективности.

Таким образом, процесс принятия решений можно описать функцией, аргументами которой являются допустимые решения, а значениями – числа, характеризующие меру достижения поставленной цели при различных аргументах. Эта функция называется целевой. Она связывает допустимые решения с показателем эффективности. Задача выбора решения сводится к нахождению экстремального значения функции (показателя эффективности) и аргумента, на котором оно достигается.

Решение, максимизирующее (минимизирующее) функцию (показатель эффективности), является оптимальным.

Эффективность операции зависит от двух групп факторов: условий ее проведения (наличие ресурсов, погодные условия, плановые задания и т.д.) $-a_1, a_2, \ldots, a_n$ и способа организации и параметров операции $x_1, x_2, \ldots, x_m$. Выбор способа организации операции и ее параметров представляет собой решение задачи, а искомые величины $x_1, x_2, \ldots, x_m$ – его элементы.

Критерием эффективности операции является функция заданных условий и элементов решения

$$W = f(a_1, a_2, \ldots, a_n;\ x_1, x_2, \ldots, x_m).$$

Заданные условия $a_1, a_2, \ldots, a_n$ могут быть функциями и ограничениями элементов решения.

Основная задача исследования операций математически может быть сформулирована следующим образом: при заданных условиях $a_1, a_2, \ldots, a_n$ найти такие элементы решения $x_1, x_2, \ldots, x_m$, при которых показатель W достигает максимума (минимума).

Это типичная математическая вариационная задача. Однако решить ее методами классической математики невозможно или затруднительно из-за большого числа переменных, наличия ограничений и иногда из-за дискретности изменения переменных. Для решения таких задач применяют методы математического программирования.

Рассмотренный выше случай представляет собой так называемый детерминированный процесс, который в чистом виде наблюдается редко.

Обычно на исход операции влияют различные случайные факторы, при этом условия ее проведения известны неточно или не полностью. Например, горно-геологические условия залегания месторождения и качественная характеристика полезного ископаемого, а также производительность оборудования являются вероятностными величинами. В подобных случаях кроме известных условий выполнения операции $a_1, a_2, \ldots, a_n$ и элементов решения $x_1, x_2, \ldots, x_m$ показатель эффективности W зависит также от неизвестных факторов $b_1, b_2, \ldots, b_l$.

$$W = f(a_1, a_2, \ldots, a_n;\ b_1, b_2, \ldots, b_l;\ x_1, x_2, \ldots, x_m).$$

Хотя $b_1, b_2, \ldots, b_l$ неизвестны вообще или известны неточно, задача принятия решений существует и формулируется следующим образом: при заданных $a_1, a_2, \ldots, a_n$ с учетом неизвестных факторов $b_1, b_2, \ldots, b_l$ найти элементы решения $x_1, x_2, \ldots, x_m$, обращающие по возможности в максимум (минимум) критерий эффективности W. Это уже не чисто математическая задача. Наличие неизвестных факторов $b_1, b_2, \ldots, b_l$ переводит ее из разряда детерминированных в разряд статистических, которые подразделяются на задачи в условиях риска и неопределенности.

К задачам в условиях риска относятся такие, в которых неизвестные факторы $b_1, b_2, \ldots, b_l$ могут быть охарактеризованы статистически, определенными законами распределения. Например, работу оборудования, ошибку подсчета запасов можно охарактеризовать статистическими показателями, закон распределения которых находят методами математической статистики.

Если $b_1, b_2, \ldots, b_l$ – случайные величины с известными (определенными) законами распределения, то оптимизация решения может осуществляться двумя способами: сведением к детерминированной задаче или оптимизацией "в среднем".

При первом способе случайные величины $b_1, b_2, \ldots, b_l$ заменяются их математическими ожиданиями, т.е. неизвестные факторы переводятся в разряд детерминированных. Этот прием достаточно груб и применим в том случае, если случайные факторы изменяются в небольшом диапазоне.

Второй способ более сложен. Критерий оптимальности выражается статистически, в него вводятся законы распределения случайных факторов и выбирается решение, оптимизирующее математическое ожидание показателя эффективности.

Если задача решается многократно (например, при сменном планировании), то результат будет известен. При разовом решении задачи (например, при проектировании предприятия) случайные величины могут существенно отклоняться от средних значений. Поэтому наряду с расчетом математического ожидания надо оценивать его дисперсию.

Когда неизвестные факторы $b_1, b_2, \ldots, b_l$ нельзя охарактеризовать статистически (нельзя получить законы распределения этих факторов), возникает задача выбора решений в условиях неопределенности. В этом случае вместо оптимизации "в среднем" следует рассматривать диапазон изменения факторов $b_1, b_2, \ldots, b_l$ и для каждого из них находить решение. В результате будут найдены локально-оптимальные решения, являющиеся оптимальными при определенных значениях неизвестных факторов $b_1, b_2, \ldots, b_l$. Совокупность локально-оптимальных решений дает полную картину для выбора оптимального варианта. Если нет решения, предпочтительного для всех условий (доминирующего), то можно принять решение, которое не является строго оптимальным ни для одного из условий, но приемлемо для целого их ряда. Такая задача относится к задачам выбора компромиссного решения.

В ряде случаев факторы $b_1, b_2, \ldots, b_l$ зависят не от объективных обстоятельств, а являются результатом действия активно противостоящего нам противника. Математическая теория конфликтных ситуаций достаточно хорошо разработана (теория игр). Она позволяет принимать наименее рискованное решение, дает основания для выбора стратегии поведения.

Несмотря на разнообразие встречающихся операционных задач, лежащие в их основе оптимизационные модели включают всегда одни

и те же составные части и имеют общий вид

$$Z = F(x_1, x_2, \ldots, x_n;\ a_1, a_2, \ldots, a_m) \to \max(\min); \tag{I.12}$$

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n;\ a_1, a_2, \ldots, a_m) \leqslant A_1, \\ f_r(x_1, x_2, \ldots, x_n;\ a_1, a_2, \ldots, a_m) \leqslant A_r; \end{cases} \tag{I.13}$$

$$\begin{cases} \varphi_1(x_1, x_2, \ldots, x_n;\ a_1, a_2, \ldots, a_m) = 0, \\ \varphi_k(x_1, x_2, \ldots, x_n;\ a_1, a_2, \ldots, a_m) = 0, \end{cases} \tag{I.14}$$

где $x_1, x_2, \ldots, x_n$ – управляемые, оптимизируемые переменные, значения которых необходимо найти в результате решения; $a_1, a_2, \ldots, a_m$ – неуправляемые переменные или параметры модели; Z – критерий оптимальности; $A_1, A_2, \ldots, A_r$ – граничные условия операции.

Критерий оптимальности – это показатель эффективности решения, выбираемый на стадии постановки оптимизационной задачи.

Функция (I.12), связывающая критерий оптимальности с параметрами системы, называется целевой функцией.

Граничные условия $A_1, A_2, \ldots, A_r$ определяются на стадии постановки, конкретизируют цель и ограничивают условия проведения операции. Величины $x_1, x_2, \ldots, x_n$ и $a_1, a_2, \ldots, a_m$ представляют собой параметры системы, существенно связанные с критерием оптимальности и граничными условиями. В качестве оптимизируемых величин $x_1, x_2, \ldots, x_n$ выбирают параметры, на которые может влиять руководитель операции.

Остальные существенные параметры системы $a_1, a_2, \ldots, a_m$ выполняют роль параметров модели. Эти параметры учитываются руководителем операции, но он не имеет средств воздействовать на них. Следует отметить, что одна и та же величина может быть управляемой переменной или параметром модели в зависимости от лица, принимающего решения.

Например, проектировщик может рассматривать газоносность рабочего пласта, как управляемую переменную, предусматривая его предварительную дегазацию. Для начальника участка это уже неуправляемый фактор, который необходимо учитывать при решении целого ряда производственных задач, но на который у него нет средств повлиять.

Система ограничений (I.13) связывает оптимизируемые величины и параметры модели с граничными условиями операции, заданными постановкой задачи. Но этих ограничений часто бывает недостаточно, так как они обычно учитывают внешние условия операции и не отражают внутренних связей существенных параметров системы. С этой целью в модель вводится система ограничений (I.14), отражающая внутренние условия проведения операции.

Например, поставлена задача: установить сменные задания участкам таким образом, чтобы план шахты выполнялся при минимальной средней зольности добываемого угля.

В этой задаче оптимизируемыми переменными являются нагрузки на участки шахты – $x_1, x_2, \ldots, x_n$, критерием оптимальности – зольность добываемого угля в среднем по шахте Z, а граничными условиями – сменное плановое задание шахте по добыче угля A. Поставленную задачу при этих обозначениях можно записать в виде модели

$$Z = \frac{\sum_{i=1}^{n} Z_i x_i}{\sum_{i=1}^{n} x_i} \to \min,$$

$$\sum_{i=1}^{n} x_i = A,$$

где Z_i – зольность угля на i-м участке шахты; n – число участков.

Эта модель дает тривиальное решение задачи $x_k = A$, если $Z_k = \min Z_i, i = 1, 2, \ldots, n$ и остальные $x_i = 0, i \neq k, i = 1, 2, \ldots, n$. То есть вся нагрузка на шахту должна выполняться участком с минимальной зольностью добываемого угля. Очевидно, что такое решение в большинстве случаев неприемлемо из-за ограниченной производительности участков, пропускной способности транспортных звеньев, обслуживающих несколько очистных забоев, условий проветривания и т.д. Следовательно, модель должна быть дополнена системой ограничений, учитывающих условия, специфические для каждой конкретной шахты.

Структура системы ограничений (I.13) определяется постановкой задачи и, как правило, формулируется без особого труда. В отличие от этого, каждое ограничение системы (I.14) – результат исследования возможных последствий изменения оптимизируемых величин в условиях конкретной производственной системы.

Критерий оптимальности выбирается на стадии постановки задачи. Как правило, это один из натуральных или экономических показателей системы, служащий для оценки ее эффективности. Реже в качестве критерия оптимальности используется функция нескольких показателей системы, отражающих разные аспекты эффективности ее функционирования. Такие критерии называют синтетическими. В отличие от натуральных и экономических, синтетические критерии обычно не имеют четкого физического или экономического смысла и не позволяют прямо определить эффект от реализации полученного решения.

На стадии проектирования горных предприятий обычно используется критерий – минимум удельных приведенных затрат, учитывающий эффективность капитальных вложений в народное хозяйство и разновременность затрат.

На стадии эксплуатации для оценки производственных решений с небольшим сроком реализации в качестве критериев обычно используют показатели: себестоимость продукции; производительность труда; удельная прибыль; рентабельность.

Следует отметить, что критерий оптимальности играет ведущую роль в согласовании работы взаимосвязанных систем. Проблема выбора критерия при проведении операции внутри сложной системы решается методами системного анализа.

Чем сложнее система, тем труднее найти единый показатель эффективности ее функционирования. Так, работа предприятия оценивается целым рядом плановых и расчетных показателей, отражающих разные аспекты производственно-хозяйственной деятельности. В результате решения многих производственных задач оказывается невозможным оценить одним критерием эффективности и возникает многокритериальная постановка задачи. Такая постановка предусматривает включение в цель операции нескольких критериев, оптимизируемых одним решением.

Нередко многокритериальная постановка возникает из-за того, что для какого-либо показателя не удается установить верхней или нижней границы на стадии постановки задачи. В этом случае из ограничивающего условия показатель переходит в роль дополнительного критерия, который должен минимизироваться или максимизироваться.

С математической точки зрения решение оптимизационной задачи сводится к отысканию условного экстремума целевой функции на множестве решений системы ограничений. В многокритериальной постановке число целевых функций равно числу критериев оптимальности, что в общем случае приводит к неразрешимости задачи из-за несовпадения точек условного экстремума функций. Поэтому при многокритериальной постановке ищется решение, которое в большей или меньшей степени приближает значение отдельных критериев к уровню условного экстремума соответствующей целевой функции.

Подобные задачи решают путем сведения их к однокритериальным одним из следующих способов:

введением комплексного критерия, объединяющего все показатели эффективности в одну функцию;

последовательным решением задачи поочередно с различными критериями и поиском некоторого компромисса между результатами.

Подробно способы решения многокритериальных задач будут рассмотрены ниже.

Одним из важнейших элементов оптимизационной модели является целевая функция (I.12), связывающая критерий оптимальности с оптимизируемыми переменными и параметрами модели. Эта функция может составляться из отдельных частных, аналитических или эмпирических зависимостей, представлять собой множественное уравнение регрессии или аналитически выводится на основе анализа системы, но в любом случае она должна отвечать следующим требованиям:

однозначно связывать все оптимизируемые переменные с критерием оптимальности;

не иметь разрывов на всем множестве допустимых значений оптимизируемых переменных.

Точность вычисления критерия должна позволять однозначно сравнивать два решения, отличающиеся на величину, близкую к погрешности определения любой из оптимизируемых переменных. Кроме этих требований желательно, чтобы целевая функция имела простейший вид и была унимодальной, т.е. имела один экстремум. Эти же требования и пожелания относятся к целевым функциям многокритериальных задач.

Появление в задаче качественных переменных может несколько изменить вид операционной модели, хотя ее структура в принципе не изменится. Прямое включение качественных переменных в оптимизационную модель невозможно, поэтому их учитывают косвенно, путем изменения структуры модели. Например, при оптимизации параметров подготовки шахтного поля могут рассматриваться различные системы разработки, каждая из которых характеризуется рядом общих и специфических переменных, входящих в различные расчетные выражения. В этом случае оптимизационная модель распадается как бы на ряд подмоделей, определяющих оптимальные варианты для каждой системы разработки по одному и тому же критерию. Окончательный выбор делается путем сравнения этих лучших для каждой системы разработки вариантов.

§ 5. ЭТАПЫ РЕШЕНИЯ ОПТИМИЗАЦИОННОЙ ЗАДАЧИ

Процесс решения любой операционной задачи на основе оптимизационной модели можно разделить на следующие этапы:

постановка задачи;

составление математической оптимизационной модели;

выбор метода и разработка алгоритма решения;

подготовка информации и вычисление оптимального решения;

анализ и проверка полученного решения;

разработка рекомендаций по использованию результатов.

Выделение этих этапов в известной мере условно, так как они теснейшим образом связаны между собой. Например, при составлении математической модели нередко приходится уточнять постановку задачи и постоянно иметь в виду возможность использования того или иного стандартного метода поиска оптимального решения. Но тем не менее, каждый из этих этапов имеет свои начальные условия, результат и определенную методику.

Результатом этапа постановки является формально записанная задача с полностью определенной целью операции. Напомним, что цель операции (I.1) содержит критерии оптимальности и граничные условия.

Для достижения этого результата необходимо: уяснить и словесно сформулировать задачу; выбрать критерии оптимальности; установить граничные условия; формально записать задачу и утвердить эту постановку.

В процессе уяснения и словесного описания задачи необходимо ответить на следующие вопросы.

Какой результат ожидается получить?

Чем его измерить?

Какие средства отпускаются для достижения результата, их источники?

Какими показателями измеряются эти средства?

Какой эффект ожидается получить и за счет чего?

Чем можно измерить эффект?

Организация операции может предусматривать два варианта:

руководитель операции сам формулирует и решает задачу, а также реализует результаты;

руководитель операции ставит задачу, принимает и реализует результаты, а формулирует и решает задачу специалист по исследованию операций, которого далее будем называть "исследователем".

В первом случае, который встречается обычно при решении небольших задач, руководителю необходимо самому себе ответить на указанные вопросы, как можно объективнее оценивая ситуацию, собственные пожелания и требования руководства.

Во втором случае ситуация сложнее, так как, во-первых, такая организация используется для решения более крупных проблем; во-вторых, в постановке задачи принимают участие два человека, что увеличивает влияние субъективных факторов.

Чтобы у исследователя сложилось правильное мнение по перечисленным вопросам, руководителю необходимо:

самому как можно полнее и объективнее ответить на эти вопросы;

невзирая на потери времени дать ответы на все вопросы исследователя;

в тактичной форме выяснить, насколько правильно ответы поняты исследователем.

В свою очередь, исследователю можно дать следующие советы:

перед очередной встречей с руководителем надо сформулировать перечень вопросов, по которым будет проходить беседа. В ходе беседы они могут уточняться;

в ходе беседы надо сразу уточнять используемые термины, названия показателей и т.д. Неправильное толкование названия или термина может исказить понимание задачи;

важно не только слушать, но и слышать, что говорят. Нередко у исследователя складывается свое представление о проблеме и он начинает пропускать мимо ушей то, что идет вразрез с его мнением;

для правильного понимания проблемы исследователю следует пос-

тараться представить себя на месте руководителя и взглянуть на ситуацию с этой точки зрения;

при изучении проблемы надо избегать стереотипов ранее решенных задач. Тем более нельзя пытаться втиснуть ее в рамки известной модели только за тем, чтобы использовать стандартный метод решений.

Для точной оценки ситуации исследователь должен проверять и уточнять все утверждения и замечания руководителя при помощи встречных вопросов. Все противоречивые высказывания должны обязательно разрешаться.

После уяснения общей постановки задачи выбираются критерии оптимальности. С этой целью составляются перечни показателей, определяющих результат и эффект. Затем на основе логического, инженерного анализа оценивается связь между результирующими показателями и дублируемые величины исключаются. То же выполняется для группы показателей эффективности. Далее устанавливается связь между результирующими показателями оценки эффективности. Критерии оптимальности выбирают среди группы показателей эффективности, тесно связанных с результатом операции. Если решаемая задача представляет собой часть более крупной проблемы, то выбор критериев оптимальности осуществляется на основе декомпозиции проблемы таким образом, чтобы достижение целей отдельных задач обеспечивало решение всего вопроса.

Граничные условия устанавливают на показатели, определяющие результат операции и выделяемые средства.

В результате этих исследований проблемы формулируется постановка задачи с количественными параметрами цели. Эту постановку целесообразно всесторонне рассмотреть с руководителем операции и только устранив все замечания приступать к построению модели.

Итог следующего этапа – оптимизационная модель задачи с полностью сформулированной системой ограничений и целевой функцией.

Для ее получения необходимо:

исследовать систему и выявить существенные для решения задачи параметры, сформулировать группы оптимизируемых и неуправляемых величин;

составить целевую функцию и систему ограничений, определенную граничными условиями;

описать ограничения, отражающие особенности моделируемого объекта.

Под существенными параметрами понимаются показатели системы, определяющие результат и эффективность операции. Для их выявления проводится всестороннее исследование факторов, влияющих на результирующие показатели и критерий оптимальности. С этой целью на основе инженерного, логического анализа выявляется исходная группа влияющих факторов, определяется логическая схема их взаимодействия, которая затем проверяется статистически или экспериментально. В

производственных задачах чаще используется статистический анализ на основе имеющихся технико-экономических показателей. В качестве переменных и параметров модели выбирают величины, существенно связанные с результатом операции и критерием эффективности и не дублирующие друг друга.

В качестве оптимизируемых переменных берут те показатели, на которые может влиять руководитель операции. Значения остальных оцениваются статистическим путем, прогнозируются или определяются заданиями вышестоящих организаций и включаются в модель в виде числовых параметров.

После выявления оптимизируемых переменных и параметров модели составляются целевая функция и ограничения. В зависимости от типа переменных, изученности системы и решаемой задачи строят аналитические или эмпирические зависимости, добиваясь примерно одинаковой точности во всех выражениях, входящих в модель.

Составив оптимизационную модель, приступают к выбору метода поиска оптимального решения, составлению алгоритма и программы расчета на ЭВМ. В настоящее время математическое программирование позволяет решать многие классы оптимизационных задач, а разработка новых методов требует больших теоретических изысканий, поэтому на практике выбор метода решения сводится к подбору стандартного алгоритма или их сочетаний. Метод оптимизации определяется видом модели и свойствами оптимизируемых переменных (непрерывные, дискретные, случайные), поэтому возможность использования того или иного алгоритма решения необходимо предвидеть еще на стадии составления модели. В отдельных случаях трудности поиска оптимального решения заставляют менять вид модели.

При решении многокритериальных задач сначала выбирают способ принятия решения (объединение критериев в комплексный критерий, поиск компромиссного решения), а затем – метод оптимизации для однокритериальных моделей, к которым сводится исходная задача.

Выбрав метод решения и уточнив вид модели, приступают к программированию задачи для ЭВМ и одновременно готовят исходную информацию. На этом этапе все параметры модели должны получить количественные значения, а для этого требуется провести сбор и обработку исходной информации. Следует отметить, что на сбор информации иногда приходится до трех четвертей всего времени моделирования, а наличие и ее достоверность определяют успех решения задачи.

После получения оптимального решения оно проверяется и всесторонне анализируется. Важность этого этапа трудно переоценить, так как именно здесь определяется возможность и целесообразность использования полученных результатов и принимается окончательное решение проблемы. Проверка модели и результатов проводится в две фазы – сначала выбирают способ проверки, а затем устанавливают адекватность модели, устойчивость и точность оптимального решения. Следует под-

черкнуть, что методически проверка результатов нередко представляет собой задачу не менее сложную, чем решение самой оптимизационной задачи.

В заключение разрабатывают рекомендации по внедрению результатов моделирования, в которых даны производственные ситуации для разработанных моделей, эффективность возможных решений и указания для практического действия.

Рекомендации должны быть написаны языком, понятным для тех, кто будет ими пользоваться. В них следует обстоятельно указывать функции каждого работника, внедряющего результаты моделирования, и порядок их выполнения. Однако на практике могут возникнуть такие обстоятельства, которые заранее трудно предвидеть. Поэтому специалисты, разрабатывающие математические модели, должны принимать непосредственное участие в реализации своих выводов и рекомендаций.

Использование полученных рекомендаций должно позволить руководителю лучше обосновать принимаемое решение. Причем использование методов математического моделирования не снимает с него обязанности принимать решения и отвечать за них.

Контрольные вопросы и задачи

1. Когда совокупность элементов становится системой?
2. Чем определяется состояние системы?
3. Что включает в себя цель операции?
4. Можно ли словесное описание считать абстрактной моделью?
5. На какие группы делятся математические модели?
6. Какие типы параметров могут описывать состояние системы?
7. Какие из перечисленных отношений имеют смысл, если A и B – детерминированные, α и β – случайные величины, а C и D – качественные признаки: $a = \alpha$; $\beta = \alpha$; $C = D$; $b = \alpha$; $c > b$; $\beta < \alpha$; $\alpha > b$?
8. Возможна ли функциональная связь между случайной и детерминированной величиной, детерминированной и случайной величиной?
9. Как определяется адекватность модели?
10. Почему в реальных моделях приходится ограничивать полноту описания системы?
11. Какие элементы обязательно присутствуют в любой оптимизационной модели?
12. Может ли критерий оптимальности быть качественным показателем?
13. Может ли оптимизационная модель включать несколько целевых функций?
14. Из каких этапов состоит решение операционной задачи?

Глава II

МЕТОДЫ И СРЕДСТВА ПОИСКА ОПТИМАЛЬНЫХ РЕШЕНИЙ

§ 1. ОСОБЕННОСТИ РЕШЕНИЯ ЗАДАЧ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Разработка методов поиска оптимального решения на заданной оптимизационной модели составляет предмет математического программирования.

В общем виде задача математического программирования формируется следующим образом: найти значения переменных $x_1, x_2, \ldots, x_n$, обеспечивающие максимум (минимум) целевой функции

$$Z = f(x_1, x_2, \ldots, x_n) \rightarrow \max \ (\min) \qquad \text{(II.1)}$$

на множестве допустимых решений системы ограничений

$$\varphi_i (x_1, x_2, \ldots, x_n) \leqslant \beta_i, \quad i = 1, 2, \ldots, m. \qquad \text{(II.2)}$$

В качестве дополнительных ограничений в задачу может включаться требование целочисленности или бинарности (значения 0 или 1) всех или некоторых переменных.

Метод решения – это теоретически обоснованная совокупность правил, обеспечивающих решение определенного типа задач. Важно отметить, что метод решения – это не только алгоритм расчета, но и правила оценки результата, анализ возможности возникновения ошибок, обоснование точности решения.

По своей постановке задачи математического программирования образуют класс задач поиска условного экстремума функции многих переменных. Их принципиальное отличие от задач поиска безусловного экстремума заключается в том, что при ограниченной области изменения переменных максимальное значение функции может достигаться не только в точке ее экстремума, но и на границе изменения переменных. Это хорошо иллюстрируется следующим примером. Пусть необходимо найти максимум линейной целевой функции $Z = c_1 x_1 + c_2 x_2 + Q \rightarrow \max$ на множестве решений системы

$$\begin{cases} a_1 x_1 + a_2 x_2 = b, \\ x_1 \geqslant 0, \\ x_2 \geqslant 0. \end{cases}$$

Из математического анализа известно, что необходимое условие

существования безусловного экстремума – наличие точки, в которой частные производные от всех переменных обращаются в ноль. В данном случае

$$\frac{\partial Z}{\partial x_1} = c_1 \neq 0 \text{ и } \frac{\partial Z}{\partial x_2} = c_2 \neq 0.$$

То есть функция не имеет точек безусловного экстремума и можно говорить только об условном максимуме на множестве допустимых значений аргументов x_1 и x_2. Этими значениями являются неотрицательные решения ограничения – равенства $a_1x_1 + a_2x_2 = b$ и данная задача сводится к исследованию этих решений и значений функции Z на них.

Таким образом, решение задач математического программирования в первую очередь требует разработки методов нахождения и перебора решений систем ограничений, состоящих из уравнений и неравенств многих переменных, и затем методов анализа изменения целевой функции на этом множестве решений.

Другой отличительной особенностью задач математического программирования является возможность разрывов в области допустимых значений переменных. Так, в задачах с дискретными переменными область допустимых значений распадается на ряд отдельных точек, что вообще исключает применение метода анализа производных.

Особое место в задачах математического программирования занимают модели со случайными функциями. Для этих задач понятие единственного оптимального решения вообще теряет смысл, так как наличие случайных функций-ограничений "размывает" область допустимых значений, делает ее неопределенной, а случайная целевая функция приводит к неопределенности выбора двух близких решений.

Для решения оптимизационных задач в настоящее время разработано большое число методов, которые можно разделить на две группы: конечные и приближенные.

Конечные методы гарантируют получение точного решения за конечное число шагов. Точность решения определяется только погрешностью вычислений. Приближенные методы позволяют получить лишь приближенное значение оптимума с погрешностью, заложенной в метод решения задачи, а число шагов в принципе может быть бесконечным.

Как конечные, так и приближенные методы могут быть универсальные, позволяющие решать некоторый класс задач, или специальные – применяемые для решения задач только определенного типа.

По принципу поиска оптимального решения все методы можно разделить на аналитические и итерационные. В первом случае за счет аналитических преобразований и вычислений получают единственное допустимое и оно же оптимальное решение. Во втором – оптимум находят путем перебора некоторого числа допустимых решений.

Возможность использования того или иного метода, его эффективность, вероятность ошибочных результатов и т.д. требуют многостороннего теоретического исследования и обоснования. Для наиболее распространенных в настоящее время итерационных методов такое обоснование строится по следующей схеме:

определяют свойства множества допустимых решений в результате исследования отдельных ограничений;

определяют условия существования оптимума и возможные исходы решения задачи исходя из этих свойств и вида целевой функции;

выбирают стратегию поиска оптимального решения. Определяют критерии перехода от одного решения к другому, обеспечивающие эту стратегию. Оценивают конечность поиска с помощью этой стратегии;

разрабатывают численную процедуру, найдя соответствующие признаки допустимого и оптимального решений.

Изложение теоретического обоснования всех приводимых ниже методов не представляется возможным в рамках учебного пособия, но описание одних алгоритмов также недопустимо, так как это затруднит понимание сути оптимизации и может привести к ошибкам в постановке задач, интерпретации результатов. Наиболее подробно изложена теория линейного программирования, имеющего большое самостоятельное практическое значение и широко применяемого в других методах оптимизации. В остальных случаях теория будет излагаться только в минимальном объеме, необходимом для понимания сути процесса оптимизации.

§ 2. ГРАФИЧЕСКОЕ РЕШЕНИЕ ОПТИМИЗАЦИОННОЙ ЗАДАЧИ

Графическое решение оптимизационных задач вряд ли можно рассматривать как метод математического программирования из-за его очень ограниченных возможностей, но он полезен для иллюстрации идей, заложенных в численные методы решения.

Графически легко решаются задачи с одной или двумя оптимизируемыми переменными.

В этом случае решение сводится к построению и анализу точек пересечения ряда линий на плоскости. Задачи с тремя переменными требуют изображения и анализа линий пересечения поверхностей в трехмерном пространстве, что весьма сложно с графической точки зрения, а большее число переменных делает задачу графически неразрешимой.

Графическое решение включает следующие операции:

в прямоугольных координатах, соответствующих оптимизируемым переменным, строят области решений всех ограничений;

определяют пересечение областей решения всех ограничений, которое представляет собой множество допустимых решений задачи;

строят ряд линий равного уровня целевой функции, покрывающих область допустимых решений;

по линиям уровня находят точку оптимума целевой функции на области допустимых решений.

Одним из случаев графического решения оптимизационной задачи является поиск на геодезическом плане точки с максимальной отметкой в пределах оконтуренного участка. В этом случае площадь участка – это множество допустимых решений, а горизонтали – линии уровня целевой функции. Полный процесс графического решения отличается от этой задачи тем, что включает в себя еще и построение "плана участка" с "горизонталями".

Так как графический метод будет использоваться далее только как иллюстрация, рассмотрим его на примерах, без строгих определений и доказательств.

Пусть имеется модель

$$W = x_1 + 3x_2 - 6 \to \max,$$

$$3x_1 + 2x_2 \leqslant 18,$$

$$(x_1 - 3)^2 + (x_2 - 2)^2 \leqslant 16,$$

$$x_1 \geqslant 0; \quad x_2 \geqslant 0.$$

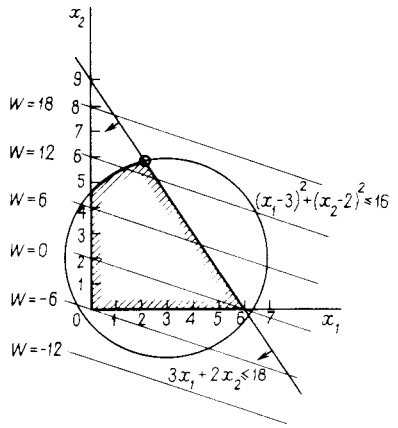
В прямоугольной системе координат $x_1\ 0\ x_2$ последовательно построим области решений ограничений-неравенств.

Из аналитической геометрии известно, что решения уравнения $F(x_1, x_2) = c$ в прямоугольной системе координат $x_1\ 0\ x_2$ образуют плоскую линию, разделяющую множества решений неравенств $F(x_1, x_2) \geqslant c$ и $F(x_1, x_2) \leqslant c$. То есть линия отсекает часть плоскости, на которой лежат точки, представляющие собой решения соответствующего неравенства. Образованная ими область может быть сплошной или разрывной в зависимости от свойств функции $F(x_1, x_2) = c$. При графическом решении задач с нелинейными ограничениями область решений неравенств можно определить, построив линию разделяющих значений $F(x_1, x_2) = c$, и затем семейство линий $F(x_1, x_2) = c - \Delta c \cdot\ \Delta c > 0$, покрывающих область предположительных значений переменных x_1 и x_2.

Для построения области решения линейных неравенств можно воспользоваться их известными свойствами. Решения уравнения $ax_1 + bx_2 = c$ в прямоугольной системе координат образуют прямую линию. Эта прямая делит плоскость на две полуплоскости. Одна из этих полуплоскостей состоит из множества точек, являющихся решениями неравенства $ax_1 + bx_2 - c \geqslant 0$, а вторая – решениями неравенства $ax_1 + bx_2 - c \leqslant 0$. Чтобы определить, является ли данная полуплоскость решением неравенства, необходимо координаты одной из точек этой полуплоскости, не лежащей на разделяющей прямой, подставить в неравенство. Если оно удовлетворяется, то данная полуплоскость составляет множество решений неравенства и, наоборот. Обычно в неравенство подставляют точку (0; 0) – начало координат. На рис. II.1 область решений неравенства $3x_1 + 2x_2 \leqslant 18$ обозначена стрелками.

Решениями второго ограничения-неравенства $(x_1 - 3)^2 + (x_2 - 2)^2 \leqslant 16$ будут точки, лежащие внутри окружности, заданной уравнением

Рис. II.1. Графическое решение оптимизационной задачи



$(x_1 - 3)^2 + (x_2 - 2)^2 = 16$. Это вытекает из известных свойств окружности, заданной уравнением $(x_1 - a)^2 + (x_2 - b)^2 = R^2$. Центр этой окружности лежит в точке (a, b), а радиус равен R. Очевидно, что при постоянных значениях параметров a и b решения неравенства $(x_1 - a)^2 + (x_2 - b)^2 < R^2$ будут лежать на концентрических окружностях радиуса меньшего R. Последние два ограничения $x_1 \geqslant 0$ и $x_2 \geqslant 0$ определяют положительный квадрант системы координат.

Общая область решений всех ограничений составляет множество допустимых решений задачи. На рис. II.1 эта область заштрихована.

Следующий этап решения – нанесение линий равного уровня целевой функции. Для этого вычисляется значение целевой функции W_0 от какого-либо допустимого решения $x_{1,0}$ и $x_{2,0}$ и строится линия $f(x_1, x_2) = W_0$; затем последовательно наносятся линии $f(x_1, x_2) = W_0 + \Delta$; $f(x_1, x_2) = W_0 + 2\Delta$ и линии $f(x_1, x_2) = W_0 - \Delta$; $f(x_1, x_2) = W_0 - 2\Delta$ и т.д., пока ими не будет покрыта вся область допустимых решений. Шаг приращения целевой функции выбирается таким, чтобы картина ее изменения на области допустимых решений стала очевидной. Оптимальное решение находится по чертежу и уточняется аналитическим расчетом, если оно лежит в точке пересечений или касания линий.

В рассматриваемом примере целевая функция линейна и линии равного уровня представляют собой ряд параллельных прямых. Из рис. II.1 видно, что оптимальное решение задачи лежит в точке пересечения прямой и окружности, ограничивающих область допустимых значений. Координаты этой точки могут быть определены графически или путем решения системы уравнений соответствующих линий

$$\begin{cases} 3x_1 + 2x_2 = 18, \\ (x_1 - 3)^2 + (x_2 - 2)^2 = 16. \end{cases}$$

Эта система может быть решена методом последовательных исключений.

1-й шаг – исключается переменная x_1

$$x_1 = \frac{18 - 2x_2}{3},$$

$$\left(\frac{18 - 2x_2}{3} - 3\right)^2 + (x_2 - 2)^2 = 16,$$

$$13x_2^2 - 72x_2 - 27 = 0.$$

2-й шаг – решается последнее уравнение относительно x_2

$$x_{2\,(1,2)} = \frac{72 \pm \sqrt{72^2 + 4 \cdot 13 \cdot 27}}{2 \cdot 13},$$

$$x_{2\,(1)} = 5{,}9, \quad x_{2\,(2)} = -0{,}35.$$

3-й шаг – вычисляются соответствующие значения переменной x_1 подстановкой в первое уравнение

$$\begin{cases} x_{2\,(1)} = 5{,}9, \\ x_{1\,(1)} = 2{,}1, \end{cases} \qquad \begin{cases} x_{2\,(2)} = -0{,}35, \\ x_{1\,(2)} = 6{,}2, \end{cases}$$

так как по условиям исходной задачи переменные $x_1 \geqslant 0$ и $x_2 \geqslant 0$, то второе решение отбрасывается.
Отсюда

$$W_{max} = 2{,}1 + 3 \cdot 5{,}9 - 6 = 13{,}8$$

при $x_1 = 2{,}1, \quad x_2 = 5{,}9$.

§ 3. СТРАТЕГИЯ ПОИСКА ОПТИМУМА

Перед чтением этого параграфа рекомендуем повторить следующие определения из курсов математического анализа и линейной алгебры: экстремум функции нескольких переменных; матрица и вектор; n-мерное евклидово пространство; точка в n-мерном пространстве.

Прежде чем обсуждать понятие стратегии оптимизации, рассмотрим пример решения конкретной задачи.

Для увеличения производительности вскрышного комплекса на 1,5 млн. м^3 в год карьеру требуется приобрести дополнительное количество большегрузных автосамосвалов одного из трех типов БелАЗ-548, БелАЗ-549 или БелАЗ-75191 грузоподъемностью 40; 75 и 110 т соответственно. На эти цели может быть израсходовано не более 4000 тыс. руб. В настоящее время на карьере уже эксплуатируются автосамосвалы грузоподъемностью 110 т, поэтому дополнительных затрат на расширение дорог и рабочих площадок в любом случае не требуется и выбор

Таблица II.1

Тип автосамосвала	Цена $Ц$, руб.	Удельные затраты на транспортирование вскрыши c, руб/тыс. м^3	Среднегодовая производительность P, тыс. м^3
БелАЗ-548	32815	1800	34
БелАЗ-549	167862	1440	65
БелАЗ-75191	270212	1200	90

типа автосамосвала может производиться по минимуму затрат на автотранспорт вскрыши. В процессе эксплуатации в условиях данного карьера эти автосамосвалы дали показатели, приведенные в табл. II.1.

Если обозначить количество приобретаемых автосамосвалов j-го типа через x, то оптимизационная модель задачи будет иметь вид

$$Z = \sum_{j=1}^{3} C_j P_j x_j \to \min,$$

$$\sum_{j=1}^{3} P_j x_j \geqslant 1500,$$

$$\sum_{j=1}^{3} Ц_j x_j \leqslant 4 \cdot 10^6,$$

$$x_j \geqslant 0, \quad j = 1, 2, 3.$$

Первое ограничение обеспечивает необходимую производительность парка автосамосвалов, второе учитывает лимит капиталовложений. Эта модель не отражает требование выбора единственного типа машин, но это проще сделать введя это условие в алгоритм решения вместе с условием целочисленности переменных.

Решить эту задачу можно путем перебора трех возможных вариантов приобретения автосамосвалов.

Потребное количество машин каждого типа определяется из первого ограничения и составит: 45 шт. БелАЗ-540; 23 шт. БелАЗ-549; 17 шт. БелАЗ-75191. Проверка требуемых капитальных вложений: $45 \cdot 32815 =$ $= 1448325$ руб; $23 \cdot 167862 = 3860826$ руб.; $17 \cdot 270212 = 4593604$ руб. показывает, что вариант приобретения БелАЗ-75191 слишком дорог.

Среднегодовые затраты на транспортирование дополнительного объема вскрыши в допустимых вариантах составит: БелАЗ-548 – 2700 тыс. руб.; БелАЗ-549 – 2160 тыс. руб. Отсюда допустимым и наиболее экономичным является вариант использования БелАЗ-549 в количестве 23 шт.

Если попытаться разложить решение задачи на характерные этапы, то можно выделить следующие:

составлена оптимизационная модель;

определено правило перебора значений переменных;

задано выражение расчета значений;

определено правило проверки допустимости значений переменных;

выбрано правило проверки оптимальности;

последовательно вычислены значения оптимизируемых переменных, каждый вариант проверен на допустимость и оптимальность.

Такая схема характерна для любого численного метода решения оптимизационной задачи и отличается только конкретным содержанием правил перебора, оценки допустимости и оптимальности значений переменных.

В дальнейшем конкретный набор значений всех оптимизируемых переменных будет называться точкой, а набор значений переменных, удовлетворяющих системе ограничений – решением. Ключевым моментом численного метода является правило перебора точек или решений, которое называется стратегией поиска оптимума.

Использованная в рассмотренном примере стратегия оптимизации – перебор всех возможных вариантов, применима и эффективна только в случае дискретных значений переменных и малом числе вариантов. Например, выбор одного из нескольких типов проектов, технологических схем, видов механизации и т.д. Появление хотя бы одной непрерывной переменной, определенной на отрезке, делает эту стратегию неприемлемой, так как непрерывная переменная заполняет отрезок бесчисленным числом значений и перебрать их все невозможно даже теоретически. Это вызывает необходимость поиска стратегий, позволяющих путем перебора ограниченного числа точек находить оптимальное решение или узкую область его существования. Математическое программирование рассматривает именно такие стратегии частичного перебора.

Проверка только части значений оптимизируемых переменных и целевой функции на допустимость и оптимальность среди бесчисленного множества возможных вариантов ставит два принципиальных вопроса.

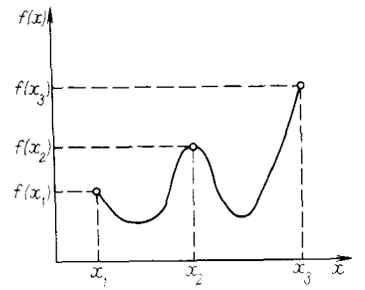
Как определить те точки, среди которых находится оптимум?

Как распознать оптимум среди ряда решений?

Очевидно, что решение этих проблем неразрывно связано с выбором стратегии поиска, так как ответ на первый вопрос определяет возможность использования той или иной стратегии, а ответ на второй вопрос зависит от ее выбора.

Исследование области допустимых значений и свойств целевой функции позволяет в отдельных случаях строго обосновать конечное множество точек, в которых может находиться искомый экстремум, например, задачи линейного программирования, или оно дает направление поиска и анализа возможных осложнений.

Рис. II.2. Глобальный и локальный максимумы функций на отрезке $x_1 \leqslant x \leqslant x_3$



В математическом программировании важную роль играют понятия глобального и локального экстремума функции. (В некоторых книгах используются термины абсолютный и относительный экстремум).

Глобальный максимум. Говорят, что функция $f(x)$, определенная на замкнутом множестве X, достигает глобального максимума в точке x^*, если $f(x) \leqslant f(x^*)$ для любой точки $x \in X$.

Локальный максимум. Пусть функция $f(x)$ определена во всех точках δ – окрестности точки x_0. Тогда говорят, что $f(x)$ имеет локальный максимум в точке x_0, если для любой точки x из δ – окрестности точки x_0 выполняется неравенство $f(x) \leqslant f(x_0)$.

Аналогично определяются глобальный и локальный минимумы функции.

Разница между глобальным и локальным максимумами показана на рис. II.2. Здесь в точках x_1, x_2, x_3 функция достигает локальных максимумов, а в точке x_3 – глобального максимума.

Следует отметить, что глобальный экстремум совпадает с одним из локальных экстремумов.

В задачах математического программирования предусматривается поиск глобального экстремума, поэтому возможное наличие нескольких локальных экстремумов резко усложняет задачу.

В дальнейшем вопрос многоэкстремальности (мультимодальности) будет особо рассматриваться при описании методов нелинейного программирования.

Все стратегии оптимизации можно разделить на две принципиально отличные группы – активные и пассивные.

Пассивные стратегии до начала расчетов определяют все точки, подлежащие анализу на допустимость и оптимальность.

Активные стратегии предусматривают выбор варианта перехода в новую точку в зависимости от ранее полученного результата. Эти стратегии еще называют последовательным поиском.

С вычислительной точки зрения последовательные стратегии более эффективны, но требуют большего теоретического обоснования и легче подвержены случайным ошибкам, так как ошибка на одном шаге может привести к искажению результатов всех последующих действий. Однако при большой размерности задач последовательный поиск – единственно

возможный путь достижения результата и поэтому преобладает в математических методах.

Стратегия поиска полностью определяет конечность и сходимость алгоритма оптимизации.

Для задач математического программирования конечный алгоритм решения должен обеспечивать отыскание глобального оптимума или указывать на его отсутствие за конечное число шагов (итераций). Но, к сожалению, в настоящее время стратегии оптимизации, гарантирующие получение глобального оптимума с точностью вычислений, разработаны только для ограниченного круга задач. Многие стратегии позволяют получить окрестность глобального оптимума, сужающуюся в зависимости от числа итераций. Хотя теоретически эти стратегии не могут дать конечного алгоритма их практическое использование позволяет сравнительно быстро получать глобальный оптимум с необходимой точностью.

Хуже обстоит дело, когда стратегия гарантирует достижение оптимума, но не гарантирует что он глобальный. Этот недостаток может присутствовать изначально или быть следствием применения известной стратегии в неполностью удовлетворяемых исходных условиях. Такие стратегии можно считать конечными только условно.

Сходимость алгоритма характеризуется числом шагов, за которое получается результат, при прочих равных условиях (точность, устойчивость от ошибок и т.д.). Стратегия поиска прежде всего влияет на этот показатель алгоритма. Практически многообразие методов решения отдельных задач порождено стремлением повысить сходимость алгоритма и за счет этого снизить объем вычислений до реально выполнимого уровня.

При выборе стратегии все же следует отдавать предпочтение конечным алгоритмам, если их сходимость практически позволяет решить задачу имеющимися вычислительными средствами.

§ 4. КРАТКАЯ ХАРАКТЕРИСТИКА ЗАДАЧ, МЕТОДОВ И ПРОГРАММНЫХ СРЕДСТВ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Математическое программирование развивается довольно бурно и в настоящее время не существует достаточно полной и обоснованной с нашей точки зрения классификации задач и методов, хотя ряд разделов приобрели устойчивое название и рассматривают вполне определенный круг моделей. Здесь мы также не будем делать попытку создать подробный классификатор, а используя устоявшуюся терминологию и введенные понятия попытаемся составить краткий указатель, с помощью которого можно будет распознать тип задачи и предположительный метод поиска экстремума.

Первая группа задач, которую часто даже не относят к математическому программированию, – это поиск экстремума функции одной или

нескольких переменных из области допустимых значений, заданной интервалами изменения переменных.

Нередко в литературе эти задачи называют оптимизацией без ограничений, но тем не менее рассматривают методы численного поиска экстремума на ограниченном интервале изменения переменных.

Такие задачи чаще всего возникают при оптимальном управлении техническими системами.

В зависимости от характера целевой функции здесь используются следующие группы методов поиска экстремума:

целевая функция аналитически определена – детермированный поиск методом исключения интервалов, градиентным методом и т.д.;

целевая функция статистически определена или имеются случайные ошибки измерений – стохастические методы поиска экстремума;

целевая функция не определена и оптимум ищется на основе активного эксперимента – статистические методы планирования экстремальных экспериментов.

Вторая группа – задачи с линейной целевой функцией и линейными ограничениями в виде уравнений и неравенств. Решаются методами линейного программирования, обеспечивающего конечность и быструю сходимость алгоритмов. Кроме универсальных, в линейном программировании имеются и специализированные алгоритмы, позволяющие решать линейные задачи определенного типа большей размерности.

Третья группа – задачи с линейной целевой функцией и линейными ограничениями, уравнениями или неравенствами, и дополнительными условиями целочисленности всех или части переменных. Эти задачи рассматриваются в целочисленном программировании. В настоящее время есть методы теоретически конечных алгоритмов поиска целочисленных решений, но практически их сходимость крайне низка из-за погрешностей вычислений. Кроме этого сейчас разрабатывается ряд более эффективных алгоритмов, но их можно считать конечными только условно из-за недостаточного теоретического обоснования.

Четвертая группа – задачи с нелинейными дифференцируемыми на всем интервале изменения переменных целевой функцией и ограничениями в виде уравнений. Оптимальное решение находится с помощью неопределенных множителей Лагранжа. Этот метод часто относят к способам классической оптимизации. Он дает конечный алгоритм решения, но его применение в задачах большой размерности вызывает значительные трудности.

Пятая группа – задачи с квадратичной целевой функцией и линейными ограничениями. Эти задачи решаются с помощью конечных алгоритмов квадратичного программирования.

Шестая группа – задачи с сепарабельной целевой функцией и сепарабельными функциями в ограничениях-равенствах или неравенствах (сепарабельная функция представляет собой сумму функций одной переменной). Эти задачи решаются приближенными методами с помощью

кусочно-линейной аппроксимации функции. В большинстве случаев получение глобального оптимума метод не гарантирует.

Седьмая группа – задачи с нелинейной целевой функцией, линейными или нелинейными ограничениями. Многие, но далеко не все задачи этой группы решаются градиентными методами. Алгоритмы поиска оптимума бесконечны и в большинстве случаев обеспечивают поиск локального оптимума.

Восьмая группа – задачи с сепарабельной целевой функцией и сепарабельными односторонними или двусторонними ограничениями, целочисленными или непрерывными переменными. Решаются методами динамического программирования. Алгоритмы конечны и обеспечивают решение на ЭВМ задач весьма большой размерности.

Приведенная группировка задач и методов весьма условна, так как многие задачи могут быть отнесены к нескольким группам и решаться разными методами. Особо отметим использование динамического программирования. В приведенной классификации применение этих методов явно рекомендуется только для решения одной группы задач, но они могут быть нередко использованы при оптимизации на сетях и графах в тех случаях, когда отдельные переменные могут последовательно менять свои значения на отдельных шагах решения.

Для использования оптимизационных математических моделей реальной размерности, как правило, необходимо численное решение задач математического программирования на ЭВМ, а их распространенность потребовала создания универсальных программных средств для решения различных классов задач.

Наиболее часто на практике встречаются задачи линейного программирования, для них имеется широкий выбор пакетов прикладных программ [ППП]. В первую очередь следует назвать системы линейного программирования ПМП, ЛП АСУ и СПО МПР-2, поставляемые НПО "Центропрограммсистем". Они позволяют решать на ЕС ЭВМ задачи с несколькими тысячами ограничений и тысячами переменных. В настоящее время все чаще ставится вопрос использования этих пакетов не только для решения конкретных задач линейного программирования, но и для построения на их базе специализированных систем оптимального управления. Для построения таких систем эти пакеты имеют специальный управляющий язык, обладающий значительными возможностями по обработке арифметических и текстовых данных, управлению ходом вычислительного процесса, вызову программ пользователя, написанных на любых языках программирования ЕС ЭВМ. В ПМП и ЛП АСУ – это так называемый управляющий язык системы математического программирования (MPSCL). СПО МПР-2 наряду с MPSCL располагает управляющим языком (ECL), представляющим собой некоторое расширение многоцелевого языка программирования ПЛ/1. В ПМП используется также система обработки данных линейного программирования "МАТФОР", в ЛП АСУ и СПО МПР-2 - пакеты "Генератор матриц и отчетов" (ГМО) и "САПФИР-2".

Еще раньше были разработаны системы линейного программирования для ЭВМ БЭСМ-6 – пакет "ОПТИМА-2" (Институт математики и механики Уральского отделения АН СССР) и система ЛП БЭСМ-6 (Вычислительный центр АН СССР).

Более широкими возможностями обладают пакеты "ОМЕГА" (ВНИИ системных исследований АН СССР) и пакет анализа оптимизационных экономических моделей (ПАОЭМ ЕС, разработанный в Центральном экономико-математическом институте АН СССР). Последний содержит модули решения задач трех типов: линейного и квадратичного программирования; нелинейного программирования (минимизации при двусторонних ограничениях на значения каждой переменной, задачи геометрического программирования и задачи минимизации при ограничениях-равенствах); транспортных задач (матричных сетевых и потоковых).

Для задач с непрерывными переменными наиболее широкими возможностями обладает диалоговая система оптимизации (ДИСО), разработанная в Вычислительном центре АН СССР. Существуют версии этой системы для БЭСМ-6, СМ ЭВМ и персональных ЭВМ типа IBM PC XT/AT. Аналогичная система для ЕС ЭВМ (ДИСО–ЕС) создана в Физико-техническом институте низких температур АН УССР. Система ДИСО имеет подсистемы линейного программирования, безусловной оптимизации, нелинейного программирования (локальные методы) и глобальной оптимизации (для задач невысокой размерности), а также многокритериальной оптимизации и оптимального управления. Система рассчитана на задачи средней размерности, зато предоставляет пользователю широкий сервис – управление ходом решения задачи с желательной подробностью, наглядное представление промежуточных результатов в числовом и графическом виде и даже возможность в диалоге описать условия нелинейной задачи средствами простого языка описания формул ("ДИФАЛГ").

Для задач с дискретными переменными также разработаны ППП. Наиболее распространенной возможностью ППП является решение целочисленных и частично-целочисленных задач линейного программирования. Кроме того, в Институте кибернетики АН УССР разработаны пакеты "ВЕКТОР-2" для решения задач комбинаторной оптимизации и "ДИСПРО" – для задач дискретного программирования; дальнейшим развитием последнего является ППП "ПЛАНЕР", имеющий также возможности решения задач математического программирования блочной структуры.

Контрольные вопросы и задачи

1. Как в общем виде формируется решение задачи математического программирования?
2. Что представляет собой оптимальное решение?
3. В чем заключается особенность поиска условного экстремума функции?
4. Почему методы классической оптимизации находят ограниченное применение в решении задач математического программирования?
5. Что такое конечный и бесконечный алгоритм поиска решения?
6. Что такое стратегия поиска оптимального решения?
7. От чего в первую очередь зависит сходимость алгоритма решения?

Глава III

РЕШЕНИЕ ОСНОВНОЙ ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

В этой и последующих главах теории линейного программирования будет уделено большое внимание прежде всего потому, что на его основе строятся некоторые другие математические методы, например, целочисленное, квадратическое программирование, а также в связи с тем, что знакомство с его теорией позволит понять логику построения других методов оптимизации. С другой стороны, учитывая прикладной характер данного курса и наличие большого числа стандартных программ для ЭВМ, приводится только простейший вариант симплекс-алгоритма как иллюстрация практической реализации оптимизационного метода.

Следуя схеме теоретического обоснования, описанной в предыдущей главе, здесь будет дана формальная постановка основной задачи линейного программирования и рассмотрены способы приведения реальных моделей к каноническому виду. На основе исследования свойств множества решений и целевой функции определены точки, в которых может находиться оптимальное решение, обоснована стратегия поиска и составлен алгоритм оптимизации.

§ 1. ОСНОВНАЯ ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

В каноническом виде основная задача линейного программирования (ЛП) формулируется следующим образом:

найти неотрицательное решение системы ограничений

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i, \quad (i = 1, 2, \ldots, m), \tag{III.1}$$

$$x_j \geqslant 0, \quad (j = 1, 2, \ldots, n), \tag{III.2}$$

обеспечивающее максимум (минимум) целевой функции

$$Z = c_1x_1 + c_2x_2 + \ldots + c_nx_n + q \to \max. \tag{III.3}$$

Кроме приведенной формы записи далее будет использоваться частично развернутая

$$Z = \sum_{j=1}^{n} a_jx_j + q \to \max,$$

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

$$x_j \geqslant 0, \quad (j = 1, 2, \ldots, n)$$

и матричная формы

$$Z = cx + q \to \max,$$

$$Ax = b$$

$$x \geqslant 0.$$

Все дальнейшие рассуждения будут проводиться только для основной задачи ЛП в каноническом виде.

Если исходная задача имеет систему ограничений неравенств или смешанную систему ограничений, то она приводится к эквивалентной системе уравнений путем введения в каждое неравенство вспомогательной, неотрицательной переменной со знаком плюс или минус в зависимости от знака отношения. Так, для системы неравенств

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \leqslant b_1,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \leqslant b_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n \leqslant b_r,\\
&x_j \geqslant 0, \quad (j = 1, 2, \ldots, n)
\end{aligned} \tag{III.4}$$

эквивалентной будет система

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + y_1 = b_1,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + y_2 = b_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned} \tag{III.5}$$

$$a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n + y_r = b_r,$$

$$x_j \geqslant 0, \quad (j = 1, 2, \ldots, n), \quad y_j \geqslant 0, \quad (i = 1, 2, \ldots, r),$$

а для системы

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \geqslant b_1, \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \geqslant b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n \geqslant b_r, \\ &x_j \geqslant 0, \quad (j = 1, 2, \ldots, n) \end{aligned} \qquad \text{(III.6)}$$

эквивалентной будет система

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n - y_1 = b_1, \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n - y_2 = b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n - y_r = b_r, \\ &x_j \geqslant 0, \quad (j = 1, 2, \ldots, n), \\ &y_i \geqslant 0, \quad (i = 1, 2, \ldots, r). \end{aligned} \qquad \text{(III.7)}$$

Доказательство эквивалентности замены системы можно найти, например, в работе [9].

Чтобы вспомогательные переменные не повлияли на искомый оптимум, в целевую функцию их вводят с нулевыми коэффициентами.

Минимизация формы Z. В дальнейшем будет рассматриваться только задача максимизации формы Z. Если же необходимо решить задачу минимизации линейной формы, коэффициенты целевой функции следует умножить на (-1) и решать эту новую задачу на максимум. Искомый минимум целевой функции получается умножением найденного максимального значения на (-1), т.е.

$$Z \min = -\max(-Z). \qquad \text{(III.8)}$$

Особо отметим требование неотрицательности всех переменных в основной задаче ЛП. Оно играет ведущую роль в определении точек возможного оптимума и исключение этого условия делает симплекс-метод неприемлемым. При решении экономических и технических задач, как правило, требование неотрицательности соответствует физической природе оптимизируемых переменных и оно естественно входит

в математическую модель. Если же в задаче какая-либо переменная по своей природе может принимать отрицательные значения, то в большинстве случаев изменение формулировки условий позволяет избавиться от неограниченных по знаку переменных.

В тех случаях, когда изменение формулировки условий невозможно или нежелательно, избавиться от неограниченности по знаку переменных можно следующим способом. Каждая неограниченная переменная исходной задачи заменяется разностью двух неотрицательных вспомогательных переменных

$$x_i = u_i - v_i; \quad u_i \geqslant 0; \quad v_i \geqslant 0.$$

Пример III.1. Приведение к каноническому виду задачи ЛП:

$$\begin{aligned} &z = x_1 - 3x_2 + 2x_3 \to \min, \\ &x_1 + x_2 - x_3 \leqslant 7, \\ &x_1 - x_2 + x_3 = 2, \\ &3x_1 - x_2 - x_3 = 5, \\ &x_1 \geqslant 0; \quad x_2 \geqslant 0, \\ &x_3 \text{ -- не ограничена по знаку.} \end{aligned}$$

1. Заменяем x_3 на $u_3 - v_3$, где $u_3 \geqslant 0, v_3 \geqslant 0$.
2. Введем в ограничения-неравенства искусственные переменные y_1 и y_2.
3. Перейдем к задаче максимизации, умножив целевую функцию на -1.

Приведенная к каноническому виду задача имеет вид:

$$\begin{aligned} &(-Z) = -x_1 - 3x_2 + 2u_3 - 2v_3 \to \max \\ &x_1 + x_2 - u_3 + v_3 + y_1 = 7, \\ &x_1 - x_2 + u_3 - v_3 - y_2 = 2, \\ &3x_1 - x_2 - u_1 + v_3 = 5, \\ &x_1, x_2, u_3, v_3, y_1, y_2 \geqslant 0. \end{aligned}$$

Отметим, что указанный способ избавления от неограниченных переменных ведет к росту размерности задачи.

Другой способ заключается в том, что неограниченные переменные выражаются через неотрицательные и исключаются из системы ограничений, как это делается при решении систем линейных уравнений методом исключений Гаусса. Оптимальное решение находится для модели, в которой остались только неотрицательные переменные. Оптимальное значение неограниченных переменных получают обратной подстановкой. Такой способ осуществим, если ранг системы ограничений больше числа неограниченных переменных.

Этот способ менее "элегантен" с теоретической точки зрения, более трудоемок и усложняет процесс решения, но с практической точки зрения он имеет одно важное преимущество – сокращает размерность оптимизационной модели.

Пример III.2. Приведение к каноническому виду задачи ЛП из примера III.1 способом исключения переменных.

1. В ограничения-неравенства вводятся искусственные неотрицательные переменные y_1 и y_2.

2. Переходим к задаче максимизации, умножив целевую функцию на -1.

3. Первое уравнение решается относительно x_3

$$x_3 = x_1 + x_2 + y_1.$$

4. Это решение подставляется в целевую функцию и остальные ограничения, после приведения подобных членов оптимизационная модель примет вид

$$\begin{aligned}(-Z) &= x_1 - x_2 + 2y_1 \to \max,\\ 2x_1 &\quad + y_1 - y_2 = 2,\\ 2x_1 &- 2x_2 - y_1 = 5,\\ x_1, x_2, y_1, y_2 &\geqslant 0.\end{aligned}$$

§ 2. СВОЙСТВА МНОЖЕСТВА РЕШЕНИЙ

Исследование свойств множества решений должно дать ответ на вопросы о том, в каких точках может находиться оптимум и какие исходы возможны при решении основной задачи ЛП.

Оптимальное решение основной задачи ЛП представляет собой одно из неотрицательных решений системы линейных уравнений (III.1), при поиске которых могут возникнуть следующие случаи:

система уравнений несовместна, т.е. не имеет ни одного решения;

система имеет единственное решение, но хотя бы одна переменная имеет отрицательное значение;

система имеет единственное неотрицательное решение;

имеет бесчисленное множество неотрицательных решений.

Несовместность системы – вполне возможный результат моделирования, поэтому этот случай будет подробно изучен в последующих главах.

Случай единственного неотрицательного решения не представляет особого интереса, так как здесь задача поиска максимума перестает существовать.

Наибольший интерес представляет случай бесчисленного множества решений, возникающий тогда, когда число линейно независимых уравнений системы меньше количества оптимизируемых переменных.

Из линейной алгебры [9] известно, что в этом случае система уравнений (III.1) имеет общее решение

$$\begin{aligned}x_1 &= \alpha_{11}x_{k+1} + \alpha_{12}x_{k+2} + \ldots + \alpha_{1n-k}x_n + \beta_1,\\ x_2 &= \alpha_{21}x_{k+1} + \alpha_{22}x_{k+2} + \ldots + \alpha_{2n-k}x_n + \beta_2,\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ x_k &= \alpha_{k1}x_{k+1} + \alpha_{k2}x_{k+2} + \ldots + \alpha_{kn-k}x_n + \beta_k,\end{aligned} \qquad \text{(III.9)}$$

в котором базисные переменные $x_1, x_2, \ldots, x_k$ соответствуют базисным, линейно независимым столбцам матрицы системы A, а независимые переменные – $x_{k+1}, x_{k+2}, \ldots, x_n$ соответствуют остальным линейно зависимым столбцам матрицы A. Число базисных переменных в решении системы равно рангу матрицы A, т.е. $k = r(A)$.

Произвольно задавая значения независимым переменным $x_{k+1}, x_{k+2}, \ldots, x_n$ и вычисляя соответствующие значения базисных переменных можно получить бесчисленное множество решений, среди которых могут быть такие, что все $x_1, x_2, \ldots, x_n \geqslant 0$. Именно они образуют множество решений основной задачи ЛП.

Среди решений системы (III.1) особо выделяются так называемые базисные, которые получаются путем приравнивания независимых переменных нулю, т.е.

$$x_{k+1}, x_{k+2}, \ldots, x_n = 0; \quad x_1 = \beta_1; \quad x_2 = \beta_2; \ \ldots; \ x_k = \beta_k.$$

Отметим, что система (III.1) может иметь несколько общих и соответствующих им базисных решений. Так, если в решении (III.9) одну из базисных переменных перенести в правую часть, а на ее место влево вынести одну независимую переменную, то получится новое общее и базисное решение. Сделав еще одну замену, можно получить еще одно решение и т.д.

Пример III.3. Рассмотрим варианты базисных решений системы уравнений, полученной в примере III.2.

Система

$$\begin{aligned}2x_1 &\quad + y_1 - y_2 = 2;\\ 2x_1 &- 2x_2 - y_1 = 5\end{aligned}$$

легко решается относительно переменных

$$\begin{aligned}y_2 &= 2x_1 + y_1 - 2,\\ x_2 &= x_1 - 1/2\, y_1 - 5,\end{aligned}$$

этому общему решению соответствует базисное

$$x_1 = 0; \quad y_1 = 0; \quad y_2 = -2; \quad x_2 = -5.$$

Новое общее решение можно получить заменив, например, y_2 на x_1. Для этого надо решить первое уравнение относительно x_1 и подставить его во второе. В результате новое решение будет иметь вид

$$\begin{aligned}x_1 &= 1/2\, y_2 - 1/2\, y_1 + 1,\\ x_2 &= 1/2\, y_2 + y_1 - 4,\end{aligned}$$

а базисное решение будет

$$y_1 = 0; \quad y_2 = 0; \quad x_1 = 1; \quad x_2 = -4$$

Следующее решение можно получить из предыдущего, заменив, например, x_1 на y_1 и т.д.

Не обсуждая пока способов получения общего решения, будем считать, что исходная задача имела систему из m ограничений-неравенств вида (III.4), была приведена к эквивалентной (III.5), решение которой относительно вспомогательных переменных находится совершенно очевидно

$$\begin{aligned} y_1 &= -a_{11}x_1 - a_{12}x_2 \ldots - a_{1n}x_n + b_1, \\ y_2 &= -a_{21}x_1 - a_{22}x_2 \ldots - a_{2n}x_n + b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= -a_{m1}x_1 - a_{m2}x_2 \ldots - a_{mn}x_n + b_m. \end{aligned} \qquad \text{(III.10)}$$

Если обозначить коэффициенты $-a_{ij}$ через α_{ij} и учесть неотрицательность всех переменных, то множество допустимых решений основной задачи ЛП будет описываться системой линейных неравенств

$$\begin{aligned} y_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 \ldots + \alpha_{1n}x_n + b_1 \geqslant 0, \\ y_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 \ldots + \alpha_{2n}x_n + b_2 \geqslant 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= \alpha_{m1}x_1 + \alpha_{m2}x_2 \ldots + \alpha_{mn}x_n + b_m \geqslant 0, \\ x_1, x_2, \ldots, x_n &\geqslant 0. \end{aligned} \qquad \text{(III.11)}$$

Каждое решение этой системы представляет собой набор конкретных значений переменных $x_1, x_2, \ldots, x_n$, который можно рассматривать как соответствующие компоненты n-мерного вектора x из n-мерного евклидова пространства E^n, т.е.

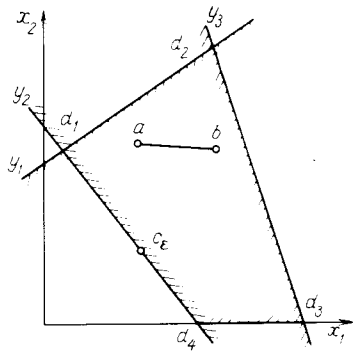
$$x = [x_1, x_2, \ldots, x_n].$$

Геометрически решение представляет собой точку в пространстве с прямоугольной системой координат с осями, соответствующими независимым переменным $x_1, x_2, \ldots, x_n$. Все решения системы (III.11) образуют точечное множество X. Так как решения задачи ЛП могут содержать только неотрицательные значения переменных, то множество X может располагаться только в положительном квадранте пространства E^n и образовывать некоторую геометрическую фигуру. Важнейшим свойством множества решений является его выпуклость, определяющая условия существования экстремума.

Выпуклое множество. Множество точек называется выпуклым, если вместе с любыми двумя точками, принадлежащими этому множеству, оно содержит и отрезок их соединяющий.

Из аналитической геометрии известно, что прямая, проходящая через две точки x_1 и x_2, есть геометрическое место точек a – линейных комбинаций x_1 и x_2, т.е.

Рис.III.1. Многогранник решений основной задачи ЛП



$$a = (1 - \lambda)x_1 + \lambda x_2.$$

Из понятия выпуклой линейной комбинации вытекает аналитическое определение выпуклого множества.

Множество точек $x \in E^m$ называется выпуклым, если вместе с любыми точками x_1 и $x_2 \in X$ оно содержит и их произвольную выпуклую линейную комбинацию $x = (1 - \lambda)x_1 + \lambda x_2$, $0 \leqslant \lambda \leqslant 1$.

Пересечением множеств называется множество точек X_3, принадлежащих одновременно множествам X_1 и X_2. Геометрически это выглядит как общая область двух пересекающихся фигур (рис. III.1). При этом справедливо утверждение, что пересечение выпуклых множеств также является выпуклым множеством.

Множество решений основной задачи ЛП, заданное (III.11), представляет собой пересечение выпуклых множеств решений отдельных неравенств [9], из чего следует очень важный вывод: множество допустимых решений основной задачи линейного программирования выпукло.

Графически множество решений основной задачи ЛП выглядит как выпуклый многогранник, образованный пересечением полуплоскостей – решениями отдельных неравенств (см. рис. III.1).

Среди множества точек многогранника решений следует различать точки, лежащие на границе и в вершинах многогранника. Аналитически эти точки могут быть определены следующим образом.

Точка множества является граничной, если шар сколь угодно малого радиуса ϵ, построенный вокруг этой точки, содержит как точки, принадлежащие множеству, так и точки, не принадлежащие ему. Например, точка c_ϵ на рис. III.1.

Точка выпуклого множества является крайней точкой, если ее нельзя выразить в виде выпуклой линейной комбинации двух других произвольных точек. Например, крайними точками являются вершины d_1, d_2, d_3 и d_4 многогранника решений, изображенного на рис. III.1.

Аналитически крайние точки множества решений задачи ЛП могут быть определены на основании теоремы о том, что неотрицательное

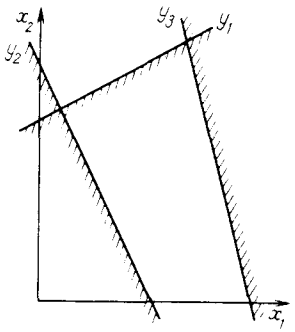


Рис. III. 2. Множество решений пусто

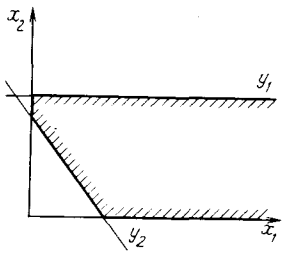


Рис. III. 3. Неограниченное множество решений

базисное решение системы линейных ограничений (III.1), (III.2) есть крайняя точка множества решений основной задачи ЛП.

Особо отметим, что данная теорема справедлива только в случае наличия условия неотрицательности всех переменных.

Рассматривая графическое изображение области допустимых решений задачи ЛП, можно легко представить случаи несовместности системы ограничений и неограниченного множества решений.

Случай несовместности (рис. III.2) возникает тогда, когда нет общей области пересечения полуплоскостей или оно попадает в отрицательный квадрант. В этом случае говорят, что множество решений пусто.

По определению множество называется неограниченным, если нельзя построить сферу конечного радиуса из какой-либо точки множества, включающую все это множество.

Графически неограниченное множество представляет собой незамкнутый многоугольник (рис. III.3). Этот рисунок также показывает, что для неограниченности множества решений достаточно наличия возможности бесконечного роста хотя бы одной независимой переменной.

Если теперь подвести некоторый итог, то можно сделать следующие выводы:

множество решений основной задачи ЛП, если оно не пусто, представляет собой замкнутое или неограниченное выпуклое множество;

крайними точками множества решений являются неотрицательные базисные решения системы ограничений.

§ 3. СТРАТЕГИЯ ПОИСКА ОПТИМАЛЬНОГО РЕШЕНИЯ СИМПЛЕКС-МЕТОДОМ

Выбор стратегии целиком зависит от того, до какой степени определены точки возможного оптимума среди множества решений задачи. В данном случае путь к теоретическому исследованию этой проблемы

становится более очевидным, если рассмотреть пример графического решения задачи ЛП.

Пример III.4. Графическое решение задачи ЛП.

$$Z = x_1 + x_2 + 1 \to \max,$$
$$5x_1 + x_2 \geqslant 5,$$
$$3x_1 + 2x_2 \geqslant 6,$$
$$x_1 - \frac{1}{2} x_2 \leqslant 3,$$
$$x_1 + 2x_2 \leqslant 8,$$
$$x_1,\ x_2 \geqslant 0.$$

Приведем задачу к каноническому виду, введя в каждое неравенство вспомогательную неотрицательную переменную и решив полученную систему уравнений. После этого исходная задача будет иметь вид

$$Z = x_1 + x_2 + 1 \to \max,$$
$$y_1 = 5x_1 + x_2 - 5 \geqslant 0,$$
$$y_2 = 3x_1 + 2x_2 - 6 \geqslant 0,$$
$$y_3 = -x_1 + \frac{1}{2} x_2 + 3 \geqslant 0,$$
$$y_4 = -x_1 - 2x_2 + 8 \geqslant 0,$$
$$x_1,\ x_2 \geqslant 0.$$

В положительном октате системы координат x_1, 0, x_2 построим полуплоскости, соответствующие решениям неравенств, и выделим область их пересечения.

Построим линии уровня целевой функции, покрывающие полученный многогранник решений.

Из рис. III.4 видно, что целевая функция достигает максимального значения $Z = 7$ в точке пересечения линий y_3 и y_4, т.е. в вершине многогранника решений.

Из обобщения геометрического решения задачи ЛП на многомерный случай естественно вытекает важнейшая для численного решения теорема о том, что линейная функция задачи линейного программирования достигает своего единственного максимального значения в крайней точке множества решений.

Из вышеупомянутых теорем следует очень важный вывод: оптимальное решение основной задачи ЛП следует искать среди допустимых базисных решений системы ограничений.

В дальнейшем допустимое базисное решение основной задачи ЛП будет называться опорным решением.

Если еще раз обратиться к рис. III.4, то может возникнуть мысль, что при определенных значениях коэффициентов целевой функции линии уровня могут стать параллельны ребру y_4 или y_3 и тогда оптимальное решение станет не единственным.

Пример III.5. Графическое решение задачи из примера III.4 с целевой функцией $W = x_1 + 2x_2 + 1 \to \max$ изображено на рис. III 5.

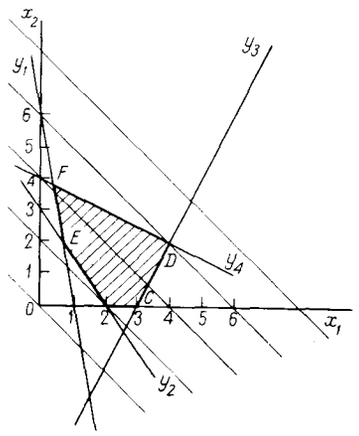


Рис. III.4. Графическое решение задачи из примера III.4

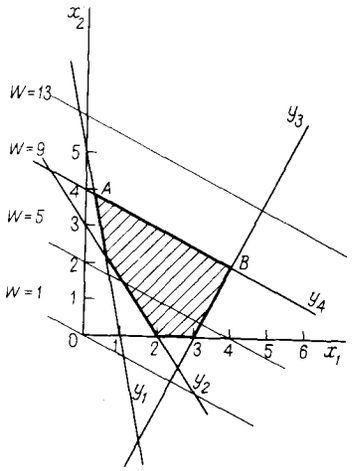


Рис. III.5. Графическое решение задачи из примера III.5

Здесь целевая функция достигает максимального значения $W = 9$ не только в вершинах многоугольника A и B, но и во всех точках ребра AB.

Этот пример показывает, что при решении задачи ЛП может возникнуть случай, когда оптимум достигается в двух и более крайних точках и тогда он достигается и во всех остальных точках, являющихся их выпуклой линейной комбинацией.

В предыдущем параграфе было показано, что множество решений может быть неограниченно, и теперь надо выяснить, как это может отразиться на поиске оптимума целевой функции. Для этого рассмотрим два примера графического решения задач.

Пример III.6. Графическое решение задачи, приведенной к каноническому виду

$$Z = x_1 + x_2 + 1 \to \max,$$

$$y_1 = 5x_1 + x_2 - 5 \geqslant 0,$$

$$y_2 = 3x_1 + 2x_2 - 6 \geqslant 0,$$

$$y_3 = -x_1 + \frac{1}{2} x_2 + 3 \geqslant 0,$$

$$x_1, x_2 \geqslant 0.$$

показано на рис. III.6. Здесь целевая функция не имеет условного максимума, так как многогранник решений незамкнут в направлении роста целевой функции. В этом случае говорят, что целевая функция неограничена.

Пример III.7. На рис. III.7 приведено решение задачи

$$Z' = -x_1 - x_2 + 1 \to \max,$$

$$y_1 = 5x_1 + x_2 - 5 \geqslant 0,$$

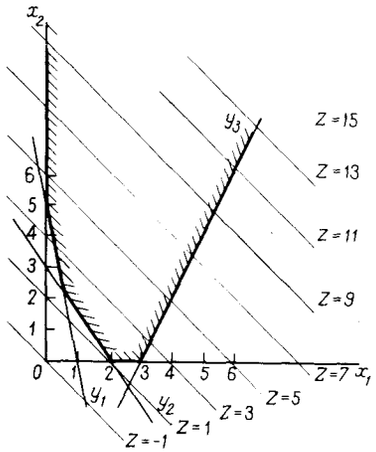


Рис. III.6. Графическое решение примера III.6

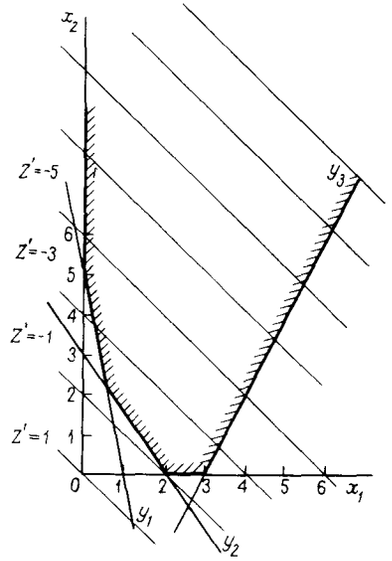


Рис. III.7. Графическое решение примера III.7

$$y_2 = 3x_1 + 2x_2 - 6 \geqslant 0,$$

$$y_3 = -x_1 + \frac{1}{2} x_2 + 3 \geqslant 0,$$

$$x_1, x_2 \geqslant 0.$$

Эта задача отличается от примера III.6 только целевой функцией, но она имеет единственное оптимальное решение $x_1 = 2$; $x_2 = 0$; $Z' = -1$, так как в данном случае многогранник решений замкнут в направлении роста целевой функции.

Если обобщить результаты исследования свойств множества решений основной задачи ЛП, можно сделать следующие выводы:

задача ЛП может иметь несовместную систему ограничений;

оптимальное решение следует искать среди допустимых базисных решений системы ограничений, т.е. опорных решений;

при неограниченном множестве решений целевая функция может бесконечно возрастать;

оптимальное решение может быть неединственным, но при этом максимум целевой функции существует хотя бы в двух опорных решениях.

После того, как полностью определены возможные исходы решения задачи и возможные точки оптимума можно выбрать стратегию поиска.

Если считать, что получение исходного базисного решения и переход от него к другому – задача с вычислительной точки зрения разрешима, то наиболее простой стратегией поиска оптимума является сплош-

ной перебор всех базисных решений с проверкой их опорности и сравнением значений целевой функции.

Одно решение от другого отличается набором базисных переменных, поэтому их количество определяется числом сочетаний из общего числа переменных n по числу базисных переменных k –

$$C_n^k = \frac{n!}{k\,(n-k)\,!}.$$

В реальных задачах это число становится настолько большим, что сплошной перебор практически неосуществим даже на ЭВМ большой мощности.

Каким же образом сократить число перебираемых решений?

Оптимальное решение может находиться только среди опорных, поэтому повысить эффективность поиска можно перебирая лишь допустимые базисные решения. Допустимые базисные решения могут содержать только неотрицательные переменные, следовательно алгоритм поиска должен отвечать условию:

при переходе от одного решения к другому должна сохраняться неотрицательность переменных.

Еще больше ускорит поиск переход от сплошного к целенаправленному перебору опорных решений. Конечной целью является получение опорного решения с максимальным значением целевой функции, отсюда естественно вытекает еще одно условие стратегии поиска;

при переходе от одного опорного решения к другому должен обеспечиваться рост целевой функции.

Данная стратегия направленного поиска предполагает наличие некоторого исходного опорного решения, поэтому весь процесс решения распадается на три этапа – получение исходного базисного решения, поиск опорного и затем оптимального решения. Конечность алгоритма поиска с применением этой стратегии, например, для случая ограниченного невырожденного множества решений, доказывается следующим образом.

Вырожденным называется базисное решение, содержащее одну (или более) базисную переменную, равную нулю. В этом случае переход от одного решения к другому может не дать изменения целевой функции и условие ее монотонного роста будет нарушено.

Любая система ограничений задачи ЛП имеет конечное число базисных решений, равное C_n^k, где n – общее число переменных, а k – ранг матрицы системы ограничений. Пусть среди этого числа содержится m опорных решений, каждому из которых соответствует единственное значение целевой функции. Множество опорных решений конечно и может быть ранжировано по уровню целевой функции. Так, что:

$$Z(X_1) < Z(X_2), \ldots, < Z(X_k) < Z(X_{k+1}), \ldots, < Z(X_m).$$

Пусть в результате поиска исходного опорного решения получено решение X_k. Тогда двигаясь по направлению роста целевой функции, можно сделать только $(m - k)$ шагов, т.е. конечное число переходов от одного решения к другому. Следовательно, предполагаемая стратегия поиска в указанных условиях конечна.

Поиск опорного решения представляет собой весьма сложную задачу и требует разработки собственной стратегии.

Опорное решение должно содержать неотрицательные значения переменных, поэтому, если в исходном базисном решении какая-либо переменная имеет значение меньше нуля, следует перейти к новому решению. При этом можно использовать различные стратегии, например, сплошной перебор, случайный поиск, но наиболее эффективна стратегия последовательного улучшения решения. Конечной целью алгоритма является получение базисного решения с неотрицательными переменными, поэтому естественным представляется следующее условие перебора решений: если на каком-либо шаге переменная стала неотрицательной, она должна сохранять неотрицательность во всех следующих решениях.

Опираясь на это условие, можно постепенно сделать все переменные неотрицательными или установить несовместимость системы ограничений, если применить следующее правило: чтобы переменная стала неотрицательной, каждый переход к новому базисному решению должен обеспечивать ее монотонный рост.

Конечность алгоритма поиска опорного решения с использованием данной стратегии доказывается так же, как было показано выше.

§ 4. РАСЧЕТ НОВОГО БАЗИСНОГО РЕШЕНИЯ. ЖОРДАНОВЫ ИСКЛЮЧЕНИЯ

Симплекс-метод решения задачи ЛП построен на направленном переборе базисных решений системы ограничений, поэтому разработка численной процедуры вычисления нового решения имеет принципиальное значение.

Решение реальных задач практически не обходится без ЭВМ, поэтому разрабатывая какую-либо численную процедуру следует предусматривать хранение информации в виде таблиц однородных данных, например, матрица числовых коэффициентов, столбец или строка обозначений переменных и т.д., а также заменять развернутые формулы циклическими вычислениями.

Практика показывает, что большая часть ограничений в реальных задачах имеет вид (III.4) или в матричной форме

$$Ax < b,$$

$$x \geqslant 0.$$

Обычно они отражают ограниченность используемых ресурсов, пропускных способностей транспортных коммуникаций, производственных мощностей и т.д.

После приведения к каноническому виду путем введения искусственных переменных и решения полученной эквивалентной системы уравнений (III.5) общее решение будет иметь вид

$$\begin{aligned} y_1 &= -a_{11}x_1 - a_{12}x_2 \ldots - a_{1n}x_n + b_1, \\ y_2 &= -a_{21}x_1 - a_{22}x_2 \ldots - a_{2n}x_n + b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= -a_{m}x_1 - a_{m2}x_2 \ldots - a_{mn}x_n + b_m, \end{aligned} \qquad \text{(III.12)}$$

а в матричной форме

$$y = -Ax + b. \qquad \text{(III.14)}$$

Чтобы исключить операцию изменения знака у элементов матрицы A, (III.13) можно записать в следующем эквивалентном виде

$$y = A(-x) + b.$$

Если принять во внимание, что в базисном решении все независимые переменные $x = 0$, то преимущество такой записи становится очевидным.

В дальнейшем полученное решение удобно хранить в так называемой симплекс-таблице (табл. III.1).

Каждая строка симплекс-таблицы несет в себе всю необходимую числовую информацию из уравнения связи базисной и независимых переменных. Так, переменная y_r получается путем умножения r-й строки таблицы на вектор независимых переменных, дополненный единицей, т.е.

Таблица III.1

Базисная переменная	Независимые переменные				Свободный член
	$-x_1$	$-x_2$	...	$-x_n$	1
$y_1 =$	a_{11}	a_{12}	...	a_{1n}	b_1
$y_2 =$	a_{21}	a_{22}	...	a_{2n}	b_2
...	...	...	...	...	...
$y_r =$	a_{r1}	a_{r2}	...	a_{rn}	b_r
...	...	...	...	...	...
$y_m =$	a_{m1}	a_{m2}	...	a_{mn}	b_m

$$y_r = (a_{r1}, a_{r2}, \ldots, a_{rn}, b_r) \begin{bmatrix} -x_1 \\ -x_2 \\ \ldots \\ -x_n \\ 1 \end{bmatrix} =$$

$$= a_{r1}(-x_1) + a_{r2}(-x_2) \ldots + a_{rn}(-x_n) + b_r.$$

Переход к новому общему решению предусматривает замену местами одной базисной и независимой переменной. При этом изменяются векторы базисных и независимых переменных и элементы симплекс-таблицы.

Алгоритм пересчета элементов симплекс-таблицы легко выводится в результате анализа выполняемых алгебраических преобразований.

Пусть имеется решение, записанное в табл. III.2.

Таблица III.2

	$-z_1$	$-x_2$ ...	$-x_s$	...	$-x_n$	1
$y_1 =$	a_{11}	a_{12} ...	a_{1s}	...	a_{1n}	b_1
$y_2 =$	a_{21}	a_{22} ...	a_{2s}	...	a_{2n}	b_2
	...	...	...	...	...	...
$y_r =$	a_{r1}	a_{r2} ...	$\boxed{a_{rs}}$	...	a_{rn}	b_r
	...	...	...	...	...	...
$y_m =$	a_{m1}	a_{m2} ...	a_{ms}	...	a_{mn}	b_m

в котором необходимо заменить y_r на x_s. Строку и столбец с заменяемыми переменными называют разрешающими строкой и столбцом, а элемент, стоящий на их пересечении – разрешающим элементом. Некоторые авторы вместо этих терминов используют названия ведущие столбец, строка и элемент.

В новой таблице на месте y_r будет располагаться переменная x_s, а на месте x_s – переменная y_r.

Элементы новой симплекс-таблицы вычисляются по следующим правилам.

1. На месте разрешающего элемента в новой таблице записывается $1/a_{rs}$.

2. Остальные элементы разрешающей строки делятся на разрешающий элемент, т.е. вместо $a_{r,j}$ в новую таблицу записывается (a_{rj}/a_{rs}).

3. Остальные элементы разрешающего столбца делятся на разрешающий элемент и записываются в новую таблицу с противоположным знаком, т.е. вместо a_{si} заносится $(-a_{si}/a_{rs})$.

4. Элементы, не принадлежащие разрешающим строке и столбцу, вычисляются по формуле

$$a_{ij} = \frac{a_{ij}a_{rs} - a_{is}a_{rj}}{a_{rs}}. \qquad \text{(III.15)}$$

Из приведенных формул пересчета элементов следует вывод:

замену переменных можно выполнить только в случае, когда разрешающий элемент не равен нулю, т.е. $a_{r,s} \neq 0$.

Описанная процедура носит название модифицированных жордановых исключений и может использоваться для решения целого ряда задач линейной алгебры.

Пример III.8. Для системы ограничений

$$x_1 + 3x_2 - x_3 + y_1 = 3,$$
$$2x_1 + 2x_2 - 3x_3 + y_2 = 2,$$
$$3x_1 - x_2 + 4x_3 + y_3 = 5$$

найти базисное решение, в котором $x_2 \neq 0$.

Система уравнений решается относительно переменных и заносится в симплекс-таблицу III.3.

Таблица III.3

	$-x_1$	$-x_2$	$-x_3$	1
$y_1 =$	1	3	−1	3
$y =$	−2	2	−3	2
	3	−1	4	5

В этом базисном решении все независимые переменные $x_1, x_2, x_3 = 0$. Чтобы $x_2 \not> 0$, достаточно ввести x_2 в базис вместо какой-либо переменной y_i. Пусть это будет y_2. Выполнив по описанным правилам шаг исключений с разрешающим элементом a_{22}, получим таблицу III.4.

Таблица III.4

	$-x_1$	$-y_2$	$-x_3$	1
$y_1 =$	4	−3/2	7/2	0
$x_2 =$	−1	1/2	−3/2	1
$y_3 =$	2	1/2	5/2	6

Выпишем из табл. III.4 общее и базисное решения

$$y_1 = -4x_1 + 3/2y_2 - 7/2x_3,$$
$$x_2 = x_1 - 1/2y_2 + 3/2x_3 + 1,$$
$$y_3 = -2x_1 - 1/2y_2 - 5/2x_3 + 6,$$
$$x_1, y_2, x_3 = 0; \quad y_1 = 0; \quad x_2 = 0; \quad y_3 = 6.$$

§ 5. ПОЛУЧЕНИЕ ИСХОДНОГО БАЗИСНОГО РЕШЕНИЯ

В задачах ЛП с ограничениями-неравенствами получение исходного базисного решения не представляет каких-либо трудностей после введения вспомогательных переменных. Это настолько подробно было показано в § 1, 2 и 4, что здесь можно сразу вывести правило записи исходного базисного решения в симплекс-таблицу.

Система ограничений-неравенств приводится к виду

$$Ax \leqslant b$$

и вписывается в симплекс-таблицу

	$-x$	1
$y =$	A	b

в левый столбец которой заносятся вспомогательные переменные y.

Значительно сложнее получить исходное базисное решение с ограничениями-равенствами.

Пусть имеется система ограничений

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1,$$
$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \qquad \text{(III.16)}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m.$$

Если в эту систему ввести вспомогательные переменные

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + \xi_1 = b_1,$$
$$a_2x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + \xi = b_2, \qquad \text{(III.17)}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n + \xi_m = b_m,$$

то получится система, легко разрешимая относительно $\xi_1, \xi_2, \ldots, \xi_m$. Ее общее решение имеет вид

$$\xi_1 = -a_{11}x_1 - a_{12}x_2 - \ldots - a_{1n}x_n + b_1,$$
$$\xi_2 = -a_{21}x_1 - a_{22}x_2 - \ldots - a_{2n}x_n + b_2, \qquad \text{(III.18)}$$
$$\ldots\ldots\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\xi_m = -a_{m1}x_1 - a_{m2}x_2 - \ldots - a_{mn}x_n + b_m.$$

Очевидно, что система (III.17) будет эквивалентна (III.16) тогда и только тогда, когда $\xi_1 = 0;\ \xi_2 = 0;\ \ldots,\ \xi_m = 0$. Поэтому записывая решение (III.18) в симплекс-таблицу на месте переменных ξ_i, можно сразу записать нули или так называемые 0-переменные

	$-x$	1
$0 =$	A	b

Полученная симплекс-таблица не содержит базисного решения, если хотя бы одно $b_i \neq 0$, чтобы его получить необходимо все 0-переменные вывести из базиса. Эту операцию можно осуществить последовательными шагами жордановых исключений. Ставшая независимой 0-переменная не должна возвращаться в базис, поэтому столбец под ней можно исключить из таблицы.

В процессе замены переменных может возникнуть случай, когда в строке с 0-переменной под $-x$ стоят нули, а свободный член отличен от нуля (табл. III.5). Это указывает на несовместность системы, так как при любых значениях x равенство $0 = 0 \cdot x + \beta$ не может обратиться в тождество.

Если в результате замены появилась строка с 0-переменной, содержащая одни нули, то ее следует исключить из таблицы.

Таким образом, получение исходного базисного решения задачи ЛП со смешанной системой ограничений сводится к следующим операциям.

Таблица III.5

	$-x_{k+1}$	$\ldots$	$-x_n$	1
$x_1 =$	a_{11}	$\ldots$	a_{1n-k}	β_1
	$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$			
$x_k =$	a_{k1}	$\ldots$	a_{kn-k}	β_k
$0 =$	0	$\ldots$	0	β_{k+1}

Ограничения-неравенства приводят к виду (III.4) и вся задача заносится в симплекс-таблицу.

В строках, соответствующих неравенствам, записываются вспомогательные переменные, а в строках с уравнениями – 0-переменные.

Последовательными шагами жордановых исключений 0-переменные выводятся из базиса, а столбец под ними исключается из таблицы.

Если в таблице остались 0-переменные, которые нельзя вывести из базиса из-за появления нулей под независимыми переменными и свободный член в этой строке отличен от нуля, то система несовместна. В противном случае строки с оставшимися 0-переменными исключаются.

Пример III.9. Найти исходное базисное решение задачи

$$Z = 4x_1 + 5x_2 - 5x_3 + x_4 + 4 \to \max,$$
$$-2x_2 + x_4 = 5,$$
$$-2x_1 + 5x_2 + x_3 - 2x_4 = 6,$$
$$2x_1 - 7x_2 - x_3 + 3x_4 = 2,$$
$$-7x_1 + 5x_2 - 6x_3 + 5x_4 \leqslant 4,$$
$$x_1, x_2, x_3, x_4 \geqslant 0.$$

Задача без преобразований записывается в симплекс-таблицу (табл. III.6).

Таблица III.6

	$-x_1$	$-x_2$	$-x_3$	$-x_4$	1
$0 =$	0	–2	0	1	5
$0 =$	–2	5	–1	–2	6
$0 =$	2	–7	–1	3	2
$y_1 =$	–7	5	–6	–5	4
$Z =$	–4	–5	5	1	4

Заметим, что при записи в симплекс-таблицу коэффициенты целевой функции поменяли знаки, так как в ней все независимые переменные стоят со знаком минус.

Последовательными шагами 0-переменные выводятся из базиса (табл. III.7).

Таблица III.7

	$-x_1$	$-x_2$	$-x_3$	1
$x_4 =$	0	–2	0	5
$0 =$	–2	1	1	16

Продолжение табл. III.7

$0 =$	2	−1	−1	12
$y_1 =$	−7	−5	−6	29
$Z =$	−4	−3	5	−1

Таблица III.8

	$-x_1$	$-x_2$	1
$x_4 =$	0	−2	5
$x_3 =$	−2	1	16
$0 =$	0	0	28
$y_1 =$	−19	1	125
$Z =$	6	−8	−81

После второго шага решение прекращается из-за появления признака несовместимости системы в третьей строке табл. III.8.

Пример III.10. Найти исходное базисное решение задачи

$$Z = -6x_1 + 3x_2 - x_3 - 5x_4 \to \max,$$
$$-x_1 + 6x_2$$
$$-2x_1 + x_3 - 5x_4 = 3,$$
$$4x_1 + 2x_2 - 2x_3 + 10x_4 \geqslant 6,$$
$$x_1 + 12x_2 - x_3 + 5x_4 \leqslant 5,$$
$$x_1, x_2, x_3, x_4 \geqslant 0.$$

Ограничения-неравенства преобразуются к виду (III.4), а первое уравнение умножается на −1, чтобы свободный член стал положительным

$$Z = -6x_1 + 3x_2 - x_3 - 5x_4 \to \max,$$
$$x_1 - 6x_2 = 1,$$
$$-2x_1 + x_3 - 5x_4 = 3,$$
$$-4x_1 - 2x_2 + 2x_3 - 10x_4 \leqslant -6,$$
$$x_1 + 12x_2 - x_3 + 5x_4 \leqslant 5,$$
$$x_1, x_2, x_3, x_4 \geqslant 0.$$

Задача записывается в симплекс-таблицу и 0-переменные выводятся из базиса.

Таблица III.9

	$-x_1$	$-x_2$	$-x_3$	$-x_4$	1
$0 =$	1	−6	0	0	1
$0 =$	−2	0	1	−5	3
$y_1 =$	−4	−2	2	−10	−6
$y_2 =$	1	12	−1	5	5
$Z =$	6	−3	1	5	0

Таблица III.10

	$-x_2$	$-x_3$	$-x_4$	1
$x_1 =$	−6	0	0	1
$0 =$	−12	1	−5	5
$y_1 =$	−26	2	−10	2
$y_2 =$	18	−1	5	4
$Z =$	33	1	5	−6

Таблица III.11

	$-x_2$	$-x_4$	1
$x_1 =$	−6	0	1
$x_3 =$	−12	−5	5
$y_1 =$	−2	0	−6
$y_2 =$	6	0	9
$Z =$	21	10	−11

Таблица III.11 содержит исходное базисное решение задачи: $x_2 = x_4 = 0$; $x_1 = 1$; $x_3 = 5$; $y_1 = -6$; $y_2 = 9$.

§ 6. ПОИСК ОПОРНОГО РЕШЕНИЯ

Пусть имеется исходное базисное решение, помещенное в симплекс-таблицу III.12.

Таблица III.12

	$-x_1$	$-x_2$	...	$-x_s$	...	$-x_n$	1
$y_1 =$	a_{11},	a_{12}	...	a_{1s}	...	a_{1n}	b_1
$y_2 =$	a_{21},	a_{22}	...	a_{2s}	...	a_{2n}	b_2
.....	.....	.....	.....	.....	.....	.....	.....
$y_k =$	a_{k1}	a_{k2}	...	a_{ks}	...	a_{kn}	b_k
.....	.....	.....	.....	.....	.....	.....	.....
$y_n =$	a_{n1}	a_{n2}	...	a_{ns}	...	a_{nn}	b_n
.....	.....	.....	.....	.....	.....	.....	.....
$y_m =$	a_{m1}	a_{m2}	...	a_{ms}	...	a_{mn}	b_m
$Z =$	γ_1	γ_2	...	γ_s	...	γ_n	q

Допустимым, т.е. опорным, это решение будет в том случае, если все переменные имеют неотрицательное значение. В базисном решении независимые переменные приравниваются к нулю, а базисные оказываются равными свободным членам. Так, в имеющемся решении x_1, $x_2, \ldots, x_n = 0$; $y_1 = b_1$; $y_2 = b_2$; $\ldots, y_m = b_m$.

Отсюда можно сформулировать правило проверки допустимости базисного решения.

Решение является опорным, если элементы столбца свободных членов в строках ограничений симплекс-таблицы неотрицательны.

Пусть в данном решении $b_r < 0$, а все остальные свободные члены положительны.

Это решение неопорное и следует перейти к другому, в котором переменная $y_r \geqslant 0$.

В § 2 сформулирована стратегия поиска опорного решения, которая сводится к следующим условиям:

переход к новому решению должен сохранять неотрицательность переменных;

каждый шаг должен обеспечивать монотонный рост хотя бы одной отрицательной переменной.

Если использовать в качестве средства перехода к новому решению процедуру жордановых исключений, то реализация этой стратегии сведется к правилам выбора разрешающего элемента, т.е. выбора заменяемых переменных.

При переходе к новому решению выводимая из базиса переменная становится независимой и приравнивается к нулю, а новая базисная переменная определяется из соответствующей приравненной нулю линейной формы. Следовательно, чтобы переменная y_r стала неотрицательной, ее необходимо перевести в независимые, введя в базис ту переменную, которая после замены сохранит положительное значение.

Решим вопрос о возможности замены y_r на x_s. После этой замены y_r будет равна нулю, а x_s определится из уравнения

$$y_r = -a_{r1}x_1 - a_{r2}x_2 \ldots - a_{rs}x_s \ldots - a_{rn}x_n + b_r = 0, \qquad \text{(III.19)}$$

в котором $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ как независимые переменные в базисном решении, т.е.

$$x_s = \frac{b_r}{a_{rs}}. \qquad \text{(III.20)}$$

Так как по условию $b_r < 0$, то после замены переменных x_s будет больше нуля только в том случае, если $a_{rs} < 0$. Из этого можно сделать важный вывод: избавиться от неотрицательности базисной переменной, сохранив число неотрицательных переменных, можно только в том случае, если в соответствующей y_r строке симплекс-таблицы есть хотя бы один отрицательный элемент $a_{ls} < 0$. В противном случае замена y_r на любую переменную x_j не даст сокращение числа отрицательных переменных и не приблизит новое базисное решение к допустимому. Отсюда вытекают правила.

Если в строке с отрицательным свободным членом все элементы неотрицательны, то система ограничений несовместна.

В базис должна вводиться независимая переменная, у которой коэффициент в строке с отрицательным сводным членом также отрицателен.

При переходе к новому решению изменяются все базисные переменные, поэтому следует рассмотреть возможность их выхода в отрицательную область.

Пусть в базис вводится переменная x_s, у которой $a_{rs} < 0$. По условию в новом решении она должна получить некоторое положительное приращение $\Delta x_s > 0$. Рассмотрим, как будут меняться при этом базисные переменные $y_1, y_2, \ldots, y_m$, подставив Δx_s в исходное решение. Приняв во внимание, что в новом базисном решении все остальные независимые переменные сохранят свои нулевые значения, можно придти к следующим линейным функциям изменения базисных переменных

$$\begin{aligned} y_1 &= a_{1s}(-\Delta x_s) + b_1, \\ y_2 &= a_{2s}(-\Delta x_s) + b_2, \\ &\ldots\ldots\ldots\ldots \\ y_m &= a_{ms}(-\Delta x_s) + b_m. \end{aligned} \qquad \text{(III.21)}$$

Приращение Δx_s взято с отрицательным знаком, так как в симплекс-таблице независимые переменные записаны со знаком минус.

Из приведенных линейных зависимостей и условия положительности свободных членов очевидно, что при увеличении Δx_s первой обратится в ноль базисная переменная y_k, для которой

$$\frac{b_k}{a_{ks}} = \min \frac{b_i}{a_{is}} > 0,$$

т.е. отношение свободного члена к элементу разрешающего столбца минимально и положительно. Чтобы сохранить число неотрицательных переменных, эту переменную надо вывести из базиса и тем самым закрепить на нулевом уровне.

Следует отметить, что отношение свободного члена к элементу разрешающего столбца положительно для всех базисных переменных, стремящихся к нулю.

Обобщив сказанное, можно вывести следующее правило.

Из базиса следует выводить переменную, для которой отношение свободного члена к коэффициенту при вводимой в базис переменной положительно и минимально.

В результате перехода к новому решению переменная y_r может пока не достичь нулевого уровня, а некоторые отрицательные базисные переменные могут стать еще меньше. Поэтому для обеспечения монотонного роста одной из отрицательных переменных выбор разрешающего элемента надо начинать с анализа одной и той же строки, пока в ней свободный член не станет положительным или не будет установлена несовместность системы.

Пример III.11. Поиск опорного решения задачи из примера III.4.

После приведения к каноническому виду задача записывается в симплекс-таблицу.

Таблица III.13

	$-x_1$	$-x_2$	1	Отношение $\frac{b_i}{a_{i1}} > 0$
$y_1 =$	[−5]	−1	−5	−5/−5
$y_2 =$	−3	−2	−6	−6/−3
$y_3 =$	1	−1/2	3	3/1
$y_4 =$	1	2	8	8/1
$Z =$	−1	−1	1	

Эта таблица содержит исходное базисное решение, но оно не допустимо, так как в первой и второй строках свободные члены отрицательны. Выберем разрешающий элемент и перейдем к новому решению. Чтобы обеспечить монотонность роста отрицательных переменных, будем выбирать разрешающий элемент, просматривая строки сверху вниз и столбцы слева направо, т.е. в порядке возрастания индексов.

Первая сверху строка имеет отрицательный свободный член и в первом столбце она содержит отрицательный элемент. Этот столбец будет разрешающим. Минимальное положительное отношение свободного члена к элементу разрешающего столбца достигается в первой строке и она принимается за разрешающую. После шага исключений с разрешающим элементом a_{11} получается таблица III.14.

Таблица III.14

	$-y_1$	$-x_2$	1	Отношение $\frac{b_i}{a_{i1}} > 0$
$x_1 =$	−1/5	1/5	1	
$y_2 =$	[−3/5]	−7/5	−3	−3/(−3/5)
$y_3 =$	1/5	−7/10	2	2/(1/5)
$y_4 =$	1/5	9/5	7	7/(1/5)
$Z =$	−1/5	−4/5	2	

Это решение также неопорное, так как во второй строке свободный член отрицателен. Первый столбец содержит в этой строке отрицательный элемент, поэтому он назначается разрешающим. Разрешающая строка выбирается по отношению свободного члена к элементу разрешающего столбца. Это отношение будет положительным и минимальным во второй строке. После шага исключений с разрешающим элементом a_{21} получается табл. III.15.

Таблица III.15

	$-y_2$	$-x_2$	1
$x_1 =$	−1/3	2/3	2
$y_1 =$	−5/3	7/3	5
$y_3 =$	1/3	−7/6	1
$y_4 =$	1/3	4/3	6
$Z =$	−1/3	−1/3	3

В табл. III.15 все свободные члены в строках ограничений положительны; следовательно решение $y_2 = x_2 = 0$; $x_1 = 2$; $y_1 = 5$; $y_3 = 1$; $y_4 = 6$ опорное.

Пример III.12. Поиск опорного решения задачи из примера III.10. Полученное в этом примере исходное базисное решение (табл. III.16).

Таблица III.16

	$-x_2$	$-x_4$	1	Отношение $\frac{b_i}{a_{i1}} > 0$
$x_1 =$	−6	0	1	
$x_3 =$	−12	−5	5	
$y_1 =$	−2	0	−6	−6/−2
$y_2 =$	**6**	0	9	9/6
$Z =$	21	10	−11	

не является опорным, так как в третьей строке свободный член отрицателен. В этой строке первый элемент отрицателен, поэтому первый столбец назначается разрешающим. Минимальное положительное отношение b_i/a_{i1} соответствует четвертой строке и a_{41} назначается разрешающим элементом. После шага исключений получается табл. III.17.

Таблица III.17

	$-y_2$	$-x_4$	1
$x_1 =$	1	0	10
$x_3 =$	2	−5	23
$y_1 =$	1/3	0	−3/2
$x_2 =$	1/6	0	3/2
$Z =$	7/2	10	$20^1/_2$

В этой таблице в третьей строке сохранилось отрицательное значение свободного члена, но больше отрицательных элементов нет, следовательно система несовместна.

§ 7. ПОИСК ОПТИМАЛЬНОГО РЕШЕНИЯ

Предположим, что симплекс-таблица III.12 содержит опорное решение и все элементы столбца свободных членов в строках ограничений положительны, т.е. $b_1, b_2, \ldots, b_m > 0$. Сформулированная в § 3 стратегия поиска оптимума требует выполнения следующих условий:

в процессе поиска перебираются только опорные решения, т.е. на каждом шаге должна сохраняться неотрицательность всех переменных;

каждый шаг должен обеспечивать монотонный рост целевой функции.

Проанализируем возможность выполнения этих условий.

Для обеспечения первого условия переход к новому базисному решению должен сопровождаться положительным приращением вводимой в базис независимой переменной. Пусть это будет переменная, которая после ввода в базис получит значение $\Delta x_s > 0$. Так как в новом базисном решении все остальные переменные x_i сохранят свои нулевые значения, то целевая функция определится следующим выражением

$$Z_s = r_s(-\Delta x_s) + q. \tag{III.22}$$

В исходном решении, при $x_1, x_2, \ldots, x_s, \ldots, x_n = 0$, целевая функция была равна

$$Z = q. \tag{III.23}$$

Чтобы реализовать второе условие, должно выполняться отношение $Z < Z_s$ или после подстановки (III.22) и (III.23)

$$\begin{aligned} &q < \gamma_s(-\Delta x_s) + q, \\ &\gamma_s(-\Delta x_s) > 0. \end{aligned} \tag{III.24}$$

Так как по условию $\Delta x_s > 0$, то отношение (III.24) может выполняться только, если $\gamma_s < 0$.

Из условия конечности алгоритма поиска (см. § 3) и проведенных рассуждений можно сформировать условие оптимальности решения.

Опорное решение является оптимальным, если в Z-строке симплекс-таблицы все коэффициенты при переменных неотрицательны.

Пусть оптимальное решение еще не достигнуто и в Z-строке табл. III.12 $\gamma_s < 0$. Переход к новому решению, в котором x_s станет базисной переменной, обеспечит рост целевой функции, поэтому s-й столбец можно назначить решающим. Чтобы при этом выдерживалось первое условие стратегии поиска, должны выполняться отношения

$$\begin{aligned} y_1 &= a_{1s}(-\Delta x_s) + b_1 \geqslant 0, \\ y_2 &= a_{2s}(-\Delta x_s) + b_2 \geqslant 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{ms}(-\Delta x_s) + b_m \geqslant 0. \end{aligned}$$

При увеличении Δx_s первой будет обращаться в ноль переменная y_r, для которой

$$\frac{b_r}{a_{rs}} = \min \frac{b_i}{a_{is}} > 0, \quad i = 1, 2, \ldots, m$$

и ее следует вывести из базиса.

Из этого можно сформулировать правило выбора разрешающей строки.

В качестве разрешающей берется строка, в которой отношение свободного члена к элементу разрешающего столбца положительно и минимально.

Если отношение $\frac{b_k}{a_{ks}} < 0$, а это может быть только при $a_{ks} < 0$, при увеличении Δx_s переменная y_k будет расти и также будет расти целевая функция Z. Пределом роста переменной y_r было обращение в ноль одной из базисных переменных с положительным коэффициентом в разрешающем столбце, но, если все $a_{1s}, a_{2s}, \ldots, a_{ms} \leqslant 0$, это ограничение исчезает и за счет переменной x_s целевая функция может быть увеличена до ∞. Отсюда вытекает признак неограниченности линейной формы.

Если в опорном решении в столбце с отрицательным коэффициентом в Z-строке нет положительных элементов, то целевая функция не ограничена.

Пример III.13 Нахождение оптимального решения задачи из примера III.4. В примере III.12 для этой задачи было получено опорное решение (см. табл. III.15), помещенное в табл. III.18.

Таблица III.18

	$-y_2$	$-x_2$	1	Отношение $\frac{b_i}{a_{i1}} > 0$
$x_1 =$	$-1/3$	$2/3$	2	
$y_1 =$	$-5/3$	$7/3$	5	
$y_3 =$	$\boxed{1/3}$	$-7/6$	1	$1/(1/3)$
$y_4 =$	$1/3$	$4/3$	6	$6/(1/3)$
$Z =$	$-1/3$	$-1/3$	3	

Это решение не оптимально, так как в Z-строке имеются отрицательные коэффициенты при неизвестных. $\gamma_1 < 0$, поэтому первый столбец назначается

разрешающим. Минимальное отношение $b_i/a_{i1} > 0$ достигается в третьей строке и она назначается разрешающей. После шага исключений имеем табл. III.19.

Таблица III.19

	$-y_3$	$-x_2$	1	Отношение $\frac{b_i}{a_{i2}} > 0$.
$x_1 =$	1	$-1/2$	3	
$y_1 =$	5	$-7/2$	10	$10/(7/2)$
$y_2 =$	3	$-7/2$	3	
$y_4 =$	-1	$\boxed{5/2}$	5	$5/(5/2)$
$Z =$	1	$-3/2$	4	

Это решение также не оптимально и может быть улучшено введением y_4 в базис. Минимальное отношение $b_i/a_{i2} > 0$ достигается в четвертой строке. После шага исключений приходим к табл. III.20.

Таблица III.20

	$-y_3$	$-y_4$	1
$x_1 =$	4/5	1/5	4
$y_1 =$	18/5	7/5	17
$y_2 =$	8/5	7/5	10
$x_2 =$	$-2/5$	2/5	2
$Z =$	2/5	3/5	7

Z-строка табл. III.20 больше не содержит отрицательных коэффициентов, из чего следует оптимальность полученного решения:

$x_1 = 4$; $x_2 = 2$; $Z_{\max} = 7$.

Пример III.14. Решение задачи из примера III.5.

$$W = x_1 + 2x_2 + 1 \to \max,$$

$$5x_1 + x_2 \geqslant 5,$$

$$3x_1 + 2x_2 \geqslant 6,$$

$$x_1 - \frac{1}{2}x_2 \leqslant 3,$$

$$x_1 + 2x_2 \leqslant 8,$$

$$x_1, x_2 \leqslant 0.$$

После приведения к каноническому виду задача записывается в исходную симплекс-таблицу III.21.

Таблица III.21

	$-x_1$	$-x_2$	1
$y_1 =$	[−5]	−1	−5
$y_2 =$	−3	−2	−6
$y_3 =$	1	−1/2	3
$y_4 =$	1	2	8
$W =$	−1	−2	1

По описанной процедуре симплекс-метода сначала находится опорное (табл. III.21 – III.23), а затем оптимальное решения (табл. III.23 – III.25).

Таблица III.22

	$-y_1$	$-x_2$	1
$x_1 =$	−1/5	1/5	1
$y_2 =$	[−3/5]	−7/5	−3
$y_3 =$	1/5	−7/10	2
$y_4 =$	1/5	9/5	7
$W =$	−1/5	−9/5	2

Таблица III.23

	$-y_2$	$-x_2$	1
$x_1 =$	−1/3	2/3	2
$y_1 =$	−5/3	7/3	5
$y_3 =$	[1/3]	−7/6	1
$y_4 =$	1/3	4/3	6
$W =$	−1/3	−4/3	3

Таблица III.24

	$-y_3$	$-x_2$	1
$x_1 =$	1	−1/2	3
$y_1 =$	5	−7/2	10
$y_2 =$	3	−7/2	3
$y_4 =$	−1	[5/2]	5
$W =$	1	−5/2	4

Таблица III.25

	$-y_3$	$-y_4$	1
$x_1 =$	4/5	1/5	4
$y_1 =$	[18/5]	7/5	17
$y_2 =$	8/5	7/5	10
$x_2 =$	−2/5	2/5	2
$W =$	0	1	7

Из графического решения задачи (см. пример III.5) известно, что оптимальное решение $x_1 = 4$; $x_2 = 2$; $W_{max} = 7$ не единственное.

Действительно, сделав еще один шаг с разрешающим элементом r_{21}, получим табл. III.26, в которой также содержится оптимальное решение $x_1 = 2/9$; $x_2 = 35/9$; $W_{max} = 7$.

Таблица III.26

	$-y_1$	$-y_4$	1
$x_1 =$	−2/9	−1/9	2/9
$y_3 =$	5/18	7/18	85/18
$y_2 =$	−4/9	7/9	22/9
$x_2 =$	1/9	5/9	35/9
$W =$	0	1	7

Сопоставляя последние строки табл. III.25 и III.26, можно сделать важный практический вывод:

если в симплекс-таблице, содержащей оптимальное решение, в Z-строке имеются нулевые коэффициенты, то оптимальное решение неединственно. Перейти к другому оптимальному решению можно, выбрав разрешающий элемент из столбца с нулевым коэффициентом в Z-строке.

Пример III.15. Численное решение задачи из примера III.6.

После приведения к каноническому виду задача записывается в симплекс-таблицу III.27 и находится опорное решение (табл. III.27 – III.29).

Для поиска оптимального решения делается шаг модифицированных жордановых исключений с разрешающим элементом r_{31}. Получим симплекс-таблицу III.30.

Таблица III.27

	$-x_1$	$-x_2$	1
$y_1 =$	[−5]	−1	−5
$y_2 =$	−3	−2	−6
$y_3 =$	1	−1/2	3
$Z =$	−1	−1	1

Таблица III.28

	$-y_1$	$-x_2$	1
$x_1 =$	−1/5	1/5	1
$y_2 =$	[−3/5]	−7/5	−3
$y_3 =$	1/5	−7/10	2
$Z =$	−1/5	−4/5	2

Таблица III.29

	$-y_2$	$-x_2$	1
$x_1 =$	$-1/3$	$2/3$	2
$y_1 =$	$-5/3$	$7/3$	5
$y_3 =$	$\boxed{1/3}$	$-7/6$	1
$Z =$	$-1/3$	$-1/3$	3

Таблица III.30

	$-y_3$	$-x_2$	1
$x_1 =$	1	$-1/2$	3
$y_1 =$	5	$-7/2$	10
$y_2 =$	3	$-7/2$	3
$Z =$	1	$-3/2$	4

Дальнейший поиск оптимального решения следует прервать, так как во втором столбце в Z-строке имеется отрицательный коэффициент и над ним нет ни одного положительного элемента, что указывает на неограниченность целевой функции.

Для уяснения геометрической сущности симплекс-метода сопоставим графическое (см. рис. III.4) и численное (табл. III.13 – III.16, III.18 – III.20) решения задачи из примера III.4.

Началом симплекс-процесса послужило исходное базисное решение $x_1, x_2 = 0$; $y_1 = -5$; $y_2 = -6$; $y_3 = 3$; $y_4 = 8$. На рис. III.4 ему соответствует начало координат – точка 0. Затем последовал переход к решению $y_1 = x_2 = 0$; $x_1 = 1$; $y_2 = -3$; $y_3 = 2$; $y_4 = 7$, которому соответствует точка A, и далее к решению $y_2 = x_2 = 0$; $x_1 = 2$; $y_1 = 5$; $y_3 = 1$; $y_4 = 6$ в точку B. В численном решении с этого момента начинается поиск оптимального решения, а графически это выглядит как переход от одной вершины многогранника решений к соседней по соединяющему их ребру. Так, далее последовало решение $y_3 = x_3 = 0$; $y_1 = 10$; $y_2 = 3$; $y_4 = 5$ – точка C и, наконец, оптимальное решение $y_3 = y_4 = 0$; $x_1 = 4$; $x_2 = 2$; $y_1 = 17$; $y_2 = 10$ – точка D.

Следует отметить, что характерной чертой симплекс-процесса является переход в соседнюю точку пересечения разделяющих прямых на новом этапе решения задачи. При этом прогноз улучшения решения делается не более чем на один шаг. Такая стратегия обеспечивает неуклонное движение вперед, но не гарантирует самый короткий маршрут. Так из точки B можно было двинуться по пути B, E, F, D вместо B, C, D, взяв в табл. III.18 в качестве разрешающего второй столбец, что не противоречит стратегии поиска.

Из геометрической интерпретации видно, что усовершенствовать описанный вариант симплекс-метода можно, во-первых, за счет исключения остановок в некоторых промежуточных точках при поиске опорного решения, например, в точке A и, во-вторых, за счет выбора шага, который дает максимальное увеличение целевой функции.

§ 8. ВЫРОЖДЕНИЕ

При обсуждении стратегии поиска и путей ее реализации в виде симплекс-алгоритма предполагалось отсутствие вырожденных базисных решений у системы ограничений.

Напомним, базисное решение называется вырожденным, если оно содержит хотя бы одну базисную переменную, равную нулю. Реальные задачи достаточно часто содержат вырожденные решения, поэтому здесь следует обсудить связанные с вырождением осложнения процесса решения и методы борьбы с ними.

Рассмотрим графическое и численное решение задачи, содержащей вырожденные базисные решения.

Пример III.16. Графическое и численное решения задачи

$$Z = x_1 + x_2 + 1 \to \max,$$
$$-3x_1 - 2x_2 \leqslant -6,$$
$$3/2x_1 - 2x_2 \leqslant 3,$$
$$-x_1 - 2x_2 \leqslant -2,$$
$$x_1 - 1/2x_2 \leqslant 3,$$
$$x_1 + 2x_2 \leqslant 8,$$
$$x_1, x_2 \geqslant 0.$$

Графическое решение задачи, приведенное на рис. III.8, протекает по обычной схеме. Из рисунка видно, что оптимум достигается в точке D, представляющей собой точку пересечения линий

$$y_4 = x_1 - 1/2x_2 - 3 = 0,$$
$$y_5 = x_1 + 2x_2 - 8 = 0.$$

Решая эту систему уравнений, находим

$$x_1 = \frac{\begin{vmatrix} -3 & -1/2 \\ -8 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & -1/2 \\ 1 & 2 \end{vmatrix}} = 4;$$

$$x_2 = \frac{\begin{vmatrix} 1 & -3 \\ 1 & -8 \end{vmatrix}}{\begin{vmatrix} 1 & -1/2 \\ 1 & 2 \end{vmatrix}} = 2;$$

$$Z_{\max} = 7.$$

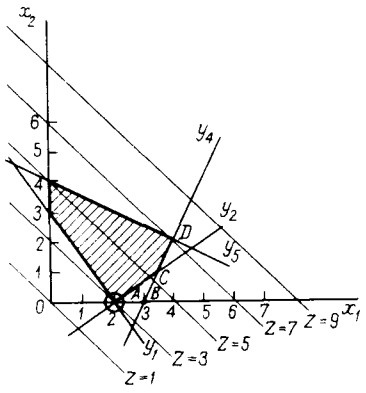


Рис. III.8. Графическое решение примера III.16

Теперь попытаемся решить эту задачу численно. После одного шага поиска опорного решения от исходной симплекс-таблицы III.31. приходим к табл. III.32, содержащей опорное вырожденное решение. На рис. III.8 ему соответствует точка A, через которую проходит четыре разделяющие линии y_1, y_2, y_3 и x_1.

Таблица III.31

	$-x_1$	$-x_2$	1
$y_1 =$	−3	−2	−6
$y_2 =$	3/2	−2	3
$y_3 =$	−1	−2	−2
$y_4 =$	1	−1/2	3
$y_5 =$	1	2	8
$Z =$	−1	−1	1

Таблица III.32

	$-y_3$	$-x_2$	1
$y_1 =$	−3	4	0
$y_2 =$	3/2	−5	0
$x_1 =$	−1	2	2
$y_4 =$	1	−5/2	1
$y_5 =$	1	0	6
$Z =$	−1	1	3

Если действовать дальше по ранее сформулированным правилам, то в качестве разрешающих следует взять первый столбец и четвертую строку табл. III.32, в которой отношение $b_i/a_{i1} > 0$ минимально. Но шаг исключений с разрешающим элементом $a_{4,1}$ приводит к потере опорного решения (табл. III.33), так как во второй строке свободный член стал отрицательным. На рисунке этому решению соответствует точка B, лежащая за пределами многогранника решений. Из этого следует, что в случае вырождения сформулированные правила выбора разрешающего элемента не гарантируют сохранение неотрицательности переменных.

Таблица III.33

	$-y_4$	$-x_2$	1
$y_1 =$	3	−7/2	3
$y_2 =$	−3/2	−5/4	−3/2
$x_1 =$	1	−1/2	3
$y_3 =$	1	−5/2	1
$y_5 =$	−1	5/2	5
$Z =$	1	−3/2	4

Проводя рассуждения аналогично изложенным в § 7, 8, можно предположить следующее.

Чтобы исключить появление отрицательных свободных членов, вместо нулей в качестве разрешающей надо брать строку, для которой отношение $b_i/a_{is} \geqslant 0$ и минимально.

Действуя по этому правилу, из табл. III.32 следует выбрать в качестве разрешающих первые столбец и строку. После шага исключений с a_1 получается табл. III.34, в которой сохранилось опорное решение, но не исчезло вырождение и не возросло значение целевой функции. То есть данный шаг не обеспечил выполнение условия монотонного роста целевой функции. На рис. III.8 данному решению по прежнему соответствует точка A.

Действуя дальше по новым правилам, из табл. III.34 в качестве разрешающих следует снова взять первые столбец и строку. Но повторный шаг с элементом $a_{1,1}$ приведет к исходной табл. III.32.

Таблица III.34

	$-y_1$	$-x_2$	1
$y_3 =$	−1/3	−4/3	0
$y_2 =$	1/2	−3	0
$x_1 =$	−1/3	2/3	2
$y_4 =$	1/3	−7/6	1
$y_5 =$	1/3	4/3	6
$Z =$	−1/3	−1/3	3

Из этого следует вывод, что такое изменение правил также не решает проблемы, так как может привести к "зацикливанию" процесса поиска.

Одним из эффективных способов борьбы с вырождением является ϵ-метод, предложенный А. Чарнсом. Идея метода заключается в следующем.

Каждый шаг симплекс-процесса представляет собой переход в соседнюю точку по одному из ребер, образующих данную вершину многогранника решений. Положительное отношение свободного члена к элементу разрешающего столбца показывает, что двигаясь по этому ребру можно достичь крайней точки, слелав шаг указанного размера.

В случае вырождения соседние крайние точки сливаются в одну и вычисление "пробного шага" не позволяет правильно сориентироваться. Чтобы вернуть возможность ориентироваться, предлагается путем параллельного сдвига ребер немного разнести крайние точки на время выбора направления движения по обычным правилам симплекс-метода.

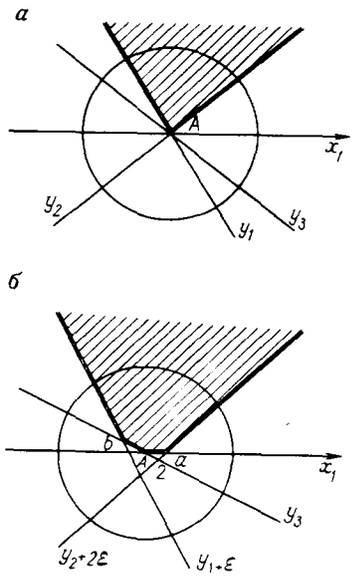


Рис. III.9. Окрестность точки A до ($а$) и после ($б$) внесения смещений

На рис. III.9 показано, что после введения ϵ-смещений в окрестности точки A появились две новые крайние точки a и b, шаг в которые однозначно определит дальнейшее движение по y_2 или y_1.

В численном решении ϵ-метод реализуется следующим образом. Перепишем табл. III.32 в табл. III.35.

Таблица III.35

	y_3	$-x_2$	1	Смещения	отношения $b_i/a_{i1} > 0$
$y_1 =$	−3	4	0	ϵ	
$y_2 =$	[3/2]	−5	0	2ϵ	$2\epsilon/(3/2)$
$x_1 =$	−1	2	2		
$y_4 =$	1	−5/2	1		1
$y_5 =$	1	0	6		6
$Z =$	−1	1	3		

В вырожденных строках вместо нулей запишем очень малые числа ϵ, умноженные на номер строки, и вычислим положительные отношения свободных членов к элементам разрешающего столбца, выбранного по обычным правилам.

Величина ϵ обычно берется на пять порядков меньше наименьшего свободного члена. На номер строки величина ϵ умножается для того, чтобы снизить вероятность получения одинаковых отношений и обеспечить однозначность выбора.

Приняв во внимание величину ϵ по отношению к остальным свободным членам в качестве разрешающей, выбираем вторую строку табл. III.35. Дальнейший шаг исключений выполняем с прежними, нулевыми значениями свободных членов и приходим к табл. III.36.

Применив опять ϵ-метод, выбираем в качестве разрешающих второй столбец и четвертую строку и после шага исключений приходим к табл. III.37. Это решение уже невырожденное и дальнейшее решение может идти по обычным правилам. В результате приходим к решению (табл. III.38) : $x_1 = 4$; $x_2 = 2$; $Z = 7$.

Аналогично ϵ-метод применяется на этапе поиска опорного решения.

В машинных алгоритмах после выбора разрешающего элемента обратное смещение в ноль свободных членов не делается. Легко показать, что при этом погрешность не превышает ϵm или отсутствует вовсе.

Таблица III.36

	$-y_2$	$-x_2$	1	Смещения отношения $b_i/a_{i2} > 0$
$y_1 =$	2	−6	0	ϵ
$y_3 =$	2/3	−10/3	0	2ϵ
$x_1 =$	2/3	−4/3	2	
$y_4 =$	−2/3	[5/6]	1	6/5
$y_5 =$	−2/3	10/3	6	6/(10/3)
$Z =$	2/3	−7/3	3	

Таблица III.37

	$-y_2$	$-y_4$	1
$y_1 =$	−14/5	36/5	36/5
$y_3 =$	−2	4	4
$x_1 =$	−2/5	8/5	18/5
$x_2 =$	−4/5	6/5	6/5
$y_5 =$	[2]	−4	2
$Z =$	−6/5	14/5	29/5

Таблица III.38

	$-y_5$	$-y_4$	1
$y_1 =$	7/5	8/5	10
$y_3 =$	1	0	6
$x_1 =$	1/5	4/5	4
$x_2 =$	2/5	−2/5	2
$y_2 =$	1/2	−2	1
$Z =$	3/5	2/5	7

Контрольные вопросы и задачи

1. Какой тип ограничений должен присутствовать в задаче ЛП, приведенной к каноническому виду?

2. Какое точечное множество называется выпуклым?

3. Чем отличаются граничные и крайние точки выпуклого множества?

4. Что представляет собой крайняя точка множества решений основной задачи ЛП:

а) с геометрической точки зрения?

б) с аналитической точки зрения?

5. Какие исходы решения задачи ЛП могут быть?

6. В каких характерных точках множества решений может находиться оптимальное решение задачи ЛП?

7. В чем состоит суть стратегии поиска оптимального решения задачи ЛП, заложенная в симплекс-метод?

8. Какое условие стратегии поиска опорного решения обеспечивается:

а) правилом выбора разрешающего столбца в симплекс-алгоритме?

б) правилом выбора разрешающей строки?

9. Какое условие стратегии поиска оптимального решения обеспечивается:

а) правилом выбора разрешающего столбца?

б) правилом выбора разрешающей строки?

10. Каков признак того, что целевая функция достигает максимума в двух или более опорных решениях?

11. Задача. Решить численно и графически следующие задачи ЛП:

а) $Z = 2x_1 + x_2 \to \max,$

$$2x_1 + 4x_2 \leqslant 16,$$
$$-4x_1 + 2x_2 \leqslant 8,$$
$$x_1 + 3x_2 \geqslant 6,$$
$$x_1, x_2 \geqslant 0.$$

б) $Z = 3x_1 + 2x_2 \to \min,$

$$x_1 + 2x_2 \leqslant 6,$$
$$2x_1 + x_2 \leqslant 8,$$
$$-x_1 + x_2 \leqslant 1,$$
$$x_1, x_2 \geqslant 0.$$

12. Задача. Решить симплекс-методом следующие задачи ЛП:

а) $Z = x_1 + x_2 + x_3 \to \max,$

$$2x_1 + 2x_2 + x_3 \geqslant 2,$$
$$3x_1 - 3x_2 + 2x_3 \leqslant 6,$$
$$3x_1 - 3x_2 - 2x_3 \leqslant 6,$$
$$2x_2 - 2x_3 \leqslant 2,$$
$$x_1, x_2, x_3 \geqslant 0.$$

б) $Z = 10x_1 - x_2 - 9x_3 - 8x_4 \to \max,$

$$2x_1 - x_2 - 3x_3 - x_4 = -2,$$

$$-5x_1 + 2x_2 + 3x_4 = 5,$$
$$7x_1 - 4x_2 + x_3 + 4x_4 \leqslant 1,$$
$$3x_1 + 2x_2 + 5x_3 + 6x_4 \leqslant 10,$$
$$x_1, x_2, x_3, x_4 \geqslant 0.$$

в) $Z = 3x_1 + 3x_2 - 4x_3 \to \max,$

$$2x_1 + x_2 - 3x_3 \geqslant 18,$$
$$4x_1 - 5x_3 \geqslant 12,$$
$$3x_1 - 2x_2 + x_3 \geqslant 14,$$
$$x_1, x_2, x_3 \geqslant 0.$$

Глава IV

ПРИМЕНЕНИЕ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ В ЗАДАЧАХ ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ ГОРНЫМ ПРОИЗВОДСТВОМ

§ 1. УСЛОВИЯ ПРИМЕНЕНИЯ И КЛАССИФИКАЦИЯ ЛИНЕЙНЫХ МОДЕЛЕЙ

Как известно, к линейным относятся модели, в которых целевая функция и все ограничения линейные, т.е. коэффициенты целевой функции и ограничений постоянны. Хотя практически обычно коэффициенты целевой функции или части ограничений (параметры моделей) зависят от оптимизируемых величин, в ряде случаев этой зависимостью можно пренебречь. Это не существенно сказывается на результате решения, но в то же время позволяет значительно снизить трудоемкость решения.

Линейные модели применяются для решения большого количества плановых задач. Причем форма задач во многих случаях повторяется, задачи одного класса возникают в различных отраслях горной промышленности, сферах деятельности горного предприятия.

Практически все линейные модели, адекватно описывающие задачи организации, планирования и управления горным производством, относятся к классу о распределении, смеси или их комбинациям.

Задачи распределения заключаются в отыскании такого распределения ресурсов (людей, материалов, оборудования, денег), имеющихся в распоряжении руководителя, при котором операция выполняется эффективно. К этому виду относится большое количество задач планирования и снабжения.

Различают следующие три типа задач распределения:

количество ресурсов соответствует их потребности на выполнение всех операций и необходимо найти их наилучшее распределение;

количество ресурсов не хватает для выполнения всех операций и необходимо выбрать те, которые следует выполнять, а также определить, каким способом;

имеется возможность регулирования ресурсов и требуется определить, какие их них следует использовать, а от каких отказаться.

Простейшие задачи первого типа – задачи о назначении. В этом случае имеется один вид ресурсов, которые следует разместить по операциям. Каждой комбинации операций и ресурсов соответствует определенная эффективность. Например, имеется несколько шоферов и различные типы машин. Каждый шофер может достичь разных показателей на машинах неодинаковых марок. Требуется так распределить шоферов по машинам, чтобы эффективность использования автотранспорта была наибольшей. К подобным типам относятся также задачи распределения: экскаваторов по местам работ (забоям, категориям пород), чтобы минимизировать затраты на погрузочные работы; машин между маршрутами; служащих по видам работ; станков между рабочими; автомашин по экскаваторам и т.д.

К первому типу относятся и так называемые транспортные задачи, широко распространенные в народном хозяйстве. Сущность их заключается в следующем. Ряд шахт и карьеров должны отправлять добываемый уголь в несколько адресов. Известно количество угля, добываемого на каждой шахте и карьере, и потребности каждого пункта потребления. Требуется так организовать перевозки угля, чтобы минимизировать затраты на транспорт или пробег груза.

К транспортной относится задача распределения поступных объемов по отвалам, видам транспорта и перегрузочным пунктам (если на карьере имеется несколько видов транспорта) и т.д.

Ко второму типу относятся практически все задачи, связанные с распределением финансов. Задача расстановки оборудования с целью максимизации выполняемых объемов также может быть отнесена к этому типу. Сюда же относятся задачи, связанные с распределением времени работы оборудования по операциям, если все необходимые операции не могут быть выполнены. Например, ЭВМ перегружена заказами. Требуется определить порядок выполнения заказов, дающий наибольший эффект.

К третьему типу относится большинство задач проектирования и перспективного планирования. Например, при проектировании карьера требуется выбрать вид карьерного транспорта. Если возможно использование автомобильного и железнодорожного транспорта и их комбинации, то следует определить, какие объемы горной массы и каким видом транспорта перевозятся, а в случае комбинации – указать горизонт расположения перегрузочного пункта. В этом случае будут определены объемы работ каждого вида транспорта и количество оборудования, минимизирующие стоимость работ.

К третьему типу относятся задачи определения места строительства нового предприятия, порядка ввода в эксплуатацию месторождений, выбор расположения обогатительной фабрики, обслуживающей группу шахт и карьеров, а также об оптимальном использовании различных ресурсов (задача об оптимальном плане выпуска продукции). На предприятиях, выпускающих различную продукцию, руководители планируют выпуск каждого ее вида. Выпуск определенной продукции ограничен как технологическими, так и экономическими возможностями и требует затрат трудовых, материальных, финансовых и других ресурсов. Причем нормы расхода отдельных ресурсов по видам продукции существенно различны. Необходимо с учетом имеющегося количества ресурсов определить выпуск различных видов продукции, обеспечивающих максимальную эффективность работы предприятия.

Этот тип задач является комбинацией всех трех типов задач о распределении.

К задачам о смеси относятся те, в которых из различных компонентов, отличающихся по ряду технологических (физических, химических и др.) свойств, требуется образовать наиболее эффективным образом смесь, отвечающую определенным требованиям. В горной промышленности к данному классу относятся задачи планирования и управления добычными работами в режиме усреднения. Полезные ископаемые, добываемые на отдельных участках месторождения, отличаются различным содержанием полезных и вредных компонентов, в то время как перерабатывающие предприятия (обогатительные фабрики, металлургические заводы, электростанции и др.) предъявляют к качеству поставляемого сырья определенные (достаточно жесткие) требования. Необходимо спланировать добычу полезного ископаемого с каждого участка месторождения таким образом, чтобы общий его поток соответствовал требованиям потребителей и добыча (а иногда добыча и переработка в комплексе) осуществлялись наиболее эффективно. Подобные задачи возникают при обеспечении центральной обогатительной фабрики углем различных шахт, обогащении руды с нескольких рудников на одной обогатительной фабрике и др.

§ 2. ЗАДАЧИ О РАССТАНОВКЕ ОБОРУДОВАНИЯ

Задачи этого вида появляются, когда используется разнотипное оборудование и имеются различные по своим характеристикам участки работы. На горных предприятиях такие задачи возникают, когда необходимо распределить бригады или отдельных рабочих по местам работы, транспортные средства по участкам, а оборудование по блокам (уступам, забоям) и др.

Расстановка оборудования должна обеспечивать либо минимальные затраты на выполнение заданного объема работ, либо максимизацию выполняемого объема работ (выпуска продукции) имеющимся обору-

дованием. Соответственно критерием оптимальности могут служить затраты (денежные или времени) или объем работ (выпуск продукции в натуральном или денежном выражении).

Приведем обобщенные постановки задач расстановки оборудования и их математические модели.

Имеется оборудование нескольких типов, которое можно использовать на различных участках. Количество оборудования каждого типа и его производительность на участках известны. Необходимо так распределить оборудование, чтобы общее время его работы для выполнения задания было минимальным.

Обозначим через $i = 1, 2, \ldots, n$ – тип оборудования; $j = 1, 2, \ldots, m$ – участки работы; P_{ij} – производительность i-го оборудования на j-м участке; V_j – плановый объем работы на j-м участке; K_i – ресурс (во времени) работы оборудования i-го типа. За управляемые переменные можно принять x_{ij} – объем работы i-го оборудования на j-м участке или y_{ij} – время работы i-го оборудования на j-м участке. Приведем модели задач для обоих типов управляемых переменных, что проиллюстрирует возможность создания различных моделей для решения одной и той же задачи.

За критерий оптимальности примем общее время выполнения всего заданного объема работ. Общее время работы слагается из времени работы каждого вида оборудования i на всех участках j и выражается следующим образом

$$W = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{x_{ij}}{P_{ij}} \to \min. \qquad \text{(IV.1)}$$

Выражение (IV.1) представляет собой целевую функцию задачи. При решении задач должны соблюдаться ограничения:

по выполнению планового объема работ на каждом участке V_j

$$\sum_{i=1}^{n} x_{ij} = V_j, \quad (j = 1, 2, \ldots, m); \qquad \text{(IV.2)}$$

по имеющемуся ресурсу работы оборудования

$$\sum_{j=1}^{m} \frac{x_{ij}}{P_{ij}} \leqslant K_i, \quad (i = 1, 2, \ldots, n); \qquad \text{(IV.3)}$$

по положительности решения

$$x_{ij} \geqslant 0, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \qquad \text{(IV.4)}$$

Количество ограничений в задаче: первого типа n, второго m, а третьего $n \times m$.

Модель этой же задачи с управляемыми переменными y_{ij} имеет вид

$$\sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} \to \min \qquad \text{(IV.5)}$$

при ограничениях:

$$\sum_{i=1}^{n} y_{ij} P_{ij} = V_j, \quad (j = 1, 2, \ldots, m); \qquad \text{(IV.6)}$$

$$\sum_{j=1}^{m} y_{ij} \leqslant K_i, \quad (i = 1, 2, \ldots, n); \qquad \text{(IV.7)}$$

$$x_{ij} \geqslant 0, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \qquad \text{(IV.8)}$$

Данные модели являются линейными. Если на определенной работе может находиться только целое количество единиц оборудования, то задача будет относиться к классу задач дискретного (целочисленного) линейного программирования.

Если стоимость 1 машино-ч работы оборудования различных типов существенно отличается и равна C_i, то в качестве критерия оптимальности могут использоваться суммарные затраты на выполнение заданного объема работ

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \frac{x_{ij}}{P_{ij}} C_i \to \min; \qquad \text{(IV.9)}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} C_i \to \min. \qquad \text{(IV.10)}$$

При решении задачи в описанной постановке может оказаться, что для выполнения заданного объема работ нет необходимости использовать все имеющееся оборудование. Поэтому часто целесообразно определять оптимальную расстановку оборудования, принимая за критерий эффективности не затраты времени (или денежные), а выполнение объема работ, который необходимо максимизировать.

В этом случае модель задачи будет иметь вид

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} \to \max \qquad \text{(IV.11)}$$

при ограничениях:

по выполнению минимально необходимого объема работ на участке

$$\sum_{i=1}^{n} x_{ij} \geqslant V_j, \quad (j = 1, 2, \ldots, m); \qquad \text{(IV.12)}$$

по количеству оборудования

$$\sum_{j=1}^{m} \frac{x_{ij}}{P_{ij}} \leqslant K_i, \quad (i = 1, 2, \ldots, n); \tag{IV.13}$$

по положительности решения

$$x_{ij} \geqslant 0, \quad (i = \overline{1, n}; \quad j = \overline{1, m}). \tag{IV.14}$$

В указанной модели объемы работ по участкам равноценны.

В ряде же случаев бывает выгоднее увеличить объемы на отдельных участках, либо необходимо пропорционально перевыполнять объемы на всех участках и т.д. В этом случае модель задачи изменяется в соответствии с конкретной обстановкой. Покажем это на примерах.

Автотранспортное хозяйство, имеющее автосамосвалы различной грузоподъемности, обслуживает несколько участков (участки карьеров, строительные участки, перевозка нерудных материалов и т.д.). Эффективность (производительность) работы автосамосвала зависит от его типа и участка работы (разница в длине откатки, типе погрузочного оборудования, характеристике дорог). Необходимо их распределить таким образом, чтобы максимизировать общий объем работ, соблюдая плановое соотношение объемов работ между участками.

За управляемые переменные примем объем груза, вывозимый автосамосвалами с каждого участка, x_{ij}, где i – номер участка ($i = 1, 2, \ldots, n$); j – тип автосамосвала ($j = 1, 2, \ldots, m$).

Исходными данными (неуправляемыми переменными) являются: Q_{ij} – производительность каждого типа автосамовала на каждом участке; N_j – количество автосамосвалов j-го типа; V_i – отношение объемов работ по каждому участку к общему объему работ.

Критерием эффективности, исходя из постановки задачи, принимаем суммарный объем работ, который надо максимизировать. В этом случае работа транспорта будет наиболее производительной, что, в свою очередь, снизит и затраты на перевозку.

Целевая функция задачи будет иметь вид

$$\sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij} \to \max. \tag{IV.15}$$

При максимизации целевой функции должны быть учтены ограничения:

по количеству имеющихся автосамосвалов

$$\sum_{i=1}^{n} \frac{x_{ij}}{Q_{ij}} \leqslant N_j, \quad j = \overline{1, m}; \tag{IV.16}$$

по соблюдению планового соотношения объемов работ между участками

$$\frac{\sum_{j=1}^{m} x_{ij}}{\sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij}} = V_i, \quad (i = 1, 2, \ldots, n), \tag{IV.17}$$

или

$$\frac{\sum_{j=1}^{m} x_{i1}}{V_1} = \ldots = \frac{\sum_{j=1}^{m} x_{ij}}{V_i} = \ldots = \frac{\sum_{j=1}^{m} x_{im}}{V_n}; \tag{IV.17a}$$

по положительности решения

$$x_{ij} \geqslant 0, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \tag{IV.18}$$

Данная модель относится к классу линейных.

На следующем примере покажем, как учитываются ценность различных участков (видов работ) и необходимость соблюдения определенных соотношений объемов работ между участками, определяемых технологией.

Карьер разрабатывает горизонтальное месторождение. Вскрышные работы отличаются по своим физико-техническим свойствам. На вскрышных работах используются экскаваторы различных типов. Известна производительность экскаваторов по участкам. Требуется так расставить оборудование по уступам, чтобы обеспечить максимальный прирост вскрытых запасов.

Обозначим через i – номера уступов (групп уступов), $i = 1, 2, \ldots, n$ (нумерация начинается с верхнего уступа); j – тип экскаватора; Q_{ij} – производительность на i-м участке j-го экскаватора; N_{ij} – количество экскаваторов j-го типа.

За управляемые переменные примем объем работ на i-м уступе j-го экскаватора, x_{ij}. Так как прирост вскрытых запасов непосредственно обеспечивается удалением вскрышных пород с нижних уступов, необходимо учитывать приоритет нижних уступов (увеличивающуюся ценность объемов работ на них). Коэффициент приоритета нижних уступов обозначим через K_i; он должен существенно возрастать с увеличением номера уступа (т.е. сверху-вниз).

Тогда целевую функцию можно представить как

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} K_i \to \max \tag{IV.19}$$

При ограничениях:
по количеству оборудования

$$\sum_{i=1}^{n} \frac{x_{ij}}{Q_{ij}} \leqslant N_j, \quad j = \overline{1, m} \tag{IV.20}$$

по положительности решения

$$x_{ij} \geqslant 0, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \tag{IV.21}$$

Если в задаче нет других ограничений, то ответ тривиален: все экскаваторы надо использовать на нижнем уступе. Естественно, что такой ответ неверен. Между уступами существует взаимосвязь, при их отработке должна сохраняться минимально допустимая правилами технической эксплуатации рабочая площадка. Таким образом, в модели должны учитываться еще технологические ограничения, делающие невозможной подработку верхних уступов нижними. Эти ограничения при параллельном подвигании уступов записываются следующим образом

$$\frac{\sum_{j=1}^{m} x_{ij}}{H_i L_i} - \frac{\sum_{j=1}^{m} x_{i-1,j}}{H_{i-1} L_{i-1}} \geqslant 0, \quad i = \overline{1, n}, \tag{IV.22}$$

где H_i – высота i-го уступа; L_i – длина фронта работ i-го уступа.

Если на момент начала работ ширина рабочей площадки отличается от нормальной, то вместо нуля в ограничение ставится величина a, равная величине превышения (уменьшения) ширины рабочей площадки по сравнению с минимально необходимой.

Модель даже с учетом технологических ограничений является линейной. Если фронт перемещается по вееру или криволинейно, то формализация технологических ограничений существенно усложняется.

Пример IV.1. На карьере имеются два вскрышных уступа и используются автосамосвалы КрАЗ-256Б и МоАЗ-522. Общее их количество по типам соответственно 8 и 12. Производительность автосамосвалов (т/смену) КрАЗ-256Б и МоАЗ-522 составляет: по уступу № 1 соответственно 240 и 324; № 2 – 180 и 234.

Определить, какой объем работ может быть выполнен, если объем вскрышных работ на уступе № 2 не должен быть меньше, чем на № 1?

Словесная формулировка задачи: определить максимальные объемы пород, вывозимых автосамосвалами с уступов, при соблюдении заданного соотношения объемов работ между уступами. Математическая формулировка: обозначим через x_1 – объем пород, вывозимый КрАЗ-256Б с уступа № 1; x_2 – то же, с уступа № 2; x_3 – объем пород, вывозимый МоАЗ-522 с уступа № 1; x_4 – то же, с уступа № 2.

Математическая модель задачи принимает вид

$W = x_1 + x_2 + x_3 + x_4$ (общие объемы работ).

Максимизировать при ограничениях:

$$\frac{x_1}{240} + \frac{x_2}{180} \leqslant 8 \text{ (по количеству КрАЗ-256Б)};$$

$$\frac{x_3}{324} + \frac{x_4}{234} \leqslant 12 \text{ (по количеству МоАЗ-522)};$$

$x_1 + x_3 \leqslant x_2 + x_4$, (по соотношению объемов работ между уступами);

$x_{1,\,2,\,3,\,4} \geqslant 0$ (условие неотрицательности переменных).

§ 3. ЗАДАЧИ ОБ ОПТИМАЛЬНОМ ИСПОЛЬЗОВАНИИ РЕСУРСОВ (ОПТИМАЛЬНОМ ПЛАНЕ ВЫПУСКА ПРОДУКЦИИ)

Многие предприятия и объединения горной промышленности являются комплексными и выпускают несколько видов продукции. Например, комбинаты и объединения цветной металлургии обычно выпускают концентраты нескольких металлов (или сами металлы), объединения угольной промышленности добывают уголь разных марок, обширна номенклатура продукции заводов горного машиностроения. Выпуск каждого вида продукции требует определенного расхода ресурсов различных видов (финансовых, трудовых, сырья, материалов и т.д.). Каждый вид продукции имеет определенную себестоимость и отпускную цену. В условиях, когда известен максимально возможный и минимально необходимый выпуск продукции каждого вида, требуется составить наилучший план работы (выпуска продукции).

Обозначим через $i = 1, 2, \ldots, n$ – виды продукции; $j = 1, 2, \ldots, m$ – типы ресурсов; C_i – себестоимость производства продукции i-го вида; $Ц_i$ – отпускная цена продукции i-го вида; a_{ij} – расход j-го ресурса на изготовление i-й продукции; d_j – имеющееся в наличии количество j-го ресурса; $b_i^{\min}$ и $b_i^{\max}$ – соответственно минимально необходимый (обязательное плановое задание) и максимально возможный (по условиям реализации) выпуск продукции i-го типа.

За неуправляемые переменные задачи примем x_i – объем выпуска продукции i-го типа. В качестве критерия оптимальности используется прибыль (объем реализации продукции), а ограничениями являются возможные пределы (минимальный и максимальный) выпуска продукции и количество ресурсов. Модель задачи имеет следующий вид.

Максимизировать целевую функцию

$$\sum_{i=1}^{n} x_i (Ц_i - C_i) \to \max \tag{IV.23}$$

при ограничениях:

по выпуску продукции

$$b_i^{\max} \geqslant x_i \geqslant b_i^{\min}, \quad i = \overline{1, n}; \tag{IV.24}$$

по расходу ресурсов

$$\sum_{i=1}^{n} x_i a_{ij} \leqslant d_j, \quad j = \overline{1, m}; \tag{IV.25}$$

по положительности решения

$$x_i \geqslant 0, \quad i = \overline{1, n}. \tag{IV. 26}$$

Если критерием оптимальности является объем реализации продукции в денежном выражении, то целевая функция записывается следующим образом

$$\sum_{i=1}^{n} x_i Ц_i \rightarrow \max. \tag{IV.27}$$

В данной модели ограничение по выпуску продукции является двусторонним, левая часть ограничения соответствует возможному объему реализации продукции, а правая – минимально необходимому (заданному) объему выпуска.

Приведенные модели являются линейными, так как величина C_i принята постоянной и независимой от объема выпуска продукции. Если же $C_i = f(x_i)$, то модель становится нелинейной.

Пример IV.2. Завод выпускает два вида деталей ($А$ и $Б$), обработка которых осуществляется на токарных и фрезерных станках. На месяц заводу выделено 30 т металла. Его расход на изготовление детали $А$ – 0,10 т, $Б$ – 0,12 т.

На изготовление детали $А$ требуется 3 ч работы токарного и 4 ч работы фрезерного станков, а детали $Б$ 3,7 и 3,8 ч соответственно. Всего на заводе 3 токарных и 4 фрезерных станка, работающих в две смены (т.е. месячный календарный фонд рабочего времени каждого станка составляет 300 ч.) Определить объем выпуска деталей каждого вида, если их цена соответственно составляет 3,5 и 4.2 руб.

Словесная формулировка задачи: определить выпуск деталей, при котором обеспечивается максимальный объем товарной продукции при соблюдении ограничений по ресурсам (металлу и станочному времени)

Математическая формулировка задачи: обозначим через $x_А$ – количество деталей вида $А$; $x_Б$ – то же, вида $Б$.

При использовании этих обозначений математически задача запишется следующим образом:

максимизировать $Z = 3{,}5x_А + 4{,}2x_Б$ (объем товарной продукции) при ограничениях:

$0{,}1x_А + 0{,}12\,x_Б \leqslant 30$ (по расходу металла);

$\left.\begin{array}{l} 3.0x_А + 3{,}7x_Б \leqslant 3 \cdot 300 \\ 4.0x_А + 3{,}8x_Б \leqslant 4 \cdot 300 \end{array}\right\}$ (по предельному времени использования токарных и фрезерных станков);

$x_А, x_Б \geqslant 0$ (условие неотрицательности переменных).

§ 4. ПЛАНИРОВАНИЕ ДОБЫЧНЫХ РАБОТ В РЕЖИМЕ УСРЕДНЕНИЯ КАЧЕСТВА

Отдельные участки (лавы, забои) горных предприятий имеют полезное ископаемое, в котором содержание полезных и вредных компонентов варьирует в широких пределах. Потребитель предъявляет к качеству сырья жесткие требования. Необходимо, учитывая возможности каждого участка, так составить план добычных работ, чтобы затраты были минимальны и выполнялись все требования потребителей к качеству.

Обозначим через i – участки добычи, $i = \overline{1, n}$; α_i – содержание 1-го компонента на i-м участке; β_i – то же, 2-го; γ_i – то же, 3-го; $Q_i^{\max}$, $Q_i^{\min}$ – максимально возможный и минимально необходимый объемы добычи на i-м участке. Максимально возможный объем определяется техническими характеристиками оборудования, а минимально необходимый – по технологическим соображениям; C_i –затраты на добычу руды на i-м участке; $Q_{\text{пл}}$ – плановый объем добычи руды; $\alpha_{\text{пл}}$ – плановое содержание 1-го компонента, т.е. содержание 1-го компонента должно строго соответствовать плановому; $\beta^{\min}$, $\beta^{\max}$ – минимально и максимально допустимые содержания 2-го компонента в руде; $\gamma^{\max}$ – максимальное содержание 3-го (вредного) компонента.

За управляемые переменные примем x_i – объем добычи на i-м участке.

За критерий оптимальности примем затраты на добычу. Тогда модель задачи будет иметь следующий вид.

Минимизировать целевую функцию

$$\sum_{i=1}^{n} x_i\ C_i \rightarrow \min \tag{IV.28}$$

при ограничениях:

по производительности каждого участка

$$Q_i^{\max} \geqslant x_i \geqslant Q_i^{\min}, \quad i = \overline{1, n}; \tag{IV.29}$$

по суммарному объему работ

$$\sum_{i=1}^{n} x_i \geqslant Q_{\text{пл}}; \tag{IV.30}$$

по качеству руды

$$\frac{\sum_{i=1}^{n} x_i\ \alpha_i}{\sum_{i=1}^{n} x_i} = \alpha_{\text{пл}}; \tag{IV.31}$$

$$\beta^{max} \geqslant \frac{\sum_{i=1}^{n} x_i \beta_i}{\sum_{i=1}^{n} x_i} \geqslant \beta^{min}; \qquad \text{(IV.32)}$$

$$\frac{\sum_{i=1}^{n} x_i \gamma_i}{\sum_{i=1}^{n} x_i} \leqslant \gamma^{max}; \qquad \text{(IV.32a)}$$

по положительности решения

$$x_i \geqslant 0; \quad i = \overline{1, n}. \qquad \text{(IV.33)}$$

Данная задача является линейной, она относится к известному в исследовании операций классу задач о смеси (диете).

Помимо стоимостного в качестве критерия оптимальности могут использоваться, например, следующие показатели.

1. Суммарный объем добычи (при планировании в течение смены)

$$\sum_{i=1}^{n} x_i \to \max \qquad \text{(IV.34)}$$

В этом случае второе ограничение (по объему добычи) становится излишним.

2. Минимальное отклонение качества по основному компоненту от планового

$$|\alpha - \alpha_{пл}| \to \min \qquad \text{(IV.35)}$$

$$\alpha = \frac{\sum_{i=1_1}^{n} x_i \alpha_i'}{\sum_{i=1}^{n} x_i}. \qquad \text{(IV.35a)}$$

В этом случае не требуется ограничение по качеству 1-го (основного) компонента.

3. Равномерный объем добычи по участкам

$$\max|x_i - x_{ср}| \to \min \text{ или } \min \max|x_i - x_{ср}|, \qquad \text{(IV.36)}$$

где

$$x_{ср} = \frac{\sum_{i=1}^{n} x_i}{n}. \qquad \text{(IV.36a)}$$

Этот же критерий может иметь и квадратичную форму

Таблица IV.1

Показатель	Рудники		
	1	2	3
Максимальный объем добычи, млн.т	1,2	0,9	6,5
Затраты на добычу 1 т, руб	7,5	8,2	2,4
Содержание меди в руде, %	1,8	2,1	1,0
Содержание серы в руде, %	0,7	0,8	0,6

$$\sum_{i=1}^{n} (x_i - x_{ср})^2 \to \min \qquad \text{(IV.37)}$$

При использовании критериев равномерности добычи или минимальных колебаний качества задача становится нелинейной. В случае модульных критериев задача может быть сведена к ряду линейных путем введения специальных ограничений.

Пример IV.3. Обогатительная фабрика перерабатывает медную руду с трех рудников. Годовой объем переработки 8 млн. т. Данные по рудникам приведены в табл. IV.1.

Содержание меди в руде, поставляемой на обогатительную фабрику, должно быть в пределах 1,2–1,3 %, а серы не превышать 0,7 %.

Определить годовые плановые задания рудникам, обеспечивающие выполнение планового объема переработки руды установленного качества с минимальными затратами.

Словесная формулировка задачи: определить объемы добычи руды, обеспечивающие установленные объемы и качество перерабатываемой руды с минимальными затратами.

Математическая формулировка задачи: обозначим через x_1 объем добычи руды 1-м рудником; x_2 – то же, 2-м; x_3 – то же, 3-м.

С учетом введенных обозначений математическая модель задачи принимает следующий вид.

Минимизировать $Z = 7{,}5x_1 + 8{,}2x_2 + 2{,}4x_3$ (затраты на добычу руды)

при ограничениях:

$$\left.\begin{array}{l} x_1 \leqslant 1{,}2 \\ x_2 \leqslant 0{,}9 \\ x_3 \leqslant 6{,}5 \end{array}\right\} \text{(по максимальной производственной мощности рудников);}$$

$x_1 + x_2 + x_3 = 8$ (по объему переработки);

$$\left.\begin{array}{l} \dfrac{1{,}8x_1 + 2{,}1x_2 + 1{,}0x_3}{x_1 + x_2 + x_3} \geqslant 1{,}2 \\ \dfrac{1{,}8x_1 + 2{,}1x_2 + 1{,}0x_3}{x_1 + x_2 + x_3} \leqslant 1{,}3 \end{array}\right\} \text{(по минимальному и максимальному содержанию меди в руде);}$$

$$\left.\frac{0{,}7x_1 + 0{,}8x_2 + 0{,}6x_3}{x_1 + x_2 + x_3} \leqslant 0{,}7\right| \quad \text{(по максимально допустимому содержанию серы);}$$

$x_1, x_2, x_3 \geqslant 0$ (по неотрицательности переменных).

Ограничения по качеству после преобразований принимают вид

$$0{,}6x_1 + 0{,}9x_2 - 0{,}2x_3 \geqslant 0,$$
$$0{,}5x_1 + 0{,}8x_2 - 0{,}3x_3 \leqslant 0,$$
$$0{,}1x_2 - 0{,}1x_3 \leqslant 0.$$

§ 5. ПЛАНИРОВАНИЕ ПЕРЕВОЗОК ГРУЗОВ ГОРНЫХ ПРЕДПРИЯТИЙ

Грузы предприятий горной промышленности составляют значительный удельный вес (до 40 %) в грузообороте страны. Поэтому рациональное планирование их перевозок является важной задачей. Они называются транспортными и относятся к линейным.

Транспортная задача линейного программирования заключается в определении объемов перевозок от поставщиков к потребителям, если известны выпуск продукции у каждого поставщика и потребности потребителей; затраты на перемещение грузов от поставщиков к потребителям. План грузоперевозок должен иметь минимальные затраты. Причем это могут быть затраты денежных средств, транспортной работы (грузооборот в тоннокилометрах) или времени перевозок.

Приведем модель транспортной задачи в общем виде.

Имеется несколько n производителей (поставщиков) однородного продукта ($i = 1, 2, 3, ..., n$) и несколько m потребителей ($j = 1, 2, 3, ..., m$). Объем груза, который необходимо вывезти от i-го поставщика – A_i, а потребность j-го потребителя – B_i. Затраты на перевозку единицы груза от i-го поставщика j-му потребителю составляют c_{ij}.

Необходимо составить план перевозок, имеющий минимальные затраты, при котором все грузы вывозятся, полностью удовлетворяются запросы потребителей.

За управляемые переменные примем x_{ij} – объем груза, перевозимого от i-го поставщика j-му потребителю.

Тогда целевая функция задачи примет вид

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}\ c_{ij} \to \min \qquad \text{(IV.38)}$$

Ограничениями задачи (помимо требований положительности переменных, которые здесь и в дальнейшем не указываются) являются:

обязательность вывоза грузов от каждого поставщика

$$\sum_{j=1}^{m} x_{ij} = A_i, \quad i = 1, 2, ..., n; \qquad \text{(IV.39)}$$

полное удовлетворение запросов потребителей

$$\sum_{i=1}^{n} x_{ij} = B_j, \quad j = 1, 2, ..., m. \qquad \text{(IV.40)}$$

В данной модели предполагается, что суммарный объем производства полностью соответствует общей потребности, т.е.

$$\sum_{i=1}^{n} A_i = \sum_{j=1}^{m} B_j. \qquad \text{(IV.41)}$$

Подобная модель называется закрытой. Если $\sum_{i=1}^{n} A_i \neq \sum_{j=1}^{m} B_j$, то модель называется открытой, для ее решения она приводится к закрытой введением фиктивного поставщика или потребителя.

Транспортные задачи широко используются в горной промышленности как в чистом виде, так и в качестве составных частей более сложных задач планирования работы группы горных предприятий и размещения промышленности.

Транспортные задачи – это также и задачи планирования грузоперевозок внутри предприятия, например, планирование карьерных грузопотоков. Приведем модель этой задачи.

На карьере имеется несколько вскрышных забоев. Объемы работ по каждому забою известны. Имеется также несколько отвалов, приемная способность каждого известна. Требуется так спланировать перевозку горной массы из забоев на отвалы, чтобы затраты на транспорт были минимальными.

Обозначим через i – забои карьера (участки горного предприятия), $i = \overline{1, n}$; j – отвалы (приемные пункты), $j = \overline{1, m}$; Q_i – объем работ по i-му участку; Q_j – максимальную приемную способность j-го отвала; c_{ij} – затраты на перевозку единицы объема с i-го забоя на j-й отвал.

В зависимости от типа задачи и способа измерения затрат критерий оптимальности может быть различным (грузооборот, стоимость перевозок, время перевозок).

За управляемые переменные примем x_{ij} – объем грузоперевозок с i-го участка на j-й отвал, тогда целевая функция примет вид

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}\ c_{ij} \to \min \qquad \text{(IV.42)}$$

при ограничениях:

по объему работ по участкам

$$\sum_{j=1}^{m} x_{ij} = Q_i, \quad i = \overline{1, n}, \qquad \text{(IV.43)}$$

т.е. с каждого участка нужно увезти породу полностью;

Таблица IV.2

Участок	Отвалы	
	1	2
1	3,5	4,2
2	5,7	5,1
3	4,5	4,8

по приемной способности отвалов

$$\sum_{i=1}^{n} x_{ij} \leqslant Q_j, \quad j = \overline{1, m} \tag{IV.44}$$

т.е. объем пород, вывозимый на отвал, не может превышать его приемной способности;

по положительности переменных

$$x_{ij} \geqslant 0. \tag{IV.45}$$

Модель открытая, так как есть неравенства и объем, который могут принять отвалы, больше, чем добывается на участках

$$\left(\sum_{j=1}^{m} Q_j - \sum_{i=1}^{n} Q_i\right) > 0. \tag{IV.46}$$

Пример IV.4. На карьере три вскрышных участка, порода с которых вывозится на два отвала. Средние расстояния транспортирования породы с участков на отвалы (в км) приведены в табл. IV.2.

Годовой объем вскрышных работ по участкам соответственно составляет (в млн.м^3) – 2,6; 3,7; 2,9, а максимальные приемные способности отвалов соответственно 4,7 и 5,3.

Распределить объемы вскрышных пород с участков по отвалам, обеспечив минимальную работу транспорта.

Словесная формулировка задачи: составить план грузоперевозок, обеспечивающих вывозку и плановые объемы вскрышных пород с минимальными транспортными затратами, учитывая приемные способности отвалов.

Математическая формулировка задачи: обозначим через x_{11} – объем породы, транспортируемый с 1-го участка на 1-й отвал; x_{12} – то же, с 1-го участка на 2-й отвал; x_{21} – то же, со 2-го участка на 1-й отвал; x_{22} – то же, со 2-го участка на 2-й отвал; x_{31} – то же, с 3-го участка на 1-й отвал; x_{32} – то же, с 3-го участка на 2-й отвал.

При использовании этих обозначений математическая модель задачи принимает следующий вид.

Минимизировать $Z = 3{,}5x_{11} + 4{,}2x_{12} + 5{,}7x_{21} + 5{,}1x_{22} + 4{,}5x_{31} + 4{,}8x_{32}$ (годовой грузооборот)

при ограничениях:

$$\left.\begin{aligned} x_{11} + x_{12} &= 2{,}6 \\ x_{21} + x_{22} &= 3{,}7 \\ x_{31} + x_{32} &= 2{,}9 \end{aligned}\right\}$$ (условия обязательности выполнения плановых объемов работ по участкам);

$$\left.\begin{aligned} x_{11} + x_{21} + x_{31} &\leqslant 4{,}7 \\ x_{11} + x_{22} + x_{32} &\leqslant 5{,}7 \end{aligned}\right\}$$ (предельная приемная способность отвалов);

$x_{ij} \geqslant 0$, $i = \overline{1,3}$; $j = \overline{1,2}$ (условие неотрицательности переменных).

§ 6. МОДЕЛЬ ЗАДАЧИ ПЛАНИРОВАНИЯ РАБОТЫ ГРУППЫ ГОРНЫХ ПРЕДПРИЯТИЙ (ДОБЫВАЮЩИХ И ПЕРЕРАБАТЫВАЮЩИХ)

В составе комбината (объединения, подотрасли или отрасли в целом) имеется несколько добывающих и перерабатывающих предприятий. На каждом предприятии добывается полезное ископаемое определенного качества и с известными затратами. Имеется несколько перерабатывающих предприятий (потребителей). Технологические и экономические показатели переработки руды каждого типа на перерабатывающих предприятиях, а также затраты на транспорт от горного предприятия к перерабатывающему определены. Максимальная производственная мощность добывающих и перерабатывающих предприятий известна. Необходимо спланировать работу предприятий таким образом (определить объемы добычи и переработки по предприятиям и грузопотоки), чтобы прибыль была максимальна.

Обозначим через i – добывающие предприятия, $i = \overline{1, n}$; j – перерабатывающие предприятия, $j = \overline{1, m}$; a_i – содержание полезного компонента на i-м предприятии; $Q_i^{\max}$ – максимально возможный объем добычи на i-м предприятии; C_i – затраты на добычу на i-м предприятии; $Q_j^{\max}$ – максимальная производственная мощность j-го перерабатывающего предприятия; ϵ_{ij} – извлечение полезного компонента в концентрат на j-м перерабатывающем предприятии из руды i-го добывающего; C_{ij}^{o} – затраты на обогащение 1 т руды с i-го предприятия на j-м перерабатывающем; $d_j^{\min}$ – требуемое минимальное содержание полезного компонента на j-м предприятии; $d_j^{\max}$ – допустимое максимальное содержание полезного компонента на j-м предприятии; C'_{ij} – затраты на транспорт руды с i-го рудника на j-ю фабрику; $Ц$ – цена единицы конечной продукции.

Управляемыми переменными в задаче являются: y_i – плановая добыча на i-м предприятии; y_j – плановый объем переработки на j-й фабрике; x_{ij} – объем перевозок с i-го добывающего предприятия на j-е перерабатывающее; при этом

$$y_i = \sum_{j=1}^{m} x_{ij}; \tag{IV.47}$$

$$y_j = \sum_{i=1}^{n} x_{ij}. \tag{IV.48}$$

В данной постановке за критерий оптимальности можно принять прибыль комбината

$$П_р = Д - З^д - З^т - З^о,$$

где $Д$ – доход комбината; $З^д$, $З^т$, $З^о$ – затраты на добычу, транспорт, обогащение.

Целевая функция задачи имеет вид

$$П_р = \sum_{j=1}^{m} Ц\, y_i\, \alpha_j\, \epsilon_{ij} - \sum_{i=1}^{n} y_i\, C_i - \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}\, C_{ij} - \sum_{j=1}^{m} C_{ij}^{o}\, y_i \to \max \qquad \text{(IV.49)}$$

при ограничениях:

по объему добычи на каждом предприятии

$$y_i \leqslant Q_i^{\max}, \quad i = \overline{1, n}; \qquad \text{(IV.50)}$$

по мощности перерабатывающих предприятий

$$y_i \leqslant Q_j^{\max}, \quad j = \overline{1, m}; \qquad \text{(IV.51)}$$

по содержанию металла (по качеству)

$$\alpha_j^{\min} \leqslant \alpha_j \leqslant \alpha_j^{\max}; \qquad \text{(IV.52)}$$

$$\frac{\sum_{i=1}^{n} x_{ij}\, \alpha_i}{\sum_{i=1}^{n} x_{ij}} = \alpha_j. \qquad \text{(IV.52а)}$$

В простейшем случае, когда α_i, C_i, C'_{ij}, C^{o}_{ij}, ϵ_{ij} не зависят от объемов добычи, переработки и перевозок, задача является линейной. Если же эти величины являются функциями управляемых переменных, то модель существенно усложняется и становится нелинейной. Данная является моделью обобщенных задач планирования добычных работ и грузопотоков.

Пример IV.5. Горно-обогатительный комбинат включает три рудника и две обогатительные фабрики. Данные по рудникам комбината приведены в табл. IV.1 (из предыдущего примера), а по фабрикам – в табл. IV.3.

Таблица IV.3

Показатель	Фабрики	
	1	2
Производственная мощность, млн.т	3	4
Коэффициент извлечения	0,76	0,80
Затраты на переработку 1 т руды, руб.	1,7	1,9
Максимально допустимое содержание меди в перерабатываемой руде, %	1,30	1,25
Минимально допустимое содержание меди, %	1,20	1,15
Максимально допустимое содержание серы, %	0,68	0,75

Таблица IV.4

Рудник	Фабрики	
	1	2
1	1,5	2,0
2	1,9	2,7
3	2,5	2,8

Затраты на транспортирование 1 т руды с рудников на обогатительные фабрики (в руб.) приведены в табл. IV.4.

Определить плановые задания рудникам и фабрикам, если ГОК в целом должен выпустить 65 тыс.т меди (в концентрате).

Словесная формулировка задачи: определить объемы добычи руды по рудникам и переработки по обогатительным фабрикам комбината, обеспечивающие выполнение планового задания по выпуску металла с минимальными затратами.

Математическая формулировка задачи: обозначим через x_{11} – объем руды, добываемый на 1-м руднике и перерабатываемый на 1-й обогатительной фабрике; x_{12} – то же, на 1-м руднике и перерабатываемый на 2-й фабрике; x_{21} – то же, на 2-м руднике и перерабатываемый на 1-й фабрике; x_{22} – то же, на 2-м руднике и перерабатываемый на 2-й фабрике; x_{31} – то же, на 3-м руднике и перерабатываемый на 1-й фабрике; x_{32} – то же, на 3-м руднике и перерабатываемый на 2-й фабрике.

Тогда общий объем добычи руды соответственно по 1-му, 2-му и 3-му рудникам составит

$$y_1 = x_{11} + x_{12}; \quad y_2 = x_{21} + x_{22}; \quad y_3 = x_{31} + x_{32};$$

а объемы переработки руды по фабрикам:

$$Z_1 = x_{11} + x_{21} + x_{31}; \quad Z_2 = x_{12} + x_{22} + x_{32}.$$

С учетом введенных обозначений математическая модель задачи примет следующий вид.

Минимизировать $W = (7{,}5y_1 + 8{,}2y_2 + 2{,}4y_3) + (1{,}5x_{11} + 2x_{12} + 1{,}9x_{21} + 2{,}7x_{22} + 2{,}5x_{31} + 2{,}8x_{32} + (1{,}7Z_1 + 1{,}9Z_2)$ – затраты на добычу, транспорт и обогащение руды.

После преобразований целевая функция примет вид

$$W = 10{,}7x_{11} + 11{,}4x_{12} + 11{,}8x_{21} + 12{,}8x_{22} + 6{,}6x_{31} + 7{,}1x_{32} \to \min$$

при ограничениях:

$$\left.\begin{aligned} y_1 &\leqslant 1{,}2 \\ y_2 &\leqslant 0{,}9 \\ y_3 &\leqslant 6{,}5 \end{aligned}\right\}$$ (по максимальной производственной мощности рудников);

$$\left.\begin{aligned} Z_1 &\leqslant 3 \\ Z_2 &\leqslant 4 \end{aligned}\right\}$$ (по максимальной производственной мощности обогатительных фабрик);

$$\left.\begin{aligned} \frac{1{,}8x_{11} + 2{,}1x_{21} + 1{,}0x_{31}}{Z_1} &\geqslant 1{,}2 \\ \frac{1{,}8x_{11} + 2{,}1x_{21} + 1{,}0x_{31}}{Z_1} &\leqslant 1{,}3 \end{aligned}\right\}$$ по требованиям и качеству руды, поступающей на 1-ую обогатительную фабрику);

$$\frac{0{,}7x_{11} + 0{,}8x_{21} + 0{,}6x_{31}}{Z_1} \leqslant 0{,}6$$

$$\left.\begin{aligned} &\frac{1{,}8x_{12} + 2{,}1x_{22} + 1{,}0x_{32}}{Z_2} \geqslant 1{,}15 \\ &\frac{1{,}8x_{12} + 2{,}1x_{22} + 1{,}0x_{32}}{Z_2} \leqslant 1{,}25 \\ &\frac{0{,}7x_{12} + 0{,}8x_{22} + 0{,}6x_{32}}{Z_2} \leqslant 0{,}75 \end{aligned}\right\}$$ (по требованиям и качеству руды, поступающей на 2-ую обогатительную фабрику);

$0{,}76(1{,}8x_{11} + 2{,}1x_{21} + 1{,}0x_{31}) + 0{,}80(1{,}80x_{12} + 2{,}1x_{22} + 1{,}0x_{32}) \geqslant 6{,}5$ (по обязательности выполнения планового задания по выпуску металла в концентрате);

$x_{ij} \geqslant 0,\ i = \overline{1,3},\ j = \overline{1,2}$ (по неотрицательности переменных).

Все ограничения могут быть приведены к единым переменным x_{ij}, а ограничения по качеству существенно упрощены, как в предыдущем примере.

Пример IV.6. Решить задачу предыдущего примера, если сбыт концентрата не ограничен, а цена 1 т металла в концентрате 1200 руб.

В этом случае все ограничения задачи останутся без изменений, изменится лишь целевая функция. За критерий оптимальности следует принять прибыль комбината. Тогда целевая функция (с учетом размерности переменных и параметров модели) примет вид.

Максимизировать $П = 12\,[0{,}76\ (1{,}8x_{11} + 2{,}1x_{21} + 1{,}0x_{31}) + 0{,}80\,(1{,}80x_{12} + 2{,}1x_{22} + 1{,}0x_{32})] - [(7{,}5y_1 + 8{,}2y_2 + 2{,}4y_3) + (1{,}5x_{11} + 2{,}0x_{12} + 1{,}9x_{21} + 2{,}7x_{22} + 2{,}5x_{31} + 2{,}8x_{32}) + (1{,}7Z_1 + 1{,}9Z_2)]$.

Выражение в первых квадратных скобках – выручка от реализации металла в концентрации, а во вторых – затраты на его производство, включающие затраты на добычу, транспорт и обогащение (соответственно выражения в круглых скобках).

После преобразований целевая функция примет вид

$П = 5{,}7x_{11} + 5{,}9x_{12} + 7{,}4x_{21} + 7{,}4x_{22} + 2{,}4x_{31} + 2{,}5x_{32} \to \max.$

§ 7. МОДЕЛИ ЗАДАЧ РАЗМЕЩЕНИЯ

Задачи размещения – это определение мест строительства и мощности предприятий. Приведем простейший пример задачи размещения.

Имеется несколько месторождений полезного ископаемого. Затраты на добычу на каждом месторождении, а также на транспорт от каждого месторождения к потребителям продукции известны. Необходимо определить мощность предприятия на каждом месторождении. Если на некоторых месторождениях имеются действующие предприятия, по ним следует определить, насколько необходимо снизить или увеличить мощность.

Обозначим через i – месторождения полезного ископаемого, $i = \overline{1,n}$; C_i – затраты на добычу на i-м месторождении; $Q_i^{\max}$ – максимально возможная добыча с i-го месторождения; $Q_i^{\min}$ – минимальная добыча с i-го месторождения (для эксплуатируемых месторождений); j – потребитель продукции, $j = \overline{1,m}$; Q_j – потребность j-го потребителя; C_{ij} – затраты на транспорт с i-го месторождения к j-му потребителю.

Задача заключается в определении мощности предприятия на каждом месторождении и объемов перевозок с каждого месторождения до потребителя.

За управляемые переменные задачи примем: x_{ij} – объем продукции, перевозимой с i-го месторождения j-му потребителю; y_i – мощность предприятия на i-м месторождении,

$$y_i = \sum_{j=1}^{m} x_{ij}. \qquad \text{(IV.53)}$$

Критерий оптимальности – затраты на добычу полезного ископаемого

$$Z = \sum_{i=1}^{n} y_i\ C_i + \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}\ C_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}\ (C_i + C_{ij}) \to \min \qquad \text{(IV.54)}$$

при ограничениях:

по мощности добывающих предприятий

$$Q_i^{\max} \geqslant y_i \geqslant Q_i^{\min},\ i = \overline{1,n}; \qquad \text{(IV.55)}$$

по обеспечению потребителей

$$\sum_{i=1}^{n} x_{ij} = Q_j,\ j = \overline{1,m}. \qquad \text{(IV.56)}$$

В самом простом случае; т.е. когда затраты постоянны и не зависят от мощности предприятия (что является обычно грубым допущением), задача является линейной и может быть сведена к чисто транспортной в результате объединения затрат на добычу и транспорт, т.е. введение новых стоимостных показателей

$$C'_{ij} = C_i + C_{ij}. \qquad \text{(IV.57)}$$

Пример IV.7. В связи с ростом промышленного и жилищного строительства в области намечается реконструкция двух заводов железобетонных изделий и увеличение их мощности. Для обеспечения потребности заводов в щебне могут быть реконструированы два действующих карьера и освоено еще одно месторождение.

Потребность заводов в щебне составляет соответственно 1 и 1,5 млн.м^3.

На действующих карьерах достигнуты мощности 0,5 и 0,6 млн.м^3 и могут быть увеличены до 1 млн.м^3 на каждом. На вновь осваиваемом месторождении может быть построен карьер и дробильно-сортировочная фабрика мощностью до 1,2 млн.м^3.

Затраты на выпуск 1 м^3 щебня после реконструкции и строительства соответственно по карьерам составят 3,4; 3,7; 4,1 руб. Затраты на транспорт от карьера до железобетонных заводов приведены в табл. IV.5.

Таблица IV.5

Поставщик (карьер)	Потребители (заводы)	
	1	2
1	2,5	2,8
2	2,3	2,9
3 (новый)	2,6	3,0

Определить мощности предприятий на эксплуатируемых и намечаемом к освоению месторождениях.

Словесная формулировка задачи: определить мощности предприятий на каждом месторождении и объемы перевозок продукции потребителям при обеспечении плановой потребности в продукции с минимальными затратами.

Математическая формулировка задачи: обозначим через x_{11} – объем выпуска щебня на 1-м карьере, отправляемый 1-му заводу; x_{12} – то же, на 1-м карьере, отправляемый 2-му заводу; x_{21} – то же, на 2-м карьере, отправляемый 1-му заводу; x_{22} – то же, на 2-м карьере, отправляемый 2-му заводу; x_{31} – то же, на новом месторождении, отправляемый 1-му заводу; x_{32} – то же, на новом месторождении, отправляемый 2-му заводу.

При условии этих обозначений математическая модель принимает следующий вид.

Минимизировать $Z = [3{,}4(x_{11} + x_{12}) + 3{,}7(x_{21} + x_{22}) + 4{,}1(x_{31} + x_{32})] + [2{,}5x_{11} + 2{,}8x_{12} + 2{,}3x_{21} + 2{,}9x_{22} + 2{,}6x_{31} + 3{,}0x_{32}]$,

где Z – общие затраты на выпуск и транспорт щебня потребителю при ограничениях:

$$\left.\begin{array}{l} x_{11} + x_{12} \geqslant 0{,}5 \\ x_{21} + x_{22} \geqslant 0{,}6 \end{array}\right\}$$ (по загрузке имеющихся мощностей)

$$\left.\begin{array}{l} x_{11} + x_{12} \leqslant 1{,}0 \\ x_{21} + x_{22} \leqslant 1{,}0 \\ x_{31} + x_{32} \leqslant 1{,}2 \end{array}\right\}$$ (по максимальной мощности предприятий);

$$\left.\begin{array}{l} x_{11} + x_{21} + x_{31} = 1{,}0 \\ x_{12} + x_{22} + x_{32} = 1{,}5 \end{array}\right\}$$ (по обеспечению потребителей);

$x_{ij} \geqslant 0$; $i = \overline{1,3}$; $j = \overline{1,2}$ (по неотрицательности решения).

Контрольные вопросы и задачи

1. Условия использования линейных экономико-математических моделей при решении задач планирования, организации и управления горным производством?
2. В чем особенность задач о расстановке оборудования?
3. Общая характеристика задач оптимального использования ресурсов.
4. В чем особенность задач планирования добычных работ в режиме усреднения качества? К какому типу задач они относятся?
5. Каким образом задачи о размещении предприятий сводятся к транспортным?
6. К какому классу задач относятся задачи планирования работы группы горных предприятий?

Таблица IV.6

Вид металла	Содержание металла в концентрате на рудниках, %			
	1	2	3	4
1-й	2,8	–	–	6,8
2-й	1,2	2,4	–	3,6
3-й	–	0,3	0,8	1,6
Затраты на транспортирование и переработку 1 т руды, руб.	1,15	0,72	1,36	0,94
Максимальный объем поставок с рудника, т	7000	12 000	5000	6000

7. Задача. Определить объемы поставок руды на обогатительную фабрику с четырех рудников, обеспечивающие выполнение задания по количеству металлов в концентрате при минимальных затратах на транспорт и переработку руды.

Варианты решения задачи отличаются набором рассматриваемых (находящихся в работе) рудников и заданием по количеству каждого вида металла в концентрате. Исходные данные по содержанию металла в концентрате по рудникам и их технико-экономические показатели приведены в табл. IV.6.

Дополнительные данные, определяющие вариант задачи, приведены в табл. IV.7.

Таблица IV.7

№ варианта	Рассматриваемые рудники	Задания по количеству металлов в концентрате по видам, т			
		1-й	2-й	3-й	4-й
1	1 2 3	490	300	–	–
2	1 2 4	450	500	–	–
3	1 3 4	500	–	–	500
4	2 3 4	400	–	120	–
5	1 2; 3 4	500	500	140	1230
6	1 2 3	–	320	70	–
7	1 2 4	–	550	140	800
8	1 3 4	550	250	80	–
9	2 3 4	300	450	110	250
10	1 2; 3 4	500	500	–	–
11	1 2 3	140	290	70	1220
12	1 2 4	300	–	100	–
13	1 3 4	–	–	120	400
14	2 3 4	290	420	–	–
15	1 2; 3 4	460	–	100	–

8. Задача. На руднике добыча руды ведется на двух участках. Содержание металла в блоках участков соответственно 1,2 и 1,4 %. Производительность (в тыс.т/смену) погрузочного оборудования 0,1 и 0,12, а доставочного – 0,035 и 0,03. В распоряжении начальника рудника 10 погрузочных и 30 транспортных машин. Максимальная добыча с участков 0,7 и 0,8 тыс.т соответственно. Содержание металла в добываемой руде должно быть в пределах 1,27–1,32 %. Рассчитать, как следует распределить оборудование по участкам, чтобы максимизировалась добыча.

9. Задача. На карьере два добычных участка, максимальная производительность которых в смену 1,5 и 1,8 тыс.т соответственно. Содержание полезного компонента в руде на участке № 1–28 % и на № 2 – 32 %. Производительность автомашин на участках 0,3 и 0,25 % тыс.т/смену. На обогащение должна поступать руда содержанием не менее 30 %. Всего в работе 10 автомашин. Как распределить транспорт между участками, чтобы максимизировать: добычу руды; объем металла в добытой руде.

10. Задача. Добыча угля на шахте осуществляется из трех лав. Максимальная сменная добыча каждой лавы составляет 800, 1100 и 1200 т, а зольность добываемого угля в лавах 14, 18 и 21 % соответственно. Какое максимальное количество угля может быть добыто, чтобы его зольность не превышала 20 %?

На сколько уменьшится объем добычи, если зольность угля должна быть не менее 16 %?

11. Задача. На обогатительную фабрику годовой производственной мощностью до 3 млн.т руда поступает из двух рудников. Фабрика перерабатывает руду с содержанием меди 1,2–1,3 %, а серы – не более 1 %. Минимальный объем выпуска металла 37 тыс.т. Максимальная производственная мощность рудников 1,7 и 1,8 млн.т/год, затраты на добычу 1 т руды на руднике № 1–5 руб, на № 2–6 руб. Содержание меди и серы составляет на руднике № 1 - 1.1 % и 0,8 %, на № 2–1,4 % и 1,1 %. Определить:

какой объем руды необходимо добыть на каждом руднике для выполнения плана по выпуску металла;

план, оптимальный по двум критериям (объем выпуска металла и затраты);

какой максимальный объем руды может быть переработан на фабрике.

12. Задача. Завод выпускает два вида изделий (А и Б), которые обрабатываются на токарных и фрезерных станках. Затраты времени на обработку детали А на токарном станке – 1 ч, на фрезерном – 1,2 ч, детали Б – соответственно 1,3 и 1,4 ч. Отпускная цена детали А –17 руб., Б – 22 руб.

Всего на заводе 3 токарных и 4 фрезерных станка (двухсменные). Определить, какое количество деталей А и Б целесообразно выпускать в неделю (6 рабочих дней).

13. Задача. Требуется определить годовые объемы добычи руды по трем предприятиям, данные о которых приведены в табл.IV.8.

Всего в работе находится 13 составов. Руда со всех трех рудников перерабатывается на одной обогатительной фабрике, причем среднее ее содержание должно быть в пределах 6,4–6,6 %.

Таблица IV.8

Показатель	№ 1	№ 2	№ 3
Максимальная добыча, млн.т/год	16,7	17,4	15,9
Годовая производительность состава, млн.т	2,5	2,35	2,5
Содержание металла, %	5,9	6,4	7,7
Прибыль, получаемая от добычи и переработки 1 млн.т руды, млн.руб.	8,8	9,6	12,12

Таблица IV.9

Химический элемент сплава	Содержание элемента в сплаве, кг	Содержание (в %) элементов в отходах плавок				Цена 1 кг покупаемого элемента, руб.
		1-й	2-й	3-й	4-й	
Олово	70	78	85	72	75	0,81
Сурьма	15	10	6	16	9	0,12
Медь	11	9	7	11	14	0,17
Свинец	4	3	2	1	2	0,25
Масса отходов, кг	–	23	20	18	80	–
Цена 1 кг отходов, руб.	–	0,25	0,31	0,27	0,26	

14. Задача. Для получения сплава используются отходы четырех ранее проведенных плавок (химический состав сплава и содержание элементов в отходах по каждой плавке приведены в табл. IV.9). Могут быть использованы и покупаемые чистые химические элементы, цены на которые также приведены в табл. IV.9.

Определить, из какого количества каждого вида отходов и чистых химических элементов нужно составить сплав массой 100 кг, чтобы стоимость его была минимальной.

15. Задача. Горно-обогатительный комбинат имеет в своем составе три рудника, характеристика которых приведена в табл. IV.10.

Среднее содержание никеля в добываемой руде должно быть в пределах 0,70–0,75 %, меди 1,62–1,66 %. Минимальный план добычи (по металлам): никеля – 30 тыс., меди – 40 тыс.т.

Каков должен быть объем добычи по рудникам, чтобы:

извлечь максимум меди;

извлечь максимум никеля;

выполнить план с минимальными затратами;

выполнить план с минимальной численностью.

Таблица IV.10

Показатель	№ 1	№ 2	№ 3
Максимальная производственная мощность, млн.т	1,7	0,9	1,5
Себестоимость 1 т руды, руб.	3.6	4,2	2,9
Содержание никеля, %	0.72	0,81	0,64
Содержание меди, %	1,6	1,5	1.8
Производительность труда, тыс.т	1,4	1,2	1.3

Таблица IV.11

Ресурсы	Затраты на 1 ед. изделия			
	Р1	Р2	Р3	Р4
Сырье 1-го вида	4	2	8	11
Сырье 2-го вида	2	4	3	1
Трудовые	4	5	7	9
Прибыль, руб/шт	12	6	14	11

16. Задача. Предприятие выпускает четыре вида продукции: Р1, Р2, Р3 и Р4. При их изготовлении используются два вида основного сырья, причем предприятие имеет в наличии 1500 т одного вида сырья и 1000 т другого. Кроме того, трудовые ресурсы составляют 800 чел-дней.

Нормативы затрат каждого из видов ресурсов на изготовление 1 ед. изделия и прибыль на 1 ед. изделия приведены в табл. IV.11.

Требуется:

а) определить оптимальный ассортимент, максимизирующий прибыль;

б) решить ту же задачу при дополнительных условиях, налагаемых на ассортимент (в ед.): продукции Р1 $\geqslant$42, Р2 $\geqslant$125, Р3 $\geqslant$30; Р4 $\geqslant$14;

в) решить задачу "а" при условии комплектности: Р1 : Р2 = 1 : 6;

г) заданы дополнительно цены (в руб.): Р1 – 30; Р2 – 16; Р3 – 11; Р4 –80. Определить оптимальный ассортимент, максимизирующий товарную продукцию, при единственном ограничении на ассортимент – условии комплектности Р1 : Р2 – 1 : 6.

17. Задача. Угольная шахта работает в комплексе с обогатительной фабрикой. Среднесуточная добыча 3300 т, плановая зольность угля 19,2 %. Весь уголь с шахты передается на обогащение, поэтому ежесуточно для обеспечения постоянной зольности перерабатываемого сырья задания по качеству угля корректируются. В результате поступления партии угля с высокой зольностью обогатительная фабрика потребовала снизить на следующие сутки зольность добываемого угля до 17 %. В связи с этим требуется скорректировать суточные задания добычным участкам шахты таким образом, чтобы снижение добычи в целом по шахте было минимальным. Показатели работы участков шахты приведены в табл. IV.12.

Нагрузки на участок можно увеличить за счет перераспределения порожних вагонеток и людских ресурсов. Уголь с участков № 1 и 2 транспортируется по конвейерной линии, имеющей суточную производительность не более 1870 т, с участков № 3 и 4 – не более 1680 т.

Таблица IV.12

№ участка	Плановая суточная нагрузка на участок т	Зольность добываемого угля, %	Максимальная возможная суточная нагрузка на участок, т
1	950	21	1100
2	800	15	940
3	850	22	1000
4	800	18	900

Таблица IV.13

Ресурсы	П1	П2	П3	П4	Объем ресурса
Трудовые, чел-недели	2	1	1	2	18
Полуфабрикаты, кг	7	5	6	4	115
Станочное оборудование, станко-смены	5	6	11	15	110
Цена 1 шт., руб.	65	70	130	120	–

18. Задача. Предприятие может выпускать четыре вида продукции – П1, П2, П3 и П4 без ограничения на количественный ее выпуск. Для изготовления этой продукции предприятие в течение месяца располагает ресурсами. Для изготовления 1 шт. продукции известен объем каждого ресурса. Данные приведены в табл. IV. 13.

Требуется спланировать выпуск таким образом, чтобы получаемый доход был максимальным.

19. Задача. Определить суточные нагрузки на группу добычных участков угольной шахты, обеспечивающие перевыполнение задания по суммарной добыче угля, с учетом требований по минимально необходимым подвиганиям очистных забоев и установленной средней зольности добываемого угля.

Суточное подвигание очистного забоя σ определяется из выражения

$$\sigma = A/(Pl),$$

где A – нагрузка на забой, т; P – производительность пласта, т/м^2; l – длина лавы, м.

Показатели очистных участков шахты приведены в табл. IV.14; дополнительные данные, определяющие вариант задачи – в табл. IV.15.

20. Задача. Определить месячную производительность рудников, обеспечивающую план поставки сырой руды на обогатительную фабрику с заданным содержанием компонентов в руде и с минимальными суммарными эксплуатационными

Таблица IV.14

№ участка	Производительность пласта, т/м^2	Длина лавы, м	Зольность угля, %	Минимально необходимое подвигание забоя, м/сут	Допустимая нагрузка на забой, т/сут
1	1,69	120	25,0	1.26	1110
2	1,82	100	31.3	1.26	1200
3	1,95	110	26.6	1.26	900
4	1,56	125	24.8	1.89	700
5	1,95	130	32.8	0.63	2000
6	1,49	100	35,0	0,00	500
7	1,82	115	29.5	1,10	800

Таблица IV.15

№ варианта	№ рассматриваемых участков	Суммарная добыча угля участками, т	Установленная зольность угля, %
1	1, 2, 3, 4	3500	27
2	1, 2, 3, 5	4500	28
3	1, 2, 3, 6	3200	26
4	1, 2, 4, 5	4000	29
5	1, 2, 4, 6	2800	26
6	1, 2, 3, 7	3500	28
7	1, 2, 4, 7	3300	26
8	1, 2, 5, 6	4400	27
9	1, 2, 5, 7	2700	28
10	1, 2, 6, 7	4600	27
11	1, 3, 4, 5	4100	27
12	1, 3, 4, 6	2900	26

затратами на ее добычу. Обогатительной фабрикой предъявлены следующие требования к качеству исходного сырья:

среднее содержание 1-го компонента $\alpha_1 = \alpha_{пл}$;
" " 2-го " $\alpha_{2\,min} \leqslant \alpha_2 \leqslant \alpha_{2\,max}$;
" " 3-го " $\alpha_3 \leqslant \alpha_{3\,доп}$.

Технико-экономические показатели рудников приведены в табл. IV.16; дополнительные данные, определяющие вариант задачи – в табл. IV.17.

21. Задача. На обогатительной фабрике имеются три склада сырья объемом A_1, A_2 и A_3, на которых в данный момент находится уголь средней зольности Z_1, Z_2 и Z_3. Уголь поступает на две технологические линии пропускной способностью $П_1$ и $П_2$, рассчитанные на переработку угля средней зольности Z_1' и Z_2'. Составить план переработки угля на следующие сутки, обеспечивающий полную загрузку технологических линий для максимального освобождения указанного угольного склада.

Таблица IV.16

№ рудника	Содержание компонента в руде, %			Себестоимость добычи, руб.	Допустимая нагрузка на рудник, тыс.т
	1-го	2-го	3-го		
1	1,2	0,5	3,2	1,25	50
2	2,1	0,4	3,1	1,10	45
3	1,6	0,3	4,1	0,81	36
4	2,6	0,8	3,6	1,05	55
5	3,0	0,4	2,8	1,51	60
6	1,8	0,6	3,8	0,72	30
7	2,3	0,7	3,5	0,95	40

Таблица IV.17

№ варианта	Рассматриваемая группа рудников	Допустимые уровни содержания компонента в усредненной руде, %				Объем поставок руды, тыс.т
		$\alpha_{пл}$	$\alpha_{2\,min}$	$\alpha_{2\,min}$	$\alpha_{3\,доп}$	
1	1, 2, 3, 4	1,3	0,4	0,5	–	160
2	1, 2, 3, 5	2,3	0,4	0,5	–	155
3	1, 2, 3, 6	1,9	0,6	0,7	–	140
4	1, 2, 3, 7	2,0	–	0,6	3,9	150
5	3, 4, 6, 7	1,9	–	0,7	3,9	125
6	3, 5, 6, 7	2,4	0,4	–	3,8	140
7	4, 5, 6, 7	2,5	0,5	0,7	–	130

Исходные данные по вариантам приведены в табл.IV.18.

22. Задача. Обогатительная фабрика имеет запас рядового угля на трех складах в объемах A_1, A_2, A_3 средней зольностью Z_1, Z_2 и Z_3. Уголь подается на две технологические линии пропускной способностью $П_1$ и $П_2$.

Требуется составить план переработки угля на следующие сутки, обеспечивающий максимальный выпуск концентрата, если выход концентрата Δ т/т из усредненного угля с зольностью Z' известен для каждой технологической линии. Исходные данные приведены в табл.IV.19.

Таблица IV.18

№ варианта	Количество угля на складах, тыс.т			Зольность угля на складах, %			Производительность линий, тыс. т/сут		Расчетная зольность угля, %		№ максимально освобождаемого склада
	A_1	A_2	A_3	Z_1	Z_2	Z_3	$П_1$	$П_2$	Z_1'	Z_2'	
1	5	6	4	12	18	27	3	3	15	20	3
2	4	2	5	13	19	27	3	2	15	22	2
3	6	5	5	12	19	27	2	2	15	24	2
4	4	5	6	14	18	25	3	4	16	23	1
5	8	10	3	15	20	25	2	4	16	23	3
6	4	3	2	13	20	25	2	1,5	15	22	3
7	2	5	1	14	21	30	2	2	17	25	3
8	1	4	4	15	21	27	1	2	19	24	1
9	3	2	4	16	22	30	2	1	19	27	1
10	1	2	5	16	22	30	2	2	18	25	2
11	6	7	4	13	19	28	2	4	16	21	2

Таблица IV.19

№ варианта	Количество угля на складах, тыс. т			Зольность угля на складах, %			Пропускная способность линий, тыс. т/сут		Требования по усреднению угля, %		Выход концентрата, т/т	
	A_1	A_2	A_3	Z_1	Z_2	Z_3	$П_1$	$П_2$	y	y_1	Δ	Δ'
1	4	3	4	12	18	25	1,2	3,0	15	23	0,80	0,63
2	5	3	2	12	17	25	1,5	2,0	15	23	0,81	0,65
3	4	3	3	12	18	27	2,0	2,0	15	23	0,81	0,65
4	2	1	3	12	18	27	2,2	2,2	16	23	0,79	0,65
5	3	2	2	13	17	28	1,0	1,4	16	23	0,82	0,63
6	1	3	5	12	19	28	2,0	2,0	15	24	0,80	0,61

23. Задача. На разрезе работают три экскаватора. Уголь отгружается на склад, на котором имеются свободные штабеля: два пустых, вместимостью по 30 тыс.т и один полупустой – 12 тыс.т. Последний загружен 18 тыс.т угля с зольностью $A^c \leqslant 42$ % и низшей рабочей теплотой сгорания $Q^p = 17\,253$ кДж/кг. На смену запланировано добыть 45 тыс.т угля с $A^c \leqslant 41{,}2$ % и $Q^p \geqslant 17\,052$ кДж/кг. Эти показатели должны выдерживаться в каждом штабеле. Каждый экскаватор может грузить во все штабеля.

Рассчитать такие нагрузки на экскаваторы и объемы, отгружаемые в каждый из штабелей, чтобы добыча по разрезу была максимальной при условии отгрузки экскаваторами угля качеством: первым – $A^c = 43{,}8$ % и $Q^p = 16\,056$ кДж/кг; вторым – $A^c = 40{,}2$ % и $Q^p = 17\,438$ кДж/кг; третьим – $A^c = 38{,}2$ % и $Q^p = 17\,984$ кДж/кг.

Глава V

ДВОЙСТВЕННОСТЬ В ЛИНЕЙНОМ ПРОГРАММИРОВАНИИ И ОЦЕНКА РЕЗУЛЬТАТОВ РЕШЕНИЯ ЛИНЕЙНЫХ ЗАДАЧ

§ 1. ПРЯМАЯ И ДВОЙСТВЕННАЯ ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

С любой задачей линейного программирования, называемой исходной, связана другая линейная, именуемая двойственной. Решая одну из задач, мы получаем также и решение другой.

Пусть исходная задача линейного программирования записана в форме матрицы для модифицированных жордановых исключений (табл.V.1).

Таблица V.1

	x_1	$-x_2$	...	$-x_j$	...	$-x_n$	1
$y_1 =$	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}	b_1
$y_2 =$	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}	b_2
...	...	...	...	...	...	...	...
$y_i =$	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}	b_i
...	...	...	...	...	...	...	...
$y_m =$	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}	b_m
$Z =$	$-P_1$	$-P_2$	...	$-P_j$	...	$-P_n$	0

Сопоставим независимым переменным этой задачи x_j новые зависимые переменные v_j, а каждой зависимой переменной y_i новые независимые переменные u_i. Столбцу свободных членов сопоставим функцию W, а Z строке сопоставим единицу. В результате исходная матрица примет вид табл.V.2.

Новая таблица содержит одновременно условия двух сопряженных задач – исходной и двойственной.

Для получения двойственной задачи отбросим старые переменные x_j и y_i и транспонируем матрицу (V.2).

В результате получим новую матрицу (табл.V.3).

В этой матрице записаны условия следующей задачи линейного программирования

Таблица V.2

	v_1	v_2	...	v_j	...	v_n	W
	$-x_1$	$-x_2$	...	$-x_j$	...	$-x_n$	1
$u_1 y_1 =$	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}	b_1
$u_2 y_2 =$	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}	b_2
...	...	...	...	...	...	...	...
$u_i y_i =$	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}	b_i
...	...	...	...	...	...	...	...
$u_m y_m =$	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}	b_m
$Z =$	$-P_1$	$-P_2$	...	$-P_j$	...	$-P_n$	0

Таблица V.3

	u_1	u_2	...	u_i	...	u_m	1
$v_1 =$	a_{11}	a_{21}	...	a_{i1}	...	a_{m1}	$-P_1$
$v_2 =$	a_{12}	a_{22}	...	a_{i2}	...	a_{m2}	$-P_2$
...	...	...	...	...	...	...	...
$v_j =$	a_{1j}	a_{2j}	...	a_{ij}	...	a_{mj}	$-P_j$
...	...	...	...	...	...	...	...
$v_n =$	a_{1n}	a_{2n}	...	a_{in}	...	a_{mn}	$-P_n$
$W =$	b_1	b_2	...	b_i	...	b_m	0

$$W = b_1u_1 + b_2u_2 + \ldots + b_iu_i + \ldots + b_mu_m \to \min,$$
$$v = a_{1j}\,u_1 + a_{2j}\,u_2 + \ldots + a_{ij}\,u_i + \ldots + a_{mj}\,u_m \geqslant P_j,$$
$$j = (1, 2, \ldots, n).$$

Эта задача является двойственной по отношению к исходной

$$Z = P_1 x_1 + P_2 x_2 + \ldots + P_j x_j + \ldots + P_n x_n \to \max,$$
$$a_{i1} x_1 + a_{i2} x_2 + \ldots + a_{ij} x_j + \ldots + a_{in} x_n \leqslant b,$$
$$i = (1, 2, \ldots, m),$$

записанной в матрице V.1. При сопоставлении матриц V.1 и V.3 видно, что записывая прямую задачу в виде матрицы для модифицированных жордановых исключений, мы одновременно записываем и двойственную в виде матрицы для обыкновенных жордановых исключений. Если исходная задача решается с помощью модифицированных жордановых исключений, то одновременно решается и двойственная задача, но уже с помощью обыкновенных жордановых исключений.

Прямая и двойственная задачи могут быть объединены в одну симплексную матрицу (например, табл.V.2).

Особенности двойственной задачи линейного программирования заключаются в следующем.

1. Если исходная является задачей максимизации, то двойственная – задачей минимизации.

2. Коэффициенты целевой функции исходной задачи становятся свободными членами ограничений двойственной.

3. Свободные члены ограничений исходной задачи становятся коэффициентами целевой функции двойственной.

4. Матрицу ограничений двойственной задачи получают транспонированием матрицы ограничений исходной.

5. Знаки неравенств ограничений меняются на обратные.

6. Число ограничений исходной задачи равно числу ограничений двойственной, а число ограничений двойственной равно числу ограничений исходной задачи.

Пример V.1. Составить задачу, двойственную следующей

$$Z = 4x_1 + 5x_2 + 9x_3 \to \max,$$
$$y_1 = x_1 + x_2 + 2x_3 \leqslant 16,$$
$$y_2 = 7x_1 + 5x_2 + 3x_3 \leqslant 25,$$
$$x_{1,2,3} \geqslant 0.$$

Так как прямая задача на максимум, следовательно, двойственная будет на минимум, а коэффициенты целевой функции образуются из правых частей ограничений (соответственно 16 и 25). Следовательно, целевая функция двойственной задачи принимает вид

$$W = 16y_1 + 25y_2 \to \min.$$

Так как в прямой задаче три переменных, следовательно, двойственная задача имеет три ограничения-неравенства. Причем, если в прямой задаче ограничения типа $\leqslant$, то в двойственной они становятся $\geqslant$. Правые части ограничений двойственной задачи равны коэффициентам целевой функции прямой задачи, соответственно 4, 5 и 9.

Учитывая, что коэффициенты в ограничениях двойственной задачи – результат транспонирования матрицы ограничений прямой, получим систему ограничений двойственной задачи

$$y_1 + 7y_2 \geqslant 4,$$
$$y_1 + 5y_2 \geqslant 5,$$
$$2y_1 + 3y_2 \geqslant 9.$$

Целевая функция W и система ограничений дают в целом модель двойственной задачи.

§ 2. ТЕОРЕМЫ ДВОЙСТВЕННОСТИ И ИХ ЭКОНОМИЧЕСКИЙ СМЫСЛ

Первая теорема двойственности. Если исходная задача имеет оптимальное решение, то и двойственная имеет оптимальное решение, причем оптимальные значения линейных форм равны: $\max Z = \min W$.

Если одна из двойственных задач имеет неограниченную линейную форму, то она противоречива.

Последняя часть теоремы не допускает обращения. Если в одной из задач условия несовместны (противоречивы), то и в другой они могут быть также противоречивы.

Вторая теорема двойственности. Для того чтобы два допустимых решения

$$x = (x_1, x_2, \ldots, x_j, \ldots, x_m) \text{ и } u = (u_1, u_2, \ldots, u_i, \ldots, u_m)$$

пары двойственных задач были оптимальны, необходимо и достаточно выполнения условий дополняющей нежесткости

$$x_j \left(\sum_{i=1}^{m} a_{ij} u_i - p_j \right) = 0, \; j = \overline{1, m}; \qquad \text{(V.1)}$$

$$u_i = \left(\sum_{j=1}^{m} a_{ij} x_j - b_i \right) = 0, \; i = \overline{1, m}, \qquad \text{(V.2)}$$

т.е., чтобы равнялись нулю произведения переменных одной из задач на разность правой и левой частей соответствующего ограничения двойственной задачи.

Если хотя бы одно оптимальное решение одной из двойственных задач обращает i-е ограничение этой задачи в строгое неравенство, то i-я компонента каждого оптимального решения второй двойственной задачи равна 0.

Если же i-я компонента хотя бы одного оптимального решения двойственной задачи больше 0, то каждое оптимальное решение другой двойственной задачи обращает i-е ограничение этой задачи в строгое равенство.

Рассмотрим экономический смысл теории двойственности на примере задачи оптимального планирования выпуска продукции и использования ресурсов. Объединение может выпускать несколько видов продукции ($i = 1, 2, ..., n$). Каждая продукция требует затрат определенных ресурсов (трудовых, материальных, сырьевых). Количество ресурсов каждого вида ($j = 1, 2, ..., m$) ограничено и составляет b_j.

Пусть p_i – прибыль от реализации соответствующего вида продукции.

Определим, какое количество продукции каждого вида целесообразно произвести, чтобы получить максимальную прибыль, с учетом имеющегося количества ресурсов.

За управляемые переменные примем объем выпуска каждого вида продукции x_i. Тогда целевая функция примет вид

$$Z = p_1 x_1 + p_2 x_2 + \ldots + p_i x_i + \ldots + p_n x_n \to \max. \qquad \text{(V.3)}$$

Ограничением при решении задачи выступает наличное количество ресурсов. Если a_{ij} – расход ресурсов j-го типа на выпуск i-й продукции, то ограничения можно записать следующим образом

$$a_{1j} x_1 + a_{2j} x_2 + \ldots + a_{ij} x_i + a_{nj} x_n \leqslant b_j, \; j = 1, 2, \ldots, m. \qquad \text{(V.4)}$$

Данная задача является исходной и позволяет найти оптимальный план выпуска продукции.

Предположим, что предприятие может дополнительно приобретать ресурсы (т.е. за счет каких-либо мероприятий увеличивать их количество). Примем цену j ресурса за u_j. Тогда общие затраты на ресурсы составят

$$u_1 b_1 + u_2 b_2 + \ldots + u_j b_j + \ldots + u_m b_m. \qquad \text{(V.5)}$$

Естественно, что эти затраты необходимо минимизировать. Цены на ресурсы должны быть выбраны таким образом, чтобы прибыль от выпуска единицы каждого вида продукции не уменьшалась

$$a_{i1} u_1 + a_{i2} u_2 + \ldots + a_{ij} u_j + \ldots + a_{im} u_m \geqslant p_i, \; (i = 1, 2, \ldots, n). \qquad \text{(V.6)}$$

Таким образом получаем новую модель линейной задачи

$$u_1 b_1 + u_2 b_2 + \ldots + u_j b_j + \ldots + u_m b_m \to \min, \qquad \text{(V.7)}$$

$$a_{1j} u_1 + a_{2j} u_2 + \ldots + a_{ij} u_j + \ldots + a_{nj} u_n \geqslant p_i, \; (i = 1, 2, ..., n), \qquad \text{(V.8)}$$

являющуюся двойственной по отношению к первой.

Таким образом, исходной задаче оптимального использования ресурсов (оптимального плана выпуска продукции) соответствует двойственная задача обоснования оптимальных затрат (цен) на используемые ресурсы. При этом следует подчеркнуть, что оценка ресурсов имеет относительный характер только с точки зрения прибыльности получаемой объединением продукции. Эти оценки называются неявными, теневыми ценами.

Между исходной и двойственной задачами существует не только математическая, но и экономическая связь. Если исходная задача имеет оптимальное решение, то и двойственная, согласно первой теореме двойственности, имеет оптимальное решение и значение целевых функций совпадает. Значит, оптимальному плану выпуска продукции соответствует система минимальных оценок ресурсов, причем прибыль, получаемая при составлении оптимального плана, равна суммарной оценке ресурсов. Следовательно, оценки ресурсов, получаемые при решении двойственной задачи, являются экономическим инструментом, обеспечивающим выполнение оптимального плана, балансирующим затраты и расходы.

Проанализируем экономический смысл второй теоремы двойственности, согласно которой произведение переменной одной из задач на разность левой и правой частей соответствующего ограничения двойственной задачи равно нулю. Рассмотрим переменные двойственной задачи – оценки ресурсов. Оценки ресурсов не равны 0 только в том случае, если соответствующее ограничение по количеству ресурсов обращается в строгое равенство, т.е. для данного объединения представляют ценность только дефицитные полностью используемые ресурсы. Если ресурсы используются не полностью, то их оценка равна нулю, т.е. нет смысла приобретать ресурсы и платить за них средства, если имеющиеся у нас в наличии мы используем не полностью. Таким образом, оценка ресурсов в оптимальном плане характеризует меру их дефицитности. Для переменных основной задачи – объемов вы-

пуска продукции, в качестве условий дополняющей нежесткости выступает разность между оценкой ресурсов, затрачиваемых на данную продукцию, и прибылью от ее выпуска.

Если затраты на ресурсы больше прибыли, то выпуск такой продукции будет равен 0, т.е. в оптимальный план попадает только рентабельная (с точки зрения оценки ресурсов) продукция. В этом случае ограничение двойственной задачи – строгое равенство (затраты не превышают прибыльности продукции).

Экономическое содержание двойственных задач делает их важным инструментом социалистического хозяйствования.

§ 3. АНАЛИЗ ЧУВСТВИТЕЛЬНОСТИ РЕШЕНИЯ ЛИНЕЙНЫХ ЭКОНОМИЧЕСКИХ ЗАДАЧ

При решении любой оптимизационной задачи важно не только найти оптимальное значение управляемых переменных, но и проанализировать, насколько устойчиво оптимальное решение к изменению параметров задачи.

При решении задачи линейного программирования удельные стоимостные показатели $(c_1, c_2, \ldots, c_j, \ldots, c_n)$, входящие в целевую функцию, уровень запасов ресурсов b_i и нормы их расхода a_{ij} приняты строго постоянными. На практике же технико-экономические показатели (например, c_j и a_{ij}) определяются с некоторой погрешностью. Кроме того, может меняться и уровень запаса ресурсов b_i.

Поэтому, исследуя линейную экономическую модель, желательно также определять:

изменится ли оптимальное решение при колебании стоимостных показателей или норм расхода ресурсов;

до каких пределов изменения стоимостных показателей и объема расходуемых ресурсов не влияют на решение задачи;

количество каких ресурсов следует увеличивать для максимального повышения эффективности решения всей задачи.

Специальные расчеты, позволяющие ответить на эти и подобные вопросы, представляют собой анализ чувствительности модели.

Если поиск оптимума по линейной экономической модели в чисто детерминированной постановке представляет решение так называемой обратной задачи исследования операций, то анализ чувствительности сводится к решению прямой задачи исследования операций, когда параметры c_j, a_{ij}, b_i изменяются в определенном диапазоне.

Теория и методы анализа чувствительности, а также решения задач с переменными величинами c_j, a_{ij}, b_i пока еще не разработаны в полной мере, но некоторые этапы исследования модели могут быть проведены путем изучения симплекс-таблиц.

При решении линейной модели наряду с исходными переменными (x_j) получают значения вспомогательных переменных y_i. В тех

случаях, когда величина y_i не равна 0, ее значение показывает объем неиспользуемых ресурсов и, следовательно, насколько без ущерба для решения задачи они могут быть уменьшены.

Важное значение имеет также решение двойственных задач. Как показано в гл. V § 3, переменные этой задачи характеризуют дефицитность ресурсов и величину изменения целевой функции при увеличении их на единицу. Таким образом, решение двойственной задачи позволяет ответить на вопрос о целесообразности изменения запасов ресурсов для повышения эффективности ее решения.

Кроме того, разрабатываются специальные методы решения задач, когда параметры модели (c_j, a_{ij}, b_i) изменяются в некотором диапазоне. Теоретически достаточно хорошо изучены случаи, когда c_j и b_i являются аналитическими (обычно линейными) функциями. Эти задачи составляют специальный раздел – параметрическое программирование.

Различают прямую и двойственную задачи параметрического программирования.

В прямой задаче рассматривается случай, когда коэффициенты целевой функции c_j линейно зависят от параметра t, изменяющегося в определенных пределах. Модель прямой задачи параметрического линейного программирования имеет следующий вид.

Максимизировать целевую функцию

$$Z = \sum_{j=1}^{n} (P_j + t q_i) x_j \to \max, \; t \in [\lambda_1, \lambda_2] \qquad \text{(V.9)}$$

при ограничениях:

$$\sum_{j=1}^{n} a_{ij} x_j + b_i \geqslant 0, \; (i = 1, 2, \ldots, m). \qquad \text{(V.10)}$$

Прямая задача параметрического программирования формулируется следующим образом.

Требуется разбить сегмент $[\lambda_1, \lambda_2]$ на конечное число подынтервалов таким образом, чтобы для всех значений параметра t из каждого подынтервала максимум значения целевой функции достигался в одной и той же вершине многогранника ограничений.

В реальных задачах величины P_j часто являются коэффициентами целевой функции, а q_j – долями погрешности их определения. Параметр t в этом случае изменяется от нуля в сторону увеличения или уменьшения при анализе влияния снижения точности оценок P_j. Например, величина прибыли от реализации 1 т угля, добываемого на j-й шахте, зависит от зольности, которая на стадии планирования нагрузок на шахты может быть рассчитана только с определенной долей погрешности. Поэтому составляя план производства группе шахт, целесообразно проверить, насколько он устойчив к ошибкам в прогнозе качества добываемого угля.

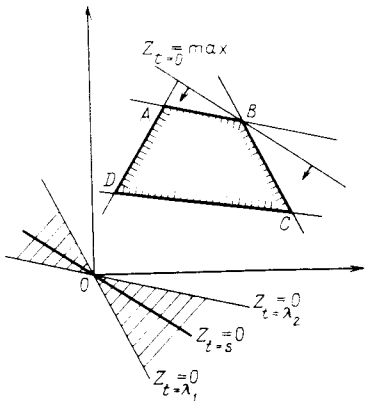


Рис. V.1. Геометрический смысл анализа чувствительности

В других задачах величины P_j, q_i и t могут иметь самостоятельный экономический смысл. Например, в задаче планирования поставок с фиксированными ценами продукта и расстояниями перевозок до потребителя, но неустойчивыми затратами на перевозки из-за использования разнородного транспорта P_j – цена единицы продукта; q_i – расстояние до j-го поставщика; t – переменный тариф за перевозку единицы продукта на 1 км. В этом случае величина S – средний уровень тарифа.

Устойчивость решения определяется в предположении возможных изменений тарифов из-за перераспределения долей различных видов транспорта.

Рассмотрим геометрический смысл задачи параметрического программирования. Как известно, оптимальное решение задачи линейного программирования соответствует вершине многогранника ограничений, удаление которой от плоскости целевой функции, приравненной к нулю, максимально.

Меняя t, мы изменяем наклон плоскости, соответствующей целевой функции. При этом оптимальное решение остается без изменений до тех пор, пока Z не станет параллельна какому-либо ребру (в этом случае – бесчисленное множество решений).

На рис. V.1 оптимальному решению задачи линейного программирования соответствует вершина B (предположим $t = s$). Эта же вершина соответствует оптимуму до тех пор, пока Z не становится параллельной ребрам AB или BC (при этом $t = [\lambda_1; \lambda_2]$). Следовательно, при повороте целевой функции в определенных пределах (на рисунке эта область заштрихована) решение задачи не меняется.

В двойственной задаче параметрического программирования рассматривается случай переменных запасов ресурсов. Модель двойственной задачи параметрического программирования имеет вид

$$Z = \sum_{j=1}^{n} c_j x_j \to \max \qquad (V.11)$$

при ограничениях:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \leqslant b_i + t_{q_i}, \ (i = 1, 2, \ldots, m), \lambda_1 \leqslant t \leqslant \lambda_2. \qquad (V.12)$$

То есть требуется разбить сегмент $[\lambda_1; \lambda_2]$ на конечное число подынтервалов таким образом, чтобы для всех значений параметра t из каждого подынтервала максимальное значение целевой функции достигалось в вершинах, определяемых одной и той же подсистемой системы ограничений, или достигалось бы в вершинах, отличающихся лишь параллельным сдвигом определяющих их плоскостей. Решение двойственной задачи позволяет указать пределы изменения ресурсов, не влияющих на положение оптимума.

Идея алгоритма решения прямой задачи параметрического программирования основана на отыскании вершины многогранника ограничений, соответствующей началу интервала t. Затем с помощью шагов модифицированных жордановых исключений переходим в соседние вершины многогранника, при которых решение остается еще без изменений.

Решение параметрической задачи в данной постановке можно разбить на три этапа. На первом предварительном этапе находится оптимальное решение задачи при фиксированном значении $t = s$. Далее определяется отрезок $[S; \lambda]$, на котором сохраняется оптимальное решение. На последнем третьем этапе определяется новое оптимальное решение, к которому следует перейти после выхода t за установленную границу отрезка.

Первый предварительный этап решения параметрической задачи заключается в поиске оптимального решения X_s при фиксированном значении $t = s$. Для этого исходная задача приводится к виду (V.9) и (V.10), записывается в специальную табл. V.4.

По сравнению с обычной линейной задачей в матрице добавлены две строки p_j и q_j. Решаем эту задачу до получения оптимального решения (табл. V.5).

Анализ устойчивости оптимального решения осуществляется следующим образом.

По условию оптимальности решения

$$\beta_i \geqslant 0, \ i = 1, 2, \ldots, m, \qquad (V.13)$$

$$(p'_j + s q'_j) \geqslant 0, \ j = 1, 2, \ldots, n.$$

Решение может измениться только тогда, когда это условие будет нарушено при изменении t от значения $t = s$ до $t_1 = \lambda_1 < s$ или $t_2 =$

Таблица V.4

	$-x_1$	...	$-x_n$	1
$y_1 =$	a_{11}	...	a_{1n}	b_1
...	...	...	...	...
$y_m =$	a_{m1}	...	a_{mn}	b_m
$Z_s =$	$(p_1 + sq_1)$	...	$(p_n + sq_n)$	Q
$p =$	p_1	...	p_n	–
$q =$	q_1	...	q_n	–

Таблица V.5

	$-x_1$	...	$-x_k$	$-y_{k+1}$	...	$-y_m$	1
$y_1 =$	α_{11}	...	α_{1k}	$\alpha_{1,k+1}$	...	α_{1m}	β_1
...	...	...	...	...	...	...	...
$y_k =$	α_{k1}	...	α_{kk}	$\alpha_{k,k+1}$	...	α_{kn}	β_k
$x_{k+1} =$	$\alpha_{k+1,1}$	...	$\alpha_{k+1,k}$	$\alpha_{k+1,k+1}$	...	$\alpha_{k+1,n}$	β_{k+1}
...	...	...	...	...	...	...	...
$x_m =$	α_{m1}	...	α_{mk}	$\alpha_{m,k+1}$	...	α_{mn}	β_m
$Z_s =$	$(p'_1 + sq'_1)$	...	$(p'_k + sq'_k)$	$(p'_{k+1} + sq'_{k+1})$	...	$(p'_n + sq'_n)$	Q'
$p =$	p'_1	...	p'_k	p'_{k+1}	...	p'_n	–
$q =$	q'_1	...	q'_k	q'_{k+1}	...	q'_n	–

$= \lambda_2 > s$. При этом хотя бы один член $(p'_j + tq'_j)$ должен стать отрицательным.

Множество всех решений параметра t, для которых сохранится данное решение x_s, задается множеством решений системы

$$p'_j + tq_j \geqslant 0,\ j = 1, 2, \ldots, n,\ \lambda_1 \leqslant t \leqslant \lambda_2, \qquad \text{(V.14)}$$

составленной так, что $s > \lambda_1$ и $s < \lambda_2$.

Отсюда задача анализа устойчивости сводится к отысканию верхней λ_2 и нижней λ_1 границ интервала $[\lambda_1, \lambda_2]$, включающего значение s.

Определение верхней границы. Пусть точка $t = s \in [\lambda_1, \lambda_2]$ и параметр t увеличивается на положительную величину $\Delta t > 0$. Текущие значения параметра определяются выражением $(s + \Delta t)$. После его

подстановки в (V.14)

$$p'_j + (s + \Delta t)q'_j \geqslant 0,\ j = 1, 2, \ldots, n, \qquad \text{(V.15)}$$

или, что равносильно,

$$(p'_j + sq'_j) + \Delta t q'_j \geqslant 0,\ j = 1, 2, \ldots, n. \qquad \text{(V.16)}$$

Так как по условию оптимальности

$$(p'_j + sq'_j) \geqslant 0,\ j = 1, 2, \ldots, n$$

и начальному условию $\Delta t > 0$, то очевидно, что решение системы (V.16) определяется знаками элементов q'_j.

Если все $q'_j \geqslant 0, j = 1, 2, \ldots, n$, то верхней границы для t не существует, $\lambda_2 = +\infty$.

Если же имеется хотя бы одно $q'_j < 0$, то

$$\lambda_2 = \min - \frac{p'_j}{q'_j} \qquad \text{(V.17)}$$

для всех $q'_j < 0$.

Определение нижней границы. Пусть также точка $t = s \in [\lambda_1, \lambda_2]$, но параметр уменьшается на положительную величину $\Delta t > 0$. Тогда текущее значение параметра определяется выражением $(s - \Delta t)$, подстановка которого в систему (V.18) приведет к системе

$$p'_j + (s - \Delta t)q'_j \geqslant 0,\ j = 1, 2, \ldots, n \qquad \text{(V.18)}$$

или

$$(p'_j + sq'_j) - \Delta t q'_j \geqslant 0,\ j = 1, 2, \ldots, n. \qquad \text{(V.19)}$$

Учитывая условие оптимальности $(p_j + sq'_j) \geqslant 0, j = 1, 2, \ldots, n$ и условие $\Delta t > 0$, из анализа системы (V.19) можно установить следующее:

если все $q'_j \leqslant 0, j = 1, 2, \ldots, n$, то нижней границы для t не существует, $\lambda_1 = -\infty$;

если же хотя бы одно $q'_j > 0$, то

$$\lambda_1 = \max - \frac{p'_j}{q'_j} \qquad \text{(V.20)}$$

для всех $q'_j > 0$.

Из условия определения границ λ_1 и λ_2 для t видно, что дальнейшее решение имеет два способа. В первом, когда граница отсутствует, следует вывод о неизменности решения x_s независимо от увеличения t при $\lambda_2 = +\infty$ или уменьшения t при $\lambda_1 = -\infty$.

В случае существования границы λ_1 или λ_2 дальнейший анализ должен определить возможные изменения решения x_s.

Пусть решается задача анализа устойчивости x_s при увеличении $t > s$ и существует верхняя граница t, равная λ_2.

В результате подстановки λ_2 в целевую функцию в табл. V.5, содержащей оптимальное решение x_s, изменятся элементы Z-строки и один из них обратится в ноль. Этот элемент должен соответствовать столбцу с $\min \frac{p'_j}{q'_j}$ при $q'_j < 0$. Пусть это будет k-й столбец, а в табл. V.6 пересчитана Z-строка.

Таблица V.6

	$-x_1$	...	$-x_k$	$-y_{k+1}$	...	y_n	1
$y_1 =$	α_{11}	...	α_{1k}	$\alpha_{1\,k+1}$	...	α_{1n}	β_1
...	...	...	...	...	...	...	...
$y_k =$	α_{k1}	...	α_{kk}	$\alpha_{k,k+1}$	...	$\alpha_{k,n}$	β_k
$x_{k+1} =$	$\alpha_{k+1,1}$	...	$\alpha_{k+1,k}$	$\alpha_{k+1,k+1}$	...	$\alpha_{k+1,n}$	β_{k+1}
...	...	...	...	...	...	...	...
$x_m =$	α_m	...	α_{mk}	$\alpha_{m,k+1}$	...	α_{mn}	β_m
$Z_\lambda =$	$(p'_1 + \lambda_2 q'_1)$	...	0	$(p'_{k+1} + \lambda_2 q'_{k+1})$	...	$(p'_n + \lambda_2 q'_n)$	Q'
$p =$	p'_1	...	p'_k	p'_{k+1}	...	p'_n	–
$q =$	q'_1	...	q'_k	q'_{k+1}	...	q'_n	–

При дальнейшем увеличении от $t = \lambda_2$ к $t = \lambda_2 + \epsilon$ в k-м столбце $p'_k + (\lambda_2 + \epsilon) q'_k < 0$, т.е. решение x_s уже не будет оптимальным. Здесь возможны два случая.

В k-м столбце имеются положительные элементы α_{ik} и тогда существует новое оптимальное решение x_λ.

Это решение можно получить, действуя по правилам симплекс-алгоритма, приняв за разрешающий k-й столбец и выбрав разрешающую строку по минимальному положительному отношению элементов столбца свободных членов к k-му столбцу, т.е. выбирается l-я строка, для которой

$$\frac{\beta_l}{\alpha_{lk}} = \min \frac{\beta_i}{\alpha_{ik}} > 0,\ i = 1, 2, \ldots, m. \qquad (V.21)$$

Если в k-м столбце все элементы $\alpha_{ik} \leqslant 0$, $i = 1, 2, \ldots, m$, то при $t > \lambda_2$ целевая функция становится неограниченной и бесконечно возрастает за счет независимой переменной, соответствующей k-му столбцу.

Задача анализа устойчивости решения x_s при уменьшении $t < s$ решается аналогично, но определяется нижняя граница изменения параметра λ_1.

Пример V.2. Исходные условия заданы в табл. V.7 и $t \geqslant 5$.

Таблица V.7

	$-x_1$	$-x_2$	$-x_3$	$-x_4$	1
$y_1 =$	0,70	1,30	–0,30	0	2,00
$y_2 =$	0.40	1,00	–0,20	0,20	1.40
$y_3 =$	6.00	11.0	–2,00	[1,00]	2.00
$p =$	1.40	6.40	5.70	0.40	3.80
$q =$	0	–3.50	0,35	–0,50	

На предварительном шаге решается задача линейного программирования с $t = 5$ (табл. V.8, V.9).

Таблица V.8

	$-x_1$	$-x_2$	$-x_3$	$-x_4$	1
$y_1 =$	0.70	1.30	–0,30	0	2,00
$y_2 =$	0.40	1.00	–0.20	0,20	1.40
$y_3 =$	6,00	11.0	–2.00	[1.00]	2.00
$Z =$	1.40	–11,1	7.45	–2,10	3,80
$p =$	1,40	6,40	5,70	0,40	3,80
$q =$	0	–3,50	0.35	–0.50	

Таблица V.9

	$-x_1$	$-x_2$	$-x_3$	$-y_3$	1
$y_1 =$	0.70	1,30	–0.30	0	2
$y_2 =$	–0,80	–1.20	0,20	–0,20	1
$x_3 =$	6.00	11,0	–2,00	1,00	2
$Z =$	14.0	12.0	3,25	2,10	8
$p =$	–1.00	2,00	6.50	–0,40	8
$q =$	3,00	2,00	–0.65	0,50	

Содержащееся в табл. V.9 оптимальное решение $x_1 = x_2 = x_3 = 0$; $x_4 = 2$; $Z_{max} = 8$ будет сохраняться, пока все элементы Z-строки имеют неотрицательность при увеличении t. Верхняя граница изменения параметра t определится из анализа отношений $-(p_j/q_i)$, вычисленных для $q < C$. В данном примере $q_3 = -0{,}65 < 0$, а $-(p_3/q_3) = \frac{-6{,}50}{-0{,}65} = 10$. Это значение является верхней границей λ изменения параметра t, до которой сохраняется полученное решение.

Для нового решения следует принять $t = \lambda = 10$, пересчитать элементы Z-строки и сделать шаг жордановых исключений с разрешающим элементом, выбранным из столбца с обратившимся в ноль коэффициентом в Z-строке (табл. V.10, V.11).

Таблица V.10

	$-x_1$	$-x_2$	$-x_3$	$-y_3$	1
$y_1 =$	0,70	1,30	−0,30	0	2
$y_2 =$	−0,80	−1,20	**0,20**	−0,20	1
$x_4 =$	6,00	11,0	−2,00	1,00	2
$Z =$	29,0	22,0	0	4,60	8
$p =$	−1,00	2,00	6,50	−0,40	8
$q =$	3,00	2,00	−0,65	0,50	

Таблица V.11

	$-x_1$	$-x_2$	$-y_2$	$-x_4$	1
$y_1 =$	−0,50	−0,50	1,50	−0,30	3,5
$x_3 =$	−4,00	−6,00	5,00	−1,00	5,0
$y_3 =$	−2,00	−1,00	10,0	−1,00	12,0
$Z =$	29,0	22,0	0,00	4,60	8,0
$p =$	25,0	41,0	−32,5	6,10	8,0
$q =$	0,40	−1,90	3,25	−0,15	

Вывод: при $5 \leqslant t \leqslant 10$ задача имеет решение: $x_1 = x_2 = x_3 = 0$; $x_4 = 2$; $Z_{max} = 8$, а при $t \geqslant 10$ $x_1 = x_2 = x_4 = 0$; $x_3 = 5$; $Z_{max} = 8$.

§ 4. ДВОЙСТВЕННЫЙ СИМПЛЕКС-МЕТОД

Как ранее указывалось, когда с матрицей исходной задачи при помощи модифицированных жордановых исключений решается прямая задача, двойственная – решается с помощью обыкновенных жордановых исключений. На этой основе построен двойственный симплекс-метод

решения задач линейного программирования. Его обычно целесообразно использовать, если количество ограничений больше числа переменных или строка целевой функции в исходной матрице положительна.

Рассмотрим задачу минимизации линейной формы

$$W = b_1 y_1 + b_2 y_2 + \ldots + b_j y_j + \ldots + b_m y_m \to \min$$

при ограничениях:

$$v_i = a_{11} y_1 + a_{12} y_2 + \ldots + a_{ij} y_j + \ldots + a_{1n} y_n \geqslant c_i .$$

Как известно, к этой форме могут быть сведены любые другие.

Алгоритм двойственного симплекс-метода можно представить в следующем виде.

1. Запишем задачу в форме табл. V.12.

Таблица V.12

	y_1	y_2	...	y_m	1
$v_1 =$	a_{11}	a_{21}	...	a_{m1}	$-c_1$
$v_2 =$	a_{12}	a_{22}	...	a_{m2}	$-c_2$
.	.	.	...	.	.
.	.	.	...	.	.
.	.	.	...	.	.
$v_n =$	a_{1n}	a_{2n}	...	a_{mn}	$-c_n$
$W =$	b_1	b_2	...	b_m	0

2. Рассмотрим W-строку:

если все коэффициенты W-строки неотрицательны, то 0 является оценкой W снизу и переходим к отысканию оптимального решения;

если среди коэффициентов W-строки есть отрицательные, то выбираем любой из них. Если все коэффициенты выбранного столбца неотрицательны, то либо W неограничена снизу, либо задача противоречива;

находим в выбранном столбце какой-нибудь из отрицательных коэффициентов, строку, содержащую этот коэффициент, берем в качестве разрешающей;

находим неотрицательные отношения коэффициентов W-строки к коэффициентам разрешающей строки. Разрешающим берем тот из коэффициентов разрешающей строки, для которого это отношение наименьшее;

если наименьшее отношение равно нулю, то его знаменатель берем разрешающим элементом лишь когда он положителен;

выбрав разрешающий элемент, делаем шаг обыкновенного жорданова исключения.

3. Рассматриваем столбец свободных членов. Если все коэффициенты $b_1, \ldots, b_m$ неотрицательны, а среди свободных членов есть отрицательные, то решение не является оптимальным. В этом случае:

в качестве разрешающей берем строку, содержащую отрицательный свободный член;

отбираем все положительные коэффициенты этой строки, делим на них соответствующие коэффициенты W-строки и в качестве разрешающего берем тот из коэффициентов разрешающей строки, для которого это отношение имеет наименьшее значение;

если в разрешающей строке нет положительных коэффициентов, то ограничения задачи противоречивы.

После конечного числа шагов обыкновенных жордановых исключений мы либо найдем минимальное значение функции W, либо убедимся в противоречивости ограничений задачи.

Пример V.3. Минимизировать функцию

$$W = 4y_1 - y_2 \to \min$$

при ограничениях:

$$v_1 = y_2 - 2 \geqslant 0,$$

$$v_2 = -y_1 + y_2 + 1 \geqslant 0,$$

$$v_3 = y_1 - y_2 + 1 \geqslant 0,$$

$$v_4 = y_1 - \frac{3}{2} \geqslant 0.$$

Ход решения показан в табл. V.13–V.15.

Таблица V.13

	y_1	y_2	1
$v_1 =$	0	1	−2
$v_2 =$	−1	1	1
$v_3 =$	1	−1	1
$v_4 =$	1	0	−3/2
$W =$	4	−1	0

Таблица V.14

	y_1	v_3	1
$v_1 =$	1	−1	−1
$v_2 =$	0	−1	2
$y_2 =$	1	−1	1
$v_4 =$	1	0	3/2
$W =$	3	1	−1

Таблица V.15

	v_4	v_3	1
$v_1 =$			1/2
$v_2 =$			2
$y_2 =$			5/2
$y_1 =$			3/2
$W =$	3	1	7/2

$\min W = \frac{7}{2}$ и достигается при $y_1 = \frac{3}{2}$ и $y_2 = \frac{5}{2}$.

Контрольные вопросы и задачи

1. В чем отличительные особенности двойственных задач?
2. Каков экономический смысл первой и второй теорем двойственности?
3. Дайте экономическую интерпретацию задачи – двойственной задачи оптимальной расстановки оборудования. Каков экономический смысл двойственных оценок?
4. То же, транспортной задачи?
5. То же, задачам размещения?
6. То же, задачам планирования добычных работ в режиме усреднения качества?
7. В каких случаях эффективно используется двойственный симплекс-метод?
8. Постройте двойственные модели к задачам 11–12 из гл. III.
9. Решите задачи 11–12 из гл. III двойственным симплекс-методом.
10. В задаче 8 из гл. IV на основе построения и решения двойственной задачи определите, какое оборудование целесообразно приобретать в первую очередь.
11. В задаче 9 из гл. IV определите, в каких пределах может изменяться содержание металла в руде по участкам, чтобы оптимальное распределение машин по ним, обеспечивающее максимальный объем металла в руде, не менялось.
12. В задаче 12 из гл. IV определите, какое оборудование и за какую цену целесообразно дополнительно приобретать для увеличения выпуска продукции.
13. В задаче 16 из гл. IV определите, какое сырье и за какую цену целесообразно приобретать для увеличения выпуска продукции и получаемой прибыли.
14. В задаче 18 из гл. IV определите, какие ресурсы и за какую цену целесообразно приобретать для увеличения дохода предприятия.

Глава VI

ТРАНСПОРТНАЯ ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

§ 1. ОСОБЕННОСТИ ТРАНСПОРТНОЙ ЗАДАЧИ

Симплексный метод – универсальный для решения любых задач линейного программирования. Однако отдельные типы задач линейного программирования (например, о назначении и транспортные) в силу специфичности своей структуры могут решаться более простыми методами. К транспортным относится большое количество задач планирования перевозок грузов. Кроме того, ряд других задач (например, задача о размещении промышленности) может быть сведен к транспортным.

Напомним классическую формулировку и модель транспортной задачи. Имеется n поставщиков однородного груза, i-й поставщик имеет

запас груза a_i $(i = 1, 2, ..., n)$ и m потребителей, j-й потребитель имеет потребность в грузе b_j $(j = 1, 2, ..., m)$. Затраты на перевозку груза от i-го поставщика к j-му потребителю составляют c_{ij}. Требуется определить объемы грузоперевозок от i-го поставщика к j-му потребителю x_{ij}, при которых достигается минимальная суммарная стоимость перевозок

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

При этом весь запас груза должен быть вывезен, т.е.

$$x_{i1} + x_{i2} + x_{i3} + \ldots + x_{ij} + \ldots + x_{in} = a_i, \ (i = \overline{1, n}), \qquad \text{(VI.1)}$$

а потребности всех потребителей удовлетворены, т.е.

$$x_{1j} + x_{2j} + \ldots + x_{ij} + \ldots + x_{mj} = b_j, \ (j = \overline{1, m}). \qquad \text{(VI.2)}$$

Рассматривается закрытая модель, при которой $\sum_{j=1}^{m} b_j = \sum_{i=1}^{n} a_i$, т.е. суммарные запасы груза равны потребности в нем. Любая задача может быть сведена к закрытой путем введения фиктивного поставщика (если $\sum_{i=1}^{n} a_i \leqslant \sum_{j=1}^{m} b_j$) или фиктивного потребителя (если $\sum_{i=1}^{n} a_i > \sum_{j=1}^{m} b_j$).

Транспортная задача имеет следующие особенности:

все ограничения имеют вид равенств;

каждая переменная входит всего в два ограничения;

коэффициенты при переменных в ограничениях равны единице.

Благодаря этим особенностям транспортная задача может быть решена более простым способом, чем симплексный, причем симплексная матрица для транспортной задачи состоит из нулей и единиц (табл. VI.1).

Таблица VI.1

	$-x_{11}$ ...	x_{1j} ...	$-x_{1m}$ ...	$-x_{i1}$ ...	$-x_{ij}$ ...	x_{im} ...	$-x_{n1}$ ...	$-x_{nj}$ ...	x_{nm} ...	1
u_1 y_1	−1 ...	−1 ...	−1 ...	0 ...	0 ...	0 ...	0 ...	0 ...	0 ...	c_1
u_i y_i	0 ...	0 ...	0 ...	−1 ...	−1 ...	−1 ...	0 ...	0 ...	0 ...	a_i
u_n y_n	0 ...	0 ...	0 ...	0 ...	0 ...	0 ...	−1 ...	−1 ...	−1 ...	a_n
v_1 y_1	+1 ...	0 ...	0 ...	+1 ...	0 ...	0 ...	+1 ...	0 ...	0 ...	$-b_1$
v_j y_j	0 ...	+1 ...	0 ...	0 ...	+1 ...	0 ...	0 ...	+1 ...	0	$-b_j$
v_m y_m	0 ...	0 ...	+1 ...	0 ...	0 ...	+1 ...	0 ...	0 ...	+1	$-b_m$
	c_{11} ...	c_{1j} ...	c_{1m} ...	c_{i1} ...	c_{ij} ...	c_{im} ...	c_{n1} ...	c_{nj} ...	c_{nm}	

Таблица VI.2

Поставщик	Потребитель					Запасы
	B_1	...	B_j	...	B_m	
A_1	c_{11} x_{11}	...	c_{1j} x_{1j}	...	c_{1m} x_{1m}	a_1
...	...	...	...	...	...	...
A_i	c_{i1} x_{i1}	...	c_{ij} x_{ij}	...	c_{im} x_{im}	a_i
...	...	...	...	...	...	...
A_n	c_{n1} x_{n1}	...	c_{nj} x_{nj}	...	c_{nm} x_{nm}	a_n
Потребность	b_1	...	b_j	...	b_m	$\sum_{j=1}^{m} b_j = \sum_{i=1}^{n} a_i$

Для решения транспортной задачи составляют специальную матрицу (табл. VI.2).

Транспортная матрица по размеру $(n + 1)(m + 1)$ существенно меньше симплексной матрицы $(nm + 1)(n + m + 1)$ и совмещает матрицу стоимостей $\|c_{ij}\|$ и объемов перевозок $\|x_{ij}\|$, причем в этой матрице построчно записаны ограничения по запасам сырья у поставщиков (VI.1), а по столбцам – по потребности каждого потребителя (VI.2).

В отличие от задач линейного программирования транспортная задача всегда имеет оптимальное решение. Для доказательства этого утверждения покажем, что система ограничений совместна (т.е. имеется опорное решение) и что целевая функция ограничена.

Обозначим суммарный объем перевозок – M.

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

Предположим, что $x_{ij} = \dfrac{a_i b_j}{M}$.

Подставляя это значение x_{ij} в ограничения (VI.1) и (VI.2)

$$\sum_{j=1}^{m} x_{ij} = \sum_{j=1}^{m} \frac{a_i b_j}{M} = \frac{a_i}{M} \sum_{j=1}^{m} b_j = a_i;$$

$$\sum_{i=1}^{n} x_{ij} = \sum_{i=1}^{n} \frac{a_i b_i}{M} = \frac{b_j}{M} \sum_{i=1}^{n} a_i = b_j$$

получим, что все они выполняются, и, следовательно, $x_{ij} = \frac{a_i b_j}{M}$ дает опорное решение задачи.

Для доказательства ограниченности целевой функции выберем максимальную и минимальную стоимости перевозок и обозначим $\max\ c_{ij} = c_1$, а $\min\ c_{ji} = c_2$. Тогда значение целевой функции

$$W = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}\, x_{ij} \leqslant c_1 \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij},$$

$$W = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij}\, x_{ij} \geqslant c_2 \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij}.$$

Так как $\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} = M$, то $c_1 M \geqslant W \geqslant c_2 M$, т.е. общее решение задачи ограничено сверху и снизу.

Общее количество базисных переменных в транспортной задаче равно $(m + n - 1)$, так как среди $(m + n)$ ограничений-равенств одно является линейно-зависимым [так как суммарные потребности и запасы равны, то, складывая равенства систем (VI.1) и (VI.2), можно получить одно и то же уравнение].

Решение транспортной задачи линейного программирования включает два основных этапа: построение опорного решения (начального плана перевозок) и построение оптимального решения.

Разберем последовательно методы получения опорного и оптимального решений.

§ 2. ПОСТРОЕНИЕ ОПОРНОГО РЕШЕНИЯ

Опорное решение транспортной задачи может быть получено методами северо-западного угла, наименьших стоимостей и двойного предпочтения.

Наиболее простым и легко формализуемым является метод северо-западного угла, однако он дает обычно решение, далекое от оптимального. При построении опорного решения методами наименьших стоимостей и двойного предпочтения анализируется матрица затрат и начальный план обычно близок к оптимальному.

Метод северо-западного угла используется, как правило, при расчетах на ЭВМ. При его применении данные о стоимостях не нужны. Рассматривается клетка 11 (северо-западный угол) и в нее на основании сравнения величин a_1 и b_1 проставляется максимальный объем перевозок (минимальная из величин a_1 и b_1). Если $a_1 > b_1$, то первый столбец из рассмотрения исключается, так как потребности первого потребителя удовлетворены ($x_{11} = b_1$), и определяется разность $a_1 - b_1$ (остаток запасов у первого поставщика). Далее рассматривается клетка 12 (новый северо-западный угол) и сравниваются величины

b_2 и $(a_1 - b_1)$. В клетку 12 проставляется минимальная из этих величин. Если $(a_1 - b_1) \leqslant b_2$, то далее рассматривается клетка 22 (новый северо-западный угол) и сравниваются величины a_2 и $[b_2 - (a_1 - b_1)]$ (оставшаяся потребность у второго потребителя). Подобная процедура повторяется вплоть до заполнения всей матрицы.

Пример VI.1. Пусть условия транспортной задачи заданы матрицей (табл. VI.3).

Таблица VI.3

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	4	6	5	7	50
A_2	6	4	8	9	90
A_3	5	7	4	6	250
Потребность	30	200	70	130	

Перед построением опорного решения проверим баланс запасов и потребления, т.е. является ли модель закрытой.

Суммарные запасы ($\sum_{i=1}^{3} a_i$) составляют 50 + 90 + 250 = 390 единиц груза, а суммарная потребность ($\sum_{j=1}^{4} b_j$) будет 30 + 200 + 70 + 130 = 430 единиц груза.

Так как суммарная потребность превышает объем запасов, вводим четвертого фиктивного поставщика таким образом, чтобы модель стала закрытой, т.е.

$$a_4 = \sum_{j=1}^{4} b_j - \sum_{i=1}^{3} a_i = 430 - 390 = 40.$$

В результате получим новую матрицу для закрытой модели (табл. VI.4), причем стоимости перевозок от фиктивного поставщика равны 0.

Таблица VI.4

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	4	6	5	7	50
A_2	6	4	8	9	90
A_3	5	7	4	6	250
A_4	0	0	0	0	40
Потребность	30	200	70	130	

Таблица VI.5

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	30 4	20 6	5	7	50
A_2	6	90 4	8	9	90
A_3	5	90 7	70 4	90 6	250
A_4	0	0	0	40 0	40
Потребности	30	200	70	130	

Приступим теперь к построению опорного решения.

1. Проставляем в клетку 11 минимальный из объемов a_1 и b_1

$$\min\left\{a_1\ b_1\right\} = \min\left\{50, 30\right\} = 30.$$

2. Находим разность $a_1 - b_1 = 50 - 30 = 20$.

3. Рассматриваем новый северо-западный угол – клетку 12 и сопоставляем b_2 и $(a_1 - b_1)$.

Минимум $a_1 - b_1 = 20$ проставляем в клетку 12.

4. Находим разность $b_2 - (a_1 - b_1)$, она равна $200 - 20 = 180$.

Минимальную из этих величин (a_2) проставляем в клетку 22.

5. Находим разность $[b_2 - (a_1 - b_1 - a_2)] = 200 - 20 - 90 = 90$.

6. Рассматриваем клетку 32 и так далее до заполнения всей матрицы.

Опорное решение показано в табл. VI.5.

Суммарная стоимость перевозок при этом решении составит $W = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}x_{ij} = 30 \cdot 4 + 20 \cdot 6 + 90 \cdot 4 + 90 \cdot 7 + 70 \cdot 4 + 90 \cdot 6 + 40 \cdot 0 = 2050.$

При использовании метода наименьших стоимостей последовательно заполняются клетки, содержащие наименьшие стоимостные показатели. При этом в эти клетки проставляется максимально допустимый объем перевозок, для чего сопоставляются объемы запасов и потребности. Процедура повторяется до получения допустимого решения.

Пример VI.2. Составить опорное решение для прежней задачи, условия которой приведены в матрице (см. табл. VI.4).

Минимальные стоимости (4) имеют перевозки, содержащиеся в клетках 11, 22, 33. Рассмотрим последовательно эти клетки и проставим в них возможные объемы перевозок.

Сопоставляя a_1 и b_1, проставляем в клетку 11 (табл. VI.6) минимальный объем 30; сопоставляя a_2 и b_2 проставляем в клетку 22 объем 90; a_3 и b_3 – в клетку 33 ставим объем 70.

Столбец B_1 и B_3, а также строку A_2 далее не рассматриваем, так как ограничения по ним выполнены.

Таблица VI.6

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	30 4	20 6	5	7	50
A_2	6	90 4	8	9	90
A_3	5	50 7	70 4	130 6	250
A_4	0	40 0	0	0	40
Потребность	30	200	70	130	

При этом условии с наименьшей стоимостью перевозок – клетки 34 и 12. Сопоставляя $B_4 = 130$ – остаток запасов третьего поставщика и $a_3 - b_3 = 250 - 70 = 180$, проставляем в клетку 34 объем 130. Столбец B_4 далее не рассматривается. Аналогично, сопоставляя $a_1 - b_1 = 50 - 30 = 20$ и $b_2 - a_2 = 200 - 90 = 110$, т.е. остаток запасов у первого поставщика и неудовлетворенную потребность второго потребителя, проставляем в клетку 12 минимальную из этих величин (20) и первую строку в дальнейшем не рассматриваем.

Следующая клетка с наименьшей стоимостью 32. Проставляем в нее минимальный из объемов $[b_2 - a_2 - (a_1 - b_1)] = 200 - 90 - (50 - 30) = 90$ и $a_3 - b_3 - b_4 = 250 - 130 - 70 = 50$. Последнюю клетку 42 заполняем из условия соблюдения ограничений по четвертой строке и второму столбцу.

Затраты на перевозку при плане, полученном методом наименьших стоимостей, составляют $(30 \cdot 4 + 20 \cdot 6 + 90 \cdot 4 + 50 \cdot 7 + 70 \cdot 4 + 130 \cdot 6 + 40 \cdot 0) = 2010$, т.е. меньше, чем при плане, составленном по методу северо-западного угла.

При построении опорного решения транспортной задачи методом двойного предпочтения поочередно просматриваются все строки и столбцы матрицы. Клетки, содержащие минимальные удельные затраты на перевозку в каждой строке и столбце, помечаются каким-нибудь знаком. Первоначально объемы перевозок проставляются в клетках, отмеченных дважды, т.е. предпочтительных как в строке, так и в столбце (отсюда и название метода). После того, как клетки, помеченные дважды, заполнены полностью, объемы проставляются в клетки, отмеченные один раз. Затем, с учетом ограничений по a_i и b_j, заполняются остальные клетки.

Пример VI.3. Составить опорное решение для рассматриваемой ранее транспортной задачи (см. табл. VI.4) методом двойного предпочтения.

Просматриваем строки и отмечаем клетки, в которых затраты минимальны (табл. VI.7). Это клетки 11, 22, 33. Затем просматриваем столбцы и также отмечаем клетки с минимальными затратами (11, 22, 33, 34). Таким образом, клетки 11, 22, 33 отмечены дважды – проставляем в них возможные объемы перевозок (соответственно 30, 90, 70). Далее рассматриваем клетку 34 и отмечаем в ней перер зку.

Таблица VI.7

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	30 4	20 6	5	7	50
A_2	6	90 4	8	9	90
A_3	5	50 7	70 4	130 6	250
A_4	0	40 0	0	0	40
Потребитель	30	200	70	130	

Больше помеченных клеток нет и дальнейшее заполнение матрицы можно производить по методу наименьших стоимостей. Опорное решение показано в табл. VI.7. Как видно из составления табл. VI.6 и VI.7, в рассматриваемом примере опорные решения, полученные методом наименьших стоимостей и двойного предпочтения, совпадают.

§ 3. УСЛОВИЯ И МЕТОД ПОСТРОЕНИЯ ОПТИМАЛЬНОГО РЕШЕНИЯ ТРАНСПОРТНОЙ ЗАДАЧИ

Один из наиболее простых и распространенных методов оптимизации транспортной задачи – метод потенциалов. Условия оптимальности плана транспортной задачи вытекают из теорем двойственности и заключаются в следующем.

Для того чтобы допустимый план транспортной задачи $X = \| x_{ij} \|_{m,n}$ был оптимальным, необходимо и достаточно, чтобы ему соответствовала система чисел $U_1, U_2, ..., U_n, V_1, V_2, ..., V_m$, удовлетворяющая условиям:

$$\text{I. } V_j - U_i \leqslant c_{ij}, \text{ если } x_{ij} = 0, \tag{VI.3}$$

$$\text{II. } V_j - U_i = c_{ij}, \text{ если } x_{ij} > 0; \; i = \overline{1, n}, \; j = \overline{1, m}. \tag{VI.4}$$

(VI.3) и (VI.4) называются условиями потенциальности оптимального плана транспортной задачи, и теорема может формулироваться следующим образом: для оптимального решения транспортной задачи необходимо и достаточно, чтобы оно было потенциально.

Условия потенциальности показывают, что разность потенциалов потребителя и поставщика равна стоимости перевозки между ними, если перевозка осуществляется (условие VI.3), и меньше стоимости перевозки при ее отсутствии (условие VI.4).

Алгоритм решения транспортной задачи методом потенциалов состоит из двух этапов: предварительного и общего.

Первый (предварительный) этап включает:

I – 1 – построение опорного решения;

I – 2 – присвоение и расчет системы потенциалов;

I – 3 – проверка первоначального плана на оптимальность.

Если опорное решение не является оптимальным, то переходят ко второму (общему) этапу, который включает:

II – 1 – улучшение плана перевозок;

II – 2 – исправление системы потенциалов;

II – 3 – проверка улучшенного плана на оптимальность.

Общий шаг циклически выполняется до получения оптимального плана перевозок.

Рассмотрим подробнее составные части метода потенциалов.

Построение опорного решения (I–1) осуществляется методами северо-западного угла, наименьших стоимостей или двойного предпочтения.

Расчет системы потенциалов (I–2) производится на основании второго условия оптимальности.

Всего необходимо присвоить $(m + n)$ потенциалов, а допустимый невырожденный план включает $(m + n - 1)$ перевозку.

Используя занятые клетки для расчета потенциалов, можно составить систему из $(m + n - 1)$ уравнений с $(m + n)$ переменными. Такая система имеет неограниченное множество решений. Поэтому для построения системы потенциалов надо задаться потенциалом одного поставщика или потребителя (строки или столбца). Обычно полагают $U_1 = 0$ либо присваивают нулевой потенциал строке, имеющей занятую перевозкой клетку с наибольшей стоимостью. Далее потенциалы всех остальных столбцов и строк вычисляют через занятые клетки, используя второе условие оптимальности.

Проверка плана на оптимальность (шаги I – 3 и II – 3) проводится для клеток, не занятых перевозками. Очевидно, что для клеток с перевозками второе условие оптимальности выполняется, так как оно было использовано для присвоения системы потенциалов. Для всех незанятых клеток проверяется выполнение первого условия оптимальности. Для этого вычисляют невязку $v_{ij} = V_j - U_i - c_{ij}$. Если $v_{ij} \leqslant 0$ для всех клеток, то план является оптимальным, так как первое условие оптимальности соблюдено. Если имеются положительные превышения, то первое условие оптимальности не соблюдено и план может быть улучшен. Величина v_{ij} показывает, какая экономия на перевозку единицы груза может быть достигнута при изменении плана перевозок.

Улучшение плана перевозок заключается в выборе (II–1) клетки с максимальным превышением и составлении замкнутого контура (цикла), вершинами которого являются клетка с нарушением и клетка с перевозками. Контур может иметь различную конфигурацию (рис. VI.1), но на нем всегда четное количество вершин. Начиная с клетки,

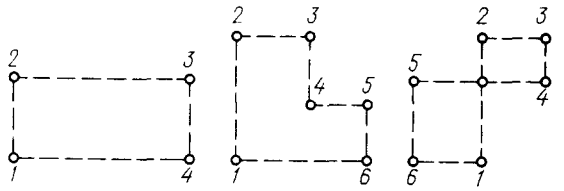


Рис. VI.1. Примеры циклов

имеющей нарушение, вершины контура нумеруются (по часовой стрелке или против). Нечетные вершины составляют положительную полуцепь, а четные – отрицательную. При улучшении плана в вершинах положительной полуцепи (нечетных) объемы перевозок увеличатся, а в вершинах отрицательной полуцепи (четных) уменьшатся. Далее просматривают объемы перевозок в четных клетках и находят минимальный из них $q = \min\ x_{ij}$. После этого перераспределяют объемы перевозок внутри контура, для чего к объемам перевозок в нечетных вершинах прибавляют объем q, а из объемов перевозок в четных вершинах вычитают эту величину. В результате этой процедуры получают улучшенный план, в котором общие затраты на перевозку меньше, чем в предыдущем плане, на величину $q v_{ij}$. Новый план является допустимым, так как в каждой строке и столбце замкнутого контура в одной клетке объем увеличился, а в другой уменьшился на одну и ту же величину.

Проверке улучшенного плана на оптимальность (шаг II–3) предшествует исправление системы потенциалов (шаг II–2). Для этого не обязательно строить новую систему потенциалов, можно исправить прежнюю.

Предположим, что второе условие оптимальности нарушено для клетки, в которой ранее не было перевозки, т.е. клетки с максимальным нарушением (первой вершины контура). Следовательно, необходимо исправить потенциал V_j для этой клетки на величину нарушения v_{ij}. Если в данном столбце нет больше клеток (занятых), то можно переходить к проверке первого условия оптимальности. Если же в столбце есть клетки с перевозками, то, исправив потенциал V_j столбца на v_{ij}, нарушим условия оптимальности на величину v_{ij} для остальных занятых клеток столбца. Чтобы исправить потенциалы этих клеток, достаточно соответствующую им величину U_i уменьшить на v_{ij}, т.е. новые потенциалы строк составят $U_i' = U_i - v_{ij}$. Затем просматриваем строки, где потенциалы изменились. Если в этих строках имеются занятые клетки (кроме клеток столбца V_j, у которого уже изменился потенциал), то потенциалы столбцов, соответствующих этим клеткам, исправляются на величину v_{ij}. Далее аналогично просматриваем новые столбцы и т.д. Процесс перерасчета потенциалов заканчивается, когда при просмотре строк или столбцов в них не окажется клеток с поставками.

Пример VI.4. Найти оптимальное решение задачи примера VI.1. Опорное решение, полученное методом северо-западного угла (см. табл. VI.6), повторено в табл. VI.8.

Таблица VI.8

	$V_1 = 5$	$V_2 = 7$	$V_3 = 4$	$V_4 = 6$	
Поставщик	**Потребители**				**Запасы**
	B_1	B_2	B_3	B_4	
$U_1 = 1$ A_1	30 4	20 6	5	7	50
$U_2 = 3$ A_2	6	90 4	8	9	90
$U_3 = 0$ A_3	5	2 90 7	70 4	90 6 3	250
$U_4 = 6$ A_4	0	1 0	0	40 0 4	40
Потребность	30	200	70	130	

Строим систему потенциалов. Третья строка имеет наибольшую стоимость в занятой клетке, поэтому ей присваивается потенциал, равный 0. Потенциалы обычно проставляются слева и вверху матрицы.

Зная потенциал третьей строки, вычисляем потенциал остальных столбцов и строк

$$V_2 = U_3 + c_{32} = 0 + 7 = 7,$$
$$V_3 = U_3 + c_{33} = 0 + 4 = 4,$$
$$V_4 = U_3 + c_{24} = 0 + 6 = 6,$$
$$U_1 = V_2 - c_{21} = 7 - 6 = 1,$$
$$U_2 = V_2 - c_{22} = 7 - 4 = 3,$$
$$U_4 = V_4 - c_{44} = 6 - 0 = 6,$$
$$V_1 = U_1 + c_{11} = 1 + 4 = 5.$$

Проверим опорное решение на оптимальность. Для этого определим отрицательны ли нарушения $v_{ij} = V_j - U_i - c_{ij}$ для незанятых клеток.

$$v_{13} = V_3 - U_1 - c_{13} = 4 - 1 - 5 = -2 < 0,$$
$$v_{14} = V_4 - U_1 - c_{14} = 6 - 1 - 7 = -2 < 0,$$
$$v_{21} = V_1 - U_2 - c_{21} = 5 - 3 - 6 = -4 < 0,$$
$$v_{23} = V_3 - U_2 - c_{23} = 4 - 3 - 8 = -7 < 0,$$

$$v_{24} = V_4 - U_2 - c_{24} = 6 - 3 - 9 = -6 < 0,$$

$$v_{31} = V_1 - U_3 - c_{31} = 5 - 0 - 5 = 0 = 0,$$

$$v_{41} = V_1 - U_4 - c_{41} = 5 - 6 - 0 = -1 < 0,$$

$$v_{42} = V_2 - U_4 - c_{42} = 7 - 6 - 0 = 1 > 0,$$

$$v_{43} = V_3 - U_4 - c_{43} = 4 - 6 - 0 = -2 < 0.$$

План не является оптимальным, так как имеется нарушение в клетке 42, ибо $v_{42} > 0$.

Переходим к общему этапу.

Улучшим первоначальный план. Для этого, начиная от клетки 42, построим замкнутый контур, вершинами которого (помимо 42) будут клетки с перевозками. Прямоугольный контур образует клетки 42, 32, 34, 44 (в табл. VI.8 показан пунктирной линией). Нумеруем по часовой стрелке вершины контура, начиная с клетки 42 (см. табл. VI.8). Четные клетки (32 и 44) образуют вершины *отрицательной полуцепи*, а нечетные (42 и 34) – *положительной*. Наименьший объем перевозок в четных клетках составляет $x_{44} = 40$ единиц груза (клетка 45). Отнимая 40 от объемов перевозок в клетках с четными номерами (32 и 44) и прибавляя к объемам перевозок в нечетных клетках, получим улучшенный план (табл. VI.9).

При этом суммарные затраты уменьшились на величину $v_{42}x_{42} = 40 \cdot 1$ и составили $2050 - 40 = 2010$.

Исправление системы потенциалов начинаем с потенциала второго столбца (V_2), так как $v_{42} = 1 > 0$,

$$V_2' = V_2 - v_{42} = 6 - 1 = 5.$$

Таблица VI.9

$V_1' = 4$ $V_2' = 6$ $V_3' = 3$ $V_4' = 5$

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
$U_1' = 0$ A_1	30 4	20 6	5	7	50
$U_2' = 2$ A_2	6	90 4	8	9	90
$U_3' = -1$ A_3	5	50 7	70 4	130 6	250
$U_4' = 5$ A_4	0	40 0	0	0	40
Потребность	30	200	70	130	

Так как во втором столбце помимо x_{42} имеются положительные поставки x_{12}, x_{22}, x_{32}, исправляем потенциалы U_1', U_2' и U_3'

$$U_1' = U_1 - v_{42} = 1 - 1 = 0,$$

$$U_2' = U_2 - v_{42} = 3 - 1 = 2,$$

$$U_3' = U_2 - v_{42} = 0 - 1 = -1.$$

Далее, так как в первой строке помимо x_{12} есть перевозка x_{11}, а в третьей помимо x_{32} имеются перевозки x_{33} и x_{34}, исправляем потенциалы V_1, V_3 и V_4

$$V_1' = V_1 - v_{42} = 5 - 1 = 4,$$

$$V_3' = V_3 - v_{42} = 4 - 1 = 3,$$

$$V_4' = V_4 - v_{42} = 6 - 1 = 5.$$

Без изменений в данной задаче остается только потенциал четвертой строки (U_4).

Новые потенциалы проставляем слева и сверху от матрицы (см. табл. VI.9).

Проверим новый план на оптимальность. Для этого вычислим невязки в клетках без перевозок:

$$v_{13} = V_3' - U_1' - c_{13} = 3 - 0 - 5 = -2 < 0,$$

$$v_{14} = V_4' - U_1' - c_{14} = 5 - 0 - 7 = -2 < 0,$$

$$v_{21} = V_1' - U_2' - c_{21} = 4 - 2 - 6 = -4 < 0,$$

$$v_{23} = V_3' - U_2' - c_{23} = 3 - 2 - 8 = -7 < 0,$$

$$v_{24} = V_4' - U_2' - c_{24} = 5 - 2 - 9 = -6 < 0,$$

$$v_{31} = V_1' - U_3' - c_{31} = 4 - (-1) - 5 = 0,$$

$$v_{41} = V_1' - U_4' - c_{41} = 4 - 6 - 0 = -2 < 0,$$

$$v_{43} = V_3' - U_4' - c_{43} = 3 - 6 - 0 = -3 < 0,$$

$$v_{44} = V_4' - U_4' - c_{44} = 5 - 6 - 0 = -1 < 0.$$

План оптимален, так как все невязки отрицательны. Как ранее указывалось, суммарная стоимость перевозок $(30 \cdot 4 + 20 \cdot 6 + 90 \cdot 4 + 50 \cdot 7 + 70 \cdot 4 + 130 \cdot 6 + 40 \cdot 0) = 2010$.

Между прочим, в данной задаче начальный план, составленный методами наименьших стоимостей и двойного предпочтения (см. табл. VI.7), сразу является оптимальным.

При построении опорного решения или при его улучшении количество клеток, занятых перевозками, может оказаться меньше, чем $(m + n - 1)$. В этом случае нельзя построить систему потенциалов и такой план называется вырожденным. Для решения этой задачи в клетку (или клетки) без перевозок проставляют фиктивную перевозку малого объема ϵ. При этом объемы перевозок добавляют во столько клеток, чтобы план стал невырожденным, т.е. количество заполненных клеток равнялось бы $(m + n - 1)$. После этого решают задачу как невырожденную, а в оптимальном плане ϵ заменяют нулями. Если клетка,

в которой имеется перевозка ϵ, будет четной вершиной контура, то объем перевозок, на который улучшается план, равен ϵ.

§ 4. АЛГОРИТМ РЕШЕНИЯ ТРАНСПОРТНОЙ ЗАДАЧИ НА СЕТИ

На ряде случаев транспортную задачу целесообразно решать в сетевой постановке, отличающейся наглядностью.

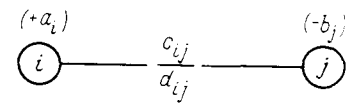
Транспортная сеть – это совокупность вершин или узлов (пункты отправления и приема грузов, а также промежуточные пункты) и соединяющих их транспортных коммуникаций или звеньев. На каждом звене проставляются удельная стоимость перевозки груза по данному транспортному участку c_{ij}, а при ограниченности пропускной способности звена также и ее максимальное значение d_{ij}. Около каждой вершины в скобках со знаком "+" или "–" проставляются объемы отправления и приема грузов. В случае отправления грузов ставится знак "+", приема его – знак "–".

На рис. VI.2 показана коммуникация, соединяющая поставщика i с потребителем j. Объем запасов у поставщика $(+a_i)$ и потребность j-го потребителя $(-b_j)$ проставлены у соответствующих вершин. В разрыве ребра, показывающего транспортную коммуникацию, в числителе указаны стоимость перевозки по звену (c_{ij}), а в знаменателе – пропускная способность транспортной коммуникации (d_{ij}). Нахождение оптимального при сетевой постановке транспортной задачи плана осуществляется методом потенциалов в порядке, рассмотренном ранее для матричной постановки.

1. Первоначально строится опорное решение. Четкий алгоритм построения допустимого плана при сетевой постановке транспортной задачи отсутствует. При построении первоначального плана стремятся все объемы грузов вывезти от поставщиков (т.е. из вершин-поставщиков должны выходить грузопотоки, равные запасам груза в данном пункте) и доставить потребителю (в вершины-потребители должны входить грузопотоки, равные потребности в грузе). При построении плана объем перевозок по звену не должен превышать его пропускной способности. Перевозки на сети обозначаются стрелками, показывающими направление потока. Над стрелками дан объем перевозимого груза.

2. Для построения системы потенциалов любой вершине присваивается какой-либо потенциал. Обходя сеть по звеньям с перевозками в любом направлении, на основании второго условия оптимальности последовательно определяются потенциалы вершин контура. Потенциал последующей вершины контура равен сумме потенциала предыдущей вершины и удельной стоимости перевозки единицы груза по данному транспортному участку, если направление обхода звена совпадает с направлением грузопотока, и их разности, если направление обхода звена не совпадает с направлением

Рис. VI.2. Элемент транспортной сети



грузопотока, т.е. $V_i = U_i + c_{ij}$ или $U_i = V_j - c_{ij}$. Потенциалы присваиваются только с использованием звеньев, по которым имеются перевозки и пропускная способность их не используется полностью.

3. Проверка плана на оптимальность осуществляется для всех звеньев без грузопотока, т.е. выполнение условий оптимальности проверяется на всех звеньях, которые не использовались при построении системы потенциалов.

Если отдельные коммуникации имеют ограниченную пропускную способность, то в задаче помимо ограничений VI.1 и VI.2 имеется еще система ограничений по пропускной способности

$$x_{ij} \leqslant d_{ij} \quad (i = \overline{1, n},\ j = \overline{1, m}). \tag{VI.5}$$

В этом случае используются не два (VI.1 и VI.2), а три условия оптимальности плана. Третье условие оптимальности относится к перевозкам, равным пропускной способности

$$V_j - U_i > c_{ij}, \text{ если } x_{ij} = d_{ij}, \tag{VI.6}$$

т.е. разность потенциалов должна быть больше стоимости перевозок по звену, если по нему осуществляется перевозка, объемом равной его пропускной способности. Это же условие оптимальности можно записать следующим образом

$$V_j - U_i = c_{ij} + \delta_{ij}.$$

Величина δ_{ij} называется прокатной оценкой звена и характеризует перерасход средств из-за ограниченности его пропускной способности.

Если имеются нарушения первого или третьего условия оптимальности, то план не является оптимальным и его необходимо улучшить.

4. При улучшении плана по звену, на котором нарушено первое условие оптимальности, должна пройти перевозка; если нарушено третье условие оптимальности, то на этом звене перевозка должна уменьшиться. Для улучшения плана выбирается звено с максимальной величиной нарушения и затем строится замкнутый контур, который образуется звеном с максимальным нарушением и звеньями с грузопотоками. Далее выбирается направление обхода контура. Контур обходят в направлении от меньшего потенциала к большему на звене с максимальным нарушением, если не соблюдено первое условие оптимальности, и от большего потенциала к меньшему, если нарушено третье условие оптимальности.

Объемы перевозок для улучшения плана равны минимальному

объему встречной перевозки или минимальной разности между пропускной способностью и объемом перевозки по звену; для звеньев с ограниченной пропускной способностью и попутной перевозкой

$$q = \min \begin{cases} x_{ij} & \text{встречная} \\ (d_{ij} - x_{ij}) & \text{попутная.} \end{cases} \qquad \text{(VI.7)}$$

Найденную величину улучшения плана добавляют к объемам звеньев замкнутого контура с попутной перевозкой и отнимают от объемов звеньев с встречной перевозкой.

После улучшения плана необходимо исправить систему потенциалов.

Исправление потенциалов производится у тех вершин, к которым изменилось направление грузопотока.

После исправления потенциалов проверяют выполнение условий оптимальности и шаги 2–4 повторяются до получения оптимального решения.

Пример VI.5. Найти оптимальный план перевозок для транспортной задачи, условия для которой показаны на рис. VI.3.

1. Начальный план показан на рис. VI.3 стрелками.

2. Для построения системы потенциалов присвоим вершине 1 потенциал 10 ($U_1 = 10$).

Тогда потенциалы остальных составят

$$U_6 = U_1 + c_{16} = 10 + 4 = 14,$$

$$U_4 = U_6 + c_{46} = 14 + 3 = 17,$$

$$U_5 = U_4 - c_{45} = 17 - 3 = 14,$$

$$U_2 = U_5 - c_{52} = 14 - 3 = 11,$$

$$U_3 = U_5 + c_{53} = 14 + 4 = 18.$$

Потенциалы присвоены всем вершинам сети. При этом звено 23 не применялось для построения системы потенциалов, так как на нем полностью использованы пропускные способности. Потенциалы показаны на рис. VI.3 рядом с вершинами.

3. Проверяем выполнение условия оптимальности на звеньях 12, 15 и 43 (без перевозок) и звене 23 (с "заполненной" пропускной способностью), которые не использовались при построении системы потенциалов.

На звеньях 12, 15, 34 проверяем выполнение первого условия оптимальности, а на звене 32 – третьего условия.

На звене 15 имеется нарушение первого условия оптимальности и план необходимо улучшить.

4. Для улучшения плана составляем замкнутый контур из звена 15 и звеньев с перевозками 14, 46, 61. При этом на звеньях 15 и 54 попутные перевозки, а на звеньях 46 и 61 – встречные. Минимальный объем встречной перевозки 170, а минимальная разность между пропускной способностью и объемом перевозки составляет (100 – 30) = 70 единиц. Следовательно, объем улучшения плана составляет 70 единиц. Прибавляя этот объем к попутным перевозкам, получим улучшенный план (рис. VI.4).

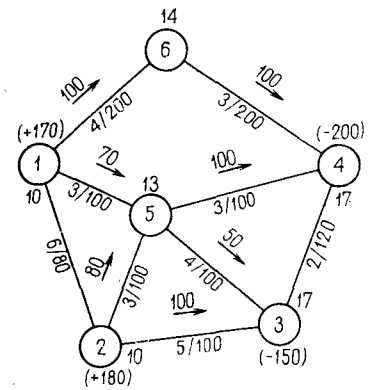


Рис. VI.3. Условия решения и начальный план транспортной задачи

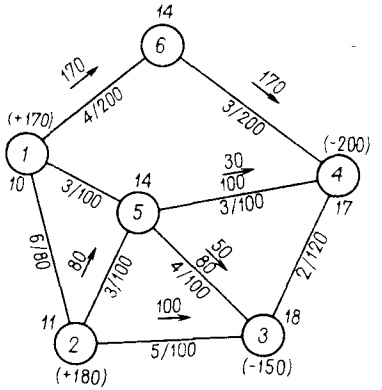


Рис. VI.4. Улучшенный план транспортной задачи

5. Исправляем систему потенциалов

$$U_5 = U_1 + c_{15} = 10 + 3 = 13,$$

$$U_2 = U_5 - c_{52} = 13 - 3 = 10,$$

$$U_6 = U_1 + c_{16} = 10 + 4 = 14,$$

$$U_4 = U_6 + c_{64} = 14 + 3 = 17,$$

$$U_3 = U_5 + c_{53} = 13 + 4 = 17.$$

6. Проверяем выполнение условий оптимальности для улучшенного плана. Для этого проверяем выполнение первого условия оптимальности на ребрах 12 и 34 и третьего условия на ребрах 54 и 23

$$\delta_{12} = U_2 - U_1 - c_{12} = 10 - 10 - 6 = -6 < 0,$$

$$\delta_{34} = U_4 - U_3 - c_{34} = 17 - 17 - 2 = -2 < 0,$$

$$\delta_{54} = U_4 - U_5 - c_{45} = 17 - 13 - 3 = 1 > 0,$$

$$\delta_{23} = U_3 - U_2 - c_{32} = 17 - 10 - 5 = 2 > 0.$$

Условия оптимальности выполнены, следовательно, улучшенный план является оптимальным.

Прокатная оценка $\delta_{23} = 2$ показывает, что если увеличить пропускную способность этой коммуникации, то экономия на единицу груза составит 2. Действительно, мы везем часть груза из пункта 2 в пункт 3 через пункт 5 с общими затратами ($c_{25} + c_{53} = 3 + 4$) = 7, в то время как при перевозке непосредственно по коммуникации 23 они составляют 5.

Прокатные оценки позволяют экономически оценивать эффективность мероприятий по увеличению пропускной способности коммуникаций сети.

§ 5. ПОСТАНОВКА И РЕШЕНИЕ ТРАНСПОРТНОЙ ЗАДАЧИ ПО КРИТЕРИЮ ВРЕМЕНИ

Иногда (перевозка возгорающихся полезных ископаемых, оперативное управление работой транспорта и др.) перевозку грузов от поставщиков $i = \overline{1, n}$ до потребителей $j = \overline{1, m}$ необходимо спланировать и организовать за минимальное время.

В этом случае должна быть построена матрица затрат времени на перемещение груза от i-го поставщика j-му потребителю, т.е. $\|t_{ij}\|$.

Среди всех допустимых планов транспортной задачи, определяемых ограничениями по вывозу грузов от поставщиков ($i = \overline{1, n}$) и обеспечением ими всех потребителей ($j = \overline{1, m}$), следует выбрать тот, у которого время доставки груза минимально.

Каждому допустимому плану $X = \{x_{ij}\}_{mn}$ соответствует набор времени перевозок $t_{ij} \in X$, состоящий из элементов матрицы $T = \|t_{ij}\|$, соответствующих положительным элементам x допустимого плана X. Общее время перевозки грузов, соответствующее плану X, $t_x = \max\{t_{ij}\} \in X$.

Оптимальный план по критерию времени должен обеспечивать минимальное время перевозок. Целевая функция задачи имеет вид

$$\min t_x = \min \max\{t_{ij}\}. \qquad \text{(VI.8)}$$

Ограничения задачи по критерию времени полностью аналогичны ограничениям транспортной задачи по критерию стоимости.

Условия транспортной задачи по критерию времени записываются в таблицу, аналогичную таблице транспортной задачи по критерию стоимости, только вместо стоимостей – c_{ij} в таблице записывается время доставки грузов – t_{ij}.

Решение транспортной задачи по критерию времени осуществляется в следующем порядке.

1. Методом северо-западного угла или наименьшего элемента строится опорное решение транспортной задачи.

2. Из всех клеток, занятых перевозками, выбирается наибольшее время t_{ij}; $t'_{ij} = \max\{t_{ij}\}$ x для $x_{ij} > 0$.

3. Клетки исходной матрицы, в которых время перевозки превышает максимальное для допустимого плана t'_{ij}, исключаются из рассмотрения (зачеркиваются).

4. Улучшается допустимый план, для этого:

а) строится контур с клеткой с максимальным временем перевозки t'_{ij}, клетками с перевозками и клеткой (клетками) без перевозок, в которых $t_{ij} < t'_{ij}$.

Нечетные вершины контура (отрицательная полуцепь) образуются клетками с перевозками, в том числе и клетка с t'_{ij}.

Четные вершины контура (положительная полуцепь) – клетки с t_{ij} меньшими, чем t'_{ij};

б) определяется объем перевозок, на величину которого улучшается план. Для этого среди вершин отрицательной полуцепи (нечетных) определяется минимальный объем перевозок;

в) найденный минимальный объем перевозок вычитается из всех вершин отрицательной полуцепи и прибавляется к вершинам положительной полуцепи.

5. Улучшение плана осуществляется до тех пор, пока не остается незанятых клеток, для которых время перевозок меньше, чем максимальное в допустимом плане, или клеток, которые могут быть включены в контур. В этом случае становится невозможным построение контура и полученный допустимый план является оптимальным.

Пример VI.6. Найти план перевозок, обеспечивающий минимальное время доставки груза для транспортной задачи, условия которой заданы матрицей (табл. VI.10).

Таблица VI.10

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	2	4	8	13	36
A_2	3	9	10	9	20
A_3	7	2	5	6	24
Потребность	22	18	26	14	

1. В данной задаче сумма запасов равна общей потребности, т.е. задача является закрытой.

2. Построим опорное решение задачи методом северо-западного угла (табл. VI.11).

Таблица VI.11

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	22 2	14 4	8	13	36
A_2	3	4 9	16 11	9	20
A_3	7	2	10 5	14 6	24
Потребность	22	18	26	14	

3. Находим клетку с перевозкой, имеющую наибольшее время перевозок: $\max t_{ij} = \{2, 4, 9, 11, 5, 6\} = 11$ принадлежит клетке 23. Вычеркиваем (исключаем из рассмотрения) клетки, имеющие время перевозок больше 11, так как $t_{14} = 13 > t_{23}$, то клетку 14 из рассмотрения исключаем и в последующих матрицах она зачеркнута.

4. Строим замкнутый контур с клеткой с максимальным временем перевозок. Замкнутый контур проходит через клетки 23, 13, 11, 21, причем в отрицательную полуцепь входят клетки 23 и 11. Минимальный объем перевозок в этих клетках равен 16. Отнимая этот объем (величину улучшения плана) от объемов в вершинах отрицательной полуцепи и добавляя его к вершинам положительной полуцепи (13, 21), получим новый план (табл. VI.12).

Таблица VI.12

Поставщик	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	6 2	14 4	16 8	13	36
A_2	16 3	4 9	11	9	20
A_3	7	2	10 5	14 6	24
Потребность	22	18	26	14	

5. Находим в новом плане клетку с максимальным временем перевозки $\max t_{ij} = \{2, 4, 3, 9, 5, 6\} = 9$ и клетки, превышающие это время, исключаем из рассмотрения $t_{23} = 11 > 9$.

6. Составим новый контур, включающий клетку с максимальным временем (22). Контур пройдет через клетки 22, 12, 1, 21. Минимальный объем в клетках отрицательной полуцепи равен 4. Отнимая этот объем из объемов перевозок в вершинах 22 и 11 и добавляя к объемам в вершинах положительной полуцепи 12 и 21, получим новый план (табл. VI.13).

Таблица VI.13

Поставщик	*Потребители*				Запасы
	B_1	B_2	B_3	B_4	
A_1	2 2	18 4	16 8	13	36
A_2	20 3	9	11	9	20
A_3	7	2	10 5	14 6	24
Потребность	22	18	26	14	

7. Минимальное время перевозок для нового плана составляет 8 и соответствует перевозке в клетке 13. Анализ полученного решения показывает, что невозможно составить новый контур, включающий клетку 13 и, следовательно, полученный план является оптимальным, а наименьшее время перевозки груза равно 8.

Контрольные вопросы и задачи

1. Каковы особенности транспортной задачи?

2. В чем отличия транспортной задачи от классических задач линейного программирования?

3. В чем экономический смысл потенциалов?

4. Преимущества и недостатки различных методов построения опорного решения?

5. Когда целесообразно использовать сетевую постановку транспортной задачи?

6. Задача. Имеется четыре карьера, производящих строительные материалы, и четыре потребителя сырья (строительных материалов). Известны объемы производства на каждом карьере, потребности в их продукции каждого из потребителей, а также стоимость перевозки 1 т продукции с i-го карьера j-му потребителю. Определить, при каких объемах грузоперевозок от i-го поставщика j-му потребителю суммарная стоимость перевозок будет минимальной.

Стоимости перевозок, потребности потребителей и запасы поставщиков приведены в табл. VI.14. Дополнительные данные, определяющие вариант решения задачи, указаны в табл. VI.15. В таблице необходимо найти номер варианта, слева в этой же строке стоят номера поставщиков, сверху в этом же столбце – номера потребителей. Например: по табл. VI.15 находим номер варианта – 38. Слева в первой колонке стоят числа 4–7, значит, номера поставщиков – с 4 по 7 (4, 5, 6 и 7). Далее сверху в этом же столбце находим номера потребителей – с 8 по 11 (8, 9, 10, 11). Затем из табл. VI.14 берем исходные данные, соответствующие потребителям 8, 9, 10 и 11 и поставщикам 4, 5, 6 и 7. Составляем матрицу для решения транспортной задачи. Она будет состоять из четырех строк и четырех столбцов.

7. Задача. На рис. VI.5 приведена транспортная сеть. В табл. VI.16 приведены стоимости транспортировки угля, пропускные способности коммуникаций, номера пунктов-поставщиков (шахт), объемы добычи по каждой шахте, а также номера пунктов-потребителей угля (ТЭЦ и обогатительные фабрики) и их потребность в угле.

Требуется:

составить оптимальный план перевозок;

определить, по какой коммуникации выгоднее всего увеличивать пропускную способность.

8. Задача. Составить оптимальный план перевозок по критерию времени. Условия транспортной задачи заданы матрицей (табл. VI.17).

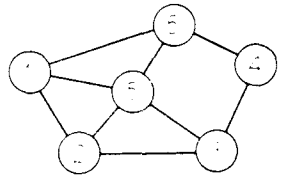


Рис. VI.5. Транспортная сеть

Таблица VI.14

Поставщик	Потребители													Запасы
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	2	1	3	2	4	6	8	5	7	6	1	4	3	120
2	1	4	1	2	2	3	5	4	6	5	4	3	5	140
3	1	2	4	3	1	4	3	2	5	6	5	4	2	230
4	3	4	2	1	5	6	4	7	3	2	4	5	6	200
5	5	3	4	2	4	3	5	6	3	5	2	3	3	180
6	1	2	3	4	5	6	7	5	3	2	4	6	3	150
7	3	4	5	3	2	6	7	3	5	4	2	3	6	190
8	4	2	5	6	3	4	2	6	4	5	2	5	4	220
9	2	3	6	4	5	7	3	5	2	3	7	4	3	100
10	4	2	5	3	4	6	7	4	3	4	5	2	7	170
11	3	2	5	3	6	7	5	8	3	5	4	2	5	130
12	5	7	3	4	8	2	4	3	6	4	5	2	3	250
13	2	4	3	7	5	4	5	3	2	6	4	3	4	210
Потребность	250	180	200	150	120	300	160	190	210	240	130	170	230	

Таблица VI.15

A	*B*									
	1–4	2–5	3–6	4–7	5–8	6–9	7–10	8–11	9–12	10–13
1–4	1	2	3	4	5	6	7	8	9	10
2–5	11	12	13	14	15	16	17	18	19	20
3–6	21	22	23	24	25	26	27	28	29	30
4–7	31	32	33	34	35	36	37	38	39	40
5–8	41	42	43	44	45	46	47	48	49	50
6–9	51	52	53	54	55	56	57	58	59	60
7–10	61	62	63	64	65	66	67	68	69	70
8–11	71	72	73	74	75	76	77	78	79	80
9–12	81	82	83	84	85	86	87	88	89	90
10–13	91	92	93	94	95	96	97	98	99	100

Таблица VI.16

№ варианта	Коммуникации: стоимость транспортирования/пропускная способность								Поставщики пункта/объем добычи	Потребители пункта/потребность
	1–2	1–5	1–6	2–3	2–5	3–4	3–5	4–6	5–6	
1	3/70	2/100	3/80	4/60	3/900	2/100	6/70	5/120	2/70	1/100 2/800 3/90 4/90
2	4/200	3/100	6/80	3/200	3/100	4/100	5/100	3/100	2/120	1/170 6/180 3/200 4/150
3	2/100	4/200	5/350	2/150	5/190	3/250	2/160	6/200	5/100	2/120 3/80 5/100 6/100
4	3/80	6/80	4/60	3/200	3/90	3/100	2/100	4/100	6/70	5/100 1/50 2/80 4/70
5	5/150	5/150	2/100	5/100	4/140	5/80	3/200	2/80	3/200	2/150 4/100 1/120 3/130
6	6/140	2/110	3/200	4/170	2/250	6/200	4/90	3/150	4/150	3/110 5/90 4/80 6/120
7	2/90	4/120	5/150	3/140	4/160	3/150	3/80	4/180	2/90	4/80 6/150 3/130 5/100
8	6/70	4/100	2/100	3/100	3/90	3/200	4/60	6/80	3/80	3/160 4/110 5/140 6/130
9	2/100	3/100	4/200	6/80	5/150	2/110	4/120	4/100	3/100	1/90 2/120 4/140 5/70
10	4/60	3/200	2/150	3/200	5/100	4/70	3/140	3/90	6/120	2/100 3/150 5/110 6/140
11	3/80	6/100	5/350	4/100	2/80	3/200	5/150	2/100	4/200	3/150 4/150 1/170 6/130
12	2/100	4/90	3/250	3/100	5/80	6/20	3/150	4/180	6/180	4/120 5/180 2/150 3/150
13	5/120	3/100	6/200	4/100	2/80	3/140	4/180	3/90	5/100	1/80 2/120 3/100 4/100
14	4/170	2/80	4/150	3/70	6/150	3/140	2/150	5/200	2/90	5/90 6/80 1/60 3/110
15	2/100	4/130	5/160	3/150	4/170	3/160	4/70	6/100	3/90	1/130 2/130 4/140 6/120
16	3/60	5/150	3/90	5/200	3/120	4/180	3/120	3/150	4/170	1/100 3/150 2/120 4/130

Таблица VI.17

Поставщик	Потребители				Запасы
	1	2	3	4	
1	7	11	9	5	10
2	10	12	8	5	18
3	6	9	8	4	7
Потребность	5	12	10	8	

Глава VII

ЦЕЛОЧИСЛЕННОЕ ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ И ЕГО ПРИМЕНЕНИЕ ПРИ РЕШЕНИИ ЗАДАЧ ПЛАНИРОВАНИЯ ГОРНОГО ПРОИЗВОДСТВА

§ 1. ОСОБЕННОСТИ ЦЕЛОЧИСЛЕННЫХ ЛИНЕЙНЫХ ЗАДАЧ И МЕТОДОВ ИХ РЕШЕНИЯ

При решении вопросов планирования горного производства, формулируемых в рамках задач математического программирования, довольно часто требуется получить целочисленное решение. Условие целочисленности решения характерно для задач:

распределения трудовых ресурсов и оборудования, планирования дискретного производства, а также ряда других, которые содержат переменные, обозначающие количество неделимых ресурсов или продукции;

выбора одного из ряда вариантов (например, механизации горных работ, структуры управления, маршрута, размещения производства и т.д.), где переменные (все или некоторые) могут принимать значение, равное только 0 или 1 (0 – соответствующий вариант отвергается, 1 – принимается).

Данная глава посвящена изучению методов линейного целочисленного программирования и их применению при решении задач планирования горного производства.

Задача линейного целочисленного программирования (ЦЛП) – это задача математического программирования, в которой все или некоторые переменные должны принимать только целочисленные значения, а целевая функция и функции, входящие в ограничения – ли-

нейные. То есть это такая нелинейная задача, которая могла быть линейной, если бы не требования целочисленности ряда переменных.

В общем виде модель задачи линейного целочисленного программирования может быть представлена следующим образом

$$Z = \bar{c}\bar{x} \to \max(\min), \qquad \text{(VII.1)}$$

$$A\bar{x} + \bar{B} = 0,\ \bar{x} \geqslant 0,\ x_j \ -\ \text{целые числа},\ j \in J,$$

где J – множество значений индекса j, соответствующих целочисленным переменным.

Причем задача называется полностью целочисленной, если все значения j принадлежат J, и частично целочисленной – в противном случае.

Если целочисленные переменные могут принимать значения, равные только 0 или 1, то (VII.1) является моделью задачи с булевыми переменными.

Заметим, что любая целочисленная переменная может быть выражена через булевые переменные. Допустим $x(0 \leqslant x \leqslant n)$ – целочисленная переменная, значения которой не превышают некоторого целого положительного числа n. Тогда, если $\delta_1, \delta_2, \ldots, \delta_n$ – булевые переменные, все допустимые значения x представляются в виде

$$x = \delta_1 + \delta_2 + \ldots + \delta_n. \qquad \text{(VII.2)}$$

Другое (более экономичное) двоичное представление, в котором количество булевых переменных меньше n, задается формулой

$$x = \delta_1 + 2\delta_2 + 2^2\delta_3 + \ldots + 2^{k-1}\delta_k, \qquad \text{(VII.3)}$$

где k – наименьшее целое число, удовлетворяющее условию $2^k - 1 \geqslant n$.

Таким образом, любую целочисленную задачу можно записать в булевых переменных.

Рассмотрим возможные подходы к решению задач ЦЛП. Допустим, что имеется задача линейного программирования (ЛП) с двумя основными переменными, ограничения которой определяют область (многоугольник) допустимых решений (рис. VII.1). Требования целочисленности x_1 и x_2 приводят к вырождению множества допустимых решений в систему точек, причем указанное множество не является выпуклым.

Задачу ЦЛП можно решить, например, как задачу ЛП без учета условий целочисленности переменных, а затем округлить полученное решение с избытком или недостатком. При этом будет получено некоторое целочисленное решение. Однако использование такого подхода требует обязательной проверки допустимости решения. Таким методом часто пользуются при решении практических задач, особенно, когда значения переменных настолько велики, что можно пренебречь ошибками округления. При решении задач, в которых целочисленная переменная принимает малые значения, округление может привести

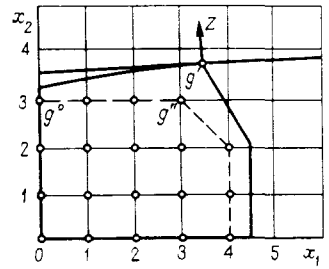


Рис. VII.1. Множество допустимых решений задачи ЦЛП, (g^0 – оптимальное решение задачи ЦЛП, g' – оптимальное решение задачи ЛП, g'' – ближайшая к g' целочисленная точка области допустимых решений задачи ЦЛП)

к далекому от истинного оптимума целочисленному решению. Кроме того, при решении задач большой размерности такой метод требует слишком много машинного времени. Например, пусть оптимальное решение соответствующей задачи ЛП имеет вид $x_1 = 3{,}4$; $x_2 = 3{,}7$ (см. рис. VII.1). Для получения приближенного оптимального целочисленного решения необходимо рассмотреть четыре точки (4; 3), (3; 3), (3; 4) и (4; 4) и выбрать среди них допустимую точку с наилучшим значением целевой функции. В рассматриваемом случае единственной допустимой является точка (3; 3). Если в задаче имеется 10 целочисленных переменных, то следует рассмотреть $2^{10} = 1024$ варианта целочисленного решения. Заметим, что даже перебор всех вариантов не гарантирует получения оптимального решения задачи ЦЛП. Так, в рассматриваемом случае, как следует из рис. VII.1, оптимальным целочисленным решением будет $x_1 = 0$; $x_2 = 3$.

Таким образом, эвристический метод решения задачи ЦЛП, основанный на округлении оптимального решения соответствующей задачи ЛП, позволяет получать в общем случае только приближенное решение. Причем в случае задачи с булевыми переменными оперирование с дробными значениями бессмысленно и, следовательно, округление является логически неприемлемым.

Для определения истинных значений целочисленных переменных применяют метод отсечений и метод ветвей и границ.

Первый базируется на идее деформации области допустимых решений задачи ЛП таким образом, чтобы от нее было отсечено оптимальное нецелочисленное решение, но сохранены все допустимые целочисленные решения. То есть была бы построена область, для которой оптимальное решение соответствующей задачи ЛП совпадало с оптимальным решением задачи ЦЛП. Указанная деформация осуществляется путем введения специальных дополнительных ограничений. Для условий нашего примера – это ограничение $x_2 \leqslant 3$.

Второй метод основан на идее проверочной подстановки всех допустимых целочисленных значений в целевую функцию и сравнении полученных результатов. Естественно, что удобной для использования эта идея становится лишь тогда, когда полный перебор может быть сокращен дополнительными рассуждениями, основанными на прин-

цине разветвления комбинаторного программирования. Так, для условий рассматриваемого примера ветвление базируется на том факте, что целочисленное значение x_1 должно удовлетворять неравенству либо $x_1 \leqslant 3$, либо $x_1 \geqslant 4$.

В настоящее время оба метода решения задач ЦЛП интенсивно разрабатываются, так что постоянно появляются новые варианты. Однако ни один из вариантов не обеспечивает желательной эффективности соответствующих вычислений, что особенно проявляется при увеличении размерности задачи. Одна из основных трудностей, присущих обоим методам при использовании ЭВМ, заключается в том, что необходимо решить, рассматривать ли число, несущее в себе погрешности округления, как целое или нет. Даже наличие специальных алгоритмов, применяемых при решении задач с целочисленными параметрами и позволяющих обойтись без оперирования с дробями, не упрощает ситуацию, поскольку в ряде случаев такие алгоритмы сходятся крайне медленно. Таким образом, в отличие от задач ЛП, время решения которых относительно невелико, реализация методов отсечений и ветвей и границ в ряде случаев весьма затруднительна. Количество переменных в задаче ЦЛП – основной фактор, влияющий на процесс вычислений. Поэтому при построении модели необходимо ограничиться меньшим числом целочисленных переменных, причем с минимальными интервалами их допустимых значений. При решении конкретных задач этого добиваются путем приближенного представления целочисленных переменных с помощью непрерывных, а также сужением интервалов допустимых значений. Кроме того, в ряде случаев целесообразно отказаться от введения дополнительных булевых переменных в пользу применения прямых методов решения. Добавление новых ограничений, особенно включающих целочисленные переменные, обычно уменьшает время решения задач ЦЛП. Повышение эффективности методов решения целочисленных задач в настоящее время по всей вероятности больше связано с совершенствованием средств вычислительной техники (увеличение быстродействия и объема памяти, появление возможности распараллеливания вычислений), чем с дальнейшим развитием теории.

Перед тем как приступить к подробному рассмотрению указанных методов, продемонстрируем практическую ценность использования булевых переменных при построении моделей целочисленных задач и рассмотрим типичные целочисленные модели планирования горного производства.

§ 2. ИСПОЛЬЗОВАНИЕ БУЛЕВЫХ ПЕРЕМЕННЫХ ПРИ ПОСТРОЕНИИ МОДЕЛЕЙ ЦЕЛОЧИСЛЕННЫХ ЗАДАЧ ПЛАНИРОВАНИЯ

Практическую ценность представляют открываемые целочисленным программированием возможности приведения "некорректных" задач к стандартному виду задач математического программирования. Использование специальных приемов позволяет получить решение в ряде случаев, когда его поиск с помощью прямых способов оказывается затруднительным.

Рассмотрим, например, задачу планирования перевозок полезного ископаемого с группы карьеров на несколько объектов. Причем будем предполагать, что эта задача формализуется в рамках транспортной модели, особенностью которой является наличие разрыва функции затрат $c_{ij}(x_{ij})$ на перевозку объема x_{ij} полезного ископаемого от i-го поставщика j-му потребителю

$$c_{ij}(x_{ij}) = \begin{cases} k_{ij} + c_{ij} x_{ij}, & x_{ij} > 0, \\ 0, & x_{ij} = 0, \end{cases} \qquad \text{(VII.4)}$$

где k_{ij} – капитальные затраты на строительство или реконструкцию дороги, связывающей i-й карьер с j-м объектом; c_{ij} – стоимость перевозки единичного объема полезного ископаемого с i-го карьера на j-й объект.

Так как в этом случае целевая функция не является линейной по x_{ij}, то транспортная задача с постоянными элементами затрат оказывается непригодной для дальнейшего исследования.

Введем вспомогательные булевые переменные

$$\delta_{ij} = \begin{cases} 0, & x_{ij} = 0, \\ 1, & x_{ij} > 0, \end{cases} \qquad \text{(VII.5)}$$

связанные с основными переменными линейным неравенством $x_{ij} \leqslant M \delta_{ij}$. Причем $M > 0$ достаточно велико, чтобы условие $x_{ij} \leqslant M$ выполнялось для всех допустимых объемов перевозок.

Представим теперь целевую функцию исходной задачи как

$$Z = \sum_i \sum_j (c_{ij} x_{ij} + k_{ij} \delta_{ij}) \to \min, \qquad \text{(VII.6)}$$

а систему ограничений дополним требованиями

$$\left.\begin{array}{l} x_{ij} \leqslant M \delta_{ij} \text{ для всех } i \text{ и } j, \\ 0 \leqslant \delta_{ij} \leqslant 1, \delta_{ij} - \text{целое для всех } i \text{ и } j. \end{array}\right\} \qquad \text{(VII.7)}$$

Рассмотрим подробнее первое из дополнительных требований (VII.7). При $x_{ij} > 0$ имеем $\delta_{ij} = 1$ и целевая функция включает постоянные затраты k_{ij}. Если $x_{ij} = 0$, то δ_{ij} может принимать значения 0

или 1. Однако поскольку $k_{ij} > 0$ и Z требуется минимизировать, переменная δ_{ij} должна быть равна нулю.

Таким образом, введение дополнительных булевых переменных позволило преобразовать задачу к виду, "более приемлемому" с аналитических позиций. Преобразованная задача является частично целочисленной с булевыми переменными.

При решении задач планирования горного производства может возникнуть ситуация, когда требуется выполнение нескольких из общего количества линейных ограничений. Причем заранее неизвестно, какие именно ограничения должны выполняться. Аналогичная ситуация возникает, когда правая часть некоторого линейного ограничения может принимать одно из нескольких значений. Это так называемые задачи с дихотомией. Прямой способ решения заключается в следующем. Определяется количество возможных наборов ограничений и решается соответствующее количество различных линейных задач. Окончательное решение выбирается по лучшему значению целевой функции. Аналогично поступают и в случае с несколькими возможными значениями правой части одного из ограничений. Однако, если ввести специальным образом булевые переменные, то решение можно получить в рамках одной частично целочисленной задачи.

Допустим, что планирование добычных работ на карьере ведется для условий двух альтернативных вариантов требований обогатительной фабрики по содержанию одного из компонентов в суммарном рудопотоке

$$\frac{\sum\limits_i x_i a_i}{\sum\limits_i x_i} \leqslant a' \text{ или } a'', \qquad \text{(VII.8)}$$

где x_i – объем добычи руды на i-м участке; a_i – содержание указанного компонента в руде, добываемой на i-м участке; a' и a'' – альтернативные варианты максимально допустимого содержания указанного компонента в суммарном рудопотоке.

Введем вспомогательную булеву переменную δ, принимающую значение 1, если в правой части ограничения (VII.8) будет стоять a', и значение 0, если – a''.

Затем представим VII.8 следующим образом

$$\left.\begin{aligned} &\frac{\sum\limits_i x_i a_i}{\sum\limits_i x_i} \leqslant (1-\delta)M + a', \\ &\frac{\sum\limits_i x_i a_i}{\sum\limits_i x_i} \leqslant \delta M + a'', \\ &0 \leqslant \delta \leqslant 1,\ \delta \text{ – целое.} \end{aligned}\right\} \qquad \text{(VII.9)}$$

где M – достаточно большое положительное число.

В системе (VII.9) одно из двух первых ограничений становится избыточным в зависимости от значений булевой переменной δ.

Аналогично можно построить модель частично целочисленной задачи и в случае планирования добычных работ на руднике, когда вследствие принятой технологической схемы разработки и допустимых вариантов организации погрузочных и доставочных работ производственные возможности каждого эксплуатационного блока могут принимать одно из нескольких дискретных значений

$$x_i = b_{i1}, b_{i2}, ..., b_{ij}, ..., \text{ или } b_{ik} \text{ для всех } i, \qquad \text{(VII.10)}$$

где x_i – планируемый объем добычи руды в i-м блоке; b_{ij} – j-й вариант производственных возможностей i-го блока по добыче руды.

Введем вспомогательные булевы переменные δ_{ij} и представим (VII.10) следующим образом

$$\left.\begin{aligned} &x_i = \sum_j b_{ij}\delta_{ij} \text{ для всех } i, \\ &\sum_j \delta_{ij} = 1 \text{ для всех } i, \\ &0 \leqslant \delta_{ij} \leqslant 1,\ \delta_{ij} \text{ – целое для всех } i \text{ и } j. \end{aligned}\right\} \qquad \text{(VII.11)}$$

Причем в (VII.11) $\delta_{ij} = 1$, если принят вариант b_{ij}, и $\delta_{ij} = 0$ в противном случае.

§ 3. МОДЕЛЬ ПЛАНИРОВАНИЯ РАЗМЕЩЕНИЯ УГЛЕОБОГАТИТЕЛЬНЫХ ФАБРИК*

Серьезным препятствием в использовании одного из важнейших энергоносителей – угля, с точки зрения экологии, является наличие в нем серы. Топки по сжиганию угля – один из основных источников загрязнения воздушного бассейна диоксидом серы. С помощью механического процесса уголь с различным содержанием серы может быть переработан в смесь, которая удовлетворяет установленным природоохранным нормам в районе его потребления.

Предположим, что известны районы добычи и потребления угля, а также возможного расположения углеобогатительных фабрик. Необходимо решить задачу размещения фабрик с целью обеспечения условий для получения угля, удовлетворяющего одновременно запросам потребителей и стандартам по охране окружающей среды. В результате решения должны быть получены ответы на следующие вопросы. Сколько нужно построить фабрик? Где их нужно строить? Каких потребителей

*Модель разработана А. Рейвиндраном (Г. Рейклиет, А. Рейвиндран, К. Рэгсдел. Оптимизация в технике. М., Мир, 1986).

должна обеспечивать каждая фабрика и с каких шахт получать уголь? Каковы оптимальные соотношения различных углей, предназначенных для переработки, с точки зрения соответствия природоохранным нормам в районах его потребления?

Поиск решений указанной задачи осуществляется на основе частично целочисленной модели с булевыми переменными. В модели индексы i, j, k используются для обозначения шахт, обогатительных фабрик и потребителей соответственно. Переменные модели определены следующим образом: x_{ij} – количество угля, отправляемого с i-й шахты на j-ю обогатительную фабрику (фабрику, расположенную в j-м районе); y_{ij} – количество угля, отправляемого с j-й обогатительной фабрики k-му потребителю; λ_{ij} – булева переменная; $\lambda_{ij} = 1$, если уголь отправляется с шахты на j-ю обогатительную фабрику, в противном случае $\lambda_{ij} = 0$; δ_{jk} – булева переменная; $\delta_{jk} = 1$, если уголь отправляется с j-й обогатительной фабрики k-му потребителю, в противном случае $\delta_{jk} = 0$; ϵ_j – булева переменная; $\epsilon_j = 1$, если для строительства обогатительной фабрики используется j-й район, в противном случае $\epsilon_j = 0$.

Критерием оптимальности являются суммарные затраты системы, в состав которых входят следующие.

1. Затраты на добычу угля.

2. Затраты на переработку угля, включающие постоянные (капитальные) на строительство обогатительной фабрики и переменные, пропорциональные объему перерабатываемого угля. Расходы на строительство вместе с другими капитальными затратами (на приобретение оборудования и др.) приводятся путем дисконтирования за средний срок работы фабрики к погодовой стоимости.

3. Транспортные затраты, зависящие от места добычи, назначения и количества груза. Предполагается, что транспортные затраты включают постоянную составляющую, не зависящую от количества перевозимого угля, и переменную, пропорциональную расстоянию и объему перевозимого угля. Постоянные транспортные затраты учитывают расширение существующей транспортной системы или строительство новых транспортных коммуникаций. Постоянная составляющая также приводится путем дисконтирования к погодовой стоимости за средний срок пользования введенными транспортными коммуникациями.

Таким образом, целевая функция системы имеет вид

$$Z(x, y, \lambda, \delta, \epsilon) = \sum_i D_i \left(\sum_j x_{ij}\right) + \sum_j c_j \epsilon_j + \\ + \sum_j B_j \left(\sum_i x_{ij}\right) + \sum_i \sum_j (S_{ij} \lambda_{ij} + T_{ij} x_{ij}) + \\ + \sum_j \sum_k (F_{jk} \delta_{jk} + G_{jk} y_{jk}) \to \min, \qquad \text{(VII.12)}$$

где D_i – затраты на добычу 1 т угля на i-й шахте; S_{ij} – постоянные

расходы на перевозку угля с i-й шахты на j-ю обогатительную фабрику; T_{ij} – затраты на перевозку 1 т угля с i-й шахты на j-ю обогатительную фабрику; c_j – ежегодные постоянные затраты на переработку 1 т угля на j-й фабрике; B_j – переменные затраты на переработку 1 т угля на j-й фабрике; F_{jk} – постоянные расходы на перевозку с j-й фабрики k-му потребителю; G_{jk} – затраты на перевозку 1 т угля с j-й фабрики k-му потребителю.

Ограничения модели содержат следующие требования.

1. Количество поставляемого шахтой угля не может превысить ее мощность по его добыче

$$\sum_j x_{ij} \leqslant A_i \text{ для всех } i, \qquad \text{(VII.13)}$$

где A_i – мощность i-й шахты по добыче угля.

2. Объем поставок потребителю должен удовлетворять его потребности

$$\sum_j y_{jk} \geqslant N_k \text{ для всех } k. \qquad \text{(VII.14)}$$

3. Объем перевозок с фабрики не может превысить мощность этой фабрики

$$\sum_k y_{jk} \leqslant P_j \epsilon_j \text{ для всех } j, \qquad \text{(VII.15)}$$

где P_j – мощность j-й фабрики (предполагается, что нет запасов обогащенного угля).

4. Общий объем перевозок обогащенного угля с фабрики не может превысить (с учетом поправочного коэффициента) количества угля, доставляемого на нее с шахт

$$E \sum_i x_{ij} \geqslant \sum_k y_{jk} \text{ для всех } j, \qquad \text{(VII.16)}$$

где E – показатель эффективности процесса переработки.

5. Смесь углей, производимая на обогатительной фабрике, должна удовлетворять требованиям потребителя по содержанию серы в угле

$$\sum_i H_i x_{ij} - L_k \sum_i x_{ij} \leqslant M(1 - \delta_{jk}) \text{ для всех } j \text{ и } k, \qquad \text{(VII.17)}$$

где H_i – содержание серы в угле, добываемом на i-й шахте; L_k – максимально допустимое содержание серы в угле для k-го потребителя; M – достаточно большое положительное число.

Если $\delta_{jk} = 1$, то (VII.17) можно представить как

$$\frac{\sum_i H_i x_{ij}}{\sum_i x_{ij}} \leqslant L_k. \qquad \text{(VII.18)}$$

В противном случае ограничение (VII.17) становится избыточным.

6. Необходимо обеспечить введение в модель постоянных транспортных издержек

$$y_{jk} - M\delta_{jk} \leqslant 0 \text{ для всех } j \text{ и } k, \qquad \text{(VII.19)}$$

$$x_{ij} - M\lambda_{ij} \leqslant 0 \text{ для всех } i \text{ и } j. \qquad \text{(VII.20)}$$

7. Переменные модели должны отвечать условиям

$$0 \leqslant \lambda_{ij}, \delta_{jk}, \epsilon_j \leqslant 1; \ \lambda_{ij}, \delta_{jk}, \epsilon_j \text{ – целые для всех } i, j, k; \qquad \text{(VII.21)}$$

$$x_{ij}, y_{jk} \geqslant 0 \text{ для всех } i, j, k. \qquad \text{(VII.22)}$$

§ 4. МОДЕЛЬ ОПЕРАТИВНОГО ПЛАНИРОВАНИЯ РАССТАНОВКИ САМОХОДНОГО ОБОРУДОВАНИЯ ПО ОЧИСТНЫМ БЛОКАМ РУДНИКА

Одно из прогрессивных направлений дальнейшего совершенствования технологии и организации добычи руд подземным способом – использование самоходного оборудования. При этом наряду с улучшением технологических и экономических показателей повышается управляемость объекта. Первоначальное планирование расстановки оборудования осуществляется при составлении месячной программы работы рудника исходя из сложившейся технологической ситуации, возможностей блоков, содержания металлов в руде по блокам, заданий по добыче руды и количеству металлов в ней, требований обогатительной фабрики, количества имеющегося оборудования и т.д. В процессе реализации программы из-за влияния ряда случайных факторов (поломок оборудования, изменения горно-геологических условий и др.) фактически выполненные объемы работ обычно отклоняются от запланированных и в случае угрозы срыва выполнения месячного плана по добыче руды и количеству металлов в ней появляется необходимость в корректировке расстановки оборудования. При этом, естественно, следует стремиться к минимальным изменениям в первоначальной его расстановке, обеспечивающей запланированное календарное распределение добываемых объемов руд из отдельных очистных блоков и, следовательно, соблюдение принятой концентрации горных работ. Кроме того, такая постановка соответствует стремлению минимизировать суммарные затраты времени на перестановку машин по блокам.

Задача оперативного планирования расстановки оборудования формулируется в рамках целочисленной задачи и решается на основе типовой модели для каждого из видов оборудования – доставочного, погрузочного и бурового. В модели индексы i и s используются для обозначения очистных блоков и видов металла. За оптимизируемые переменные приняты x_i – количество единиц самоходного оборудования (здесь и ниже имеется в виду определенный вид оборудования)

в i-м блоке, которое должно работать с момента перестановки до конца месяца. Критерий оптимальности – количество перемещений оборудования. Целевая функция модели имеет вид

$$\sum_i |k_i - x_i| \to \min, \qquad \text{(VII.23)}$$

где k_i – первоначальное (плановое) количество оборудования в i-м блоке.

Ограничения модели содержат следующие требования.

1. Месячный план по добыче руды из каждого очистного блока должен быть выполнен

$$x_i q_i t \geqslant d_i \text{ для всех } i, \qquad \text{(VII.24)}$$

где q_i – сменная производительность единицы оборудования в i-м блоке (для бурового оборудования в пересчете на 1 т взорванной рудной массы); t – количество рабочих смен, оставшихся до конца месяца; d_i – объем руды, который необходимо добыть в i-м блоке за оставшееся до конца месяца количество смен для выполнения месячного плана.

2. Месячный план по количеству металлов в руде, добытой по руднику в целом, должен быть выполнен

$$\sum_i x_i q_i a_{si} \epsilon_s t \geqslant M_s \text{ для всех } s, \qquad \text{(VII.25)}$$

где a_{si} – содержание s-го металла в руде i-го блока; ϵ_s – коэффициент извлечения металла s-го вида; M_s – количество металла s-го вида, необходимое для выполнения месячного плана по руднику за оставшееся количество смен.

3. Количество машин, предназначенных для работы в i-м блоке, должно быть не больше, чем его возможности по размещению оборудования

$$x_i \leqslant B_i \text{ для всех } i, \qquad \text{(VII.26)}$$

где B_i – максимальное количество самоходного оборудования, которое может работать в i-м блоке.

4. Общее количество расставляемого оборудования не может превышать общего его количества на руднике

$$\sum_i x_i \leqslant R, \qquad \text{(VII.27)}$$

где R – общее количество единиц самоходного оборудования на руднике.

5. Неотрицательные переменные модели могут принимать только целые значения

$$x_i \geqslant 0, \ x_i \text{ – целое для всех } i. \qquad \text{(VII.28)}$$

Целевая функция (VII.23) модели не является линейной вследствие наличия в ней модуля. Для того чтобы построить линейную целочисленную модель оперативного планирования расстановки самоходного оборудования, к ограничениям (VII.24)–(VII.28) добавляется система

$$k_i - y_i \leqslant x_i \leqslant k_i + y_i,\ y_i \geqslant 0 \text{ для всех } i, \qquad \text{(VII.29)}$$

где y_i – вспомогательная переменная.

Введение системы (VII.29) позволяет представить целевую функцию следующим образом

$$\sum_i y_i \to \min. \qquad \text{(VII.30)}$$

Заметим, что в силу (VII.29)–(VII.30) в решении задачи (VII.23)–(VII.30) y_i всегда будет иметь целочисленное значение.

§ 5. МОДЕЛЬ ПЛАНИРОВАНИЯ РАЗВИТИЯ ШАХТНОГО ФОНДА ПРОИЗВОДСТВЕННОГО ОБЪЕДИНЕНИЯ*

Совершенствование структуры шахтного фонда объединения заключается в выполнении ряда мероприятий по поддержанию мощностей действующих шахт и повышению эффективности их работы, реконструкции действующих и строительству новых. Обычно существуют альтернативные варианты совершенствования структуры шахтного фонда, отличающиеся динамикой мощностей шахт, вариантами их технического переоснащения и развития горных работ. Естественно, что эффективность вариантов различна.

Рациональное планирование мощностей шахт осуществляется поэтапно. Первоначально разрабатываются варианты развития по каждой шахте (шахтоварианты). Для формирования шахтовариантов анализируют мощности по отдельным технологическим процессам и участкам шахтного поля, на основе анализа выявляют "узкие места" и разрабатывают мероприятия по их ликвидации. Построение шахтовариантов может осуществляться на основе имитационного моделирования работы отдельных участков и шахты в целом. Затем (на втором этапе) рассчитываются технико-экономические показатели (себестоимость, капитальные вложения, приведенные затраты и др.) по каждой шахте (по годам).

Технико-экономические показатели рассчитываются отдельно для действующих, реконструируемых и строящихся (или намечаемых к строительству) шахт с различными периодами освоения мощности.

На третьем этапе на основе экономико-математического моделирования определяются оптимальные варианты развития шахт. Построим экономико-математическую модель для решения данной задачи, обозначив через i – номер шахты, $i = 1, 2, ..., m$; j – вариант развития шахты, $j = 1, 2, ..., n$; t – год планового периода T, $t = 1, 2, ..., T$; A_{ijt} – производственная мощность i-й шахты в год t при j-м варианте; C_{ijt} – эксплуатационные расходы (без реновационных отчислений) на добычу 1 т на i-й шахте в год t при j-м варианте развития; K_{ijt} – то же, капитальные вложения; r^k_{ijt} – расход ресурсов k-го типа (трудовых, энергетических, оборудования и др.) на i-й шахте в год t, при j-м варианте развития; A_t – плановый объем добычи в год t по группе шахт; K_t – объем капиталовложений на развитие всех шахт объединений в год t; R^k_t – количество ресурсов k-го типа в год t.

Управляемыми переменными являются: x_{ij} – булевые переменные; $x_{ij} = 1$, если для i-й шахты принимается j-й вариант развития и $x_{ij} = 0$ в противном случае.

За критерий оптимальности принимаются суммарные (приведенные к одному моменту оценки) затраты на плановый период времени T.

Тогда целевая функция задачи имеет вид

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{t=1}^{T}(C_{ijt} + K_{ijt})A_{ijt}x_{ij}(1 + E_н)^{-t} \to \min, \qquad \text{(VII.31)}$$

где $E_н$ – нормативный коэффициент экономической эффективности для приведения разновременных затрат ($E_н = 0{,}8$).

При решении задачи необходимо учесть ограничения:

по выполнению плана по добыче угля

$$\sum_{i=1}^{n}\sum_{j=1}^{m} A_{ijt}x_{ij} \geqslant A_t,\ t = 1, 2, ..., T; \qquad \text{(VII.32)}$$

по объему капитальных вложений

$$\sum_{i=1}^{n}\sum_{j=1}^{m} A_{ijt}K_{ijt}x_{ij} \leqslant K_t,\ t = 1, 2, ..., T; \qquad \text{(VII.33)}$$

по ресурсам

$$\sum_{i=1}^{n}\sum_{j=1}^{m} A_{ijt}r^k_{ijt}x_{ij} \leqslant R^k_t,\ t = 1, 2, ..., T,\ k = 1, 2, ..., K; \qquad \text{(VII.34)}$$

по выполнению единственного варианта развития шахты

$$\sum_{j=1}^{m} x_{ij} = 1,\ i = 1, 2, ..., n. \qquad \text{(VII.35)}$$

§ 6. МОДЕЛЬ ЗАДАЧИ О РАСКРОЕ

К линейным целочисленным задачам относится известная задача о раскрое, впервые сформулированная акад. В.В. Канторовичем. Задачи о раскрое заключаются в обосновании, из каких заготовок и каким

*Модель разработана В.М. Еремеевым, И.П. Ромашкиным, В.Л. Браверманом.

образом получить требуемую продукцию, чтобы минимизировать количество отходов (или из имеющегося сырья получить максимум продукции).

Приведем модель задачи о раскрое.

На предприятии имеются различные заготовки сырья $i = 1, 2, 3, \ldots, m$ (бревна разной длины на лесоскладе шахты, неодинаковых размеров мраморные блоки на камнеобрабатывающем предприятии, листы стали – на машиностроительном заводе). Каждая заготовка имеется в количестве b_i. Из имеющегося сырья требуется изготовить продукцию $j = 1, 2, 3, \ldots, n$ в количестве N_j с минимальным количеством отходов. Каждую заготовку можно раскроить различным способом $k = 1, 2, 3, \ldots, l$. При этом известны выход продукции a_{ijk} из любых заготовок при различных способах раскроя и количество отходов сырья, получаемое при этом, c_k.

Примем за управляемые переменные – количество заготовок, раскраиваемых k-м способом, x_{ik}. Тогда из i-й заготовки при k-м способе раскроя будет получаться $a_{ijk}\, x_{ik}$ видов j-й продукции, а количество отходов при этом составит $x_{ik}\, c_k$. За критерий оптимальности примем минимум отходов, тогда целевая функция запишется следующим образом

$$\sum_{i=1}^{m} \sum_{k=1}^{l} x_{ik}\, c_k \to \min \qquad \text{(VII.36)}$$

Задачу необходимо решить при соблюдении ограничений:

по числу заготовок

$$\sum_{k=1}^{l} x_{ik} \leqslant b_i, \quad i = 1, 2, \ldots, m; \qquad \text{(VII.37)}$$

по плану выпуска продукции

$$\sum_{i=1}^{m} \sum_{k=1}^{l} x_{ik}\, a_{ijk} \geqslant N_j, \quad j = 1, 2, \ldots, n; \qquad \text{(VII.38)}$$

по неотрицательности и целочисленности переменных

$$x_{ik} \geqslant 0, \quad x_{ik} - \text{целое}, \quad i = 1, 2, \ldots, m, \quad k = 1, 2, \ldots, l. \qquad \text{(VII.39)}$$

§ 7. РЕШЕНИЕ ЦЕЛОЧИСЛЕННЫХ ЗАДАЧ МЕТОДОМ ОТСЕЧЕНИЙ

Основная идея решения линейных целочисленных задач методом отсечений, первоначально предложенная Дж. Данцигом, Д. Фулкерсоном и С. Джонсоном, заключается в следующем (рис. VII.2). Сначала решается задача ЛП-1, полученная из задачи ЦЛП (VII.1) путем исключения ограничений по целочисленности. Если полученное решение целочисленно, то оно является оптимальным решением задачи ЦЛП (VII.1), так как ее любое допустимое решение – допустимое решение задачи ЛП-1. Если полученное решение нецелочисленно, то к задаче ЛП-1 присоединяют дополнительные ограничения, обладающие тем свойством, что множество допустимых решений новой задачи линейного программирования не содержит оптимального решения предыдущей задачи ЛП-1, и содержит все допустимые решения задачи ЦЛП (VII.1). Затем решается новая задача ЛП-2. Если ее решение допустимо для задачи ЦЛП (VII.1), то оно и оптимально для нее. Если нет, то вводится новое ограничение, обладающее уже указанным свойством, решается соответствующая задача ЛП-3 и т.д. Дополнительные ограничения, вводимые при переходе от одной задачи линейного программирования к другой, называются сечениями. Каждое из вновь вводимых ограничений отсекает некоторую часть множества допустимых решений исходной задачи ЛП-1. Сечения, вводимые указанным выше способом, вообще говоря, не обязательно должны привести к оптимальному решению задачи за конечное количество шагов. Однако при построении сечений по рассмотренным ниже правилам гарантируется сходимость к оптимальному решению задачи целочисленного линейного программирования, если только такое решение существует.

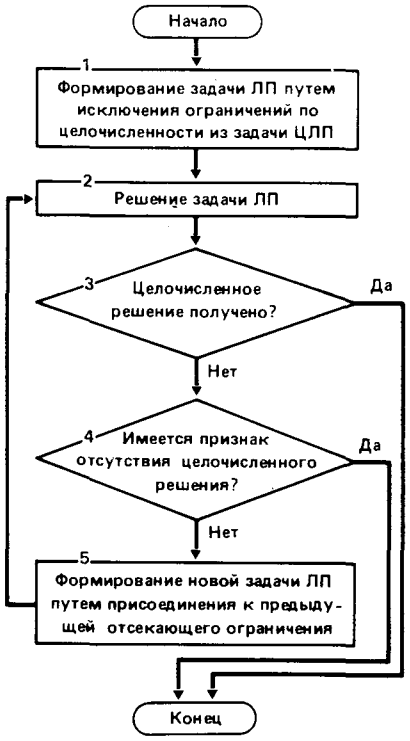


Рис. VII.2. Блок-схема алгоритма решения задач ЦЛП методом отсечений

Рассмотрим случай полностью целочисленной задачи линейного программирования, т.е. случай, когда в задаче ЦЛП (VII.1) множество J содержит все значения j.

Любое допустимое решение задачи ЛП-1 должно удовлетворять соотношению

$$\bar{x}_B = \bar{a}_0 - \sum_{j \in R} \bar{a}_j\, \bar{x}_j, \qquad \text{(VII.40)}$$

где $\bar{x}_B$ – вектор, компонентами которого являются базисные переменные в оптимальном решении задачи ЛП-1; $\bar{a}_0$ – вектор, компонентами

которого являются элементы столбца свободных членов таблицы с оптимальным решением задачи ЛП-1; j – индекс небазисной переменной; R – множество значений индексов небазисных переменных таблицы с оптимальным решением задачи ЛП-1; x_j – j-я небазисная переменная; a_j – вектор, компонентами которого являются элементы столбца, соответствующего j-й небазисной переменной в таблице с оптимальным решением задачи ЛП-1.

Существует единственное базисное решение $\bar{x}_B = \bar{a}_0$, $x_j = 0$ (для всех $j \in R$), удовлетвояющее (VII.40). Если это решение не удовлетворяет требованию целочисленности – все компоненты вектора $\bar{a}_0$ неотрицательные целые числа, то очевидно, в любом оптимальном решении задачи (VII.1) по крайней мере одна из небазисных переменных x_j должна быть положительной. Так как по условию все переменные целые, то необходимо, чтобы по крайней мере одна из небазисных переменных была бы не меньше 1

$$\sum_{j \in R} x_j \geqslant 1. \qquad \text{(VII.41)}$$

Полученное решение $x_j = 0$ (для всех $j \in R$) не удовлетворяет ограничению (VII.41). Следовательно, если сформировать новую задачу линейного программирования ЛП-2, присоединив ограничение (VII. 41) к задаче ЛП-1, то множество ее допустимых решений будет "меньше", чем соответствующее множество задачи ЛП-1, но тем не менее будет содержать все допустимые решения задачи ЦЛП (VII.1).

Предположим, что в результате решения задачи ЛП-2 получены целые значения всех переменных. Тогда это решение является оптимальным для задачи (VII.1). В противном случае для нового множества небазисных переменных вводится новое ограничение типа (VII.41) и процесс повторяется. Метод решения полностью целочисленных задач, базирующийся на введении сечений типа (VII.41), впервые был предложен Дж. Данцигом. Для небольшого количества задач, на которых этот метод испытывался, он оказался сходящимся. Было, однако, доказано, что существуют случаи, когда сходимости нет. С другой стороны, Р. Гомори разработал систематический способ построения сечений, который за конечное количество шагов приводит к оптимальному решению задачи (VII.1).

Рассмотрим метод Р. Гомори. Предположим, что в оптимальном решении задачи ЛП-1 s-я компонента вектора $\bar{a}_0$ не является целым числом. Соотношение (VII.40) для s-й компоненты можно представить следующим образом

$$x_{Bs} = a_{0s} - \sum_{j \in R} a_{sj} x_j. \qquad \text{(VII.42)}$$

Кроме того, представим a_{sj} и a_{s0} в виде

$$\left.\begin{aligned} a_{sj} &= [a_{sj}] + f_{sj}, \quad j \in R, \\ a_{s0} &= [a_{s0}] + f_{s0}, \end{aligned}\right\} \qquad \text{(VII.43)}$$

где $[a_{sj}]$ – наибольшее целое число, не превосходящее a_{sj} ($j \in R$ и $j = 0$); f_{sj} – положительная собственная дробная часть числа a_{sj} ($j \in R$ или $j = 0$).

Так как, по предположению, a_{s0} не целое, то $f_{s0} > 0$, а $f_{sj} \geqslant 0$ по определению. Осуществив подстановку (VII.43) в (VII.42), получим

$$x_{Bs} = [a_{0s}] - \sum_{j \in R} [a_{sj}] x_j + f_{s0} - \sum_{j \in R} f_{sj} x_j. \qquad \text{(VII.44)}$$

Для любого целочисленного решения, удовлетворяющего (VII.40), величина

$$x_{Bs} - [a_{0s}] + \sum_{j \in R} [a_{sj}] x_j \qquad \text{(VII.45)}$$

должна быть целой. Тогда, как следует из (VII.44), для любого целочисленного решения, удовлетворяющего (VII. 40), величина

$$f_{s0} - \sum_{j \in R} f_{sj} x_j \qquad \text{(VII.46)}$$

также должна быть целой. Величина $\sum_{j \in R} f_{sj} x_j$ не может быть отрицательной. Следовательно, в силу $0 < f_{s0} < 1$, величина (VII.46) не может быть положительным целым числом. Поэтому любое допустимое решение задачи ЦЛП (VII.1) должно удовлетворять неравенству

$$f_{s0} - \sum_{j \in R} f_{sj} x_j \leqslant 0,$$

или, что то же

$$\sum_{j \in R} f_{sj} x_j - f_{s0} \geqslant 0. \qquad \text{(VII.47)}$$

Очевидно, что оптимальное решение, полученное для задачи ЛП-1, не удовлетворяет (VII.47), так как $x_j = 0$ для всех $j \in R$. Таким образом, если сформулировать новую задачу линейного программирования ЛП-2, присоединив (VII.47) к задаче ЛП-1, то множество ее допустимых решений окажется "меньше" соответствующего множества задачи ЛП-1, но будет содержать все допустимые решения задачи ЦЛП (VII.1). Ограничение (VII.47) и представляет собой сечение, предложенное Р. Гомори.

Отметим, что если в (VII.42) все a_{sj} ($j \in R$) целые, то, по определению, $f_{sj} = 0$ для всех $j \in R$ и ограничение (VII.47) не может быть выполнено при любых значениях x_j ($j \in R$). Указанный случай является

признаком отсутствия допустимого целочисленного решения. В связи с этим надо помнить, что необходимым условием применения сечения (VII.47) при решении полностью целочисленных задач является целочисленность параметров исходной модели. Это объясняется тем, что наличие в ограничениях дробных коэффициентов приводит к нарушению целочисленности дополнительных переменных и позволяет обнаружить отсутствие допустимого решения даже тогда, когда задача, записанная только в основных переменных, имеет такое решение.

Рассмотрим случай частично целочисленной задачи. По-прежнему будем предполагать, что в оптимальном решении задачи ЛП-1 значение s-й базисной переменной не удовлетворяет требованию целочисленности. Однако теперь среди небазисных переменных x_j, $j \in R$ могут встречаться как целочисленные x_j, $j \in R \cap J$, так и нецелочисленные x_j, $j \in R^*$ (R^* подмножество R, не содержащееся в $R \cap J$). Кроме того, среди небазисных нецелочисленных переменных будем различать переменные x_j, $j \in R_+^*$ и x_j, $j \in R_-^*$ (R_+^*, R_-^* – подмножества R^*), для которых в (VII.42) $a_{sj} \geqslant 0$ и $a_{sj} < 0$ соответственно. С учетом указанных особенностей сечение, предложенное Р. Гомори для решения частично целочисленных задач, имеет вид

$$\sum_{j \in R} d_{sj} x_j - f_{s0} \geqslant 0, \qquad \text{(VII.48)}$$

где

$$d_{sj} = \begin{cases} a_{sj}, & j \in R_+^*, \\ \dfrac{f_{s0}}{1 - f_{s0}} |a_{sj}|, & j \in R_-^*, \\ f_{sj}, & j \in R \cap J,\ f_{sj} \leqslant f_{s0}, \\ \dfrac{f_{s0}}{1 - f_{s0}} (1 - f_{sj}), & j \in R \cap J,\ f_{sj} > f_{s0} \end{cases} \qquad \text{(VII.49)}$$

Вывод формулы (VII.48), а также доказательство сходимости алгоритма, реализующего метод отсечений по Р. Гомори, за конечное количество шагов можно найти, например, в монографии Дж. Хедли.

Приведем несколько рекомендаций, представляющих интерес с точки зрения практической реализации метода отсечений.

Во-первых, в процессе решения задачи ЦЛП (VII.1) может возникнуть ситуация, когда в оптимальном решении задачи линейного программирования требование целочисленности не выполняется для ряда базисных переменных. В этом случае, согласно разработанным эмпирическим правилам, сечение Р. Гомори целесообразно вводить на основе строки, соответствующей

$$\max_{i \in Q} \{ f_{i0} \} \quad \text{или} \quad \max_{i \in Q} \left\{ \frac{f_{i0}}{\sum\limits_{j \in R} f_{ij}} \right\}, \qquad \text{(VII.50)}$$

где Q – множество значений индекса i базисных целочисленных переменных, которым соответствуют нецелые значения a_{i0}.

Во-вторых, присоединив, например, к задаче ЛП-1 дополнительное ограничение–сечение типа (VII.47) или (VII.48), получаем задачу ЛП-2. При этом совсем не обязательно решать ее с начала. Достаточно сечение ввести в таблицу с оптимальным решением задачи ЛП-1 и затем найти оптимальное решение задачи ЛП-2.

В-третьих, если на каком-либо этапе в таблице с оптимальным нецелочисленным решением задачи линейного программирования среди базисных переменных появилась дополнительная, связанная с введенным ранее сечением, то в дальнейшем строку с данной переменной можно исключить из рассмотрения (так как нет необходимости, чтобы это ограничение всегда выполнялось). Поэтому очевидно, что максимальное количество дополнительных ограничений, которое может содержать любая задача линейного программирования в процессе поиска целочисленного решения, не может превышать $n + 1$, где n – количество основных переменных задачи ЦЛП (VII.1).

Завершая рассмотрение метода отсечений, заметим, что выше были рассмотрены лишь два типа сечений. Между тем, известен целый ряд способов построения сечений других типов. Однако ни один тип дополнительных ограничений по целочисленности не обеспечивает высокой эффективности соответствующих вычислительных процедур. Несмотря на то, что метод отсечений является надежным средством решения некоторых задач, имеющих специальную структуру, общая точка зрения специалистов-практиков такова – метод отсечений не подходит для решения задач ЦЛП большой размерности. Более того, известны случаи, когда изменение порядка следования ограничений в модели задачи приводило к резкому увеличению объема вычислений.

Пример VII.1. Используя метод отсечений, решим следующую задачу ЦЛП:

$$2x_1 + 3x_2 \to \max,$$
$$-x_1 + 11x_2 \leqslant 44,$$
$$3x_1 + x_2 \leqslant 21,$$
$$x_i \geqslant 0,\ x_i \text{ – целое},\ i = 1, 2.$$

Решим сначала соответствующую задачу ЛП, для чего введем в модель дополнительные переменные $y_i \geqslant 0$ ($i = 1, 2$) и составим исходную симплекс-таблицу (табл. VII.1). В результате решения (табл. VII.1 – VII.3) получаем нецелочисленное значение переменных $x_1 = 5\frac{1}{2}$, $x_2 = 4\frac{1}{2}$ (табл. VII.3, рис. VII.3). Так как исходная таблица содержит только целочисленные коэффициенты, то для построе-

Таблица VII.1

	$-x_1$	$-x_2$	1
$y_1 =$	−1	11	44
$y_2 =$	3	1	21
$Z =$	−2	−3	0

Таблица VII.2

	$-x_1$	$-y_1$	1
$x_2 =$	$-\frac{1}{11}$	$\frac{1}{11}$	4
$y_2 =$	$3\frac{1}{11}$	$-\frac{1}{11}$	17
$Z =$	$-2\frac{3}{11}$	$\frac{3}{11}$	12

Таблица VII.3

	$-y_2$	$-y_1$	1
$x_2 =$	$\frac{1}{34}$	$\frac{3}{34}$	$4\frac{1}{2}$
$x_1 =$	$\frac{11}{34}$	$-\frac{1}{34}$	$5\frac{1}{2}$
$Z =$	$\frac{25}{34}$	$\frac{7}{34}$	$24\frac{1}{2}$

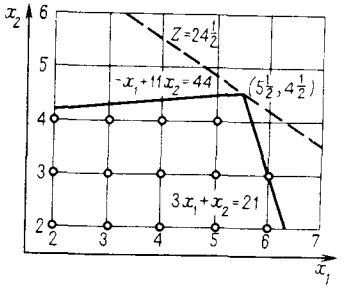


Рис. VII.3. Решение задачи примера VII.1 без учета требований целочисленности переменных

ния сечений можно выбрать правило (VII.47). В оптимальном решении положительные собственные дробные части значений переменных равны между собой, поэтому для определения строки, на основе которой в дальнейшем будет построено сечение, воспользуемся вторым из правил (VII.50):

$$\max\left\{-\frac{1}{2}\left(\frac{1}{34}+\frac{3}{34}\right)^{-1};\ \frac{1}{2}\left(\frac{11}{34}+\frac{33}{34}\right)^{-1}\right\}=\max\left\{4\frac{1}{4};\ \frac{17}{44}\right\}=$$

$$=4\frac{1}{4}.$$

Так как максимальное значение соответствует строке с переменной x_2, то сечение будет иметь вид

$$\frac{1}{34}y_2+\frac{3}{34}y_1-\frac{1}{2}\geqslant 0.$$

Введем дополнительно переменную $y_3 \geqslant 0$ и запишем сечение в симплекс-таблицу (табл. VII.4). В результате дальнейшего решения получаем $x_1 = 5\frac{2}{3}$, $x_2 = 4$ (табл. VII.5, рис. VII.4). На основе строки с переменной x_1 строим второе

Таблица VII.4

	$-y_2$	$-y_1$	1
$x_2 =$	$\frac{1}{34}$	$\frac{3}{34}$	$4\frac{1}{2}$
$x_1 =$	$\frac{11}{34}$	$-\frac{1}{34}$	$5\frac{1}{2}$
$y_3 =$	$-\frac{1}{34}$	$-\frac{3}{34}$	$-\frac{1}{2}$
$Z =$	$\frac{25}{34}$	$\frac{7}{34}$	$24\frac{1}{2}$

Таблица VII.5

	$-y_2$	$-y_3$	1
$x_2 =$	0	1	4
$x_1 =$	$\frac{1}{3}$	$-\frac{1}{3}$	$5\frac{2}{3}$
$y_1 =$	$\frac{1}{3}$	$-11\frac{1}{3}$	$5\frac{2}{3}$
$Z =$	$\frac{2}{3}$	$2\frac{1}{3}$	$23\frac{1}{3}$

сечение, которое будет иметь вид

$$\frac{1}{3}y_2+\frac{2}{3}y_3-\frac{2}{3}\geqslant 0.$$

Введем дополнительную переменную $y_4 \geqslant 0$, запишем сечение в симплекс-таблицу (табл. VII.6). Продолжим поиск решения (табл. VII.6 – VII.8). В результате получаем оптимальное целочисленное решение $x_1 = 5$, $x_2 = 4$, $Z = 22$ (табл. VII.8, рис. VII.5).

Пример VII.2. Используя метод отсечений, решим следующую задачу ЦЛП:

Таблица VII.6

	$-y_2$	$-y_1$	1
$x_2 =$	0	1	4
$x_1 =$	$\frac{1}{3}$	$-\frac{1}{3}$	$5\frac{2}{3}$
$y_1 =$	$\frac{1}{3}$	$-11\frac{1}{3}$	$5\frac{2}{3}$
$y_4 =$	$-\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{2}{3}$
$Z =$	$\frac{2}{3}$	$2\frac{1}{3}$	$23\frac{1}{3}$

Таблица VII.7

	$-y_2$	$-y_4$	1
$x_2 =$	$-\frac{1}{2}$	$1\frac{1}{2}$	3
$x_1 =$	$\frac{1}{2}$	$-\frac{1}{2}$	6
$y_1 =$	6	−17	17
$y_3 =$	$\frac{1}{2}$	$-1\frac{1}{2}$	1
$Z =$	$-\frac{1}{2}$	$3\frac{1}{2}$	21

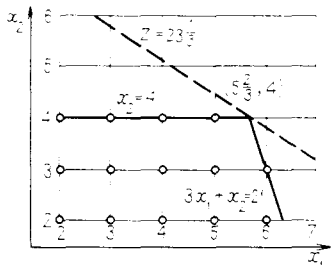


Рис. VII.4. Решение задачи примера VII.1 с дополнительным ограничением

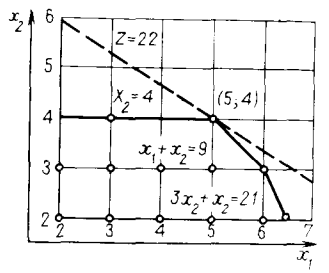


Рис. VII.5. Оптимальное целочисленное решение задачи примера VII. 1

$$3x_1 + 2x_2 \to \max,$$

$$x_1 + x_2 \leqslant 5{,}5,$$

$$0 \leqslant x_i \leqslant 4,\ x_i - \text{целое},\ i = 1, 2.$$

Решим сначала соответствующую задачу ЛП, для чего введем в модель дополнительные переменные $y_i \geqslant 0$ $(i = 1, 2, 3)$ и составим исходную симплекс-таблицу (табл. VII.9). В результате решения имеем $x_1 = 4$, $x_2 = 1\frac{1}{2}$ (табл. VII.10, промежуточные таблицы, в целях краткости изложения, не приводятся).

Так как исходная таблица содержит нецелочисленные коэффициенты, то для построения сечения воспользуемся правилом (VII.48). Причем будем считать y_1 целочисленной, а y_3 – нецелочисленной переменной. Основанием для этого является наличие целочисленных решений уравнения $y_1 = -x_1 + 4$ и, напротив, отсутствие таковых для уравнения $y_3 = 5\frac{1}{2} - x_1 - x_2$.

Сечение будет иметь вид $y_3 - \frac{1}{2} \geqslant 0$.

Введем дополнительную переменную $y_4 \geqslant 0$ и запишем сечение в симплекс-таблицу (табл. VII.11). Продолжив поиск решения, получим $x_1 = 4$, $x_2 = 1$, $Z = 14$ (табл. VII.12).

Таблица VII.8

	$-y_3$	$-y_4$	1
$x_2 =$	1	0	4
$x_1 =$	−1	1	5
$y_1 =$	−12	1	5
$y_2 =$	2	−3	2
$Z =$	1	2	22

Таблица VII.9

	$-x_1$	$-x_2$	1
$y_1 =$	1	0	4
$y_2 =$	0	1	4
$y_3 =$	1	1	$5\frac{1}{2}$
$Z =$	−3	−2	0

Таблица VII.10

	$-y_1$	$-y_2$	1
$x_1 =$	1	0	4
$y_2 =$	1	−1	$2\frac{1}{2}$
$x_2 =$	−1	1	$1\frac{1}{2}$
$Z =$	1	2	15

Таблица VII.11

	$-y_1$	$-y_3$	1
$x_1 =$	1	0	4
$y_2 =$	1	−1	$2\frac{1}{2}$
$x_2 =$	−1	1	$1\frac{1}{2}$
$y_4 =$	0	−1	$-\frac{1}{2}$
$Z =$	1	2	15

Таблица VII.12

	$-y_1$	$-y_4$	1
$x_1 =$	1	0	4
$y_2 =$	1	−1	3
$x_2 =$	−1	1	1
$y_3 =$	0	−1	$\frac{1}{2}$
$Z =$	1	2	14

Пример VII.3. Из бревен длиной 5 м для крепления выработок необходимо изготовить стойки длиной 1,4; 2,2; 2,8 м в количестве 12, 10 и 6 шт. соответственно. Определим минимальное количество бревен, необходимое для выполнения заказа.

Прежде всего составим варианты распила бревен. Как правило, составление вариантов – наиболее трудоемкий процесс при решении задач о раскрое материалов. В данном случае возможны пять вариантов распила:

1-й – изготовление трех элементов длиной 1,4 м;

2-й – изготовление двух элементов длиной 2,2 м;

3-й – изготовление одного элемента длиной 2,2 м и двух элементов длиной 1,4 м;

4-й – изготовление по одному элементу длиной 1.4 и 2,8 м;

5-й – изготовление по одному элементу длиной 2,2 и 2,8 м.

Теперь можно составить матрицу выхода готовых элементов при различных вариантах распила (табл. VII.13).

Таблица VII.13

Вариант распила	Выход элементов (шт.) длиной, м		
	1,4	2,2	2,8
1-й	3	0	0
2-й	0	2	0
3-й	2	1	0
4-й	1	0	1
5-й	0	1	1

После составления матрицы построим модель задачи. В качестве переменных примем x_i – количество бревен, распиливаемых по i-му варианту ($i = 1, 2, ..., 5$), а за критерий оптимальности – общее количество распиливаемых бревен.

Тогда целевая функция примет вид

$x_1 + x_2 + x_3 + x_4 + x_5 \to \min.$

Задача решается при выполнении ограничений по изготовлению требуемого количества бревен различной длины. Используя матрицу выхода готовой продукции при различных способах распила (см. табл. VII.13), составим ограничения:

по выпуску бревен длиной 1,4 м $3x_1 + 0x_2 + 2x_3 + 1x_4 + 0x_5 \geqslant 12$;

по выпуску бревен длиной 2,2 м $0x_1 + 2x_2 + 1x_3 + 0x_4 + 1x_5 \geqslant 10$;

по выпуску бревен длиной 2,8 м $0x_1 + 0x_2 + 0x_3 + 1x_4 + 1x_5 \geqslant 6$.

Кроме того, переменные x_i ($i = 1, 2, ..., 5$) – неотрицательные целые числа.

Таким образом, экономико-математическая модель задачи имеет вид

$x_1 + x_2 + x_3 + x_4 + x_5 \to \min,$

$3x_1 + 2x_3 + x_4 \geqslant 12, \quad 2x_2 + x_3 + x_5 \geqslant 10, \quad x_4 + x_5 \geqslant 6,$

$x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_3 \geqslant 0,\ x_4 \geqslant 0,\ x_5 \geqslant 0,$

x_1, x_2, x_3, x_4, x_5 – целые числа.

Таблица VII.14

	$-x_1$	$-x_2$	$-x_3$	$-x_4$	$-x_5$	1
$y_1 =$	−3	0	−2	−1	0	−12
$y_2 =$	0	−2	−1	0	−1	−10
$y_3 =$	0	0	0	−1	−1	−6
$Z =$	1	1	1	1	1	0

Таблица VII.15

	$-x_1$	$-x_2$	$-y_3$	$-y_1$	$-y_2$	1
$x_3 =$	1	$\frac{2}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$5\frac{1}{3}$
$x_5 =$	−1	$1\frac{1}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$4\frac{2}{3}$
$x_4 =$	1	$-1\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$1\frac{1}{3}$
$Z =$	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$-11\frac{1}{3}$

Таблица VII.16

	$-x_1$	$-x_2$	$-y_3$	$-y_1$	$-y_2$	1
$x_3 =$	1	$\frac{2}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$5\frac{1}{3}$
$x_5 =$	−1	$1\frac{1}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$4\frac{2}{3}$
$x_4 =$	1	$-1\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$1\frac{1}{3}$
$y_4 =$	0	$-\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$
$Z =$	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$-11\frac{1}{3}$

Таблица VII.17

	$-x_1$	$-x_2$	$-y_4$	$-y_1$	$-y_2$	1
$x_3 =$	1	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	5
$x_5 =$	−1	$1\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	5
$x_4 =$	1	−1	−1	0	1	2
$y_3 =$	0	$\frac{1}{2}$	$-1\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
$Z =$	0	0	1	0	0	−12

Поиск решения задачи иллюстрируется симплекс-таблицами: исходной (табл. VII.14), с оптимальным нецелочисленным решением (табл. VII.15), с введенным сечением по правилу (VII.47) (табл. VII.16) и с оптимальным целочисленным решением (табл. VII.17). Таким образом, для выполнения заказа необходимо 12 бревен, причем 5 из них следует распилить по 3-му варианту, 2 – по 4-му и 5 – по 5-му.

§ 8. РЕШЕНИЕ ЦЕЛОЧИСЛЕННЫХ ЗАДАЧ МЕТОДОМ ВЕТВЕЙ И ГРАНИЦ

Метод ветвей и границ впервые был предложен для решения линейных целочисленных задач А. Лэндом и А. Дойгом и получил название "алгоритм Лэнда и Дойга". Рассмотрим указанный алгоритм (рис. VII.6).

Допустим, имеется задача ЦЛП (VII.1), предусматривающая максимизацию целевой функции. В соответствии с алгоритмом сначала необходимо решить задачу ЛП-1, полученную из задачи ЦЛП (VII.1) путем исключения требования целочисленности. Предположим, что в оптимальном решении задачи ЛП-1 некоторые целочисленные переменные принимают дробные значения, а значение целевой функции равно Z_1 (рис. VII.7). Следовательно, оптимальное решение задачи ЛП-1 не является оптимальным решением задачи ЦЛП (VII.1). Величина Z_1

представляет собой верхнюю границу оптимального значения Z исходной задачи ЦЛП (VII.1), так как при выполнении требования целочисленности значение Z может лишь уменьшиться. Затем производится ветвление по одной из целочисленных переменных, имеющих дробное значение в оптимальном решении задачи ЛП-1. Выбор переменной, по которой производится ветвление, осуществляется с использованием эмпирических правил. Рассмотрим некоторые из них:

1. Выбирается переменная, имеющая значение собственной дробной части, наиболее близкое к 0,5.

2. Выбирается переменная с наибольшим приоритетом. Приоритет переменной может определяться, например, следующими соображениями:

важностью определения значения переменной с точки зрения решения, принимаемого в рамках рассматриваемой модели;

величиной ее коэффициента стоимости или прибыли в целевой функции.

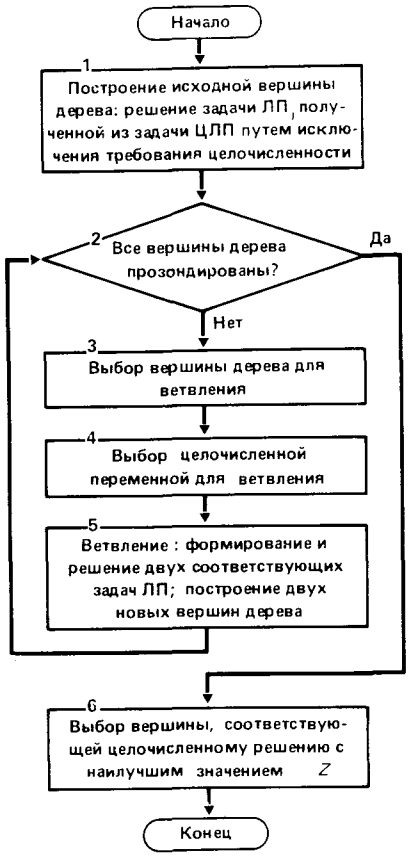


Рис. VII.6. Блок-схема алгоритма А. Лэнда и А. Дойга

3. Переменная выбирается произвольным образом.

Допустим, что ветвление происходит по целочисленной переменной x_s, дробное значение которой в оптимальном решении задачи ЛП-1 равно a_s. Тогда далее формируются две новые задачи ЛП-2 и ЛП-3. Они получаются путем введения в задачу ЛП-1 ограничений $x_s \leqslant [a_s]$ и $x_s \geqslant [a_s] + 1$ соответственно.

Предположим, что оптимальные решения задач ЛП-2 и ЛП-3 также содержат дробные значения целочисленных переменных и поэтому не являются допустимыми решениями задачи ЦЛП (VII.1). Далее необходимо выбрать задачу ЛП-2 или ЛП-3 и произвести ветвление в соответствующей вершине вводя новые ограничения. Выбор вершины (задачи ЛП), как и выбор переменной, осуществляется с использованием

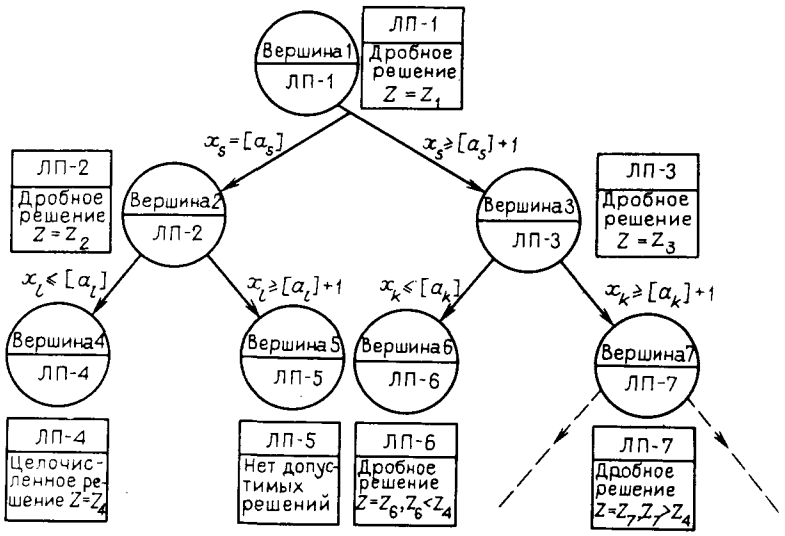


Рис. VII.7. Дерево, иллюстрирующее решение задачи ЦЛП методом ветвей и границ (вершины 4 и 5 прозондированы явным образом, 6 – неявным)

специальных эмпирических правил. Рассмотрим некоторые из них.

1. Выбирается вершина, которой соответствует задача ЛП с наибольшим значением целевой функции в оптимальном решении.

2. Вершина выбирается произвольным образом.

После определения вершины для дальнейшего ветвления выбирается целочисленная переменная, которая имеет в оптимальном решении соответствующей задачи ЛП дробное значение, и производится ветвление по этой переменной. Процесс ветвления и решения задач ЛП продолжается до получения целочисленного оптимального решения одной из задач ЛП.

Значение Z в этом решении представляет собой нижнюю границу целевой функции в оптимальном решении исходной задачи ЦЛП (VII.1). На этом этапе фиксируются все вершины (задачи ЛП), для которых значение Z в оптимальном решении не превосходит полученной нижней границы. Эти вершины называются прозондированными, поскольку области допустимых решений соответствующих им задач ЛП не содержат решений лучших, чем полученное. Естественно, что в дальнейшем ветвлении они не участвуют.

В качестве иллюстрации рассмотрим ситуацию, изображенную на рис. VII.7. В вершине 4 получено значение Z_4 нижней границы. Задача ЛП-5, соответствующая вершине 5, не имеет допустимых решений. Следовательно, вершины 4 и 5 прозондированы. Кроме того, так как значение Z_6 целевой функции в оптимальном решении задачи ЛП-6 меньше Z_4, вершина 6 является также прозондированной. Дальнейшее

ветвление следует производить в вершине 7, поскольку в области допустимых решений задачи ЛП-7 может существовать целочисленное решение, для которого $Z_4 < Z < Z_7$.

Таким образом, промежуточная вершина (задача ЛП) является прозондированной (явным или неявным образом) в том случае, если она удовлетворяет хотя бы одному из следующих условий.

1. Соответствующее ей оптимальное решение целочисленно и, следовательно, допустимо для исходной задачи ЦЛП (VII.1).

2. Задача ЛП, соответствующая рассматриваемой вершине, не имеет допустимых решений.

3. Значение Z в оптимальном решении соответствующей задачи ЛП не превосходит текущей нижней границы.

При использовании алгоритма А. Лэнда и А. Дойга выбор вершин для дальнейшего ветвления осуществляется до тех пор, пока остается хотя бы одна непрозондированная вершина.

Прозондированная вершина с наилучшим значением Z дает оптимальное решение исходной задачи ЦЛП (VII.1). Следовательно, эффективность алгоритма существенно зависит от скорости последовательного зондирования вершин. Обычно для выполнения условий 1 и 2 требуется значительное время. Условие 3 нельзя применять до получения нижней границы. Однако нижняя граница определяется только после просматривания вершины с допустимым (целочисленным) решением исходной задачи, т.е. вершины, удовлетворяющей условию 1. Поэтому получение перед реализацией алгоритма целочисленного решения исходной задачи может оказаться весьма полезным. Такое решение дает начальную нижнюю границу, которая используется до получения лучшей нижней границы в процессе ветвления. Для многих практических задач достигнутые показатели функционирования оптимизируемой системы можно использовать в качестве начального решения.

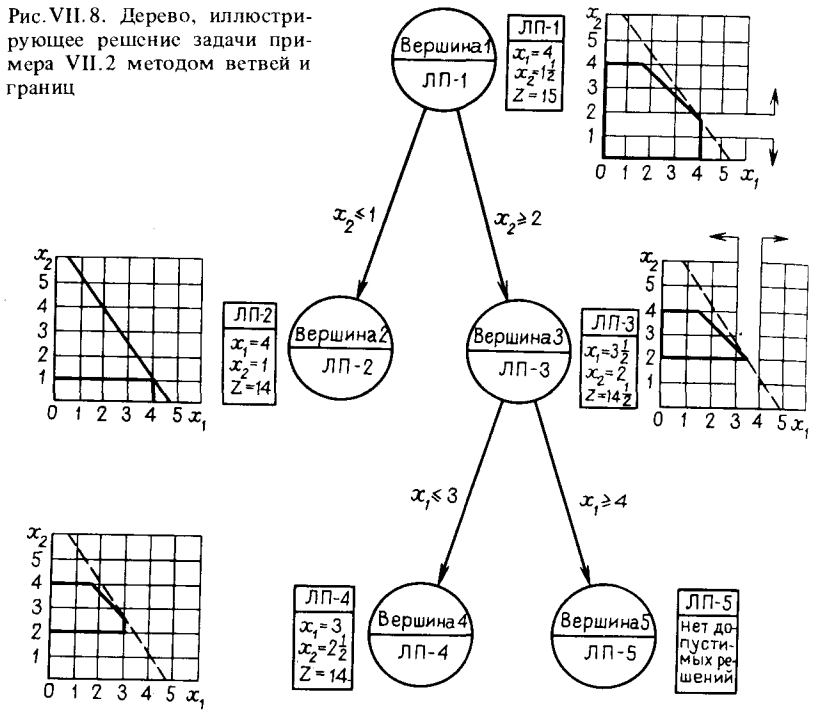
Основной недостаток рассмотренного выше алгоритма заключается в том, что каждая из соответствующих вершинам дерева задач ЛП должна решаться полностью. Для задач большой размерности это требует значительных затрат времени работы ЭВМ. Кроме того, по сравнению с алгоритмом, реализующим метод отсечений, как правило, требуется значительно больший объем машинной памяти. Однако численно алгоритм А. Лэнда и А. Дойга устойчивее. Поэтому несмотря на указанный недостаток, в настоящее время модификации алгоритма, реализующего метод ветвей и границ, являются наиболее надежным средством решения целочисленных линейных задач.

Пример VII.4. Используя метод ветвей и границ, решим задачу, рассмотренную в примере VII.2.

Оптимальное решение задачи ЛП-1 уже найдено в процессе разбора примера VII.2 (см. табл. VII.10). Поэтому произведем ветвление в вершине 1 по переменной x_2 (рис. VII.8).

Для решения последующих задач ЛП воспользуемся графическим методом. В вершине 2 определена нижняя граница значения целевой функции в оптимальном

Рис. VII.8. Дерево, иллюстрирующее решение задачи примера VII.2 методом ветвей и границ



решении исходной задачи. Вершина 2 прозондирована. Поскольку значение целевой функции в оптимальном решении задачи ЛП-3 больше нижней границы, то продолжаем процесс ветвления по переменной x_1. В результате получаем задачи ЛП-4 и ЛП-5.

Соответствующие им вершины являются прозондированными, так как значение целевой функции в оптимальном решении задачи ЛП-4 не превышает нижней границы, а задача ЛП-5 не имеет допустимых решений. Таким образом, все вершины прозондированы, целочисленное решение ($x_1 = 4$, $x_2 = 1$, $Z = 14$) соответствует только вершине 2, следовательно, оно и является оптимальным решением исходной задачи ЦЛП.

Таблица VII.18

Договор на выполнение НИР	Затраты (тыс. руб.) на выполнение НИР по годам			Ожидаемый экономический эффект от внедрения результатов НИР, тыс. руб.
	1-й	2-й	3-й	
1	60	40	90	350
2	50	100	120	600
3	60	50	30	350
4	80	120	60	450
5	70	60	120	500

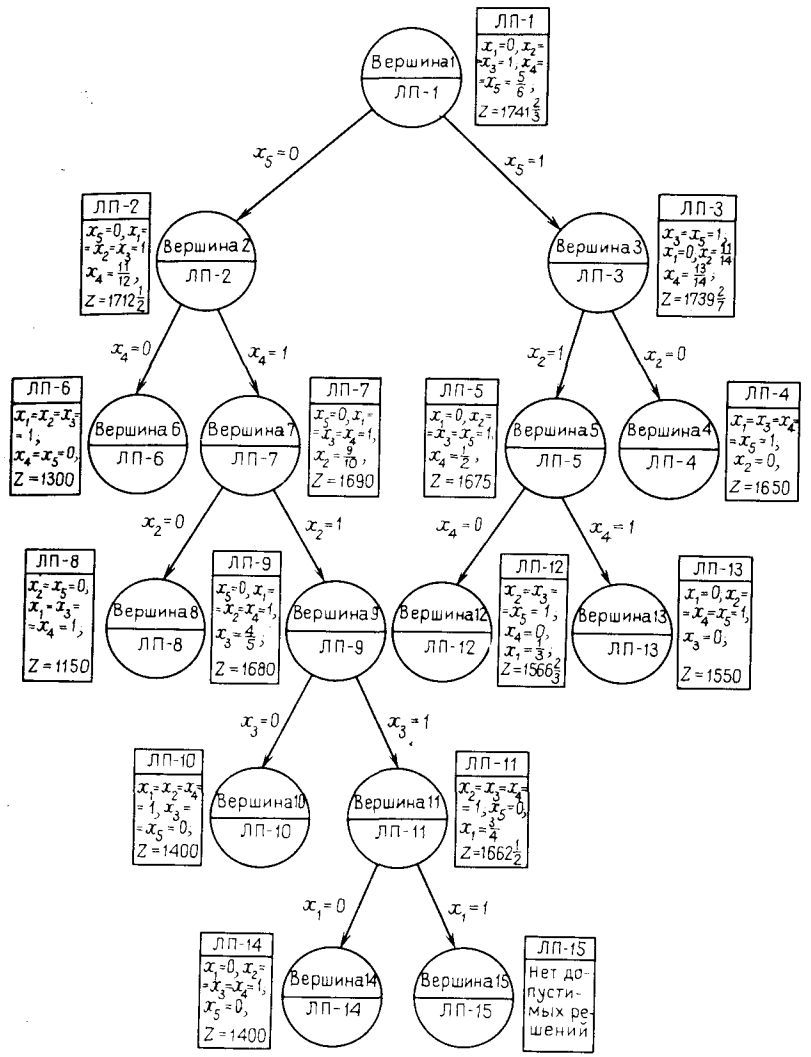


Рис. VII.9. Дерево, иллюстрирующее решение задачи о заключении договоров методом ветвей и границ

Пример VII.5. Объем финансирования научно-исследовательских работ (НИР) на горном предприятии составляет 300 тыс.руб. в год. Имеется пять предложений по заключению со сторонними организациями различных (не совпадающих по тематике) трехлетних договоров на выполнение этих работ. По каждому из дого-

воров известны годовые затраты на проведение работ и ожидаемый экономический эффект внедрения их результатов (табл. VII.18). Требуется выбрать совокупность договоров, обеспечивающих максимальный суммарный экономический эффект от внедрения на предприятии результатов НИР.

Составим экономико-математическую модель задачи.

В модели используем булевы переменные x_i, $i = 1, 2, ..., 5$ ($x_i = 1$, если i-й договор заключается, и $x_i = 0$ в противном случае). Критерий решения известен по условиям задачи. Тогда целевая функция будет иметь вид $Z = 350x_1 + 600x_2 + 350x_3 + 450x_4 + 500x_5 \to \max$.

Так как затраты на проведение НИР по всем договорам не могут превышать установленного объема годового финансирования, то соответствующие ограничения имеют вид

$$60x_1 + 50x_2 + 60x_3 + 80x_4 + 70x_5 \leqslant 300,$$

$$40x_1 + 100x_2 + 50x_3 + 120x_4 + 60x_5 \leqslant 300,$$

$$90x_1 + 120x_2 + 30x_3 + 60x_4 + 120x_5 \leqslant 300.$$

Кроме того, переменные модели, по определению, должны отвечать требованиям: $0 \leqslant x_i \leqslant 1$, x_i – целое, $i = 1, 2, ..., 5$. Определим оптимальные целочисленные значения переменных модели методом ветвей и границ. Решение задачи ЛП-1 (процесс поиска решения данной задачи ЛП, а также других в примере опущен) содержит дробные значения переменных $x_4 = x_5 = \frac{5}{6}$ (рис. VII.9). Учитывая, что коэффициент при x_5 в целевой функции модели больше, произведем ветвление по данной переменной. В силу специфики булевых переменных задачи ЛП-2 и ЛП-3 формируются путем присвоения переменной x_5 в задаче ЛП-1 значения 0 и 1 соответственно. Решения указанных задач содержат дробные значения переменных. Для дальнейшего ветвления выбираем вершину 3, так как ей соответствует лучшее значение целевой функции. В результате ветвления по переменной x_2 получаем вершины 4 и 5. В вершине 4 найдено целочисленное решение и, следовательно, определена нижняя граница значения – целевой функции в оптимальном решении исходной задачи. Вершина 4 прозондирована. Сравнение значений целевых функций в вершинах 2 и 5 определяет направление дальнейшего ветвления (вершина 5). Ветвление производим по переменной x_4. Новые вершины 6 и 7 являются прозондированными, поскольку значение Z в этих вершинах не превышает нижней границы. Дальнейшее ветвление осуществляем в вершине 2. В результате реализации метода в конечном итоге имеем прозондированные вершины 4, 6, 7, 8, 10, 12, 14 и 15. Среди них вершины 4, 7, 8, 10, 12 и 14 содержат целочисленное решение. Лучшее из них соответствует вершине 4 ($x_1 = x_3 = x_4 = x_5 = 1$, $x_2 = 0$, $Z = 1650$).

Таким образом, предприятию следует заключить договора 1, 3, 4, 5; при этом ожидаемый эффект от внедрения результатов НИР составит 1650 тыс.руб.

§ 9. ЧАСТИЧНЫЙ ПЕРЕБОР В ЗАДАЧАХ С БУЛЕВЫМИ ПЕРЕМЕННЫМИ

Как было показано (§ 1), за счет подстановки (VII.2) или (VII.3) любую целочисленную задачу можно записать в булевых переменных, использование которых упрощает вычислительные процедуры, так как каждая переменная принимает всего два значения. Данное обстоятельство учтено в рассматриваемом ниже аддитивном алгоритме решения задач ЦЛП с булевыми переменными, который базируется на идеях

более общего метода ветвей и границ, но не требует решения задач линейного программирования. Вычислительные операции ограничиваются лишь сложением и вычитанием, что и объясняет название алгоритма. Его основная идея заключается в переборе всех возможных решений целочисленной задачи. Процедура перебора осуществляется специальным образом, что позволяет отвергать некоторые решения без непосредственной проверки.

Для того чтобы можно было использовать аддитивный алгоритм, исходная задача должна быть представлена следующим образом

$$\left.\begin{aligned} &Z = \sum_{j=1}^{n} c_j x_j \to \min, \\ &\sum_{j=1}^{n} a_{ij} x_j + y_i = B_i, \quad i = 1, 2, ..., m, \\ &x_j = 0 \text{ или } 1, \ j = 1, 2, ..., n, \\ &y_i \geqslant 0, \ i = 1, 2, ..., m, \end{aligned}\right\} \qquad \text{(VII.51)}$$

где c_j – неотрицательный коэффициент при j-й переменной в целевой функции; y_i – дополнительная переменная, ассоциированная с i-м ограничением.

Заметим, что согласно (VII.51) исходная задача, записанная только в основных переменных, не должна содержать точных равенств. Если некоторые значения c_j в исходной задаче отрицательные, то ее можно привести к требуемому виду путем замены соответствующих переменных; $x_j = 1 - x_j'$, где x_j' – булева переменная.

Допустим, что в задаче VII.51 все $b_i \geqslant 0$. Тогда искомый минимум находится путем приравнивания к нулю всех основных переменных. В противном случае применяется аддитивный алгоритм.

В основу алгоритма положена итеративная процедура выбора основных переменных, которым надлежит присвоить значения 1 с целью обеспечения допустимости решения: $y_i \geqslant 0$ для всех i. На начальной итерации все $x_j = 0$ и, следовательно, для некоторых i $y_i < 0$. Ветвление (выбор переменных) и зондирование вершин осуществляются с помощью специальных тестов, в которых используются понятия "свободная переменная" и "частичное решение".

Булева переменная называется свободной в некоторой вершине дерева, если ее значение не фиксируется ни на одной из ветвей, ведущих в эту вершину. Обычно предполагается, что значения свободных переменных равны 0, но могут быть увеличены до 1, если необходимо, в целях приближения к допустимому решению.

Частичное решение представляет собой набор булевых переменных, в котором значение одной или нескольких переменных фиксированы и равны нулю или единице. Удобным для применения алгоритма является

представление частичного решения в виде упорядоченного множества. Допустим J_t – частичное решение, рассматриваемое в t-й вершине (или на t-й итерации), а запись $+j$ ($-j$) означает, что $x_j = 1$ ($x_j = 0$). Тогда элементами множества J_t являются значения индексов фиксированных переменных, взятые с плюсом (минусом), если значение соответствующей переменной равно нулю (единице). Множество упорядочено в том смысле, что его расширение производится путем приписывания каждого нового элемента справа от уже имеющихся элементов.

Каждой вершине дерева, порожденного исходной задачей, соответствует определенное частичное решение, так как при движении по ветвям, приводящим в данную вершину, фиксируются значения некоторых булевых переменных (рис. VII.10).

Частичное решение считается прозондированным, если оно не может привести к уменьшению значения целевой функции и получению допустимого решения. Если некоторое частичное решение прозондировано, то, следовательно, нецелесообразно осуществлять ветвление в соответствующей вершине, так как все вновь полученные решения либо не окажутся допустимыми, либо не обеспечат уменьшение целевой функции.

Использование частичных решений при реализации аддитивного алгоритма позволяет обходиться без представления совокупности порожденных задач в виде дерева. Это означает, что каждое последующее

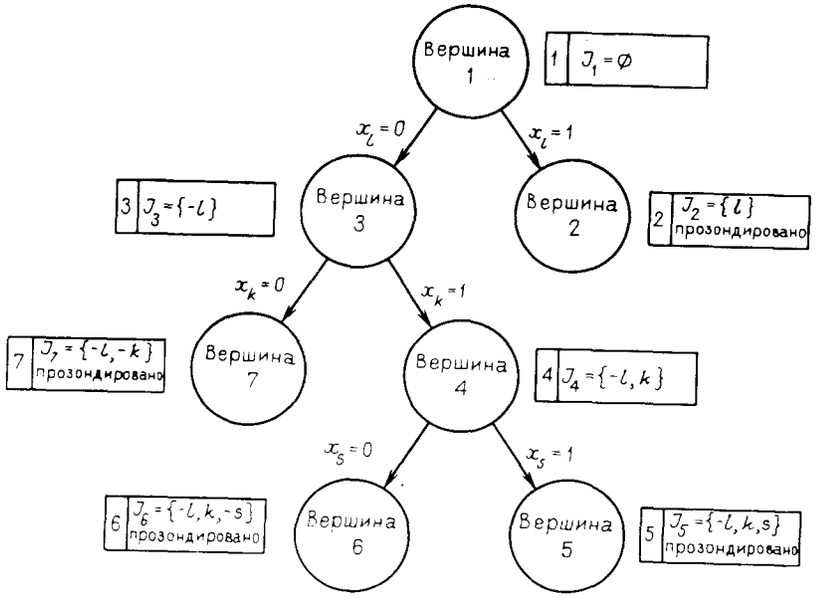


Рис. VII.10. Частичные решения, ассоциированные с вершинами дерева при решении задачи ЦЛП с булевыми переменными

решение может быть естественным образом получено из предыдущего. Возможны два случая.

Первый, когда построение частичного решения осуществляется путем введения новой переменной (а это всегда связано с присвоением ей единичного значения). В этом случае частичное решение строится за счет расширения предыдущего, т.е. приписывания нового элемента справа от уже имеющихся. Таким образом, были получены частичные решения J_2, J_4 и J_5 на основе J_1, J_3 и J_4 соответственно (см. рис. VII.10).

Второй, когда построение частичного решения осуществляется на основе прозондированного. В этом случае пользуются следующим правилом. Если все элементы прозондированного частичного решения отрицательные, то перебор заканчивается. В противном случае следует изменить знак крайнего правого положительного элемента и затем отбросить все (отрицательные) элементы, расположенные справа от него. Таким образом, были получены частичные решения J_3, J_6 и J_7 на основе прозондированных J_2, J_5 и J_6 соответственно (см. рис. VII.10). Заметим, что наличие отрицательного элемента в частичном решении означает, что предшествующее решение (в котором указанный элемент был положительным) прозондировано. Таким образом, если все элементы прозондированного частичного решения отрицательные, то соответствующие переменные рассмотрены как при нулевых, так и единичных значениях. В результате новые ветви не появляются и процесс перебора следует считать завершенным.

Рассмотрим теперь специальные тесты, с помощью которых осуществляются зондирование и построение новых частичных решений при реализации аддитивного алгоритма (рис. VII.11).

Допустим имеется задача ЦЛП с булевыми переменными VII.51. Пусть J_t – частичное решение, соответствующее вершине t (полученное на итерации t; на начальной итерации $J_1 \equiv \phi$, следовательно, все переменные являются свободными), Z_t – значение Z на итерации t. Через $\bar{Z}$ обозначим текущее значение верхней границы (на начальной итерации $\bar{Z} = \infty$).

Тест 1. Если x_s – свободная переменная и неравенство

$$a_{is} \geqslant 0 \qquad \text{(VII.52)}$$

выполняется при всех значениях i, для которых $y_i^t < 0$, то присвоение x_s значения 1 не может привести к получению допустимого решения, и значение x_s должно остаться равным нулю.

Тест 2. Если x_s – свободная переменная и выполняется неравенство

$$c_s + Z_t \geqslant \bar{Z}, \qquad \text{(VII.53)}$$

то присвоение x_s значения 1 не может уменьшить значения целевой функции, и значение s_x должно остаться равным нулю.

В результате применения тестов 1 и 2 определяются свободные переменные, присвоение которым значения 1 не приведет ни к получению допустимого решения, ни к улучшению значения Z соответственно.

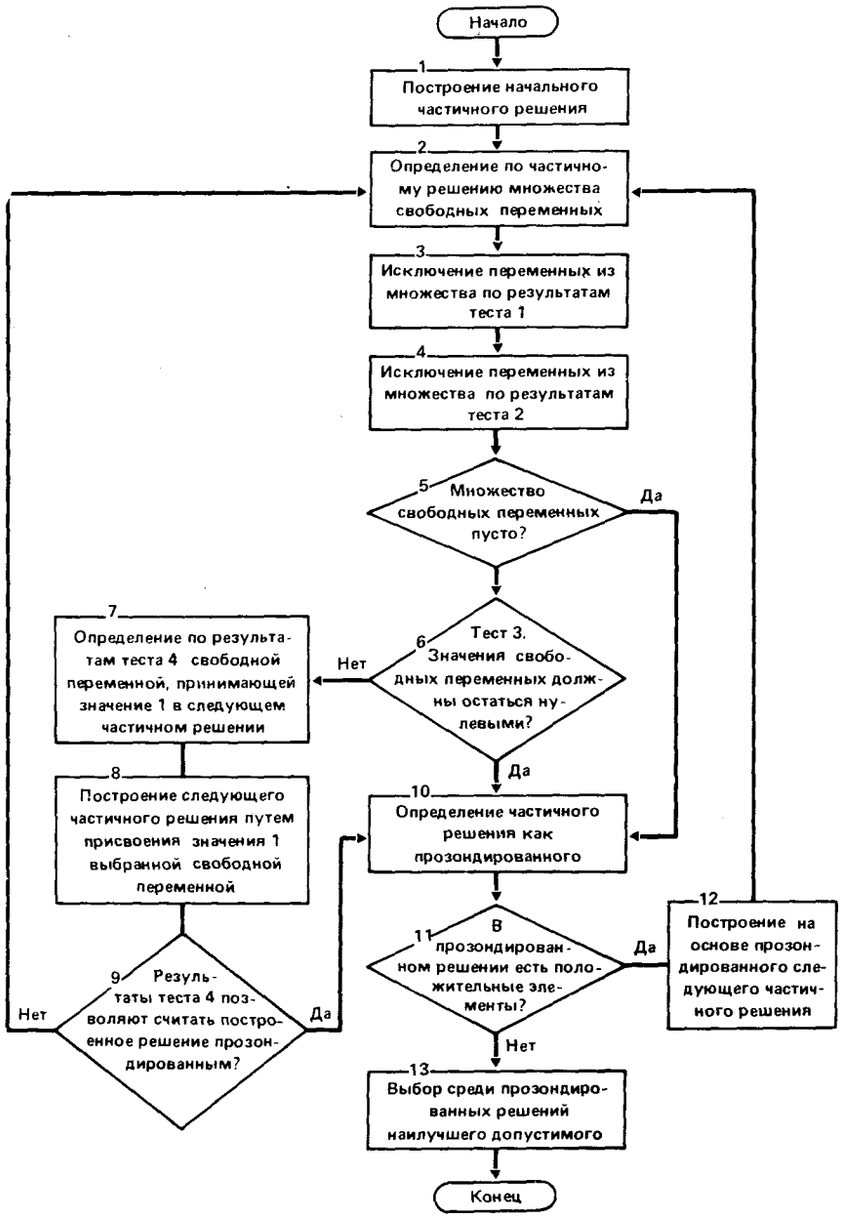


Рис. VII.11. Блок-схема аддитивного алгоритма решения задач ЦЛП с булевыми переменными

Обозначим через N_t множество свободных переменных, для которых не выполняются на итерации t условия, сформулированные в указанных тестах. Если в результате применения теста (ов) 1 или (и) 2 $N_t = \phi$, то J_t прозондировано. В противном случае применяют следующий тест.

Тест 3. Рассмотрим i-е ограничение $\sum_{j=1}^{n} a_{ij} x_j + y_i = B_i$, которому соответствует $y_i^t < 0$. Если по крайней мере для одного значения $y_i^t < 0$ выполняется неравенство

$$\sum_{j \in N_t} \min\left\{0, a_{ij}\right\} > y_i^t, \tag{VII.54}$$

то значения всех переменных из множества N_t должны остаться равными нулю, так как никакой допустимый набор их фиксированных значений не приводит к получению допустимого решения.

Если условие, сформулированное в тесте 3, выполняется, то J_t прозондировано. В противном случае определяется переменная, инициирующая процесс ветвления.

Тест 4. Если

$$U_k^t = \max_{j \in N_t}\left\{U_j^t\right\} = \max_{j \in N_t}\left\{\sum_{i=1}^{m} \min\left\{0, y_i^t - a_{ij}\right\}\right\}, \tag{VII.55}$$

то переменной x_k следует присвоить значение 1. Величина U_j^t является оценкой меры общей недопустимости присвоения значения 1 свободной переменной x_j на итерации t.

В случае $U_k^t = 0$ объединение условия $x_k = 1$ с частичным решением J_t обеспечивает получение допустимого решения с меньшим, чем Z значением целевой функции. При этом частичное решение J_{t+1}, построенное путем приписывания к J_t справа элемента k, можно считать прозондированным. В противном случае необходимо провести проверку J_{t+1} с помощью описанных выше тестов.

Если прозондированное частичное решение содержит только отрицательные элементы, то процедура перебора завершается. Успешная реализация аддитивного алгоритма в значительной степени зависит от качества тестовых процедур. Лучшие из принятых на практике программ для ЭВМ основываются на различных уточненных процедурах, что позволяет уменьшить затраты машинного времени. Однако с ростом количества переменных задачи продолжительность вычислений увеличивается почти экспоненциально. Это существенно ограничивает количество переменных в задачах, решаемых с использованием аддитивного алгоритма (при разумных затратах машинного времени – до 100 переменных).

Необходимо отметить, что существует зависимость между продолжительностью вычислений и конкретным видом постановки задачи. Порядок, в котором рассматриваются переменные и ограничения, в ряде

случаев оказывает непосредственное влияние на эффективность алгоритма. Предпочтительными являются упорядочение ограничений по мере убывания их "жесткости" и ранжирование переменных в соответствии с порядком увеличения неотрицательных коэффициентов целевой функции. Упорядочение рассмотрения переменных и ограничений ускоряет процедуру зондирования частичных решений.

В случае решения задачи большой размерности предпочтение, как правило, отдается методу ветвей и границ.

Пример VII.6. Используя аддитивный алгоритм, решим задачу о финансировании научно-исследовательских работ на горном предприятии, сформулированную в примере VII.5.

Сначала представим модель задачи в виде (VII.51)

$$Z' = 350\,x_1' + 600\,x_2' + 350\,x_3' + 450\,x_4' + 500\,x_5' \to \min,$$

$$-60\,x_1' + 50\,x_2' - 60\,x_3' - 80\,x_4' - 70\,x_5' + y_1 = -20,$$

$$-40\,x_1' - 100\,x_2' - 50\,x_3' - 120\,x_4' - 60\,x_5' + y_2 = -70,$$

$$-90\,x_1' - 120\,x_2' - 30\,x_3' - 60\,x_4' - 120\,x_5' + y_3 = -120,$$

$y_i \geqslant 0,\ i = 1, 2, 3,$

$x_j' = 0$ или $1,\ j = 1, 2, \ldots, 5,$

где $x_j' = 1 - x_j,\ j = 1, 2, \ldots, 5,$

$Z' = -Z + 2250.$

Затем осуществим процедуру частичного перебора в соответствии с аддитивным алгоритмом.

Итерация 1.

$J_1 = \phi,\ \bar{Z} = \infty,\ (y_1', y_2', y_3') = (-20, -70, -120),\ Z_1 = 0.$

По результатам первых двух тестов не выявлены переменные, значения которых должны остаться равными нулю. Тест 3 свидетельствует, что множество $N_1 = \{1, 2, 3, 4, 5\}$ нельзя исключить из рассмотрения, так как

$y_1 - 60 - 50 - 60 - 80 - 70 < -20,$

$y_2 - 40 - 100 - 50 - 120 - 60 < -70,$

$y_3 - 90 - 120 - 30 - 60 - 120 < -120.$

Применяя тест 4, получим

$U_1' = 0 - 30 - 30 = -60,$

$U_2' = 0 + 0 + 0 = 0,$

$U_3' = 0 - 20 - 90 = -110,$

$U_4' = 0 + 0 - 60 = -60,$

$U_5' = 0 - 10 + 0 = -10.$

Следовательно, $K = 2$.

Итерация 2.

$J_2 = \{2\},\ \bar{Z} = 600,\ (y_1^2, y_2^2, y_3^2) = (-20 + 50, -70 + 100, -120 + 120) = (30, 30, 0),\ Z_2 = 600.$

Так как получено допустимое решение, то $\bar{Z} = Z_2 = 600$. Частичное решение J_2 прозондировано.

Итерация 3.

$J_3 = \{-2\},\ \bar{Z} = 600,\ (y_1^3, y_2^3, y_3^3) = (-20, -70, -120),\ Z_3 = 0.$

По результатам первых двух тестов не выявлены переменные, значения которых должны остаться равными нулю. Тест 3 свидетельствует, что множество $N_3 = \{1, 2, 3, 4, 5\}$ нельзя исключить из рассмотрения. По результатам теста 4 получаем

$U_1^3 = -60,\ U_3^3 = -110,\ U_4^3 = -60,\ U_5^3 = \boxed{-10.}$

Следовательно, $K = 5$

Итерация 4.

$J_4 = \{-2,5\},\ \bar{Z} = 600,\ (y_1^4, y_2^4, y_3^4) = (-20 + 70, -70 + 60, -120 + 120) = (50, -10, 0),\ Z_4 = 500.$

Тест 1 не исключает из рассмотрения ни одну из свободных переменных. Тест 2 исключает из рассмотрения переменные 1, 3 и 4. Таким образом, $N_4 = \phi$, следовательно J_4 прозондировано.

Итерация 5.

$J_5 = \{-2, -5\},\ \bar{Z} = 600,\ (y_1^5, y_2^5, y_3^5) = (-20, -70, -120)),\ Z_5 = 0.$

После применения первых трех тестов имеем $N_5 = \{1, 3, 4\}$.
По результатам теста 4 получаем: $U_1^5 = \boxed{-60,}\ U_3^5 = -110,\ U_4^5 = \boxed{-60,}$
Выбираем первую переменную $(K = 1)$.

Итерация 6.

$J_6 = \{-2, -5, 1\},\ \bar{Z} = 600,\ (y_1^6, y_2^6, y_3^6) = (-20, +60, -70 + 40, -120 + 90) = (40, -30, -30),\ Z_6 = 350.$

Тест 2 исключает из рассмотрения переменные 3 и 4 $(N_6 = \phi)$. Частичное решение J_6 прозондировано.

Итерация 7.

$J_7 = \{-2, -5, -1\},\ \bar{Z} = 600,\ (y_1^7, y_2^7, y_3^7) = (-20, -70, -120),\ Z_7 = 0.$

Тест 3 исключает из рассмотрения множество $N_7 = \{3, 4\}$, так как $y_3 - 30 - 60 > -120$.
Следовательно J_7 прозондировано. Все элементы J_7 отрицательные – процедура перебора завершена. Частичное решение J_2 определяет оптимальное решение задачи: $x_2' = 1,\ x_1 = x_3 = x_4 = x_5 = 0,\ Z' = 600$.

С учетом преобразований, выполненных в начале примера, решение задачи в исходных переменных будет иметь вид $x_1 = x_3 = x_4 = x_5 = 1,\ x_2 = 0,\ Z = 1650$ тыс. руб.

Заметим, что в ходе решения задачи был осуществлен непосредственный просмотр только 7 из $2^5 = 32$ возможных решений.

1. Для каких задач планирования горного производства характерно условие целочисленности решения?

2. Назовите методы решения задач ЦЛП и укажите их характерные особенности, достоинства и недостатки.

3. Объясните, при выполнении каких условий округление значений непрерывных переменных является удовлетворительным способом получения целочисленного решения. Приведите примеры, когда этот метод неприемлем, и объясните почему.

4. В чем заключается практическая ценность использования булевых переменных при построении моделей целочисленных задач планирования?

5. Какова принципиальная последовательность действий при решении задач ЦЛП методом отсечений?

6. В чем заключается необходимое условие применения сечения Р. Гомори, предназначенного для решения полностью целочисленных задач? Что является признаком отсутствия допустимого целочисленного решения?

7. Каким образом следует выбрать строку, на основе которой строится сечение Р. Гомори?

8. Объясните, почему максимальное количество дополнительных ограничений (сечений), которое может содержать любая задача ЛП в процессе поиска целочисленного решения, не может превышать $n + 1$, где n – количество основных переменных исходной задачи ЦЛП.

9. Какова принципиальная последовательность действий при решении задач ЦЛП методом ветвей и границ?

10. Объясните значения понятий "ветвление" и "граница", используемых при описании метода ветвей и границ.

11. Каким образом выбираются вершины и переменные для осуществления процесса ветвления?

12. Что такое прозондированная вершина? При выполнении каких условий при реализации метода ветвей и границ вершина объявляется прозондированной?

13. Каким образом на основании выбранной вершины и переменной формируются две новые вершины (задачи ЛП) в процессе реализации метода ветвей и границ?

14. Почему получение перед реализацией метода ветвей и границ допустимого целочисленного решения может оказаться весьма полезным?

15. Какова принципиальная последовательность действий при решении задач ЦЛП с булевыми переменными с помощью аддитивного алгоритма?

16. Каким образом должна быть представлена исходная задача ЦЛП с булевыми переменными для того, чтобы можно было использовать аддитивный алгоритм?

17. Объясните значения понятий "свободная переменная" и "частичное решение", используемых при реализации аддитивного алгоритма.

18. Каким образом осуществляется построение частичных решений?

19. Объясните сущность тестов, применяемых в процессе реализации аддитивного алгоритма.

20. Почему наличие в прозондированном частичном решении только отрицательных элементов является признаком окончания процедуры перебора?

21. Дана задача

$$Z = x_1 + x_2 \to max,$$
$$-3x_1 + x_2 \leqslant 0,$$
$$3x_1 + x_2 \leqslant 9,$$
$$x_i \geqslant 0,\ x_i \text{ – целое},\ i = 1, 2.$$

Решите указанную задачу без учета целочисленности переменных и покажите, что путем округления результата нельзя получить допустимое решение исходной задачи ЦЛП.

22. Задача. Представьте следующие условия в виде системы линейных неравенств с булевыми переменными:

выполняется неравенство $x_1 - 3x_2 \leqslant 2$ или $2x_1 - x_2 \geqslant 8$;

если $x_3 \leqslant 5$, то $x_4 \geqslant 7$. В противном случае $x_4 \leqslant 2$;

выполняется по крайней мере два из следующих четырех неравенств $x_5 + x_6 \leqslant 6$, $x_5 \leqslant 1$, $x_6 \leqslant 7$, $x_5 + x_6 \leqslant 3$.

23. Задача. На руднике имеется группа из трех блоков, расположенных рядом на одном горизонте. Указанные блоки должны быть вовлечены в обработку, причем следует определить момент начала x_i и длительность отработки t_i каждого i-го блока соответственно ($i = 1, 2, 3$). Принятая система разработки не позволяет вести одновременно работы в соседних блоках, поэтому возможны следующие варианты последовательности добычных работ: сначала отрабатываются блоки с номерами 1 и 3 или 2, а затем 2 или 1 и 3 соответственно.

Формализуйте в рамках систем линейных ограничений технологические требования по возможной последовательности отработки блоков.

24. Задача. На горно-металлургическом комбинате группа рудников обеспечивает поставки руды на обогатительные фабрики, непосредственно в цеха металлургического предела и внешним потребителям. Принято решение о строительстве рудных складов, через которые будет осуществляться отгрузка внутренним и внешним потребителям. Склады могут быть размещены на пяти площадках, причем затраты на строительство на i-й площадке составляют c_i, а ожидаемый экономический эффект за счет усреднения руды на складе – f_i. С учетом существующей транспортной сети и необходимости обеспечения отгрузки по всем направлениям возможны следующие три варианта размещения на площадках: 1, 2, 3; 1, 3, 5 и 2, 4, 5. На строительство выделены средства в размере K.

Формализуйте задачу планирования размещения складов в рамках модели ЦЛП, учитывающей максимизацию экономического эффекта от создания рудных складов.

25. Задача. На производственном участке работают m рабочих, каждый из которых обладает такой квалификацией, что может выполнять одну или несколько операций из общего числа n. Эффективность выполнения i-м рабочим j-й операции f_{ij}. Требуется все операции распределить между рабочими с целью достижения максимальной суммарной эффективности.

Формализуйте задачу о назначениях в рамках модели ЦЛП.

26. Задача. Имеется n объектов, на которые необходимо распределить однородный ресурс.

На каждый объект может быть поставлен ресурс в одном из объемов b_i ($i = 1, 2, ..., m$). Эффективность использования i-го объема на j-м объекте составляет a_{ij}. Суммарный объем распределяемого ресурса d.

Формализуйте задачу распределения дискретных объемов ресурса с целью получения максимального эффекта от его использования.

27. Задачу 26 решить при условии, что не требуется обязательного распределения ресурсов по всем объектам.

28. Задача. Покажите, используя графический метод решения, что задача

$$Z = x_1 + x_2 \to max,$$
$$10x_1 + 10x_2 \leqslant 8,$$
$$10x_1 + 5x_2 \geqslant 1,$$
$$x_i \geqslant 0,\ x_i \text{ – целое},\ i = 1, 2$$

не имеет допустимых целочисленных решений. Проверьте это с помощью метода отсечений и метода ветвей и границ.

29. Решите полностью целочисленную задачу

$$Z = -2x_1 - 3x_2 \to \max,$$
$$5x_1 + 4x_2 \leqslant 20,$$
$$2x_1 + x_2 \geqslant 1,$$
$$x_i \geqslant 0,\ x_i - \text{целое},\ i = 1, 2.$$

В процессе решения используйте правило построения сечения Р. Гомори для полностью целочисленных задач.

30. Задача. Покажите, что если к условиям задачи 29 добавить ограничение $x_2 \geqslant 0{,}5$ и не преобразовать его в ограничение с полностью целочисленными параметрами, то обнаружится признак отсутствия допустимого решения.

31. Решите задачу 29 с дополнительным ограничением $2x_2 \geqslant 1$, используя правило построения сечения Р. Гомори для полностью целочисленных задач, и с ограничением $x_2 \leqslant 5$ – для частично целочисленных задач.

32. Решите задачу 29 с дополнительным ограничением $x_2 \geqslant 0{,}5$, используя правило построения сечения Р. Гомори для частично целочисленных задач. Причем требование целочисленности распространите только на переменную x_2.

33. Покажите, что в случае полностью целочисленной задачи сечение Р. Гомори для частично целочисленной задачи оказывается более эффективным.

34. Решите задачу методом ветвей и границ

$$Z = -x_1 - x_2 \to \min,$$
$$2x_1 + 11x_2 \leqslant 38,$$
$$4x_1 - 5x_2 \leqslant 5,$$
$$x_1 + x_2 \leqslant 7,$$
$$x_i \geqslant 0,\ x_i - \text{целое},\ i = 1, 2.$$

35. Дана задача

$$Z = -x_1 + 4x_2 \to \max,$$
$$x_1 + x_2 \leqslant 3,$$
$$-x_1 + 2x_2 \leqslant 2,$$
$$x_i \geqslant 0,\ x_i - \text{целое},\ i = 1, 2.$$

Представьте указанную задачу в булевых переменных и решите ее с использованием аддитивного алгоритма.

Глава VIII

НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ И ЕГО ИСПОЛЬЗОВАНИЕ В ПЛАНИРОВАНИИ И УПРАВЛЕНИИ ГОРНЫМ ПРОИЗВОДСТВОМ

§ 1. ОБЩАЯ ХАРАКТЕРИСТИКА, ОСНОВНЫЕ ТИПЫ И ОСОБЕННОСТИ ЗАДАЧ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Как было показано ранее, общая задача математического программирования состоит в оптимизации целевой функции

$$W = f(a_1, ..., a_j, ..., a_n;\ x_1, ..., x_j, ..., x_m)$$

при ограничениях

$$g_i = f_i(a_1, ..., a_j, ..., a_n;\ x_1, ..., x_j, ..., x_m) \leqslant A_i.$$
$$i = \overline{1, I}$$

Когда все ограничения и целевая функция были линейны, задача решалась методами линейного программирования. К линейным моделям сводится задача о расстановке оборудования, планирования добычи в режиме усреднения, распределения ресурсов, транспортная, размещения предприятий и др. Более детальное изучение практических ситуаций показывает, что большинство указанных задач по существу является нелинейными и сводится к линейным с достаточной степенью округления. Нелинейность этих задач прежде всего обусловлена зависимостью затрат или прибыли от объема выпуска продукции (т.е. коэффициентов целевой функции или ограничений c_i, $a_{ij} \neq \text{const}$), а также изменением показателей качества и использования сырья в зависимости от объемов работ (в задачах планирования работы горных предприятий и размещения промышленности).

Когда целевая функция или ограничения нелинейны, то возникает широкий класс задач нелинейного программирования. Общий метод решения нелинейных задач отсутствует. Предложенные алгоритмы позволяют эффективно решать лишь весьма узкий класс задач.

Задачи нелинейного программирования делятся на задачи с ограничениями и без ограничений, выпуклого и невыпуклого программирования (в зависимости от типа целевой функции). Весьма важными в практическом отношении являются задачи, в которых целевая функция квадратичная или сепарабельная (является суммой ряда функций).

Основную трудность при решении нелинейных задач составляет наличие множества экстремальных точек (рис. VIII.1). Большинство же классических методов дают точки локального оптимума и не позволяют установить, является ли он глобальным, так как равенство нулю при

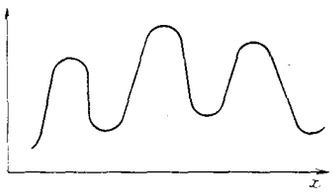


Рис. VIII.1. Многоэкстремальность целевой функции

всех производных в определенной точке необходимое, но недостаточное условие получения глобального оптимума. Перебор же всех экстремальных точек затруднен.

Эффективный симплексный алгоритм решения задач линейного программирования удалось построить, так как доказывается, что область допустимых значений в линейных задачах выпукла, а решение находится в одной из вершин многогранника ограничений. В нелинейных задачах область допустимых решений не всегда выпукла, она может быть и вогнута, и даже разрывна. Кроме того, оптимум целевой функции может достигаться не только в вершине многогранника, но и на его границе и даже внутри многогранника. Проиллюстрируем это примерами.

Пример VIII.1. Найти экстремальные (минимальное и максимальное) значения сепарабельной квадратичной целевой функции

$$Z = (x_1 - 5)^2 + (x_2 - 4)^2$$

при ограничениях:

$$x_1 + x_2 \geqslant 2,$$

$$3x_1 + 2x_2 \leqslant 12,\ x_{1,2} \geqslant 0.$$

Построив систему прямых, соответствующих ограничениям, получим область допустимых решений в виде многоугольника $ABCD$ (рис. VIII.2). Целевая функция представляет собой уравнение окружности с центром в точке О (координаты 5, 4). С уменьшением (увеличением) квадрата радиуса значение целевой функции уменьшается (увеличивается).

Минимум и максимум целевой функции получим, проводя из точки О окружности различных радиусов. Минимальное значение ($Z = 9{,}32$) целевая функция принимает в точке касания окружностью многогранника ограничений (точка К) с координатами (2,4; 2,4), т.е. минимум не находится в вершине многогранника, а лежит на ребре. Максимальное же значение ($Z = 29$) целевая функция достигает в точке A (с координатами 0, 2), т.е. в вершине многогранника.

Пример VIII.2. Найти максимум и минимум целевой функции

$$(x_1 - 1)^2 + (x_2 - 3)^2$$

при тех же ограничениях, что и в предыдущем примере.

Область допустимых решений по-прежнему представляет выпуклый многоугольник $ABCD$, а целевая функция – уравнение окружности с центром в точке O_1 (рис. VIII.3) с координатами 1, 3. Минимальное значение ($Z = 0$) целевая функция принимает в центре окружности (точка O_1), т.е. внутри области допустимых значений функции, а максимальное ($Z = 18$) – в вершине многогранника ограничений в точке C с координатами 4, 0.

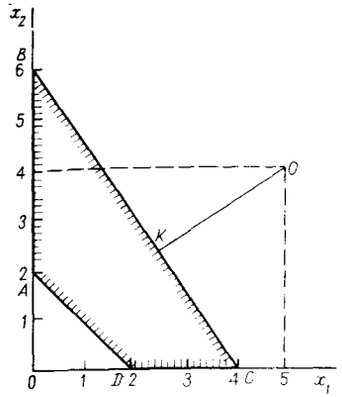


Рис. VIII.2. Расположение максимума целевой функции в вершине многогранника ограничений (A), а минимума на ребре (D)

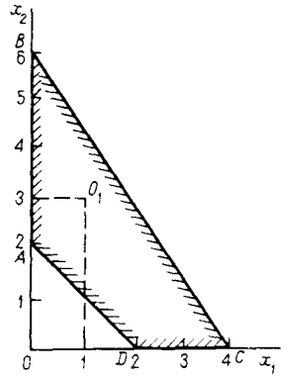


Рис. VIII.3. Расположение минимума целевой функции внутри многогранника ограничений

Пример VIII.3. Найти максимум и минимум функции

$$(x_1 - 2)^2 + (x_2 - 1)^2$$

при ограничениях:

$$x_1 x_2 \leqslant 3,$$

$$x_1 + x_2 \geqslant 4,$$

$$0 \leqslant x_1 \leqslant 5,\ 0 \leqslant x_2 \leqslant 5.$$

Первое ограничение соответствует гиперболе, а остальные – прямым линиям. Построив графически систему ограничений (рис. VIII.4), убеждаемся, что область допустимых значений не только не выпукла, но и разрывна (состоит из двух отдельных частей $ABCD$ и $EFGH$).

Целевая функция представляет собой уравнение окружности с центром в точке О с координатами 2, 1. Максимум и минимум целевой функции достигаются в точках касания окружности угловых точек области ограничений, соответственно B (0, 5) и E (3, 1).

При этом $Z_{\max} = 20$, а $Z_{\min} = 1$.

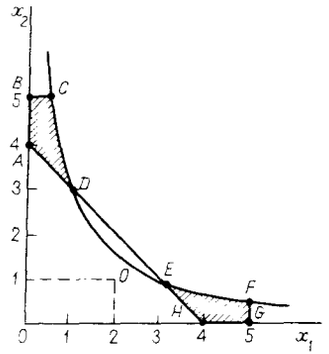


Рис. VIII.4. Разрывность области допустимых значений управляемых переменных

§ 2. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ БЕЗУСЛОВНОЙ ОПТИМИЗАЦИИ

Задача минимизации функции многих переменных без дополнительных ограничений относится к числу простейших задач нелинейного программирования. Применяемые для ее решения методы являются итерационными. Это значит, что взяв некоторое начальное приближение – n-мерный вектор $x^0 = \{x^0_1, x^0_2, ..., x^0_n\}$, строят последовательность n-мерных векторов $x^0, x^1, ...$ (в дальнейшем будем называть их просто точками), сходящуюся к оптимальной точке x. Сходимость за конечное количество итераций ($x^m = x$) может иметь место лишь чисто случайно, за исключением задач, в которых функция f является квадратичной, т.е. может быть представлена в виде

$$f(x_1, ..., x_n) = \sum_{i=1}^{n} \sum_{j=1}^{m} d_{ij} x_i x_j + \sum_{i=1}^{n} c_i x_i + b.$$

Наиболее распространенными методами безусловной оптимизации являются методы: покоординатного спуска, градиентные, сопряженных направлений и метод Ньютона. Применимость той или иной группы методов зависит от наличия у функции непрерывных производных первого и второго порядка и трудоемкости вычисления функции и ее производных.

Методы покоординатного спуска – простейшие, хотя по эффективности они могут превосходить и более сложные градиентные. Для их применения требуется производить вычисление только самой функции f_0 в разных точках. Пусть в качестве начальной взята точка $\{x^0_1, ..., x^0_n\}$. Тогда следующая точка x^1 отличается от x^0 значением лишь первой координаты $x^1 = \{x^1_1, x^0_2, ..., x^0_n\}$, которая находится из решения задачи $f(x_1, x^0_2, ..., x^0_n) \to \min$.

На следующей итерации изменяется лишь значение второй координаты, которая является решением задачи $f(x^1_1, x_2, x^1_3, ..., x^1_n) \to \min$, на n-й итерации определяется новое значение n-й координаты, на $(n+1)$-й – снова первой и т.д. Таким образом, исходная сложная задача минимизации функции многих переменных сводится к последовательности задач минимизации функции одной переменной.

Рассмотрим применение метода покоординатного спуска на примере.

Пример VIII.4. Методами математической статистики установлена следующая зависимость годовой прибыли горно-обогатительного комбината $П$ (млн. руб.) от содержания двух основных компонентов в руде x_1 и x_2 (в %)

$$П(x) = x_1 + x_2 + 3x_1x_2 - 2x_1^2 - 2x_2^2.$$

Требуется найти оптимальное значение содержания компонентов, при котором прибыль максимальна, т.е. значения x_1 и x_2, максимизирующие $П(x)$.

Равносильная постановка задачи: минимизировать $f(x) = -П(x)$. Пусть в качестве начальной взята точка $\{1, 0\}$. Величину x^1_1 находим из решения задачи

$$-(x_1 + 0 + 3 \cdot 0 \cdot x_1 - 2x_1^2 - 2 \cdot 0^2) = -x_1 + 2x_1^2 \to \min.$$

Приравнивая производную к нулю, получаем

$$x^1_1 = \frac{1}{4} = 0{,}250, \quad x^1_2 = x^0_2 = 0.$$

Подобным образом x^2_2 находим из решения задачи

$$-x_2 - (-\frac{3}{4})x_2 + 4x_2^2,$$

откуда имеем $x^2_2 = \frac{7}{16} = 0{,}438$.

Повторяя подобные выкладки на последующих итерациях, получим

$$x^3_1 = \frac{37}{64} = 0{,}578; \; x^4_2 = \frac{175}{256} = 0{,}684; \; x^5_1 = \frac{781}{1024} = 0{,}763; \; x^6_2 = \frac{3367}{4096} = 0{,}847;$$

$$x^7_1 = 0{,}885; \; x^8_2 = 0{,}907; \; x^9_1 = 0{,}930; \; x^{10}_2 = 0{,}947; \; x^{11}_1 = 0{,}960; \; x^{12}_2 = 0{,}970;$$

$$x^{13}_1 = 0{,}977; \; x^{14}_2 = 0{,}982; \; x^{15}_1 = 0{,}986; \; x^{16}_2 = 0{,}989; \; x^{17}_1 = 0{,}992; \; x^{18}_2 = 0{,}994;$$

$$x^{19}_1 = 0{,}995; \; x^{20}_2 = 0{,}996; \; x^{21}_1 = 0{,}997; \; x^{22}_1 = 0{,}998; \; x^{23}_1 = 0{,}998; \; x^{24}_2 = 0{,}998.$$

Таким образом, производя вычисления с точностью до трех значащих цифр, мы получили близкие к оптимальным значения $x_1 = x_2 = 0{,}998$. Взяв другое начальное приближение, мы получим не в точности те же, но близкие значения x_1, x_2. Вычисления на ЭВМ обычно производятся по крайней мере с 6 десятичными цифрами, что обеспечивает для подобных простых задач достаточную точность получения оптимума. Но, как можно было видеть, итерационный процесс сходился сравнительно медленно.

В градиентных методах переход от предыдущей точки x^k к последующей x^{k+1} производится вдоль направления наибольшего убывания целевой функции (в точке x^k). Это направление противоположно градиенту функции $\nabla f(x^k)$, т.е. вектору, i-я компонента которого есть $\frac{\partial f(x^k)}{\partial x^k_i}$. Метод успешно работает, если только градиент является непрерывной вектор-функцией.

Таким образом, $x^{k+1} = x^k - \alpha \nabla f(x^k)$. Величина шага α_k может определяться, как и в методе покоординатного спуска, из решения задачи одномерной минимизации: найти положительное число α, минимизирующее $f(x^k - \alpha \nabla f(x^k))$.

Рассмотренный пример решается градиентным методом следующим образом. Пусть взята та же начальная точка 1, 0. Градиент целевой функции $\nabla П = \{1 + 3x_2 - 4x_1; \; 1 + 3x_1 - 4x_2\}$ в начальной точке равен $\{-3, 4\}$. Следовательно, точку x^1 ищем среди точек вида $\{1 - 3\alpha; \; 4\alpha\}$, $\alpha > 0$; для этих точек получаем выражение $П(x) = -1 + 25\alpha - 86\alpha^2$, достигающее максимума при $\alpha = 2\frac{25}{86} = 0{,}145$. Таким образом:

$$x^1 = \left\{\frac{97}{172}; \; \frac{100}{172}\right\} = \{0{,}564; \; 0{,}581\};$$

$$\nabla \Pi(x^1) = \left\{ \frac{84}{172}; \frac{63}{172} \right\} = \left\{ 0{,}488;\ 0{,}366 \right\}, x^2 = \left\{ 1{,}000;\ 0{,}908 \right\},$$
$$x^3 = \left\{ 0{,}960;\ 0{,}962 \right\}, x^4 = \left\{ 1{,}000;\ 0{,}992 \right\}, x^5 = \left\{ 0{,}996;\ 0{,}996 \right\},$$
$$x^6 = \left\{ 1{,}000;\ 0{,}999 \right\}, x^7 = \left\{ 1{,}000;\ 1{,}000 \right\}.$$

В рассмотренном примере одномерная минимизация (максимизация) целевой функции не представляла трудностей, так как минимизировалась квадратичная функция. В более сложных случаях эта задача сама может потребовать использования итерационного метода. В то же время существуют варианты градиентного метода, когда α вычисляется гораздо проще. Например, Ю.М. Данилин и Б.Н. Пшеничный предложили в качестве α_k выбирать первое число из последовательности $\alpha_0, \alpha_0 \cdot \lambda, \alpha_0 \cdot \lambda^2$, для которого выполнено условие

$$f(x^k - \alpha \nabla f(x^k)) - f(x^k) \leqslant \epsilon \alpha \sum_{i=1}^{n} \left(\frac{\partial f(x^k)}{\partial x_i} \right)^2;$$

здесь $\alpha_0 > 0, 1 > \epsilon > 0$ – заранее выбранные числа.

Недостаток градиентных методов состоит в том, что вблизи точки минимума градиент близок к нулю и потому линейное приближение теряет точность. Если требуется получить значение оптимума с высокой точностью, на завершающем этапе оптимизации используются другие методы.

Метод Ньютона основан на квадратичной аппроксимации целевой функции. Обозначим $A(x)$ квадратную матрицу размерности n, элементы которой

$$A_{ik}^{(x)} = \frac{\partial^2 f(x)}{\partial x_i \partial x_k}.$$

Для нахождения x^{k+1} решают задачу минимизации функции, являющейся квадратичным приближением $f(x)$

$$\psi(x) = f(x^k) + \nabla f^T(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^T A(x^k)(x - x^k),$$

что дает $x - x^k = p_k = -(A(x^k))^{-1} \cdot \nabla f(x^k)$. В обычном методе Ньютона берут $x^{k+1} = x^k + p_k$, в методе Ньютона с регулировкой $x^{k+1} = x^k + \alpha_k p_k$, $0 < \alpha_k \leqslant 1$ (существуют многочисленные способы вычисления α_k). Необходимость регулировки шага вызывается тем, что обычный метод Ньютона при производном начальном приближении не гарантирует сходимости последовательности $\{x^k\}$.

В нашем примере квадратичная аппроксимация является точной, т.е. $\psi(x) = f(x) = -\Pi(x)$, и метод Ньютона дает оптимальную точку за одну итерацию ($x^1 = x$) при любом начальном приближении x^0. Убедимся в

этом. Действительно, при любом x^k

$$A(x^k) = A = \begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}, A^{-1} = \begin{pmatrix} 4/7 & 3/7 \\ 3/7 & 4/7 \end{pmatrix}$$

и при любом начальном приближении

$$x_1^1 = x_1^0 + 4/7(1 + 3x_2^0 - 4x_1^0) + 3/7(1 + 3x_1^0 - 4x_2^0) = 1,$$
$$x_2^0 = 3/7(1 + 3x_2^0 - 4x_1^0) + 4/7(1 + 3x_1^0 - 4x_2^0) = 1.$$

В полученной точке $\nabla f(x^1) = \{0, 0\}$ и, следовательно, выполнено необходимое условие оптимальности.

Не только в рассматриваемом примере, но и для любой квадратичной функции метод Ньютона дает решение за одну итерацию. Однако каждая итерация метода Ньютона требует обращения матрицы, что при больших n означает неприемлемо большой объем вычислений. По этой причине даже для квадратичных задач большой размерности применяют другую группу методов – методы сопряженных направлений, в которых идея квадратичной аппроксимации осуществляется менее трудоемким способом. Эти методы для квадратичных задач дают решение не более чем за n итераций и могут применяться в произвольных нелинейных задачах.

Методы сопряженных направлений основаны на построении такой системы векторов $p_0, \ldots, p_{n-1}$, связанной с матрицей вторых производных A, что всякую точку x^0 можно представить в виде

$$x^0 = x + \alpha_0 p + \alpha_1 p_1 + \ldots + \alpha_{n-1} p_{n-1}. \qquad \text{(VIII.1)}$$

Для квадратичных задач это представление является точным, а для остальных – приближенным, но погрешность формулы (VIII.1) – второго порядка относительно величины $(x^0 - x)$, что и означает n-шаговую квадратичную аппроксимацию, а как следствие – быструю сходимость к оптимальной точке.

Величины α_i – это уже известные нам множители, определяемые на i-й итерации из решения задачи одномерной минимизации, так что формула (VIII.1) не является чисто теоретической, а отражает действительный порядок получения решения

$$x^1 = x^0 - \alpha_0 p_0;\ x^2 = x^1 - \alpha_1 p_1, \ldots,$$
$$x^n = x^{n-1} - \alpha_{n-1} p_{n-1}.$$

В качестве вектора p_0 берется антиградиент в начальной точке: $p_0 = -\nabla f(x^0)$, вычисляется $x^1 = x^0 + \alpha_0 p_0$ (α_0, как и ранее, находится из решения задачи безусловной минимизации). Затем при k от 1 до $n - 1$ производят следующие вычисления:

$$\beta_k = -\frac{\nabla f^T(x^k)[\nabla f(x^k) - \nabla f(x^{k-1})]}{f^T(x^{k-1}) p_{k-1}},$$

$$p_k = -\nabla f(x^k) + \beta_k p_k,\ x^{k+1} = x^k + \alpha_k p_k,$$

где

$$f(x^k + \alpha_k p_k) = \min\left\{ f(x^k + \alpha p_k) : \alpha \geqslant 0 \right\}.$$

Для задачи квадратичного программирования x^n есть оптимальная точка, в остальных случаях x^n берется в качестве x^0 и процесс повторяется до тех пор, пока длина градиента не уменьшится до заданной малой величины.

Все перечисленные методы гарантируют получение лишь локального минимума, т.е. точки, доставляющей минимум функции в некоторой области n-мерного пространства, которая может быть уже области поиска решения. Существующие методы поиска глобального минимума требуют гораздо большего объема вычислений и практически могут быть применены, если размерность задачи (количество переменных) не более 10. Они делятся на две группы: методы случайного поиска и методы неравномерного перебора и покрытия. В первом случае с помощью специальных процедур имитируется случайная последовательность точек из области поиска. Чем длиннее последовательность, тем больше вероятность того, что наименьшее значение целевой функции на последовательности достаточно близко к оптимальному. Для большей надежности следует комбинировать случайный поиск с методами локальной минимизации, выбирая для них в качестве начальных приближений лучшие из полученных точек. При высокой размерности задачи комбинированная стратегия является предпочтительной.

Методы неравномерного перебора и покрытия основаны на знании некоторых характеристик функции, в частности, константы Липшица L – величины, при которой при любых $x_1, ..., x_n;\ x_1', ..., x_n'$

$$|f(x_1, x_2, ..., x_n) - f(x_1', x_2', ..., x_n')| \leqslant L\sqrt{(x_1 - x_1')^2 + ... + (x_n - x_n')^2}.$$

Если следует найти оптимум с точностью до ϵ и вычислено значение функции в точке x_1, то n-мерный шар радиуса ϵ/L вокруг этой точки не следует рассматривать. Если же в другой точке $f(x_2) < f(x_1)$, то из рассмотрения можно выбросить шар радиуса $\dfrac{f(x_1) - f(x_2) + \epsilon}{L}$.

Таким образом, область поиска последовательно сужается и в конце концов покрывается этими шарами полностью.

§ 3. ПРЯМЫЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ УСЛОВНОЙ ОПТИМИЗАЦИИ

Задачи условной оптимизации – это задачи о нахождении минимума (максимума) целевой функции $f(x_1, ..., x_n)$ при условии, что величины $x_1, ..., x_n$ удовлетворяют ограничениям типа равенства

$$g_i(x_1, ..., x_n) = 0 \quad (i = 1, ..., m) \tag{VIII.2}$$

или неравенства

$$r_j(x_1, ..., x_n) \leqslant 0 \quad (j = 1, ..., s). \tag{VIII.3}$$

Часто встречаются задачи, в которых имеются ограничения обоих видов. При решении рассматриваемых задач применяются либо прямые методы, явно учитывающие ограничения на каждой итерации, либо методы преобразования, сводящие решение исходной задачи к последовательности задач безусловной оптимизации.

Рассмотренные ранее задачи линейного программирования, а также квадратичного программирования – в ней целевая функция квадратична, а ограничения линейны – также принадлежат к числу задач условной оптимизации, причем методы их решения наиболее развиты. По этой причине прямые методы решения других задач условной оптимизации основываются на линейной или квадратичной аппроксимации целевой функции и, возможно, ограничений и на решении задач линейного или квадратичного программирования для определения приращений $x_1, ..., x_n$ на каждой итерации.

Метод возможных направлений, в котором линеаризуются и целевая функция, и ограничения, служит для решения задач с ограничениями типа неравенства (VIII.3). В качестве начального приближения берут допустимую точку x^0; строится последовательность допустимых точек x^k, на которой целевая функция монотонно уменьшается до оптимального значения. На произвольной k-й итерации направление p^k находят из решения задачи линейного программирования: минимизировать скалярную величину η при ограничениях

$$\nabla f^T(x^k)p \leqslant \eta;$$

при любом j, при котором j-е ограничение с точностью до δ_k обращается в равенство $(r_j(x^k) \geqslant -\delta_k)$

$$\nabla r_j^T(x_k)p \leqslant \eta;$$

при любом i от 1 до n

$$|p_i| \leqslant 1. \tag{VIII.4}$$

Смысл этой задачи прост: находится направление, обеспечивающее убывание целевой функции и ненарушение ограничений. Найденное направление задает луч $x^k + \alpha p^k$, $\alpha \geqslant 0$, точки которого принадлежат допустимой области при малых α. В качестве значения δ_k обычно выбирается наибольшее из чисел последовательности $1, \frac{1}{2}, \frac{1}{4}$ ..., для которого

$$\eta \leqslant -\delta_k. \tag{VIII.5}$$

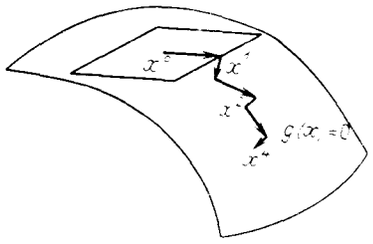


Рис.VIII.5. Пример последовательности, построенной с помощью метода проекции градиента

По теоретическим оценкам более быстрая сходимость к оптимальному значению целевой функции будет достигнута, если вместо (VIII.5) использовать условие $\eta \leqslant -\sqrt{\delta_k}$.

Величина α_k выбирается среди всех значений α, при которых точка $x^k + \alpha p^k$ допустима, по тем же правилам, что и в градиентных методах: в качестве x^{k+1} принимается $x^k + \alpha_k p^k$.

Метод возможных направлений может использоваться также для решения задач, в которых наряду с ограничениями (VIII.3) имеются линейные ограничения (VIII.2).

Методы условного градиента и прямой метод Ньютона основаны на решении на каждой итерации для определения p_k задач с исходными ограничениями при соответственно линейной и квадратичной аппроксимации целевой функции. Метод Ньютона особенно эффективен для задач, в которых ограничения линейны. В целом область применения этих методов уже, чем у метода возможных направлений, и они менее конструктивны, так как задачи определения p_k для них требуют специальных подходов.

Для задач с ограничениями типа равенства наиболее часто используется метод проекции градиента. В этом методе каждая итерация разделена на два шага. Первый – сдвиг по касательной к ограничениям в сторону убывания целевой функции. Второй – определение точки, ближайшей к найденной на первом шаге и удовлетворяющей ограничениям, иными словами – проектирование найденной точки на допустимую поверхность. При этом приращение целевой функции в первом приближении определяется первым шагом, который и является основным для оптимизации. На этом шаге фактически используется градиентный метод на касательной гиперплоскости. Если же ограничения являются линейными, то операция проектирования выполняется только на первой итерации.

На рис. VIII.5 изображен пример последовательности, построенной с помощью метода проекции градиента.

Наконец, метод линеаризации Б.Н. Пшеничного может быть применен к задачам с ограничениями обоих видов. Он не требует, чтобы начальная точка была допустимой. Ввиду того, что для этого метода ограничение-равенство (VIII.2) можно заменить на два неравенства

$g_i(x_1, ..., x_n) \leqslant 0,$

$g_i(x_1, ..., x_n) \geqslant 0,$

достаточно рассмотреть случай ограничений (VIII.3). Для нахождения направления спуска p_k решается задача: минимизировать

$$\nabla f^T(x^k)p + \frac{1}{2}\sum_{i=1}^{n} p_i^2 \qquad \text{(VIII.6)}$$

при ограничениях

$$r_j(x^k) + \nabla r_j^T(x^k)p \leqslant 0$$

при любом j, для которого $r_j(x^k) \geqslant -\delta_k$.

Смысл этой задачи тот же, что и в методе возможных направлений, только ограниченность вектора p обеспечивается не прямо [как в условии (VIII.4)], а через второе ("штрафное") слагаемое в (VIII.6). При выборе $x^{k+1} = x^k + \alpha_k p^k$ требуется обеспечить не убывание целевой функции и соблюдение ограничений, как в методе возможных направлений, а убывание функции

$$f(x) + N\max r_j(x),\ 1 \leqslant j \leqslant s, \qquad \text{(VIII.7)}$$

где N – специальным образом подбираемое положительное число. С определенной точки зрения выражение (VIII.7), включающее целевую функцию и ограничения, является штрафной функцией. Метод линеаризации – пограничный между прямыми методами и методами преобразования.

§ 4. МЕТОДЫ ПРЕОБРАЗОВАНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ УСЛОВНОЙ ОПТИМИЗАЦИИ

Большинство прямых методов, описанных в предыдущем параграфе, не отличается универсальностью: одни предназначены для задач с ограничениями-равенствами (VIII.2), другие – с неравенствами (VIII.3). Правда, как и в линейном программировании, неравенство может быть преобразовано в равенство, но цена этому преобразованию – увеличение размерности задачи. Есть и другие причины, по которым непрямые методы часто предпочтительнее.

Наиболее распространенный непрямой метод – метод штрафных функций. Задача минимизации функции $(x_1, x_2, ..., x)$ на области Ω n-мерного пространства равносильна задаче нахождения минимума

$$F(x) = f(x) + G(x),$$

если мы определим функцию $G(x)$ из условия

$$G(x) = \begin{cases} 0, \text{ если } x \in \Omega \text{ (т.е. все условия выполнены)} \\ \infty, \text{ если } x \in \Omega \text{ (т.е. хотя бы одно из условий нарушено)}. \end{cases}$$

Непосредственно минимизировать $F(x)$ известными методами невозможно, так как $G(x)$ не является непрерывной функцией. Решают последовательность задач минимизации

$$F_k(x) = f(x) + G_k(x),$$

где $G_k(x) \xrightarrow[k \to \infty]{} G(x)$ при любом x. Функция $G_k(x)$ является штрафом за выход за ограничения.

Чаще всего $G_k(x)$ представляется в виде

$$G_k(x) = r_k G_1(x) = r_k \left[\sum_{i=1}^{m} (g_i(x))^2 + \sum_{j=1}^{s} (\max \{ r_i(x), 0 \})^2 \right]$$

и тогда процесс решения задачи определяется последовательностью $r_k \xrightarrow[k \to \infty]{} +\infty$. Точка x^k, найденная из решения задачи минимизации $F_k(x)$, служит начальным приближением для отыскания точки x^{k+1} на следующей итерации. Решать задачу минимизации $F_k(x)$ с высокой точностью нет необходимости. Теоретически обоснована сходимость варианта метода штрафа, в которой в качестве x^k может быть принята точка с длиной градиента $\nabla F_k(x^k)$ не больше l_k, а $l_k \xrightarrow[k \to \infty]{} 0$ (если бы задача решалась точно, эта длина должна была бы равняться нулю).

Достоинством метода штрафов является возможность использовать начальные приближения, весьма далекие от оптимальной точки. В то же время получить решение задачи с высокой точностью этим методом невозможно. Дело в том, что с ростом r все труднее становится решать вспомогательную задачу, поскольку всюду, кроме окрестности искомой точки $\bar{x}_k$, первое слагаемое $F_k(x)$ пренебрежимо мало по сравнению со штрафным слагаемым и, следовательно, практически не учитывается. Сама же окрестность, в которой происходит поиск x^k, становится все меньше по мере увеличения r_k.

Метод модифицированных функций Лагранжа является обобщением классического метода множителей Лагранжа. Из теории известно, что в точке x, являющейся решением задачи условной минимизации при ограничениях (VIII.2), существуют числа (множители Лагранжа) $\bar{\lambda}_i$ ($i = 1, \ldots, m$), при которых

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \bar{\lambda}_i \nabla g_i(\bar{x}) = 0. \qquad \text{(VIII.8)}$$

Условие (VIII.8) – это система из n уравнений и, кроме того, имеется m уравнений (VIII.2).

Таким образом, из этих уравнений могут быть найдены $n + m$ величин – $\bar{x}_1, \ldots, \bar{x}_n, \bar{\lambda}_1, \ldots, \bar{\lambda}_m$. Если ввести функцию Лагранжа

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

то уравнения (VIII.2), (VIII.8) означают, что ее частные производные по всем аргументам в точке $\{x, \bar{\lambda}\}$ равны нулю. Применение классического метода множителей Лагранжа покажем на примере задачи оптимального распределения ресурсов.

Пример VIII.5. На развитие трех шахт ($i = 1, 2, 3$) дополнительно выделено 29 млн. руб. Эффективность использования капиталовложений по этим шахтам выражается зависимостями:

по № 1 $q_1 = 0{,}2\sqrt{x_1}$,
по № 2 $q_2 = 0{,}3\sqrt{x_2}$,
по № 3 $q_3 = 0{,}4\sqrt{x_3}$,

где x_1, x_2, x_3 – объем капиталовложений, выделяемых соответственно первой, второй и третьей шахте.

Снижение темпа роста эффективности по мере возрастания объемов капиталовложений – результат вовлечения в эксплуатацию участков с худшими горно-геологическими условиями.

Необходимо выделенные капиталовложения распределить таким образом, чтобы суммарная эффективность их использования была максимальной.

Кажущееся на первый взгляд правильным решение о выделении всех средств шахте № 3, где эффективность их использования максимальна, не является наилучшим.

Решим задачу с помощью метода неопределенных множителей Лангранжа.

Суммарная эффективность использования ресурсов выражается целевой функцией

$$Z = f(x_1, x_2, x_3) = 0{,}2\sqrt{x_1} + 0{,}3\sqrt{x_2} + 0{,}4\sqrt{x_3}.$$

Необходимо максимизировать эту функцию с учетом ограничения по выделенным капиталовложениям

$$x_1 + x_2 + x_3 = 29.$$

Составим функцию Лагранжа

$$F(x, \lambda) = 0{,}2\sqrt{x_1} + 0{,}3\sqrt{x_2} + 0{,}4\sqrt{x_3} - \lambda(x_1 + x_2 + x_3 - 29).$$

Беря частные производные по x_1, x_2, x_3 и λ, получим систему уравнений

$$\frac{\partial F}{\partial x_1} = \frac{0{,}2}{2\sqrt{x_1}} - \lambda = 0,$$

$$\frac{\partial F}{\partial x_2} = \frac{0{,}3}{2\sqrt{x_2}} - \lambda = 0,$$

$$\frac{\partial F}{\partial x_3} = \frac{0{,}4}{2\sqrt{x_3}} - \lambda = 0, \quad x_1 + x_2 + x_3 - 29 = 0.$$

Выразив x_1, x_2, x_3 через λ из первых трех уравнений и подставив эти выражения в четвертое уравнение, получим следующие значения искомых переменных: $x_1 = 4$; $x_2 = 9$; $x_3 = 16$, при этом $\lambda = 0{,}05$.

Таким образом, на развитие шахты № 1 надо выделить 4 млн., № 2 – 9 млн., а № 3 – 16 млн. руб. Как видно из решения, деление выделенных ресурсов на части рациональнее, чем передача всех средств предприятию, где эффективность их использования максимальна. Действительно, при полученном решении общая эффективность составит $(0{,}2\sqrt{4} + 0{,}3\sqrt{9} + 0{,}4\sqrt{16}) = 2{,}9$, а при передаче всех средств шахте № 3 – $(0{,}4\sqrt{29}) = 2{,}15$.

В задаче с ограничениями-неравенствами функция Лагранжа будет

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{s} \mu_j r_j(x).$$

По-прежнему справедливо условие равенства нулю всех частных производных в оптимальной точке $(\bar{x}, \bar{\lambda}, \bar{\mu})$, но, кроме того, $\bar{\mu}_j \geqslant 0$ при любом j. Для того чтобы сформулированное условие стало системой уравнений, положим $\bar{\mu}_j = \bar{\nu}_j^2$. Тогда величины $x_1, ..., x_n, \lambda_1, ..., \lambda_m, \nu_1, ..., \nu_s$ находятся из системы уравнений, объединяющей условия (VIII.1)

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^{s} \nu_j^2 \nabla r_j(\bar{x}) = 0$$

и

$$\bar{\nu}_j r_j(\bar{x}) = 0 \quad (j = 1, ..., s).$$

Существуют и другие формы модифицированной функции Лагранжа. Таким образом, метод модифицированных функций Лагранжа сводит исходную задачу к системе уравнений, для решения которой могут использоваться известные методы Ньютона, простой итерации и т.д.

§ 5. ПРИБЛИЖЕННЫЕ МЕТОДЫ РЕШЕНИЙ НЕЛИНЕЙНЫХ ЗАДАЧ

К приближенным методам относят методы решения нелинейных задач с сепарабельными целевыми функциями, основанные на их кусочно-линейной аппроксимации. При этом решается задача, несколько отличающаяся от исходной (приближенная к ней), обычно значительно большая по размерности, но линейная. Это позволяет использовать стандартный алгоритм симплекс-метода. Точность решения задачи зависит от точности аппроксимации исходной нелинейной целевой функции – кусочно-линейной. Причем точность аппроксимации регулируется количеством отрезков (интервалов). В некоторых случаях можно использовать методы непосредственной линеаризации целевой функции. При этом вместо одной нелинейной задачи решают несколько линейных. В линейных моделях коэффициенты целевой функции постоянны и не зависят от значения переменных. Обычно же это условие не всегда соблюдается. Например, затраты обычно гиперболически зависят от объема (рис. VIII.6,*а*). Для решения подобных задач зависимость коэффициентов целевой функции от переменных аппроксимируют линейными отрезками. Для этого выделяют интервалы изменения переменных, внутри которых коэффициенты можно считать постоянными (рис. VIII.6,*б*).

Тогда решают задачу линейного программирования, причем первоначально берут коэффициенты при максимальных значениях переменных. Если полученное решение содержит переменные, которым соответствуют взятые коэффициенты, то задача решена. Если же нет, то для полученных значений переменных изменяют коэффициенты целевой функции. Эта процедура повторяется до тех пор, пока дважды не будет получено одно и то же решение, т.е. когда принятые коэффициенты целевой функции и значения переменных соответствуют друг другу. Поясним этот метод на примере решения нелинейной целочисленной задачи определения мощностей предприятия нерудных строительных материалов.

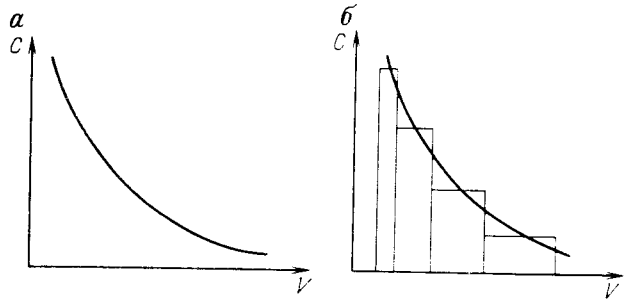


Рис. VIII.6. Гиперболическая зависимость затрат от объемов (*а*) и ее аппроксимация (*б*)

В пределах некоторого экономического района имеется n пунктов производства (разведанные месторождения, существующие карьеры $i = 1, 2, 3, ..., n$) и m пунктов потребления нерудного сырья (заводы железобетонных изделий, строительные и дорожные организации и т.д. $j = 1, 2, 3, ..., m$). Требуется определить объем производства продукции на каждом месторождении карьера x_i (т.е. мощность карьера) и объем перевозок готовой продукции x_{ij} с i-го карьера j-му потребителю.

При этом известны нужды каждого потребителя в нерудном сырье b_j, приведенные производственные (эксплуатационные и капитальные) затраты на выпуск единицы продукции на i-м карьере C_i и их функциональная зависимость от мощности карьера, стоимость перевозки 1 м^3 продукции с i-го карьера j-му потребителю C_{ij}.

Кроме того, определены минимальные $x_{i\min}$ и максимальные $x_{i\max}$ объемы производства на каждом карьере. Минимальный объем производства для действующих карьеров $x_{i\min}$ – это достигнутый уровень добычи.

Для намечаемых к строительству карьеров $x_{i\min} = 0$. Максимальный объем производства на карьере $x_{i\max}$ может быть ограничен темпом углубления или скоростью подвигания фронта, пропускной способностью железных и автомобильных дорог, производительностью оборудования, общей потребностью ($\sum_{j=1}^{m} b_j$) в инертных материалах и т.д.

Следует обратить в минимум функцию

$$S = \sum_{i=1}^{n} C_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{m} C_{ij} x_{ij}$$

при ограничениях:

производимая на каждом карьере продукция должна вывозиться

$$x_i = \sum_{j=1}^{m} x_{ij}, \quad (i = 1, 2, 3, ..., n),$$

доставляемая каждому потребителю продукция должна использоваться

$$b_j = \sum_{i=1}^{n} x_{ij}, \quad (j = 1, 2, 3, ..., m),$$

объемы производства инертных материалов должны удовлетворять потребности в них

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{m} b_j,$$

по объему производства

$$x_{i\,\min} < x_i < x_{i\,\max},$$

по положительности поставок

$$x_{ij} \geqslant 0.$$

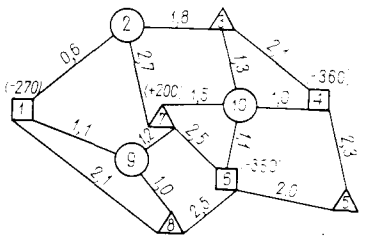
Задача оптимального районирования карьеров и определения их производственной мощности может быть решена в сетевой постановке, которая отличается наглядностью, простотой и экономным алгоритмом.

Ниже на условном примере показан один из возможных приближенных способов решения (в сетевой постановке) нелинейной и целочисленной задачи оптимального размещения производства щебня; описываемый алгоритм этого способа состоит из последовательного решения ряда линейных задач с изменяющимися условиями.

Первоначально оптимизируется план линейной задачи (этап 1), после чего в зависимости от полученных производственных мощностей карьеров меняются производственные затраты на единицу продукции (этап 2) и ее решение осуществляется вновь. Оба этапа последовательно повторяются до тех пор, пока отпадает необходимость в пересчете удельных производственных затрат. Полученное решение является минимумом (иногда локальным) нелинейной задачи. Если мощности карьеров не соответствуют типовым, то постепенно они округляются до заданных типовых (см. ниже). В результате получается план, являющийся минимумом (иногда локальным) целочисленной задачи.

Пример VIII.6. Пусть на какой-то период времени имеется транспортная сеть, состоящая из системы транспортных узлов и связывающих их коммуникаций (рис. VIII.7). Кружками обозначены транспортные узлы – вершины сети. Линии – ребра сети – показывают транспортные коммуникации, перевозка по которым возможна в обоих направлениях, каждое ребро характеризуется затратами (C_{ij}) на перевозку единицы продукции (1 м^3 щебня) по изображаемой ребром коммуникации. На рис. VIII.7 затраты показаны в разрывах ребер (тыс. руб. на 1 тыс. м^3 щебня). Например, на ребре 1,8 $C_{18} = 2{,}1$.

Рис. VIII.7. Транспортная сеть



Вершины 1, 4, 6 (пункты потребления) имеется потребность в щебне, равная соответственно 270, 380 и 330 тыс. м^3, которая показана рядом с вершинами в скобках со знаком (–).

В вершине 7 существует карьер производственной мощностью 200 тыс. м^3. Учитывая дефицитность оборудования, наличие кадров, освоенность месторождения, карьер в пункте 7 должен быть сохранен, даже если затраты на нем выше, чем на намечаемых к строительству. В этом случае стоимостные показатели этого карьера для решения задачи не имеют значения. Наличие обязательного производства в пункте 7 показано на рис. VIII.7 числом (+200).

На месторождениях, расположенных в пунктах 3, 5, 8, возможно строительство дробильно-сортировочных заводов с типовой годовой производственной мощностью 200, 400, 600 тыс. м^3. Строительство заводов большей мощности не рассматривается в связи с ограниченностью запасов месторождений и пропускной способности дорог. В случае нецелесообразности освоения месторождения мощность карьера равна 0.

Приведенные удельные затраты (капитальные и эксплуатационные) на выпуск 1000 м^3 щебня в зависимости от мощности карьера показаны в табл. VIII.1.

Так как производственные затраты изменяются нелинейно в зависимости от мощности карьера, то в соответствии с описанным алгоритмом необходимо знать расходы для рассматриваемого интервала объема производства (0–600 тыс. м^3). В нашем примере затраты даны для типовых дискретных значений мощностей (200, 400, 600 тыс. м^3). Поэтому будем считать, что указанные удельные затраты соответствуют не только типовым мощностям, но и некоторым интервалам возможных значений, т.е. объемам 1–250 тыс. м^3 соответствуют затраты, характерные для типовой мощности 200 тыс. м^3; 251–450 тыс. м^3 – характерные для типовой мощности 400 тыс. м^3 и 450–600 тыс. м^3 – характерные для типовой мощности 600 тыс. м^3.

Таким образом, задача сводится к определению типовой мощности предприятий в пунктах 3, 5, 8 и плана перевозки щебня потребителям (пункты 1, 4, 6), чтобы при полном удовлетворении потребности в щебне суммарные затраты на его производство и перевозку были минимальны.

Таблица VIII.1

Вершина	Мощность карьера, тыс. м^3		
	200	400	600
3	3,1	2,5	2,3
5	3,3	2,7	2,5
8	2,9	2,3	2,1

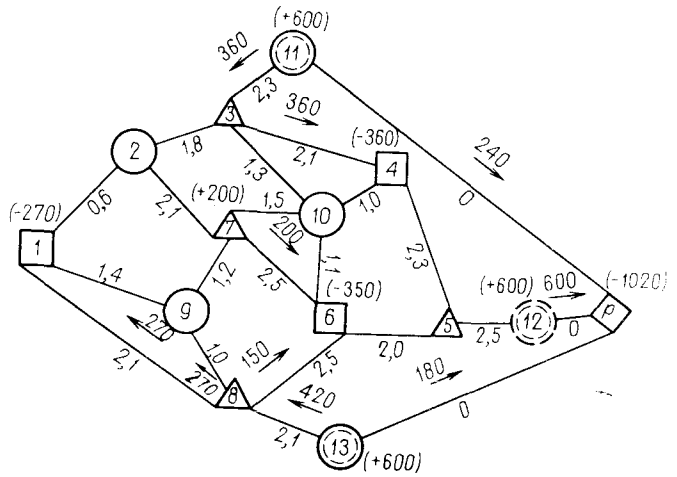


Рис. VIII. 8. Преобразованная транспортная сеть

Для решения поставленной задачи в сетевой постановке имеющуюся сеть (см. рис. VIII.7) необходимо преобразовать таким образом, чтобы производство изображалось аналогично перевозкам на соответствующих ребрах. Для этого возможность добычи в пунктах 3, 5, 8 изобразим в виде дополнительных вершин 11, 12, 13, вынесенных за пределы сети (рис. VIII.8). Тогда процесс производства изображается соответствующими ребрами 3–11, 5–12, 8–13, а затраты на выпуск щебня указываются на этих ребрах.

Описанное преобразование позволяет свести задачу определения оптимального плана производства и перевозок к нахождению плана только "перевозок" на расширенной сети.

Объем производства в пунктах 11, 12, 13 принимаем равным максимальному (600 тыс. $м^3$), при этом производственные затраты, показываемые на ребрах 3–11, 5–12, 8–13, будут наименьшими и равны соответственно 2,3; 2,5; 2,1.

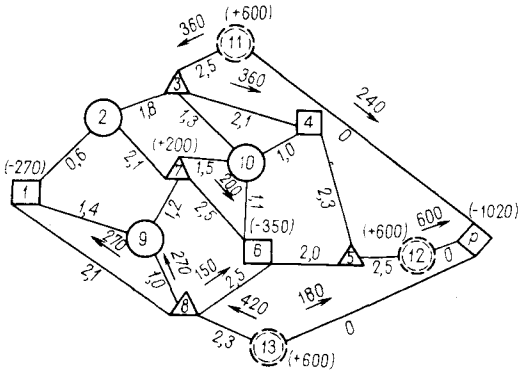
Так как общий объем добычи щебня превышает потребность, то для сбалансирования вводится фиктивный потребитель (точка p на рис. VIII.8) со спросом, равным разности между суммарным объемом производства и потребностью, т.е. $600 \cdot 3 + 200 - (270 + 350 + 360) = 1020$ тыс. $м^3$. К фиктивному потребителю направляем перевозки из пунктов 11, 12, 13, причем стоимость этих перевозок ($C_{11\,p}$, $C_{12\,p}$, $C_{13\,p}$) принимается равной нулю.

В результате описанных преобразований получена в сетевой постановке (см. рис. VIII.8) закрытая транспортная модель, которую решаем методом потенциалов.

Решение первой линейной задачи, где все мощности приняты максимальными, а издержки соответственно минимальными, дает план с наименьшими затратами из всех планов *исходной нелинейной и целочисленной задачи* (рис. VIII.9).

Если при решении первой задачи полученные объемы производства во всех пунктах равны максимальной (600 тыс. $м^3$) или минимальной (0) мощности карьера, то выполнены все условия исходной задачи, и она решена. В противном случае осуществляется переход к этапу 2: решению новой линейной задачи с измененными в соответствии с объемом производства (полученным на этапе 1) затратами на ребрах. При нулевом потоке затраты могут быть оставлены без изменения. В нашем примере на основе оптимального плана, полученного на этапе 1, меняем затраты на ребрах 2–11 и 8–13, т.е. соответственно вместо затрат 2,3 и 2,1 проставляем 2,5 и 2,3. Решаем новую полученную линейную задачу. Ее план (рис. VIII.10) полностью повторяет план первой линейной задачи, показанный на рис. VIII.8. Это означает, что найден по крайней мере *локальный минимум нелинейной задачи*.

Рис. VIII. 9. Оптимальный план нелинейной задачи



После решения нелинейной задачи может оказаться, что мощности нескольких карьеров не соответствуют типовым. Поэтому их необходимо округлять. Из числа карьеров, объемы добычи которых положительны и не равны типовым, выбирается такой, у которого отклонение одной из типовых мощностей или от нуля наименьшее: это отклонение устраняется. В дальнейшем объемы производства, равные типовым, не меняются. Вновь пересчитываются затраты на ребрах, аналогично алгоритму решения нелинейной задачи, и находится минимум новой задачи. Округление объема производства одного из карьеров до типового и решение вновь полученной задачи продолжается до тех пор, пока мощность будет отличаться от типовой не более чем на одном карьере.

На рис. VIII.10 показан план решения задачи, при котором мощности карьеров в пунктах 11 и 13 (все равно, что 3 и 8) не равны типовым. Из них наиболее близка к типовой мощность карьера в пункте 8 (на 20 тыс. $м^3$) превышает типовую. Поэтому, согласно описанному алгоритму, ее целесообразно ликвидировать и наилучшим образом увеличить мощность других карьеров. Такая возможность имеется, так как мощность карьера в пункте 3 недогружена до типовой на 40 тыс. $м^3$. В этом случае в окончательном плане останется нетиповой (но очень близкой к ней) мощность карьера в пункте 3.

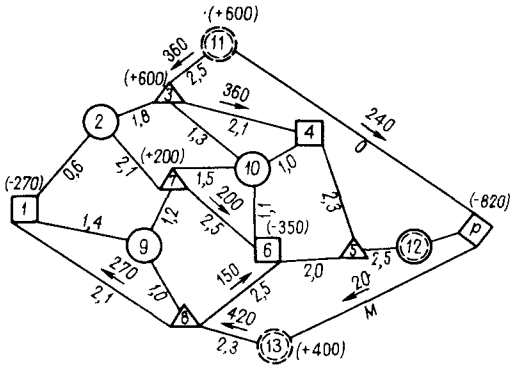


Рис. VIII. 10. Порядок округления производственной мощности карьеров до типовых

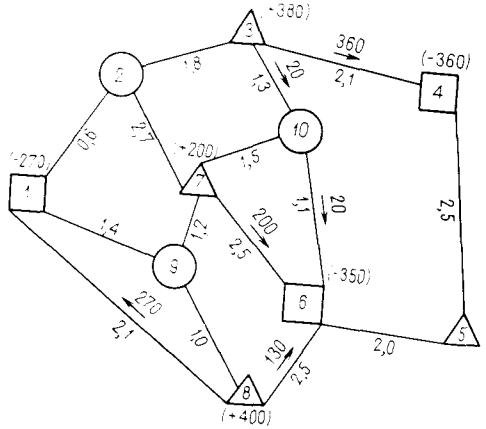


Рис. VIII. 11. Оптимальный план производства и поставок щебня

Алгоритм округления заключается в нахождении оптимального плана новой задачи.

Если округление до ближайшей типовой мощности осуществляется в сторону уменьшения (как в нашей задаче), то по ребру 1–3 вместо потока в 180 тыс. м^3 направляем в противоположную сторону разность между типовой и полученной мощностями. Причем затраты принимаются заведомо большими – M (см. рис. VIII.10). При решении этой задачи, естественно, этот поток должен исчезнуть, тогда получаем оптимальный план первоначальной нелинейной целочисленной задачи (рис. VIII.11).

В случае, если округление производится в сторону увеличения, то по ребру из соответствующего пункта производства в точку направляем с очень большими затратами поток M, равный величине прироста объема. После устранения этого потока в процессе решения новой задачи получается округленный план.

Согласно окончательному решению, из существующего карьера в пункте 7 должно поставляться 200 тыс. м^3 щебня потребителю 6. Кроме того, необходимо построить дробильно-сортировочные заводы в пунктах 3 и 8 мощностью 400 тыс. м^3, причем в пункте 3 мощность карьера будет недогружена на 200 тыс. м^3. Из карьера 3 поставку щебня целесообразно производить потребителям 4 (360 тыс. м^3) и 6 (20 тыс. м^3), а из карьера 8 – потребителям 1 (270 тыс. м^3) и 6 (130 тыс. м^3). Месторождение в пункте 5 вообще осваивать в рассматриваемый период нецелесообразно.

Интересно заметить, что если для первого плана задачи, приведенной к линейной (см. рис. VIII.7), общие затраты составили 3908 тыс. руб., то для оптимального плана нелинейной задачи (см. рис. VIII.8) – 4064 тыс. руб., а для окончательного оптимального решения нелинейной целочисленной задачи – 4066 тыс. руб., т.е. даже если окончательный план и является приближенным, то он отличается от затрат первого плана всего на 3,8 %, что вполне допустимо.

Контрольные вопросы и задачи

1. Каковы основные особенности нелинейных задач?

2. Для того чтобы точки x^k и x^{k+1}, полученные на следующих друг за другом итерациях градиентного метода, не совпадали, требуется, очевидно, чтобы градиент $\nabla f(x^k)$ не был нулевым вектором. Если же $\nabla f(x^k) = 0$, то $x^k = x^{k+1} = x^{k+2} = \ldots$, т.е. мы не сдвинемся из полученной точки x^k. Что же делать?

3. Может ли точка минимума квадратичной функции от n-мерного вектора быть найдена методом сопряженных направлений менее чем за n итераций? Если да, то приведите пример.

4. Методом координатного спуска найдите точку минимума функции $x^2 - 8x + 2yx - 8y + 16$.

5. Сформулируйте алгоритм одной итерации метода сопряженных направлений для минимизации квадратичной функции общего вида.

Пояснение. На луче, образованном векторами вида $x = x^k + \alpha p^k$, значение функции квадратично зависит от α. Следовательно, для нахождения величины α_k следует вычислить коэффициенты этого квадратного трехчлена и найти его максимум.

6. Очевидно, что любое ограничение-равенство (VIII.2) можно заменить на два неравенства: $g_i(x) \geqslant 0$ и $g_i(x) \leqslant 0$. Можно ли к преобразованной задаче применить метод возможных направлений? Почему?

7. Условие (VIII.4) в задаче определения направления спуска не задает ни условие "возможности" направления, ни условие убывания целевой функции. Следовательно, задача без этого условия также может служить для определения p^k. Так ли это и почему?

8. Может ли условие $|x_1 + x_3| \leqslant 5$ включаться в задачу, решаемую одним из перечисленных методов?

9. Нарисуйте на плоскости x, y "траекторию" (последовательность точек) метода возможных направлений для задачи

$$4x + y \to \max,$$
$$1 - (x-4)^2 \geqslant y,$$
$$(x-4)^2 - 1 \leqslant y,$$

10. Пусть $x^1_1 = 1$, $x^1_2 = 2$, $x^1_3 = 1$. Найти координаты точки x^2, построенной методом проекции градиента для задачи

$$x_1^2 + x_3^2 + 3x_2x_3 + x_2 \to \min,$$
$$x_1 + x_2 + x_3 = 4.$$

11. Найти значение N, при котором задача минимизации функции $-x^2 + 2yx - 2y^2 + N(x^2 + y^2 - 3)$ эквивалентна задаче: минимизировать $-x^2 + 2yx - 2y^2$ при условии $x^2 + y^2 - 3 \leqslant 0$.

12. Всякое условие-неравенство вида VIII.3, очевидно, можно записать в виде

$$\max\left\{g_i(x), 0\right\} = 0.$$

Левая часть этого соотношения есть непрерывная функция (попробуйте это доказать). Можно ли применить метод множителей Лагранжа после такого преобразования?

13. Пусть полученная на k-й итерации по методу штрафов точка x^k является допустимой в задаче. Что еще можно сказать об x^k?

14. В методе штрафов строят последовательность точек x^k. Тем не менее в задаче определения x^k не участвуют точки $x^1, \ldots, x^{k-1}$. Зачем нужно было их вычислять?

15. Преобразуйте задачу минимизации функции $x + \ln(y+1)^2$ при ограничениях

$$\max\left\{3 - y, 4x + 5y^2\right\} \leqslant 5,$$
$$\ln|x + 7y^2| \leqslant 7$$

к виду, пригодному для применения метода модифицированных функций Лагранжа, и запишите выражение для модифицированной функции Лагранжа.

16. Найти решение задачи: минимизировать $3x + y^2$ при ограничении $-1 \leqslant x \leqslant 1$ методом штрафных функций, полагая $r_k = 2^{k-1}$. За какое количество итераций решение получено?

17. На развитие трех шахт ($i = 1, 2, 3$) дополнительно выделены ресурсы. Эффективность использования капиталовложений по этим шахтам выражается зависимостью

$$g_i = A_i \sqrt{x_i},$$

где x_i – капиталовложения, выделенные i-й шахте, A_i – коэффициент функции эффективности для i-й шахты.

Суммарная эффективность использования ресурсов выражается целевой функцией

$$Z = \sum_i A_i \sqrt{x_i}.$$

Необходимо так распределить выделенные капитальные вложения, чтобы суммарная эффективность их использования была максимальной. Решить задачу с помощью неопределенных множителей Лагранжа.

Варианты решения задачи отличаются набором рассматриваемых предприятий и общим количеством ресурсов, выделенных всем трем предприятиям. Коэффициенты A функций эффективности, соответствующие конкретному предприятию, следующие.

№ предприятия	1	2	3	4	5	6	7	8	9	10
Коэффициент A	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0

Таблица VIII.2

№ варианта	№ предприятия	Общее количество ресурсов	№ варианта	№ предприятия	Общее количество ресурсов
1	1, 2, 3	14,0	14	1, 2, 5	300,0
2	1, 2, 3	2,1	15	1, 2, 6	41,0
3	2, 3, 5	3,8	16	4, 5, 7	3,6
4	2, 3, 6	9,8	17	4, 5, 8	21,0
5	2, 3, 7	12,4	18	4, 5, 9	6,1
6	2, 3, 4	58,0	19	4, 6, 7	10,1
7	2, 4, 5	90,0	20	4, 6, 8	11,6
8	2, 4, 6	14,0	21	4, 6, 9	2,66
9	2, 4, 7	138,0	22	4, 7, 8	25,8
10	2, 5, 7	39,0	23	5, 6, 7	55,0
11	2, 6, 7	8,9	24	4, 7, 9	73,0
12	1, 3, 4	78,0	25	4, 8, 9	161,0
13	1, 3, 5	7,0	26	5, 6, 8	62,5

Дополнительные данные, определяющие вариант задачи (номер предприятий и общее количество ресурсов) приведены в табл. VIII.2.

Глава IX

ДИНАМИЧЕСКИЕ ОПТИМИЗАЦИОННЫЕ МОДЕЛИ ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ ГОРНЫМ ПРОИЗВОДСТВОМ

§ 1. ОБЩАЯ ПОСТАНОВКА И ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ ДИНАМИЧЕСКИХ ЗАДАЧ ОПТИМИЗАЦИИ

В экономических задачах, задачах организации и управления производством управление обычно заключается в распределении на каждом этапе производственных заданий, ресурсов, финансов и т.д. При этом этапом планирования (решения задачи) является временной отрезок (год, месяц, сутки и т.д.), т.е., по существу, большинство задач планирования и управления являются динамическими.

Например, планируется работа объединения, включающего ряд производственных единиц $П_1$, $П_2$,..., $П_i$, ..., $П_k$ на период времени $t = 1, 2, ..., T$ лет. В начале периода (состояние S_0) система характеризуется определенным количеством оборотных и основных фондов, трудовыми ресурсами и т.д. На развитие объединения дополнительно выделено K средств, которое характеризует (определяет) конечное состояние системы S_k.

Необходимо так распределить дополнительные средства по годам и производственным единицам, чтобы к концу рассматриваемого периода T суммарный доход объединения W был бы максимальным.

Примем за управляемые переменные x_{it} – количество средств, выделенное i-му предприятию в год t. Распределение средств по производственным единицам в год t – $x_{1t}, x_{2t}, ..., x_{it}, ..., x_{kt}$ характеризует управление на этапе t – U_t. Совокупность управлений по годам выразится системой

$$\begin{cases} U_1(x_{11}, x_{21}, ..., x_{k1}), \\ U_2(x_{12}, x_{22}, ..., x_{k2}), \\ \quad \ldots\ldots\ldots\ldots \\ U_t(x_{1t}, x_{2t}, ..., x_{kt}), \\ \quad \ldots\ldots\ldots\ldots \\ U_T(x_{1T}, x_{2T}, ..., x_{kT}). \end{cases} \qquad \text{(IX.1)}$$

Каждому распределению средств по годам соответствует определенный доход w_t, причем общий доход W за T лет зависит от совокупности управлений, т.е.

$$W = f(U_1, U_2, ..., U_t, ..., U_T) = w_1 + w_2 + ... + w_t + ... + w_T = \sum_{t=1}^{T} w_t. \qquad \text{(IX.2)}$$

Итак, задача состоит в определении управления на каждом этапе, чтобы общий доход был максимален. Найдя оптимальное управление на каждом этапе $t = 1, 2, ..., T$ (в каждый год), мы найдем и оптимальное решение всей задачи.

В общем виде динамическая оптимизационная задача формулируется следующим образом.

Пусть некоторая система S находится в первоначальном состоянии $S_0 \in \widetilde{S}_0$, где $\widetilde{S}_0$ – область начальных состояний системы. Состояние системы S характеризуется какими-то параметрами (наличие фондов, ресурсов, мощность предприятия, запасы и т.д.), которые называются координатами системы. Состояние системы характеризуется точкой в n-мерном пространстве, где n – количество параметров. С течением времени под влиянием управления состояние системы меняется и в результате она переходит в конечное состояние $\widetilde{S}_k \in S_k$, где $\widetilde{S}_k$ – область конечных состояний системы. Переход системы из одного состояния в другое связан с выбором траектории перемещения точки, причем все траектории перемещения должны находиться в некоторой области допустимых состояний системы. Различным траекториям перемещения соответствуют разные эффективности. Необходимо таким образом спланировать процесс, выбрать такую траекторию перемещения системы из начального состояния в конечное, чтобы критерий эффективности достигал экстремального значения. Таким образом, в динамических задачах из множества допустимых управлений $U = (U_1, U_2, ..., U_t, ..., U_T)$ надо найти оптимальное, при котором система переходит из начального состояния $S_0 \in \widetilde{S}_0$ в конечное $S_k \in \widetilde{S}_k$ таким образом, что процесс протекает наиболее эффективно (т.е. критерий оптимальности W достигает максимума или минимума).

Общую математическую постановку динамической оптимизационной задачи с аддитивным критерием можно представить следующим образом

$$W = w(x_1, u_1) + ... + w(x_t, u_t) + ... + w(x_T, u_T) \to \min, \qquad (IX.3)$$

$$x_{t+1} = f(x_t, u_t), \quad t = \overline{1, T}, \qquad (IX.4)$$

$$\{x_t, u_t\} \in \widetilde{S}_t, \quad t = \overline{1, T}, \qquad (IX.5)$$

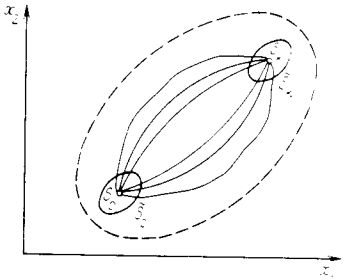
где x_t – вектор состояния, а u_t – вектор управления.

Выражение (IX.4) показывает, что последующее состояние x_{t+1} определяется предшествующим x_t и управлением u_t на данном этапе.

В частных случаях условия (IX.5) могут представлять ограничения на вектор управления $u_t \in U_t$ и вектор состояния $x_t \in X_t$.

Если состояние системы характеризуется двумя координатами (например, x_1 и x_2), то можно дать графическую интерпретацию, показанную на рис. IX.1. Областью возможных состояний системы является плоскость $x_1 O x_2$ или ее часть (на рисунке показана пунктирной линией). Управление изображается на плоскости линией, по которой точка S перемещается из начального состояния S_0 в конечное S_k. Таких траекторий перемещения точки S из S_0 в S_k – множество. Среди них надо выбрать такую, при которой критерий, характеризующий эффективность протекания процесса, максимизируется (минимизируется).

Рис. IX.1. Геометрическая интерпретация задачи динамического программирования



Таким образом, динамическая задача оптимального управления имеет следующую геометрическую интерпретацию. Найти такую траекторию перемещения точки S в фазовом пространстве, при которой перемещение точки из начального состояния S_0 в конечное S_k произойдет внутри области возможных состояний системы и критерий эффективности W примет оптимальное значение.

§ 2. ПРИНЦИП ОПТИМАЛЬНОСТИ, ОСНОВНОЕ ФУНКЦИОНАЛЬНОЕ УРАВНЕНИЕ И ПОРЯДОК РЕШЕНИЯ ЗАДАЧ МЕТОДОМ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Динамические оптимизационные задачи можно решать двумя путями: искать сразу оптимальное управление для всей системы (как в ранее рассмотренных методах линейного и нелинейного программирования) или же определять его постепенно, этап за этапом, оптимизируя каждый раз управления за один этап. Динамическое программирование базируется на втором принципе, т.е. на поэтапной оптимизации. При этом вместо решения сложной многовариантной задачи многократно (на каждом этапе) решаются более простые. Динамическое программирование используется при оптимизации как линейных, так и нелинейных задач. Причем часто для их решения не могут использоваться другие математические методы. Существенный недостаток динамического программирования – трудоемкость решения задач большой размерности.

Таким образом, в отличие от всех других методов математического программирования динамическое программирование представляет собой метод оптимизации многошаговых процессов, при котором оптимальное решение находится постепенно, поэтапно (по шагам). Наиболее естественное разделение процессов по этапам – их рассмотрение во вре-

мени. Этот способ разбивки по шагам наиболее характерен для задач планирования и управления производством. Однако возможны и другие варианты деления процесса на этапы: изменение координат объекта, мощности, ресурсов и т.д.

Динамическое программирование предназначено для оптимизации многошаговых управляемых процессов, т.е. когда на каждом этапе протекания процесса, принимая определенное решение, мы можем влиять на его результат. Совокупность решений, принятых на определенных этапах, и представляет (дает) управление процессом в целом. Таким образом, управление процессом состоит из ряда элементарных управлений на каждом этапе.

Следует заметить, что задачи динамического программирования решаются (за исключением отдельных частных случаев) тогда, когда критерий оптимальности аддитивен, т.е. представляет сумму показателей эффективности за отдельные этапы.

Суть метода динамического программирования состоит в постепенной, пошаговой оптимизации процесса. При этом оптимум на каждом этапе должен определяться не только с позиций данного этапа, а с учетом эффективности процесса в целом. Действительно, какой смысл в оптимальном управлении на данном этапе, если это ведет к резкому ухудшению показателей на других. Сумма локальных оптимумов в результате независимой (без учета последствий на других этапах) оптимизации на каждом этапе не только не дает глобального оптимума, но и ведет к серьезным ошибкам. Например, такой подход при планировании горных работ может привести к первоочередной хищнической отработке богатых участков месторождений, к недопустимому обеднению и обесцениванию остающихся запасов полезного ископаемого.

Таким образом, при пошаговой оптимизации динамических процессов управление на каждом шаге необходимо выбирать с учетом перспективы и последствий на предстоящих этапах, т.е. с учетом интересов процесса в целом.

Однако в каждом процессе имеется этап, который можно планировать без учета перспективы, ибо решение на нем не влияет на остальные. Обычно это – последний этап. Поэтому на нем принимают решение только исходя из принципов данного этапа, т.е. решение, позволяющее получить максимальный эффект.

После того, как последний T-й этап спланирован, к нему можно добавить $(T-1)$-й этап, к которому затем – $(T-2)$ и так далее до начального. Поэтому обычно задача динамического программирования рассчитывается от конца к началу.

Управление на последнем шаге выбирают исходя из различных предположений об исходе предпоследнего этапа (о положении системы к началу последнего этапа). Для планирования следующего, предпоследнего $(T-1)$ этапа надо сделать различные предположения об исходе предпоследнего $(T-2)$ этапа и т.д. Оптимальное управление на этапах находится для

каждого из возможных состояний системы на начало этапа. Такое управление называется условно-оптимальным; при достижении определенного предшествующего состояния оно является оптимальным.

Таким образом, если процесс в целом разбит на T этапов, то делаем предположения о состояниях системы на предпоследнем этапе $S_{T-1}^{(1)}$, $S_{T-1}^{(2)}$, ..., $S_{T-1}^{(k)}$ и для каждого из них находим условно-оптимальные управления $u_T^{(1)}$, $u_T^{(2)}$, ..., $u_T^{(k)}$. Таким образом, планирование последнего T-го этапа завершено. Действительно, как бы ни закончился предпоследний $(T-1)$-й этап, мы знаем, какое управление надо использовать на последнем. Далее делаем предположения об исходе предпоследнего $(T-2)$-го этапа и с учетом найденных условно-оптимальных уравнений на последнем T-м этапе определяем для каждого из предшествующих состояний $S_{T-2}^{(1)}$, $S_{T-2}^{(2)}$, ..., $S_{T-2}^{(m)}$ условно-оптимальные управления $u_{T-2}^{(1)}$, $u_{T-2}^{(2)}$, ..., $u_{T-2}^{(m)}$ и т.д. Аналогичным образом поступает вплоть до начального состояния системы. Если начальное состояние известно (задано), то для первого этапа предположения о возможных состояниях системы не делаются.

Таким образом, проходя последовательно от последнего этапа к первому, получаем совокупность условно-оптимальных управляющих воздействий

$$\begin{cases} u_T^{(1)}, u_T^{(2)}, \ldots, u_T^{(k)}, \\ u_{T-1}^{(1)}, u_{T-1}^{(2)}, \ldots, u_{T-1}^{(m)}, \\ \ldots\ldots\ldots\ldots\ldots \end{cases}$$

$$u_t^{(1)}, u_t^{(2)}, \ldots, u_t^{(n)},$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$u_1^{(1)}, u_1^{(2)}, \ldots, u_1^{(l)}.$$

Найденные условно-оптимальные управления позволяют определить для любых возможных состояний оптимальное продолжение процесса. Поэтому, начиная с начального состояния S_0, можно определить уже не условно-оптимальное, а просто оптимальное управление на каждом этапе. Для этого, используя условно-оптимальное управление, найденное для начального состояния S_0, находим состояние, в которое перейдет система S_1. Применяя условно-оптимальное управление для данного состояния, получаем новое состояние системы. Но так как известны условно-оптимальные управления для всех возможных состояний системы, постепенно находим совокупность оптимальных управлений на каждом шаге и, следовательно, оптимальное управление процессом в целом $U = (U_1, U_2, \ldots, U_t, \ldots, U_T)$, дающее наибольший эффект W.

Таким образом, при решении задачи методом динамического программирования первоначальные вычисления проводятся от последнего этапа к начальному, а затем в обратном направлении.

На первой, наиболее трудоемкой стадии расчетов находятся условно-оптимальные управления на каждом этапе и значения критерия эффективности, им соответствующие. На второй стадии (от начала к концу процесса) постепенно строится система не условных, а оптимальных поэтапных управлений.

В основе расчетов методом динамического программирования лежит принцип оптимальности, впервые сформулированный американским математиком Р. Беллманом: оптимальное управление обладает тем свойством, что каковы бы ни были достигнутые состояния и решения до данного момента, последующее решение должно составлять оптимальное поведение относительно состояния, достигнутого на данный момент.

Символически метод динамического программирования можно записать следующим образом.

Обозначим через $W_t(S)$ условно-оптимальный выигрыш (значение критерия) с t-го этапа до конца процесса, если система в начале находится в состоянии S; $u_t(S)$ – условно-оптимальное управление на t-м шаге, которое совместно с оптимальными управлениями на последующих этапах дает нам условно-оптимальный выигрыш; w_t – выигрыш на одном t-м этапе, w_t зависит от начального состояния и управления на данном этапе.

Под влиянием управления u_t система на t-м шаге переходит из состояния S в состояние (S').

Таким образом S' – производное от начального состояния S_0 управления u_t

$$S' = f(S, u_t).$$

Условно-оптимальный выигрыш выражается следующим образом

$$W_t(S) = w_t(S, u_t) + W_{t+1}(S'), \qquad \text{(IX.6)}$$

или

$$W_t(S) = w_t(S, u_t) + W_{t+1}(f(S, u_t)). \qquad \text{(IX.6а)}$$

Первое слагаемое представляет значение критерия на данном этапе, а второе – от данного этапа до конца процесса.

В соответствии с принципом оптимальности на каждом шаге для достигнутого состояния S должны выбирать такое управление u_t, при котором критерий оптимальности (условно-оптимальный выигрыш) максимизируется.

Тогда основное функциональное уравнение динамического программирования, представляющее собой рекуррентное соотношение, имеет вид

$$W_t(S) = \max[w_t(S, u_t) + W_{t+1}(f(S, u_t))]. \qquad \text{(IX.7)}$$

Рекуррентное соотношение позволяет определить функцию $W_t(S)$, если известна следующая по порядку функция $W_{t+1}(S')$. Причем при определении максимума $W_t(S)$ учитываются только допустимые состояния системы.

Первоначально находятся условно-оптимальный выигрыш $W_t(S)$ на последнем шаге и соответствующее ему условно-оптимальное управление u_t:

$$W_T(S) = \max_{u_T(S)} \{w_T(S, u_T)\}. \qquad \text{(IX.8)}$$

Зная $W_T(S)$ и u_T и используя основное функциональное уравнение, находим W_{T-1} и u_{T-1}. Управление u_T, при котором максимизируется $W_T(S)$, представляет условно-оптимальное управление. Подобная процедура повторяется вплоть до первого этапа.

Затем переходят ко второй стадии расчетов – нахождению оптимального управления $U = (U_1, U_2, ..., U_t, ..., U_T)$.

Начинают с первого шага. Для начального состояния S_0 определяют $W_{\max} = W_1(S_0)$ и управление $U_1 = U_1(S)$, переводящее систему в состояние S_1, ибо $S_1 = f(S_0, u_1)$. Зная состояние S_1, находят оптимальное управление на втором шаге $U_2 = U_2(S_1)$ и затем $S_2 = f(S_1, U_2)$ и так последовательно получают шаговые оптимальные управления и поэтапные состояния системы по следующей цепочке

$$S_0 \to u_1(S_0) \to S_1 \to u_2(S_1) \to \ldots \to S_{t-1} \to u_t(S_{t-1}) \to S_t.$$

Если начальное состояние задано не точкой, а областью S_0, то на первом этапе принимают за основу несколько начальных положений, а далее оптимизация проводится в описанном порядке.

Решение задачи динамического программирования состоит из следующих этапов.

1. Выбирают способы описания процесса (параметры, характеризующие систему) и деления на шаги (этапы).

2. Вычисляют значение критерия на t-м шаге в зависимости от управления $W_t(S, u_t)$.

3. Записывают функцию изменения состояния системы при переходе от состояния S под влиянием управления u_t

$$S' = f(S, u_t).$$

4. Составляют основное функциональное уравнение

$$W_t(S) = \max[w_t(S, u_t) + W_{t+1}(f(S, u_t))].$$

5. Находят значение критерия на последнем шаге

$$W_T(S) = \max\{w_T(S, u_T)\}.$$

6. Зная значения критерия на последнем шаге и используя функциональное уравнение, находят последовательно одно за другим значе-

ния критерия и условно-оптимальное управление на последующих шагах $u_{T-1}(S), u_{T-2}(S), \ldots, u_1(S)$.

7. Для начального состояния S_0 находят оптимальный выигрыш $W_1(S_0)$ и далее по цепочке систему безусловно-оптимальных управлений:

$$S_0 \to U_1(S_0) \to S_1 \to U_2(S_1) \to \ldots \to S_{T-1} \to U_T(S_{T-1}) \to S_T.$$

Следует отметить, что в ряде случаев при использовании динамического программирования процесс решения может разворачиваться и в прямом направлении (на первой стадии от начала к концу, а на второй – от конца к началу).

§ 3. ЗАДАЧА ОБ ОПРЕДЕЛЕНИИ ОПТИМАЛЬНОЙ ТРАЕКТОРИИ ПЕРЕМЕЩЕНИЯ СИСТЕМЫ И ЕЕ РЕШЕНИЕ МЕТОДОМ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Наглядный пример использования метода динамического программирования – задача об определении траектории изменения состояния системы, к которой сводится большое количество прикладных задач. При этом координатами системы могут являться различные параметры (пространственные координаты, технические или технико-экономические характеристики).

Для простоты первоначально рассмотрим случай, когда состояние системы характеризуется двумя параметрами (координатами) и изменение ее происходит за шаг (этап) по одному из параметров (только по вертикали или горизонтали).

Пусть система (рис. IX.2) из начального состояния (точка A) должна перейти в конечное (точка B), причем каждой траектории перемещения соответствует определенный расход средств. Все множество траекторий, по которым система перемещается из начального состояния в конечное, лежит в пределах прямоугольника $ACBD$.

Прежде всего необходимо разбить все пространство (весь процесс изменения состояний системы) на этапы. Так как для упрощения принято, что изменение координат объекта происходит только по горизонтали или вертикали, то для выделения этапов разделим интервал изме-

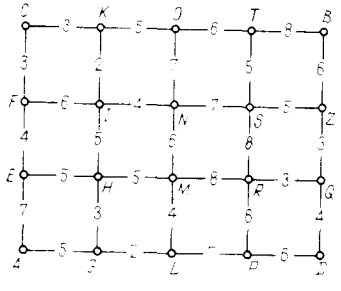


Рис.IX.2. Определение траектории перемещения системы

нения координат x_1 и x_2 соответственно на n и m равных частей. Общее количество этапов изменения состояний системы составит $n + m$ (в нашем случае $3 + 4 = 7$). При этом количество выделенных этапов принципиального значения не имеет и определяется требуемой точностью расчетов.

Выделив этапы, получим все возможные промежуточные состояния системы, которые на рис. XI.2 изображены точками A, E, F, C, G, H, I, K, L, M, N, O, P, R, S, T, D, Q, Z, B. Для оптимизации процесса необходимо знать значение критерия эффективности на каждом этапе, т.е. расход средств при переходе из одного состояния в другое. Предположим, что расходы определены и записаны (в условных единицах) на соответствующих элементах траектории (см. рис. IX.2).

Хотя данная задача невелика по объему и может быть в принципе решена перебором вариантов, использование метода динамического программирования проще и эффективнее. Следует заметить, что с увеличением размерности задачи метод перебора вариантов становится громоздким и не может быть применен.

Оптимизацию начинаем с последнего этапа. В конечное состояние системы (точка B) можно перейти из точки T или Z, при этом для каждой из них траектория единственная, что и предопределяет условно-оптимальное управление. Если на предпоследнем (шестом) этапе мы попали в точку T, то в последний этап изменение состояния системы (управление) происходит по горизонтали, а из точки Z – по вертикали. Соединим точки T и Z с точкой B дугами (рис. IX.3), в кружках, соответствующих этим точкам, проставим значения критерия (условно-оптимальный выигрыш от данного положения до конца процесса). Запись цифры 8 в кружке у точки T означает, что если мы попали в точку T, то минимальный расход средств до конца процесса равен 8, а стрелка из точки T в точку B показывает, каким образом изменяется состояние системы.

Найдя условно-оптимальные управления на последнем (седьмом) этапе, переходим к планированию процесса на предпоследнем (шестом) этапе. Для этого рассмотрим все возможные состояния системы на начало шестого, конец предшествующего – пятого этапа. После пятого этапа можно оказаться в одной из точек O, S, Q (см. рис. IX.3). Для каждой из этих точек должны найти путь в конечное состояние (точка B). Для точки O управление единственно, система должна перемещаться в точку T и значение критерия при этом составит $6 + 8 = 14$ усл. ед. Подобная картина наблюдается для точки Q. Здесь также переход системы единственен (в точку Z), а значение критерия составляет $3 + 6 = 9$ усл.ед. Значение критерия запишем в соответствующих кружках у данных точек, а направления перехода (условно-оптимальное управление) покажем дугами (см. рис. IX.3).

Для точки S есть выбор управления. Из нее в точку B можно попасть через точки T или Z. В первом случае расход средств составит $5 + 8 = 13$ единиц, а во втором $5 + 6 = 11$ единиц. Естественно, что выбираем

траекторию с минимальным расходом средств и, следовательно, условно-оптимальным управлением из точки S будет переход в точку Z. Это управление показано стрелкой на рис. IX.3, а минимальное значение критерия (11) в кружке у точки S.

Аналогичным образом находятся условно-оптимальное управление и условно-оптимальный выигрыш для всех оставшихся этапов (точек) вплоть до начального состояния (точка A). Для этого, переходя от точки к точке и от конца процесса к его началу (справа—налево и сверху—вниз), последовательно для каждой точки находятся условно-оптимальное управление (изображаемое дугой, соединяющей данную точку с соседней) и минимальный расход средств на переход от данной точки в конечное состояние (записываемый в кружке у данной точки). Для выбора условно-оптимального управления сравниваются значения критерия для двух путей (направо и вверх), причем критерий определяется как сумма расходов средств на переход в соседнюю точку и из соседней точки в конечное состояние. Последняя величина определена при оптимизации процесса на предыдущем этапе расчетов. Выполнив указанную процедуру расчетов для всех точек, доходим до начального состояния системы (точка A). Ход и результаты расчетов для рассматриваемого примера показаны на рис. IX.3.

Закончив этап условной оптимизации (процедуру расчетов от конца к началу), переходим к этапу определения просто оптимального управления. Для этого, начиная от начальной точки, по стрелкам строим траекторию перемещения системы. На рис. IX.4 оптимальная траектория выражается непрерывной линией от точки A к точке B. Эта оптимальная траектория проходит через точки A, G, L, M, R, Q, Z и B. Минимальный расход средств при этом равен 31.

Интересной особенностью решения задач методом динамического программирования является определение условно-оптимальных управлений для всех возможных состояний системы. На рис. IX.4 условно-оптимальное управление можно получить, перемещаясь по стрелкам, ведущим из данной точки в конечное состояние. Поэтому, если по ка-

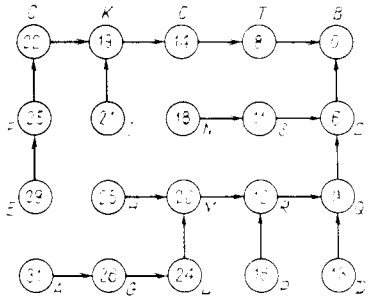


Рис.IX.3. Условно-оптимальные и оптимальная траектории перемещения системы

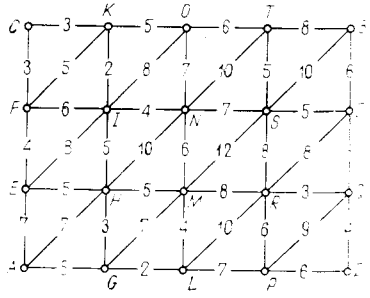
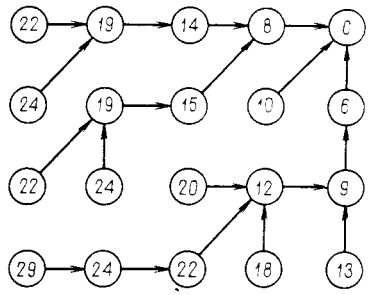


Рис.IX.4. Определение траектории перемещения системы

Рис.IX.5. Ход решения задачи и оптимальная траектория перемещения системы



кой-либо причине ход процесса отклонится от оптимального, а ожидаемые характеристики процесса в будущем (значение критерия на оставшихся шагах) не изменятся, мы, не производя дополнительных расчетов, имеем оптимальную траекторию для любой точки, в которую может прийти система.

В рассмотренной задаче процесс изменения траектории предельно упрощен. Было допущено, что одновременно происходит изменение только одной из координат состояния объекта (перемещение системы производится только по горизонтали или вертикали). Если же одновременно возможно изменение двух координат, то ход решения задачи остается без изменений, лишь несколько увеличивается объем вычислений.

Допустим, что в рассмотренной задаче помимо движения системы параллельно одной из координатных осей возможно и перемещение по диагонали, которому соответствует изменение обеих координат. Расходы средств при движении по диагонали показаны на рис. IX.4. Приведенная задача решается аналогично рассмотренной ранее. Начиная с конечной точки B последовательно находятся условно-оптимальные уравнения для всех точек. Сначала для точек T, S, Z, затем Q, R, M, N, O и так далее вплоть до начальной точки A. Отличие заключается лишь в том, что при выборе условно-оптимального управления для точек, не лежащих на границе области допустимых состояний системы (прямых AC, CB, BD, AD), сравниваются не два возможных пути (направо и вверх), а три (еще и по диагонали). Ход решения задачи показан на рис. IX.5. Оптимальная траектория проходит через точки A, G, L, R, Q, Z, B, а минимальный расход средств при этом составляет 29 условных единиц.

§ 4. ИСПОЛЬЗОВАНИЕ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ ДЛЯ РЕШЕНИЯ СТАТИЧЕСКИХ ЗАДАЧ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ

Статические задачи распределения ресурсов возникают при одноразовом распределении ограниченных ресурсов X (капитальные вложения, сырье, оборудование, фонды и т.д.) между несколькими объектами

$i = 1, 2, ..., n$ (отраслями, объединениями, предприятиями, технологическими процессами и т.д.). Эффективность использования ресурсов по объектам $q_i(x_i)$ различна и для каждого объекта зависит только от объема выделенных ресурсов x_i. Необходимо таким образом распределить ресурсы между объектами, чтобы суммарная эффективность их использования была максимальной.

Математическая модель задачи имеет вид

$$\sum_{i=1}^{n} q_i(x_i) \to \max, \quad (i = 1, 2, .., n), \tag{IX.9}$$

$$\sum_{i=1}^{n} x_i = X, \quad x_i \geqslant 0. \tag{IX.10}$$

При изучении линейного программирования мы решали подобные задачи при линейной целевой функции и линейных ограничениях. Ряд задач распределения ресурсов (например, при аналитическом задании непрерывных целевых функций) может быть решен методами нелинейного программирования. Динамическое программирование позволяет решать статические задачи распределения ресурсов практически для любого класса целевых функций, в том числе и для целочисленных, разрывных, недифференцируемых.

Приведем решение данной задачи в общем виде и на конкретных примерах.

При решении статической задачи распределения ресурсов методом динамического программирования за этап (шаг) принимается выделение ресурсов дополнительно одному объекту.

Решение задачи начинается с одного объекта. При этом находятся условно-оптимальное управление и значение критерия при выделении определенного количества ресурсов этому объекту.

Обозначим значение критерия (условно-оптимального выигрыша) при распределении X ресурсов по k объектам $W_k(X)$.

Тогда на первом этапе решения задачи получим

$$W_1(x_1) = \max q_1(x_1) = q_1(x_1).$$

На втором этапе определяем условно-оптимальное распределение ресурсов при их выделении на два объекта. При общем количестве ресурсов X второму объекту выделяется x_2, а первому, следовательно, $(X - x_2)$.

Тогда $W_2(X) = W_2(x_1, x_2) = q_2(x_2) + W(X - x_2)$.

Значение критерия $W_2(X)$ зависит только от одной переменной X (величина X характеризует состояние системы, ее значение принимается постоянным и задается вариантами). Поэтому можно отыскивать оптимум $W_2(X)$ в интервале $(0, X)$, причем этот оптимум определит условно-оптимальное управление (x_1, x_2) и условно-оптимальный выигрыш на

втором этапе $W_2(X) = \max W_2(x_1, x_2) = \max\{q_2(x_2) + W_1(X - x_2)\}$, $X \geqslant x_2 \geqslant 0$.

Рассуждая аналогично, получим основное функциональное уравнение динамического программирования для данной задачи

$$W_k(X) = \max W_k(x_1, x_2, ..., x_k) = \max\{q_k(x_k) + W_{k-1}(X - x_k))\},$$

$$X \geqslant x_k \geqslant 0.$$

Пользуясь основным функциональным уравнением, постепенно находят условно-оптимальные значения x_k вплоть до подключения всех объектов $(k = 1, 2, ..., n)$.

Соответствующее значение x_n, максимизирующее $W_n(X)$, дает безусловно оптимальное управление. Далее, возвращаясь к $(n - 1)$-му этапу и зная x_n, определяем оптимальное значение x_{n-1} и так, постепенно, дойдя до первого этапа, найдем совокупность $x_n, x_{n-1}, ..., x_k$, дающую максимальный эффект от распределения X ресурсов $(\max W_n)$.

Приведем примеры решения рассмотренной статической задачи распределения ресурсов при аналитическом и табличном задании функций эффективности.

Пример IX.1. Рассмотрим задачу распределения капитальных вложений (всего 29 млн. руб.) между тремя шахтами. Эффективность их использования на каждой из них составляет соответственно $q_1(x_1) = 0{,}2\sqrt{x_1}$, $q_2(x_2) = 0{,}3\sqrt{x_2}$; $q_3(x_3) = 0{,}4\sqrt{x_3}$. Ранее (см. § 4 гл. VIII, где дана математическая модель задачи) указанная задача решалась методом множителей Лагранжа. Продемонстрируем теперь возможность и эффективность использования для ее решения метода динамического программирования.

На первом этапе решения рассматриваем распределение ресурсов одной (первой) шахты. Очевидно, условно-оптимальное управление здесь состоит в полном использовании имеющихся ресурсов (так как $q_1(x_1)$ возрастает с увеличением x_1), а значение критерия

$$W_1(x_1) = \max q_1(x_1) = q_1(x_1) = 0{,}2\sqrt{x_1}.$$

Далее переходим ко второму этапу, на котором рассматриваем распределение ресурсов между двумя (первой и второй) шахтами. Эффективность использования X ресурсов на этих шахтах составляет

$$W_2(X) = q_2(x_2) + W_1(x_1) = q_2(x_2) + q_1(X - x_2) = 0{,}3\sqrt{x_2} + 0{,}2\sqrt{X - x_2}.$$

Для определения условно-оптимального распределения ресурсов между шахтами найдем оптимум функции $W_2(X)$. Для этого возьмем производную $\frac{dW_2}{dx_2}$ и приравняем ее к нулю

$$\frac{dW_2}{dx_2} = \frac{0{,}3}{2\sqrt{x_2}} - \frac{0{,}2}{2\sqrt{X - x_2}} = 0.$$

Откуда

$$3\sqrt{X - x_2} = 2\sqrt{x_2}; \qquad 9(X - x_2) = 4x_2;$$

$$x_2 = \frac{9}{13}X;$$

$$x_1 = X - x_2 = \frac{4}{13}X.$$

Таким образом, условно-оптимальное распределение между двумя шахтами состоит в том, что при любом количестве ресурсов, выделяемых на первую и вторую шахты, их следует делить в отношении $x_1 : x_2 = 4 : 9$. При этом максимальная эффективность использования ресурсов составит

$$\max W_2(X) = 0{,}3\sqrt{\frac{9}{13}X} + 0{,}2\sqrt{\frac{4}{13}X} = 1{,}3\sqrt{\frac{X}{13}} = 0{,}1\sqrt{13X}.$$

Переходим к третьему, для данной задачи последнему, этапу.

Рассмотрим распределение ресурсов между тремя шахтами. Используя полученное условно-оптимальное значение критерия при распределении ресурсов между двумя шахтами, составим зависимость для расчета суммарной эффективности использования ресурсов по объединению

$$W_3(X) = q_3(x_3) + W_2(X - x_3) = 0{,}4\sqrt{x_3} + 0{,}1\sqrt{13(X - x_3)}.$$

Для определения максимума x_3 (условно-оптимальное управление на третьем этапе) возьмем производную $\frac{dW_3}{dx_3}$ и приравняем ее к нулю

$$\frac{dW_3}{dx_3} = \frac{0{,}4}{2\sqrt{x_3}} - \frac{0{,}1\sqrt{13}}{2\sqrt{(X - x_3)}}; \quad 0{,}4\sqrt{(X - x_3)} = \sqrt{13x_3};$$

$$4^2(X - x_3) = 13x_3;$$

$$x_3 = \frac{16}{29}X.$$

Следовательно, $x_1 + x_2 = \frac{13}{29}X$, т.е. третьей шахте надо выделить 16/29 всех ресурсов, а первой и второй 13/29.

Суммарный эффект при этом составит

$$\max W_3(X) = 0{,}4\sqrt{\frac{16}{29}X} + 0{,}1\sqrt{13\,\frac{13}{29}X} = 0{,}1\sqrt{29X}.$$

Переходим ко второй стадии расчетов, решаем задачу в противоположном направлении.

Третьей шахте надо выделить 16/29 всех капитальных вложений, или 16 млн. руб. Первой и второй шахтам следует выделить 13/29 всех капитальных вложений, или 13 млн. руб. На втором этапе (шаге) оптимизации установлено, что оптимально имеющиеся ресурсы между первой и второй шахтами следует распределять в пропорции $x_1 : x_2 = 4 : 9$. Следовательно, на развитие первой шахты надо выделить 4 млн. руб., а второй – 9 млн. руб.

Аналогичный ответ был получен в § 4 гл. VIII при решении задачи методом неопределенных множителей Лагранжа. В целом же при любом другом количестве ресурсов их необходимо распределять между шахтами в пропорции $x_1 : x_2 : x_3 = 4/13\ 13/29 : 9/13\ 13/29 : 16/29 = 4 : 9 : 16$.

Пример IX.2. Горно-обогатительному комбинату дополнительно выделено 8 комплектов горного оборудования, которое надо распределить между тремя

Таблица IX.1

Количество дополнительно выделяемого оборудования x_i	Прирост добычи по рудникам		
	№ 1 $q_1(x_1)$	№ 2 $q_2(x_2)$	№ 3 $q_3(x_3)$
1	2	3	4
1	1,2	1,1	1,2
2	2,1	2,1	2,2
3	3,0	3,1	3,2
4	–	3,9	4,0

рудниками таким образом, чтобы общий прирост добычи был максимален. По горнотехническим условиям на руднике № 1 можно дополнительно установить не более 3 комплектов оборудования, а на № 2 и № 3 – не более 4 (на каждом отдельно). Прирост добычи по рудникам (в условных единицах) в зависимости от количества дополнительно выделяемого оборудования (соответственно x_1, x_2, x_3) рассчитан методом вариантов и задан таблично (табл. IX. 1).

Математическая модель рассматриваемой задачи имеет вид

$$\sum_{i=1}^{3} q_i(x_i) = q_1(x_1) + q_2(x_2) + q_3(x_3) \to \max;$$

$$x_1 + x_2 + x_3 = 8; \quad 0 \leqslant x_1 \leqslant 3; \quad 0 \leqslant x_2 \leqslant 4; \quad 0 \leqslant x_3 \leqslant 4.$$

Первый этап решения задачи методом динамического программирования, как и в ранее рассмотренном примере, начинается с определения эффективности использования ресурсов на одном объекте. Рассмотрим выделение ресурсов руднику № 1. Эффективность использования ресурсов при различных состояниях системы выражается при этом данными табл. IX.1 (столбец 2).

На втором этапе рассмотрим распределение ресурсов между двумя (№ 1 и № 2) рудниками. Сначала определим состояния системы, характеризуемые количеством ресурсов, выделяемым на два рудника ($X = x_1 + x_2$). Максимальное количество ресурсов, которое можно выделить на первые два рудника, составит $(x_1^{\max} + x_2^{\max} = 3 + 4) = 7$, а минимальное $(x_1 + x_2 + x_3 - x_3^{\max} = 8 - 4) = 4$. Таким образом, на первые два рудника можно выделить 4, 5, 6 или 7 комплектов оборудования. Далее для каждого из этих состояний системы определяем возможные варианты распределения оборудования по рудникам. Например, при $X = x_1 + x_2 = 4$ возможны следующие варианты: 1) $x_1 = 0, x_2 = 4$; 2) $x_1 = 1, x_2 = 3$; 3) $x_1 = 2, x_2 = 2$; 4) $x_1 = 3, x_2 = 1$.

Все состояния системы и возможные варианты распределения оборудования заносим в табл. XI.2 (столбцы 1, 2, 3). В следующих столбцах этой таблицы даем составляющие слагаемые основного функционального уравнения, т.е. значения эффективности использования дополнительно выделяемого оборудования на каждом руднике и суммарную по обоим рудникам.

Для каждого из возможных состояний системы находим условно-оптимальное управление. Для этого в столбце 6 выбираем максимальное значение критерия,

Таблица IX.2

$x=x_1+x_2$	x_1	x_2	$q_1(x_1)$	$q_2(x_2)$	$W_2(X)=q_1(x_1)+q_2(x_2)$
1	2	3	4	5	6
4	0	4	0	3,9	3,9
	1	3	1,2	3,1	**4,3**
	2	2	2,1	2,1	4,2
	3	1	3,0	1,1	4,1
5	1	4	1,2	3,9	5,1
	2	3	2,1	3,1	**5,2**
	3	2	3,0	2,1	5,1
6	2	4	2,1	3,9	6,0
	3	3	3,0	3,1	**6,1**
7	3	4	3,0	3,9	6,9

которое соответственно составляет (при $X=4, 5, 6, 7$) 4,3; 5,2; 6,1; 6,9. Полученное максимальное значение критерия используем на следующем этапе решения задачи.

Рассмотрим теперь распределение 8 комплектов оборудования между тремя рудниками. Все расчеты сведем в табл. IX.3. Причем в столбце 5 табл. IX.3 значения критерия эффективности при выделении оборудования рудникам № 1 и 2 берутся по результатам оптимизации на предыдущем этапе (из столбца 6 табл. IX.2).

В столбце 6 находим значение критерия, соответствующее не только условно, но и просто оптимальному распределению оборудования.

Максимальная суммарная эффективность (8.4) достигается при выделении руднику № 3 комплектов оборудования, а рудникам № 1 и 2 – 5 комплектов.

Таблица IX.3

$X=x_1+x_2+x_3$	x_3	x_1+x_2	$q_3(x_3)$	$W_2(x_1, x_2)$	$W_3(x_1, x_2, x_3)$
1	2	3	4	5	6
8	4	4	4,0	4,3	8,3
	3	5	3,2	5,2	8,4
	2	6	2,2	6,1	8,3
	1	7	1,2	6,9	8,1

Возвращаясь назад, из данных табл. IX.2 находим распределение ресурсов между рудниками № 1 и 2, если $x_1+x_2=5$. В столбце 6 находим, что при $x_1+x_2=5$ максимальное значение критерия 5,2 соответствует выделению на рудник № 1 2 комплектов оборудования, а на № 2 – 3. Таким образом, оптимальное распределение оборудования следующее: $x_1=2$, $x_2=3$, $x_3=3$, максимальный прирост добычи при этом составит 8,4 условных единиц.

Табличное решение задачи методом динамического программирования применяется чаще аналитического при оптимизации широкого круга задач распределения ресурсов, управления запасами, замены и ремонта оборудования и т.д.

§ 5. ДИНАМИЧЕСКАЯ ЗАДАЧА РАСПРЕДЕЛЕНИЯ РЕСУРСОВ

Преимущества метода динамического программирования наглядно проявляются при решении динамических задач планирования распределения ресурсов между несколькими объектами управления за ряд лет. Причем эффективность использования ресурсов на объектах различна и в общем случае меняется во времени. В процессе работы ресурсы частично уменьшаются (тратятся, амортизируются). Функции изменения количества ресурсов при их использовании на различных объектах известны. На начало каждого нового планового периода остающиеся ресурсы могут перераспределяться между объектами. Задача заключается в определении количества ресурсов, выделяемых каждому объекту в различные периоды времени таким образом, чтобы общая (за ряд лет по всем объектам) эффективность использования ресурсов была максимальной.

Рассмотрим простейшую динамическую задачу распределения ресурсов между двумя объектами управления.

Необходимо в течение T-периодов времени ($t=1, 2, ..., T$) распределить между двумя объединениями ресурсы в количестве A_0. Ресурсы, вложенные в определенный объект к концу периода, приносят прибыль, равную на 1-м объекте $q_1(x_{1t})$, а на 2-м – $q_2(x_{2t})$. Здесь и далее x_1t и x_{2t} – управляемые переменные задачи, обозначающие соответственно количество ресурсов, выделенное 1-му и 2-му объединениям в период t. В результате использования количество ресурсов постоянно уменьшается на 1-м объекте до величины $y_1(x_{1t})$, а на 2-м – $y_2(x_{2t})$. По истечении каждого периода оставшееся количество ресурсов может вновь перераспределяться между объектами. Требуется распределить ресурсы между объектами (объединениями) по периодам времени, т.е. найти количество ресурсов, выделяемое ежегодно (если плановый период – год) на каждый объект управления, чтобы суммарная прибыль была максимальна.

Оптимизируемая система в данном случае включает два объединения. Состояние системы характеризуется количеством ресурсов, выделенных каждому объединению – x_{1t} и x_{2t}.

Управление U_t на шаге t состоит в определении количества ресурсов x_{1t} и x_{2t}, выделяемых 1-му и 2-му объектам.

Управление всей операцией включает совокупность поэтапных управлений

$$U = (U_1, U_2, ..., U_t, ..., U_T) = (x_{11}, x_{21}; x_{12}, x_{22}; ..., x_{1t}, x_{2t}; ..., x_{1T}, x_{2T}).$$

Требуется найти оптимальное управление U_{opt}, максимизирующее суммарную прибыль

$$W = \sum_{t=1}^{T} w_t = \sum_{t=1}^{T} [q_1(x_{1t}) + q_2(x_{2t})].$$

Состояние системы перед шагом t характеризуется суммарным количеством ресурсов A_t, сохранившихся после предшествующих $(t-1)$ шагов.

Управление на этапе состоит в определении одной из переменных x_{1t} или x_{2t}, поскольку, зная A_t, вторая легко определяется (например, $x_{2t} = A_t - x_{1t}$).

Прибыль на этапе t зависит от суммарного количества и распределения ресурсов и равна

$$w_t(A_t, x_{1t}) = q_1(x_{1t}) + q_2(A_t - x_{1t}).$$

Остаток ресурсов после завершения этапа t составит

$$A_{t+1} = y_1(x_{1t}) + y_2(A_t - x_{1t}).$$

Зная зависимости для расчета прибыли на этапе и суммарной, а также для определения остатка ресурсов, получим основное функциональное уравнение задачи

$$W_t(A_t) = \max\left\{q_1(x_{1t}) + q_2(A_t - x_{1t}) + W_{t+1}[y_1(x_{1t}) + y_2(A_t - x_{1t})]\right\}, \quad 0 \leqslant x_{1t} \leqslant A.$$

Условно-оптимальное управление на каждом этапе x_{1t} должно при определенной величине A_t обращать в максимум условно-оптимальный выигрыш $W_t(A_t)$.

Решение задачи, как и обычно в динамическом программировании, начинается с последнего этапа.

Условно-оптимальный выигрыш на последнем этапе (T) при различных значениях A_t составит

$$W_T(A_T) = \max\left\{q_1(x_{1T}) + q_2(A_T - x_{1T})\right\};$$
$$A_T \geqslant x_{1T} \geqslant 0.$$

Определив $W_T(A_T)$ и пользуясь приведенным выше основным функциональным уравнением, последовательно находим условно-оптимальную прибыль на $T-1, T-2, ..., t, ..., 2, 1$ этапах и соответствующее ей условно-оптимальные управления

$$U_{T-1}, U_{T-2}, ..., U_t, ..., U_2, U_1.$$

Зная начальное количество ресурсов A_0, находится максимальная прибыль $W_{max} = W_1(A_0)$ и оптимальное управление на первом шаге $X_{11} = X_{11}(A_0)$ и $X_{21} = A_0 - X_{11}$.

Определив распределение ресурсов по объектам в первый период времени, рассчитывается остаток ресурса на начало второго периода

$$A_2 = y_1(x_{11}) + y_2(A_0 - x_{11}).$$

Зная состояние системы перед началом второго этапа, характеризуемое величиной A_2, определяется оптимальное управление на этом этапе

$$x_{12} = x_{12}(A_2);$$
$$x_{22} = A_2 - x_{12}.$$

Аналогичным образом, последовательно по цепочке определяются состояния системы и оптимальные управления на всех остальных этапах

$$A_0 \to x_{11}(A_0) = x_{11} \to A_2 \to x_{12}(A_2) = x_{12} \to ... \to A_t \to x_{1t}(A_t) =$$
$$= x_{1t} \to ... \to A_T \to x_{1T}(A_T) = x_{1T}.$$

Совокупность ресурсов $x_{11}, x_{12}, ..., x_{1t}, ..., x_{1T}$ выделенных по периодам 1-му объединению, совместно с легко подсчитываемой совокупностью $x_{21}, x_{22}, ..., x_{2t}, ..., x_{2T}$, характеризующей количество ресурсов, выделяемых по периодам 2-му объединению, представляет оптимальное распределение ресурсов.

Продемонстрируем описанный метод решения динамической задачи распределения ресурсов на численном примере.

Пример IX.3. Необходимо составить укрупненный (по пятилеткам) перспективный (на 15 лет) план распределения капитальных вложений на оборудование между двумя объединениями. Эффективность использования капитальных вложений в 1-м объединении выражается зависимостью $q_1(x_{1t}) = 0{,}5\,x_{1t}^2$, а на 2-м $q_2(x_{2t}) = 0{,}4\,x_{2t}^2$.

После эксплуатации оборудование частично амортизируется, кроме того, эффективность его работы снижается. Остающаяся стоимость оборудования после его эксплуатации в течение периода на 1-м объединении составляет $0.6\,x_1$, а после эксплуатации на 2-м – $0{,}8\,x_2$. Начальное количество капитальных вложений A_0. Перед началом каждой пятилетки имеющееся оборудование может перераспределяться между объединениями.

Рассматриваемая задача включает три периода времени (три пятилетки).

Решение задачи начинаем с последнего третьего периода. Эффективность использования оборудования в последний период (за третью пятилетку) составит

$$W_3(A_3) = q_1(x_{13}) + q_2(A_3 - x_{13}) = 0{,}5x_{13}^2 + 0.4\,(A_3 - x_{13})^2.$$

Для отыскания экстремума W_3 в интервале $(0, A_3)$ применим метод дифференциального исчисления. Возьмем производную $\frac{dW_3}{dx_{13}}$ и приравняем ее к нулю.

$$\frac{dW_3}{dx_{13}} = x_{13} - 0{,}8\,(A_3 - x_{13}) = 0;\quad x_{13} = 0{,}44\,A_3;\quad W_3 = 0{,}22\,A_3^2.$$

Проверим, является ли полученная экстремальная точка точкой максимума. Для этого возьмем вторую производную

$$\frac{d^2W_3}{dx_{13}^2} = 1 + 0{,}8 = 1{,}8 > 0.$$

Так как вторая производная положительна, то получена точка минимума и, следовательно, точку максимума надо искать на концах интервала $(0, A_3)$

$W_3 = 0{,}4\,A_3^2$ при $x_{13} = 0$;

$W_3 = 0{,}5\,A_3^2$ при $x_{13} = A_3$.

Так как $0{,}5\,A_3^2 > 0{,}4\,A_3^2 > 0{,}22\,A_3^2$, то максимальное значение эффективности за третий период наблюдается при $x_1 = A_3$, т.е. когда все средства отдаются 1-му объединению, при этом эффект составит $0{,}5\,A_3^2$.

Далее определяем оптимальное распределение средств во вторую пятилетку.

Эффективность использования оборудования за две пятилетки составит

$$W_2(A_2) = [0{,}5\,x_{12}^2 + 0{,}4\,(A_2 - x_{12})^2] + 0{,}5\,[0{,}6\,x_{12} + 0{,}8\,(A_2 - x_{12})]^2.$$

В этой формуле выражение, стоящее в первых квадратных скобках, представляет собой эффективность использования капитальных вложений во вторую пятилетку, а во вторых квадратных скобках – остаток средств на начало третьей пятилетки.

Найдем значение x_{12}, при котором эффект от использования средств за две пятилетки $W_2(A_2)$ достигает максимума. Для этого поступим так же, как и при определении x_{13}

$$\frac{dW_2}{dx_{12}} = x_{12} - 0{,}8\,(A_2 - x_{12}) + 0{,}5 \cdot 2\,[0{,}6x_{12} + 0{,}8\,(A_2 - x_{12})]\,(0{,}6 - 0{,}8) = 0.$$

Откуда $1{,}84\ x_{12} = 0{,}96A_2$; $x_{12} = 0{,}52A^2$; $W_2 = 0{,}47\,A_2^2$.

Вторая производная $\dfrac{d^2W_2}{dx_{12}^2} = 1 + 0{,}8 + 0{,}04 = 1{,}84 > 0$, поэтому найденная точка является минимумом функции, а ее максимум следует искать на концах интервала $(0, A_2)$.

При $x_{12} = 0$ $W_2 = 0{,}72\,A_2^2$, а при $x_{12} = A_2$ $W_2 = 0{,}68\,A_2^2$.

Так как $0{,}72\,A_2^2 > 0{,}68\,A_2^2 > 0{,}47\,A_2^2$, то максимальное значение функции W_2 (максимальная эффективность использования средств за две пятилетки) достигается при $x_{12} = 0$, и следовательно, $x_{22} = A_2$, т. е. во второй период все имеющиеся средства эффективно отдавать 2-му объединению, при этом общий за две пятилетки максимальный эффект от их использования $0{,}72\,A_2^2$.

Составим основное функциональное уравнение для первого периода

$$W_1(A_0) = \max\left\{[0{,}5\,x_{11}^2 + 0{,}4\,(A_0 - x_{11})^2] + 0{,}72\,[0{,}6\,x_{11} + 0{,}8\,(A_0 - x_{11})]^2\right\},$$

$$A_0 \geqslant x_{11} \geqslant 0.$$

Первое слагаемое представляет собой эффект от использования средств в количестве A_0 в первую пятилетку, а второе – условно-оптимальный эффект за две остальные пятилетки. Выражение, стоящее во вторых квадратных скобках $0{,}72\,[0{,}6\,x_{11} + 0{,}8\,(A_0 - x_{11})]^2$, представляет собой остаток средств на начало второй пятилетки.

Воспользуясь методом дифференциального исчисления, найдем оптимальное распределение капитальных вложений в первую пятилетку x_{11}

$$\frac{dW_1}{dx_{11}} = x_{11} - 0{,}8\,(A_0 - x_{11}) + 0{,}72 \cdot 2\,[0{,}6\,x_{11} + 0{,}8(A_0 - x_{11})] \times$$

$$\times\,(0{,}6 - 0{,}8) = 0.$$

Откуда $1{,}86\,x_{11} - 1{,}03\,A_0 = 0$, $x_{11} = 0{,}55\,A_0$, $W_1 = 0{,}57\,A_0^2$.

Для проверки, является ли найденная экстремальная точка точкой максимума, возьмем вторую производную

$$\frac{d^2W_1}{dx_{11}^2} = 1 + 0{,}8 + 0{,}06 = 1{,}86 > 0.$$

Следовательно, экстремальная точка не является точкой максимума, которую надо искать на границах интервала $(0, A_0)$.

При $x_{11} = 0$, $W_1 = 0{,}861\,A_0^2$,

а при $x_{11} = A_0$, $W_1 = 0{,}759\,A_0^2$.

Так как $0{,}861\,A_0^2 > 0{,}759\,A_0^2 > 0{,}57\,A_0^2$, то максимальный эффект от использования ресурсов достигается, когда $x_{11} = 0$, а $x_{12} = A_0$, т.е. все средства в первом периоде отдаются 2-му объединению.

Таким образом, получена следующая оптимальная стратегия распределения средств между объединениями по пятилеткам. В первую и вторую пятилетки все капиталовложения на оборудование следует отдавать 2-му объединению, а в третью все оставшиеся средства передать 1-му. Эта стратегия справедлива при любом начальном количестве средств A_0.

При вложении всех средств в первый период во 2-е объединение эффективность их использования составит $w_1 = 0{,}4\,A_0^2$, а остаток на начале второй пятилетки $A_2 = 0{,}8\,A_0$. Эти средства надо передать опять 2-му объединению, тогда эффект от их использования во вторую пятилетку составит $w_2 = 0{,}4\,(0{,}8\,A_0)^2 = 0{,}256\,A_0^2$, а остаток средств на третью пятилетку $A_3 = 0{,}8 \times 0{,}8\,A_0 = 0{,}64A_0$.

В третью пятилетку эти средства целесообразно передать 1-му объединению, при этом эффект от их использования составит $w_3 = 0{,}5\,(0{,}64\,A_0)^2 = 0{,}205\,A_0^2$. При подобном распределении суммарный максимальный эффект составит $W = w_1 + w_2 + w_3 = 0{,}861\,A_0^2$, а после окончания трех пятилеток остаток средств будет равен $A_4 = 0{,}6 \cdot 0{,}64\,A_0 = 0{,}384\,A_0$, т.е. в активе их сохранится около 38 %.

Метод динамического программирования позволяет решать и значительно более сложные задачи распределения (при большем количестве объектов управления, когда часть эффекта используется в производстве, при переменной по периодам и зависящей от уровня ресурсов эффективности использования ресурсов и т.д.).

§ 6. ЗАДАЧА ОТЫСКАНИЯ КРАТЧАЙШИХ РАССТОЯНИЙ НА СЕТИ

Описываемая ниже задача является усложненной ранее рассмотренной задачей об определении оптимальной траектории изменения состояний системы (см. § 3). При ее решении наряду с определением оптима-

льной траектории перемещения системы из начальной точки в конечную, находились также траектории перемещения системы от всех промежуточных точек до конечной. Это свойство используется при решении задачи отыскания кратчайших расстояний на сети, которая, помимо самостоятельного значения, является также первой стадией решения широко распространенных транспортных задач линейного программирования. При ее решении находятся коэффициенты целевой функции транспортной задачи: расстояния (стоимости перевозок) между поставщиками и потребителями.

Как известно, сетью называется совокупность точек некоторой поверхности, соединенных пересекающимися отрезками – звеньями. Каждое звено характеризуется длиной или стоимостью перевозок по звену. Если звено P_{ij} соединяет точки i и j, то его длина l_{ij}. Длина звена обычно проставляется непосредственно на звене.

От каждой точки сети до всех остальных необходимо найти кратчайшее расстояние и пути, им соответствующие. Точки сети имеют определенные названия (наименования станций, разъездов и т.д.) либо нумеруются в произвольном порядке.

При решении задачи методом динамического программирования путь между точками делится на две части. При этом первая часть – это уже пройденный путь, к которому должны оптимальным образом подобрать вторую часть – следующий кратчайший отрезок пути от промежуточной точки до конечной.

При нахождении расстояний между всеми точками сети задача решается поэтапно, а внутри этапа – по шагам. Этап решения задачи состоит в определении кратчайших расстояний от какой-либо фиксированной точки до всех остальных, а шаг – в определении кратчайших расстояний от любой точки до каждой через соседние, для которых кратчайшие расстояния уже рассчитаны. При нахождении кратчайших расстояний последовательно определяют характеристики точек – расстояния от них до некоторой начальной точки. Алгоритм решения задачи включает следующее:

фиксируем точку P_i, до которой необходимо рассчитать все расстояния. Характеристика точки $P_i = 0$. Характеристику точки удобно записывать на сети непосредственно рядом с точкой;

определяем характеристики точек, соседних с i, $P_j = P_i + l_{ij} = l_{ij}$, на звеньях сети отмечаем стрелкой направление, по которому определена данная характеристика;

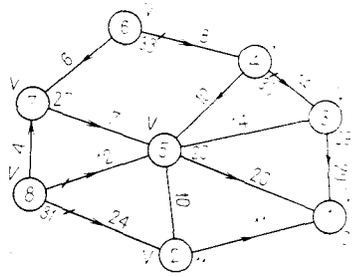
когда характеристики всех соседних точек подсчитаны, точка P_i помечается каким-либо символом (например, V);

переходим к соседней с P_j точке, для которой найдена характеристика, рассчитываем характеристики соседних с ней точек $P_{j'}$

$P_{j'} = P_j + l_{ij'}$;

точку P_j фиксируем (отмечаем симолом V);

Рис. IX.6. Определение кратчайших расстояний на транспортной сети



если характеристики точки ранее уже были рассчитаны, то сравниваем новую характеристику со старой. Если новая характеристика больше старой (например, $P_{j_2} > P_{j_1}$), то все остается без изменений. Если же характеристика меньше старой ($P_{j_2}' < P_{j_1}'$), то меняем ее у точки j, направление связи помечаем знаком в том случае, если точка, у которой изменилась характеристика, не была отмечена. Уменьшение характеристики точки означает, что определено более короткое расстояние от нее до начальной;

если изменилась характеристика у отмеченной точки, то пересчитываем характеристики соседних точек и процедура повторяется до тех пор, пока не будут отмечены все точки.

Поясним основные положения алгоритма на примере.

Пример IX.4. Дана транспортная сеть (рис. IX.6), состоящая из восьми транспортных узлов (точек). Номера точек стоят в кружках, а длины звеньев – непосредственно у звеньев. Маршрут, по которому расстояние минимально, будем показывать стрелками. Характеристики точек записываются рядом с ними. Необходимо найти кратчайшее расстояние от точки 1 до всех остальных.

Записываем возле исходной точки 1 ее характеристику – 0 (пункт 1 алгоритма).

Определяем характеристики точек 2, 3, 4, соседних с исходной точки 1. Характеристики равны длинам звеньев 12, 13, 14. Стрелкой из точек 2, 3, 5 показываем направление, по которому подсчитаны характеристики. Далее переходим к 3 пункту алгоритма.

Точку 1 помечаем символом V.

Переходим к соседней точке 2 (пункт 4 алгоритма). Определим характеристики точек 1, 5, 8 соседних с точкой 2

$P_1 = P_2 + l_{21} = 11 + 11 = 22 > P_1 = 0$,

$P_5 = P_2 + l_{25} = 11 + 10 = 21 > P_5 = 20$,

$P_8 = P_2 + l_{28} = 11 + 24 = 35$.

Так как новые характеристики точек 1 и 4 больше ранее подсчитанных, то прежние характеристики оставляем без изменения, у точки 8 записываем определенную характеристику. Точку 2 отмечаем символом V.

Рассмотрим теперь точку 3 и определим характеристики соседних с ней точек 1, 4, 5

$P_1 = P_3 + l_{31} = 29 + 29 > 0$.

$P_4 = P_3 + l_{34} = 29 + 14 = 43$.

$P_5 = P_3 + l_{35} = 29 + 14 = 43 > 20$.

Характеристики точек 1 и 5 остаются без изменений, а найденную характеристику точки 4 записываем около нее и ставим стрелку от точки 4 к точке 3. Точку 3 помечаем знаком V.

Переходим к точке 4. Характеристики соседних с ней точек 3, 5, 6 составляют

$$P_3 = P_4 + l_{48} = 43 + 14 > 29,$$

$$P_5 = P_4 + l_{45} = 43 + 10 > 20,$$

$$P_6 = P_4 + l_{46} = 43 + 8 = 51.$$

Новые характеристики у точек 3 и 5 больше ранее подсчитанных, поэтому прежние характеристики остаются без изменений. У точки 6 проставляем найденную для нее начальную характеристику и стрелкой на звене 64 отмечаем направление, откуда определена эта характеристика. Точку 4 отмечаем символом V.

Переходим далее к точке 5. Характеристики соседних с ней точкек 2, 4, 7, 8 составляют

$$P_2 = P_5 + l_{52} = 20 + 10 > 11,$$

$$P_4 = P_5 + l_{54} = 20 + 10 = 30 < 43,$$

$$P_7 = P_5 + l_{57} = 20 + 7 = 27,$$

$$P_8 = P_5 + l_{58} = 20 + 12 = 32 < 35.$$

Характеристики у точки 2 остаются без изменений, у точек 7 и 8 проставляем вновь найденные характеристики. Определенная через точку 5 характеристика у точки 4 и 8 меньше ранее определенных ($30 < 43$; $32 < 35$), поэтому у этих точек записываем новые характеристики и стрелкой на звене 45 и 48 отмечаем направление, по которому определена характеристика, а стрелку на звеньях 43 и 28 зачеркиваем (так как теперь кратчайшее расстояние от точки 4 не проходит по этим звеньям).

Так как характеристика у точки 4 изменилась, а она уже была помечена символом V, пересчитываем характеристики соседних с ней точек 3, 5, 6

Как видно, меняется характеристика только у точки 6, направление же, по

$$P_3 = P_4 + l_{43} = 30 + 14 > 29,$$

$$P_5 = P_4 + l_{45} = 30 + 10 > 20,$$

$$P_6 = P_4 + l_{46} = 30 + 8 = 38.$$

Как видно, меняется характеристика только у точки 6, направление же, по которому определена характеристика, остается без изменения.

Переходим теперь к точке 6. Характеристики соседних с ней точек 4 и 7 составляют

$$P_4 = P_6 + l_{64} = 38 + 8 > 30,$$

$$P_7 = P_6 + l_{67} = 38 + 6 > 27.$$

Следовательно, первоначально определенные характеристики у этих точек остаются без изменений. Закончив рассмотрение точек, соседних с точкой 6, отмечаем эту точку знаком V.

Рассмотрим теперь точку 7 и определим характеристики соседних с ней точек 6 и 8

$$P_6 = P_7 + l_{76} = 27 + 6 = 33 < 38,$$

$$P_8 = P_7 + l_{78} = 27 + 4 = 31 < 32.$$

Характеристики у точек 6 и 8 меньше ранее определенных. Записываем новые характеристики у этих точек и стрелкой отмечаем направления, откуда они рассчитаны (звенья 76 и 78), прежние направления (стрелки на звеньях 46 и 58) зачеркиваем. Точку 6 помечаем символом V.

Таблица IX.4

Точка	Кратчайшее расстояние до точки 1	Оптимальный маршрут
2	11	2–1
3	29	3–1
4	30	4–5–1
5	20	5–1
6	33	6–7–5–1
7	27	7–5–1
8	31	8–7–5–1

Так как вновь изменилась характеристика у точки 6, проверим, не изменятся ли характеристики соседних с ней точек 4 и 7. Легко убедиться, что характеристики этих точек остаются без изменения ($P_4 = 33 + 8 > 30$, а $P_7 = 33 + 6 > 27$).

Переходим к точке 8. Характеристики соседних точек 2, 5 и 7 равны

$$P_2 = P_8 + l_{82} = 31 + 24 > 11,$$

$$P_5 = P_8 + l_{85} = 31 + 12 > 20,$$

$$P_1 = P_8 + l_{87} = 31 + 4 > 27.$$

Характеристики у этих точек остаются без изменений. Итак, мы рассмотрели все точки сети. Полученные характеристики точек дают кратчайшие расстояния от них до начальной точки 1 и по стрелкам можно определить оптимальный маршрут.

Например, характеристика точки 8 – 31, показывает, что минимальное расстояние от нее до точки 1 равно 31. Эта характеристика точки 8 определена через точку 7 (на звене 78 имеется стрелка), характеристика точки 7 – через 5, а точки 5 – через 1. Следовательно, кратчайший маршрут из точки 8 в точку 1 проходит по звеньям 87 – 75 – 51, т.е. через точки 8, 7, 5, 1.

Кратчайшие расстояния от всех точек до точки 1 и маршрут, им соответствующий, приведен в табл. IX.4.

§ 7. ИСПОЛЬЗОВАНИЕ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ ПРИ ОПТИМИЗАЦИИ АЛЬТЕРНАТИВНЫХ ГРАФОВ

Альтернативные графы широко используются для решения различных экономических, организационных и технических задач.

Граф – это множество точек плоскости, соединенных различными отрезками. Множество точек – это вершины графа, а отрезки, их соединяющие, – его ребра (дуги, звенья).

Граф называется **сильно связанным**, если много вершин соединены между собой.

Путь – это последовательность дуг, в которой конец каждой совпадает с началом следующей.

Цикл на графе (контур) – это путь, в котором начальная и конечная вершины совпадают.

Графы бывают ориентированные и неориентированные.

Ориентированные — если каждое ребро имеет определенное направление, которое обозначается стрелкой.

В альтернативных графах из множества путей, ведущих из начального состояния в конечное, надо выбрать один, предпочтительный по какому-либо критерию.

По существу из геометрической интерпретации задачи динамического программирования следует, что все их можно интерпретировать с помощью альтернативных графов. Однако во многих случаях альтернативные графы получаются очень сложными, особенно при большом количестве координат системы. Обычно они эффективно используются при решении экономических задач, если система характеризуется двумя параметрами..

Вершины альтернативного графа могут отображать состояние объекта (мощность предприятия, технологическую схему, наличие ресурсов, пространственные характеристики, например, положение горных работ и т.д.). Ребра графа позволяют определять возможные пути перехода из одного состояния в другое (из одной вершины в другую). На ребрах (а иногда и у вершин) обычно проставляют технико-экономические показатели при переходе из одного состояния в другое, например, получаемую прибыль или затраты.

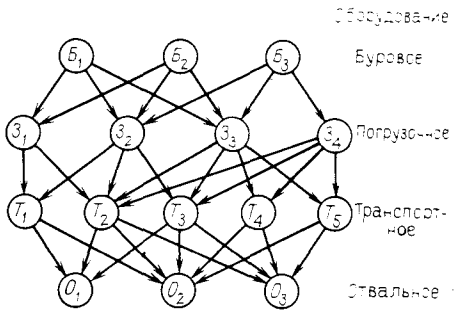
Задача оптимизации состоит обычно в отыскании на графе самого длинного или самого короткого пути (например, пути с максимальной прибылью или с минимальными затратами).

С помощью альтернативных графов эффективно решаются задачи обоснования структуры комплексной механизации, выбора различных технологических схем с учетом динамики развития объекта (например, с учетом изменения параметров горного предприятия в пространстве и времени), планирования развития горных работ, обоснования мощности предприятий на перспективу, распределения ресурсов. Рассмотренные ранее задачи изменения координат объекта и определения кратчайшего пути на сети в принципе тоже решались на альтернативном графе (специфической структуры).

При обосновании с помощью альтернативных графов структуры комплексной механизации (типов и количества оборудования по процессам) вершины графа представляют определенные виды оборудования (рис. IX.7) по процессам. Ребра графа показывают возможные сочетания оборудования (например, типы самосвалов в зависимости от марки экскаватора и т.д.). У вершин обычно выписывают стоимость оборудования, а на ребрах — затраты на его эксплуатацию в зависимости от типа, используемого в смежных процессах. За этап решения задачи здесь принимается определение эффективности работы оборудования, возможных его сочетаний в смежных звеньях.

При использовании альтернативных графов для обоснования технологических схем с учетом динамики развития объекта вершины графа соответствуют определенным технологическим схемам, а ребра показывают возможные варианты их изменения (рис. IX.8). У вершин (или внутри их) проставляются затраты на эксплуатацию определенной схемы, а у ребер — капитальные вложения на переход от одной схемы к другой (или на развитие существующей). Этапом в данном случае является временной период, обычно год.

Рис. IX.7. Альтернативный граф для выбора структуры комплексной механизации



При планировании горных работ вершины альтернативного графа характеризуют пространственное состояние объекта по годам разработки (положение горных выработок, контуры уступов и т.д.). При этом каждая вершина характеризуется и определенными объемами горных работ. Ребра графа показывают переходы из одного состояния горных работ в другое (отработку определенных участков, проведение горных выработок и т.д.). Зная объемы горных работ по положениям, легко определяются и стоимостные показатели по ребрам (например, прибыль при переходе из одного положения горных работ в другое). Этапом решения задачи здесь также выступает временной период — месяц, квартал, год, пятилетка — в зависимости от вида и общего периода планирования.

При использовании альтернативных графов для обоснования поряд-

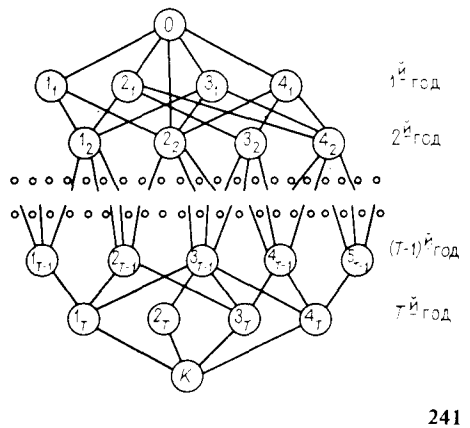


Рис. IX.8. Альтернативный граф для выбора технологической схемы: 1_1, $2_1, \ldots, 1_T, 2_T, \ldots, K$ — варианты технологических схем

ка развития предприятий (плановых заданий) или решения задач распределения ресурсов вершины графа характеризуют состояние предприятия (мощность, плановое задание, количество ресурсов и т.д.), а ребра показывают динамику объекта. Каждому варианту для отдельного предприятия (при планировании деятельности объединения варианты работ шахт, отображаемые вершинами альтернативного графа, носят название шахтовариантов) соответствуют определенные технико-экономические показатели, на основе которых проводится оптимизация графа. Как и в предыдущих двух задачах за этап принимается временная единица (чаще всего год).

Покажем, как оптимизируется альтернативный граф методом динамического программирования на конкретном примере.

Пример IX.5. При разработке месторождения (рассматривается четыре (I – IV) этапа развития горных работ) может быть использовано пять (1–5) вариантов технологических схем. Возможные по периодам варианты технологических схем (вершины) и их изменения (ребра) показаны в виде графа на рис. IX.9.

В вершине графа (за исключением вершин I этапа) показаны эксплуатационные затраты при различных технологических схемах, а на ребрах – капитальные затраты для поддержания работоспособности или изменение схем. Внутри вершин I этапа проставлены суммарные затраты на приобретение оборудования и его эксплуатацию в течение I этапа. Необходимо определить технологические схемы и когда их использовать.

Альтернативные графы обычно оптимизируются методом динамического программирования в прямом порядке (от начала к концу).

Оптимизацию начинаем с I этапа. На этом этапе возможно использование трех технологических схем (у системы допустимы три состояния). Значение критерия оптимальности для каждого из вариантов записано внутри вершины.

Переходим ко II этапу. Для каждой из вершин этого этапа сравниваем затраты при различных вариантах изменения технологических схем. Для этого складываются затраты: в вершине (для остальных этапов у вершины) предыдущего I этапа, на ребре, соединяющем вершину предыдущего этапа с рассматриваемой, и стоящие внутри вершины данного этапа. За условно-оптимальный берется вариант с наименьшими затратами, которые записываются у вершины. На ребре, при переходе по которому затраты минимальны, ставится стрелка.

Например, в вершину, характеризующую использование во II этапе третьей технологической схемы, можно попасть из первой, второй и третьей вершин I этапа. Суммарные затраты при этом при использовании в I этапе технологических

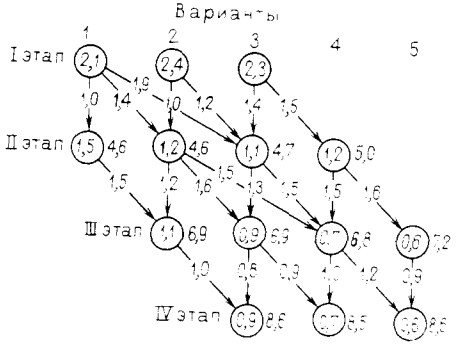


Рис. IX.9. Оптимизация альтернативного графа методом динамического программирования:
1, 2, 3, 4, 5 – варианты технологических схем

схем составят: первой 2,1 + 1,9 + 1,1 = 5,1; второй 2,4 + 1,2 + 1,1 = 4,7; третьей 2,3 + 1,4 + 1,1 = 4,8. Минимальные затраты 4,7 – при использовании в I этапе второй технологической схемы. Ставим эти затраты у третьей вершины II этапа и помечаем стрелкой ребро, соединяющее вторую вершину I этапа с рассматриваемой третьей вершиной II этапа.

Аналогичные расчеты проводят последовательно для всех вершин II–IV этапов. Результаты расчетов показаны на рис. IX.9.

Закончив расчеты для вершин последнего IV этапа, находим вершину, у которой затраты минимальны. Это – вершина, характеризующая использование в IV этапе технологической схемы. Начиная от этой вершины, проходим процесс в противоположном направлении (в данном случае от конца к началу) и по стрелкам восстанавливаем оптимальную динамику использования технологических схем. В вершину 44 (первая цифра вариант технологической схемы, а вторая – номер этапа) ведет стрелка из вершины 43, а в вершину 43 – из вершины 22, в вершину 22 – из вершины 21.

Следовательно, оптимальная стратегия заключается при использовании на I и II этапах второй технологической схемы, а на III и IV – четвертой.

При расчете альтернативных графов, как и обычно при применении динамического программирования (кроме аналитических методов), наряду с оптимальным могут быть определены и субоптимальные пути, значения критериев на которых незначительно (в пределах заданной величины – точности расчетов) отличаются от оптимального. Иногда субоптимальные пути принимаются с учетом других критериев в качестве окончательного решения.

§ 8. ЗАДАЧИ ДИСКРЕТНОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Среди рассмотренных (в этой главе) задач, для решения которых применялся метод динамического программирования, встречались задачи с непрерывной областью изменения управлений и состояний оптимизируемого процесса. Задачи оптимизации процессов, описываемых непрерывными переменными, имеют важное значение для экономико-математического моделирования горного производства: объемы добычи, величины грузопотоков и многие другие величины в большинстве задач планирования следует считать непрерывными, но изменяющимися в дискретные интервалы времени. Задачи такого вида изучает математическая теория оптимальных дискретных процессов, иначе называемая теорией дискретного оптимального управления.

Рассмотрим несколько примеров моделей таких задач.

Пример IX.6. Планирование открытых горных работ.

Месторождение полезного ископаемого разрабатывается одним уступом, высота его H_y. Задано направление развития горных работ. Найти последовательность положений горных работ (на конец 1, 2,..., T-го календарного периода), обеспечивающую минимум среднего квадратичного отклонения производственной мощности карьера по полезному ископаемому от плановой величины $Q_п$. Считается, что за каждый период извлекается объем горной массы Q.

При моделировании разобьем карьерное поле на m одинаковых полос шириной D и будем описывать положение горных работ величинами подвигания $Z_1, \ldots, Z_m$ на каждой полосе. Для решения задачи необходимо знать зависимость мощно-

сти вскрышных пород от подвигания в каждой полосе – $H_i(Z_i)$. При этом за период, в течение которого фронт горных работ в i-й полосе переместился от Z_{i1} к Z_{i2}, извлечены объемы горной массы и вскрыши, равные соответственно

$$H_y D(Z_{i2} - Z_{i1}) \text{ и } B_i(Z_{i2}) - B_i(Z_{i1}), \text{ где } B_i(Z) = D\int_0^Z H_i(\xi)\,d\xi.$$

Обозначим через X_{it} подвигание в i-й полосе за t-й период, а через Z_{it} – за первые t периодов. Требуется минимизировать величину

$$\sum_{t=1}^{T}\Big[Q - \sum_{i=1}^{m} B_i(Z_{it-1} + x_{it}) - B_i(Z_{it}) - Q_{\text{п}}\Big]^2 \tag{IX.11}$$

при условии, что при $t = 1, \ldots, T$

$$x_{it} \geqslant 0 \text{ при } i = 1, \ldots, m, \tag{IX.12}$$

$$H_y D \sum_{i=1}^{m} x_{it} = Q, \tag{IX.13}$$

причем x_{it} и Z_{it} связаны соотношениями

$$Z_{i0} = 0,\ Z_{it} = Z_{it-1} + x_{it} \text{ при } t = 1, \ldots, T;\ \ i = 1, \ldots, m. \tag{IX.14}$$

Пример IX.7. Планирование добычных работ.

Добычной участок включает m забоев. Для каждого забоя зависимость содержания металла в руде от извлеченного объема известна (для i-го забоя она обозначена $\beta_i(Z)$). Плановый период разбит на T подпериодов одинаковой продолжительности, в течение каждого из которых в i-м забое может быть добыто полезного ископаемого не более x_{iB} м^3. Плановый объем добычи участка за подпериод $x_{\text{п}}$ и объемы добычи из каждого забоя за весь период ($Z_{1\text{п}}, \ldots, Z_{m\text{п}}$) заданы. Требуется запланировать такие объемы добычи из каждого забоя по подпериодам, чтобы среднее квадратичное отклонение среднего содержания металла в руде

$$\sum_{t=1}^{m} (\beta_{ct} - \beta_{\text{п}})^2 \tag{IX.15}$$

было минимальным. При формализации этой задачи по существу возникает единственный вопрос: как записать выражение для β_{ct}. Обозначим через x_{it} объем добычи руды в i-м забое за t-й подпериод, а через Z_{it} – за первые t подпериодов. Эти величины связаны уравнениями (IX.14). Общее количество металла в первых Z м^3 руды из i-го забоя, очевидно, равно

$$B_i(Z) = \int_0^Z \beta_i(\xi)\,d\xi.$$

Тогда общее количество металла в руде, извлеченной за t-й период составит

$$\sum_{i=1}^{m} \left[B_i(Z_{it}) - B_i(Z_{it-1})\right],$$

а общее количество руды – $\sum_{i=1}^{m} x_{it}$; таким образом с учетом (IX.14) представим β_{ct} в виде

$$\beta_{ct} = \frac{\sum_{i=1}^{m} \left[B_i(Z_{it-1} + x_{it}) - B_i(Z_{it-1})\right]}{\sum_{i=1}^{m} x_{it}}. \tag{IX.16}$$

Итак, требуется минимизировать (IX.15) с учетом соотношения (IX.16), уравнений (IX.14), при условии, что справедливы ограничения

$$0 \leqslant x_{it} \leqslant x_{iB} \text{ при } i = 1, \ldots, m;\ \ t = 1 \ldots, T; \tag{IX.17}$$

$$\sum_{i=1}^{m} x_{it} = x_{\text{п}} \text{ при } t = 1, \ldots, T; \tag{IX.18}$$

$$Z_{iT} = Z_{i\text{п}} \text{ при } i = 1, \ldots, m \tag{IX.19}$$

Пример IX.8. Календарное планирование работы экскаваторов.

На добычном борту карьера имеется m уступов, причем заходки параллельны, начальное положение экскаватора на i-м уступе задается координатой $y_{i\text{н}}$. Расстояние по горизонтали между экскаваторами на соседних уступах должно быть меньше Δy. Установлены сроки планово-предупредительных ремонтов: i-й экскаватор не ведет работ с t_{i1} по t_{i2} сутки. Известны: максимальный объем суточной добычи i-го экскаватора q_{iB}, ширина заходки D_i и высота уступа H_i. Требуется спланировать такие объемы добычи экскаваторов по суткам q_{it}, при которых их колебания были бы минимальны, т.е. минимальна величина

$$\sum_{t=1}^{T} \Big(\sum_{i=1}^{m} q_{it} - q_{\text{c}}\Big)^2. \tag{IX.20}$$

Выражение (IX.20) можно переписать в виде

$$\sum_{t=1}^{T} \Big(\sum_{i=1}^{m} q_{it}\Big) - \Big(\sum_{i=1}^{m} Z_{iT}\Big)/T. \tag{IX.21}$$

Задача состоит в минимизации (IX.21) при условии, что

$$0 \leqslant q_{it} \leqslant q_{iB} \text{ при } i = 1, \ldots, m,\ t = 1, \ldots, T; \tag{IX.22}$$

$$q_{it} = 0 \text{ при } t = t_{1i}, \ldots, t_{2i}; \tag{IX.23}$$

$$y_{i-1\,t} \geqslant y_{it} + \Delta y \text{ при } i = 1, \ldots, m-1,\ t = 1, \ldots, T, \tag{IX24}$$

где

$$y_{i0} = y_{i\text{н}},\ y_{it} = y_{it-1} + q_{it}/(H_i D_i) \text{ при } i = 1, \ldots, m,\ t = 1, \ldots, T. \tag{IX.25}$$

Все приведенные задачи имеют однотипную математическую форму. Имеется T периодов, управление системой в течение t-го периода описывается m-мерным вектором x_t (как правило, его компоненты обозначались $x_1 t, \ldots, x_m t$), состояние системы к концу t-го периода характеризуется n-мерным вектором Z_t. Управление и состояние связаны разностными уравнениями

$$Z_{i0} = Z_{i\text{н}},\ \ Z_{it} = f_i(Z_{it-1}, x_{it}) \text{ при } t = 1, \ldots, T,\ \ i = 1, \ldots, m \tag{IX.26}$$

Требуется найти последовательность векторов управления $x_1, \ldots, x_T$, доставляющую минимальное (или, напротив, максимальное) значение целевой функции

$$\sum_{t=1}^{T} f_{0t}(Z_{t-1}, x_t) + F(Z_T)$$

при условии соблюдения некоторых ограничений. Это могут быть ограничения на текущее управление вида $g(x_t) \leqslant 0$ или $g(x_t) = 0$, на вектор состояния $h(Z_t) \leqslant 0$ или $h(Z_t) = 0$ и смешанные ограничения типа $r(Z_{t-1}, x_t) \leqslant 0$ или $r(Z_{t-1}, x_t) = 0$. Если однотипные ограничения вводятся при всех t, они называются траекторными, если только при $t = T$ – терминальными, а если при отдельных $t < T$ – промежуточными. Единая форма задач позволяет применить для их решения одни и те же методы: либо это методы математического программирования, учитывающие специфику этого класса задач, либо локальные варианты динамического программирования, либо методы, аналогичные методам теории оптимальных процессов в непрерывном времени.

Соотношения динамического программирования также могут быть записаны для этих задач единообразно. Но можно ли рассчитывать на их успешное использование?

Метод динамического программирования предполагает построение последовательности функций Беллмана $W_T(Z_T)$, $W_{T-1}(Z_{T-1})$, ..., $W_1(Z_1)$ (обратный ход), после чего при $t_1, \ldots, T$ производятся следующие действия: находится вектор управления x_t из условия минимизации $W_t(f_1(Z_{t-1}, x_t), \ldots, f_n(Z_{t-1}, x_t))$ при всех ограничениях для данного t; определяется вектор Z_{t+1} из разностных уравнений (прямой ход). При этом даже задачи прямого хода – найти глобальный минимум по вектору управления – являются трудноразрешимыми при размерности управления $m \geqslant 5$. Неизмеримо труднее задачи обратного хода: ведь требуется найти не вектор, а функцию, притом, как утверждает теория, не обладающую хорошими с точки зрения математического анализа свойствами. Даже в простейших случаях производные функции Беллмана разрывны. Поэтому метод динамического программирования может быть успешно применен для решения лишь отдельных задач за счет их специфики, например, если разностные уравнения и ограничения линейны, критерий оптимальности квадратичный и отсутствуют ограничения типа неравенства (этот класс задач не имеет серьезного практического значения). По этой причине стандартных программ для решения на ЭВМ таких задач методом динамического программирования нет, тогда как его локальные варианты успешно запрограммированы и работают.

§9. ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ И ЕГО ЛОКАЛЬНЫЕ ВАРИАНТЫ В ЗАДАЧАХ ДИСКРЕТНОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Все описываемые в этом параграфе подходы к решению задач методом динамического программирования, объединяемые общим названием "метод вариаций в фазовом пространстве" (пространстве состояний), рассмотрим на примере конкретной задачи календарного планирования работы экскаваторов.

Пусть число уступов равно 2, $D_{1,2} = H_{1,2} = 10$ (м), $\Delta y = 20$ (м), $y_{1н} = 0$, $y_{2н} = 30$, $T = 4$, $t_{11} = t_{21} = 3$, $t_{12} = t_{22} = 2$ (т.е. первый нижний экскаватор останавливается в 3-ю, а второй – во 2-ю смену), $q_{1,2\,B} = 3000$ (м^3), в дополнение к условиям задачи IX.22 – IX.24 вводится условие на конечное положение экскаваторов

$$y_{iT} = y_{iк},\ i = 1, 2,$$

где $y_{1к} = 60$, $y_{2к} = 90$.

Тогда $q_c = (y_{1к} - y_{1н} + y_{2к} - y_{2н})/4 = 3000$.

В рассматриваемых методах из переменных, описывающих задачу, основную роль играют переменные состояния. Последовательность векторов состояния для периодов 0, 1, ..., T называется траекторией процесса. Траектория может быть изображена в фазовом n-мерном пространстве (в нашем случае – это плоскость, на которой заданы декартовы координаты y_1, y_2).

Ограничение IX.24 выделяет допустимую область, в которой должна находиться траектория (полуплоскость $y_2 \geqslant y_1 + \Delta y$).

Пример допустимой траектории изображен на рис. IX.10. Численное решение задачи требует дискретизации допустимой области. Вводится прямоугольная сетка, образованная линиями вида $y_1 = n_1 h_1$ и $y_2 = n_2 h_2$, рассматриваются только траектории, проходящие через узлы этой сетки. Соответственно и функцию Беллмана нужно вычислять лишь для узлов сетки. Положим $W_4(y_{1к}, y_{2к}) = 0$.

Функцию W_3 будем вычислять в точках $\{n_1 h_1, n_2 h_2\}$, входящих в допустимую область, из которых достижима точка $y_к = \{y_{1к}, y_{2к}\}$, т.е. $0 \leqslant y_{iк} - n_i h_i \leqslant q_{iB}/(H_i B_i)$ при $i = 1, 2$. Управление q_4, которое соответствует переходу из такой точки в $y_к$ за четвертый период, определяется однозначно: $q_{i4} = (y_{iк} - n_i h_i)/(H_i B_i)$. Значение W_3 есть не что иное, как четвертое слагаемое целевой функции (IX.20):

$$W_3 = (n_1 h_1, n_2 h_2) = [(y_{1к} - n_1 h_1)(H_1 B_1) + (y_{2к} - n_2 h_2)(H_2 B_2) - q_c]^2.$$

Уравнение Беллмана превратилось в конечное соотношение благодаря отмеченной единственности управления, переводящего из $\{n_1 h_1, n_2 h_2\}$ в y_K.

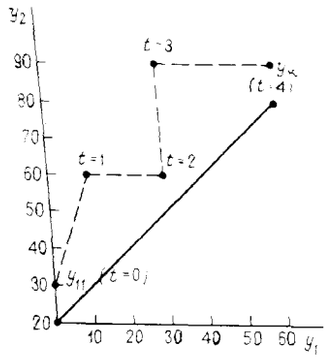


Рис.IX.10. Пример допустимой траектории

Точки $\{ n_1' h_1, n_2' h_2 \}$, в которых вычисляем W_2, в силу условия (IX.23), гласящего $y_{12} = y_{13}$, определяются соотношениями:

$$n_1' = n_1;$$

$$0 \leqslant (n_2 - n_2') \leqslant q_{2B}/(H_2 B_2 h_2). \qquad \text{(IX.27)}$$

Значения $W_2(n_1' h_1, n_2' h_2)$ определяется из решения переборной задачи: найти минимум величины

$$W_3(n_1' h_1, n_2 h_2) + [(n_2 - n_2') h_2 H_2 B_2 - q_c]^2$$

при ограничении (IX.27).

Аналогично $W_1(n_1'' h_1, n_2'' h_2)$ определяется из решения задачи: найти минимум величины

$$W_2(n_1' h_1, n_2'' h_2) + [(n_2' - n_1'') h_1 H_1 B_1 - q_c]^2$$

при ограничении

$$0 \leqslant n_1' - n_1'' \leqslant q_{1B}/(H_1 B_1 h_1). \qquad \text{(IX.28)}$$

Наконец, нас интересует значение $W_0(y_{1н}, y_{2н})$ – минимум величины

$$W_1(n_1'' h_1, n_2'' h_2) + [(n_1'' h_1 - y_{1н}) H_1 B_1 + (n_2'' h_2 - y_{2н}) \times$$
$$\times H_2 B_2 - q_c]^2$$

при ограничениях (IX.27), (IX.28). Области определения и значения функций Беллмана W_3, W_2, W_1 показаны на рис. IX.11.

Получается, что $W_0(y_{1н}, y_{2н}) = 0$. Восстановление траектории уже не составляет труда: вектор Z_t – это точка, в которой $W_t(Z_t) = 0$.

Обсуждаем полученные результаты. Число точек, в которых была вычислена функция W_1, оказалось равным 29. Решение искали на очень грубой сетке, соответствующей скудному набору вариантов плана (правда, получено абсолютно точное решение – так "удачно" были подо-

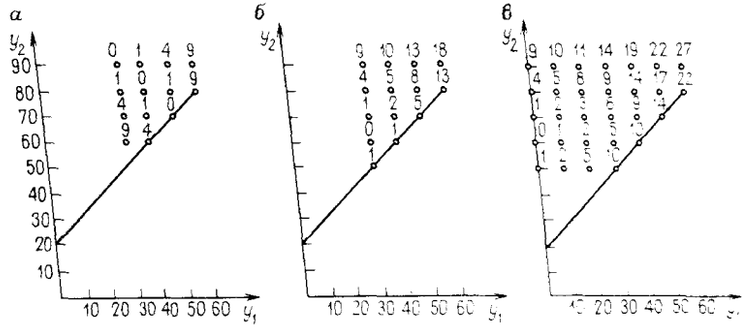


Рис.IX.11. Области определения и значения функций Беламана:
$а - W_3$; $б - W_2$; $в - W_1$

браны числовые параметры задачи планирования). Законное желание уточнить решение за счет уменьшения плана сетки всего лишь вдвое увеличит число этих точек примерно вчетверо, до 103. Нетрудно понять также, что зависимость числа точек от размерности задачи – числа экскаваторов – есть степенная функция. Если бы число экскаваторов было увеличено до 8 (вполне реальная цифра!), то число точек области определения W_1 можно было бы по порядку величины оценить как $103^{8/2} \approx$ $\approx 10^8$. А ведь каждое вычисление требует не менее десятка арифметических операций (более того, число операций также быстро растет с размерностью задачи). Миллиарды операций – часы счета на ЭВМ. А повысить точность еще вдвое (при $m = 8$) мы определенно не сможем – это увеличит время расчета еще примерно в $2^8 = 256$ раз.

Предельно грубый расчет при средней размерности задачи – вот что дает полномасштабное применение метода динамического программирования на прямоугольной сетке. Но есть возможность получить лучшее решение вблизи от полученного с использованием более мелкой сетки. Этому служат локальные варианты динамического программирования – метод блуждающей трубки и метод локальных вариаций.

В первом случае при каждом t выделяется совокупность точек Ξ_t вблизи полученного Z_t (Z_t тоже включается в Ξ_t). Например, для $m = 2$ это могут быть 9 ближайших точек сетки с половинным шагом. Теперь методом динамического программирования находится наилучшая траектория $Z' = Z_0', Z_1', Z_T'$, такая, что при любом t Z_t' относится к Ξ_t. Затем блуждающая трубка – последовательность совокупностей точек $\Xi_1', \Xi_2', \ldots, \Xi_{T-1}'$ – строится вокруг траектории Z, подобным образом находится траектория Z'. Процесс продолжается до тех пор, пока новая траектория не совпадет с предыдущей.

На рис. IX.12, а показана допустимая траектория Z, на рис. IX.12, б, в, г множества Ξ_t и значения W_t в точках Ξ_t. Оптимальная траектория задачи попадает в блуждающую трубку, и поэтому получаем ее на первой

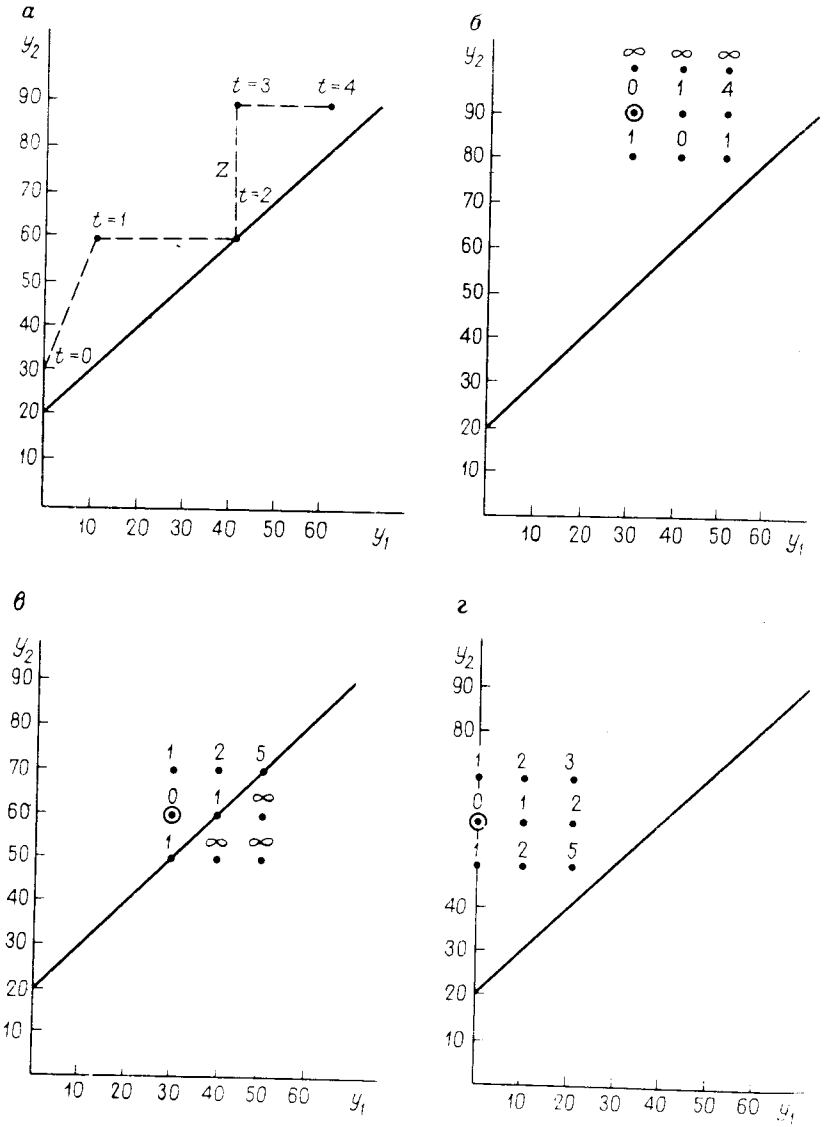


Рис. IX.12. Допустимые траектории (*а*) и значения функций полезности: *б* – W_3; *в* – W_2; *г* – W_1

же итерации. Метод всегда дает по крайней мере локально оптимальную траекторию.

Метод локальных вариаций – наименее трудоемкий из рассмотренных. На итерации этого метода по известной допустимой траектории Z строится траектория Z' с меньшим значением целевой функции, отличающаяся от Z значением вектора состояния при единственном t (это значение Z'_t должно входить в окрестность Ξ_t вектора Z_t). Если ни одного такого Z'_t (ни при одном t) не найдется, метод останавливается. К сожалению, полученная траектория может не быть и локально оптимальной. В примере, рассмотренном для метода блуждающей трубки, начальное управление за счет такого локального варьирования может быть только ухудшено, а не улучшено, и тем не менее оно не является оптимальным.

В заключение отметим, что рассмотрено применение нескольких методов в максимально благоприятных для них условиях: во-первых размерности векторов управления и состояния совпадали ($m = n$), благодаря чему для пары векторов Z_t, Z_{t+1} существует не более одного вектора x_{t+1}, обеспечивающего переход из Z_t в Z_{t+1} за $(t+1)$-й период; во-вторых компоненты x_{t+1} вычисляются по предельно простым формулам; в-третьих, в задаче нет смешанных ограничений. В случае, когда $n > m$, эти методы не могут быть применены, а наличие смешанных ограничений затрудняет их использование.

§ 10. МЕТОД ВОЗМОЖНЫХ НАПРАВЛЕНИЙ В ЗАДАЧАХ ДИСКРЕТНОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

В задачах с ограничениями-неравенствами метод возможных направлений имеет следующее достоинство: если известен допустимый план, то на всех последующих итерациях будем получать допустимые планы, причем лучшие, чем предыдущие. Таким образом, процесс оптимизации не будет безрезультатным, если просчитана хотя бы одна итерация; это важно, когда время решения ограничено, а размерность задачи высока.

Метод возможных направлений применяется в задачах дискретного оптимального управления с учетом их специфики. Особенно благоприятны условия его применения в практически важном случае, когда отсутствуют ограничения на состояние и смешанные ограничения.

Рассмотрим идею метода возможных направлений для этого случая. Пусть вектор управления для t-го периода получил приращение Δx_t. Тогда изменятся значения ограничений для текущего периода – только для него; изменится и значение критерия оптимальности J. Можно принять данное Δx_t, если оно задает возможное направление по отношению к ограничениям текущего периода

$$(\text{т.е. при } \delta > 0 \quad r_{jx}^{T}(Z_{t-1}, x_t)\,\Delta x_t \leqslant 0, \text{ если } r_j(Z_{t-1}, x_t) \geqslant -\delta) \qquad \text{(IX.29)}$$

и обеспечивает убывание целевой функции. Чтобы проверить последнее

требование, нужно знать вектор частных производных (градиент) критерия оптимальности по переменным управления $x_{1t}, \ldots, x_{mt}$, считая переменные состояния зависимыми от них (эта зависимость выражается разностными уравнениями).

Зная вектор $\nabla_{x_t} J$, можно найти Δx_t из решения задачи: минимизировать $(\nabla_{x_t} J)^T \Delta x_t$ при ограничениях (IX.29) и условии

$$-1 \leqslant \Delta x_{it} \leqslant 1, \quad i = 1, \ldots, m.$$

Таким образом, задача определения возможного направления убывания целевой функции распалась на T задач невысокой размерности. Это дает колоссальный выигрыш по времени счета на ЭВМ (порядка T^2 раз) по сравнению с непосредственным применением метода возможных направлений без учета особенностей динамической задачи.

Для вычисления вектора $\nabla x_t J$ нужно воспользоваться правилами дифференцирования сложных функций. Получаемую формулу удобнее всего выразить с помощью сопряженных переменных P_t, которые определяются из системы разностных уравнений

$$\left.\begin{aligned} &P_T = \nabla_Z F(Z_T), \\ &P_t = \nabla_Z f_0(Z_t, x_{t+1}) + f_Z^T(Z_t, x_{t+1}) P_{t+1}, \\ &t = T-1, T-2, \ldots, 1. \end{aligned}\right\} \qquad \text{(IX.30)}$$

В этой системе используются матрицы $f_x(Z, x)$, $f_Z(Z, x)$, элементы которых – соответственно

$$\frac{df_i(Z, x)}{dx_j}, j = 1, \ldots, m, \quad \frac{df_i(Z, x)}{dZ_j}, \ j = 1, \ldots, n; \ i = 1, \ldots, n.$$

Система (IX.30) решается от конца к началу: последовательно определяются вектора $P_T, P_{T-1}, \ldots, P_1$. Вектор $\nabla_{x_t} J$ выражается формулой

$$\nabla_{x_t} J = \nabla_x f_0(Z_{t-1}, x_t) + P_t^T f_x(Z_{t-1}, x_t).$$

Рассмотрим применение метода возможных направлений на примере задачи планирования открытых горных работ (см. § 8).

Сопряженная система в данном случае принимает вид

$$P_{it} = P_{it+1} + \tau_t \left(H_i(Z_{it}) - H_i(Z_{it} + x_{it+1})\right),$$

где $\tau_t = 2D\left(Q + \sum_{i=1}^{m} \left\{B_i(Z_{it}) - B_i(Z_{it} + x_{it+1})\right\} - Q_{\text{п}}\right)$,

а задача определения приращения управления для t-го периода

$$\left.\begin{aligned} &\sum_{i=1}^{m} \pi_{it} y_{it} \to \min, \ y_{it} \geqslant 0, \text{ если } x_{it} \leqslant \epsilon, \\ &\sum_{i=1}^{m} y_{it} = 0, \ -1 \leqslant y_{it} \leqslant 1, \ i = 1, \ldots, m. \end{aligned}\right\} \qquad \text{(IX.31)}$$

Здесь

$$\pi_{it} = P_{it} - \tau_{t-1} H_i(Z_{it-1} + x_{it}).$$

Задача (IX.31) может быть решена симплекс-методом. Однако для ее решения может быть предложен и более простой алгоритм. Переставим величины $\pi_{1t}, \ldots, \pi_{mt}$ в порядке возрастания: $\pi_{j_1 t} \leqslant \pi_{j_2 t} \leqslant \ldots \leqslant \pi_{j_m t}$. Удалим из этого ряда π_{it}, если $x_{it} \leqslant \epsilon$ и оставшиеся q элементов снова занумеруем $\pi_{l_1 t} \leqslant \pi_{l_2 t} \leqslant \ldots \leqslant \pi_{l_q t}$. Образуем ряд $\pi_{j_1 t} - \pi_{l_q t}, \pi_{j_2 t} - \pi_{l_{q-1} t}, \ldots, \pi_{j_q t} - \pi_{l_1 t}$; он является монотонно возрастающим. Если он не содержит ни одного отрицательного элемента, вектор y_t – нулевой; если он содержит ν отрицательных элементов, вектор y_t имеет следующие ненулевые компоненты:

$$y_{j_s t} = 1, \ y_{l_{q+1-s} t} = -1, \ s = 1, \ldots, \nu.$$

Специализированный вариант метода возможных направлений предложен и для задач со смешанными ограничениями. Обозначим $\nabla_x r_j(Z_{t-1}, x_t)$ и $\nabla_Z r_j(Z_{t-1}, x_t)$ градиенты r_j по x_t и Z_{t-1} соответственно. Представим приращение управления в виде

$$\Delta x_t = \Delta_1 x_t + C_t \Delta Z_{t-1},$$

определив матрицу C_t как решение системы, включающей уравнения

$$\left(\nabla_x r_j(Z_{t-1}, x_t)\right)^T C_t = -\nabla_Z r_j^T(Z_{t-1}, x_t)$$

для тех j, при которых $r_j(Z_{t-1}, x_t) \geqslant -\epsilon$ и уравнения $\left(\nabla_x r_j(Z_{t-1}, x_t)\right)^T \times C_t = 0$ для тех j, при которых $q_i(x_t) \geqslant -\epsilon$.

Тогда активные (обращающиеся с точностью до $\epsilon \geqslant \delta > 0$ в равенство) смешанные ограничения в линейном приближении уподобляются простейшим ограничениям $q_i(x_t) \leqslant 0$: их приращения зависят только от x_t. Задача определения вектора $\Delta_1 x_t$ в этом случае отличается от рассмотренной только тем, что сопряженные переменные в ней являются решением системы разностных уравнений

$$\left.\begin{aligned} &P_T = \nabla_Z F(Z_T), \\ &P_t = \nabla_Z f_0(Z_t, x_{t+1}) + f_Z(Z_t, x_{t+1}) P_{t+1} + C_t^T \left\{\nabla_x f_0(Z_t, x_{t+1}) + \right. \\ &\left. + f_x(Z_t, x_{t+1}) P_{t+1}\right\}, \\ &t = T-1, T-2, \ldots, 1. \end{aligned}\right.$$

§ 11. МЕТОДЫ, ОСНОВАННЫЕ НА УСЛОВИЯХ ОПТИМАЛЬНОСТИ

Методы возможных направлений заключаются в последовательном улучшении допустимого плана. Описываемые здесь методы основаны на другой идее: строится последовательность управлений и двойственных переменных (множителей Лагранжа, сопряженных переменных), на которой все более точно выполняются соотношения, выражающие необходимые условия оптимальности. Здесь будут рассмотрены две группы методов: двойственные и дискретного принципа максимума.

При использовании двойственных методов существуют две возможности выбора независимых переменных в задаче.

В первом случае считаем независимыми как переменные управления, так и переменные состояния, а разностные уравнения считаем ограничениями типа равенства, связывающими их. Множители Лагранжа для этих ограничений будем обозначать P_{ti} (как и сопряженные переменные в предыдущем параграфе); такое обозначение оправданно, поскольку, как доказано в теории, на оптимальном управлении эти величины совпадают. Функция Лагранжа имеет вид

$$F(x, Z, P) = \sum_{t=1}^{T} f_0(Z_t x_t) + P_{t+1}^T (Z_{t+1} - f(Z_t, x_t)) =$$

$$= \sum_{t=1}^{T} \left\{ f_0(Z_t x_t) + P_t^T Z_t - P_{t+1}^T f(Z_t, x_t) \right\} + P_{T+1}^T Z_{T+1} - P_1^T Z_1,$$

т.е. представляется в виде суммы $T + 2$ слагаемых, в каждое из которых входит вектор управления x_t и вектор состояния Z_t только при одном t.

Метод состоит в следующем. При известной сопряженной траектории $P = \{ P_2, ..., P_{T+1} \}$ находим вектор $x_1 = x_1(P)$, минимизируя первое слагаемое по x_1 при ограничениях на x_1; при $t = 2, ..., T$ находим пару $x_t = x_t(P)$, $Z_t = Z_t(P)$, минимизируя t-е слагаемое при ограничениях: на управление x_t, состояние Z_t и смешанных — на x_t, Z_t; находим $Z_{T+1} = Z_{T+1}(P)$, минимизируя $(T + 1)$-е слагаемое при ограничениях на конечное состояние Z_{t+1}. Затем находим новую сопряженную траекторию по формуле

$$P'_{t+1} = P_{t+1} + \alpha \left[Z_{t+1}(P) - f(Z_t(P), x_{t+1}(P)) \right], t = 1, ... T,$$

причем $\alpha > 0$ выбираем таким образом, чтобы функция Лагранжа возрастала. Таким образом, вместо того, чтобы решать исходную сложную задачу, на каждой итерации решается $T + 1$ простая задача. Доказано, что метод приводит к успеху, если задача удовлетворяет следующим требованиям: разностные уравнения и ограничения-равенства — линейны; функции $r_j(Z_{t-1}, x_t)$, $h_l(Z_t)$, $g_s(x_t)$ для ограничений-неравенств и $f_{0t}(Z_t, x_t)$ — выпуклы. В этом случае последовательность сопряженных траекторий сходится к P^*, оптимальные управление и траектория — $x(P^*)$, $Z(P^*)$. Напомним, что разностные уравнения на промежуточных итерациях не соблюдаются и только на $x(P^*)$, $Z(P^*)$ они выполняются. Ограничения же соблюдаются всегда.

Можно пойти еще дальше, отказавшись от соблюдения ограничений в процессе оптимизации. Тогда на каждой итерации пары $\{ x_t, Z_t \}$ будут определяться из решения задач безусловной минимизации, в которых вместо ограничений добавлены новые слагаемые в целевую функцию.

Одним из важнейших результатов теории оптимальных непрерывных процессов является принцип максимума Понтрягина. Для дискретных процессов он справедлив не всегда, но при тех условиях, которые гарантировали сходимость двойственного метода, он верен. В простейшем случае — в задаче с ограничениями типа неравенства на переменные управления — принцип максимума гласит: если X^* оптимальное управление, а Z^* — соответствующая ему траектория, то при любом t вектор X_t^* является решением задачи

$$f_0(Z_t^*, x_{t+1}) + (\psi_{t+1}^*)^T f(Z_t^*, x_{t+1}) \to \max, g_s(x) \leqslant 0; \quad s = 1, ..., n_{g1t}$$

(здесь $\psi_{t+1}^* = -P_{t+1}^*$). Таким образом задача устанавливает зависимость управления x_t от P_{t+1} и Z_t, т.е. $x_t(Z_t, P_{t+1})$.

Если эту зависимость подставить в разностные уравнения для траектории и сопряженной траектории, получим так называемую краевую задачу для систем $2n$ разностных уравнений: условия на Z заданы при $t = 0$, а на P — при $t = T$ (на "краях" рассматриваемого интервала времени). Методы решения таких задач разработаны, однако плохие дифференциальные свойства зависимости $x(Z_t, P_{t+1})$ осложняют их применение. Несмотря на то, что методы, основанные на принципе максимума, известны давно, теоретически они изучены еще недостаточно.

Контрольные вопросы и задачи

1. В чем отличительная особенность метода динамического программирования?
2. Почему в методе динамического программирования оптимизация ведется обычно от последнего этапа?
3. Назовите основные этапы решения задачи динамического программирования.
4. Определить оптимальную траекторию перемещения системы из точки 0 в точку 11 (рис. IX.13):

максимальный путь, если перемещение возможно только по вертикали и горизонтали;

минимальный путь, если перемещение возможно только по вертикали и горизонтали

максимальный путь, если перемещение возможно еще и по диагонали;

кратчайший путь, если перемещение возможно еще и по диагонали.

Исходные данные по вариантам приведены в табл. IX.5.

5. Горно-обогатительному комбинату дополнительно выделено 8 комплектов

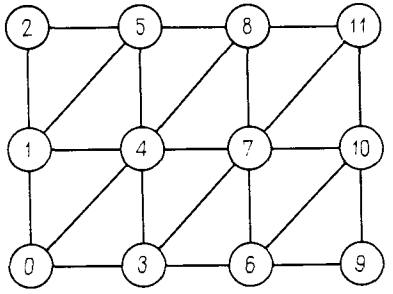


Рис. IX.13. Траектории перемещений системы

горного оборудования, которое надо распределить между тремя рудниками таким образом, чтобы общий прирост добычи был максимален. По горнотехническим условиям на руднике № 1 можно дополнительно установить не более 3 комплектов оборудования, а на № 2 и № 3 – не более 4 (на каждом отдельно). Прирост добычи по рудникам (в условных единицах) в зависимости от числа дополнительно выделенного оборудования приведен в табл. IX.6.

Таблица IX.5

Звено	Значение критерия эффективности на звеньях													
	Варианты													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0–1	5	8	4	3	5	8	4	8	3	6	8	3	6	7
0–3	7	6	3	4	5	6	3	6	4	8	6	4	8	4
0–4	7	8	8	7	9	8	8	7	8	8	7	8	7	10
1–2	4	3	5	4	3	4	3	3	4	4	3	4	5	6
1–4	5	4	5	7	4	7	4	5	7	7	4	7	8	7
1–5	8	7	9	6	10	6	11	7	6	6	7	6	8	10
2–5	6	7	4	4	6	4	6	7	4	4	8	5	4	4
3–4	2	4	4	5	6	4	5	4	5	4	3	6	4	3
3–6	3	3	6	5	9	3	6	3	5	3	6	9	6	3
3–7	7	8	10	9	11	8	10	8	9	8	10	11	10	8
4–5	5	5	6	3	5	3	6	5	6	3	5	3	5	9
4–7	5	5	9	4	8	4	9	5	9	4	8	4	8	7
4–8	10	9	11	10	9	10	11	9	11	10	9	8	10	11
5–8	4	4	8	6	7	6	6	4	8	6	7	6	7	6
6–7	7	4	7	6	7	4	8	7	6	7	8	7	8	5
6–9	4	6	7	9	4	6	10	7	7	2	4	4	7	5
6–10	9	10	10	11	8	10	7	10	9	8	8	11	10	11
7–8	8	6	7	5	6	5	7	7	7	6	8	6	6	8
7–10	6	9	4	8	5	8	7	4	4	5	5	6	5	6
7–11	12	11	8	10	11	10	10	8	9	9	11	10	8	11
8–11	7	8	5	7	9	7	6	5	5	9	8	9	7	5
9–10	4	7	4	7	3	7	5	4	4	3	3	4	4	9
10–11	3	7	3	6	5	6	11	3	3	5	5	6	5	7

Таблица IX.6

Число дополнительно выделенного оборудования	Прирост добычи по рудникам		
	№ 1	№ 2	№ 3
1	1,2	1,5	1,9
2	2,1	1,8	2,8
3	3,0	2,6	3,0
4	–	3,2	3,1

6. Задача. На три разреза выделено 10 экскаваторов. Распределить экскаваторы между разрезами таким образом, чтобы прирост добычи был бы максимальным. Прирост добычи на разрезах в зависимости от числа дополнительно выделенного оборудования показан в табл. IX.7. При этом разрезу № 1 не может быть выделено более 3 экскаваторов, № 2 – не более 4, № 3 – не более 5.

Таблица IX.7

Число дополнительно выделенного оборудования	Прирост добычи по разрезам		
	№ 1	№ 2	№ 3
1	1,1	1,2	1,1
2	2,3	2,2	2,1
3	3,2	3,2	3,2
4	–	4,1	4,0
5	–	–	5,0

7. Задача. Решить задачу 17 из гл. VIII методом динамического программирования.

Глава X
СЕТЕВОЕ ПЛАНИРОВАНИЕ И УПРАВЛЕНИЕ РЕАЛИЗАЦИЕЙ ПРОГРАММ

§ 1. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ И ЭТАПЫ СЕТЕВОГО ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ

Среди методов, предназначенных для составления, планирования и управления реализацией программ производства или исследований, сетевые методы планирования и управления (СПУ), базирующиеся на

идее критического пути, занимают ведущее место благодаря простоте, эффективности и большим возможностям.

Под программой понимается совокупность взаимосвязанных операций, которые необходимо выполнить в определенном порядке, чтобы достигнуть поставленной цели. Операции связаны отношением порядка в том смысле, что один из них нельзя начать, прежде чем не будут завершены другие. Операция программы обычно рассматривается как работа, для выполнения которой требуются затраты времени и ресурсов.

В результате высокой эффективности сетевых методов они нашли широкое применение при составлении и реализации таких программ горного производства, как:

строительство новых и реконструкция действующих горных предприятий;

решение отдельных вопросов эксплуатационной деятельности горнодобывающих предприятий (вскрытие и подготовка новых горизонтов или участков, подготовка отдельных очистных блоков,, внедрение новой техники и технологии, строительно-монтажные и пуско-наладочные работы и капитальный ремонт крупного технологического оборудования, ликвидация аварий и т.д.);

эксплуатационная деятельность горнодобывающего предприятия в целом (отработка блоков; капитальные и подготовительные работы, необходимые для отработки этих блоков; создание резерва подготовленных и готовых к выемке запасов в других блоках);

ряд других производственных и научно-исследовательских программ.

Использование сетевых методов дает возможность увязать во времени программу производства работ, планировать последовательность их выполнения, выявлять и устранять возникающие в процессе реализации нарушения. Анализ сетевых моделей позволяет: определить, от каких операций и в какой степени зависит срок выполнения всей программы; предвидеть появление "узких мест"; выделить второстепенные, с точки зрения времени реализации программы, работы, сокращение продолжительности которых не только не приведет к уменьшению времени выполнения всей совокупности операций, но и может привести к увеличению стоимости работ и простоев рабочих и оборудования. Кроме того, сетевые методы позволяют дать обоснованные, в том числе и с экономической точки зрения, ответы на вопросы организации выполнения работ программы.

Основой сетевых методов планирования и управления является представление программы в виде сетевой модели или сети, которая отражает отношение порядка, существующее на множестве операций программы.

Сетью называется ориентированный связанный граф без контура, в котором существует лишь только одна вершина, не имеющая входя-

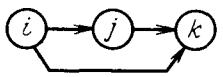


Рис. X.1. Сетевая модель

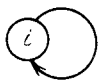


Рис. X.2. Петля

щих дуг, и только одна вершина, не имеющая выходящих дуг.

Следовательно, математической основой сетевых методов служит теория графов, которая, в свою очередь, является важной частью теории множеств. Однако в прикладных методах сетевого планирования принята несколько иная терминология и система обозначений. На ней сконцентрируем основное внимание в этом параграфе.

"Сетевая модель" или "сеть" – это граф, состоящий из вершин – "событий" и направленных дуг – "работ" или "операций" (рис. X.1).

Событие в сетевой модели задается индексом и обозначает момент начала или окончания работ. Событие, обозначающее только начальные моменты работы именуется исходным, а событие, состоящее только в окончании работ, – завершающим. На рис. X.1 исходное событие – *i* и завершающее – *k*. Промежуточное событие, или просто событие, например, на рис. X.1 событие *j*, состоит в окончании одних и начале других работ. Если событие завершает несколько работ, то оно наступает только в момент окончания последней по времени, входящей в него работы.

Каждая работа в модели однозначно определяется начальным и конечным событиями и обозначается через индексы, взятые в круглые скобки. Например, (i, j) – работа с i-м начальным и j-м конечным событием. Из определения сети, а также, условия однозначности отображения и характера описываемого объекта вытекают следующие требования:

в сети не должно быть петель (рис. X.2), т.е. одно и то же событие не может быть начальным и конечным событием работы, что соответствует невозможности существования операции (i, i);

в сети не должно быть параллельных работ, имеющих общее начальное и общее конечное события (рис. X.3, *а*), от которых избавляются путем введения различных начальных и конечных событий работ (рис. X.3, *б, в*).

Работа в сетевой модели может означать:

действительную работу – реальный процесс, требующий затрат времени и ресурсов;

ожидание – процесс, имеющий длительность, но не требующий ресурсов;

фиктивную работу, означающую логическую связь между двумя событиями.

Если события i и j связаны фиктивной работой (i, j), то из этого следует, что событие j наступит только после свершения события i. Фиктивные работы имеют нулевые длительность и расход ресурсов.

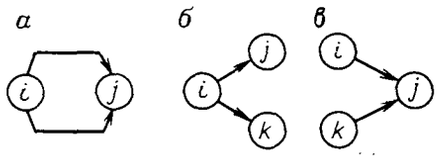


Рис. X.3. Параллельные работы (*а*) и способы их представления в сети (*б*) и (*в*)

Графически действительные работы и ожидания обозначаются сплошными, а фиктивные работы – пунктирными стрелками, которые указывают на конечное событие работы.

Последовательность работ, соединяющие два события i и f, образуют путь, обозначаемый $L\ (i, j, ..., f)$, где $i, j, ..., f$ – индексы событий, через которые проходит этот путь.

Сетевая модель отражает последовательность выполнения работ и в комплексе с временными параметрами работ и событий образует план реализаций совокупности операций программы.

В сетевом планировании оперируют следующими временными параметрами:

t_{ij} – длительность работы (i, j); T_{pi} – ранний срок наступления i-го события; $T_{пi}$ – поздний срок наступления i-события; R_i – резерв времени наступления i-го события; r_{ij} – полный резерв времени выполнения работы (i, j); Δ_{ij} – свободный резерв времени выполнения работы (i, j); $T_L\ (i, j, ..., f)$ – продолжительность пути $L\ (i, j, ..., f)$; s – продолжительность критического пути.

Порядок расчета и смысл этих величин будет определен ниже.

Сетевое планирование и управление реализацией программ включает три основных этапа;

структурное планирование;

календарное планирование;

оперативное управление.

На этапе структурного планирования выполняются: разбиение программы на операции и определение условий их выполнения; построение сетевой модели и нумерация ее событий. Построение сетевой модели на этапе структурного планирования позволяет детально проанализировать программу и внести улучшения в ее структуру еще до начала реализации.

Целью календарного планирования (планирования во времени) является построение плана, определяющего моменты начала и окончания каждой операции, а также ее взаимосвязь с другими операциями программы. Кроме того, календарный план должен обеспечить возможность выявления критических (по времени) операций (которым необходимо уделять особое внимание, чтобы выполнить программу в директивный срок), а также резервов времени некритических операций. Этап включает: расчет временных параметров событий и работ сетевой модели и определение критического пути; построение календарного плана и его оптимизацию.

Этап (заключительный) оперативного управления процессом реализации программы включает использование структурного (сетевой модели) и календарного планов для составления периодических отчетов о ходе ее выполнения. Сетевая модель подвергается анализу и, в случае необходимости, корректируется, при этом составляется новый календарный план реализации оставшейся части программы.

§ 2. СЕТЕВОЕ ПРЕДСТАВЛЕНИЕ ПРОГРАММ (СЕТЕВАЯ МОДЕЛЬ)

Исходной информацией для построения сетевой модели являются перечень работ или операций программы, сведения об их технологической последовательности, длительности и необходимых ресурсах для их выполнения. Подготовка такой информации требует знаний технологии, способов и условий ведения работ, основ нормирования, требований техники безопасности и т.д. Эта наиболее творческая и наименее формализованная часть этапа структурного планирования. После сведения информации в таблицу определенного вида все дальнейшие действия выполняются по строгим формальным правилам независимо от характера реализуемой программы.

Исходная информация и результаты последующих расчетов заносятся в специальную форму (форма X.1), которая в дальнейшем будет именоваться "таблица работ".

В таблице работ исходная информация занимает 1, 2, 3, 6 столбцы, а остальные заполняются по мере построения и расчета временных параметров модели. Число строк должно превышать количество работ в исходном списке примерно в полтора раза, чтобы обеспечить все дальнейшие записи без увеличения таблицы.

Наименования работ заносятся в произвольной последовательности и им присваиваются неповторяющиеся номера. Обычно для нумерации используется последовательный ряд натуральных чисел, начиная с единицы.

Далее, исходя из принятой последовательности работ, в столбце 3

Форма X.1

Наименование работы	Номер работы	Номера непосредственно предшествующих работ	Индекс начального события работы i	Индекс конечного события работы j	Длительность работы $t_{i,j}$	Полный резерв времени работы $r_{i,j}$	Свободный резерв времени работы $\Delta_{i,j}$
1	2	3	4	5	6	7	8

перечисляются номера непосредственно предшествующих работ. Это наиболее ответственная процедура, от которой зависит соответствие сетевой модели реальной последовательности операций. Составление перечня предшествующих работ выполняется по следующим правилам.

1. Для каждой работы в произвольной последовательности указываются номера операций, от которых зависит ее начало. Если предшествующих работ нет, то графа не заполняется.

2. Если нельзя с уверенностью определить, является ли рассматриваемая работа непосредственно предшествующей или влияет через другие операции, ее номер следует включить в перечень. Избыточные номера предшествующих работ, т.е. влияющих через другие операции, не могут исказить сетевой модели, а лишь увеличивают трудоемкость ее построения.

3. Если две работы могут выполняться частично параллельно, то их части, выполняемые последовательно и параллельно, рассматриваются как самостоятельные операции и им присваиваются собственные названия и номера.

4. Если в процессе выполнения работы есть периоды со значительным колебанием расхода ресурсов, то эта работа разделяется на последовательные операции с постоянными расходами ресурсов, им присваиваются собственные названия и номера.

При составлении "таблицы работ" следует избегать излишней детализации перечня операций и включения избыточной информации о предшествующих работах.

На этом этапе могут быть допущены следующие ошибки.

1. При составлении перечня условий выполнения работы ошибочно включен номер параллельно выполняемой или пропущен номер предшествующей работы. Эта ошибка может быть выявлена только в результате тщательного сравнения записи условий и последовательности работ в программе.

2. В таблицу включен комплекс замкнутых циклических операций. Например, в табл. X.1 работа В зависит от выполнения работы А, С – от В, а А – от С. Такой комплекс работ выявляется и ликвидируется в процессе построения сетевой модели.

Таблица X.1

Наименование работы	Номер работы	Номера непосредственно предшествующей работы
A	1	3
B	2	1
C	3	2

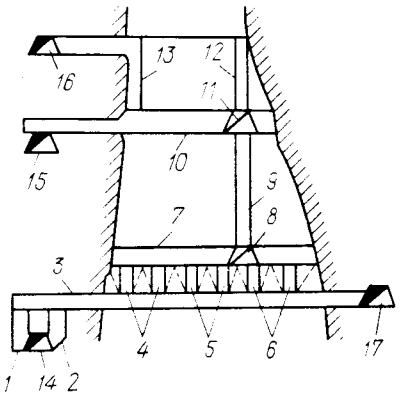


Рис. X.4. Схема расположения выработок в камере. Выработки, которые необходимо провести: *1* – ходовой восстающий; *2* – рудоспуск; *3* – скреперный орт; *4, 5, 6* – пары дучек, соответственно, 1-я, 2-я, 3-я; *7* – орт подсечки; *8, 11* – отрезной штрек § 1, 2; *9, 12* – отрезной восстающий № 1, 2; *10* – буровой орт; *13* – закладочный ходок. Проведенные штреки: *14* – откаточный; *15* – откаточный верхнего горизонта; *16* – закладочный (ходок); *17* – вентиляционный

*Пример X.1**. Составить перечень условий выполнения работ программы ввода камеры на руднике в эксплуатацию, включая сдачу ее под бурение (рис. X.4).

Для решения поставленной задачи составим табл. X.2, представляющую фрагмент таблицы работ размером в 18 строк, и в нее занесем перечень операций. Свободные строки могут понадобиться в дальнейшем при построении сети в случае появления фиктивных работ.

Таблица X.2

Наименование работы	Номер работы	Номера непосредственно предшествующих работ
1	2	3
Проведение:		
ходового восстающего	1	–
орта скреперования	2	1, 3
рудоспуска	3	–
1-й пары дучек и орта подсечки над ними	4	2
2-й пары дучек и орта подсечки над ними	5	2
3-й пары дучек и орта подсечки над ними	6	2
отрезного штрека № 1	7	8
отрезного восстающего № 1	8	6
Бурение скважин на подсечке	9	4, 5, 7, 11
Проведение:		
закладочного ходка	10	8, 11
бурового орта	11	–
отрезного штрека № 2	12	8, 11
отрезного восстающего № 2	13	12
Бурение скважин на буровом орте	14	13
Фиктивная работа		

* В основу взаимосвязанных примеров 1–2 и 7–11 положены условия примера, заимствованного у Ф.Н. Рогинского (Планирование развития горных работ на рудниках. М., Недра, 1978).

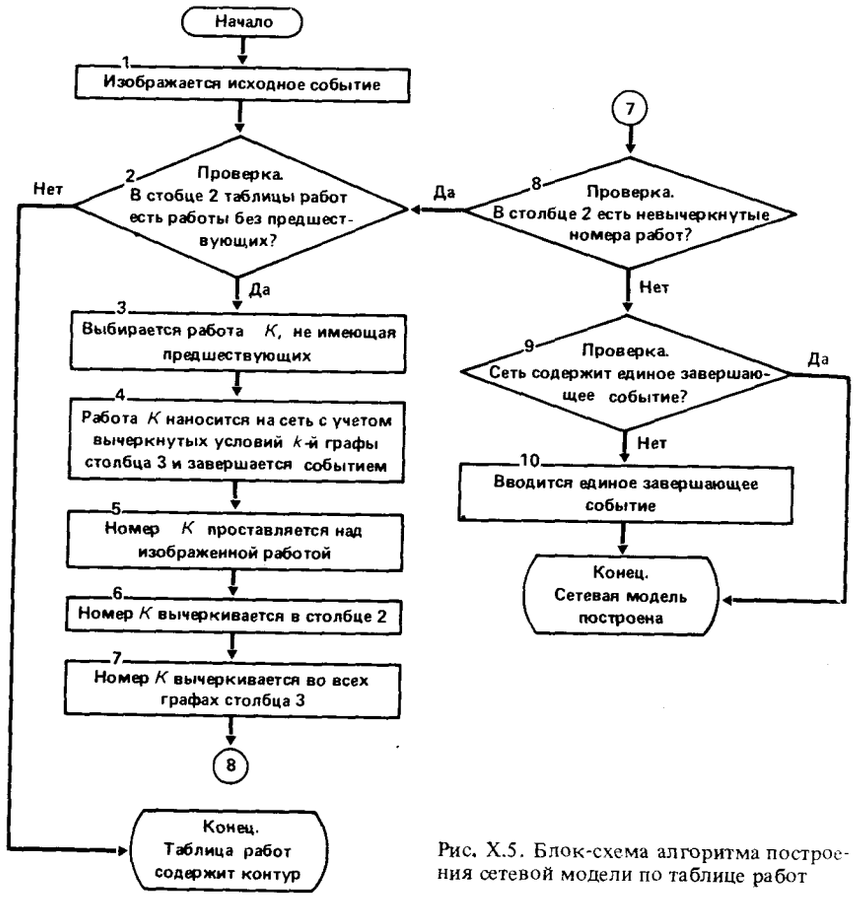
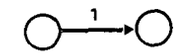
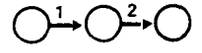
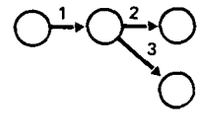
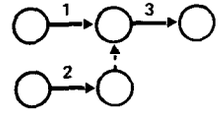
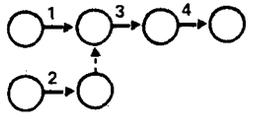
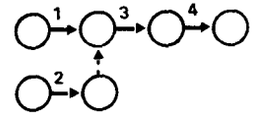
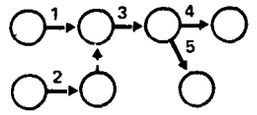
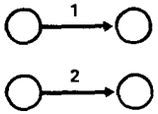
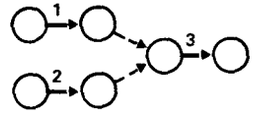
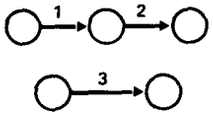
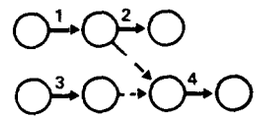


Рис. X.5. Блок-схема алгоритма построения сетевой модели по таблице работ

Затем работам, перечисленным в столбце 1 табл. X.2, присвоим порядковые номера, начиная с 1 (столбец 2). Против номера каждой работы в столбце 3 укажем номера непосредственно предшествующих работ исходя из следующих условий проведения выработок в камере: скреперный орт проводится после ходового восстающего и рудоспуска; дучки, орт подсечки, отрезной штрек и восстающий № 1 проводятся после того, как будет обеспечена сквозная вентиляция по горизонту скреперования; отрезные штреки и восстающие не должны проводиться одновременно; буровой орт проводится из откаточного штрека верхнего горизонта; все выработки выше бурового орта проводятся, а подсечка разбуривается после того, как будет обеспечена сквозная вентиляция камеры за счет сбойки орта с отрезным восстающим № 1; закладочный ходок проводится с бурового орта, одновременно можно бурить скважины с орта.

Построение сетевой модели осуществляется на основании информации, записанной в столбцах 2 и 3 табл. X.2, по специальному алгоритму (рис. X.5), сущность которого заключается в циклическим процессе

Таблица X.3.

Изображено на сети	Указано в таблице работ		Следует изобразить на сети
	номер работы	номера предшествующих работ	
1. 	2	1	
2.	3	1	
3. 	4	1,2,3	
4. 	5	2,3	
5. 	3	1,2	
6. 	4	1,3	

нанесения на сеть одной работы за другой до тех пор, пока не будет построена модель или выявлен контур в таблице работ.

Построение начинается с изображения исходного события, затем идет циклический процесс нанесения работ от ранее построенных событий. В течение одной итерации выбирается работа, не имеющая предшествующих по условию или вследствие того, что номера предшествующих работ вычеркнуты. Эта работа и ее конечное событие наносятся на сеть, а ее номер вычеркивается в столбцах 2 и 3. Если такой работы найти не удалось, то в таблице имеется контур и ее следует скорректировать.

Работа наносится на сеть с учетом всех предшествующих операций, указанных в соответствующей графе столбца 3.

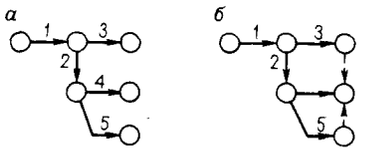


Рис.X.6. Введение в сеть *а* единого завершающего события *б*

Работы, имеющие только одну предшествующую, строятся от ее конечного события (табл. X.3, строки 1 и 2).

Если предшествующие работы лежат на путях, имеющих одно и то же конечное событие, то рассматриваемая работа строится от этого события (табл. X.3, строки 3 и 4). Этот случай возникает при избыточной информации в таблице работ.

Работы, зависящие от нескольких параллельных операций, изображаются от единого конечного события фиктивных работ, отложенных

Таблица X.4.

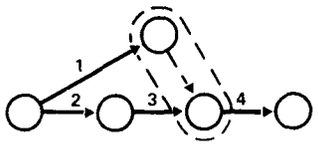
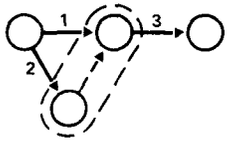
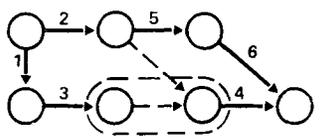
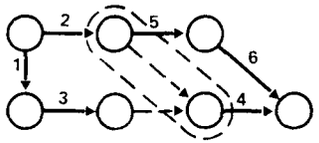
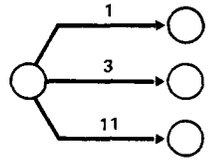
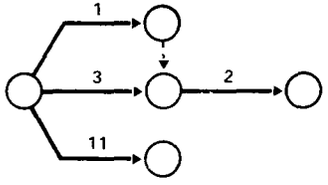
Вариант исключения фиктивной работы	Условия выполнения последующих работ			Вывод
	номер работы	номера последующих работ		
		до исключения	после исключения	
1. 	4	1,3	1,3	Исключить
2. 	3	1,2	1,2	Не исключать — работы 1 и 2 будут параллельны
3. 	4	2,3	2,3	Исключить
4. 	4 5	2,3 2	2,3 2,3	Не исключать — нарушаются условия работы 5

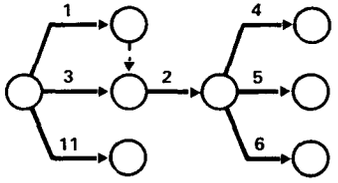
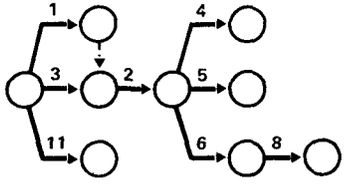
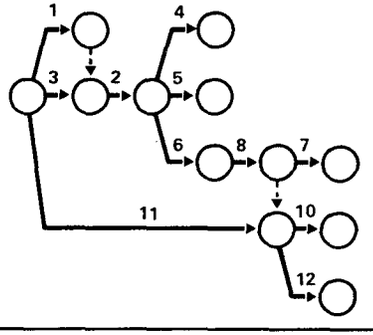
Таблица X.5.

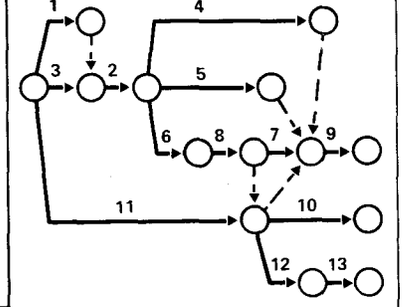
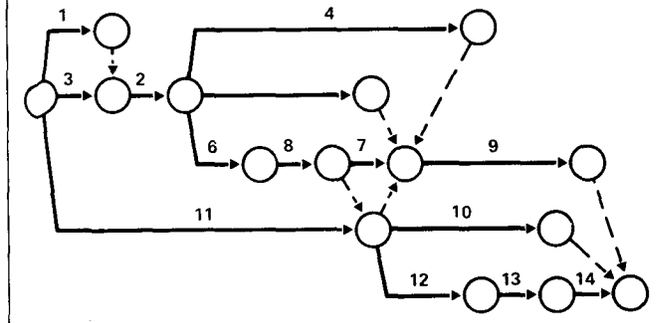
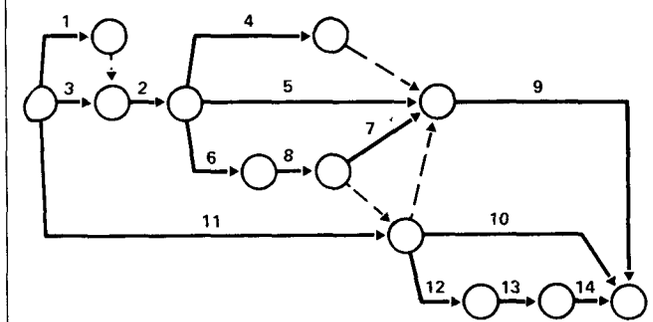
Номер этапа	Результаты построения, полученные на этапе		Примечания
1	2		3
1	Номер работы / Номера непосредственно предшествующих работ: ~~1~~ — 2 — ~~1~~, ~~3~~ ~~3~~ — 4 — 2 5 — 2 6 — 2 7 — 8 8 — 6 9 — 4, 5, 7, ~~11~~ 10 — 8, ~~11~~ ~~11~~ — 12 — 8, ~~11~~ 13 — 12 14 — 13 		Представлены результаты построения модели и таблица работ после первых трех последовательных итераций по изображению работ 1,3 и 11
2	~~1~~ — ~~2~~ — ~~1~~, ~~3~~ ~~3~~ — 4 — ~~2~~ 5 — ~~2~~ 6 — ~~2~~ 7 — 8 8 — 6 9 — 4, 5, 7, ~~11~~ 10 — 8, ~~11~~ ~~11~~ — 12 — 8, ~~11~~ 13 — 12 14 — 13 		—

от конечных событий указанных предшествующих операций (табл. X.3, строки 5 и 6).

Для введения в сеть единого завершающего события одно из конечных событий выбирается завершающим, а все остальные соединяются с ним фиктивными работами (рис. X.6).

Построенная по указанным правилам сетевая модель может содержать лишние фиктивные работы, без которых не нарушается установленная последовательность выполнения операций программы. Для проверки возможности исключения фиктивной работы объединяют ее начальное и конечное события в одно и проверяют сохранение условий

1	2	3
3	1 — 2 1, 3 3 — 4 2 5 2 6 2 7 8 8 8 9 4, 5, 7, 11 10 8, 11 11 — 12 8, 11 13 12 14 13 	Представлены результаты построения модели и таблица работ после трех последовательных итераций по изображению работ 4, 5, 6
4	1 — 2 1, 3 3 — 4 2 5 2 6 2 7 8 8 8 9 4, 5, 7, 11 10 8, 11 11 — 12 8, 11 13 12 14 13 	—
5	1 — 2 1, 3 3 — 4 2 5 2 6 2 7 8 8 8 9 4, 5, 7, 11 10 8, 11 11 — 12 8, 11 13 12 14 13 	Представлены результаты построения модели и таблица работ после трех последовательных итераций по изображению работ 7, 10, 12

1	2	3
6	1 — 2 1, 3 3 — 4 2 5 2 6 2 7 8 8 8 9 4, 5, 7, 11 10 8, 11 11 — 12 8, 11 13 12 14 13 	Представлены результаты построения модели и таблица работ двух последовательных итераций по изображению работ 9 и 13
7		Изображением работы 14 и введением конечного события заканчивается построение сетевой модели
8		Сетевая модель после исключения "лишних" фиктивных работ

выполнения последующих работ. При этом не должно появляться параллельных работ. Примеры проведения анализа условий исключения фиктивных работ приведены в табл. X.4.

Таблица X.6

Наименование работы	Номер работы	Номера непосредственно предшествующих работ
Планировка рабочей площадки блока	1	–
Отгон оборудования и снятие коммуникаций с блока	2	–
Маркшейдерская съемка блока	3	1
Составление проекта буровых работ	4	3
Вынесение проекта на блок (''разбивка'' скважин)	5	4
Монтаж отпайки ЛЭП, установка трансформаторного киоска	6	–
Подгон буровых станков на блок	7	–
Подключение станков и подготовка их к работе	8	6, 7
Бурение скважин (создание минимально необходимого опережения)	9	2, 5, 8
Бурение скважин (создание максимально допустимого опережения)	10	11
Маркшейдерская съемка обуренной части блока	11	9
Расчет зарядов скважин	12	11
Доставка одной партии взрывчатых веществ на блок	13	2, 5, 8
Доставка средств взрывания на блок	14	2, 5, 8
Подготовка арматуры для воздушных промежутков, боевиков и т.д.	15	13, 14
Замеры фактической глубины скважин	16	9
Зачистка (разбуривание) скважин	17	16
Разноска взрывчатых веществ по блоку	18	13
Зарядка скважин	19	12, 17, 18
Доставка забоечного материала	20	12, 17, 18
Забойка заряженных скважин	21	20
Ограждение заряженной части блока	22	19

Пример X.2. Построить сетевую модель программы ввода камеры в эксплуатацию на руднике по условиям, записанным в табл. X.2. (см. пример X.1).

Этапы построения сетевой модели в соответствии с алгоритмом представлены в табл. X.5.

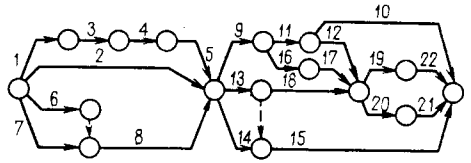
*Пример X.3**. Построим сетевую модель программы цикла работ на карьере по подготовке блока к взрыванию по условиям, записанным в табл. X.6. Указанные условия подразумевают параллельное ведение буровых и зарядных работ. В цикл входят операции по подготовке (бурению и зарядке) группы скважин, заряжаемых одной партией взрывчатых веществ.

Результаты построения модели представлены на рис. X.7.

Пример X.4. Построить сетевую модель программы работ проходческого цикла в горизонтальной выработке по условиям, записанным в табл. X.7.

* Пример заимствован у В.И. Ганицкого (Организация производства на карьерах. М., Недра, 1983).

Рис. X.7. Сетевая модель программы цикла работ на карьере по подготовке блока к взрыванию



Результаты построения модели представлены на рис. X.8.

*Пример X.5**. Построить сетевую модель программы монтажа оборудования механизированного комплекса в лаве угольной шахты по условиям, записанным в табл. X.8.

Результаты построения модели представлены на рис. X.9.

*Пример X.6***. Построить сетевую модель программы тушения пожара в начальной стадии в выработанном пространстве лавы (шахты опасной по газу при применении сплошной или столбовой системы разработки, возвратноточном проветривании участков и восходящем направлении движения воздуха) по условиям, записанным в табл. X.9.

Результаты построения модели представлены на рис. X.10.

Особенностью рассматриваемой программы является то, что работы, заключающиеся в тушении пожара определенными средствами, считаются завершенными к моменту выполнения всех операций, входящих в конечное для указанных работ

Таблица X.7

Наименование работы	Номер работы	Номера непосредственно предшествующих работ
Подготовка рабочего места и бурового оборудования	1	–
Бурение шпуров	2	1
Подготовка забойки	3	–
Подготовка ВМ	4	–
Заряжение шпуров	5	2, 3, 4
Взрывные работы	6	5
Проветривание забоя	7	6
Приведение забоя в безопасное состояние	8	7
Подготовка погрузочной машины к работе	9	8
Обеспечение забоя порожними вагонетками	10	–
Уборка породы	11	9, 10
Обеспечение забоя крепью и рельсами	12	–
Установка крепи	13	11, 12
Настилка рельсового пути	14	11, 12
Заключительные работы	15	13, 14

* Пример заимствован у Б.Ф. Негруцкого (Интенсификация монтажа оборудования угольных шахт. М., Недра, 1983).

** В основу примера положена типовая тактико-технологическая схема устава ВГСЧ (Устав ВГСЧ по организации и ведению горноспасательных работ. М., Недра, 1986).

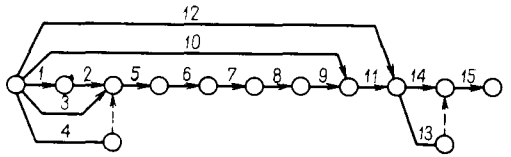


Рис. X.8. Сетевая модель программы работ проходческого цикла

Таблица X.8

Наименование работы	Номер работы	Номера непосредственно предшествующих работ
Установка монтажных приспособлений и устройств	1	–
Устройство настила, укладка направляющих, настилка рельсовых путей	2	–
Устройство приемной площадки	3	–
Устройство комплектовочной площадки	4	–
Доставка:		
концевого привода конвейера	5	1, 2, 3, 4
крепи верхнего сопряжения	6	1, 2, 3, 4
комбайна	7	1, 2, 3, 4
секций крепи	8	1, 2, 3, 4
забойного конвейера	9	1, 2, 3, 4
оборудования энергопоезда	10	1, 2, 3, 4
монтажного гидрооборудования	11	1, 2, 3, 4
головного привода конвейера	12	1, 2, 3, 4
крепи нижнего сопряжения	13	1, 2, 3, 4,
штрекового оборудования	14	1, 2, 3, 4
Монтаж энергопоезда	15	10
Установка монтажного гидрооборудования	16	11
Монтаж крепи нижнего сопряжения	17	13
Установка головного привода конвейера	18	12
Монтаж штрекового оборудования	19	14
Монтаж линейных секций конвейера	20	9, 15, 16, 17, 18
Укладка нижней цепи конвейера	21	9, 15, 16, 17, 18
Установка навесного оборудования	22	9, 15, 16, 17, 18
Прокладка магистральных трубопроводов	23	9, 15, 16, 17, 18
Монтаж секций крепи	24	8, 20, 21, 22, 23
Установка направляющих балок	25	8, 20, 21, 22, 23
Подсоединение домкратов	26	8, 20, 21, 22, 23
Крепление трубопроводов и подсоединение рукавов	27	8, 22, 20, 21, 23
Демонтаж и выдача из лавы монтажной оснастки	28	24, 25, 26, 27, 19
Сборка траковой цепи кабелеукладчика	29	24, 25, 26, 27, 19
Монтаж аппаратуры управления, связи, сигнализации и освещения	30	24, 25, 26, 27, 19
Пуско-наладочные работы	31	24, 25, 26, 27, 19
Монтаж комбайна	32	7, 28, 29
Укладка верхней цепи конвейера	33	7, 28, 29
Установка концевого привода конвейера	34	5, 32, 33
Монтаж крепи верхнего сопряжения	35	6, 32, 33

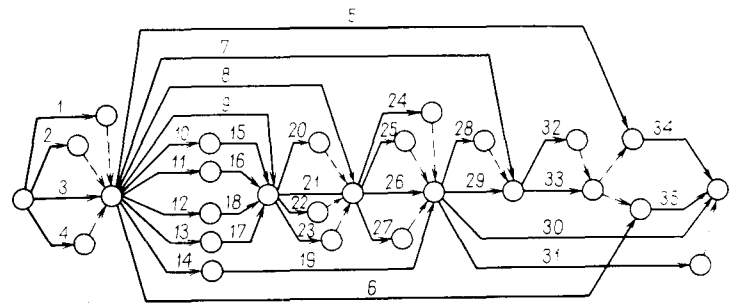


Рис. X.9. Сетевая модель программы монтажа оборудования механизированного комплекса

Таблица X.9

Наименование работы	Номер работы	Номера непосредственно предшествующих работ
Тушение пожара огнетушителями и водой со стороны свежей струи воздуха силами вспомогательной горноспасательной команды (ВГК) аварийного участка	1	–
Доставка к месту аварии членов ВГК других участков	2	–
Доставка средств пожаротушения с других участков	3	–
Введение в действие водяной завесы на вентиляционном штреке	4	–
Обеспечение мер безопасности (усиление дегазации, контроль за содержанием метана, выявление скоплений пыли и принятие необходимых мер)	5	–
Выполнение подготовительных работ по дальнейшему тушению пожара (обеспечение бесперебойной подачи воды, определение главного направления работ)	6	–
Тушение пожара огнетушителями и водой со стороны свежей струи силами ВГК аварийного и других участков	7	1, 2, 3
Доставка с поверхности оборудования и мощных порошковых средств тушения и локализации пожара	8	1, 2, 3
Тушение и локализация пожара водой и огнегасящим порошком, подаваемых непосредственно на очаг горения из всех доступных мест	9	1, 2, 3, 4, 5, 6, 7, 8
Установка водоразбрызгивателя в вентиляционном штреке и лаве для локализации пожара	10	1, 2, 3, 4, 5, 6, 7, 8
Принятие мер по увеличению подачи воды на аварийный участок	11	1, 2, 3, 4, 5, 6, 7, 8

Продолжение табл. X.9

Наименование работы	Номер работы	Номера непосредственно предшествующих работ
Обеспечение мер безопасности (контроль развития пожара, ликвидация угольной пыли, доставка средств противотепловой защиты, принятие мер по ликвидации нарушений проветривания из-за подтопления водой)	12	1, 2, 3, 4, 5, 6, 7, 8
Оборудование площадок для размещения средств объемного дистанционного тушения	13	1, 2, 3, 4, 5, 6, 7, 8
Доставка с поверхности средств объемного дистанционного тушения	14	1, 2, 3, 4, 5, 6, 7, 8
Подготовка оборудования и организация бригады для выполнения работ по аварийной дегазации	15	1, 2, 3, 4, 5, 6, 7, 8
Доставка на шахту материалов и оборудования, необходимых при различных вариантах локализации пожара	16	1, 2, 3, 4, 5, 6, 7, 8
Осуществление мероприятий при отключении дегазации	17	1, 2, 3, 4, 5, 6, 7, 8

событие. Так, например, работа 1 считается выполненной, когда выполнены работы 2 и 3. Это отражает принцип последовательного ввода сил и средств в тушении, предусмотренный уставом ВСГЧ.

Если на каком-либо этапе пожар удалось потушить, то программа считается выполненной.

В противном случае в результате выполнения работ программы подготавливается тушение развившегося пожара.

Для проведения последующих расчетов временны́х параметров все события сетевой модели должны иметь неповторяющиеся номера. Обычно указанные номера составляют сплошной ряд натуральных чисел, что объясняется удобством реализации вычислительных процедур СПУ на ЭВМ. Нумерацию выполняют со строгим соблюдением условия – каждое последующее событие должно иметь номер больше, чем все предыдущие.

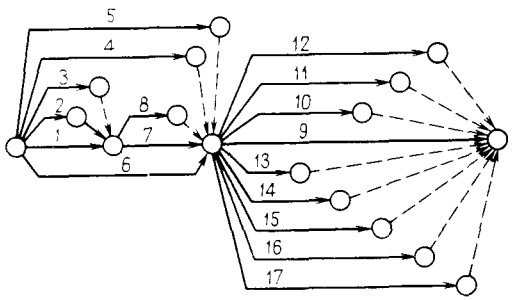


Рис. X.10. Сетевая модель программы работ по тушению пожара в начальной стадии в выработанном пространстве лавы

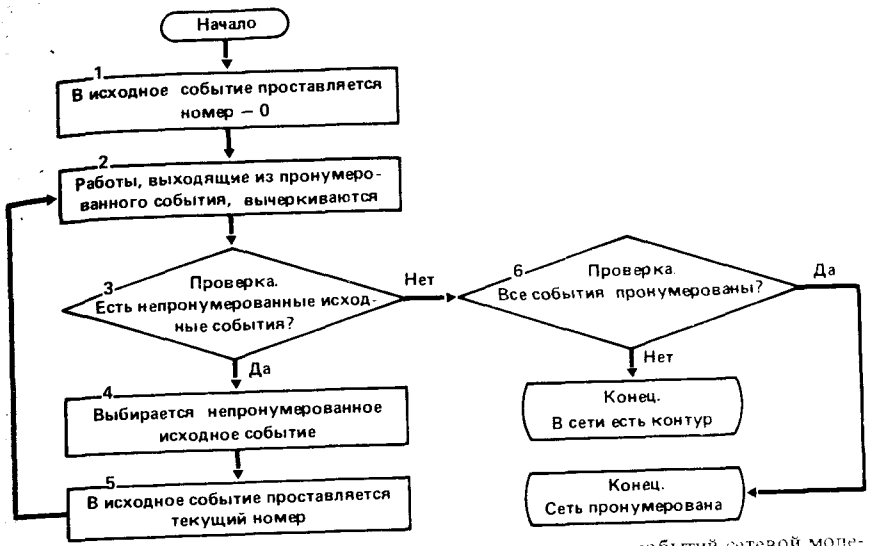


Рис. X.11. Блок-схема алгоритма упорядоченной нумерации событий сетевой модели методом вычеркивания работ

Упорядоченную нумерацию событий проводят по алгоритму (рис. X.11), основанному на использовании метода вычеркивания работ.

Алгоритм предусматривает последовательную нумерацию исходных событий, образующихся после вычеркивания (условного исключения) работ, выходящих из уже пронумерованных событий.

Если в процессе нумерации обнаружился контур, то следует проверить правильность изображения на сети работ, ограничивающих непронумерованный участок.

Нумерацию событий следует начинать с нуля.

После завершения нумерации, индексы начальных и конечных событий работ, включая фиктивные, заносятся в столбцы 4 и 5 таблицы работ.

Пример X.7. Пронумеровать события сетевой модели программы ввода камеры в эксплуатацию (см. пример X.2).

Последовательность действий по присвоению событиям сетевой модели номеров в соответствии с рассмотренным алгоритмом иллюстрируется на рис. X.12.

§ 3. РАСЧЕТ ВРЕМЕННЫХ ПАРАМЕТРОВ СЕТЕВОЙ МОДЕЛИ

Конечная цель применения сетевых методов планирования – построение календарного плана программы, определяющего сроки начала и окончания каждой операции. Определение указанных сроков осуществляется путем специальных расчетов на основе информации о взаимо-

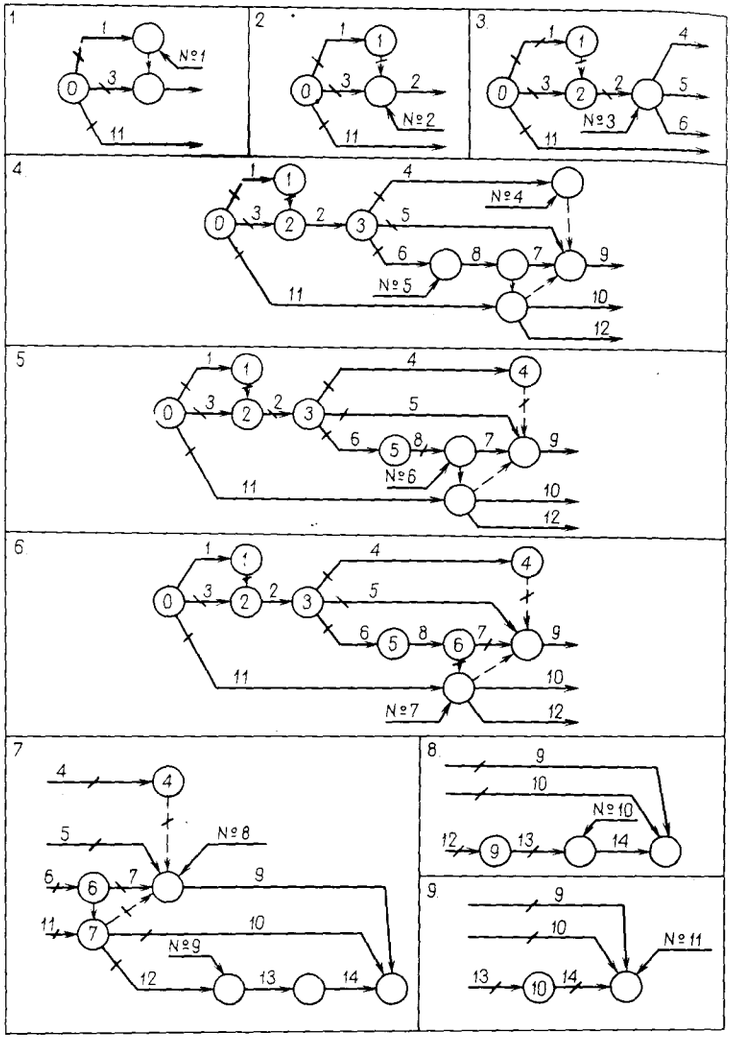


Рис.X.12. Этапы нумерации сетевой модели программы ввода камеры в эксплуатацию по методу вычеркивания работ

связи и длительности операции программы по построенной сетевой модели.

В результате расчетов определяются временные параметры сети, к числу которых относятся: ранние и поздние сроки наступления и резервы времени событий, полные и свободные резервы времени работ.

Остановимся сначала на проблеме оценки продолжительности каждой из работ программы.

Наиболее простым является случай, когда длительности операций известны и точно определены, т.е. детерминированы. В такой ситуации каждой работе программы ставится в соответствие ее продолжительность и осуществляется расчет временных параметров модели. Однако этот наиболее простой случай встречается крайне редко.

На практике для подавляющегося количества программ длительности операций недостаточно известны и неопределены. Причем здесь могут иметь место два случая:

неопределенные операции не являются новыми и закон распределения длительности каждой из них приблизительно известен;

неопределенные операции являются новыми или, по крайней мере, малоизученными; законы распределения неизвестны.

Первый случай характерен, как правило, для опытно-конструкторских программ в промышленности и программ производства, второй — для программ научно-исследовательских работ.

Если известны законы распределения длительностей $\widetilde{t}_{i,\,j}$ операций, то на их основе определяют среднее значение длительности – математическое ожидание $\overline{t}_{i,\,j}$ и дисперсию σ^2_{ij} времени выполнения каждой из операций.

Если распределения длительностей неизвестны, то при применении методов СПУ принято условно считать, что длительности работ имеют β-распределение. Вопрос о допустимости такого предложения остается до настоящего времени открытым. Причем, β-распределение выбирается таким образом, что

$$\overline{t}_{i,\,j} = \frac{1}{6}\,(a_{i,\,j} + 4\,m_{i,\,j} + b_{i,\,j}), \qquad \text{(X. 1)}$$

$$\sigma^2_{i,\,j} = \left(\frac{b_{i,\,j} - a_{i,j}}{6}\right)^2, \qquad \text{(X.2)}$$

где $a_{i,\,j}$ и $b_{i,\,j}$ – нижняя и верхняя границы случайной величины соответственно; $m_{i,\,j}$ – наиболее вероятное значение или мода $\widetilde{t}_{i,\,j}$ (рис. X.13).

Следовательно, исходной информацией для расчета параметров распределения длительности по формуле (X.1) и (X.2) являются значения

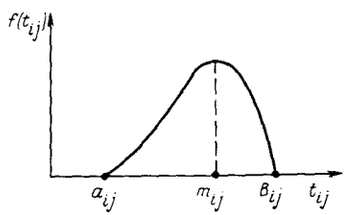


Рис.X.13. Плотность β-распределения, используемая для оценки длительности работ

трех различных несмещенных оценок продолжительности каждой операции (i, j):

$\hat{a}_{i,\,j}$ – оптимистической (минимальной), соответствующей наиболее благоприятным условиям выполнения операции;

$\hat{b}_{i,\,j}$ – пессимистической (максимальной), соответствующей самым неблагоприятным условиям выполнения операции;

$\hat{m}_{i,\,j}$ – наиболее вероятной (нормальной), характеризующей усредненные условия выполнения операций.

Значения $\hat{a}_{i,\,j}$, $\hat{b}_{i,\,j}$, $\hat{m}_{i,\,j}$ получают с помощью специалистов, ответственных за выполнение соответствующих операций, и полагают, что при их подстановке в формулы (X.1) и (X.2) будут определены оценки математического ожидания $\hat{t}_{i,\,j}$ и дисперсии $\hat{\sigma}^2_{ij}$.

Таким образом, каждой работе (i, j) программы соответствует оценка математического ожидания длительности ее выполнения $t_{i,\,j}$, которая в дальнейшем будет называться просто длительностью ее выполнения и обозначаться $t_{i,\,j}$ (как и в детерминированном случае), и оценка дисперсии $\hat{\sigma}^2_{i,\,j}$, в дальнейшем σ^2_{ij}. Указанные оценки играют важную роль при вероятностном анализе сетевой модели, выполняемом с целью определения надежности реализации всей программы или отдельных ее частей в директивный срок.

Перейдем теперь к рассмотрению методики определения значений временных параметров сетевой модели.

Первоначально вычисляются параметры событий, а затем на их основе временные параметры работ.

С организационной точки зрения ранний срок наступления события – это минимально необходимое время для выполнения всех предшествующих работ. Отсюда ранний срок наступления конечного события представляет собой минимальное время, необходимое для выполнения программы. Это время необходимо для реализации самого длинного полного пути, соединяющего начало и конец сети, который называется **критическим путем.**

Поздний срок наступления события – это максимально допустимое время, исходя из условия сохранения длительности критического пути, для выполнения всех предшествующих работ. Под предшествующими работами здесь понимаются совокупности работ, составляющие пути от начального события до данного.

Напомним, что событие наступает только тогда, когда выполнены все входящие в него работы. Следовательно, определение раннего срока есть задача нахождения длительности максимального по времени пути *из начального события до данного*, а *определение позднего срока* – задача нахождения разности между длительностями критического пути и максимального пути от данного события до конечного. Причем, понятно, что последняя задача может быть сведена к определению максимальной разности между длительностями критического пути и каждого из путей, соединяющего данное событие с конечным.

Следовательно, поиск значений сроков наступления событий может быть интерпретирован как задача определения оптимальной траектории на сети, которая, как известно, решается методами динамического программирования, базирующегося на применении рекуррентного соотношения. Рассматриваемые далее формулы и представляют собой запись рекуррентного соотношения в обозначениях, принятых в СПУ.

Расчет временных параметров событий начинается с ранних сроков наступления событий, определяемых по формуле

$$T_{\text{р}\,j} = \max_{i \in U_i} (T_{\text{р}\,i} + t_{ij});\ \ j = 1, 2, \ldots, n, \qquad \text{(X.3)}$$

где $T_{\text{р}\,i}$ – ранний срок свершения события j; U_i – множество событий, непосредственно предшествующих событию j; $T_{\text{р}\,i}$ – ранний срок наступления события i, $i \in U_i$; $t_{i,\,j}$ – продолжительность работы (i, j); n – номер завершающего события сетевой модели.

Определение $T_{\text{р}\,i}$ ведут строго по порядку номеров, начиная с исходного события, приняв ранний срок поступления исходного события равным нулю

$$T_{\text{р}0} = 0. \qquad \text{(X.4)}$$

Вычисления продолжают вплоть до завершающего n-го события, ранний срок которого определяет длительность S критического пути сетевой модели

$$S = T_{\text{р}\,n}. \qquad \text{(X.5)}$$

Значения поздних сроков, как было указано выше, определяются из условий сохранения срока выполнения программы. Отсюда поздний срок наступления завершающего события равен

$$T_{\text{п}\,n} = T_{\text{р}\,n} = S. \qquad \text{(X.6)}$$

Дальнейшие вычисления выполняют начиная с последнего события в порядке строгого убывания номеров по формуле

$$T_{\text{п}\,i} = \max_{i \in U_i} (T_{\text{п}j} - t_{i,\,j}),\ \ i = n-1,\ n-2, \ldots, 0, \qquad \text{(X.7)}$$

где $T_{\text{п}\,i}$ – поздний срок наступления события i; U_j – множество событий, непосредственно следующих за событием i; $T_{\text{п}\,j}$ – поздний срок наступления событий j, $j \in U_j$. Причем, поздний срок исходного события должен быть равен 0. В противном случае в расчетах допущена ошибка.

Резерв времени события, по своему организационному смыслу, определяет максимально допустимую с точки зрения неизменности длительности критического пути задержку срока наступления события по отношению к его раннему сроку и, следовательно, вычисляется по формуле

$$R_i = T_{\text{п}\,i} - T_{\text{р}i},\ \ i = 0, 1, \ldots, n. \qquad \text{(X.8)}$$

Из определений резерва и критического пути следует, что для всех событий последнего резерв времени равен нулю

$$R_i = 0, \; i \in Z'_{кр} \; , \qquad (X.9)$$

где $Z'_{кр}$ – множество событий, входящих в критический путь.

Результаты расчетов параметров событий заносятся в специальную форму (форма X.2), которая в дальнейшем будет именоваться "таблица событий".

Форма X.2

Индекс события	Ранний срок наступления события $T_{рi}$	Поздний срок наступления события $T_{пi}$	Резерв времени наступления события R_i
1	2	3	4

Рассмотрим сущность и методику расчета временных параметров работ.

С организационной точки зрения резервы времени работы (свободный и полный) показывают, насколько можно увеличить ее длительность или задержать срок ее начала, причем свободный – при условии неизменности сроков выполнения других работ, а полный – при сохранении длительности критического пути.

Следовательно, свободные резервы времени определяются по формуле

$$\Delta_{i,j} = T_{рj} - T_{рi} - t_{i,j}, \; (i,j) \in Q, \qquad (X.10)$$

а полные резервы времени по формуле

$$r_{i,j} = R_j + \Delta_{i,j} = T_{пj} - T_{рi} - t_{i,j}, \; (i,j) \in Q, \qquad (X.11)$$

где $\Delta_{i,j}$ – свободный резерв времени работы (i,j); $r_{i,j}$ – полный резерв времени работы (i,j); Q – множество работ сетевой модели.

Как следует из формулы (X.3) и определения критического пути для его работ, справедливо соотношение

$$T_{рj} = T_{рi} + t_{i,j}, \; (i,j) \in Z_{кр}. \qquad (X.12)$$

Кроме того, как уже было показано выше, для событий критического пути справедливо равенство (X.9). Следовательно, все резервы времени критических работ равны нулю

$$r_{i,j} = \Delta_{i,j} = 0; \; (i,j) \in Z_{кр}, \qquad (X.13)$$

где $Z_{кр}$ – множество индексов работ критического пути.

Расчет резервов времени работ позволяет определить критический путь сетевой модели, что важно с практической точки зрения, так как изменение длительности любой из критических работ отражается на сроке выполнения программы.

Результаты расчета резервов времени работ заносятся в соответствующие столбцы таблицы работ.

В дальнейшем, при улучшении (оптимизации) календарного плана программы, моменты начала некоторого множества работ Q' могут быть сдвинуты во времени, т.е. для работ указанного множества введены задержки начала их выполнения $g_{i,j}$ ($g_{i,j} \geqslant 0$; $(i,j) \in Q'$). В этом случае расчет временных параметров осуществляется по формулам (X.3), (X.7), (X.10) и (X.11), в которых вместо $t_{i,j}$ используется сумма $t_{i,j} + g_{i,j}$.

После расчета параметров осуществляется временной анализ сетевой модели, который заключается в определении критических и некритических операций программы, а также сравнении директивного времени реализации программы T с продолжительностью критического пути S.

Информация о критических и некритических операциях, а также о резервах времени последних имеет огромное значение для процесса оптимизации календарного плана.

Сравнение значений величин T и S дает ответ на вопрос о возможности реализации программы за установленный директивный срок.

В случае, если вероятностным характером процесса реализации программы нельзя пренебречь, дополнительно оценивается надежность (вероятность) P реализации программы за время T

$$P\left\{\widetilde{S} < T\right\} = P\left\{\sum_{(i,j) \in Z_{кр}} \widetilde{t}_{i,j} < T\right\} = \Phi\left\{\frac{T - \sum_{(i,j) \in L_{кр}} t_{i,j}}{\sqrt{\sum_{(i,j) \in L_{кр}} \sigma^2_{i,j}}}\right\}, \qquad (X.14)$$

где Φ – функция нормированного нормального распределения; $\widetilde{t}_{i,j}$ – случайная длительность операции (i,j) с параметрами распределения $\overline{t}_{i,j}$ и $\sigma_{i,j}^2$.

Использование функции нормального распределения для определения надежности базируется на центральной предельной теореме теории вероятностей. Такой подход является оправданным при достаточно большом количестве работ критического пути и независимости их длительностей. Заметим, что при условии β-распределения времени выполнения операций, принятом выше, сумма даже небольшого числа независимых случайных величин довольно точно удовлетворяет нормальному распределению.

При использовании формулы (X.14) необходимо помнить, что при наличии нескольких критических путей в сетевой модели, множество $Z_{кр}$ включает только индексы критических работ пути с максимальной суммарной дисперсией.

Пример X.8. Определить временные параметры сетевой модели программы ввода камеры на руднике в эксплуатацию (рис. X.14) при длительностях работ, заданных в табл. X.10.

Первоначально вычислим параметры событий (в сут):

$$T_{p0} = 0$$

$$T_{p1} = T_{p0} + t_{0,1} = 0 + 3 = 3,$$

$$T_{p2} = \max \begin{Bmatrix} T_{p1} + t_{1,2} \\ T_{p0} + t_{0,2} \end{Bmatrix} = \max \begin{Bmatrix} 3 + 0 \\ 0 + 3 \end{Bmatrix} = 3,$$

$$T_{p3} = T_{p2} + t_{2,3} = 3 + 16 = 19,$$

$$T_{p4} = T_{p3} + t_{3,4} = 19 + 6 = 25,$$

$$T_{p5} = T_{p3} + t_{3,5} = 19 + 6 = 25,$$

$$T_{p6} = T_{p5} + t_{5,6} = 25 + 12 = 37,$$

$$T_{p7} = \max \begin{Bmatrix} T_{p0} + t_{0,7} \\ T_{p6} + t_{6,7} \end{Bmatrix} = \max \begin{Bmatrix} 0 + 17 \\ 37 + 0 \end{Bmatrix} = 37,$$

$$T_{p8} = \max \begin{Bmatrix} T_{p4} + t_{4,8} \\ T_{p3} + t_{3,8} \\ T_{p6} + t_{3,6} \\ T_{p7} + t_{7,8} \end{Bmatrix} = \max \begin{Bmatrix} 25 + 0 \\ 19 + 6 \\ 37 + 7 \\ 37 + 0 \end{Bmatrix} = 44,$$

$$T_{p9} = T_{p7} + t_{7,9} = 37 + 7 = 44,$$

$$T_{p10} = T_{p9} + t_{9,10} = 44 + 9 = 53,$$

$$T_{p11} = \max \begin{Bmatrix} T_{p3} + t_{8,11} \\ T_{p7} + t_{7,11} \\ T_{p10} + t_{10,11} \end{Bmatrix} = \max \begin{Bmatrix} 44 + 35 \\ 37 + 18 \\ 53 + 30 \end{Bmatrix} = 83,$$

$$T_{п11} = T_{p11} = 83,$$

$$T_{п10} = T_{п+1} - t_{10,11} = 83 - 30 = 53,$$

$$T_{п9} = T_{п10} - t_{9,10} = 53 - 9 = 44,$$

$$T_{п8} = T_{п11} - t_{8,11} = 83 - 36 = 48,$$

$$T_{п7} = \min \begin{Bmatrix} T_{п8} - t_{7,8} \\ T_{п11} - t_{7,11} \\ T_{п9} - t_{7,9} \end{Bmatrix} = \min \begin{Bmatrix} 48 - 0 \\ 83 - 16 \\ 44 - 7 \end{Bmatrix} = 37,$$

$$T_{п6} = \min \begin{Bmatrix} T_{п8} - t_{6,8} \\ T_{п7} - t_{6,7} \end{Bmatrix} = \min \begin{Bmatrix} 48 - 7 \\ 37 - 0 \end{Bmatrix} = 37,$$

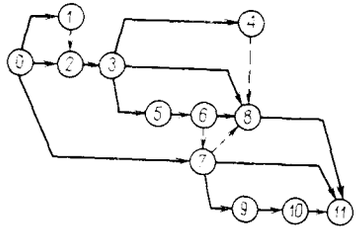


Рис. X.14. Сетевая модель программы ввода камеры в эксплуатацию

Таблица X.10

Наименование работы	Номер работы	Номера непосредственно предшествующих работ	i	j	$t_{i,j}$	$r_{i,j}$	$\Delta_{i,j}$
Проведение:							
ходового восстающего	1	–	0	1	3	0	0
орта скреперования	2	1, 3	2	3	16	0	0
рудоспуска	3	–	0	2	3	0	0
1-й пары дучек и орта подсечки над ними	4	2	3	4	6	23	0
2-й пары дучек и орта подсечки над ними	5	2	3	8	6	23	19
3-й пары дучек и орта подсечки над ними	6	2	3	5	6	0	0
отрезного штрека № 1	7	8	6	8	7	4	0
отрезного восстающего № 1	8	6	5	6	12	0	0
Бурение скважин на подсечке	9	4, 5, 7, 11	8	11	35	4	4
Проведение:							
закладочного ходока	10	8, 11	7	11	18	28	28
бурового орта	11	–	0	7	17	25	25
отрезного штрека № 2	12	8, 11	7	9	7	0	0
отрезного восстающего № 2	13	12	9	10	9	0	0
Бурение скважин на буровом орте	14	13	10	11	30	0	0
Фиктивная работа			1	2	0	0	0
То же			4	8	0	23	19
» »			7	8	0	11	7
» »			6	7	0	0	0

Примечание. Столбцы 7 и 8 заполнены по результатам расчетов, выполненных в примере.

$$T_{п\,5} = T_{п\,6} - t_{5,6} = 37 - 12 = 25,$$

$$T_{п\,4} = T_{п\,8} - t_{4,8} = 48 - 0 = 48,$$

$$T_{п\,3} = \min\begin{Bmatrix} T_{п\,4} - t_{3,4} \\ T_{п\,8} - t_{3,8} \\ T_{п\,5} - t_{3,5} \end{Bmatrix} = \min\begin{Bmatrix} 48 - 6 \\ 48 - 6 \\ 25 - 6 \end{Bmatrix} = 19,$$

$$T_{п\,2} = T_{п\,3} - t_{2,3} = 19 - 16 = 3,$$

$$T_{п\,1} = T_{п\,2} - t_{1,2} = 3 - 0 = 0,$$

$$T_{п\,0} = \min\begin{Bmatrix} T_{п\,1} - t_{0,1} \\ T_{п\,2} - t_{0,2} \\ T_{п\,7} - t_{0,7} \end{Bmatrix} = \min\begin{Bmatrix} 3 - 3 \\ 3 - 3 \\ 37 - 17 \end{Bmatrix} = 0.$$

Результаты вычислений ранних и поздних сроков наступления событий (в сут) запишем во 2 и 3 столбец таблицы событий (табл. X.11). Значения резервов времени событий определим путем нахождения построчных разностей между поздними и ранними сроками.

Анализ данных таблицы событий позволяет сделать следующие выводы:

длительность критического пути сетевой модели S равна 83 сут;

события с индексами 0–3, 5–7, 9–11 принадлежат критическим путям сетевой модели.

На основе вычисленных значений временных параметров событий рассчитаем свободные и полные резервы времени работ (в сут.):

$$\Delta_{0,1} = T_{р\,1} - T_{р\,0} - t_{0,1} = 3 - 0 - 3 = 0,$$
$$\Delta_{2,3} = T_{р\,3} - T_{р2} - t_{2,3} = 19 - 3 - 16 = 0,$$
$$\Delta_{0,2} = T_{р2} - T_{р\,0} - t_{0,2} = 3 - 0 - 3 = 0,$$
$$\Delta_{3,4} = T_{р\,4} - T_{р\,3} - t_{3,4} = 25 - 19 - 6 = 0,$$
$$\Delta_{3,8} = T_{р\,8} - T_{р\,3} - t_{3,8} = 44 - 19 - 6 = 0,$$
$$\Delta_{3,5} = T_{р\,5} - T_{р\,3} - t_{3,5} = 25 - 19 - 6 = 19,$$
$$\Delta_{6,8} = T_{р\,8} - T_{р\,6} - t_{6,8} = 44 - 37 - 7 = 0,$$

Таблица X.11

i	$T_{рi}$	$T_{пi}$	R_i
1	2	3	4
0	0	0	0
1	3	3	0
2	3	3	0
3	19	19	0
4	25	48	23
5	25	25	0
6	37	37	0
7	37	37	0
8	44	48	4
9	44	44	0
10	53	53	0
11	83	83	0

$$\Delta_{5,6} = T_{р\,6} - T_{р\,5} - t_{5,6} = 37 - 25 - 12 = 0,$$
$$\Delta_{8,11} = T_{р\,11} - T_{р\,8} - t_{8,11} = 83 - 44 - 35 = 4,$$
$$\Delta_{7,11} = T_{р\,11} - T_{р\,7} - t_{7,11} = 83 - 37 - 18 = 28,$$
$$\Delta_{0,7} = T_{р\,7} - T_{р\,0} - t_{0,7} = 37 - 0 - 17 = 20,$$
$$\Delta_{7,9} = T_{р\,9} - T_{р\,7} - t_{7,9} = 44 - 37 - 7 = 0,$$
$$\Delta_{9,10} = T_{р\,10} - T_{р\,9} - t_{9,10} = 53 - 44 - 9 = 0,$$
$$\Delta_{10,11} = T_{р\,11} - T_{р\,10} - t_{10,11} = 83 - 53 - 30 = 0,$$
$$\Delta_{1,2} = T_{р\,2} - T_{р\,1} - t_{1,2} = 3 - 3 - 0 = 0,$$
$$\Delta_{4,8} = T_{р\,8} - T_{р\,4} - t_{4,8} = 44 - 25 - 0 = 19,$$
$$\Delta_{7,8} = T_{р\,8} - T_{р\,7} - t_{7,8} = 44 - 37 - 0 = 7,$$
$$\Delta_{6,7} = T_{р\,7} - T_{р\,6} - t_{6,7} = 37 - 37 - 0 = 0,$$
$$r_{0,1} = R_1 + \Delta_{0,1} = 0 + 0 = 0,$$
$$r_{2,3} = R_3 + \Delta_{2,3} = 0 + 0 = 0,$$
$$r_{0,2} = R_2 + \Delta_{0,2} = 0 + 0 = 0,$$
$$r_{3,4} = R_4 + \Delta_{3,4} = 23 + 0 = 23,$$
$$r_{3,8} = R_8 + \Delta_{3,8} = 4 + 19 = 23,$$
$$r_{3,5} = R_5 + \Delta_{3,5} = 0 + 0 = 0,$$
$$r_{6,8} = R_8 + \Delta_{6,8} = 4 + 0 = 4,$$
$$r_{5,6} = R_6 + \Delta_{5,6} = 0 + 0 = 0,$$
$$r_{8,11} = R_{11} + \Delta_{8,11} = 0 + 4 = 4,$$
$$r_{7,11} = R_{11} + \Delta_{7,11} = 0 + 28 = 28,$$
$$r_{0,7} = R_7 + \Delta_{0,7} = 0 + 20 = 20,$$
$$r_{7,9} = R_9 + \Delta_{7,9} = 0 + 0 = 0,$$
$$r_{9,10} = R_{10} + \Delta_{9,10} = 0 + 0 = 0,$$
$$r_{10,11} = R_{11} + \Delta_{10,11} = 0 + 0 = 0,$$
$$r_{1,2} = R_2 + \Delta_{1,2} = 0 + 0 = 0,$$
$$r_{4,6} = R_8 + \Delta_{4,8} = 4 + 19 = 23,$$
$$r_{7,8} = R_8 + \Delta_{7,8} = 4 + 7 = 11,$$
$$r_{6,7} = R_7 + \Delta_{6,7} = 0 + 0 = 0.$$

Результаты вычислений заносим в соответствующие столбцы таблицы работ (см. табл. X.10). Нулевые значения полных резервов времени определяют принадлежность работ к критическому пути модели.

Анализ показывает, что имеются два критических пути: (0, 2, 3, 5, 6, 7, 9, 10, 11) и (0, 1 2, 3, 5, 6, 7, 9, 10, 11). Следовательно, наиболее ответственными, с позиции времени выполнения программы ввода камеры в эксплуатацию, являются работы: по проведению ходового и отрезных восстающих, рудоспуска, орта скреперования, отрезного штрека № 2, 3-й пары дучек и орта подсечки над ними, а также бурение скважин на буровом орте. Любое увеличение длительности хотя бы одной из указанных работ или задержка ее выполнения немедленно отразится на общем времени реализации программы.

Оценим надежность реализации программы при условии, что директивный срок составляет 89 сут, а среднее квадратичное отклонение длительности любого из двух критических путей – 6 сут

$$P\left\{\bar{S} < T\right\} = \Phi\left\{(T - \bar{S})\,\sigma_S^{-1}\right\} = \Phi\left\{(89 - 83)\,6^{-1}\right\} = \Phi\left\{1,0\right\} = 0{,}84.$$

Таким образом, существует достаточно высокая надежность реализации программы за директивный срок.

§ 4. ПОСТРОЕНИЕ КАЛЕНДАРНОГО ПЛАНА РЕАЛИЗАЦИИ ПРОГРАММ

Календарный план или ленточно-сетевой график представляет собой ленточную диаграмму временных параметров работ, построенную в масштабе длительности реализации программы с указанием связи между работами. В его основе лежат сетевая модель и расчет временных параметров событий и работ.

Ленточно-сетевой график строится в соответствии со специальной формой (форма X.3), совмещающей в себе график и диаграмму потребления ресурса.

В верхней части формы слева фиксируются информация из таблицы работ и сведения о расходе ресурса на каждой операции, справа выбирается масштаб и строится сетка времени, охватывающая отрезок $[0, S]$. На сетке вертикальными линиями обозначают ранние сроки наступления событий и под ними проставляют их номера. Затем на сетке времени в соответствующих строках строят работы программы.

Для построения работы (i, j) на отрезке между вертикальными отметками начального i и конечного j событий в масштабе откладывают длительность задержки $g_{i,j}$ работы и время ее выполнения $t_{i,j}$. Оставшаяся часть отрезка соответствует свободному резерву времени. При построении обозначают: задержку $g_{i,j}$ – символами "ооо"; время выполнения работы $t_{i,j}$ – сплошной чертой (для критических работ утолщенной линией); свободный резерв $\Delta_{i,j}$ – символами "+++".

Фиктивные работы на графике обозначаются жирной точкой.

По законченному графику строится диаграмма расхода ресурса в следующей последовательности.

На сетке диаграммы вертикальными линиями обозначаются все моменты начала или завершения работ. Эти отметки определяют границы интервалов, в течение которых расход ресурса не меняется. Если обозначить через $P_{i,j}$ количество однородного ресурса, требуемого для выполнения операции (i, j) за время $t_{i,j}$, то его суммарный расход на одновременно выполняемые работы в k-м интервале определяется по формуле

$$M_k = \sum_{(i,j) \in V_k} P_{i,j} \qquad \text{(X.15)}$$

и записывается в соответствующую строку таблицы.

Множество одновременно выполняемых работ V_k определяется по ленточно-сетевому графику как совокупность работ, попадающих в границы интервала.

По ряду значений M_k определяется размах суммарного расхода ресурса

$$\Delta M = \max_k \{M_k\} - \min_k \{M_k\} = M_{\max} - M_{\min} \qquad \text{(X.16)}$$

и масштаб диаграммы.

Форма X.3

Наименование работы	Номер работы	Индекс начального события i	Индекс конечного события j	Длительность работы $t_{i,j}$	Расход ресурса $P_{i,j}$	Сроки выполнения работ от начала реализации программы в течение времени S: 1	2	3	…	$S-2$	$S-1$	S
Номера наступающих событий												
Границы интервалов с постоянным числом работ												
Суммарный расход ресурса					$M_{\max}$ … $M_{\min}$							
Суммарный расход ресурса в интервале времени												

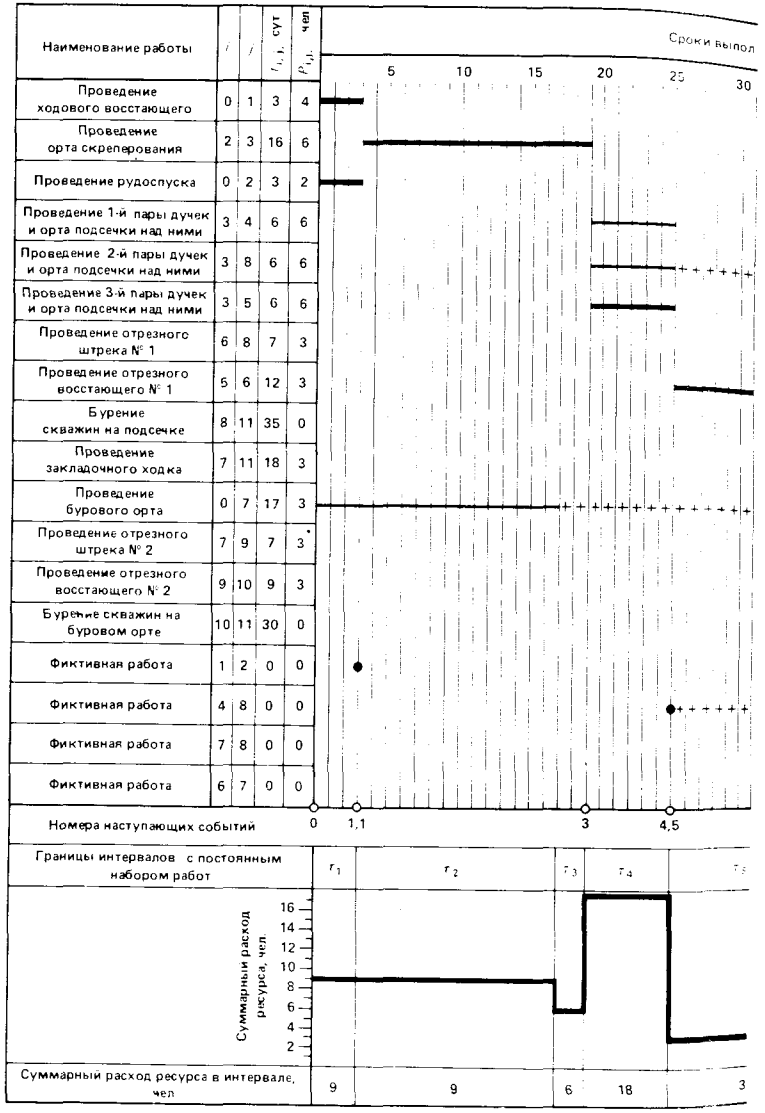
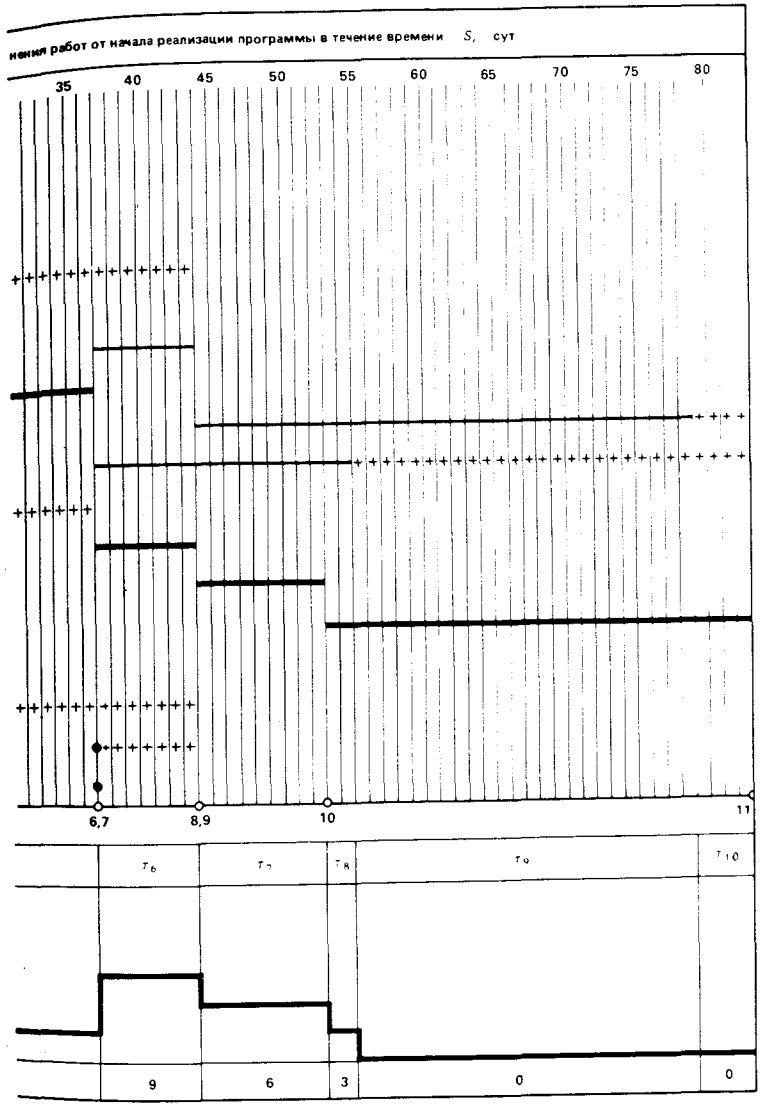


Рис. X.15. Исходный календарный план программы ввода камеры в эксплуатацию

Значения $(M_k - M_{\min})$ откладываются в принятом масштабе от уровня $M_{\min}$ в границах соответствующих интервалов.

Первоначально строится календарный план при условии, что ни одна из работ не имеет задержек $(g_{i,\ j} = 0; \quad (i, j) \in Q)$. Полученный план,



который называют исходным, подвергают анализу с точки зрения его допустимости.

Во-первых, исходный план может быть недопустим по времени его реализации, т.е. $S < T$, где T – установленный срок выполнения программы.

Во-вторых, для подавляющего большинства производственных программ характерно условие ограниченного количества ресурса, выделенного для их реализации. Наличие в исходном плане параллельных (одновременно выполняемых) операций может привести к ситуации, когда требуемое количество ресурса больше, чем имеющееся в наличии.

Кроме того, если исходный план допустим, он может быть не оптимален с точки зрения суммарных затрат на его реализацию.

Далее будут рассмотрены две основные задачи оптимизации, применяемые при построении допустимых календарных планов и их улучшении:

оптимизация плана по времени при ограниченных ресурсах;

оптимизация плана по затратам.

Пример X.9. Построить исходный календарный план программы ввода камеры на руднике в эксплуатацию по исходным данным, записанным в табл. X.10, X.11, в соответствии с сетевой моделью, изображенной на рис. X.14 (см. пример X.8).

При построении календарного плана учтем, что для выполнения работ по проведению выработок требуется следующее число проходчиков: 2 – для рудоспуска; 4 – для ходового восстающего; 6 – для орта скреперования, любой пары дучек и орта подсечки под ними; 3 – для всех остальных проходческих работ.

В результате построения получен исходный календарный план (рис. X.15), для реализации которого за критическое время ($S = 83$ сут) требуется 18 проходчиков. Причем их максимальное число необходимо для параллельно выполняемых работ по проведению дучек и орта подсечки над ними в период с 19 по 25 сутки реализации программы. В остальное время потребность в проходчиках не превышает 9.

§ 5. ОПТИМИЗАЦИЯ КАЛЕНДАРНОГО ПЛАНА ПО ВРЕМЕНИ ПРИ ОГРАНИЧЕННЫХ РЕСУРСАХ

Задача оптимизации календарного плана по времени при ограниченном суммарном расходе однородного ресурса формулируется следующим образом: определить для каждой операции сроки выполнения и расход ресурса, обеспечивающие минимальное время реализации программы при условии, что суммарное потребление его в каждый текущий момент времени не должно превышать заданный уровень. Последовательность работ, возможные расходы ресурса и их связь с длительностями операций известны.

Такая задача возникает, например, при организации работы бригады, выполняющей один подряд. В этом случае расход ресурса представляет собой число рабочих, выполняющих работу, а суммарный расход – общую численность бригады, которая не может быть превышена при расстановке ее членов на параллельные операции. Допустимая численность рабочих на одной операции определяется организационными и технологическими условиями, а также требованиями техники безопасности.

Другим примером может служить составление графика подачи бетона на сложный строительный объект с собственного растворного завода. Здесь расход ресурса – это поставки бетона в единицу времени на технологически связанные участки бетонирования, а суммарный расход – производственная мощность завода.

В общем виде задача формулируется следующим образом:

найти $g_{i,j} \geqslant 0$, $t_{i,j} \geqslant 0$, для всех $(i, j) \in Q$, обеспечивающие

$$S = \sum_{(i,j) \in L_{кр}} (g_{i,j} + t_{i,j}) \to \min \qquad (X.17)$$

при ограничениях:

$$\sum_{(i,j) \in V_k} P_{i,j} \leqslant P, \quad k = 1, 2, \ldots, r; \qquad (X.18)$$

$$a_{i,j} \leqslant P_{i,j} \leqslant b_{i,j}, \quad (i,j) \in Q, \qquad (X.19)$$

$$\text{если } P_{i,j} = f_{i,j}(t_{i,j}), \qquad (X.20)$$

где Q – множество работ (i, j) сетевой модели программы; $L_{кр}$ – множество работ (i, j) критического пути сетевой модели; $t_{i,j}$ – длительность работы (i, j), $g_{i,j}$ – задержка начала выполнения работы (i, j); P – допустимый суммарный расход ресурса; V_k – множество работ, одновременно выполняемых на k-м интервале времени, $k = 1, 2, \ldots, r$; r – число интервалов, на которые делится время выполнения программы из условия, что в течение каждого интервала состав одновременно выполняемых работ постоянен; $a_{i,j}$, $b_{i,j}$ – допустимые уровни расход ресурса на операции (i, j).

Постановка задачи предполагает, что каждая работа программы (i, j) имеет длительность выполнения $t_{i,j}$ и задержку ее $g_{i,j}$ начала от момента наступления начального события i до момента $H_{i,j}$ начала операции. Величины $t_{i,j}$ и $g_{i,j}$ выполняют роль оптимизируемых переменных. Срок начала работы (i, j) определяется по формуле

$$H_{i,j} = T_{р\,i} + g_{i,j}, \qquad (X.21)$$

где $T_{р\,i}$ – ранний срок наступления i-го события.

Введение задержек изменяет сроки окончания операций, что должно учитываться при расчете временных параметров.

Основное ограничение – допустимый суммарный расход ресурсов в каждый текущий момент времени. При постоянном расходе ресурса в течение всей операции суммарный расход может измениться только в момент изменения перечня одновременно выполняемых работ. Поэтому интегральная оценка текущего расхода ресурса по всем работам заменена системой ограничений (X.18), в которой каждое неравенство соответствует временному интервалу с постоянным числом одновременно выполняемых операций. Изменение длительности или введение задержки меняет сроки выполнения работ, а следовательно, и интервалы с постоянным количеством операций. Это создает специфическую осо-

бенность данной задачи – непостоянное число ограничений в процессе оптимизации, что не позволяет применить известные методы математического программирования.

Возможны две модификации постановки задачи:

с фиксированными расходами ресурсов и длительностями работ [ограничения (X.19) и (X.20) отсутствуют];

с расходами ресурсов, зависящими от длительности операции, но неизменными в процессе ее выполнения [ограничения (X.19) и (X.20) сохраняют свой вид].

К первой постановке сводятся задачи организации работ, выполняемых одним комплектом оборудования или звеньями рабочих, численность которых регламентирована технологией или правилами техники безопасности.

Вторая постановка включает задачи, например, распределения оборудования по стадиям производственного цикла.

Процесс решения состоит из следующих этапов:

построение исходного графика работ;

получение допустимого или опорного плана;

улучшение опорного плана.

Здесь под опорным планом подразумевается календарный план с суммарным расходом ресурса, не превышающим установленного уровня в течение времени выполнения программы.

Вид постановки задачи определяет метод поиска опорного и оптимального решений и возможность формализации этого процесса.

Методы решения в основном отличаются используемыми способами регулирования суммарного расхода ресурсов.

Заметим, что в настоящее время отсутствуют строгие аналитические методы решения указанной задачи в общем виде. Причем введение нескольких видов ресурсов значительно усложняют процедуру поиска решения. Поэтому решение практических задач, как правило, основано на применении приближенных графо-аналитических методов, реализуемых в диалоговом режиме на ЭВМ.

Рассмотрим простейшую постановку задачи с фиксированными длительностями работ и расходами ресурса. В этом случае единственное средство решения – изменение сроков начала работ, на выполнение которых не хватает ресурсов.

В основе алгоритма поиска решения лежит итерационный процесс построения календарного плана, его оценки по диаграмме расхода ресурса и корректировки путем изменения параметров отдельных работ. Этот графо-аналитический метод может быть использован для решения небольших по размерности практических задач и хорошо иллюстрирует пути оптимизации календарного плана.

Процесс оптимизации при фиксированных расходах ресурса на отдельных операциях сводится к построению исходного и получению опорного плана, требующего минимального увеличения критического пути.

Исходный план строится по заданным длительностям работ и расходам ресурса. Далее на его основе строится оптимальный план, блок-схема алгоритма которого показана на рис. X.16. В его основе лежит циклический процесс проверки опорного плана по диаграмме расхода ресурса и переноса на более поздние сроки работ, на выполнение которых не хватает ресурсов в установленный период.

Для проверки опорности плана на диаграмме расхода ресурсов откладывается допустимый уровень P. Если превышений допустимого уровня нет, то план опорный и задача решена. В противном случае план перестраивается.

Изменение сроков осуществляется за счет ввдения задержек начала работ – $g_{i,j}$ при условии минимального увеличения критического пути.

Каждая итерация построения оптимального плана начинается с проверки допустимости графика работ по суммарному расходу ресурса (см. рис. X.16, блок 1).

Если исходный план не допустим, то на первой итерации проверяется совместность условий по расходу ресурса на отдельных работах и допустимому суммарному расходу (блок 2).

$$P_{i,j} \leqslant P \text{ для всех } (i,j) \in Q. \qquad \text{(X.22)}$$

В случае, когда условие (X.22) не выполняется, задача решения не имеет.

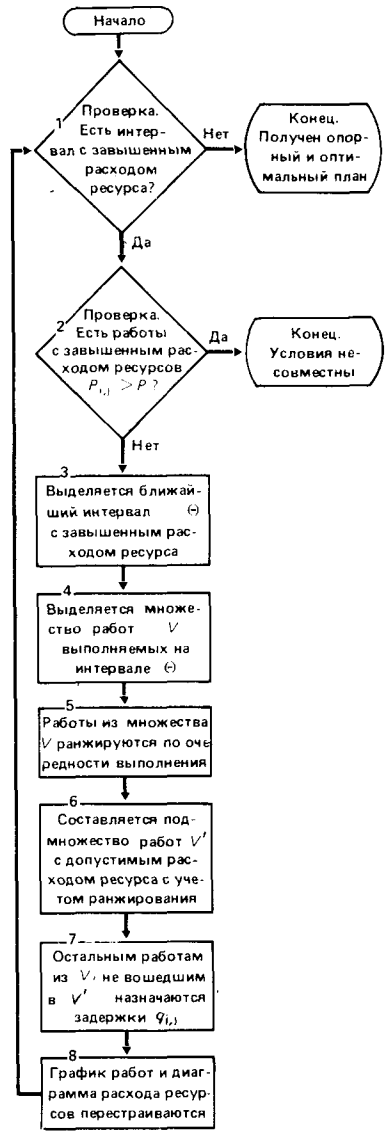


Рис. X.16. Блок-схема алгоритма получения оптимального по времени плана при фиксированном расходе однородного ресурса

Если условия совместны, но есть превышения суммарного расхода ресурса, то выбирается ближайший к началу реализации программы интервал θ с завышенным расходом и определяется множество работ V, выполняемых в этот период времени (блоки 3–4).

Эти работы ранжируются (блок 5) по следующим правилам.

1. Всем работам в соответствии с очередностью их выполнения присваиваются возрастающие номера, начиная с 1.

2. В первую очередь номера присваиваются работам, начатым в предшествующие периоды. Эти работы нумеруются в произвольной последовательности.

3. Остальные работы нумеруются по мере возрастания полного резерва. Работы с одинаковым полным резервом времени нумеруются в порядке убывания расхода ресурса.

Эти правила составлены из условий непрерывности работ, минимального роста критического пути и максимальной интенсивности использования ресурса.

После ранжирования формируется перечень работ, составляющих подмножество V', из условия

$$\sum_{(i,\,j)\,\in\,V'} P_{i,\,j} \leqslant P, \qquad \text{(X.23)}$$

причем работы включаются в перечень по мере роста ранговых номеров (блок 6). Если после включения в перечень работы (i, j) с номером k условие (X.23) нарушается, то данная операция исключается из перечня. Далее проверяется сохранение условия (X.23) при включении в список операции с номером $k + 1$ и так далее, вплоть до работы с последним номером. Работам, не включенным в перечень, назначаются задержки (блок 7)

$$g'_{i,\,j} = g_{i,\,j} + \tau, \qquad \text{(X.24)}$$

где $g_{i,\,j}$ – задержка работы (i, j) в данном плане; τ – длительность интервала θ, определяемая по формуле

$$\tau = \tau_{\text{к}} - \tau_{\text{н}}. \qquad \text{(X.25)}$$

Начальную $\tau_{\text{н}}$ и конечную $\tau_{\text{к}}$ границы интервала определяют по календарному плану и таблицам временных параметров работ и событий.

После введения задержек график работ перестраивают (блок 8) и процесс повторяется до получения опорного плана, являющегося и оптимальным.

Если в оптимальном плане критический путь ни по продолжительности, ни по составу работ не отличается от критического пути исходного плана, то значения надежности их реализации за директивный срок равны. В противном случае надежность выполнения оптимального плана оценивается заново с использованием соотношения (X.14). Следует помнить, что значение надежности изменится при появлении в оптимальном плане критического пути, равноценного по длительности, но обла-

Таблица X.12

Работы интервала в порядке возрастания ранговых номеров (i,j)	Требуемое число проходчиков для выполнения работы $P_{i,\,j}$	Требуемое число проходчиков в нарастающем итоге $\Sigma P_{i,\,j}$	Задержка времени начала работы $g_{i,\,j}$, сут
(3, 5)	6	6	0
(3, 8)	6	12	6
(3, 4)	6	18	6

дающего большей суммарной дисперсией длительностей составляющих его работ.

Пример X.10. Решить задачу оптимизации по времени исходного календарного плана программы ввода камеры в эксплуатацию (см. рис. X.15, пример 9) при условии, что для выполнения проходческих работ программы выделено 9 чел.

Анализ исходного календарного плана показывает, что он не является допустимым, так как на интервале τ_4 завышено потребление ресурса. Сравнение величин $P_{i,\,j}$ с допустимым уровнем расхода ресурса ($P = 9$ чел) показывает, что для всех работ программы выполняется условие (X.22), т.е. условия задачи совместны и следует перейти к корректировке плана.

В единственном интервале с завышенным расходом ресурса τ_4 в соответствии с исходным планом выполняются работы: (3, 4), (3, 8) и (3, 5). Полные резервы времени указанных работ составляют 23, 23 и 0 сут соответственно (см. пример X.8; табл. X.10).

Следовательно, операция (3, 5) будет иметь первый ранговый номер, а операции (3, 8) и (3, 4) – второй и третий. Причем, для двух последних работ ранговые номера можно поменять, так как они находятся в одинаковых условиях с точки зрения полных резервов времени и расхода ресурса.

Для определения перечня работ, которым будут назначены задержки, заполняется построчно табл. X.12. Величина задержки определяется длительностью интервала $\tau_4 = T_{\text{р}\,4} - T_{\text{р}\,3} = 25 - 19 = 6$ сут.

Затем корректируются временные параметры сетевой модели с учетом введенных задержек.

Так как $g_{3,\,8} < \Delta_{3,\,8}$, то по определению свободного резерва введения задержки начала работы (3, 8) на 6 сут приведет только к уменьшению на величину задержки временных параметров этой работы

$$\Delta_{3,\,8} = (T_{\text{р}\,8} - T_{\text{р}\,3} - t_{3,\,8} - g_{3,\,8}) = 19 - 6 = 13 \text{ сут},$$

$$r_{3,\,8} = (R_8 + \Delta_{3,\,8}) = 4 + 13 = 17 \text{ сут}.$$

Введение задержки $g_{3,\,4} > \Delta_{3,\,4}$ приведет к изменению раннего срока свершения события 4 и полного резерва времени работы (3, 4)

$$T_{\text{р}\,4} = (T_{\text{р}\,3} + t_{3,\,4}) + g_{3,\,4} = 25 + 6 = 31 \text{ сут},$$

$$r_{3,\,4} = (T_{\text{п}\,4} - T_{\text{р}\,3} - t_{3,\,4}) - g_{3,\,4} = 23 - 6 = 17 \text{ сут}.$$

Свободный резерв времени работы (3, 4) уже в исходном плане был нулевым.

Изменение значения $T_{\text{р}\,4}$ на величину $g_{3,\,4}$ означает сдвиг вправо по временной оси работы (4, 8) на 6 сут. Так как указанный сдвиг лежит в пределах $\Delta_{4,\,8}$, то

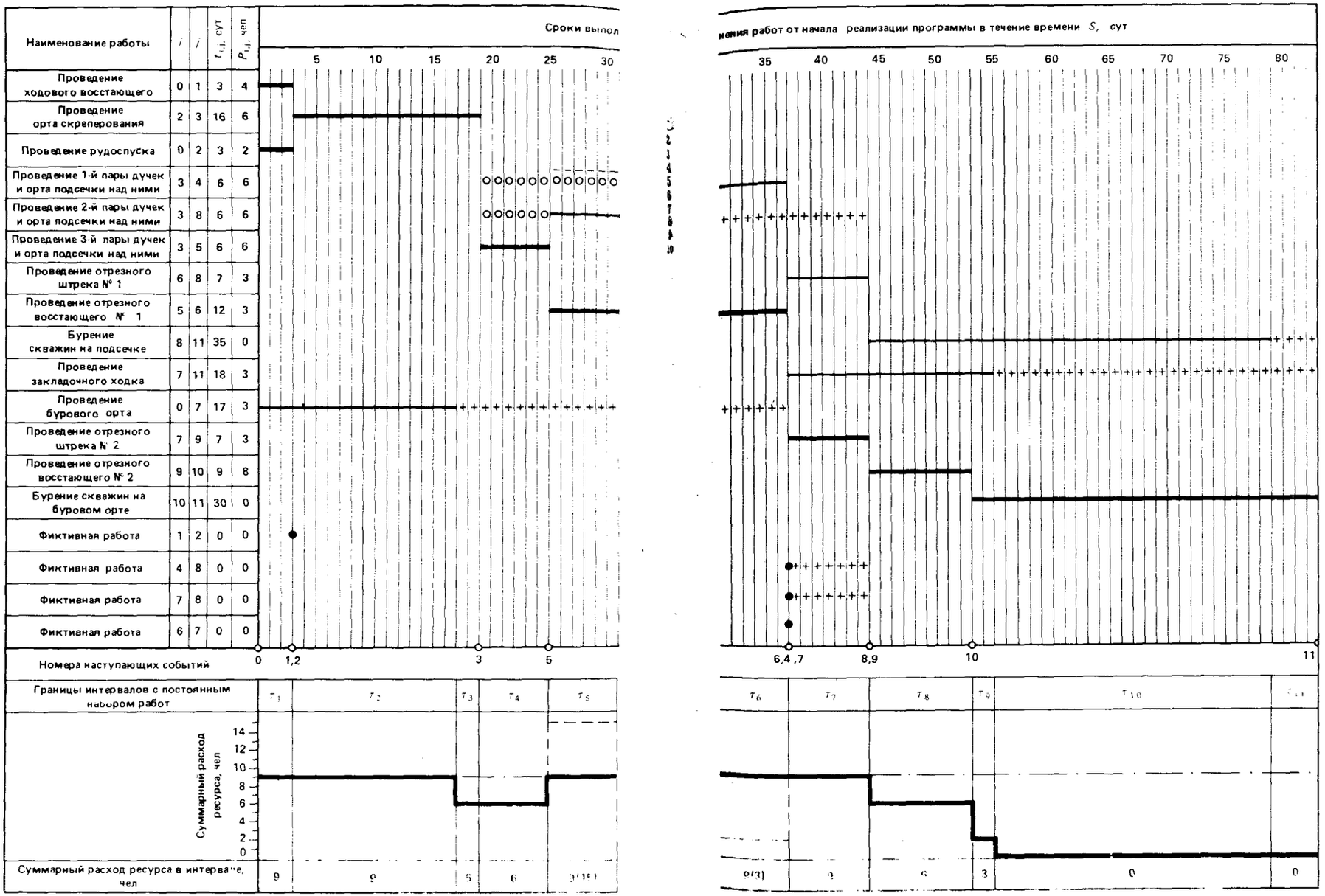


Рис. X.17. Оптимальный по времени календарный план программы ввода камеры в эксплуатацию при ограниченном числе проходчиков (пунктиром показано расположение соответствующих работ и расход ресурса после первой корректировки исходного плана)

Таблица X.13

Работы интервала в порядке возрастания ранговых номеров (i, j)	Требуемое число проходчиков для выполнения работы $P_{i,j}$	Требуемое число проходчиков в нарастающем итоге $\Sigma P_{i,j}$	Задержка времени начала работы $g_{i,j}$, сут
(5, 6)	3	3	0 + 0 = 0
(3, 8)	6	9	6 + 0 = 6
(3, 4)	6	15	6 + 6 = 12

изменяются только временные параметры работы (4, 8)

$$\Delta_{4,8} = T_{p8} - T_{p4} - t_{4,8} = 44 - 31 - 0 = 13 \text{ сут},$$

$$r_{4,8} = R_8 + \Delta_{4,8} = 4 + 13 = 17 \text{ сут}.$$

Временные параметры остальных работ и событий останутся без изменений.

С учетом введенных задержек и изменений значений временных параметров перестраивается календарный план (рис. X.17).

Анализ календарного плана, полученного после первой корректировки, показывает, что он не является допустимым, так как в интервале времени τ_5 завышено потребление ресурса. На указанном интервале выполняются работы (3, 4), (3, 8) и (5, 6) с полными резервами времени 17, 17 и 0 сут соответственно. Причем, как уже отмечалось выше, операции (3, 4) и (3, 8) равноценны по величине расхода ресурса. Поэтому теоретически ни одна из этих двух работ не имеет приоритета с точки зрения присвоения более высокого рангового номера. Однако с технологических позиций, учитывая, что на предыдущем временном интервале τ_4 выполняется работа по проведению 3-й пары дучек и орта подсечки над ними, более высокий ранговый номер правильнее присвоить работе по проведению 2-й пары дучек и орта подсечки над ними.

Для определения работ, которым будут назначены задержки, составляется табл. X.13. (интервал $\tau_5 = 6$ сут).

Увеличение задержки операции (3, 4) на 6 сут приведет к изменению значений временных параметров

$$T_{p4} = (T_{p3} + t_{3,4}) - g_{3,4} = 25 + 12 = 37 \text{ сут}.$$

$$r_{3,4} = (T_{п4} - T_{p3} - t_{3,4}) - g_{3,4} = 23 - 12 = 11 \text{ сут},$$

$$\Delta_{4,8} = T_{p8} - T_{p4} - t_{3,4} = 44 - 37 - 0 = 7 \text{ сут},$$

$$r_{4,8} = R_8 + \Delta_{4,8} = 4 + 7 = 11 \text{ сут}.$$

С учетом задержки работы (3, 4) на 12 сут и скорректированных значений временных параметров перестраивается календарный план (см. рис. X.17). В результате получается допустимый план, являющийся и оптимальным, предусматривающий реализацию программы ввода камеры в эксплуатацию за время $S = 83$ сут (продолжительность критического пути, как и состав его работ, не изменился). Надежность реализации программ за директивный срок ($T = 89$ сут) равна 0,84 (см. пример X.8).

Полученный оптимальный календарный план предусматривает выполнение проходческих работ двумя звеньями из 6 и 3 чел соответственно.

Первое звено выполняет критические работы: проведение ходового восстающего и рудоспуска, орта скреперования, 3-й пары дучек с ортом подсечки над ними. Тем временем второе звено проводит буровой орт, затем отрезной восстающий

№ 1, тогда как первое кончает проведение дучек и орта подсечки. Затем образуется три звена по 3 чел. Одно завершает проведение отрезного штрека № 1, второе закладочного ходка, третье – отрезных восстающего и штрека № 2.

§ 6. ОПТИМИЗАЦИЯ КАЛЕНДАРНОГО ПЛАНА ПО ЗАТРАТАМ

Задача оптимизации календарного плана по затратам формулируется следующим образом: определить длительности выполнения операций, обеспечивающие реализацию программы в установленный срок при минимальных суммарных затратах, если известны функции изменения затрат на выполнение работ в заданных пределах их длительностей.

В основе рассматриваемой постановки задачи лежит предположение о том, что затраты на отдельную операцию непостоянны и зависят от длительности ее выполнения. Следовательно, меняя длительности работ, можно получить план, обеспечивающий выполнение программы в установленный срок с минимальными затратами.

Кривая зависимости затрат на выполнения операции от ее длительности показана на рис. X.18.

Такой характер кривой объясняется тем, что затраты на выполнение любой производственной операции складываются из заработной платы повременно оплачиваемого персонала, амортизационных отчислений и других цеховых расходов, растущих прямо пропорционально времени работ, а также из заработной платы сдельно оплачиваемых рабочих, стоимости материалов и других затрат, зависящих только от объема работ. Эта вторая группа затрат может возрастать при ускоренном выполнении работ за счет привлечения избыточного числа работников, использования дополнительного оборудования и т.д. В результате для каждой операции (i, j) программы существует оптимальный срок выполнения $D_{i,j}$, при котором затраты на ее выполнение минимальны и равны $C_{i,j\,\min}$.

Суммарные расходы на программу в целом складываются из затрат на отдельные операции, поэтому при отсутствии ограничений на длительности операций и срок выполнения программы оптимальным будет план с $t_{ij} = D_{i,j}$ для всех $(i, j) \in Q$, где Q – множество работ сетевой модели.

В дальнейшем такой план будет называться "нормальным", так

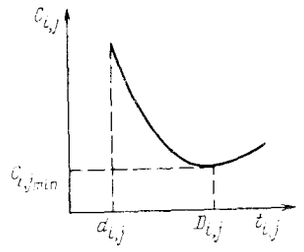


Рис. X.18. Зависимость затрат $C_{i,j}$ на выполнение операции (i, j) от ее длительности $t_{i,j}$

как он соответствует обычным срокам выполнения операций. Если же сократить длительность выполнения операций до значения $t_{i,j} = d_{i,j}$, то это приведет к сокращению срока реализации программы, но увеличит затраты. Такой план будет называться "срочным".

После введения ограничений на длительность работ и сроки выполнения программы, а также линеаризации функций затрат зависимостями $C_{i,j} = C_{i,j\min} + \delta_{i,j}(D_{i,j} - t_{i,j})$ задача примет вид:

$$Z = \sum_{(i,j)\in Q} \delta_{i,j}(D_{i,j} - t_{i,j}) \rightarrow \min, \qquad (X.26)$$

$$d_{i,j} \leqslant t_{i,j} \leqslant D_{i,j}, \text{ для всех } (i,j) \in Q \qquad (X.27)$$

$$\sum_{(i,j)\in L_k} t_{i,j} \leqslant T, \quad k = 1, 2, \ldots, r, \qquad (X.28)$$

где Z – дополнительные затраты, связанные с ускорением сроков реализации программы по сравнению с "нормальным" планом"; Q – множество работ сетевой модели; $\delta_{i,j}$ – удельные затраты на сокращение длительности выполнения операции (i, j) на одну единицу; $d_{i,j}$ и $D_{i,j}$ – нижний и верхний пределы изменения длительности работы (i, j), соответственно, выбранные по организационным и технологическим соображениям таким образом, что при $t_{i,j} = D_{i,j}$ затраты на выполнение операции минимальны; L_k – множество работ k-го полного пути сетевой модели; r – число полных путей модели; T – установленный срок реализации программы.

Данная постановка представляет собой задачу линейного программирования, но возможности ее решения симплекс-методом ограничены из-за большой размерности симплекс-таблицы, включающей двухсторонние ограничения длительностей работ. Для решения практических линейных задач большой размерности, а также при определенных нелинейных функциях изменения затрат используются специальные методы, например, алгоритм Фулкерсона.

Эффективность использования симплекс-метода при решении указанной задачи можно повысить, если сократить размерность системы ограничений и число итераций поиска опорного и оптимального решений. Этого можно достичь, если в качестве оптимизируемой переменной использовать $\Delta t_{i,j}$ – приращение длительности операции (i, j) при переходе от "срочного" к оптимальному плану реализации программы.

Пусть имеется "срочный" план с критическим путем $S_{\text{кр}} \leqslant T$. В случае $S_{\text{кр}} > T$ задача неразрешима, так как "срочный" план обеспечивает минимально возможные сроки выполнения программы.

Оптимизируемая переменная Δt_{ij} представляет разность между длительностью работы $t_{i,j}$ и минимально допустимым сроком ее выполнения $d_{i,j}$

$$\Delta t_{i,j} = t_{i,j} - d_{i,j}. \qquad (X.29)$$

В исходном "срочном" плане $\Delta t_{i,j} = 0$ для всех $(i, j) \in Q$. Увеличение $\Delta t_{i,j}$ снижает затраты от уровня "срочного" плана, поэтому в качестве критерия оптимальности можно взять максимум суммарной экономии

$$\Delta C = \sum_{(i,j)\in Q} \delta_{i,j}\,\Delta t_{i,j} \rightarrow \max. \qquad (X.30)$$

Из (X.29) следует, что

$$t_{i,j} = d_{i,j} + \Delta t_{i,j}. \qquad (X.31)$$

После подстановки (X.31) в исходные ограничения (X.27) условия допустимой длительности работ примут вид

$$\begin{cases} \Delta t_{i,j} \geqslant 0, \\ -\Delta t_{i,j} + (D_{i,j} - d_{i,j}) \geqslant 0, \\ \text{для всех } (i,j) \in Q. \end{cases} \qquad (X.32)$$

Аналогично выводится система ограничений по директивному сроку реализации программы

$$-\sum_{(i,j)\in L_{\text{кр}}} \Delta t_{i,j} + \left(T - \sum_{(i,j)\in L_{\text{кр}}} d_{i,j}\right) \geqslant 0, \quad k = 1, 2, \ldots, r. \qquad (X.33)$$

Необходимо отметить, что в модель целесообразно включать только переменные $\Delta t_{i,j}$, соответствующие работам с переменной длительностью.

Остановимся подробнее на методике формирования систем ограничений (X.33), требующих соблюдения установленного срока по всем полным путям сетевой модели.

Основа для построения системы ограничений – множество полных путей сети и их длительность в "срочном" плане. Эта информация представляется в виде специальной формы (форма X.4).

В первой строке перечислены все действительные работы сетевой модели, во второй – их минимально допустимые длительности. В следующих строках нулями и единицами указаны принадлежность работ полным путям, а в последней – интервалы изменения длительностей работ.

После вычисления допустимых приращений полных путей из формы вычеркиваются столбцы, соответствующие операциям с неизменной длительностью (нулевым интервалом изменения длительности). Оставшаяся часть со столбцом допустимых приращений полных путей составляет фрагмент симплекс-таблицы, соответствующий системе ограничений (X.33).

Остальные ограничения в симплекс-таблице формируются из единичной матрицы коэффициентов и соответствующих интервалов изменения длительностей работ, составляющих столбец свободных членов.

Исходное базисное решение, соответствующее "срочному" плану,

Форма X.4

Индексы работы (i, j)						Длительность k-го полного пути $\sum_{(i,j)\in L_k} d_{i,j}$	Допустимое приращение длительности k-го полного пути $T - \sum_{(i,j)\in L_k} d_{i,j}$
Минимальная длительность работы $d_{i,j}$							
Отметка о принадлежности работы к k-му полному пути	$k=1$						
	$k=2$						
	$k=r$						
Интервал изменения длительности работы $(D_{i,j} - d_{i,j})$							

в котором все $\Delta t_{i,j} = 0$, является допустимым, поэтому процесс сразу начинается с поиска оптимального решения по обычной процедуре симплекс-метода.

После получения решения задачи (X.30), (X.32)+ (X.33) определяются значения $t_{i,j}$ по соотношению (X.31), затем рассчитываются значения временных параметров сетевой модели и строится оптимальный календарный план.

Оценка надежности реализации оптимального плана осуществляется при помощи соотношения (X.14), в котором используются соответствующие параметры распределения оптимальных длительностей работ.

Пример X.11. Построить календарный план ввода камеры в эксплуатацию на руднике, обеспечивающий минимальные затраты на выполнение программы, представленной сетевой моделью (см. рис. X.14). Длительности операций (в сутках) и удельные затраты (в руб/сут) на ускорение темпов работ приведены в табл. X.14. Установленный срок реализации программы $T = 79$ сут.

Для расчета параметров линейной модели составляется табл. X.15.

Из табл. X.15 исключим столбцы, соответствующие работам (5, 6), (8, 11) и (10, 11), а также столбец значений длительностей полных путей срочного плана. Оставшаяся часть таблицы содержит коэффициенты при оптимизируемых переменных и свободные члены в системе ограничений по допустимым приращениям длительностей полных путей

$$\Delta t_{0,1} + \Delta t_{2,3} + \Delta t_{3,4} \leqslant 25,$$
$$\Delta t_{0,1} + \Delta t_{2,3} + \Delta t_{3,8} \leqslant 25,$$
$$\Delta t_{0,1} + \Delta t_{2,3} + \Delta t_{3,5} + \Delta t_{6,8} \leqslant 7,,$$
$$\Delta t_{0,1} + \Delta t_{2,3} + \Delta t_{3,5} \leqslant 13,$$

Таблица X.14

i	j	$d_{i,j}$	$D_{i,j}$	$\delta_{i,j}$
0	1	2	3	215
2	3	12	16	121
0	2	2	3	105
3	4	5	6	65
3	8	5	6	65
3	5	5	6	65
6	8	6	7	254
5	6	12	12	–
8	11	35	35	–
7	11	12	18	162
0	7	15	17	92
7	9	6	7	212
9	10	8	9	82
10	11	30	30	–

$$\Delta t_{0,1} + \Delta t_{2,3} + \Delta t_{3,5} + \Delta t_{1,11} \leqslant 36,$$
$$\Delta t_{0,1} + \Delta t_{2,3} + \Delta t_{3,5} + \Delta t_{7,9} + \Delta t_{9,10} \leqslant 4,$$
$$\Delta t_{0,2} + \Delta t_{2,3} + \Delta t_{3,4} \leqslant 25,$$
$$\Delta t_{0,2} + \Delta t_{2,3} + \Delta t_{3,8} \leqslant 25,$$
$$\Delta t_{0,2} + \Delta t_{2,3} + \Delta t_{3,5} + \Delta t_{6,8} \leqslant 7,$$
$$\Delta t_{0,2} + \Delta t_{2,3} + \Delta t_{3,5} \leqslant 13,$$
$$\Delta t_{0,2} + \Delta t_{2,3} + \Delta t_{3,5} + \Delta t_{7,11} \leqslant 36,$$
$$\Delta t_{0,2} + \Delta t_{2,3} + \Delta t_{3,5} + \Delta t_{7,9} + \Delta t_{9,10} \leqslant 4,$$
$$\Delta t_{0,7} \leqslant 29,$$
$$\Delta t_{0,7} + \Delta t_{7,11} \leqslant 52,$$
$$\Delta t_{0,7} + \Delta t_{7,9} + \Delta t_{9,10} \leqslant 20.$$

Сформируем систему ограничений по допустимым приращениям длительностей работ

$$\Delta t_{0,1} \leqslant 1, \quad \Delta t_{3,5} \leqslant 1,$$
$$\Delta t_{0,1} \leqslant 2, \quad \Delta t_{6,8} \leqslant 1,$$
$$\Delta t_{0,7} \leqslant 2, \quad \Delta t_{7,11} \leqslant 6$$
$$\Delta t_{2,3} \leqslant 4, \quad \Delta t_{7,9} \leqslant 1$$
$$\Delta t_{3,4} \leqslant 1, \quad \Delta t_{9,10} \leqslant 1$$
$$\Delta t_{3,8} \leqslant 1$$

Целевая функция будет иметь вид

$$\Delta C = 215\,\Delta t_{0,1} + 105\,\Delta t_{0,2} + 92\,\Delta t_{0,7} + 121\,\Delta t_{2,3} + 65\,\Delta t_{3,4} + 65\,\Delta t_{3,8} +$$
$$+ 65\,\Delta t_{3,5} + 254\,\Delta t_{6,8} + 162\,\Delta t_{7,11} + 212\,\Delta t_{7,9} + 82\,\Delta t_{9,10} \to \max.$$

(i, j)		0,1	0,2	0,7	2,3	3,4	3,8	3,5
$d_{i,j}$		2	2	15	12	5	5	5
Отметка о принадлежности работы k-му полному пути	$k=1$	1	0	0	1	1	0	0
	$k=2$	1	0	0	1	0	1	0
	$k=3$	1	0	0	1	0	0	1
	$k=4$	1	0	0	1	0	0	1
	$k=5$	1	0	0	1	0	0	1
	$k=6$	1	0	0	1	0	0	1
	$k=7$	0	1	0	1	1	0	0
	$k=8$	0	1	0	1	0	1	0
	$k=9$	0	1	0	1	0	0	1
	$k=10$	0	1	0	1	0	0	1
	$k=11$	0	1	0	1	0	0	1
	$k=12$	0	1	0	1	0	0	1
	$k=13$	0	0	1	0	0	0	0
	$k=14$	0	0	1	0	0	0	0
	$k=15$	0	0	1	0	0	0	0
$(D_{i,j} - d_{i,j})$		1	1	2	4	1	1	1

Анализ системы ограничений показывает, что ряд из них можно исключить. что приведет к упрощению линейной модели.

Например, ограничение по допустимому приращению 13-го полного пути ($\Delta t_{0,7} \leqslant 29$) менее "жесткое", чем по допустимому приращению длительности работы (0,7) ($\Delta t_{0,7} \leqslant 2$) и, следовательно, может быть исключено.

Решив задачу линейного программирования (процесс поиска решения опущен), находим оптимальные значения переменных и максимум линейной формы

$\Delta t^0_{0,1} = \Delta t^0_{0,2} = \Delta t^0_{3,4} = \Delta t^0_{3,8} = \Delta t^0_{6,8} = \Delta t^0_{7,9} = 1$ сут, $\Delta t^0_{2,3} = \Delta t^0_{0,7} = 2$ сут

$\Delta t^0_{3,5} = \Delta t^0_{9,10} = 0$; $\Delta t^0_{7,11} = 6$ сут, $\Delta C^0 = 2314$ руб.

На основании соотношения (X.31) вычисляем оптимальные значения длительностей работ (в сут)

$t^0_{0,1} = d_{0,1} + \Delta t^0_{0,1} = 2 + 1 = 3$

$t^0_{0,2} = d_{0,2} + \Delta t^0_{0,2} = 2 + 1 = 3$,

$t^0_{2,3} = d_{2,3} + \Delta t^0_{2,3} = 12 + 2 = 14$,

$t^0_{3,4} = d_{3,4} + \Delta t^0_{3,4} = 5 + 1 = 6$,

$t^0_{3,8} = d_{3,8} + \Delta t^0_{3,8} = 5 + 1 = 6$,

$t^0_{3,5} = d_{3,5} + \Delta t^0_{3,5} = 5 + 0 = 5$,

$t^0_{6,8} = d_{6,8} + \Delta t^0_{6,8} = 6 + 1 = 7$,

$t^0_{5,6} = d_{5,6} = D_{5,6} = 12$,

$t^0_{8,11} = d_{8,11} = D_{8,11} = 35$.

Таблица X.15

5,6	6,8	8,11	7,11	7,9	9,10	10,11	$\sum_{(i,j) \in L_k} d_{ij}$	$79 - \sum d_{i,j}$
12	6	35	12	6	8	30		
0	0	1	0	0	0	0	54	25
0	0	1	0	0	0	0	54	25
1	1	1	0	0	0	0	72	7
1	0	1	0	0	0	0	66	13
1	0	0	1	0	0	0	43	36
1	0	0	0	1	1	1	75	4
0	0	1	0	0	0	0	54	25
0	0	1	0	0	0	0	54	25
1	1	1	0	0	0	0	72	7
1	0	1	0	0	0	0	66	13
1	0	0	1	0	0	0	43	36
1	0	0	0	1	1	1	75	4
0	0	1	0	0	0	0	50	29
0	0	0	1	0	0	0	27	52
0	0	0	0	1	1	1	59	20
0	1	0	6	1	1	0	–	–

$t^0_{7,11} = d_{7,11} + \Delta t^0_{7,11} = 12 + 6 = 18$,

$t^0_{0,7} = d_{0,7} + \Delta t^0_{0,7} = 15 + 2 = 17$,

$t^0_{7,9} = d_{7,9} + \Delta t^0_{7,9} = 6 + 1 = 7$,

$t^0_{9,10} = d_{9,10} + \Delta t^0_{9,10} = 8 + 0 = 8$,

$t^0_{10,11} = d_{10,11} = D_{10,11} = 30$.

Таблица X.16

i	T_{pi}	T_{pi}	R_i
0	0	0	0
1	3	3	0
2	3	3	0
3	17	17	0
4	23	44	21
5	22	22	0
6	34	34	0
7	34	34	0
8	41	44	3
9	41	41	0
10	49	49	0
11	79	79	0

Таблица X.17

i	j	$r_{i,j}$	$\Delta_{i,j}$
0	1	0	0
0	2	0	0
2	3	0	0
3	4	21	0
3	8	21	18
3	5	0	0
6	8	3	0
5	6	0	0
8	11	3	3
7	11	27	27
0	7	17	17
7	9	0	0
9	10	0	0
10	11	0	0
1	2	0	0
4	8	21	18
7	8	10	7
6	7	0	0

Наименование работы	i	j	t_{ij}, сут
Проведение ходового восстающего	0	1	3
Проведение орта скреперования	2	3	14
Проведение рудоспуска	0	2	3
Проведение 1-й пары дучек и орта подсечки над ними	3	4	6
Проведение 2-й пары дучек и орта подсечки над ними	3	8	6
Проведение 3-й пары дучек и орта подсечки над ними	3	5	5
Проведение отрезного штрека № 1	6	8	7
Проведение отрезного восстающего № 1	5	6	12
Бурение скважин на подсечке	8	11	35
Проведение закладочного ходка	7	11	18
Проведение бурового орта	0	7	17
Проведение отрезного штрека № 2	7	9	7
Проведение отрезного восстающего № 2	9	10	8
Бурение скважин на буровом орте	10	11	30
Фиктивная работа	1	2	0
Фиктивная работа	4	8	0
Фиктивная работа	7	8	0
Фиктивная работа	6	7	0
Номера наступающих событий			

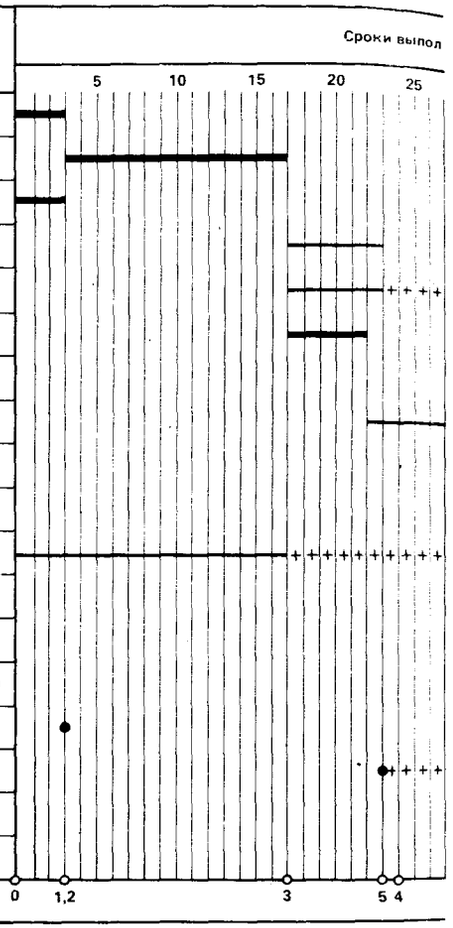
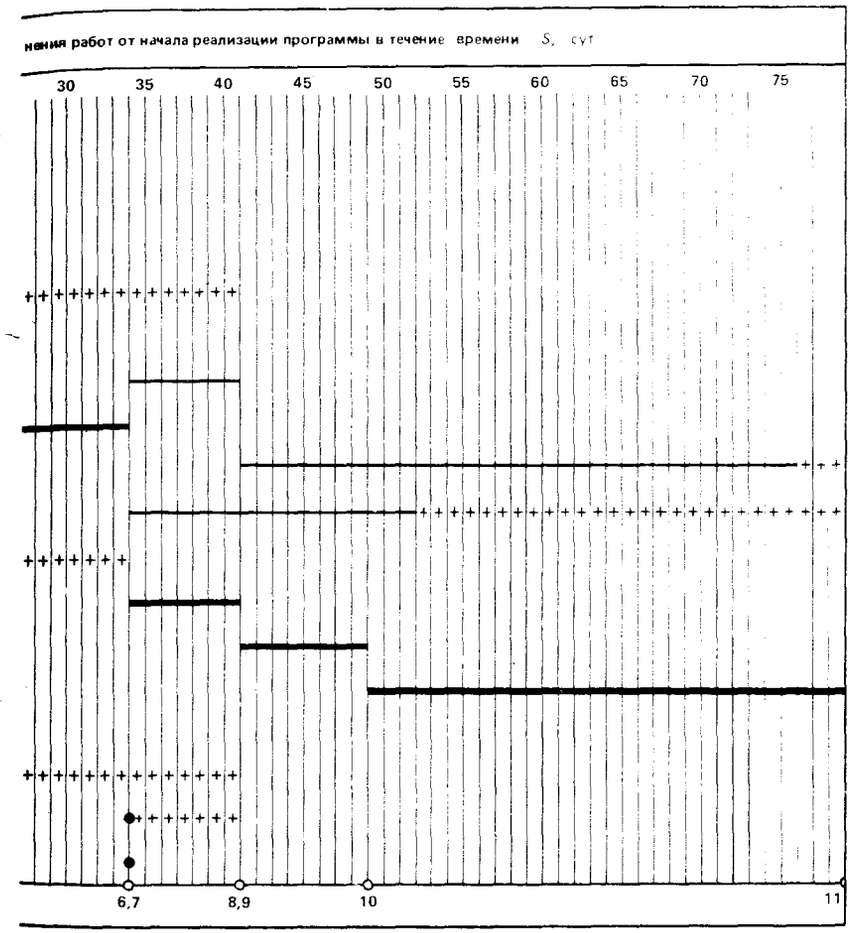


Рис. X.19. Оптимальный по затратам календарный план реализации программы ввода камеры в эксплуатацию при директивном сроке 79 сут

Затем определяем временные параметры сетевой модели при оптимальной длительности работ (табл. X.16 – X.17) и строим оптимальный календарный план (рис. X.19). Затраты на реализацию оптимального плана на 2314 руб. меньше, чем на выполнение срочного.

Заканчивая обсуждение проблем оптимизации планов, построенных сетевыми методами, заметим, что задача оптимизации по времени (при ограниченных ресурсах) наиболее характерна для программ производственной деятельности горных предприятий, а задача оптимизации по затратам – для проектных программ. Указанные задачи имеют большое самостоятельное значение и практическую ценность.



Вместе с тем, ряд задач календарного планирования производственных программ может быть сформулирован следующим образом: определить сроки выполнения каждой работы, обеспечивающие реализацию программы в установленный срок с минимальными затратами, при ограниченном суммарном расходе однородного ресурса.

В этом случае поиск, как правило, осуществляется путем последовательного решения рассмотренных выше задач оптимизации плана по времени и затратам.

Первоначально определяются оптимальные длительности работ, обеспечивающие минимальные затраты на реализацию программы в ус-

тановленный срок. Затем вычисляется необходимый расход ресурса для каждой операции, соответствующий ее оптимальной длительности, и оптимизируется календарный план по времени при фиксированных расходах ресурса.

Если время реализации полученного оптимального плана не превышает установленный срок, то задача решена.

В противном случае можно рекомендовать попытаться, ухудшая календарный план с точки зрения суммарных затрат, найти решение, удовлетворяющее ограничениям по установленному времени реализации и суммарному расходу ресурса. Однако в общем случае эта задача строго аналитического метода решения не имеет, а следовательно, приходится пользоваться эвристическими методами.

При планировании проектных программ возможна следующая постановка: определить сроки выполнения каждой операции, обеспечивающие реализацию программы в установленное время при минимальных затратах и максимальной равномерности потребления ресурса.

В этом случае поиск также, как правило, осуществляется путем последовательного решения двух оптимизационных задач по той же схеме, причем вторая задача (способом решения которой является рассредоточение работ с фиксированными расходами ресурса во времени) имеет целевую функцию, минимизирующую сумму квадратов отклонений требуемого количества ресурса от его среднего уровня.

§ 7. УПРАВЛЕНИЕ ПРОЦЕССОМ РЕАЛИЗАЦИИ ПРОГРАММ

Случаи, когда удается построить план, точно соблюдаемый при реализации программы, встречаются крайне редко. Обычно некоторые операции выполняются быстрее или медленнее, чем предусмотрено планом. Указанные отклонения могут поставить под угрозу выполнение программы в установленный срок и вызвать необходимость управления процессом реализации программы. Управление процессом реализации программы можно представить как задачу периодической выработки решений, определяющих, какие работы и с какой интенсивностью следует выполнять в каждый период в будущем с целью получения максимального эффекта (с позиций выбранного критерия оптимальности). Процесс выработки решений называется периодически, если время, через которое вырабатывается следующее решение, постоянно или кратно определенному периоду. Причинами, вызывающими выработку новых управляющих воздействий, являются отклонения фактических значений управляемых факторов (сроков начала и длительностей выполнения операций), от плановых или расчетных. Таким образом, в основу управления процессом реализации программы положена задача регулирования по отклонениям.

Заметим, что в принципе в связи с изменением условий возможны случаи, когда в процессе реализации программы могут произойти и

структурные изменения: отпадет необходимость выполнения некоторых операций программы и (или) появятся новые операции. Очевидно, что в этом случае также необходимо вмешательство в процесс реализации.

Построенный календарный план выполнения программы позволит на момент начала ее реализации определить задания ответственным исполнителям операций по интенсивности выполнения работ, срокам их начала и распределить между ними ресурсы. Далее на основе календарного плана осуществляется периодический контроль за соблюдением сроков начала и окончания работ. Выбор периода контроля зависит от особенностей объекта, в условиях которого реализуется программа. В процессе контроля особое внимание уделяется операциям критического пути.

Ход выполнения программы отображается на календарном плане. Влияние задержки какой-либо операции на остальную часть программы легче оценивается по сетевой модели.

Отклонения от календарного плана, возникающие в процессе реализации программы, можно условно разделить на два типа: первый – не приводящие к изменению длительности критического пути; второй – влекущие за собой увеличение времени выполнения программы.

Второй тип характерен для задержек выполнения критических операций или задержек некритических работ, превышающих их полные резервы времени.

При наличии отклонения первого типа календарный план корректируется и, если это необходимо, изменяются моменты начала выполнения отдельных операций.

При наличии отклонения второго типа возникает задача построения нового календарного плана. В этом случае прежде всего вносятся изменения в сетевую модель программы – законченным операциям ставятся в соответствие нулевые длительности; частично выполненным операциям приписываются оценки продолжительности, соответствующие их незавершенному объему. В сеть, при необходимости, вносят и структурные изменения. Затем осуществляется расчет временных параметров, строится и оптимизируется новый календарный план. На его основе ответственным исполнителем операций определяются интенсивность их ведения и моменты начала, а также распределяются ресурсы.

Наиболее часто случаи построения нового календарного плана встречаются в начальный период реализации программы, а затем, как правило, их количество уменьшается в связи с переходом процесса реализации программы в установившийся режим.

Контрольные вопросы и задачи

1. Что такое программа работ?
2. Приведите примеры программ горного производства, при составлении и реализации которых могут быть эффективно использованы сетевые методы.

3. Дайте определение понятию сетевая модель.

4. Что понимается в сетевом моделировании под работой (операцией) и событием?

5. Какие основные этапы включает в себя сетевое планирование и управление реализацией программ? Каково содержание этих этапов?

6. Какая информация необходима для построения сетевой модели программы работ?

7. Какова принципиальная последовательность действий при построении сетевой модели?

8. Поясните, каким образом можно выявить наличие "контура" в условиях выполнения работ при использовании рассмотренного алгоритма построения сетевых моделей?

9. Как проверяется возможность исключения фиктивной работы (предположительно "избыточной") в сетевой модели? Приведите примеры, когда такое исключение невозможно.

10. Какова принципиальная последовательность действий при нумерации событий сетевой модели?

11. Выполнение какого условия гарантирует нумерация событий сетевой модели в соответствии с алгоритмом, реализующим метод вычеркивания работ?

12. Каким образом осуществляется оценка длительности выполнения работ программы?

13. Перечислите временные параметры событий и работ сетевой модели. Определите их с организационной точки зрения

14. Дайте определение понятиям: критический путь, критическая операция. Как определяется принадлежность работы критическому пути?

15. Объясните методику расчета временных параметров событий и работ сетевой модели.

16. Если сетевая модель имеет два критических пути, то какой из них определяет надежность реализации программы за директивный срок?

17. Что представляет собой календарный план реализации программы работ, построенный сетевыми методами?

18. Как строится календарный план реализации программы работ?

19. Сформулируйте задачу оптимизации календарного плана по времени при ограниченном суммарном расходе однородного ресурса.

20. Формализуйте в рамках математической модели задачу оптимизации календарного плана по времени при ограниченном расходе однородного ресурса.

21. Какова принципиальная последовательность действий при оптимизации календарного плана по времени при ограниченном суммарном расходе ресурса (длительности работ и расход ресурса фиксирован)?

22. Как реализуются в рассмотренном алгоритме оптимизации по времени условие непрерывного ведения работ и принципы максимальной интенсивности использования ресурса?

23. Сформулируйте задачу оптимизации календарного плана реализации программы работ по затратам.

24. Формализуйте в рамках математической модели задачу оптимизации календарного плана по затратам (при условии линейности функций изменения затрат на выполнение работ в заданных пределах их длительностей).

25. Как осуществляется управление процессом реализации программы работ на основе структурного и календарного планов, построенных сетевыми методами?

26. Задача. Определите условия выполнения и постройте сетевую модель реализации программы работ по отработке очистных блоков, расположенных на двух горизонтах рудника (рис. X.20). Система разработки – горизонтальными слоями с закладкой выработанного пространства, что обусловливает запрещение на работу в соседних блоках и подработку блоков верхнего горизонта. Предполагается первоначальная отработка блоков верхнего горизонта с нечетными номерами, а затем – с четными. В дальнейшем сохраняется аналогичный порядок отработки блоков нижнего горизонта.

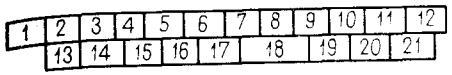


Рис. X.20. Схема расположения очистных блоков

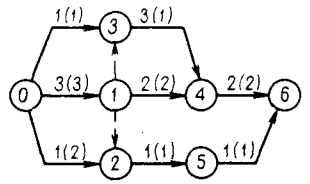


Рис. X.21. Сетевая модель реализации программы с проставленными длительностями работ и величинами фиксированного расхода однородного ресурса (в скобках)

27. Решите задачу 26 при измененном порядке отработки блоков: сначала с четными номерами, а затем – с нечетными.

28. Задача. Пронумеруйте события сетевых моделей, построенных по результатам решения задач 26, 27.

29. Задача. Покажите, что исключение фиктивной работы (7, 8) из сетевой модели реализации программы ввода камеры в эксплуатацию (см. рис. X.14) приводит к нарушению условий выполнения работ.

30. Задача. Рассчитайте временные параметры событий сетевой модели, изображенной на рис. X.21.

31. Задача. Используя результаты решения задачи 30, постройте календарный план реализации программы работ. Определите по календарному плану работы критического пути, а также полные и свободные резервы времени некритических работ.

32. Задача. Используя результаты решения задачи 30, рассчитайте временные параметры работ сетевой модели, изображенной на рис. X.21. Сравните результаты решения задач 31 и 32.

33. Задача. Используя графо-аналитический метод, оптимизируйте календарный план, построенный в результате решения задачи 31, по времени при ограниченном суммарном расходе однородного ресурса, равном 4.

34. Задача. Оцените надежность реализации исходного (результат задачи 31) и оптимального (результат 33) календарных планов за директивный срок, равный 9. Значение дисперсии времени выполнения любой из критических операций условно примите 10 % от ее длительности.

35. Решите задачу оптимизации по времени календарного плана реализации программы работ по вводу камеры в эксплуатацию (см. рис. X.15) при ограниченном суммарном расходе однородного ресурса (6 проходчиков).

36. Решите задачу оптимизации по затратам календарного плана реализации программы работ по вводу камеры в эксплуатацию (см. пример X.11) при условии, что изменение длительности возможно только для работ (0, 1), (0, 2), (0, 7) и (7, 11). При решении задачи используйте данные табл. X.18, в которой для всех работ, кроме указанных, положите $d_{i,j} = D_{i,j}$.

Глава XI

АНАЛИТИЧЕСКИЕ МОДЕЛИ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

§ 1. СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

Большое количество производственных задач и ситуаций может быть описано системами массового обслуживания. Работа системы массового обслуживания заключается в обслуживании большого количества требований. Под требованием понимается запрос на удовлетворение какой-либо потребности, а обслуживание заключается в ее удовлетворении.

Средство, осуществляющее обслуживание потока требований, называют обслуживающим аппаратом, а их совокупность – обслуживающей системой.

Схематическое изображение системы массового обслуживания показано на рис. XI.1, а примеры систем массового обслуживания даны в табл. XI.1

Совокупность требований на обслуживание представляет входящий поток, а обслуженные требования – выходящий. Требования на обслуживание образуют очередь, а порядок, в котором они поступают на обслуживание и покидают очередь, называется дисциплиной очереди. Очередь является упорядоченной, если требования на обслуживание попадают в порядке поступления, и неупорядоченной – если требования обслуживания в случайном порядке. Кроме того, если обслуживанию определенных требований отдается предпочтение перед остальными, то наблюдается обслуживание с приоритетом (например, машины большой грузоподъемности ремонтируются раньше машин малой грузоподъемности, междугородние разговоры предоставляются вне очереди и т.д.).

Рассмотрим последовательно основные составные части системы массового обслуживания.

Анализ работы систем массового обслуживания обычно начинают с входящего потока требований. Различают регулярный и случайный потоки требований (событий). Поток событий называется регулярным.

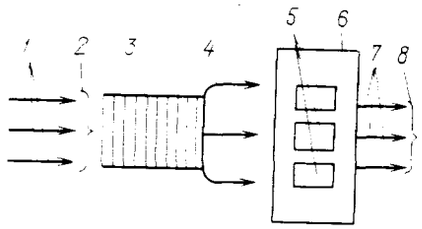


Рис. XI.1. Схема системы массового обслуживания:
1 – требования, поступающие на обслуживание; *2* – входящий поток требований; *3* – очередь; *4* – дисциплина очереди; *5* – обслуживающий аппарат; *6* – обслуживающая система; *7* – обслуженное требование; *8* – входящий поток требований

Таблица XI.1

Требование	Вид обслуживания	Обслуживающий аппарат
Телефонный разговор	Соединение с абонентом	Канал связи
Локомотивосостав (автосамосвал)	Погрузка	Экскаватор
То же	Выгрузка	Отвальный тупик
" "	Пропуск по перегону	Железнодорожный путь
Автосамосвал	Разгрузка	Бункер
Машина, требующая ремонта	Ремонт	Бригада слесарей
Решение организационной задачи	Принятие решения	Руководитель
Расчет на ЭВМ	Предоставление машинного времени	ЭВМ, оператор

если требования поступают на обслуживание в строго определенные промежутки времени (по графику). Из-за влияния на ход производственного процесса случайных факторов (природных, технических, социально-психологических) регулярный поток практически никогда не наблюдается.

Главной характеристикой потока требований является его интенсивность λ – количество требований, поступающих на обслуживание в единицу времени.

Наиболее просто задачи массового обслуживания решаются для простейшего случайного потока требований, основные свойства которого – стационарность, ординарность и отсутствие последействия.

Поток является стационарным, если его интенсивность постоянна ($\lambda = \text{const}$) в любые промежутки времени. Например, количество автосамосвалов, прибывающих на погрузку каждый час, одно и то же. Реальные производственные процессы стационарны на небольшом отрезке времени, исключая обычно время начала и конца смены. Однако при расчете систем массового обслуживания в целях упрощения процесс принимают стационарным в течение всего исследуемого промежутка времени.

Ординарность потока означает поступление требований поодиночке. Если рассматривать достаточно малый промежуток времени (Δt), условие ординарности потока обычно соблюдается. Например, под погрузку к экскаватору машины прибывают поодиночке.

Поток без последействия наблюдается в тех случаях, когда поступление требований не зависит от того, сколько их обслуживалось в системе ранее. Например, выход из строя определенного оборудования не зависит от того, какой его вид ломался ранее.

Для простейшего потока требований количество заявок, поступающих в единицу времени, подчинено пуассоновскому закону, при кото-

ром вероятность наличия в системе k требований за время t выражается формулой

$$P_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{(k)!}. \qquad \text{(XI.1)}$$

Дифференцируя это выражение, получаем плотность распределения вероятностей интервалов времени между двумя заявками

$$f(t) = \lambda e^{-\lambda t}, \qquad \text{(XI.2)}$$

которая подчиняется показательному закону.

Помимо простейшего пуассоновского в теории массового обслуживания рассматриваются и более сложные потоки (пуассоновский с переменной интенсивностью, Эрланга, Пальма и др.).

Кроме характера потока требований, работа системы массового обслуживания зависит от качества функционирования обслуживающих аппаратов, основной характеристикой которых является время обслуживания одного требования или величина, обратная времени обслуживания – интенсивность обслуживания ν. Эта величина показывает количество требований, обслуживаемых в единицу времени. Интенсивность обслуживания – обычно случайная величина.

Решение задач массового обслуживания упрощается, если предположить, что время обслуживания распределено по показательному закону, а плотность распределения определена зависимостью

$$f(t) = \nu e^{-\nu t} \qquad \text{(XI.3)}$$

На практике далеко не всегда входящий поток требований и время обслуживания можно удовлетворительно интерпретировать соответственно простейшим пуассоновским и показательным законами распределения. Однако такое допущение, существенно облегчая решение задачи, незначительно сказывается на расчетных величинах основных показателей систем массового обслуживания.

Системы массового обслуживания подразделяются на системы с потерями и ожиданием.

В системах с потерями (отказами) требование, поступившее в момент, когда все обслуживающие аппараты заняты, получает отказ и покидает систему. Такие требования считаются для данной системы потерянными.

В системах с ожиданиями требование, поступившее в момент, когда все обслуживающие аппараты заняты, становится в очередь и ожидает обслуживания. Время ожидания в очереди в общем случае может быть как ограниченным, так и неограниченным.

Классический пример системы массового обслуживания с потерями – работа автоматической телефонной станции (АТС). Если линия связи или вызываемый абонент занят, то абонент, обратившийся в АТС, кладет трубку, т.е. требование покидает систему необслуженным.

В системах с неограниченной длиной очереди поступившее требование при отсутствии свободных каналов простаивает в ожидании своего обслуживания, причем оно обязательно будет обслужено.

В системах с ограниченным ожиданием накладывается ограничение либо на длину очереди, при которой требование ждет обслуживания, либо на время его пребывания в очереди.

§ 2. ПОРЯДОК РЕШЕНИЯ ЗАДАЧ МАССОВОГО ОБСЛУЖИВАНИЯ

Задачи массового обслуживания обычно решаются в три этапа.

На первом этапе уточняют задачу и выясняют, относится ли она к задачам массового обслуживания. Далее задачу формулируют в терминах массового обслуживания, уточняют систему массового обслуживания и конкретное содержание потока требований и обслуживающих аппаратов. Для решения задачи необходимо знать характеристики входящего потока и обслуживающей системы. Начинается наиболее трудоемкий этап: сбор и обработка статистических сведений. С этой целью организуются хронометражные наблюдения или обрабатываются отчетные (исполнительные) документы. При обработке результатов обычно исключают время начала и конца смены, когда режим работы меняется (неустановившийся), т.е. считают, что работа происходит в стационарном режиме.

Вопросы сбора и обработки информации изучаются математической статистикой. После проведения хронометражных наблюдений и получения фактических функций распределения устанавливают, каким законам они подчиняются.

Прежде всего проверяют соответствие входящего потока требований пуассоновскому закону, а времени обслуживания – показательному, так как в этих случаях решение задачи осуществляется наиболее просто.

Проверку соответствия фактического распределения теоретическому (пуассоновскому, показательному, гамма-распределению, нормальному и др.) осуществляют с помощью специальных критериев, называемых критериями согласия. Наиболее распространен критерий согласия Пирсона χ^2 (хи-квадрат). При проверке по этому критерию сначала определяют меру расхождения χ^2 фактического распределения от предполагаемого теоретического

$$\chi^2 = \sum_{i=1}^{k} \frac{(m_i - Np_i)^2}{Np_i}, \qquad \text{(XI.4)}$$

где k – количество разрядов, в которые сведены результаты опытов; N – общее количество наблюдений; m_i – количество наблюдений в i-м разряде; P_i – теоретическая вероятность (по предполагаемому закону распределения) i-го разряда.

Далее определяется число степеней свободы

r	P					
	0,99	0,98	0,95	0,90	0,80	0,70
1	0,000	0,001	0,004	0,016	0,064	0,148
2	0,020	0,040	0,103	0,211	0,446	0,713
3	0,115	0,185	0,352	0,584	1,005	1,424
4	0,297	0,429	0,711	1,064	1,649	2,20
5	0,554	0,752	1,145	1,610	2,34	3,00
6	0,872	1,134	1,635	2,20	3,07	3,83
7	1,239	1,564	2,17	2,83	3,82	4,67
8	1,646	2,03	2,73	3,49	4,59	5,53
9	2,09	2,53	3,32	4,17	5,38	6,39
10	2,56	3,06	3,94	4,86	6,18	7,27
11	3,05	3,61	4,58	5,58	6,99	8,15
12	3,57	4,18	5,23	6,30	7,81	9,03
13	4,11	4,76	5,89	7,04	8,63	9,93
14	4,66	5,37	6,57	7,79	9,47	10,82
15	5,23	5,98	7,26	8,55	10,31	11,72
16	5,81	6,61	7,96	9,31	11,15	12,62
17	6,41	7,26	8,67	10,08	12,00	13,53
18	7,02	7,91	9,39	10,86	12,86	14,44
19	7,63	8,57	10,11	11,65	13,72	15,35
20	8,26	9,24	10,85	12,44	14,58	16,27
21	8,90	9,92	11,59	13,24	15,44	17,18
22	9,54	10,60	12,34	14,04	16,31	18,10
23	10,20	11,29	13,09	14,85	17,19	19,02
24	10,86	11,99	13,85	15,66	18,06	19,24
25	11,52	12,70	14,61	16,47	18,94	20,9
26	12,20	13,41	15,38	17,29	19,82	21,8
27	12,88	14,12	16,15	18,11	20,7	22,7
28	13,56	14,85	16,93	18,94	21,6	23,6
29	14,26	15,57	17,71	19,77	22,5	24,6
30	14,95	16,31	18,49	20,6	23,4	25,5

$$r = k - c - 1, \qquad \text{(XI.5)}$$

где c – число параметров распределения. Для пуассоновского и показательного закона $c = 1$.

По найденным χ^2 и r с помощью специальной таблицы (табл. XI.2) определяют вероятность совпадения предполагаемого закона распределения с теоретическим.

Если эта вероятность весьма мала, гипотеза отвергается.

Покажем, как устанавливается закон распределения на примере.

Пример XI.1. Исследовать систему массового обслуживания – станция "Породная" – отвальные тупики.

Поток требований составляют локомотивосоставы, поступающие на станцию. Статистические наблюдения показали, что прибытие составов на станцию каждые 0,5 ч распределяется таким образом, как показано в табл. XI.3.

Таблица XI.2

P							
0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
0,455	1,074	1,842	2,71	3,84	5,41	6,64	10,83
1,386	2,41	3,22	4,60	5,99	7,82	9,21	13,82
2,37	3,66	4,64	6,25	7,82	9,84	11,34	16,27
3,36	4,88	5,99	7,78	9,49	11,67	13,28	18,46
4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,50
5,35	7,23	8,56	10,64	12,59	15,03	16,81	22,5
6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,3
7,34	9,52	11,03	13,36	15,51	18,17	20,1	26,1
8,34	10,66	12,24	14,68	16,92	19,68	21,7	27,9
9,34	11,78	13,44	15,99	18,31	21,2	23,2	29,6
10,34	12,90	14,63	17,28	19,68	22,6	24,7	31,3
11,34	14,01	15,81	18,55	21,0	24,1	26,2	32,9
12,34	15,12	16,98	19,81	22,4	25,5	27,7	34,6
13,34	16,22	18,15	21,01	23,7	26,9	29,1	36,1
14,34	17,33	19,31	22,3	25,0	28,3	30,6	37,7
15,34	17,42	20,5	23,5	26,3	29,6	32,0	39,7
16,34	19,51	21,6	21,8	27,6	31,0	33,4	40,8
17,34	20,6	22,8	26,0	28,9	32,3	34,8	42,3
18,34	21,7	23,9	27,2	30,1	33,7	36,2	43,8
19,34	22,8	25,0	28,4	31,4	35,0	37,6	45,3
20,3	23,9	26,2	29,6	32,7	36,3	38,9	46,8
21,3	24,9	27,3	30,8	33,9	37,7	40,3	48,3
22,3	26,0	28,4	32,0	35,2	39,0	44,6	49,7
23,3	27,1	29,6	33,2	36,4	40,3	43,0	51,2
24,3	28,2	30,7	34,4	37,7	41,7	44,3	52,6
25,3	29,2	31,8	35,6	38,9	42,9	45,6	54,1
26,3	30,3	32,9	36,7	40,1	44,1	47,0	55,5
27,3	31,4	34,0	37,9	41,3	45,4	48,3	56,9
28,3	32,5	35,1	39,1	42,6	46,7	49,6	58,3
29,3	33,5	36,2	40,3	43,8	48,0	50,9	59,7

Таблица XI.3

Число составов n	Число наблюдений	Частота фактическая m^1	Частота теоретическая
0	11	0,11	0,135
1	29	0,29	0,270
2	27	0,27	0,270
3	19	0,19	0,180
4	10	0,10	0,090
5	4	0,04	0,036

Предположим, что полученное распределение – пуассоновское. Найдем среднее число прибытий составов за 0,5 ч

$$\overline{L} = \lambda t = \sum_{i=1}^{k} n m_i^1; \qquad \text{(XI.6)}$$

$$\lambda t = 0 \cdot 0{,}11 + 1 \cdot 0{,}29 + 2 \cdot 0{,}27 + 3 \cdot 0{,}19 + 4 \cdot 0{,}10 + 5 \cdot 0{,}04 = 2,$$

или в среднем каждый час на станцию прибывает 4 состава ($\lambda = 4$).

Далее вычисляем вероятности прибытия по пуассоновскому закону

$$P_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \qquad (XI.7)$$

т.е. последовательно вычисляем вероятность прибытия на станцию соответственно 0, 1, 2, 3, 4 и 5 составов. Результаты вычислений рассматриваемого примера показаны в графе 4 табл. XI.3.

Далее определяем критерий Пирсона χ^2

$$\chi^2 = \frac{11 - 0{,}135 \cdot 100^2}{0{,}135 \cdot 100} + \frac{29 - 0{,}270 \cdot 100^2}{0{,}270 \cdot 100} + \frac{27 - 0{,}27 \cdot 100^2}{0{,}270 \cdot 100} +$$

$$+ \frac{19 - 0{,}18 \cdot 100^2}{0{,}18 \cdot 100} + \frac{10 - 0{,}09 \cdot 100^2}{0{,}09 \cdot 100} + \frac{4 - 0{,}036 \cdot 100^2}{0{,}036 \cdot 100} = 0{,}82.$$

Число степеней свободы в рассматриваемом примере при 6-ти разрядах $r = 6 - 1 - 1 = 4$.

Далее по таблице находим, что вероятность соответствия фактического распределения теоретическому $P = 0{,}92$, т.е. распределение потока составов с вероятностью 0,92 подчиняется пуассоновскому закону. Полученная вероятность весьма высокая.

После установления законов распределения входящего потока требований и времени обслуживания на втором этапе с помощью теории массового обслуживания в зависимости от способа организации обслуживающей системы и входящего потока требований (числа обслуживающих каналов, максимального количества требований в системе и др.), а также параметров законов распределения входящего потока требований и времени обслуживания определяют основные характеристики системы.

Для оценки функционирования системы массового обслуживания используют следующие основные ее характеристики:

абсолютную пропускную способность системы (количество требований, обслуживаемых системой в единицу времени);

относительную пропускную способность системы (отношение количества обслуживаемых требований к их общему количеству, поступившему в систему);

средний процент необслуженных требований;

среднее время простоя системы из-за отсутствия требований;

среднюю длину очереди;

среднее время ожидания.

Последние два показателя используют для расчета систем массового обслуживания с ожиданием.

На третьем этапе, используя полученные характеристики систем массового обслуживания и зная необходимые стоимостные данные, находят оптимум решаемой задачи.

При этом, так как обычно вариантов немного, используется метод

их перебора. В более сложных случаях решается задача целочисленного программирования.

Помимо оценки по экономическим показателям выбор варианта организации системы массового обслуживания может осуществляться по требуемому значению одной из ее характеристик.

§ 3. МОДЕЛИРОВАНИЕ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С ОТКАЗАМИ

Первоначально рассмотрим простейшую систему массового обслуживания — работу одноканальной системы массового обслуживания с потерями. На примере этой системы покажем основные принципы получения формул для расчета систем массового обслуживания.

На вход системы поступает пуассоновский поток требований с интенсивностью λ, а время обслуживания $T_{об}$ распределяется по показательному закону, интенсивность которого $\nu = 1/T_{об}$. Требование, поступающее в систему в тот момент, когда обслуживающий канал занят, покидает систему.

Необходимо определить основные характеристики системы: абсолютную пропускную способность A, т.е. ее производительность, и относительную пропускную способность q, эквивалентную коэффициенту полезного действия системы.

Формулы для анализа и расчета систем массового обслуживания получают следующим образом:

строят граф состояний системы S;

описывают вероятности состояний системы;

определяют вероятности переходов из одного состояния в другое;

строят дифференциальные уравнения поведения системы;

решают систему дифференциальных уравнений;

на основе решения системы получают зависимости для расчета характеристик системы массового обслуживания.

Одноканальная система массового обслуживания S может находиться в двух состояниях:

S_0, когда в системе нет требований и обслуживающий канал свободен;

S_1, когда в системе имеется требование и канал занят его обслуживанием.

Из состояния S_0 система может перейти в состояние S_1 и, наоборот. Из состояния S_0 в S_1 система переходит при поступлении требования, а из состояния S_1 в S_0 — по окончании обслуживания требования. Иначе говоря, из состояния S_0 в состояние S_1 систему переводят входящий поток с интенсивностью λ, а из S_1 в S_0 — поток обслуживаний с интенсивностью ν.

Граф состояний системы приведен на рис. XI. 2. Обозначим вероятности состояний S_0 и S_1 соответственно P_0 и P_1. Очевидно, что $P_0 + P_1 = 1$.

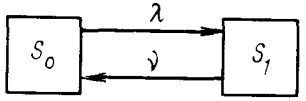


Рис. XI.2. Граф состояний системы

Определим вероятность пребывания системы в состоянии S_0 (P_0) и изменения этого состояния за малый отрезок времени Δt, т.е. вероятность того, что в момент $(t + \Delta t)$ система будет в состоянии S_0. Это событие может произойти двумя способами:

в момент t система находилась в состоянии S_0 и за время Δt не изменила состояния $(S_0 \to \Delta t \to S_0)$;

в момент t система была в состоянии S_1 и за время Δt перешла в состояние S_0 $(S_1 \to \Delta t \to S_0)$.

Вероятность первого варианта обозначим P_A, а второго – P_B. Вероятность первого варианта найдем по теореме умножения вероятностей. Она равна произведению вероятности пребывания системы в состоянии S_0 на условную вероятность того, что система из состояния S_0 не перейдет в S_1

$$P_A = P_0(t)\, P_0(\Delta t). \qquad (XI.8)$$

Так как поток пуассоновский, по формуле (XI.7) получим

$$P_0(\Delta t) = \frac{(\lambda \Delta t)^0}{0!} e^{-\lambda \Delta t} = e^{-\lambda \Delta t}. \qquad (XI.9)$$

Так как Δt – малая величина, то $P_0(\Delta t) = e^{-\lambda \Delta t} \approx 1 - \lambda \Delta t$.

Тогда вероятность того, что состояние системы не изменится, будет равна

$$P_A = P_0(t)(1 - \lambda \Delta t). \qquad (XI.10)$$

Аналогично можно получить, что

$$P_B = P_1(t)\, \nu \Delta t.$$

Тогда

$$P_0(t + \Delta t) = P_0(1 - \lambda \Delta t) + P_1 \nu \Delta t.$$

Раскроем скобки в правой части, перенесем P_0 в левую часть и разделим обе части равенства на Δt, в результате этого получим

$$\frac{P_0(t + \Delta t) - P_0}{\Delta t} = \frac{-P_0 \lambda \Delta t + P_1 \nu \Delta t}{\Delta t} = \qquad (XI.11)$$

$$= -P_0 \lambda + P_1 \nu.$$

При Δt, стремящемся к нулю, перейдем к пределам и получим

$$\lim_{\Delta t \to 0} \frac{P_0(t + \Delta t) - P_0}{\Delta t} = -P_0 \lambda + P_1 \nu. \qquad (XI.12)$$

Так как выражение в правой части представляет первую производную, получаем дифференциальное уравнение для состояния S_0

$$\frac{dP_0}{dt} = -\lambda P_0 + \nu P_1. \qquad (XI.13)$$

Аналогичным образом можно получить дифференциальное уравнение состояния S_1

$$\frac{dP_1}{dt} = -\nu P_1 + \lambda P_0. \qquad (XI.14)$$

Таким образом, дифференциальные уравнения для вероятностей состояний системы имеют вид

Так как $P_0 + P_1 = 1$, то $\frac{dP_0}{dt} = -(\lambda + \nu)P_0 + \nu.$ (XI.15)

Решая полученное уравнение для начальных условий $P_0(0) = 0$ и $P_1(t) = 1$ (т.е. в начальный момент времени канал свободен), получаем

$$P_0 = \frac{\nu}{\nu + \lambda}. \qquad (XI.16)$$

Для одноканальной системы вероятность того, что канал свободен (P_0) соответствует относительной пропускной способности системы, так как требование будет обслужено только в том случае, если в момент его поступления канал свободен.

Итак, относительная пропускная способность

$$q = P_0 = \frac{\nu}{\lambda + \nu}, \qquad (XI.17)$$

а абсолютная

$$A = \lambda q = \frac{\lambda \nu}{\lambda + \nu}. \qquad (XI.18)$$

Вероятность отказа, соответствующая среднему проценту необслуженных требований,

$$P_{отк} = 1 - P_0 = P_1 = 1 - \frac{\nu}{\lambda + \nu} = \frac{\lambda}{\lambda + \nu}. \qquad (XI.19)$$

Пример XI.2. Экскаватор обслуживает поток автосамосвалов, поступающих под погрузку. Интенсивность потока автосамосвалов $\lambda = 0{,}35$ шт/мин, т.е. в среднем автосамосвал поступает под погрузку примерно через 3 мин.

Средняя продолжительность погрузки (время обслуживания) $t_{об} = 2$ мин.

Автосамосвал, застающий экскаватор занятым, уходит под погрузку к другому (требование для данной системы теряется). Требуется определить в установившемся режиме производительность экскаватора Q (шт/ч) и вероятность потери требования.

Определяем интенсивность обслуживания

$$\nu = \frac{1}{t_{об}} = \frac{1}{2} = 0{,}5.$$

Рассчитываем относительную пропускную способность системы

$$q = \frac{\nu}{\lambda + \nu} = \frac{0{,}5}{0{,}5 + 0{,}35} = 0{,}59.$$

Следовательно, под данный экскаватор будет поступать в среднем 59 % прибывающих автосамосвалов, остальные будут уходить на погрузку к другому экскаватору.

По формуле (XI. 18) находим абсолютную пропускную способность системы, т.е. производительность экскаватора

$$Q = 60\lambda q = 60 \cdot 0{,}35 \cdot 0{,}59 = 12 \text{ шт/ч}.$$

Вероятность отказа будет равна

$$P_{отк} = 1 - q = 1 - 0{,}59 = 0{,}41,$$

т.е. в среднем 41 % автосамосвалов будет заставать экскаватор занятым.

Отсюда видно, какую большую роль играет случайность прибытия автосамосвалов под погрузку. Если бы они поступали регулярно через 2 мин один после другого, то производительность экскаватора была бы почти вдвое выше, чем тогда, когда они поступают в случайные моменты времени.

Приведем без специального вывода формулы для расчета многоканальных систем массового обслуживания с отказами. Система включает n каналов обслуживания. Входящий поток требований подчиняется пуассоновскому закону с интенсивностью λ, а время обслуживания – показательному с интенсивностью ν. Если в момент поступления требования нет свободных обслуживающих каналов (т.е. заняты все каналы, так как в системе уже находится n требований), то требование покидает систему необслуженным. Обозначим через ρ отношение λ/ν, где ρ – приведенная интенсивность потока требований, которая характеризует число требований, поступивших в систему за время обслуживания одного.

Вероятность наличия в системе k требований

$$P_k = \frac{\rho^k}{k!} P_0, \qquad (XI.20)$$

где P_0 – вероятность отсутствия требований в системе

$$P_0 = \frac{1}{1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!}} = [1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!}]^{-1}. \qquad (XI.21)$$

Формулы (XI.20) и (XI.21) получены шведским ученым Эрлангом и носят его имя.

Вероятность потери требования (отказа) соответствует вероятности наличия в системе n требований, когда все n каналов заняты

$$P_{отк} = P_n = \frac{\rho^n}{n!} P_0. \qquad (XI.22)$$

Относительная пропускная способность

$$q = 1 - P_{отк} = 1 - P_n. \qquad (XI.23)$$

Абсолютная пропускная способность

$$A = \lambda (1 - P_n). \qquad (XI.24)$$

Среднее количество заявок, находящихся в системе, соответствует среднему числу занятых каналов и определяется по формуле

$$\overline{k} = \frac{A}{\nu} = \rho (1 - P_n). \qquad (XI.25)$$

Пример XI.3. АТС горно-обогатительного комбината имеет три канала связи с городской АТС. В среднем каждые 2 мин поступает запрос на соединение между городской АТС и АТС комбината, а средняя продолжительность одного разговора 4 мин. В результате статистических наблюдений и их обработки установлено, что поток заявок подчинен пуассоновскому закону, а время обслуживания – показательному.

Найти абсолютную пропускную способность АТС, вероятность отказа в соединении, среднее количество занятых каналов, а также сколько необходимо иметь каналов связи, чтобы не менее 95 % вызовов удовлетворялись, т.е. чтобы вероятность отказа не превышала 5 %.

Интенсивность потока требований

$$\lambda = \frac{1}{t_{инт}} = \frac{1}{2} = 0{,}5.$$

Интенсивность обслуживания

$$\nu = \frac{1}{t_{об}} = \frac{1}{4} = 0{,}25.$$

Приведенная интенсивность потока требований

$$\rho = \frac{\lambda}{\nu} = \frac{0{,}5}{0{,}25} = 2.$$

По формуле Эрланга (XI. 21) подсчитываем P_0 и P_n для $n = 3$

$$P_0 = [1 + \frac{2}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!}]^{-1} = 6{,}33^{-1} \approx 0{,}16;$$

$$P_n = P_{отк} = \frac{2^3}{3!} \cdot 0{,}16 = 0{,}21.$$

Таким образом, 21 % вызовов получают отказ, а относительная пропускная способность равна

$$q = 1 - 0{,}21 = 0{,}79.$$

Абсолютная пропускная способность АТС

$$A = \lambda (1 - P_n) = 0{,}5 \cdot 0{,}79 = 0{,}4 \text{ соединений в минуту}.$$

Среднее число занятых каналов связи

$$\overline{k} = \rho (1 - P_n) = 2 \cdot 0{,}79 \approx 1{,}6.$$

Определим, сколько необходимо иметь каналов связи, чтобы $P_{отк} \leqslant 0{,}05$.

Увеличивая последовательно число каналов связи, определяем характеристики системы массового обслуживания.

При $n = 4$

$$P_0 = [1 + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!}]^{-1} = 7^{-1} = \frac{1}{7};$$

$$P_4 = P_{отк} = \frac{2^4}{4!} \cdot \frac{1}{7} = \frac{2}{21} \approx 0{,}1.$$

Следовательно, при 4 каналах связи не получают соединения около 10 % вызовов.

При $n = 5$

$$P_0 = [1 + \frac{2^1}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!} + \frac{2^5}{5!}]^{-1} = 7{,}27^{-1} \approx 0{,}14;$$

$$P_5 = P_{отк} = \frac{2^5}{5!} \cdot 0{,}14 \approx 0{,}04.$$

Следовательно, в данном случае число отказов не превышает 5 % и для обеспечения заданной надежности соединения необходимо иметь не менее 5 каналов связи.

Подсчитаем, сколько в среднем каналов будет работать

$$\overline{k} = \rho (1 - P_5) = 2(1 - 0{,}04) \approx 2.$$

Таким образом, в среднем 2 канала будут работать, а остальные 3 простаивать. Такой дорогой ценой достигается требуемый высокий уровень обслуживания (относительная пропускная способность $q = 1 - P_5 = 0{,}96$).

Пример XI.4. Для условий предыдущей задачи определить, сколько необходимо иметь каналов связи, если затраты на эксплуатацию одного канала ($C_к$) составляют 100 руб., а экономические потери от неудовлетворения 1 % требований на соединение с абонентом (C_0) равны 10 руб.

Для определения оптимального числа отказов необходимо рассчитать общие затраты на эксплуатацию каналов связи и потери от отказов в соединении с абонентом

($Z = nC_к + C_0 P_{отк} \rightarrow \min$).

Воспользуемся рассчитанными ранее значениями $P_{отк}$ и определим Z (в руб.):

при $n = 3$ $Z = 3 \cdot 100 + 10 \cdot 21 = 510$;

" $n = 4$ $Z = 4 \cdot 100 + 10 \cdot 10 = 500$;

" $n = 5$ $Z = 5 \cdot 100 + 10 \cdot 4 = 540$.

Минимальная суммарная величина затрат и потерь соответствует 4 каналам связи.

§ 4. РАЗОМКНУТЫЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ОЖИДАНИЕМ

К разомкнутым системам массового обслуживания относятся такие, в которые требования поступают извне и интенсивность их потока не зависит от состояний систем. Например, поступление порожних вагонов под погрузку готовой продукции, поступление машин в ремонт на ЦРМ, прибытие вагонов с лесом на склад шахты может интерпретироваться как входящий поток требований разомкнутой системы массового обслуживания.

Разомкнутые делятся на системы с ограниченной и неограниченной длиной очереди. Если количество мест в очереди ограничено, то

заявка, пришедшая в момент, когда все места уже заняты, покидает систему необслуженной. При неограниченной длине очереди требование в любом случае становится в очередь и дожидается обслуживания.

Зависимости для расчета основных характеристик разомкнутых систем массового обслуживания с ожиданием получают в результате построения графа и решения дифференциальных уравнений состояний системы.

Приведем без специального вывода формулы для расчета разомкнутых систем массового обслуживания.

Первоначально рассмотрим систему с ограниченным числом мест ожидания в очереди.

На вход системы поступает пуассоновский поток требований с интенсивностью λ, обслуживание осуществляется n-обслуживающими аппаратами с интенсивностью ν, время обслуживания распределено по показательному закону. Число мест в очереди l.

Вероятность простоя системы, соответствующая вероятности отсутствия в ней требований, определяется по формуле

$$P_0 = [1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!} \cdot \frac{\rho/n - (\rho/n)^{l+1}}{1 - \rho/n}]^{-1}, \quad (XI.26)$$

где ρ — приведенная интенсивность потока заявок требований; показывает среднее количество требований, приходящих в систему за среднее время обслуживания одного требования.

Обозначив ρ/n через χ, перепишем формулу (XI.26) в виде

$$P_0 = [1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!} \cdot \frac{\chi - \chi^{l+1}}{1 - \chi}]^{-1}. \quad (XI.27)$$

Вероятность отказа в обслуживании находится из расчета вероятности наличия в системе $(n + l)$ требований, т.е. когда заняты все обслуживающие аппараты n и все места в очереди l

$$P_{отк} = P_{n+l} = \frac{\rho^{l+n}}{n^l n!} P_0. \quad (XI.28)$$

Относительная пропускная способность системы

$$q = 1 - P_{отк} = 1 - \frac{\rho^{n+l}}{n^l n!} P_0. \quad (XI.29)$$

Абсолютная пропускная способность системы

$$A = q\lambda = \lambda (1 - \frac{\rho^{n+l}}{n^l n!} P_0). \quad (XI.30)$$

Среднее количество занятых каналов

$$\overline{Z} = \frac{A}{\nu} = \rho (1 - \frac{\rho^{n+l}}{n^l n!} P_0). \quad (XI.31)$$

Среднее количество требований в очереди

$$\overline{r} = \frac{\rho^{n+1} P_0}{n - n!} \cdot \frac{1 (l+1) \chi^l + l\chi^{l+1}}{(1-\chi)^2}. \quad (XI.32)$$

Среднее количество требований в системе

$$\overline{k} = \overline{r} + \overline{z}. \quad (XI.33)$$

Среднее время ожидания требования в очереди

$$\overline{t}_{ож} = \frac{\overline{r}}{\lambda} = \frac{\rho^{n+l} P_0}{\lambda n n!} \cdot \frac{1 - (l+1)\chi^l + l\chi^{l+1}}{(1-\chi)^2}. \quad (XI.34)$$

Среднее время пребывания требования в системе

$$\overline{t}_c = \overline{t}_{ож} + \overline{t}_{об} = \overline{t}_{ож} + \frac{q}{\lambda}. \quad (XI.35)$$

Для расчета разомкнутой системы массового обслуживания с неограниченной длиной очереди можно использовать приведенные зависимости, если принять $l = \infty$.

Вероятность отсутствия требований в системе с неограниченной длиной очереди

$$P_0 = [1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{n! (n - \rho)}]^{-1} \quad (XI.36)$$

Так как в системе с неограниченной длиной очереди обслуживаются все поступающие заявки, то

$$P_{отк} = 0, \quad q = 1, \quad A = \lambda q = \lambda.$$

Среднее количество требований в очереди

$$\overline{r} = \frac{\rho^{n+1} P_0}{nn! (1-\chi)^2}. \quad (XI.37)$$

Среднее время ожидания

$$\overline{t}_{ож} = \frac{\rho^n P_0}{n \nu n! (1-\chi)^2}. \quad (XI.38)$$

Среднее количество занятых каналов

$$\overline{Z} = \frac{A}{\nu} = \frac{\lambda}{\nu} = \rho. \quad (XI.39)$$

Для одноканальных систем массового обслуживания (при $n = 1$) формулы для расчета основных характеристик системы упрощаются и принимают следующий вид.

Для одноканальных разомкнутых систем с ожиданием и ограниченной длиной очереди

$$P_0 = \frac{1 - \rho}{1 - \rho^{l+2}}; \quad (XI.40)$$

$$P_{отк} = P_{l+1} = \frac{\rho^{l+1} (1 - \rho)}{1 - \rho^{l+2}}; \quad (XI.41)$$

$$q = 1 - P_{отк} = 1 - \frac{\rho^{l+1} (1 - \rho)}{1 - \rho^{l+2}}; \quad (XI.42)$$

$$\overline{r} = \frac{\rho^2 [1 - \rho^l (l + 1 - l\rho)]}{(1 - \rho^{l+2})(1 - \rho)}; \quad (XI.43)$$

$$\overline{k} = \overline{r} + \frac{\rho - \rho^{l+2}}{1 - \rho^{l+2}}; \quad (XI.44)$$

$$\overline{t}_{ож} = \frac{\overline{r}}{\lambda} = \frac{\rho [1 - \rho^l (l - l\rho)]}{\nu (1 - \rho^{l+2})(1 - \rho)}. \quad (XI.45)$$

Для одноканальных разомкнутых систем массового обслуживания с ожиданием при неограниченной длине очереди

$$P_0 = 1 - \rho; \quad (XI.46)$$

$$\overline{r} = \frac{\rho^2}{1 - \rho}; \quad (XI.47)$$

$$\overline{k} = \frac{\rho}{1 - \rho}; \quad (XI.48)$$

$$\overline{t}_{ож} = \frac{\rho^2}{\lambda (1 - \rho)} = \frac{\rho}{\nu (1 - \rho)}. \quad (XI.49)$$

Пример XI.5. На обогатительную фабрику, имеющую два бункера, поступает руда с нескольких карьеров и рудников. Проведенные хронометражные наблюдения и статистическая обработка их материалов показали, что в среднем за час прибывает под разгрузку 54 автосамосвала, а среднее время разгрузки одной машины 2 мин. Интервалы времени между прибытиями автосамосвалов и время разгрузки распределены по показательному закону.

Требуется определить:

среднее время простоя автосамосвала при условии, что все они должны разгружаться непосредственно в бункер;

процент автосамосвалов, которые будут разгружаться на складе, если время простоя автомашины в ожидании разгрузки в бункер не должно превышать 2 мин;

максимальную длину очереди, при которой автосамосвал должен разгружаться на складе.

Распределение интервалов времени между прибытием автосамосвалов под разгрузку свидетельствует о пуассоновском характере потока. Рассчитаем основные характеристики системы: интенсивность потока требований

$\lambda = 54/60 = 0{,}9$ машин/мин;

интенсивность обслуживания

$\nu = 1/2 = 0{,}5$ машин/мин;

приведенная интенсивность

$\rho = \lambda/\nu = 0{,}9/0.5 = 1{,}8.$

$\chi = \rho/n = 1{,}8/2 = 0{,}9.$

Для ответа на первый вопрос рассматриваемую систему можно интерпретировать как разомкнутую систему массового обслуживания с ожиданием и бесконечной длиной очереди. Длина очереди принимается бесконечной, так как все автосамосвалы по мере поступления должны разгрузиться в бункер. При этом имеем $n = 2$ и $\rho = 1{,}8$.

Вероятность отсутствия требований в системе рассчитываем по формуле (XI.36).

$$P_0 = [1 + \frac{1{,}8}{1!} + \frac{1.8^2}{1 \cdot 2} + \frac{1{,}8^3}{1 \cdot 2 \cdot 3}]^{-1} = 0{,}053.$$

Тогда среднее время ожидания разгрузки определяем по формуле (XI.38)

$$\overline{t}_{ож} = \frac{1.8^2 \cdot 0.053}{2 \cdot 2 \cdot 0{,}5 \cdot 0{,}2^2} = 2{,}15 \text{ мин}.$$

Если все машины разгружаются обязательно в бункер, то среднее время ожидания в очереди превышает допустимое (2 мин). Поэтому очередь необходимо ограничить и часть автосамосвалов, прибывающих на фабрику, разгружать на складе, если все места заняты.

Для ответа на второй вопрос рассматриваемую систему можно представить как разомкнутую систему массового обслуживания с ожиданием и ограниченной длиной очереди. Рассмотрим характеристики этой системы и определим, при каком количестве мест ожидания в очереди ($l = ?$) время ожидания не превысит 2 мин. При этом процент автосамосвалов, разгружаемых на складе, а не в бункер, соответствует вероятности отказа в обслуживании ($P_{отк}$). Сначала примем количество мест в очереди равным 1.

Итак, имеем $n = 2$, $l = 1$, $\rho = 1{,}8$ и $\chi = 0{,}9$.

По формулам (XI.27), (XI.28), (XI.32), (XI.34) рассчитаем P_0, $P_{отк}$, $\overline{r}$ и $\overline{t}_{ож}$

$$P_0 = [1 + \frac{1{,}8^1}{1!} + \frac{1{,}8^2}{2!} + \frac{1{,}8^2}{2!} \cdot \frac{0{,}9 - 0{,}9^{1+1}}{1 - 0{,}9}]^{-1} = 0{,}22.$$

Вероятность отказа

$$P_{отк} = P_{n+l} = P_3 = \frac{1{,}8^3}{2^3 \cdot 2} 0{,}22 = 0{,}08.$$

Среднее количество автосамосвалов в очереди

$$\overline{r} = \frac{1{,}8^3 \cdot 0{,}22}{2 \cdot 2} \cdot \frac{1 - (1 + 1)\,0{,}9^1 + 1 \cdot 0{,}9^2}{(1 - 0{,}9)^2} = 0{,}32.$$

Среднее время ожидания начала разгрузки в очереди

$\overline{t}_{ож} = 0{,}32/0.9 = 0{,}36$ мин.

Если длину очереди ограничить одним автосамосвалом, то на склад будет разгружаться 8 % машин, а среднее время простоя в очереди составит 0,36 мин.

Рассчитаем теперь характеристики системы, если в очереди могут находиться два автосамосвала. Для этого случая имеем $n = 2$, $l = 2$, $\rho = 1.8$, $\chi = 0{,}9$ и определяем P_0, $P_{отк}$, $\overline{r}$ и $\overline{t}_{ож}$ по формулам (XI.27), (XI.28), (XI.32) и (XI.34)

$$P_0 = [1 + \frac{1{,}8}{1!} + \frac{1{,}8^2}{2!} + \frac{1{,}8^2}{2!} \cdot \frac{0{,}9 - 0{,}9^{2+1}}{1 - 0{,}9}]^{-1} = 0{,}16;$$

$$P_{отк} = P_{n+l} = P_4 = \frac{1{,}8^4}{2^4 \cdot 2} 0{,}16 = 0{,}052;$$

$$\overline{r} = \frac{1{,}8^3 \cdot 0{,}16}{2 \cdot 2} \cdot \frac{1 - (2 + 1)\,0{,}9^2 + 1 \cdot 0{,}9^{2+1}}{(1 - 0{,}9)^2} = 0{,}66;$$

$$\overline{t}_{ож} = \frac{0{,}66}{0{,}9} = 0{,}73 \text{ мин}.$$

Следовательно, если длину очереди ограничить двумя автосамосвалами, то на складе будут разгружаться 5,2 % машин, а время ожидания начала разгрузки составит 0,73 мин.

Теперь рассчитаем эти же характеристики, если длину очереди, при которой автосамосвал разгружается на складе (не становится в очередь), принять равной трем автосамосвалам.

Для этого случая имеем $n = 2$, $l = 3$, $\rho = 1{,}8$, $\chi = 0{,}9$ и рассчитываем основные характеристики системы

$$P_0 = [1 + \frac{1{,}8}{1!} + \frac{1{,}8^2}{2!} + \frac{1{,}8^2}{2!} \cdot \frac{0{,}9 - 0{,}9^{3+1}}{1 - 0{,}9}]^{-1} = 0{,}12;$$

$$P_{отк} = P_{n+l} = P_5 = \frac{1{,}8^5}{2^5 \cdot 2} 0{,}12 = 0{,}035;$$

$$\overline{r} = \frac{1{,}8^3 \cdot 0{,}12}{2 \cdot 2} \cdot \frac{1 - (3 + 1)\,0{,}9^3 + 3 \cdot 0{,}9^{3+1}}{(1 - 0{,}9)^2} = 1{,}93;$$

$$\overline{t}_{ож} = \frac{1{,}93}{0{,}9} = 2{,}15.$$

Так как по условиям задачи максимальное время ожидания в очереди не должно превышать 2 мин, длину очереди, при которой машины разгружаются на складе, следует ограничить двумя автосамосвалами.

Таким образом, если при подъезде к обогатительной фабрике шофер видит, что в очереди уже стоят две машины, то он разгружает автосамосвал на складе.

§ 5. ЗАМКНУТЫЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

Системы массового обслуживания, в которых функционирует определенное ограниченное количество требований, являются замкнутыми. В этом случае интенсивность потока требований зависит от состояния системы. В принципе любая система массового обслуживания является замкнутой, но иногда количество потенциальных источников требований велико (например, количество абонентов АТС и т.д.), и влиянием работы обслуживающей системы на поток заявок пренебрегают.

Примеры замкнутых систем массового обслуживания с ожиданием — работа шахтного или карьерного транспорта, системы ремонта и др.

Процесс функционирования замкнутой системы с ожиданием формулируется в следующем виде.

На n одинаковых обслуживающих аппаратов поступает простейший поток требований интенсивностью λ_m и общее количество требований в системе m. Если в момент поступления требования имеется хотя бы один свободный обслуживающий аппарат, оно немедленно начинает обслуживаться. Если же все обслуживающие аппараты заняты, то вновь прибывшее требование становится в очередь за всеми теми, которые поступили раньше и еще не начали обслуживаться. Освободившийся аппарат немедленно приступает к обслуживанию очередного требования. Каждое требование обслуживается только одним из аппаратов, и каждый аппарат в каждый момент обслуживает не более одного требования. Время обслуживания одного требования является случайной величиной с показательным законом распределения. В этом предположении задача допускает простое решение, которое с удовлетворительной для практики точностью описывает ход интересующего процесса. Опуская теоретические выкладки, приведем выражения, определяющие некоторые характеристики системы массового обслуживания с ожиданием.

1. Вероятность P_k того, что в системе находится k требований для случая, когда их число больше, чем обслуживающих аппаратов ($k \geqslant n$)

$$P_k = \frac{m!}{n^{k-n} n! (m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k P_0, \qquad (XI.50)$$

если $k \leqslant n$,

$$P_k = \frac{m!}{k! (m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k P_0. \qquad (XI.51)$$

2. Вероятность P_0 того, что все обслуживающие аппараты свободны

$$P_0 = \left[\sum_{k=0}^{n} \frac{m!}{k! (m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k + \right. \qquad (XI.52)$$

$$\left. + \sum_{k=n+1}^{m} \frac{m}{n^{k-n} n! (m-k)!} \left(\frac{\lambda}{\nu}\right)^k \right]^{-1}.$$

3. Средняя длина очереди M_1, т.е. математическое ожидание числа требований, ждущих начала обслуживания

$$M_1 = \sum_{k=n+1}^{m} (k-n) P_k. \qquad (XI.53)$$

4. Коэффициент простоя обслуживаемого требования

$$\frac{M_1}{m} = \frac{(m-1)!}{n!} \sum_{k=n+1}^{m} \frac{k-n}{n^{k-n} (m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k P_0. \qquad (XI.54)$$

5. Среднее количество свободных обслуживающих аппаратов

$$M_2 = \sum_{k=1}^{n-1} (n-k) P_k. \qquad (XI.55)$$

6. Коэффициент простоя одного обслуживающего аппарата

$$\frac{M_2}{n} = \sum_{k=0}^{n-1} P_k - \frac{1}{n} \sum_{k=0}^{n-1} k P_k. \qquad (XI.56)$$

В этих формулах приняты следующие обозначения:

λ_m – частота вращения требования в систему; $1/\nu$ – среднее время обслуживания одного требования; m – наибольшее количество требований, находящихся в системе одновременно; k – текущее количество требований; n – количество обслуживающих аппаратов.

Для одноканальной замкнутой системы массового обслуживания ($n = 1$) расчетные формулы значительно упрощаются и принимают вид

$$P_k = \frac{m!}{(m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k P_0; \qquad (XI.57)$$

$$P_0 = \left[\frac{m!}{k! (m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k + \sum_{k=2}^{m} \frac{m!}{(m-k)!} \left(\frac{\lambda_m}{\nu}\right)^k\right]^{-1}; \qquad (XI.58)$$

$$M_1 = \sum_{k=2}^{m} (k-1) P_k; \qquad (XI.59)$$

$$K_т = 1 - \frac{M_1}{m}; \qquad (XI.60)$$

$$K_а = 1 - P_0, \qquad (XI.61)$$

где $K_т$ и $K_а$ – коэффициенты использования соответственно требования и обслуживающего аппарата.

При расчете систем массового обслуживания для удобства обычно вместо формул (XI.50) – (XI.61) пользуются рекуррентной формулой. Для этого обозначим через a_k частное

$$a_k = \frac{P_k}{P_0}, \qquad (XI.62)$$

a_k от $k = 1$ до $k = n$ вычисляются по формуле

$$a_k = \frac{m-k+1}{k} \left(\frac{\lambda_m}{\nu}\right) a_{k-1}, \qquad (XI.63)$$

а от $k = n+1$ до $k = m$ по формуле

$$a_k = \frac{m-k+1}{n} \left(\frac{\lambda_m}{\nu}\right) a_{k-1}. \qquad (XI.64)$$

Пример XI.6. Определить оптимальное количество автосамосвалов на один экскаватор при следующих условиях.

На карьере обслуживание забоев автосамосвалами осуществляется по закрытому циклу. Известно, что производительное использование экскаватора при определенном расстоянии транспортирования зависит от количества автосамосвалов,

прикрепленных к забою. С увеличением количества автосамосвалов (разумеется, до некоторых пределов) производительность экскаваторов обычно увеличивается. В то же время, чем меньше автомашин, тем лучше используется каждая из них, так как уменьшаются простои при погрузке и разгрузке, увеличивается скорость движения. Надо найти такое отношение, при котором потери от простоев экскаваторного и транспортного оборудования будут минимальны.

Данная задача относится к классу задач массового обслуживания. Автосамосвалы, поступающие на погрузку, могут интерпретироваться как входящий поток требований, а экскаватор, как обслуживающий аппарат. Так как машины работают по закрытому циклу, то рассматриваемая система массового обслуживания относится к системам с ожиданием.

Для получения характеристик входящего потока требований ко времени обслуживания, определяющих простои, необходимо рассчитать функции и параметры их распределения. Для этого проводятся статистические наблюдения с последующей их обработкой.

Предположим, что в результате сбора и анализа информации было установлено, что входящий поток требований может рассматриваться как пуассоновский, а время обслуживания (погрузки) распределяется по показательному закону. При этом математическое ожидание среднего времени погрузки 2,32 мин, а среднего времени рейса – 17,4 мин. Это означает, что параметр пуассоновского распределения входящего потока требований $\lambda_m = m/17{,}4$, а параметр показательного распределения времени обслуживания $\nu = 1/2{,}32 = 0{,}432$.

Зная параметры λ_m и ν, с помощью теории массового обслуживания можно ответить на следующие вопросы:

какое среднее количество автосамосвалов будет простаивать в ожидании погрузки?

как меняются коэффициенты использования автосамосвалов и экскаватора $K_а$ и $K_э$ в зависимости от количества автосамосвалов, прикрепленных к одному экскаватору?

По формулам (XI.57) – (XI.61) определяем основные характеристики работы системы. Результаты вычислений сведены в табл. XI.4 и показаны на графике (рис. XI.3).

Таблица XI.4

Характеристика системы	m						
	4	6	8	10	12	14	16
M_1	0,183	0,513	1,65	2,710	3,74	5,46	7,68
$K_э$	0,445	0,633	0,820	0,900	0,942	0,966	0,997
$K_а$	0,954	0,910	0,790	0,730	0,688	0,610	0,530

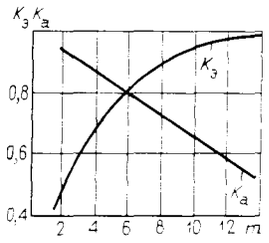
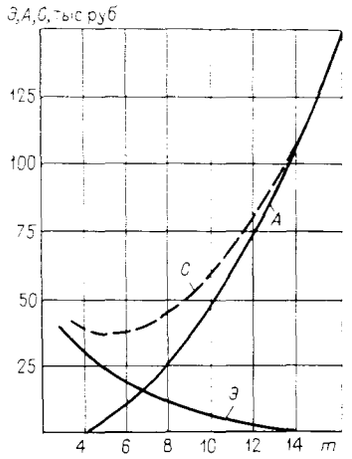


Рис. XI.3. Зависимость коэффициента использования экскаватора ($K_э$) и автосамосвалов ($K_а$) от количества автомашин, прикрепленных к одному экскаватору (m)

Рис. XI.4. Зависимость потерь от простоев экскаватора ($Э$), автосамосвалов (A) и суммарных (C) от количества автомашин, прикрепленных к одному экскаватору (m)



Далее переходим к непосредственному определению оптимального количества автосамосвалов на один экскаватор. В качестве критерия эффективности можно принять экономические потери от простоев экскаваторов и автосамосвалов C.

Естественно, что потери необходимо минимизировать. Если известны стоимостные показатели потерь от простоя экскаваторов и автосамосвалов, равные соответственно $C_э$ и $C_а$, то целевая функция имеет вид

$$C = C_э (1 - K_э) + C_а (1 - K_а) m \rightarrow \min. \qquad (XI.65)$$

Результаты вычислений для рассматриваемого примера ($C_э = 72$ руб. в смену и $C_а = 20$ руб. в смену) сведены в табл. XI.5; зависимость потерь от простоев экскаваторов $Э = C_э (1 - K_э)$, автосамосвалов $A = C_а (1 - K_а)$ и суммарных C от количества автосамосвалов, прикрепленных к одному экскаватору, приведены на рис. XI.4.

Таблица XI.5

Характеристика системы	m						
	4	6	8	10	12	14	16
$Э$	39,2	26,4	12,9	7,2	4,2	2,4	0,2
A	3,6	10,2	33,0	54,0	75,0	109,9	150,0
C	42,8	36,6	45,9	61,2	79,2	111,4	150,2

Минимум кривой C дает оптимальное количество автосамосвалов. В рассматриваемом примере оптимально за одним экскаватором закрепление 6 автосамосвалов.

Пример XI.7. На карьере вскрыша вывозится на отвал железнодорожным транспортом, отвалообразование – экскаваторное. Локомотивосоставы из забоев поступают на станцию "Породная", а оттуда – на отвал. Требуется определить оптимальное количество отвальных тупиков.

При транспортировании вскрыши на отвалы железнодорожным транспортом время рейса меняется в зависимости от времени выполнения операций погрузки, движения, разгрузки, ожидания в карьере или на станции "Породная". Время t с момента ухода состава от фиксированного экскаватора до момента прихода

на станцию "Породная" не постоянно, а колеблется около какого-то расчетного среднего значения.

Процесс прибытия груженых составов на станцию "Породная" можно рассматривать как поток требований на обслуживание (выгрузку), поступающих в систему обслуживания в случайные моменты времени. Отвальные тупики являются обслуживающей системой, а каждый отвальный экскаватор – обслуживающим аппаратом.

Поток составов от экскаваторов на отвалы для больших карьеров является простейшим (пуассоновским).

Справедливость пуассоновского закона распределения плотности потока требований (составов), кроме теоретического обоснования, проверяется также статистической обработкой потока заявок (составов), приходящих из карьера на станцию "Породная".

При этом установлено (гл. XI, § 2), что с вероятностью 0,92 поток может считаться пуассоновским и его показатель (среднее количество требований, поступающих в единицу времени) $\lambda = 4$ состава в час.

Если в системе находится всего $m = 10$ составов (требований), то частота возвращения одного требования в систему $\lambda_m = 4/10 = 0{,}4$.

При известной грузоподъемности состава пропускная способность отвального тупика обратно пропорциональна среднему времени пребывания его на отвале. Его величина равна времени входа от станции "Породная" к отвальному экскаватору и обратно, разгрузки и возможному ожиданию у приемного бункера.

Время пребывания (обслуживания) породного состава на отвальном тупике, зависящее от множества технических, организационных и природных факторов, непостоянно.

Статистическая обработка времени обслуживания составов отвальными механизмами показала, что время обслуживания заявок с достаточно высокой вероятностью подчиняется экспоненциальному закону, причем среднее время обслуживания $t_{об} = 1{,}25$ ч, а следовательно, параметр $\nu = 1/t_{об} = 0{,}8$.

Таким образом, для рассматриваемого примера справедлива следующая модель системы обслуживания: пуассоновский поток заявок с интенсивностью $\lambda_m = 0{,}4$, экспоненциальное время обслуживания с параметром $\nu = 0{,}8$.

Получив модель системы, переходим к определению ее характеристик.

Для этого прежде всего надо подсчитать вероятности наличия в системе k требований (составов) P_k при различном числе отвальных тупиков ($n = 1, 2, 3, 4, 5$).

Для расчета P_k воспользуемся формулами (XI.62) – (XI.64).

В нашем примере $\lambda_m/\nu = 0{,}5$. Определим a_k и P_k для $n = 3$.

По формуле (XI.62) рассчитываем a_k для $k = 0, 1, 2, 3$.

$$a_0 = 1;$$

$$a_1 = \frac{10-0}{1} 0{,}5 \cdot 1 = 5{,}0;$$

$$a_2 = \frac{10-1}{2} 0{,}5 \cdot 5 = 11{,}2;$$

$$a_3 = \frac{10-2}{3} 0{,}5 \cdot 11{,}2 = 15.$$

Затем по формуле (XI.63) определяем a_k для $k = 4 - 10$

$$a_4 = \frac{10-3}{3} 0{,}5 \cdot 15 = 17{,}5;$$

$$a_5 = \frac{10-4}{3} 0{,}5 \cdot 17{,}5 = 17{,}5;$$

$$a_6 = \frac{10-5}{3} 0{,}5 \cdot 17{,}5 = 14{,}5;$$

$$a_7 = \frac{10-6}{3} 0{,}5 \cdot 14{,}5 = 9{,}7;$$

$$a_8 = \frac{10-7}{3} 0{,}5 \cdot 9{,}7 = 4{,}8;$$

$$a_9 = \frac{10-8}{3} 0{,}5 \cdot 4{,}8 = 1{,}6;$$

$$a_{10} = \frac{10-9}{3} 0{,}5 \cdot 1{,}6 = 0{,}3.$$

Так как $\sum_{k=0}^{m} P_k = 1$ или $P_0 + \sum_{k=1}^{m} P_k = 1$,

$1 + \sum_{k=1}^{m} \frac{P_k}{P_0} = \frac{1}{P_0}$, откуда $P_0 = \frac{1}{1 + \sum_{k=1}^{m} a_k}$.

Подсчитаем $\sum_{k=1}^{m} a_k$

$$\sum_{k=1}^{m} a_k = 5 + 11{,}2 + 15 + 17{,}5 + 17{,}5 + 14{,}5 + 9{,}7 + 4{,}8 + 1{,}6 + 0{,}3 = 97{,}1.$$

Получив это значение, определим

$$P_0 = \frac{1}{1 + 97{,}1} = 0{,}0103.$$

Зная P_0, находим

$P_1 = a_1 P_0 = 0{,}0103 \cdot 5 = 0{,}051;$

$P_2 = a_2 P_0 = 0{,}0103 \cdot 11{,}2 = 0{,}115$ и т.д.

Результаты вычислений средней длины очереди составов M_1 и среднего количества простаивающих отвальных тупиков M_2 для $n = 1 - 5$ сведены в табл. XI.6.

Зависимость средней длины очереди и среднего количества простаивающих тупиков M_1 и M_2 от общего числа обслуживающих тупиков n показана на рис. XI.5.

Оптимальное количество отвальных тупиков определяется путем минимизации функции убытков от простоев транспортного и отвального оборудования (и сооружений) с учетом эффективности капиталовложений на транспорт и отвалообразование. При этом целесообразно применять графоаналитический метод расчета.

Сумма убытков от простоев транспортного и отвального оборудования определяется из выражения

Таблица XI.6

n	M_1	M_2
1	7,07	–
2	4,0	0,06
3	1,92	0,29
4	0,71	0,80
5	0,26	1,78

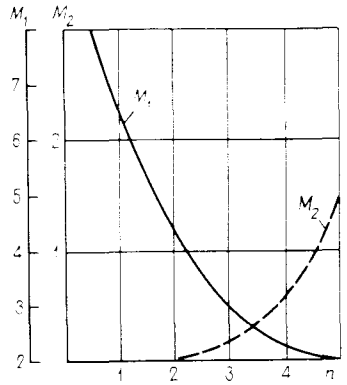


Рис. XI.5. Зависимость средней длины очереди (M_1) и среднего количества простаивающих отвальных тупиков (M_2) от их общего количества (n)

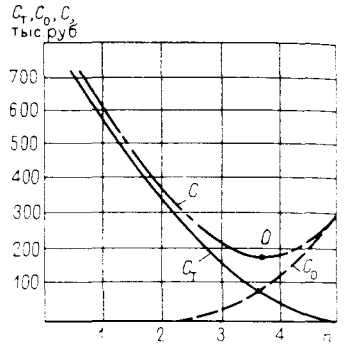


Рис. XI.6. Зависимость потерь от простоев транспорта ($C_т$), отвальных тупиков ($C_о$) и суммарных (C) от количества отвальных тупиков (n)

$$\Sigma C = C_т + C_о = \left(\frac{T_ж K_ж E}{N t_{см}} + C_ж \frac{T_ж}{t_{см}}\right) + \left(\frac{T_о K_о E}{N t_{см}} + C_о \frac{T_о}{t_{см}}\right),$$

где $C_т$ – общие (приведенные) убытки от простоев транспорта, руб.; $C_о$ – то же, для отвальных тупиков, руб.; $T_ж$ – простои локомотивосоставов в смену, ч; $T_о$ – то же, для отвальных тупиков, ч; $K_ж$ – капиталовложения на локомотивосоставы и связанные с его эксплуатацией сооружения, руб.; $K_о$ – капиталовложения на оборудование и сооружения отвального тупика, руб.; $C_ж$ – эксплуатационные потери от простоев локомотивосостава в одну смену, руб.; $C_о$ – то же, от простоев отвального тупика в одну смену, руб.; E – коэффициент эффективности капиталовложений; N – общее количество рабочих смен в году; $t_{см}$ – продолжительность смены, ч.

Очевидно, что $T_ж = M_1 t_{см}$, а $T_о = M_2 t_{см}$, поэтому формула для определения общих потерь примет следующий вид

$$\Sigma C = M_1 \left(\frac{K_ж E}{N} + C_ж\right) + M_2 \left(\frac{K_о E}{N} + C_о\right).$$

Первая часть формулы представляет потери от простоев транспорта $C_т$, а вторая – от отвального тупика $C_о$.

Результаты расчетов для рассматриваемого примера ($C_ж$ = 75 руб., $C_о$ = = 125 руб., N = 800, E = 0,15, $K_ж$ = 150 тыс. руб.) показаны на рис. XI.6.

Минимум кривой (точка О) соответствует оптимальному числу отвальных тупиков. В рассматриваемом примере оптимальным будет устройство 4 отвальных тупиков.

Пример XI.8. На участке горного предприятия имеется 6 единиц оборудования. В результате хронометражных наблюдений установлено, что каждая единица оборудования останавливается в среднем один раз в час, а устранение неисправности занимает у слесаря в среднем 12 мин. При этом интервалы времени между поломками и время ремонта распределены по показательному закону.

Таблица XI.7

k	a_k	P_k	$(k-n)P_k$
0	1	0,19	–
1	1,2	0,23	–
2	1,2	0,23	0,23
3	0,96	0,19	0,38
4	0,58	0,11	0,33
5	0,23	0,04	0,16
6	0,05	0,01	0,05

Необходимо определить: показатели использования оборудования, если на участке один слесарь; насколько улучшатся показатели использования оборудования, если на участке будут работать два слесаря; максимальное количество оборудования, которое можно прикрепить для обслуживания к одному слесарю, чтобы коэффициент использования его был не менее 0,9.

Определяем среднюю длину очереди (среднее количество станков, ожидающих ремонта), при условии, что на участке один слесарь (n = 1) обслуживает все оборудование (m = 6). Так как интервалы между поломками оборудования распределены по показательному закону, то поток требований на ремонт подчиняется пуассоновскому закону. При этом $\lambda_m = 1$; $\nu = 60/12 = 5$; $\lambda_m/\nu = 1/5$.

Расчеты a_k, P_k и $(k-n)P_k$ сведены в табл. XI.7.

Средняя длина очереди

$$M_1 = \sum_{2}^{6} (k-1) P_k = 1{,}15.$$

Следовательно, коэффициент использования оборудования составит

$$\eta_0 = 1 - \frac{M_1}{m} = 1 - \frac{1{,}15}{6} = 0{,}81.$$

Коэффициент использования рабочего времени слесарем составит

$$\eta_0 = 1 - P_0 = 1 - 0{,}19 = 0{,}81.$$

Для определения максимального количества оборудования, обслуживаемого одним рабочим, при котором коэффициент его использования будет не ниже 0,9, последовательно определим показатели работы при 5 и менее станках.

Расчеты для случая, когда к одному рабочему прикреплены 3, 4 и 5 станков (n = 1, m = 3, 4, 5), приведены в табл. XI.8.

Таблица XI.8

k	a_k			P_k			$(k-n)P_k$		
	= 3	= 4	= 5	= 3	= 4	= 5	= 3	= 4	= 5
0	1	1	1	0,53	0,40	0,28	–	–	–
1	0,60	0,80	1	0,32	0,30	0,28	–	–	–
2	0,24	0,48	0,80	0,13	0,20	0,23	0,13	0,20	0,23
3	0,05	0,19	0,48	0,02	0,08	0,14	0,04	0,16	0,28
4	–	0,04	0,19	–	0,02	0,06	–	0,06	0,18
5	–	–	0,04	–	–	0,01	–	–	0,04

175—22

Таблица XI.9

k	a_k	P_k	$(k-2)P_k$
0	1	0,32	–
1	1,2	0,38	–
2	0,6	0,19	–
3	0,24	0,08	0,08
4	0,07	0,02	0,04
5	0,01	0,007	0,03
6	0,01	0,003	0,01

Коэффициент занятости рабочего соответственно равен $1 - P_0$: при обслуживании 3 единиц оборудования – 0,47, 4 – 0,6 и 5 – 0,72.

Соответственно, среднее количество единиц оборудования, ожидающего начала ремонта, составит $M_1^3 = 0,17$, $M_1^4 = 0,42$ и $M_1^5 = 0,73$.

Коэффициент использования оборудования составит: при обслуживании 3 единиц – 0,94, 4 – 0,89 и 5 – 0,85.

Следовательно, чтобы коэффициент использования оборудования превышал 0,9, к одному слесарю надо прикрепить не более 3 единиц оборудования.

Рассчитаем показатели использования оборудования, если на участке работают два слесаря, и оборудование персонально не закреплено ($n = 2$ и $m = 6$).

Результаты расчетов a_k, P_k и $(k-n)P_k$ приведены в табл. XI.9.

Среднее число простаивающих станков составит $\sum_{3}^{6} (k-2)P_k = 0,16$, а коэффициент использования оборудования 0,97. Сопоставляя данный показатель с коэффициентом использования оборудования при закреплении за каждым рабочим 3 единиц оборудования, наглядно убеждаемся в преимуществе бригадной организации обслуживания.

Контрольные вопросы и задачи

1. Какова структура систем массового обслуживания?
2. Приведите примеры систем массового обслуживания, встречавшихся Вам на горных предприятиях.
3. Какие основные характеристики имеют входящий поток требований и механизм обслуживания?
4. Назовите основные свойства пуассоновского потока требований.
5. Сущность задач массового обслуживания и порядок их решения.
6. Как классифицируются системы массового обслуживания?
7. Какие основные показатели характеризуют функционирование систем массового обслуживания с отказами?
8. Чем разомкнутые системы массового обслуживания отличаются от замкнутых?
9. Какие основные показатели характеризуют функционирование систем массового обслуживания с ожиданием?
10. Каким путем получают основные закономерности для расчета систем массового обслуживания?
11. Задача. Система массового обслуживания представляет собой экскаватор, обслуживающий поток автосамосвалов, поступающих под погрузку. Интенсивность потока автосамосвалов $\lambda = 0,3$ в 1 мин. Средняя продолжительность времени погрузки $t_п = 4$.

Таблица XI.10

№ варианта	Интервал между двумя разговорами, мин	Средняя продолжительность разговора, мин	Вероятность отказа
1	12,2	8,3	0,1
2	10,1	4,2	0,2
3	8,1	3,0	0,07
4	12,1	6,2	0,08
5	6,2	3,1	0,05
6	15,0	4,0	0,15
7	10,8	2,5	0,01
8	15,2	7,3	0,03
9	15,3	8,1	0,05
10	16,1	7,5	0,04
11	14,2	4,3	0,04
12	14,2	4,7	0,08
13	7,5	3,8	0,10
14	7,8	2,9	0,02
15	6,5	2,4	0,02
16	6,0	1,5	0,01
17	3,7	1,0	0,02
18	4,1	1,2	0,04

Определить: интенсивность обслуживания; относительную пропускную способность системы; вероятность отказов; производительность экскаватора.

12. Задача. Сколько нужно иметь каналов связи, чтобы вероятность отказа не превышала 3 %, если в среднем каждые 2,5 мин поступает запрос на разговор, а средняя продолжительность разговора – 3, 4 мин.

13. Задача. АТС горно-обогатительного комбината имеет три канала связи с городской АТС. В среднем каждые 3 мин поступает запрос на разговор, а средняя продолжительность одного разговора – 5 мин. Поток заявок подчинен пуассоновскому закону распределения, а время обслуживания – показательному.

Определить: абсолютную пропускную способность АТС; вероятность отказа в соединении; среднее количество занятых каналов.

14. Задача. Под экскаватор каждые 4 мин прибывает машина под погрузку. Определить вероятность отказа, если время погрузки 2 мин.

15. Задача. Горное предприятие имеет собственную АТС. Среднее время (интервал) между двумя разговорами с городскими абонентами составляет $t_и$ мин, а средняя продолжительность разговора – $t_р$ мин. Распределение длительности интервала и продолжительности разговора подчинены показательному закону. Определить зависимости основных показателей работы АТС от количества каналов связи между ней и городской АТС, а также необходимое количество каналов связи, чтобы вероятность отказа не превышала $P_{отк}$. Исходные данные по вариантам приведены в табл. XI.10.

16. Для условий предыдущей задачи определить рациональное количество каналов связи, если затраты на их эксплуатацию в k раз превышают экономический ущерб от 1 % потерь требований (отказа в соединении с абонентом) $k = 8, 10, 12, 15$.

17. Задача. Система массового обслуживания представляет собой n экскаваторов, которые обслуживают поток машин, поступающих под загрузку. Поток

Таблица XI.11

№ варианта	$t_{погр}$	$t_р$	n	η	p	q
1	4,0	50	2	0,7	4	2,7
2	5,0	52	4	0,8	5	3,0
3	3,8	46	3	0,6	6	2,1
4	4,0	48	5	0,6	4	3,4
5	4,5	40	1	0,7	3	2,0
6	4,2	56	3	0,8	5	3,5
7	3,7	52	2	0,7	4	2,7
8	4,6	40	1	0,6	4	3,0
9	3,9	60	5	0,6	5	3,2
10	5,0	30	1	0,8	3	1,6
11	3,0	40	2	0,6	4	1,9
12	4,0	50	1	0,7	3,5	2,5
13	3,6	55	4	0,6	4	3,1
14	4,2	60	2	0,7	5	3,5
15	3,5	54	6	0,8	4	2,7

требований может рассматриваться как пуассоновский, а время обслуживания (погрузки) распределяется по показательному закону. При этом время погрузки составляет $t_{погр}$ (мин), а средняя продолжительность рейса $t_р$ (мин). Требуется определить:

при каком минимальном количестве автомашин обеспечивается коэффициент использования экскаватора η?

каково рациональное количество автомашин для обслуживания экскаваторов, если соотношение между убытками от простоя экскаваторов и автосамосвалов равно p?

при каком количестве автомашин обеспечиваются минимальные удельные затраты на погрузку и транспорт горной массы, если соотношение между приведенными затратами на экскаватор и автосамосвал равно q?

Исходные данные для решения задачи даны в табл. XI.11.

Таблица XI.12

№ варианта	Число вагонов в среднем за сутки	Среднее время разгрузки вагона	Количество бригад		Время ожидания
			min	max	
1	16,0	4,0	1	3	1,2
2	18,1	3,2	2	4	0,6
3	7,5	2,8	3	4	0,5
4	9,2	2,8	3	4	1,0
5	6,4	2,5	2	2	0,8
6	3,5	3,2	1	3	0,7
7	7,6	4,8	1	4	1,0
8	12,1	4,1	2	5	0,8
9	11,0	1,5	3	5	0,6
10	10,1	1,8	2	4	0,8
11	13,2	2,4	1	3	0,9
12	8,3	3,1	1	2	1,0
13	5,5	5,2	1	3	0,9
14	18,1	2,3	2	4	1,2
15	13,6	4,1	2	4	1,4

18. Задача. На склад лесных материалов шахты в сутки прибывает в среднем A вагонов (поток пуассоновский). Среднее время разгрузки одного вагона – B часов (подчиняется показательному закону). Определить:

количество простаивающих вагонов и среднее время их простоя (в ожидании погрузки и всего), если их разгрузка осуществляется $C - D$ бригадами, а также рассчитать среднее время простоя бригад;

рациональное количество бригад, если затраты на содержание бригады в k раз превышают потери от простоя одного вагона, $k =$ 0,8; 1,5; 2,5;

сколько мест для вагонов, ожидающих начала разгрузки, необходимо иметь, чтобы время ожидания не превышало $t_{ож}$ (ч), если разгрузка осуществляется $C - D$ бригадами;

построить зависимости показателей систем массового обслуживания от количества обслуживающих каналов и мест в очереди.

Исходные данные приведены в табл. XI.12.

Глава XII

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ

§ 1. ЗАДАЧИ МОДЕЛИРОВАНИЯ ПРОЦЕССОВ И КЛАССИФИКАЦИЯ ТИПОВ ВЗАИМОДЕЙСТВИЯ МАШИН И МЕХАНИЗМОВ

В настоящее время все основные технологические процессы на горных предприятиях (подготовка горных пород к выемке, погрузка, транспорт и отвалообразование) механизированы.

К числу основных задач, решаемых при организации технологических процессов, относятся:

рациональный подбор комплекта оборудования;

увязка различных процессов во времени;

определение резервов между отдельными процессами и видами работ.

Работа машин и механизмов подвержена влиянию случайных факторов трех групп.

К первой группе относятся природные факторы (изменения горно-геологических условий, физико-механические свойства пород, климатические условия); ко второй – технические (аварии и поломки оборудования, частота и продолжительность которых определяются надежностью машин), к третьей – социально-психологические факторы.

При решении различных производственных задач надо учитывать влияние случайных факторов, что вызывает необходимость применения вероятностных методов исследования. В последние годы в проектировании, планировании и управлении находит широкое применение метод статистического моделирования, который позволяет учитывать стохастический характер производственного процесса, функциональные и обратные связи между машинами и механизмами, наличие резервирования.

Статистическое моделирование – один из основных методов имитационного моделирования, цель которого состоит в воспроизведении поведения анализируемой производственной системы на основе анализа наиболее существенных связей и взаимодействия отдельных ее элементов.

Перед моделированием производственных процессов необходимо всесторонне проанализировать характер структурных связей и взаимодействий между машинами и механизмами – структуру процессов.

Под термином "структура процесса" будем понимать определенное сочетание тех или иных машин или звеньев с учетом имеющихся связей между ними.

Основным и общим во всех работах, связанных со структурными исследованиями, является стремление очистить объект исследования от всего второстепенного и рассмотреть лишь наиболее характерные его признаки.

Структурный анализ позволяет изучать объект в его наиболее общем и чистом виде. Уделяя основное внимание выявлению взаимных связей между отдельными машинами или звеньями и тем самым выдвигая на первое место структуру процесса, а не состав его отдельных частей, получаем возможность единообразно исследовать и моделировать различные по содержанию процессы.

Анализ взаимодействий между отдельными звеньями (машинами и механизмами) производственного процесса позволил выделить следующие три основных их способа:

непосредственный – обусловливает жесткую зависимость между смежными звеньями комплекса. При непосредственном взаимодействии остановка любого из механизмов, увязанных в последовательной цепочке, вызывает остановку всего комплекса. Примером такого способа может служить взаимодействие роторного экскаватора с забойным конвейером, работа става конвейеров и др.;

взаимодействие через промежуточный склад – при таком способе взаимодействия снимается непосредственная зависимость между смежными звеньями комплекса, и отдельные звенья определенное время могут работать самостоятельно (на склад или со склада). Такой способ взаимодействия часто встречается при кратной перевалке вскрышных пород, когда два экскаватора связаны между собой через промежуточный отвал. К этому же способу следует отнести и взаимодействие буровзрывных и погрузочных работ и др.;

взаимодействие как процесс обслуживания – имеет некоторые черты непосредственного взаимодействия, однако принципиально отличается от него тем, что процесс взаимодействия прерывен и носит характер обслуживания отдельных требований, поступающих через случайные промежутки времени.

Приведенная классификация способов взаимодействия машин и механизмов позволяет осуществить системный подход к проблеме моделирования производственных процессов и свести многообразие методов расчета к сравнительно небольшому числу универсальных вычислительных схем (алгоритмов моделирования), отличающихся простотой и легко реализуемых на ЭВМ.

Поскольку, как уже упоминалось, функционирование производственных процессов в значительной степени определяется вероятностными параметрами, находящимися в сложной взаимосвязи, в качестве основного математического аппарата используется метод статистических испытаний (метод Монте-Карло) или аналитические методы теории массового обслуживания и теории надежности.

Использование аналитических методов накладывает ряд весьма жестких и существенных требований и ограничений на характер процесса: отсутствие или неучет влияния емкости промежуточных складов, определенный (чаще всего простейший) поток оборудования и требований на обслуживание, показательный закон распределения времени на обслуживание, что идеализирует процесс и ухудшает результаты расчетов.

Метод статистических испытаний заключается в моделировании на ЭВМ производственных процессов по тем же законам, что и в натуре, но в измененном масштабе времени.

Учитывая важность и все возрастающее применение имитационного моделирования для решения различных проектных, организационных и управленческих задач с целью упрощения реализации имитационных моделей на ЭВМ разработаны специальные языки программирования (CPSS, Simula и др.).

Обычно при использовании метода статистических испытаний работа машин и механизмов имитируется определенными блоками программ. Блоки алгоритмов, имитирующие на ЭВМ работу машин и механизмов по определенным законам, характерным для протекания данного процесса, представляют собой датчики (генераторы) случайных чисел. Наибольшее применение нашел программный метод получения случайных чисел, для реализации которого разработаны программы для любых законов распределения. В их основе лежит первоначальное получение случайных чисел с равномерным законом распределения и их последующее преобразование в распределение с требуемым законом и параметрами.

§ 2. МОДЕЛИРОВАНИЕ НЕПОСРЕДСТВЕННОГО ВЗАИМОДЕЙСТВИЯ МАШИН И МЕХАНИЗМОВ

Для построения модели работы комплекса машин и механизмов необходимо иметь информацию о поведении каждой машины в различных режимах, возникающих вследствие аварийных простоев или изменения внешних условий. Эту информацию получают либо путем проведения натурных наблюдений, либо в результате вспомогательного моделирования отдельных процессов.

Моделирование работы комплекса машин и механизмов включает следующие основные этапы:

изучение исследуемого процесса с целью получения информации о характере его протекания;

формализация процесса, определение совокупности состояний отдельных машин и механизмов, формулировка условий переходов из состояния в состояние;

выбор совокупности параметров, характеризующих процесс функционирования;

разработка алгоритмов, моделирующих процесс функционирования комплекса;

программирование, отладка и реализация программы на ЭВМ;

обработка и анализ результатов моделирования.

Для получения информации о работе каждого механизма проводят статистические наблюдения. При этом функционирование каждой машины представляет собой чередование времени работы $\xi_{раб}$ и ремонта $\xi_{рем}$ (рис. XII. 1, *а*). На основании статистических наблюдений строят функции плотности распределения вероятностей пребывания механизма в состоянии работы и ремонта (рис. XII. 1, *б* и *в*). Зная подобные характеристики для каждого элемента схемы, формализуют процесс ее функционирования. Для этого строят матрицу возможных состояний, граф переходов и матрицу переходов.

Рассмотрим процесс непосредственного взаимодействия двух элементов. Каждый элемент может находиться в рабочем состоянии (1) или ремонте (0). Возможные состояния системы элементов приведены в табл. XII. 1.

Очевидно, что одновременно оба элемента в состоянии ремонта находиться не могут. При поломке одного элемента другой останавливается и находится в состоянии ожидания окончания ремонта вышедшего из строя элемента.

Возможные траектории блуждания системы иллюстрируются графом смены состояний (рис XII. 2). Переход из одного состояния в другое возможен при определенных условиях, которые для рассматриваемой системы сведены в табл. XII. 2.

На основании графа смены состояний и условий перехода разрабатываем алгоритм статистического моделирования процесса, реализуя который на ЭВМ, получаем вероятностные характеристики функ-

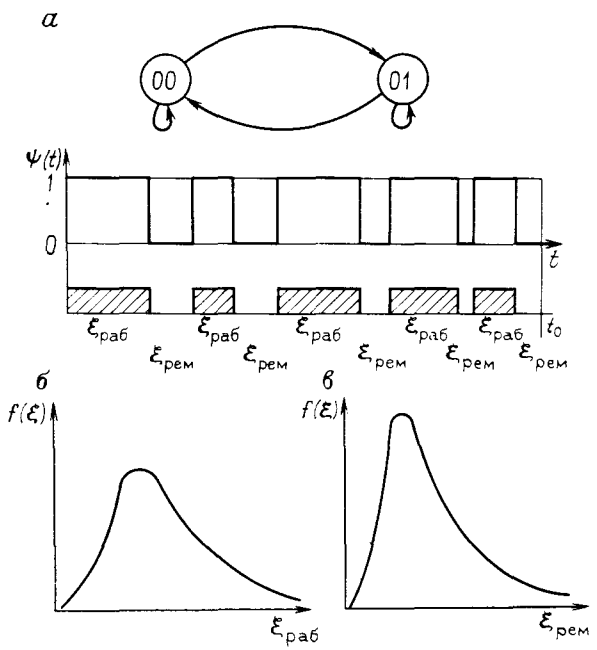


Рис. XII. 1. График работы оборудования (*а*), кривые плотности распределения времени работы (*б*) и ремонта оборудования (*в*)

ционирования схемы. Для случая непосредственного взаимодействия необходимо определить время эффективной работы и ремонта, схемы и функции плотности их распределения $f(t_{эф})$ и $f(t_{рем})$.

Блок-схема алгоритма моделирования непосредственного взаимодействия двух механизмов приведена на рис. XII. 3.

Время работы и ремонта элементов определяется по известной функции плотности распределения любым методом получения случайных чисел на ЭВМ.

Принцип работы алгоритма состоит в следующем.

В блоке 1 осуществляется засылка начальных условий работы системы (времени работы первого и второго элементов) в соответствую-

Таблица XII.1

Состояния			Номер состояния
первого элемента	второго элемента	системы	
1	1	1	1
1	0	0	2
0	1	0	3

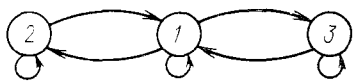


Рис.XII.2. Граф смены состояний системы

Таблица XII.2

Номер состояния системы	Условия перехода из одного состояния в другое		
	1	2	3
1	–	$\xi^1_{раб} > \xi^2_{раб}$	$\xi^1_{раб} < \xi^2_{раб}$
2	Безусловный	–	Нет
3	То же	Нет	–

щие блоки памяти ЭВМ (α и β). В блоке 2 производится операция сравнения и с учетом условий перехода определяется, в какое состояние и когда переходит система. В блоке 3 определяются исходные данные для нахождения суммарного времени пребывания системы в рабочем состоянии и плотности его распределения. В блоке 4 вычисляется текущее время работы системы. Блок 5 сравнивает текущее время функционирования схемы с заданным временем T_0 и при выполнении условия $T > T_0$ управление передается блоку 19 и процесс моделирования заканчивается обработкой данных и печатью результатов. Блоки 3–5 моделируют блуждание системы по траектории 1–2. В блоке 6 (датчике случайных чисел) производится определение времени ремонта второго элемента, в блоке 7 – вычисление текущего времени функционирования системы и в блоке 8 – его сравнение с заданным временем T_0. В блоках 9 и 10 определяется новое время работы второго и первого элементов и затем управление вновь передается блоку 2. Блоки 6–10 моделируют пребывание системы в состоянии 2 и ее переход в состояние 1. В целом левая ветвь алгоритма (блоки 3–10) моделирует блуждание системы по траектории 1–2 и обратно.

Правая ветвь алгоритма аналогична левой (так как граф переходов симметричен) и моделирует блуждание системы по траектории 1–3–1.

На печать выводятся суммарное время работы системы $t_{эф}$, а также плотности распределения времени пребывания системы в целом в рабочем и ремонтном состояниях, что вполне достаточно для оценки и анализа эффективности и ритмичности ее работы.

Описанный способ моделирования является асинхронным, текущее время смены изменяется скачкообразно, по мере смены состояния системы. Использование асинхронного способа позволяет сок-

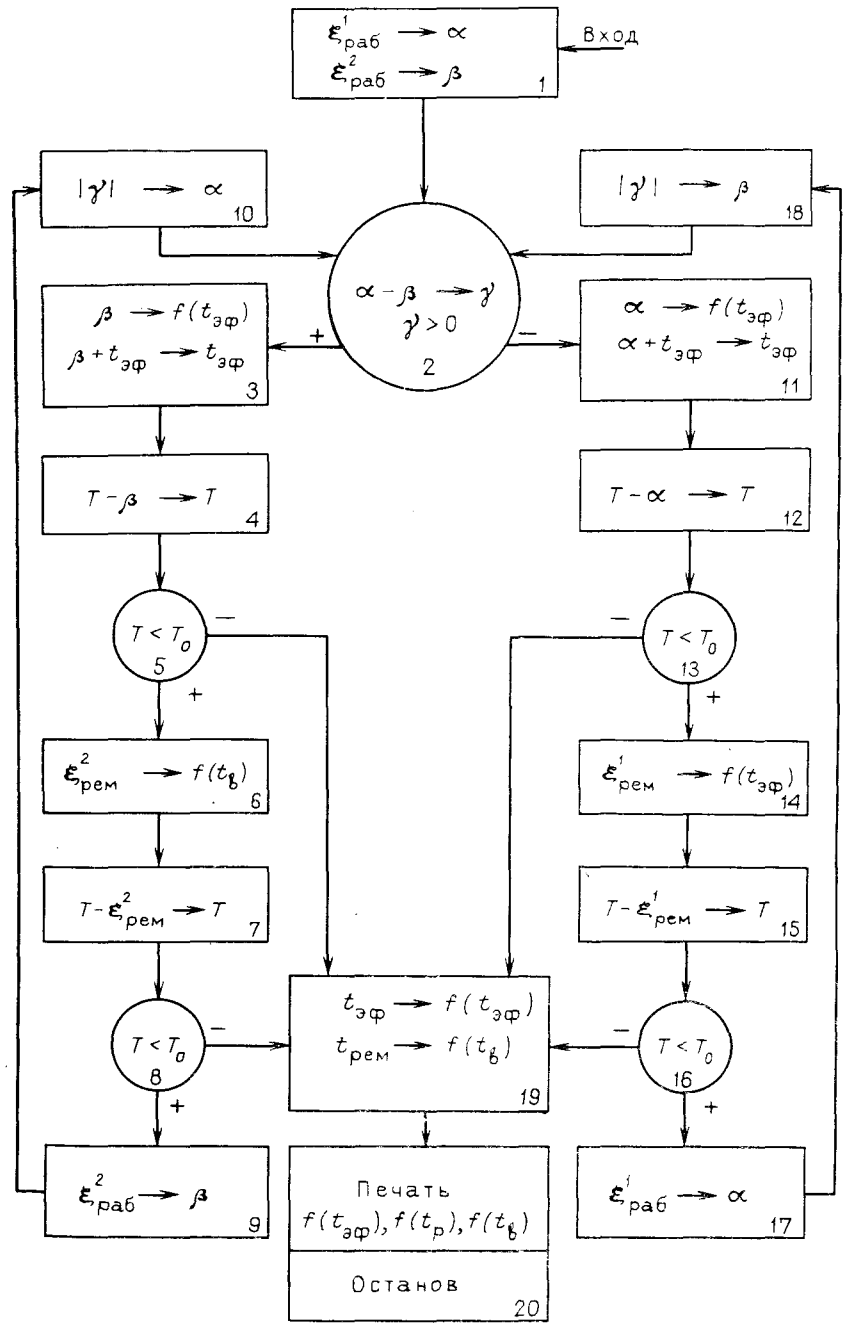


Рис.XII.3. Блок-схема алгоритма моделирования непосредственного взаимодействия двух механизмов

ратить время моделирования. При синхронном способе моделирования выбирают шаг изменения времени, и, проверяя последовательно в каждый момент времени в каком состоянии находятся отдельные элементы системы, определяют ее состояние в целом. Естественно, что этот способ моделирования более трудоемок на всех этапах.

§ 3. МОДЕЛИРОВАНИЕ ВЗАИМОДЕЙСТВИЯ ЧЕРЕЗ СКЛАД

Рассмотрим моделирование взаимодействия через склад на примере технологии горных работ с кратной перевалкой вскрыши в выработанное пространство.

Принципиальная схема комплексной механизации с кратной перевалкой вскрышных пород, получившей распространение на карьерах со значительными объемами вскрыши, приведена на рис. XII. 4.

Анализ совокупности операций при организации вскрышных работ по этой схеме показывает, что эффективность такого процесса может быть определена как эффективность функционирования Q_c некоторой стохастической системы G во времени, являющейся функциональной моделью реального процесса (рис. XII. 5).

Элементы системы $A_i (i = 1,2$ при однократной перевалке) имитируют процесс функционирования оборудования и в данный момент времени могут находиться в одном из трех возможных состояний:

$S_{A_i}(t) = 1$ – рабочее состояние;

$S_{A_i}(t) = 0$ – состояние ожидания;

$S_{A_i}(t) = -1$ – состояние ремонта.

Надежность элемента A_i определяют следующие параметры:
$f A_i^{(1)}(t)$ – плотность распределения времени пребывания в состоянии (1);
$f A_i^{(-1)}(t)$ – плотность распределения времени пребывания в состоянии (−1).

Параметр Q_{A_i} определяет пропускную способность элемента во время пребывания в состоянии (1).

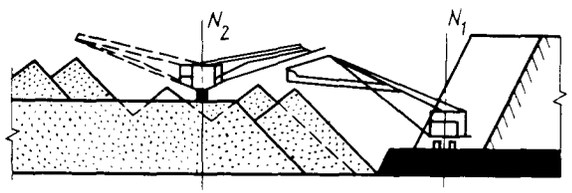


Рис. XII. 4. Схема разработки месторождения с кратной перевалкой вскрышных пород в выработанное пространство

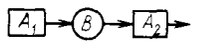


Рис. XII. 5. Графическое изображение системы, имитирующей процесс

Элемент B имитирует вместимость промежуточного склада породы, подлежащей переэкскавации.

Состояние склада $S_B(t)$ в данный момент времени t характеризуется текущим объемом породы, подлежащей переэкскавации V_t, причем $0 \leqslant V_t \leqslant V_{\max}$.

Параметром элемента B является $V_{\max}$ – максимально возможный объем породы в данных условиях производства.

Поток I_1 имитирует процесс поступления породы на склад для переэкскавации.

Поток I_2 – выходной случайный поток системы. Он имитирует выходной эффект рассматриваемой схемы.

В общем случае выходной эффект I_G системы является функцией, зависящей от времени t, начальных условий $S_{A_i}(t_0)$, $S_B(t_0)$ и параметров элементов: Q_{A_i}, $f_{A_i}^{(1)}(t)$, $V_{\max}$ и $f_{A_i}^{(-1)}(t)$, т.е.

$$I_G = F\left[S_{A_i}(t_0), S_B(t_0), Q_{A_i}, f_{A_i}^{(1)}(t), f_{A_i}^{(-1)}(t), V_{\max}, t\right]. \quad \text{(XII.1)}$$

Пропускная способность системы определяется по формуле

$$Q_G = I_G / T_0, \quad \text{(XII.2)}$$

где T_0 – время функционирования системы.

При рассматриваемой схеме максимальный выходной эффект системы за период времени T_0 может быть достигнут при условии $S_{A_i}(t_0) = (1)$,

$$\int_0^{T_0} f_{A_i}^{(-1)}(t)\, dt = 0. \quad \text{(XII.3)}$$

Это означает, что первый и второй экскаваторы к началу рассматриваемого периода находятся в рабочем состоянии и вероятность того, что какой-либо из них выйдет из строя на отрезке времени $[t_0, (t_0 + T)]$ равна нулю.

Тогда

$$\max I_G = \min\left\{Q_{A_i}\right\} T_0, \quad \text{(XII.4)}$$

а эффективность системы

$$Q_G = \min\left\{Q_{A_i}\right\}. \quad \text{(XII.5)}$$

При несоблюдении этих условий

$$Q_G < \min\left\{Q_{A_i}\right\}. \quad \text{(XII.6)}$$

В процессе функционирования происходит смена состояний элементов A_1 и A_2 системы через случайные промежутки времени $\Delta\tau$.

Случайное время пребывания элементов в состоянии (1) – $\tau_i^{(1)}$

Таблица XII. 3

Состояния			Номер состояния системы
первого элемента	второго элемента	склада	
1	1	$0 < V_t < V_{\max}$	1
1	−1	$0 < V_t < V_{\max}$	2
−1	1	$0 < V_t < V_{\max}$	3
−1	−1	$0 < V_t < V_{\max}$	4
−1	0	$V_t = 0$	5
0	−1	$V_t = \max$	6

или $(-1) - \tau_i^{(-1)}$ находится в соответствии с параметрами $f_{A_i}^{(1)}(t)$ и $f_{A_i}^{(-1)}(t)$, а случайное время пребывания в состоянии $(0) - \tau_i^{(0)}$ в соответствии с текущим объемом породы, подлежащей переэкскавации.

Система G может пребывать в шести состояниях, показанных в табл. XII. 3.

Возможные траектории блуждания системы во времени иллюстрируются графом переходов (рис. XII. 6).

Переход из одного состояния в другое обусловлен определенными условиями, которые для рассматриваемой системы сведены в матрицу и находятся на пересечении соответствующих строк и столбцов (табл. XII. 4).

С учетом этих условий перехода разработан алгоритм статистического моделирования процесса функционирования схемы, который при реализации его на ЭВМ позволит получить характеристики для оценки эффективности рассматриваемой схемы.

Блок-схема алгоритма приведена на рис. XII. 7. В блок-схеме приняты следующие обозначения: $\tau_1^{(1)}$ и $\tau_2^{(1)}$ — случайное время пребывания соответственно первого и второго экскаваторов в рабочем состоянии; $\tau_1^{(-1)}$ и $\tau_2^{(-1)}$ — случайное время пребывания соответственно первого и второго экскаваторов в состоянии ремонта; x — текущий объем породы вскрыши, подлежащей переэкскавации, отнесенной к производительности экскаватора, имеет размерность времени; y — текущий объем свободного пространства, отнесенный к производительности экскаватора, имеет размерность времени; α и β — содержимое блоков памяти ЭВМ, в которые по мере выработки случайных чисел засылаются данные о работе соответственно первого и второго экскаваторов; $t_1, t_2, \ldots, t_6$ — время пребывания системы соответственно в 1, 2, . . . , 6-м состояниях; $T_1, T_2, \ldots, T_6$ — суммарное время пребывания системы соответственно в 1, 2, . . . , 6-м состояниях; T и T_0 — текущее и заданное время функционирования системы.

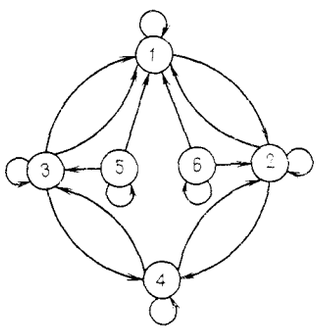


Рис. XII. 6. Граф переходов системы из одного состояния в другое

Таблица XII. 4

Номер состояния системы	1	2	3	4	5	6
1	–	$\tau_2^{(1)} \leqslant \tau_1^{(1)}$	$\tau_1^{(1)} \leqslant \tau_2^{(1)}$	Нет	Нет	Нет
2	$\tau_2^{(1)} \leqslant \tau_1^{(1)}$ $\tau_2^{(-2)} \leqslant \leqslant \frac{V_{\max} - V_t}{Q_{A_1}}$		Нет	$\tau_1^{(1)} \leqslant \tau_2^{(-1)}$ $\tau_1^{(1)} \leqslant \leqslant \frac{V_{\max} - V_t}{Q_{A_i}}$	Нет	$\min(\tau_2^{(-1)}, \tau_1^{(1)}) \geqslant \geqslant \frac{V_{\max} - V_t}{Q_{A_i}}$
3	$\tau_1^{(-1)} \leqslant \tau_2^{(1)}$ $\tau_1^{(-1)} \leqslant \frac{V_t}{Q_{A_2}}$	Нет	–	$\tau_2^{(1)} \leqslant \tau_1^{(-1)}$ $\tau_2^{(1)} \leqslant \frac{V_t}{Q_{A_2}}$	$\min(\tau_1^{(-1)}, \tau_2^{(1)}) \geqslant \frac{V_t}{Q_{A_2}}$	Нет
4	Нет	$\tau_1^{(-1)} \leqslant \tau_2^{(-1)}$	$\tau_2^{(-1)} \leqslant \tau_1^{(-1)}$		Нет	Нет
5	Безусловный переход	Нет	Нет	Нет	–	Нет
6	То же	Нет	Нет	Нет	Нет	–

В блоке 1 осуществляется ввод начальных условий. В соответствии с известными законами распределения вырабатывается время работы

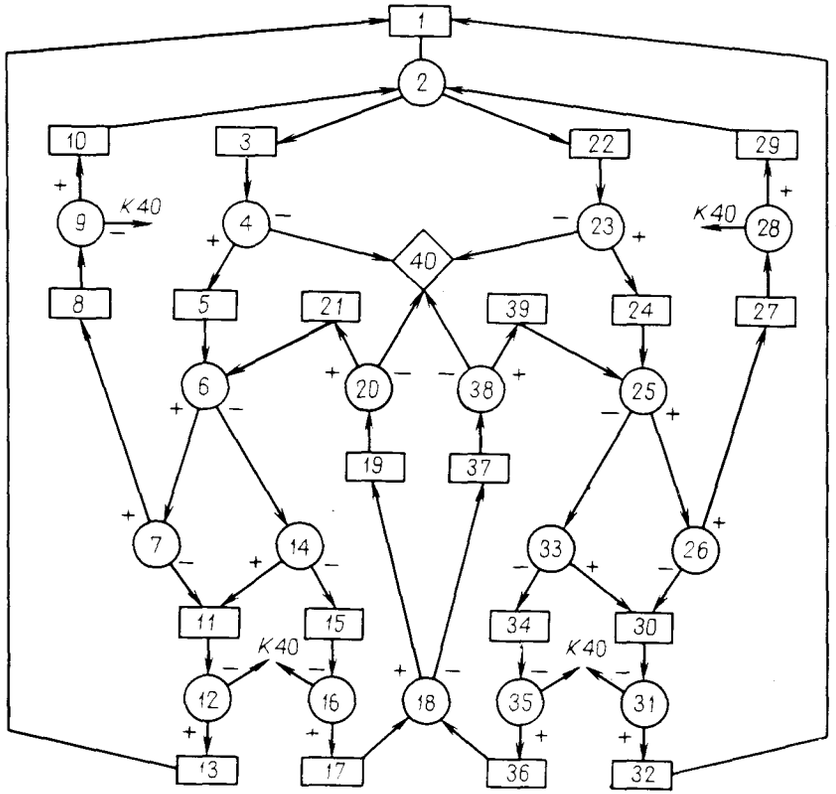


Рис.XII.7. Блок-схема алгоритма моделирования взаимодействия через склад

первого и второго экскаваторов ($\tau_1^{(1)}$ и $\tau_2^{(1)}$), которое засылается в соответствующие блоки памяти ЭВМ (α и β). Затем в блоке 2 производится сравнение и с учетом условий перехода определяется, в какое состояние (2 или 3) и когда переходит система. В результате этого находятся исходные данные для вычисления времени пребывания системы в состоянии 1 (t_1, T_1) и текущего времени функционирования системы T, которое определяется в блоке 3 (или 22). В блоке 4 сравнивается текущее время моделирования с заданным. В блоке 5 определяются параметры пребывания системы в состоянии 2, для этого вырабатывается время работы второго экскаватора $\tau_2^{(-1)}$ и рассчитывается время ремонта первого, которые засылаются соответственно в блоки памяти β и α. В блоках 6 и 7 (или 14) определяется траектория перехода системы из состояния 2. Так как система из состояния 2 может перейти в 3 (1, 4 или 5), для определения нового состояния требуется работа подряд двух логических блоков.

При выполнении условий перехода в блоках 6 и 7 система переходит из состояния 2 снова в 1; в блоке 8 определяется время пребывания системы в состоянии 2 (t_2, T_2) и текущее время функционирования. В блоке 9 снова проверяется условие окончания времени моделирования, а в блоке 10 определяются новые параметры пребывания системы в состоянии 1, для чего принимается время работы второго экскаватора $\tau_2^{(1)}$, а остальные параметры $\tau_1^{(1)}$, x и y рассчитываются. Найденные параметры засылаются в соответствующие блоки памяти ЭВМ (α, β, x и y).

При невыполнении условий перехода в блоке 7 система переходит в состояние 6. В блоке 11 определяются параметры пребывания системы в состояниях 2 и 6. В блоке 12 сравнивается текущее время работы схемы с заданным, а в блоке 13 определяются новые параметры пребывания системы в состоянии 1. Таким образом, блоки 2, 3, 4, 5, 6 и 7 соответствуют переходу системы из состояния 1 во 2, а 6, 7, 8, 9 и 10 – обратному переходу. Блоки 6, 7 и 11 соответствуют переходу системы из состояния 2 в 6, а блоки 12 и 13 – из состояния 6 в 1.

При невыполнении условий перехода в блоке 6 система переходит в состояние 4 или 6, что определяется в блоке 14. При выполнении условий перехода в блоке 14 система переходит в состояние 6 и управление передается блоку 11, а при невыполнении – блоку 15. В блоке 15 вычисляются параметры пребывания системы в состоянии 2 (t_2, T_2), текущее время работы (T), которое в блоке 16 сравнивается с заданным T_0. В блоке 17 определяются текущие параметры пребывания в состоянии 4 (оставшееся время ремонта второго экскаватора, время ремонта первого экскаватора и параметры склада), которые засылаются в соответствующие блоки памяти ЭВМ.

В блоке 18 определяется траектория блуждания системы, находящейся в состоянии 4, из которого система может вернуться в состояние 2 или перейти в 3. При выполнении условия перехода система возвращается в состояние 2 и в блоке 19 определяются параметры пребывания системы в состоянии 4 (t_4, T_4) и текущее время работы (T), которое в блоке 20 сравнивается с заданным (T_0). В блоке 21 определяются текущие параметры пребывания системы в состоянии 2, для чего моделируется время работы первого экскаватора $\tau_1^{(1)}$ и рассчитывается время ремонта второго $\tau_2^{(-1)}$.

Блоки 14, 15, 16 и 17 имитируют переход системы из 2 в 4 состояние, а 18, 19, 20 и 21 – обратный переход. Таким образом, блоки 1–21 моделируют блуждание и пребывание системы в 1, 2, 4 и 6 состояниях.

Правая ветвь алгоритма аналогична левой (так как граф перехода симметричен) и моделирует блуждание системы по траекториям 1–3 (блоки 22, 23, 24, 25, 26), 3–1 (блоки 25, 26, 27, 28, 29), 3–5 (блоки 25, 26, 30), 5–1 (блоки 31, 32), 3–4 (блоки 33, 34, 35, 36) и 4–3 (блоки 18, 37, 38, 39). Блок 40 по истечении заданного времени моделирования

$(T \geqslant T_0)$ переводит работу алгоритма на обработку результатов моделирования, останов машины и печать.

Описанный алгоритм представляет собой случай асинхронного статистического моделирования, т.е. время меняется одновременно со сменой состояния одного из элементов, при этом время всегда отсчитывается с момента смены состояния. В данном случае подобный способ позволяет ускорить процесс моделирования по сравнению с синхронным моделированием, при котором время меняется через определенные равные интервалы и в каждый момент времени необходимо просматривать состояния всех элементов системы.

Процесс моделирования включает четыре повторяющихся этапа:

определение параметров состояния системы (блоки 1, 5, 10, 13, 17, 21, 24, 29, 32, 36, 39);

определение траектории перехода системы (блоки 2, 6, 7, 14, 18, 25, 26, 33),

определение суммарного времени пребывания системы в определенном состоянии и текущего времени функционирования системы (блоки 3, 8, 11, 15, 19, 22, 27, 30, 34, 37);

сравнение текущего времени моделирования с заданным, т.е. проверка окончания времени моделирования (блоки 4, 9, 12, 16, 20, 23, 28, 31, 35, 38).

В результате моделирования получаем суммарное время (T_1, T_2, T_3, T_4, T_5, T_6) и функции распределения времени ($f(t_1)$, $f(t_2)$, $f(t_3)$, $f(t_4)$, $f(t_5)$, $f(t_6)$) пребывания системы в каждом состоянии. На основании этих функций определяем и оцениваем эффективность схемы, ритмичность процесса, влияние работы отдельных элементов на эффективность функционирования системы в целом и т.д.

При моделировании других случаев взаимодействия через склад /например, экскаваторных и буровзрывных работ, вскрышных и добычных работ, в качестве склада соответственно выступают запасы взорванной горной массы и вскрытые запасы полезного ископаемого/ алгоритм несколько изменяется и иногда усложняется.

§ 4. СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

Взаимодействие как процесс обслуживания (взаимодействие массового обслуживания) имеет широкое распространение на горных предприятиях, использующих цикличный транспорт. При этом осуществляется массовое обслуживание однородного потока требований. Например, при использовании на карьере железнодорожного и автомобильного транспорта выемочно-погрузочные работы, отвальные работы, работа карьерного транспорта могут интерпретироваться как процесс массового обслуживания, а работу всего карьера в целом можно

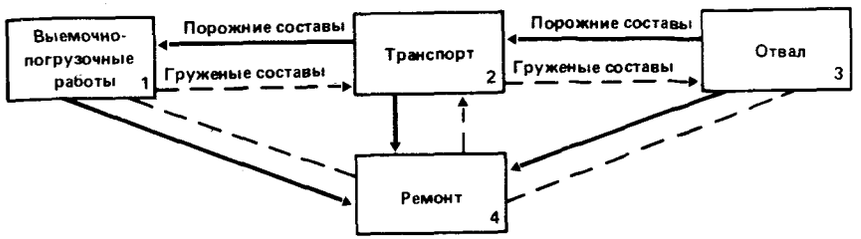


Рис. XII.8. Многофазная система массового обслуживания

моделировать как многофазную систему массового обслуживания (рис. XII.8), состоящую из подсистем: *1* – забой, *2* – транспорт, *3* – отвалы и *4* – ремонт.

В подсистеме "забой", имитирующей выемочно-погрузочные работы, требованиями (входящим потоком) являются порожние составы, поступающие из подсистемы "транспорт"; обслуживание заключается в их погрузке, а выходящим потоком являются груженые составы. В подсистеме "транспорт" требованиями на обслуживание являются составы (порожние и груженые), поступающие соответственно из подсистем "отвалы" и "забои". Обслуживание заключается в пропуске составов; в качестве обслуживающих аппаратов выступают железнодорожные пути и различные транспортные сооружения. Выходящим потоком требований являются порожние и груженые составы, которые поступают соответственно на вход подсистем "забои" и "отвалы". В целом подсистема "транспорт" может быть подразделена на две части, одна из которых обслуживает порожние, а другая – груженые составы.

В подсистеме *3*, имитирующей разгрузку поездов на отвале, входящий поток требований – груженые составы. Обслуживание заключается в их разгрузке, обслуживающие аппараты – отвальные тупики с оборудованием. Выходящий поток требований из данной подсистемы – порожние составы – поступает на один из входов подсистемы "транспорт".

Механизмы (требования или обслуживающие аппараты в любой подсистеме) могут выходить из строя и требовать ремонта. Подобные механизмы составляют входящий поток требований подсистемы "ремонт". Обслуживающими аппаратами здесь выступают бригады ремонтников с необходимым оборудованием (обслуживание заключается в ремонте механизмов), а выходящим потоком требований – исправные механизмы, которые вновь возвращаются в те подсистемы, откуда они поступили в ремонт.

Таким образом, в целом работа карьера представляет замкнутую систему: выходящий из одной системы поток является входящим для другой и т.д., что наглядно иллюстрируется рис. XII.8. Моделирование взаимодействия массового обслуживания может осуществляться аналитически или с помощью метода статистических испытаний.

Применение аналитических методов теории массового обслуживания для математического описания транспортных процессов всегда связано с принятием тех или иных допущений, например о пуассоновском характере потока требований, необходимых для применения определенных моделей.

Допустимость использования аналитических методов теории массового обслуживания в каждом конкретном случае требует особых доказательств. Более надежные результаты получаются при статистическом моделировании взаимодействия массового обслуживания.

Рассмотрим применение статистического моделирования для определения характеристик открытой многоканальной системы массового обслуживания с ожиданием, включающей забойные экскаваторы (обслуживающие аппараты) и локомотивосоставы (поток требований).

В обслуживающую систему, включающую n экскаваторов, в случайные моменты времени t_i поступают локомотивосоставы (требования на погрузку). Если в этот момент есть свободные обслуживающие аппараты, то состав становится под погрузку и занимает экскаватор на время

$$t_{зан} = t_{п} + t_{дв}, \qquad \text{(XII. 7)}$$

где $t_{п}$ и $t_{дв}$ – соответственно время погрузки и движения от распределительного устройства до экскаватора ($t_{п}$ и $t_{дв}$ – случайные величины, закон распределения которых устанавливается статистическими наблюдениями с последующей обработкой).

Если свободных экскаваторов нет, то состав становится в очередь. Экскаваторы периодически могут выходить из строя. Моменты выхода из строя экскаваторов и время их последующего ремонта являются случайными величинами, закон распределения которых известен.

Укрупненная блок-схема моделирующего алгоритма приведена на рис. XII.9. В начале моделирования все текущие параметры схемы равны нулю. Каждый оператор представляет собой подалгоритм, реализующий в процессе моделирования определенную операцию.

Приведем операторную запись моделирующего алгоритма

$$A_1\ P_2\ A_3\ A_4\ A_5\ P_6\ P_7\ A_8\ A_9\ P_{10}\ A_{11}\ A_{12}\ P_{13}\ A_{14}\ P_{15}\ A_{16}\ \Phi_{17}\ A_{18}\ A_{19}\ P_{20} \rightarrow$$
$$\rightarrow A_{21}\ A_{22}\ П_{23}\ Я_{24}.$$

(над операторами: 22, 18, 9, 17, 16, 17, 8; под операторами: 13, 10.15, 6)

Здесь и на блок-схеме (см. рис. XII.9) приняты следующие обозначения: A_1 – определение момента t_i поступления очередного требования в систему; P_2 – проверка принадлежности очередного состава к рассматриваемой смене (неравенство $t_i < T_0$, где T_0 – общее время смены). Если это условие выполнено, управление передается оператору A_3, если нет – оператору A_{22}; A_3 – счетчик общего числа

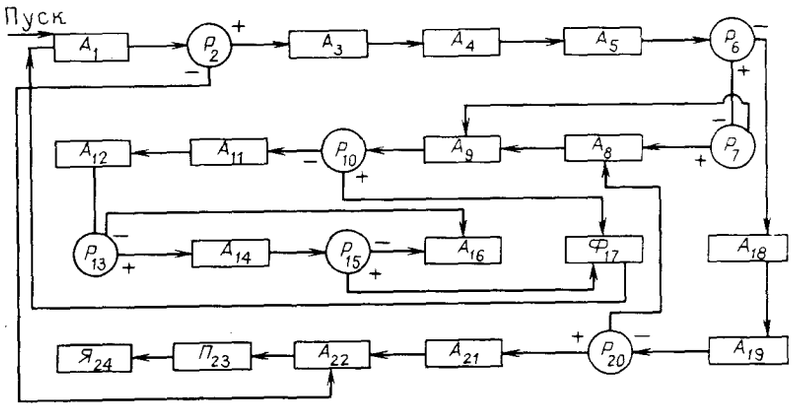


Рис. XII.9. Укрупненная блок-схема алгоритма, моделирующего работу транспорта

составов, поступивших за смену; A_4 – определение времени занятости экскаватора $t_{зан}$, обслуживающего предыдущие составы. Время занятости (как и моменты поступления составов, время ремонта и т.д.) является случайной величиной с известным законом распределения $A(x)$. Поэтому оператор A_4 преобразует случайные числа, имеющие равномерное распределение в интервале (0,1) с целью получения случайных чисел, подчиняющихся закону распределения $A(x)$. Для этого используются датчики или специальные методы получения псевдослучайных чисел; A_5 – определение времени освобождения экскаватора $t_{осв}$, для чего ко времени начала обслуживания состава прибавляется полученное оператором A_4 значение $t_{зан}$; значение фиксируется в специальных блоках программы ЭВМ; P_6 – проверка наличия свободных экскаваторов ($n_{св} > 0$). Значение момента поступления состава t_i сравнивается с $t_{осв}$ для всех экскаваторов. Если свободных экскаваторов нет, состав должен встать в очередь и управление в этом случае передается оператору A_{18}; P_7 – проверка числа свободных экскаваторов. Если имеется только один свободный экскаватор, управление передается оператору A_9; A_8 – составление перечня свободных экскаваторов и выработка условий для реализации правил распределения и очередности погрузки свободных экскаваторов; A_9 – выбор одного из свободных экскаваторов в соответствии с правилами приоритетов. Приоритет может даваться либо по мере освобождения экскаваторов, либо планом горных работ (когда диктуется первоочередная отгрузка из ряда забоев); P_{10} – проверка условий исправной работы экскаватора. Вероятностный закон выхода экскаватора из строя известен. Если t^* – момент выхода экскаватора из строя, то, зная момент окончания обслуживания $t_к = t_{осв} + t_{зан}$, можно опре-

делить, произойдет ли срыв в обслуживании $t^* < t_к$. Если экскаватор исправен и может обслуживать составы, то осуществляется переход к оператору $Ф_{17}$; A_{11} – определение времени ремонта $t_{рем}$ экскаватора, вышедшего из строя; A_{12} – счетчик числа и времени ремонтов; P_{13} – определение дальнейшего положения состава в случае выхода экскаватора из строя. Для этого время ремонта сравнивается с определенной величиной t_0, если $t_{рем} > t_0$, то состав уходит от экскаватора не полностью обслуженным, и управление передается оператору A_{16}; если $t_0 > t_{рем}$, то состав ждет конца ремонта и погружается; A_{14} – определение времени ремонта и последующего дообслуживания состава; P_{15} – проверка возможности погрузки состава после окончания ремонта до конца смены ($t_{осв} + t_{зан} + t_{рем}$); если состав будет загружен до конца смены, управление передается оператору $Ф_{17}$; A_{16} – счетчик не полностью обслуженных составов; $Ф_{17}$ – оператор подготовки алгоритма к моделированию процесса обслуживания следующего состава; A_{18} – определение времени ожидания состава в очереди $t_{ож} = t_{осв} - t_i$; A_{19} – счетчик общего времени простоя составов в очереди ожидания погрузки; P_{20} – проверка возможности обслуживания стоящего в очереди состава до конца смены; если состав может быть обслужен, то управление передается оператору A_8; A_{21} – счетчик числа необслуженных составов, A_{22} – статистическая обработка результатов моделирования; $П_{23}$ – выдача результатов; $Я_{24}$ – окончание процесса моделирования.

В результате реализации алгоритма выдаются на печать следующие характеристики процесса:

среднее значение и распределение числа составов в любой момент времени;

среднее значение и распределение продолжительности загрузки экскаваторов;

количество составов, поступивших на погрузку за смену;

количество и время ремонтов экскаваторов;

количество не полностью обслуженных составов;

количество полностью не обслуженных составов;

количество составов, загруженных каждым экскаватором.

Получив эти характеристики и зная экономические показатели, можно найти оптимальное число составов, распределение экскаваторов и локомотивосоставов по участкам работ, оптимальное размещение обменных пунктов и т.д.

Метод статистических испытаний позволяет моделировать и более сложные системы массового обслуживания (работу в переходном режиме, многофазовый процесс обслуживания, системы с неоднородным потоком требований и т.д.).

Контрольные вопросы и задачи

1. Для решения каких задач используется имитационное моделирование?
2. Какие случайные факторы функционирования производственных систем предопределяют целесообразность использования статистического моделирования?
3. В чем сущность статистического моделирования?
4. Назовите основные способы взаимодействия машин и механизмов.
5. Какой из способов взаимодействия наиболее жесткий?
6. Приведите примеры различных способов взаимодействия для условий добычи руды подземным способом.
7. Какие способы взаимодействия встречаются в следующих технологических цепочках:

уголь, отбитый комбайном, поступает на скребковый конвейер, далее через небольшой бункер загружается в вагонетки и поступает в околоствольный двор. Вагонетки скиповым подъемником выдаются на поверхность, где уголь разгружается и подается в бункер склада;

вскрышная порода экскаваторами грузится в автосамосвалы и вывозится на перегрузку, где разгружается. Далее порода грузится экскаватором в железнодорожные составы и вывозится в отвал;

на руднике используется самоходное оборудование (погрузочно-доставочные машины и самосвалы).

8. Приведите примеры различных способов взаимодействия, которые Вы наблюдали в период ознакомительных практик.
9. Какие основные этапы включает процесс статистического моделирования?
10. В чем разница между синхронным и асинхронным способами моделирования?
11. В чем преимущества асинхронного способа моделирования?
12. Как осуществляется моделирование случайных величин, характеризующих различные параметры производственных процессов?
13. Составьте алгоритм моделирования конвейерной линии, включающей три конвейера.
14. Составьте алгоритм моделирования работы экскаватора и трех автосамосвалов.
15. В чем особенность моделирования взаимодействия между буровзрывными и погрузочными работами?
16. В каких случаях системы массового обслуживания являются многофазовыми?

Глава XIII

ПРИНЯТИЕ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

§ 1. ЭЛЕМЕНТЫ ТЕОРИИ СТАТИСТИЧЕСКИХ РЕШЕНИЙ

При решении многих горно-экономических задач, задач планирования и управления горным производством, проблема принятия решений резко усложняется из-за различного вида случайных факторов, к которым чаще всего относятся условия проведения операции (климат, горно-геологическая характеристика месторождения, надежность оборудования, опыт и квалификация персонала и т.д.). Кроме того, операции часто являются многоцелевыми, при этом возникает вопрос: какому из критериев отдать предпочтение (обычно для разных критериев различно и решение). Во всех этих случаях приходится принимать решение в условиях неопределенности, возникающей из-за недостатка (отсутствия) информации (либо об условиях проведения операции, либо ее целях). Естественно, в этих случаях принятие решения более сложно, и точные математические методы не всегда дают однозначный результат. Однако и в этих условиях использование методов экономико-математического моделирования позволяет глубже разобраться в задаче, свести к минимуму элементы риска и волюнтаризма.

Задачи обоснования решений в условиях неопределенности изучаются теорией игр и статистических решений. Причем теория игр используется для анализа конфликтных ситуаций, в которых противодействуют (обычно активно) различные стороны, а теория статистических решений применяется в ситуациях, когда неопределенность рождена условиями задачи. В этих задачах нет активного противника, противодействующего нашим планам. Его роль выполняет природа, являющаяся условным противником, поведение которого неизвестно, хотя элемент противодействия отсутствует. Подобные ситуации называются "играми с природой".

Рассмотрим основную задачу теории статистических решений в общем виде и затем на конкретном примере.

Имеется несколько вариантов решения какой-либо задачи (A_1, A_2, . . . , A_m). Эффективность каждого варианта, помимо известных нам факторов, определяется также рядом факторов (условий), точное значение которых неизвестно (климат, геология месторождения, спрос на продукцию и цены на оборудование в перспективе и т.д.). Определенным предположениям о значениях (состояниях) случайных факторов (состояния природы $П_1$, $П_2$, . . . , $П_n$) соответствуют различные показатели вариантов решения задачи. Обозначим через a_{ij} показатель решения задачи при использовании варианта i и состояния

Таблица XIII. 1

Стратегия игрока	Состояния природы					
	$П_1$	$П_2$	. . .	$П_j$	. . .	$П_n$
A_1	a_{11}	a_{12}	. . .	a_{1j}	. . .	a_{1n}
A_2	a_{21}	a_{22}	. . .	a_{2j}	. . .	a_{2n}
. . .	. . .	. . .	. . .	. . .	. . .	. . .
A_i	a_{i1}	a_{i2}		a_{ij}		a_{in}
. . .	. . .	. . .	. . .	. . .	. . .	. . .
A_m	a_{m}	a_{m2}	. . .	a_{mj}	. . .	a_{mn}

природы j. Показатели для различных вариантов решений A_i при возможных состояниях природы $П_j$ задаются матрицей $\|a_{ij}\|$ (табл. XIII. 1).

Требуется найти решение задачи, т.е. такую стратегию A_i, которая более предпочтительна по сравнению с остальными. Под стратегией в теории игр и статистических решений принимается совокупность правил, определяющих выбор варианта действий; принятая стратегия однозначно определяет вариант.

Матрица $\|a_{ij}\|$ называется платежной, или матрицей выигрышей. Часто для решения задачи используют матрицу рисков (сожалений) $\|r_{ij}\|$, которая может дать более наглядную картину для оценки вариантов действий. Риск r_{ij} представляет собой разность между максимальным выигрышем при определенном состоянии природы (т.е. выигрышем, который был бы получен, если было бы известно состояние природы $П_j$) и выигрышем, полученным при использовании стратегии A_i: $r_{ij} = \beta_j - a_{ij}$, где $\beta_j = \max_i a_{ij}$.

Пример XIII.1. Составить матрицу рисков для платежной матрицы, приведенной в табл. XIII. 2.

Максимально возможные выигрыши составляют: при первом состоянии природы $\beta_1 = 12$, второй – $\beta_2 = 11$, третьем – $\beta_3 = 11$, четвертом – $\beta_4 = 9$.

Отнимая от этих значений последовательно (по столбцам) элементы платежной матрицы, получим матрицу рисков.

Например, $r_{32} = 11 - 9 = 2$.

Матрица рисков приведена в табл. XIII. 3.

Выбор решения начинают с сопоставления стратегий. При этом проверяется, не имеется ли стратегий лучших при любых состояниях природы (доминирующих).

В рассматриваемом примере (см. табл. XIII.3), сопоставляя стратегии A_3 и A_1, можно заметить, что выигрыши при A_3 больше (при $П_2$ и $П_3$ состояниях

Таблица XIII.2

Стратегия игрока	Состояния природы			
	$П_1$	$П_2$	$П_3$	$П_4$
A_1	10	8	7	6
A_2	12	7	4	9
A_3	10	9	8	6
A_4	9	10	11	5
A_5	7	11	8	8

Таблица XIII.3

Стратегия игрока	Состояния природы			
	$П_1$	$П_2$	$П_3$	$П_4$
A_1	2	3	4	3
A_2	0	4	7	0
A_3	2	2	3	3
A_4	3	1	0	4
A_5	5	0	3	1

природы) или равны (при $П_1$ и $П_4$) выигрышам при $П_1$ состоянии природы. Значит, использование стратегий A_3 лучше (или иногда одинаково, но всегда не хуже), чем A_1. В этом случае говорят, что третья стратегия доминирует над первой, которую вообще можно исключить из рассмотрения.

Возможны случаи, когда одна стратегия доминирует над всеми, тогда принятие решения тривиально. Если доминирующие стратегии отсутствуют, то в зависимости от состояния природы (которое нам не известно) эффективны и различные варианты решений. Например, при первом состоянии природы эффективен второй вариант, при втором состоянии – пятый и т.д.

В подобных случаях для принятия решения используют различные критерии оптимальности.

Наиболее просто решается задача, если имеется информация о вероятностях состояния природы (P_j, причем $\sum_{j}^{m} P_j = 1$). В этом случае в качестве критерия используется математическое ожидание выигрыша (или риска), т.е. выбирается решение, при котором

$$\sum_{j=1}^{m} P_j \, a_{ij} \to \max \ (\text{или} \ \sum_{j=1}^{m} P_j \, r_j \to \min) \ (i = 1, 2, \ldots, n). \qquad \text{(XIII.1)}$$

В такой постановке задача принятия решения в условиях неопределенности сводится к задаче принятия решений в условиях риска. Принятое решение оптимально при многократном повторении операции, т.е. "на круг", в среднем.

Пример XIII.2. Найти оптимальное решение задачи, платежная матрица которой характеризуется табл. XIII.2, если вероятности состояния природы составляют: $P_1 = 0{,}2$; $P_2 = 0{,}4$; $P_3 = 0{,}3$; $P_4 = 0{,}1$.

Подсчитаем среднее значение выигрыша, при этом первую стратегию не рассматриваем, так как она, как ранее установлено, хуже третьей

$$T_2 = 12 \cdot 0{,}2 + 7 \cdot 0{,}4 + 4 \cdot 0{,}3 + 9 \cdot 0{,}1 = 7{,}3,$$
$$T_3 = 10 \cdot 0{,}2 + 9 \cdot 0{,}4 + 8 \cdot 0{,}3 + 6 \cdot 0{,}1 = 8{,}6,$$
$$T_4 = 9 \cdot 0{,}2 + 10 \cdot 0{,}4 + 11 \cdot 0{,}3 + 5 \cdot 0{,}1 = 9{,}6,$$
$$T_5 = 7 \cdot 0{,}2 + 11 \cdot 0{,}4 + 8 \cdot 0{,}3 + 8 \cdot 0{,}1 = 9{,}0.$$

Итак, при принятом распределении вероятностей лучшей является четвертая стратегия: при ее использовании математическое ожидание выигрыша максимально.

В случаях, когда вероятности состояний природы неизвестны и их нельзя получить с достаточной степенью точности, используют критерий Лапласа, базирующийся на принципе недостаточного основания. При этом состояния природы считаются равновероятными, т.е.

$$P_1 = P_2 = P_3 = P_4.$$

В рассматриваемом примере критерий Лапласа для различных стратегий принимает значения

$$Л_2 = 12 \cdot 0{,}25 + 7 \cdot 0{,}25 + 4 \cdot 0{,}25 + 9 \cdot 0{,}25 = 8{,}0,$$
$$Л_3 = 10 \cdot 0{,}25 + 9 \cdot 0{,}25 + 9 \cdot 0{,}25 + 6 \cdot 0{,}25 = 8{,}25,$$
$$Л_4 = 9 \cdot 0{,}25 + 10 \cdot 0{,}25 + 11 \cdot 0{,}25 + 5 \cdot 0{,}25 = 8{,}75,$$
$$Л_5 = 7 \cdot 0{,}25 + 11 \cdot 0{,}25 + 8 \cdot 0{,}25 + 8 \cdot 0{,}25 = 8{,}5.$$

Итак, по критерию Лапласа в рассматриваемом примере лучшая – стратегия A_4.

Помимо критерия Лапласа используют и другие подходы к определению вероятностей состояния природы, в частности – экспертный, широко применяемый в прогнозировании. В последнее время для решения динамических задач планирования и управления все более часто используется байесовский подход (критерий Байеса), основанный на последовательном пересчете вероятностей состояния природы (апостериорных вероятностей) в зависимости от прошлых (или принятых ранее) состояний (априорных вероятностей).

Во всех случаях оценки вероятностей состояния природы решение является оптимальным только относительно принятого распределения вероятностей состояний.

Существуют и другие подходы и критерии к принятию решений в условиях неопределенности, используемые, когда нельзя получить распределение вероятностей состояний природы.

Наиболее широко распространены критерии Вальда, Сэвиджа и Гурвица.

При использовании **максиминного критерия Вальда** для каждой стратегии находят минимальное значение выигрыша, соответствующее наихудшему для нас в данном случае состоянию природы, т.е. $\min a_j$. Далее из всех возможных стратегий выбираем ту, для которой минимальный выигрыш максимален

$$B = \max_i \min_j a_{ij}. \qquad \text{(XIII. 2)}$$

Критерий Вальда является пессимистическим, при его использовании ориентируются на наихудшее для нас состояние природы, т.е. по существу природа рассматривается как активно противоборствующий противник.

Другая разновидность пессимистического подхода – использование **критерия Сэвиджа.** В этом случае находят минимальное значение риска при самом неблагоприятном состоянии природы

$$C = \min_i \max_j r_{ij}. \qquad \text{(XIII. 3)}$$

С этой целью для каждой стратегии (построчно) по матрице рисков находят максимальные значения риска, а затем выбирают из них минимальное.

Критерий Гурвица является комбинированным, учитывающим как оптимистический, так и пессимистический подходы. При использовании этого критерия состояние природы берется не самым худшим и не самым лучшим, а некоторое промежуточное. При этом за оптимальную принимается стратегия, при которой

$$\Gamma = k \min_j a_{ij} + (1-k) \max_j a_{ij} \to \max, \qquad \text{(XIII. 4)}$$

где k – коэффициент, характеризующий долю пессимизма и оптимизма (изменяется от 0 до 1).

Коэффициент k выбирается по субъективным соображениям: чем более сложнее ситуация и необходимо застраховаться, тем k ближе к единице. При $k=1$ критерий Гурвица преобразуется в критерий Вальда.

Критерии Вальда и Сэвиджа используют при принятии разовых и ответственных решений, а Гурвица, Лапласа и Байеса – при менее ответственных, когда ситуация (задача) повторяется многократно (например, при оперативном планировании).

Поясним использование различных критериев принятия решения в условиях неопределенности на примере.

Пример XIII.3. Рудное месторождение разведано редкой сетью скважин (в основном по категории C_1). В связи с дефицитностью сырья необходимо принять решение о мощности рудника, не ожидая окончания детальной разведки. Разведанные запасы месторождения (точнее их математическое ожидание) составляют 40 млн. т. Так как точность подсчета запасов по категории C_1 составляет ±100 %, реально запасы могут изменяться от 20 до 80 млн. т. Рассматриваются возможные варианты запасов 20, 30, 40, 60 и 80 млн. т (соответственно 1 – 5 состояния природы) и возможности строительства рудника мощностью 2, 3, 4 или 5 млн. т (соответственно 1 – 4 стратегии). Для каждого варианта мощности при рассматриваемых состояниях природы (вариантах запасов месторождения) подсчитаны (с учетом фактора времени) возможные значения суммарной приведенной прибыли (табл. XIII. 4).

Таблица XIII. 4

Вариант мощности (стратегия)	Прибыль, млн. руб., для вариантов запасов (состояний природы), млн. т				
	20 (1)	20 (2)	40 (3)	60 (4)	80 (5)
2 (1)	– 10	50	65	70	72
3 (2)	– 40	– 20	80	100	105
4 (3)	– 65	– 45	55	120	150
5 (4)	– 85	– 65	35	130	165

Отрицательное значение прибыли, наблюдаемое в ряде случаев, показывает, что в связи с неподтверждением запасов и большими капиталовложениями эксплуатация месторождения убыточна.

Для принятия окончательного решения о мощности рудника требуется рассчитать критерии Вальда, Сэвиджа, Гурвица, Лапласа и математическое ожидание прибыли, задаваясь вероятностями состояний природы (по аналогии с другими месторождениями).

1. Критерий Вальда. В каждой строке берем минимальную прибыль. Во всех случаях в данном примере она соответствует состоянию природы 1 (т.е. когда запасы не подтверждаются максимально). По критерию Вальда наилучшая стратегия 1, при этом потери минимальны (– 10 млн. руб.).

2. Критерий Сэвиджа. Построим матрицу рисков. Максимальное значение прибыли для различных состояний природы (по столбцам) составляет – 10, 50, 80, 130 и 165 млн. руб. Отнимая от них значения прибыли для различных стратегий, получим матрицу рисков (табл. XIII. 5).

В дополнительном столбце матрицы рисков показывается максимальное значение риска (потеря) для каждой стратегии. Минимальное значение риска (70) наблюдается при использовании стратегии 2. В этом случае (при строительстве рудника мощностью 3 млн. т) потери прибыли из-за недостатка информации не превысят 70 млн. руб. Итак, по критерию Сэвиджа наилучшая – стратегия 2.

Таблица XIII. 5

Стратегия	Риск, млн. руб., для состояний природы					Максимальная потеря, млн. руб.
	1	2	3	4	5	
1	0	0	15	60	93	93
2	30	70	0	30	50	70
3	35	95	25	10	15	95
4	75	115	45	0	0	115

3. Критерий Гурвица. Допустим, что уменьшение и увеличение запасов равновероятно, т.е. $k = 0{,}5$. Максимальное значение выигрыша в данной задаче во всех случаях соответствует максимальному количеству запасов (состояние природы 5) и соответственно для различных стратегий равно 72, 105, 150 и 165 млн. руб.

Тогда значение критерия Гурвица для различных стратегий равно

$$\Gamma_1 = -10 \cdot 0{,}5 + 72 \cdot 0{,}5 = 31{,}0,$$
$$\Gamma_2 = -40 \cdot 0{,}5 + 105 \cdot 0{,}5 = 32{,}5$$
$$\Gamma_3 = -65 \cdot 0{,}5 + 150 \cdot 0{,}5 = 42{,}5,$$
$$\Gamma_4 = 285 \cdot 0{,}5 + 165 \cdot 0.5 = 40{,}0.$$

По критерию Гурвица (при $k = 0{,}5$) наилучшая – стратегия 3, т.е. строительство рудника мощностью 4 млн. т.

4. Критерий Лапласа. При равной вероятности различных состояний природы ($P_1 = P_2 = P_3 = P_4 = P_5 = 0{,}2$) математическое ожидание прибыли при различных вариантах мощности рудника составит

$$Л_1 = -10 \cdot 0{,}2 + 50 \cdot 0{,}2 + 65 \cdot 0{,}2 + 70 \cdot 0{,}2 + 72 \cdot 0{,}2 = 49{,}4$$
$$Л_2 = -40 \cdot 0{,}2 + (-20)\,0{,}2 + 80 \cdot 0{,}2 + 100 \cdot 0{,}2 + 105 \cdot 0{,}2 = 44{,}5$$
$$Л_3 = -65 \cdot 0{,}2 + (-45)\,0{,}2 + 55 \cdot 0{,}2 + 120 \cdot 0{,}2 + 150 \cdot 0{,}2 = 43{,}0,$$
$$Л_4 = -85 \cdot 0{,}2 + (-65)\,0{,}2 + 35 \cdot 0{,}2 + 130 \cdot 0{,}2 + 165 \cdot 0{,}2 = 36{,}0.$$

Лучшая по критерию Лапласа – стратегия 1.

5. По аналогии с другими месторождениями (обычно наблюдается нормальный закон распределения ошибки подсчета запасов) примем вероятности состояния равными соответственно 0,12; 0,25; 0,3; 0,25; 0,08. Тогда математическое ожидание прибыли при различных вариантах мощности рудника составит

$$T_1 = -10 \cdot 0{,}12 + 50 \cdot 0{,}26 + 65 \cdot 0{,}3 + 70 \cdot 0{,}25 + 72 \cdot 0{,}08 = 54{,}06;$$
$$T_2 = -40 \cdot 0{,}12 + (-20)\,0{,}25 + 80 \cdot 0{,}3 + 100 \cdot 0{,}25 + 105 \cdot 0{,}08 = 47{,}4;$$
$$T_3 = -65 \cdot 0{,}12 + (-45)\,0{,}25 + 55 \cdot 0{,}3 + 120 \cdot 0{,}25 + 150 \cdot 0{,}08 = 39{,}45;$$
$$T_4 = -80 \cdot 0{,}12 + (-65)\,0{,}25 + 35 \cdot 0{,}3 + 130 \cdot 0{,}25 + 165 \cdot 0{,}08 = 29{,}75.$$

При принятом распределении вероятностей состояний природы лучшая стратегия 1.

Таким образом, по критериям Вальда, Лапласа и при принятом распределении вероятностей состояний природы лучшая – стратегия 1 (вариант строительства рудника мощностью 2 млн. т), по критерию Сэвиджа – 2, а по критерию Гурвица – 3.

Если учитывать, что принимается ответственное, разовое решение, а также ограниченность капитальных вложений и возможность дальнейшего поэтапного наращивания мощности рудника (с уточнением запасов), то для рассматриваемой задачи предпочтение в целом следует, видимо, отдать стратегии 1.

§ 2. ВЫБОР КРИТЕРИЯ ПРИНЯТИЯ РЕШЕНИЙ И ОПРЕДЕЛЕНИЕ РАЦИОНАЛЬНОЙ ТОЧНОСТИ ИСХОДНОЙ ИНФОРМАЦИИ

Как видно из рассмотренных примеров, решение, принимаемое в условиях неопределенности, существенно зависит от выбранного критерия. Кроме того, по сравнению с детерминированным подходом резко возрастает трудоемкость планирования главным образом

за счет многочисленных расчетов, необходимых для построения платежной матрицы. Поэтому весьма важно определить целесообразность использования вероятностного подхода (теории статистических решений) и область применения различных критериев оптимальности.

В Московском горном институте для задач текущего планирования открытых горных работ (месячного, квартального и годового) определены области использования различных критериев оптимальности. Для этого множество вариантов планов (стратегий) оценивалось по различным критериям оптимальности в зависимости от ошибки определения исходной информации (объема запасов и показателей качества полезного ископаемого).

Анализ расчетов показал, что при небольших ошибках исходной информации ($\delta \leqslant \pm 4$ %) оценка всех плановариантов по различным критериям совпадает и соответствует детерминированной. Поэтому при месячном планировании, когда ошибки информации невелики (не превышают ± 4 %), целесообразно использовать детерминированную постановку. С увеличением ошибки исходной информации до $\delta \leqslant \pm 7$ % детерминированная и вероятностная оценки вариантов планов по всем критериям обычно не совпадают. Учитывая совпадение оценок вариантов по ряду статистических критериев (математическое ожидание прибыли, Лапласа, Гурвица) при квартальном планировании рекомендуется наряду с детерминированными оценками использовать и вероятностную по критерию Лапласа (из-за менее трудоемких расчетов). При увеличении ошибки исходной информации ($\delta \geqslant \pm 7$ %) разница в оценке вариантов плана по всем критериям существенно увеличивается. Поэтому, на годовом уровне планирования оценку вариантов плана необходимо проводить с учетом вероятностного характера исходной информации, причем ответственное (разовое) решение следует принимать по критерию Вальда или Сэвиджа.

При планировании в условиях неопределенности принятое решение во многом зависит от объема (ошибки) имеющейся информации. Поэтому встает задача получения информации, а не использования специальных методов, позволяющих без нее найти решение. Подходы и идеи теории статистических решений позволяют определить рациональный объем и необходимую степень точности информации.

Проектные и плановые решения в горнодобывающих отраслях промышленности зависят от точности горно-геологической информации, т.е. степени изученности и разведанности месторождения.

Для конкретных месторождений ошибка определения горно-геологических показателей (запасов полезных ископаемых в определенных контурах, показателей качества и др.) зависит от особенностей строения месторождения и плотности разведочной сети. Ошибка различных горно-геологических показателей для месторождения в целом или его отдельных участков определяется либо по аналогии с дру-

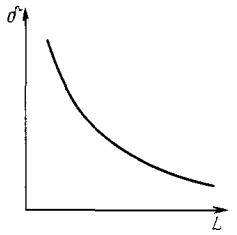


Рис. XIII.1. Зависимость величины ошибок горно-геологических показателей δ от плотности разведочной сети L

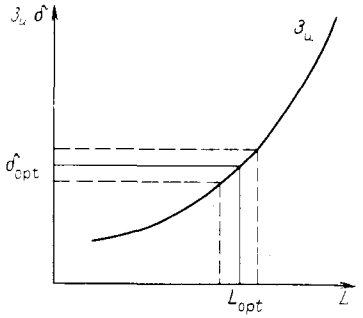


Рис. XIII.2. Зависимость затрат на получение информации (затраты на бурение и опробование) $З_и$ от плотности разведочной сети L

гими, либо специальными геолого-маркшейдерскими исследованиями. Обычно получают зависимость величины ошибок горно-геологических показателей (δ) от расстояния между скважинами (L) – плотности разведочной сети (рис. XIII. 1).

Рациональную степень точности информации (и соответствующую ей плотность разведочной сети) следует определять по минимуму общих затрат (условных) на ее получение (затраты на бурение и опробование разведочных скважин) и потерь от недостоверности.

Затраты на бурение скважин зависят от типа бурового оборудования, свойств пород, глубины разведочных скважин и расстояния между ними. Для конкретных условий строят зависимость затрат на бурение от плотности разведочной сети (рис. XIII. 2).

Потери от недостоверности исходной информации ($З_п$) определяются на основе расчета критериев принятия решений в условиях неопределенности. Для этого строится матрица рисков и необходи-

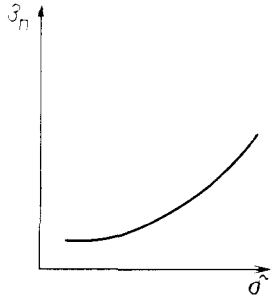


Рис. XIII.3. Зависимость экономических потерь риска $З_п$ от величины ошибок горно-геологических показателей δ

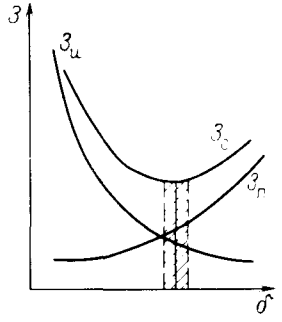


Рис. XIII.4. Обоснование рациональной точности информации

мая зависимость рассчитывается по минимальным значениям риска для определенных состояний природы, соответствующих различной точности исходной информации (рис. XIII. 3).

Зная затраты на получение информации ($З_и$) и возможные экономические потери ($З_п$) в зависимости от ее неточности, строят суммарную зависимость ($З_о$) и по ее минимуму (области минимума) определяют рациональную степень точности информации (рис. XIII. 4).

Далее, возвращаясь к кривой зависимости ошибок информации от плотности разведочной сети, находят рациональную плотность разведочной сети.

Следует заметить, что даже в пределах одного месторождения рациональная плотность зависит от сложности строения месторождения, изменчивости показателей качества и др.

§ 3. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ИГР

На практике часто возникают ситуации, в которых сталкиваются интересы двух и более сторон. Такие ситуации называются **конфликтными.** Наиболее типичны конфликтные ситуации при планировании военных действий, в спорте, в капиталистической экономике. Однако и ряд задач планирования и управления производством могут быть отнесены к конфликтным (например, обоснование цены на оборудование, планирование уровня качества и т.д.). В конфликтных ситуациях выбор решения зависит не только от нас, но и от образа действий (поведения) противоборствующей стороны.

Конфликтные ситуации изучаются **"теорией игр"**. Игра при этом представляет упрощенную модель конфликтной ситуации. Теория игр позволяет математически обосновать рекомендации по рациональному образу действий конфликтующих сторон (называемых иногда игроками).

Игры подразделяются на парные, когда сталкиваются интересы двух сторон, и множественные, когда участников конфликта несколько. Наиболее изучены парные игры. Игра как модель конфликтной ситуации отличается от нее четко сформулированными правилами, включающими:

возможные варианты поведения сторон;

объем информации о поведении противоборствующих сторон;

результат конфликта (игры) в зависимости от действий (ходов) сторон.

При этом полагается, что интересы участников конфликта могут быть оценены количественно. При этом если в результате игры одна из сторон выигрывает столько, сколько проигрывает другая, то такие игры называются **играми с нулевой суммой** (сумма выигрыша равна нулю).

Ходом в теории игр называется выбор одного из возможных по

правилам варианта действий и его осуществление. Ходы бывают личные и случайные. Личный ход осуществляется по выбору игрока, а случайный – на основании какого-либо механизма случайного выбора. В теории игр исследуются конфликтные ситуации, в которых поведение выбирается только самими участниками (т.е. игра содержит только личные ходы).

Очень важным в теории игр является понятие стратегии. **Стратегией** называется совокупность правил, по которым осуществляется выбор варианта действий в любой возможной ситуации, т.е. стратегия определена, если выбраны личные ходы для любой ситуации. Хотя решения принимаются по ходу игры, стратегия может быть выбрана заранее.

В зависимости от количества возможных стратегий игры делятся на конечные и бесконечные. Теория игр дает правила построения оптимальных стратегий. **Оптимальной** называется стратегия, обеспечивающая максимально возможный средний выигрыш при многократном повторении игры. В теории игр оптимальные стратегии вырабатываются без учета возможных ошибок противника.

Наиболее простой является парная игра с нулевой суммой. Но с ее помощью хорошо описывается большинство конфликтных ситуаций.

Рассмотрим парную игру. Сторона A может выбрать m стратегией ($i = 1, 2, \ldots, m$), а сторона B – n стратегией ($j = 1, 2, \ldots, n$). Если сторона A выбрала определенную i-ю стратегию, а сторона B – j-ю стратегию, то выбор стратегий однозначно определяет исход игры – a_{ij}, выигрыш (положительный или отрицательный) стороны A. Известные значения a_{ij} для любой пары стратегий составляют **платежную матрицу игры** $\|a_{ij}\|$ или просто матрицу игры.

Строки матрицы (табл. XIII. 6) соответствуют стратегиям стороны (игрока) A, а столбцы – стратегиям игрока B. Элемент матрицы a_{ij} показывает выигрыш игрока A, если он избрал стратегию A_i, а игрок B – стратегию B_j.

Исходя из предположения о разумности противников, определим стратегии игроков.

Проанализируем последовательно стратегии игрока A (строки платежной матрицы). При выборе стратегии необходимо учитывать, что игрок B ответит стратегией, при которой наш выигрыш минимален. Найдем минимальный выигрыш для каждой из стратегий

$$\alpha_i = \min_j a_{ij} \quad (i = 1, 2, \ldots, m). \qquad \text{(XIII.5)}$$

Числа α_i (минимумы по строкам) указываются обычно рядом с матрицей в виде добавочного столбца (табл. XIII. 7).

Таким образом, выбор стратегии A_i обеспечивает нам выигрыш α_i. Естественно, что необходимо выбрать такую стратегию, при которой наш выигрыш максимален

$$\alpha = \max_i \alpha_i$$

Таблица XIII. 6

	B_1	B_2	...	B_j	...	B_n
A_1	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}
...	...	...	...	...	...	...
A_i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}
...	...	...	...	...	...	...
A_m	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}

Таблица XIII. 7

	B_1	B_2	...	B_j	...	B_n	α_i
A_1	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}	α_1
A_2	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}	α_2
...	...	...	...	...	...	...	...
A_i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}	α_i
...	...	...	...	...	...	...	...
A_m	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}	α_m
β_j	β_1	β_2	...	β_j	...	β_n	

или

$$\alpha = \max_i \min_j \alpha_i \qquad \text{(XIII. 6)}$$

Величина α называется **нижней ценой игры**. Она показывает минимально гарантированный выигрыш игрока A. Стратегия A_i, обес-

печивающая выигрыш α, называется **максиминной.** Нижнюю цену игры называют также максимальным выигрышем или **максимином.**

Аналогичные рассуждения можно провести и за игрока B, который стремится обратить выигрыш игрока A в минимум. Поэтому просмотрев все столбцы, выделим в каждом из них максимальное значение выигрыша a_{ij}

$$\beta_i = \max_i a_{ij}.$$

Значения β_j пишутся внизу матрицы в виде дополнительной строки.

Игрок B выбирает такую стратегию B_j, при которой минимизируется значение β_j

$$\beta = \min_j \max_i a_{ij}. \qquad \text{(XIII.7)}$$

Величина β называется **верхней ценой игры** или минимаксным выигрышем (минимаксом). Стратегия B_j, дающая выигрыш β, называется минимаксной. Придерживаясь этой стратегии, игрок B не проиграет более величины β.

Нижнюю и верхнюю цены игры мы определяли исходя из предположения о разумности игроков. Этот принцип выбора стратегий называется **принципом минимакса.**

Пример XIII.4. Найти верхнюю и нижнюю цены игр, платежные матрицы которых приведены в табл. XIII. 8 и XIII. 9.

Таблица XIII. 8

	B_1	B_2	B_4
A_1	9	4	11
A_2	4	11	2
A_3	11	2	13

Таблица XIII.9

	B_1	B_2	B_3
A_1	9	8	11
A_2	6	7	9
A_3	12	6	5

Определим минимумы строк α_i и максимумы столбцов β_j (т.е. соответственно минимальные выигрыши игрока A и максимальные проигрыши игрока B).

Для игры, матрица которой задана табл. XIII. 8, получим табл. XIII. 10.

Максимальное значение α_i, равное 4, дает нижнюю цену игры.

Верхняя цена игры (равная минимуму β_j) составляет 11.

Для игры, платежная матрица которой задана табл. XIII. 9, получим табл. XIII. 11.

В данном случае нижняя и верхняя цены игры равны (8), т.е.

$$\max_i \alpha_i = \min_j \beta_j,$$

или

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = \gamma.$$

Общее значение верхней и нижней цен игры, равное γ, называется чистой

Таблица XIII. 10

	B_1	B_2	B_3	α_i
A_1	9	4	11	4
A_2	4	11	2	2
A_3	11	2	13	2
β_j	11	11	13	

Таблица XIII. 11

	B_1	B_2	B_3	α_i
A_1	9	8	11	8
A_2	6	7	9	6
A_3	12	6	5	5
β_j	12	8	11	

ценой игры, а игры, имеющие чистую цену, называются **играми с седловой точкой.**

§ 4. МЕТОДЫ РЕШЕНИЯ ПАРНЫХ ИГР

Решение игры заключается в выборе стратегии поведения каждого игрока. Оптимальной для каждого игрока является стратегия, обеспечивающая ему при многократном повторении игры максимально возможный средний выигрыш (или минимально возможный средний проигрыш).

Наиболее просто находится решение игры, имеющей седловую точку. Седловой точке соответствуют оптимальные (соответственно максиминная и минимаксная) стратегии каждого игрока. При этом, если один из игроков придерживается своей оптимальной стратегии, то для другого невыгодно отклоняться от своей оптимальной стратегии. В этом случае выигрыш постоянен и равен чистой цене игры γ. Так как чистая цена игры соответствует максимину, то отклонение игрока A от стратегии, дающей этот выигрыш, приведет только к его уменьшению. Аналогичная картина наблюдается и у игрока B. Таким образом, в играх с седловой точкой оптимальные стратегии обладают устойчивостью, т.е. каждому игроку выгодно применять какую-то одну чистую стратегию, определяемую принципом минимакса.

На практике более часто встречаются игры, у которых верхняя и нижняя цены не равны. Если в этом случае применять одну чистую стратегию по принципу минимакса, то выигрыш игрока A не будет

превышать α, а выигрыш игрока B – β. Выигрыш каждого игрока при отсутствии седловой точки может быть увеличен, если применять не одну, а несколько стратегий. Такие стратегии называются **смешанными** и заключаются в случайном чередовании чистых стратегий.

Смешанная стратегия игрока A — S_A определяется набором вероятностей $P_1, P_2, \ldots, P_i, \ldots, P_m$ использования чистых стратегий $A_i (i=1, 2, \ldots, m)$, причем $\sum_{i=1}^{m} P_i = 1$.

Аналогично, для игрока B смешанная стратегия S_B определяется набором вероятностей $q_1, q_2, \ldots, q_j, \ldots, q_n$ использования чистых стратегий $B_j (j=1, 2, \ldots, n)$, причем $\sum_{j=1}^{m} q_j = 1$.

В теории игр доказывается теорема (называется основной теоремой теории игр), согласно которой, каждая игра имеет по крайней мере одно решение в области смешанных стратегий. При этом чистые стратегии могут рассматриваться как частный случай смешанных.

Пара оптимальных стратегий S_A^* и S_B^* обладает следующим свойством: если один игрок придерживается своей оптимальной стратегии, то другому невыгодно отклоняться от своей оптимальной.

Цена игры γ при применении смешанных стратегий находится в пределах между нижней и верхней ценами игры, т.е. $\alpha \leqslant \gamma \leqslant \beta$.

Применение игроком A своей оптимальной стратегии S_A^* обеспечивает ему при любой стратегии противника B_j выигрыш не меньше γ, т.е.

$$\sum_{i=1}^{m} P_i \, a_{ij} \geqslant \gamma \quad (j = 1, 2, \ldots, n). \qquad \text{(XIII.8)}$$

Аналогично, для игрока B использование оптимальной стратегии S_B^* обеспечивает ему при любой стратегии противника A_i проигрыш, не больше γ, т.е.

$$\sum_{j=1}^{n} q_j \, a_{ij} \leqslant \gamma \quad (i = 1, 2, \ldots, m). \qquad \text{(XIII.9)}$$

Соотношения (XIII. 8) и (XIII. 9) используются для решения игры. Для получения решения матричной игры ее сводят к задаче линейного программирования. Непосредственно перед решением анализируют матрицу, исключая при наличии дублирующие и невыгодные стратегии.

При определении оптимальной стратегии игрока A должны выполняться условия

$$\left.\begin{aligned}
&P_1 a_{11} + P_2 a_{21} + \ldots + P_i a_{i1} + \ldots + P_m a_{m1} \geqslant \gamma, \\
&P_1 a_{12} + P_2 a_{22} + \ldots + P_i a_{i2} + \ldots + P_m a_{m2} \geqslant \gamma, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&P_1 a_{1j} + P_2 a_{2j} + \ldots + P_i a_{ij} + \ldots + P_m a_{mj} \geqslant \gamma, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&P_1 a_{1n} + P_2 a_{2n} + \ldots + P_i a_{in} + \ldots + P_m a_{mn} \geqslant \gamma.
\end{aligned}\right\} \qquad \text{(XIII.10)}$$

Цена игры γ, если элементы платежной матрицы неотрицательны, всегда положительное число. Если в матрице имеются отрицательные элементы, ее можно преобразовать, прибавляя ко всем элементам определенное положительное число.

Систему ограничений (XIII. 10) можно преобразовать, разделив все члены неравенств на γ и обозначив $P_i/\gamma = x_i$, $(i = 1, 2, \ldots, m)$. В результате получим

$$\left.\begin{aligned}
&a_{11} x_1 + a_{21} x_2 + \ldots + a_{i1} x_i + \ldots + a_{m1} x_m \geqslant 1, \\
&a_{12} x_1 + a_{22} x_2 + \ldots + a_{i2} x_i + \ldots + a_{m2} x_m \geqslant 1, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{1j} x_1 + a_{2j} x_2 + \ldots + a_{ij} x_i + \ldots + a_{mj} x_m \geqslant 1, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&a_{1n} x_1 + a_{2n} x_2 + \ldots + a_{in} x_i + a_{mn} x_m \geqslant 1.
\end{aligned}\right\} \qquad \text{(XIII. 11)}$$

Из условия, что $\sum_{i=1}^{m} P_i = 1$, получим

$$x_1 + x_2 + \ldots + x_i + \ldots + x_m = 1/\gamma. \qquad \text{(XIII. 12)}$$

Оптимальная стратегия S_A^* должна максимизировать значение γ, а следовательно, минимизировать $1/\gamma$.

Таким образом, задача определения выбора вероятностей P_i, составляющих оптимальную стратегию S_A^*, сводится к минимизации линейной формы (XIII. 12) при ограничениях (XIII. 11).

Решив полученную задачу линейного программирования и найдя величины x_i и $1/\gamma$, определим значения $P_i = \gamma x_i$.

Оптимальная стратегия игрока B – S_B^* находится аналогичным образом, исходя из условия минимизации проигрыша (γ), при соблюдении условий (XIII. 9).

Разделив систему неравенств (XIII. 9) на γ и заменив q_i на $y_i = \frac{q_i}{\gamma}$, получим следующую задачу линейного программирования

$$\left.\begin{aligned}
&a_{11}y_1 + a_{12}y_2 + \ldots + a_{1j}y_j + \ldots + a_{1n}y_n \leqslant 1,\\
&a_{21}y_1 + a_{22}y_2 + \ldots + a_{2j}y_j + \ldots + a_{2n}y_n \leqslant 1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{i1}y_1 + a_{i2}y_2 + \ldots + a_{ij}y_j + \ldots + a_{in}y_n \leqslant 1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{m1}y_1 + a_{m2}y_2 + \ldots + a_{mj}y_j + \ldots + a_{mn}y_n \leqslant 1,
\end{aligned}\right\} \quad \text{(XIII.13)}$$

$$y_1 + y_2 + \ldots + y_j + \ldots + y_n \to \max. \quad \text{(XIII.14)}$$

Задача максимизации линейной формы (XIII. 14) при ограничениях (XIII. 13) является двойственной по отношению к задаче, определяемой условиями (XIII. 11) и (XIII. 12).

Таким образом, задача отыскания решения матричной игры сводится к решению пары симметричных двойственных задач линейного программирования. Решение прямой задачи дает оптимальную стратегию игрока A (S_A^*), а двойственной – оптимальную стратегию игрока B (S_B^*).

Пример XIII.5. Найти оптимальные стратегии игроков A и B, т.е. решение игры, платежная матрица которой представлена табл. XIII. 8.

Найдем смешанную стратегию игрока A.

Обозначив $x_1 = P_1/\gamma$, $x_2 = P_2/\gamma$ и $x_3 = P_3/\gamma$, получим следующую модель задачи

$L = x_1 + x_2 + x_3 \to \min$,
$9x_1 + 4x_2 + 11x_3 \geqslant 1$,
$4x_1 + 11x_2 + 2x_3 \geqslant 1$,
$11x_1 + 2x_2 + 13x_3 \geqslant 1$.

Решая данную модель симплекс-методом, получим

$x_1 = 1/28$; $x_2 = 1/14$; $x_3 = 1/28$.

Цена игры $\gamma = 1/L_{\min} = 1 : (1/28 + 1/14 + 1/28) = 7$.

Соответственно вероятности использования игроком A первой, второй и третьей стратегий составят

$P_1 = \gamma x_1 = 7 \cdot 1/28 = 1/4$,
$P_2 = \gamma x_2 = 7 \cdot 1/14 = 1/2$,
$P_3 = \gamma x_3 = 7 \cdot 1/28 = 1/4$, т.е.

$S_A^* = (1/4;\ 1/2;\ 1/4)$.

Оптимальная стратегия игрока B(S_B^*) находится аналогичным образом, только ограничения (типа $\leqslant$) составляются по столбцам, а критерий оптимальности следует не минимизировать, а максимизировать. Составив модель и решив ее, получим, что оптимальная стратегия игрока B состоит в следующем

$q_1 = 1/4$; $q_2 = 1/2$; $q_3 = 1/4$ или $S_B^* = (1/4;\ 1/2;\ 1/4)$.

В рассмотренной игре каждый игрок использует все свои стратегии, т.е. все стратегии здесь **активны.** (**Активной** называется стратегия, вероятность использования которой больше нуля).

Сведение игры к задаче линейного программирования позволяет получить точное ее решение. Однако в ряде задач достаточно иметь приближенное решение, обеспечивающее выигрыш, близкий к цене игры. Поэтому, учитывая, что для больших матриц применение линейного программирования не самый простой путь решения, используют различные приближенные методы (итерационные, физические смеси стратегий и т.д.).

Контрольные вопросы и задачи

1. При решении каких задач используется теория статистических решений?
2. В каких случаях используются матрицы риска?
3. Какие критерии используются при принятии разовых, ответственных решений?
4. Какие критерии используются в часто повторяющихся ситуациях?
5. Чем критерий Сэвиджа отличается от критерия Вальда?
6. Каким образом теорию статистических решений можно использовать для обоснования необходимого объема и точности информации?
7. Что такое цена игры?
8. Чем верхняя цена игры отличается от нижней?
9. Что такое принцип минимакса?
10. Когда игра имеет седловую точку?
11. В каких случаях используются только чистые стратегии игроков?
12. Когда игрокам целесообразно использовать смешанные стратегии?
13. Как определяются вероятности использования различных стратегий?
14. Что такое доминирующая стратегия?
15. Задача. Определить значения критериев Вальда, Сэвиджа, Гурвица, Лапласа и наилучшую стратегию. Для решения задачи прежде всего необходимо сформировать платежную матрицу для определенного варианта. Для формирования матрицы находим в табл. XIII. 12 номер варианта.

Таблица XIII. 12

Стратегии	Состояния природы					
	1–4	2–5	3–6	4–7	5–8	6–9
1–4	1	2	3	4	5	6
2–5	7	8	9	10	11	12
3–6	13	14	15	16	17	18
4–7	19	20	21	22	23	24
5–8	25	26	27	28	29	30
6–9	31	32	33	34	35	36

Таблица XIII.13

Стратегия	Состояния природы								
	1	2	3	4	5	6	7	8	9
1	7	6	8	5	9	11	8	10	12
2	5	9	6	7	7	8	10	9	11
3	6	7	9	8	6	9	8	7	10
4	8	9	7	6	10	11	9	12	8
5	10	8	9	7	9	8	10	11	9
6	6	7	8	9	10	11	12	10	10
7	8	9	7	10	9	8	11	12	9
8	9	7	10	11	12	10	9	10	9
9	10	9	6	10	9	11	9	9	10

В строке с номером варианта в левой графе стоят номера строк, соответствующие определенным стратегиям (табл. XIII.13). В столбце с номером варианта в верхней строке табл. XIII. 12 указаны номера столбцов, соответствующие состояниям природы $П$ (табл. XIII. 13).

Например, номер варианта 14. В левой графе табл. XIII.12 стоит 3–6. Значит, из табл. XIII.13 берем строки с 3 по 6, соответствующие 3–6 стратегиям. В верхней строке табл. XIII. 12 над номером варианта стоит 2–5. Значит, из строк 3–6 табл. XIII. 13 берем значения, соответствующие столбцам с 2 по 5 той же таблицы. Платежная матрица будет иметь вид (табл. XIII. 14).

16. Задача. Проанализируйте конфликтную ситуацию между сторонами (игроками) А и В (парная игра). Каждый игрок может использовать по 4 стратегии. Платежную матрицу игры для выбранного (заданного) варианта построить по данным табл. XIII. 12 и XIII. 15 (как описано в задаче 14).

17. Задача. Проектируется строительство горного предприятия. Горно-геологические условия, цены на готовую продукцию и объем сбыта на перспективу точно не определены. Имеются 4 варианта технологических решений (отличающихся мощностью предприятия, структурой комплексной механизации и системой разработки). Определить наилучшее технологическое решение, если рассматривается 4 возможных состояния природы. Платежную матрицу для решения задачи для заданного (выбранного) варианта построить по данным табл. XIII.12 и XIII. 15.

Определите:

значения критериев Вальда, Лапласа, Гурвица для всех стратегий (при определении критерия Гурвица задайтесь коэффициентом, выражающим долю оптимизации от 0.3 до 0.6, с шагом 0,1);

Таблица XIII.14

Стратегия	Состояния природы			
	2	3	4	5
3	7	9	8	6
4	9	7	6	10
5	8	9	7	9
6	7	8	9	10

Таблица XIII.15

Стратегия игрока A	Стратегия игрока B								
	1	2	3	4	5	6	7	8	9
1	11	12	10	10	7	8	8	11	10
2	8	11	12	9	10	8	9	10	9
3	10	9	10	9	11	12	7	10	11
4	11	8	9	10	8	9	11	12	10
5	9	8	9	7	6	10	9	8	6
6	8	7	9	10	8	7	9	8	9
7	9	8	8	7	9	8	7	10	11
8	9	9	10	8	8	9	7	8	8
9	10	8	9	9	10	10	8	9	10

Таблица XIII.16

Вариант	Состояния природы $П$				
	1	2	3	4	5
1	0,1	0,2	0,3	0,1	0,3
2	0,3	0,1	0,2	0,3	0,1
3	0,2	0,3	0,1	0,1	0,3
4	0,2	0,1	0,4	0,2	0,1
5	0,4	0,2	0,1	0,1	0,2

насколько изменится принятое решение, если установлены вероятности состояний природы (табл. XIII. 16)

значения критериев Сэвиджа, для чего составьте матрицу рисков;

для матрицы рисков значения критериев Лапласа и Гурвица;

обоснуйте наилучшее решение, проанализировав всю совокупность критериев.

Глава XIV

ЭЛЕМЕНТЫ ТЕОРИИ ВЫБОРА И ПРИНЯТИЯ РЕШЕНИЙ

§ 1. ЗАДАЧИ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

Математические модели, формализующие свойства определенной производственной ситуации, позволяют получить строгое оптимальное решение при конкретной, четко поставленной цели, описываемой единственным критерием оптимальности. На практике же решения

обычно оцениваются с различных точек зрения, имеется как правило несколько целей и соответствующие им критерии оптимальности.

При управлении горными предприятиями, являющимися социально-экономическими системами, стремятся достичь следующие цели: выполнение плана добычи полезного ископаемого, рациональное использование недр, повышение рентабельности производства и производительности труда, наилучшее использование материальных ресурсов, снижение себестоимости продукции, улучшение качества добываемой продукции, снижение травматизма, наиболее полное удовлетворение социальных и культурных потребностей трудящихся, т.е. они являются типичными многоцелевыми объектами.

По существу практически все решаемые задачи многокритериальные. Например, при выборе вариантов плана горных работ необходимо учитывать объемы горных и подготовительных работ, качество полезного ископаемого (выражающееся также обычно комплексом показателей), использование трудовых и материальных ресурсов, расстояние откатки, уровень потерь и разубоживания.

Сложными многокритериальными задачами являются выбор проектного (технического) решения, где наряду с экономическими показателями следует учитывать вопросы надежности и безопасности, трудность эксплуатации и строительства, сроки осуществления проекта.

Другие сложные многокритериальные задачи – оценка работы предприятий, подбор кандидата на должность, распределение средств (ассигнований на научные исследования, капитальных вложений, фонда развития производства и др.) между различными объектами и т.д.

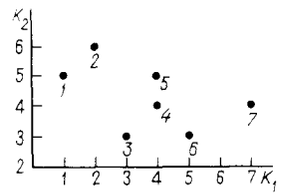
Таким образом, возникает проблема оптимизации принимаемых решений сразу по нескольким критериям. Эту проблему изучают с помощью **теории выбора и принятия решений**.

Отличительная особенность многокритериальных задач – наличие неопределенности, связанной с необходимостью сопоставления решений по нескольким критериям. Например, при построении годового плана горных работ карьера получено четыре его варианта, сопоставляемых по трем основным технологическим (для конкретной задачи) критериям (коэффициент вскрыши K_B, содержание компонентов: полезного α, вредного β).

	K_B, м³/т	α, %	β, %
1	3,7	2,29	1,60
2	4,0	2,46	1,62
3	4,1	2,35	1,52
4	3,9	2,38	1,54

В данном случае, как и обычно в многокритериальных задачах, отсутствует вариант одновременно наилучший по всем критериям (так называемый, доминирующий): вариант 1 – наилучший по первому

Рис. XIV.1. Характеристики вариантов решений (1–7) по двум критериям (K_1, K_2)



критерию ($K_B \to \min$), 2 – по второму ($\alpha \to \max$) и 3 – по третьему ($\beta \to \min$). Вариант 4 не является оптимальным ни по одному из критериев, но по всем из них стоит на втором месте, т.е. компромиссный. Обычно решение задачи многокритериальной оптимизации заключается в поиске наилучшего компромиссного решения. В области компромиссных решений нельзя достигнуть одновременного улучшения решения по всем критериям. Область компромиссов является областью оптимальных решений по Парето*. При этом оптимальными по Парето являются решения, имеющие по всем критериям не худшие, а по одному из них лучший показатель, чем другие альтернативы. Выделение решений, оптимальных по Парето, обычно позволяет сократить количество рассматриваемых альтернатив.

Приведем пример решений, оптимальных по Парето (множества Парето). Предположим, что сопоставляются 7 вариантов решений (альтернатив) по двум критериям K_1 и K_2, которые надо максимизировать. Характеристики вариантов показаны на рис. XIV. 1.

	K_1	K_2
1	1	5
2	2	6
3	3	3
4	4	4
5	4	5
6	5	3
7	7	4

Альтернативы 2, 5, 7 являются оптимальными по Парето, так как по обоим критериям лучше их нет. В то же время альтернативы 1, 3, 4, 6 не входят в множество Парето, так как, например, лучше альтернативы 1 по обоим критериям альтернатива 2, альтернатива 3 лучше 5, а 7 – лучше альтернативы 6. Альтернатива 4 не входит в множество Парето, так как альтернатива 5 не хуже ее по первому критерию и лучше по второму.

Покажем формирование области компромиссных решений на примере определения оптимального значения одной непрерывной переменной по двум критериям (рис. XIV. 2)

---

* Итальянский специалист в области политической экономии капитализма.

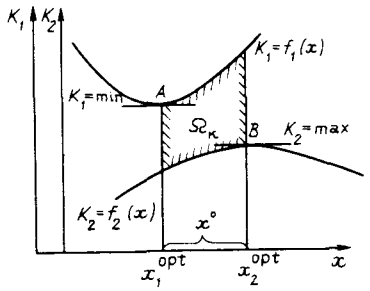


Рис.XIV.2. Область компромиссных решений

$K_1 = f_1(x) \to \min$ и $K_2 = f_2(x) \to \max$.

По первому критерию оптимальным решением является x_1^{opt}, которое соответствует минимуму функции $f_1(x)$ в точке A, по второму – x_2^{opt}, соответствующее максимуму функции $f_2(x)$. Оптимальные значения оптимизируемой переменной $x - x^0$ заключены на отрезке оси абсцисс между локальными по одному из критериев оптимами x_1^{opt} и x_2^{opt}. Заштрихованная область $\Omega_к$, заключенная между вертикальными линиями, проходящими через точки экстремальных значений целевых функций и ограниченная сверху и снизу графиками целевых функций $K_1 = f_1(x)$ и $K_2 = f_2(x)$, является областью компромиссов.

Указанная неопределенность – принципиальна для многокритериальных задач и не может быть исключена методами математического моделирования. Практически единственный путь избавления от подобной неопределенности – использование при оценке и выборе вариантов информации лица, принимающего решения (ЛПР), основанной на его понимании проблемы, знаниях, опыте, интуиции. При этом информация, получаемая на основе математической модели и ЛПР, по существу создает новую модель задачи, позволяющей формировать, оценивать и выбирать варианты решений. Хотя в новую модель привносится "субъективный" фактор – мнение ЛПР, она не может считаться полностью субъективной, так как его действия должны быть рациональными, убедительными, логичными, понятными для окружающих. Поэтому любая информация ЛПР обычно отражает политику различных групп или организации в целом. Опыт использования информации ЛПР в процессе выработки решений показывает, что для лиц, принадлежащих определенной организации, обычно свойственна идентичность взглядов на проблему и оценки вариантов решений.

Кроме того, для уменьшения роли субъективного фактора используют информацию не одного лица (ЛПР), а группы лиц – экспертов. Экспертов привлекают также для решения сложных проблем, когда ЛПР непосредственно не обладает всей информацией, необходимой для принятия решений. Для получения информации в этих случаях

привлекают экспертов, обладающих глубокими специальными знаниями по отдельным аспектам (критериям оценки) изучаемой проблемы. При этом ЛПР определяет состав критериев, формирует требования к ним, дает оценки альтернативам, основываясь на мнениях экспертов.

Таким образом, если лицом, принимающим решение (ЛПР), является человек компетентный и опытный в своей области, наделенный необходимыми полномочиями, имеющий цель, послужившей основой для постановки и формализации задачи, принимающий решение и несущий за него ответственность, то экспертом называют специалиста, имеющего информацию о рассматриваемой проблеме и дающего оценки отдельным вариантам ее решения, но не несущего непосредственной ответственности за его принятие.

§ 2. МЕТОДЫ ФОРМИРОВАНИЯ АЛЬТЕРНАТИВ

Задача принятия решений заключается в выборе из множества альтернатив (допустимых решений) Ω, наилучшего в соответствии с определенным принципом оптимальности (ОП). Наличие ряда альтернатив и принципа оптимальности – обязательное условие, ибо при отсутствии любого из этих элементов задача теряет смысл. Действительно, если нет ряда альтернатив (т.е. $\Omega = \phi$), то искать решение не из чего, а если нет принципа оптимальности, то оценить решения и найти лучшее из них невозможно. Если Ω и ОП описываются с помощью систем неравенств и уравнений, то имеем обычную оптимизационную модель (II.1), (II.2). В общем случае задачи принятия решений Ω и ОП могут быть неизвестны и информацию о них получают в процессе решения. Процесс принятия решения состоит из двух этапов: на первом формируют множество допустимых решений Ω, называемое также исходным множеством альтернатив (ИМА), а на втором выбирают из него наилучшее решение в соответствии с поставленными целями, т.е. решают задачу выбора. Различают задачи принятия решений, в которых этапы четко разделены, т.е. сначала сформированы альтернативы, которые затем оцениваются, и задачи, в которых альтернативы формируются (улучшаются) последовательно после выполнения очередного этапа их оценки. Второй способ используется обычно в задачах планирования.

Методы формирования исходного множества альтернатив зависят от конкретных условий задачи, при этом ограничения задачи (технологические, экологические, по ресурсам, качеству продукции, надежности и т.д.) определяют допустимость и возможность использования альтернатив. Во всех методах формирование ИМА включает два шага: на первом альтернатива генерируется, а на втором – проверяется ее допустимость, т.е. соответствие ограничениям задачи, иногда эти этапы совмещаются.

Различают 5 методов формирования ИМА:

формирование конечного исходного множества альтернатив на основе экспертизы;

формирование ИМА с помощью модели;

использование морфологического анализа;

поэтапное формирование ИМА в иерархических организационных структурах;

формирование бесконечного исходного множества альтернатив.

Экспертный метод формирования ИМА используют в сложных ситуациях, когда формализация проблемы невозможна или затруднительна. Например, при разработке стратегических решений, поиске выхода из непредвиденных ситуаций. В этих случаях альтернативы генерируются экспертами на основании своих знаний, опыта, интуиции. Несмотря на то, что эта деятельность носит ярко выраженный творческий характер, в последнее время разработаны специальные процедуры для выработки решений, например, системный анализ, метод "мозговой атаки" и др. При этом первоначально составляют полный перечень альтернатив, из которого затем исключаются не удовлетворяющие условиям (ограничениям) задачи. В ряде случаев конечное множество альтернатив формируется естественно (например, при решении задачи назначения на должность рассматривают всех имеющихся специалистов, удовлетворяющих требованиям по стажу работы, образованию и др.).

Формирование ИМА с помощью модели используют, когда эксперты не в состоянии перечислить все допустимые альтернативы, но известен способ (модель) их построения. Например, все варианты графиков движения локомотивосоставов, осуществляющих транспорт угля на шахте до околоствольного двора, эксперту составить практически невозможно. В то же время известны параметры графика движения (времена погрузки, разгрузки, движения на отдельных участках) и схема путевого развития, позволяющие составить модель (имитационную) формирования графика движения. Аналогичным образом можно сформировать на основе модели допустимые варианты (учитывая при этом допустимую технологию отработки, ограничения по условиям проветривания, поддержания горных выработок и т.д.) порядка отработки столбов шахтного поля, расстановки оборудования на карьере, распределения ресурсов по участкам работ и т.д.

Этот же метод можно использовать и при непрерывном изменении переменной.

В этом случае в пределах допустимых (минимальных и максимальных) величин переменной, задаваясь шагом изменения величины переменной, определяют ее возможные значения. Естественно, что количество вариантов существенно зависит от выбранного шага изменения переменной. Например, если оптимизируется длина лавы и известно, что она должна находиться в пределах 120–200 м, то при шаге изменения 10 м, рассматривается 9 альтернатив

(соответственно 120, 130, 140, 150, 160, 170, 180, 190 и 200 м), а при шаге изменения 5 м – 17 альтернатив и т.д.

Метод морфологического анализа используется, когда каждая альтернатива состоит из ряда составных частей (компонент). Например, при выборе структуры комплексной механизации горных работ на карьере рассматриваются варианты бурового, экскаваторного, транспортного и отвального оборудования, т.е. вся структура состоит из 4 компонент. С помощью модели (а при небольшом числе компонент и их вариантов – экспертным путем) на основании заданных ограничений (например, допустимые сочетания оборудования, требуемый в целом уровень надежности, производительность схемы и т.д.) формируют допустимые альтернативы – приемлемые для рассматриваемых условий структуры комплексной механизации. Метод морфологического анализа используют при проектировании сложных технических систем: шахты в целом; уникального оборудования, комплекса технических средств АСУ и т.д.

При разработке программ, планов и решения различных технико-экономических задач в иерархических многоуровневых организационных структурах, например, отрасль – производственное объединение (комбинат) – производственная единица (цех) используют **метод поэтапного формирования ИМА.** При этом на каждом из уровней иерархии вырабатываются свои альтернативы, которые затем обобщаются (детализируются) на следующем уровне и т.д., т.е. каждая последующая экспертиза основывается на результатах предыдущей. Различают два метода проведения иерархической экспертизы: снизу вверх (от низших уровней иерархии к высшим) и сверху вниз (от высшего уровня иерархии к низшим).

При первом способе сначала формируют альтернативные варианты на уровне производственных единиц (цехов). При этом альтернативы могут отличаться объемами плановых заданий, различными размерами используемых ресурсов (материальных, финансовых, трудовых и т.д.). Сформированные альтернативы (планварианты) передаются в объединение (комбинат), где формируют альтернативы развития объединения. При этом, естественно, часть альтернатив производственных единиц может и не войти в варианты, разработанные объединением и представленные на рассмотрение на отраслевом уровне, в министерство. Министерство на основе вариантов, представленных объединениями (комбинатами) с учетом общего количества ресурсов, выделенных отрасли в целом, формирует варианты плана отрасли.

При втором способе формирования ИМА в иерархических организационных структурах, т.е. сверху вниз, результаты разработки планов на более высоких уровнях служат исходными данными (ограничениями по объемам выпуска продукции, расходу ресурсов) для разработки планов на более низких иерархических уровнях. Т.е. первоначально разрабатываются допустимые планы развития отраслей, за-

тем на их основе – планы развития объединений и на следующем этане – альтернативы развития производственных единиц. При этом планы объединения должны быть допустимы с позиций отрасли, а планы производственных единиц – с позиций объединения.

Обычно при планировании и решении других задач в иерархических организационных структурах последовательно итерационно используют оба вида процедур, при этом на каждой следующей итерации конкретизируются ранее намеченные альтернативы.

§ 3. ОБЩИЕ СВЕДЕНИЯ О МЕТОДАХ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

После того, как сформированы альтернативы (ИМА), приступают ко второму этапу задачи принятия решений – решению задачи выбора. Метод решения этой задачи зависит, прежде всего, от требований, предъявляемых к окончательному решению. Различают три типа задач принятия решений:

разбиение рассматриваемых вариантов решений (альтернатив) на несколько классов – задача классификации объектов;

упорядочение альтернатив по качеству, т.е. расстановка их по местам (рангам) в порядке их ценности (значимости) – задача ранжирования;

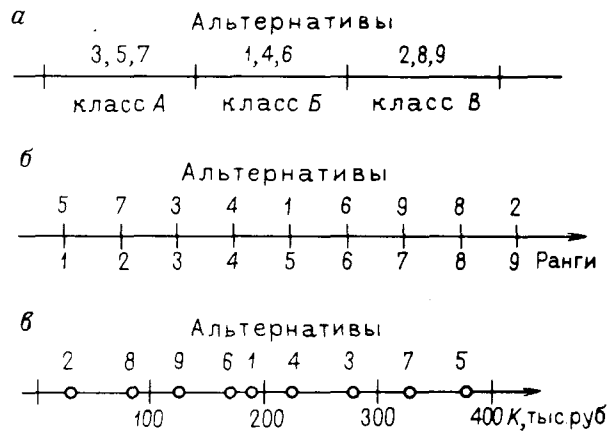
выбор наилучшего варианта решения на основе количественной оценки.

Задачи классификации применяются для разделения горных предприятий по классам в зависимости от горно-геологических, технологических и организационных условий, для разработки нормативов для групп однородных предприятий, использования типовых технологических и организационно-управленческих решений и объективной оценки их работы (внутри однородной группы нивелируется влияние природных условий и других независящих от предприятия факторов).

Задачи ранжирования используются для оценки альтернатив, выбора наилучшей из них. Конечно, количественная оценка решений – наиболее желательна. Однако строгое решение может быть получено в том случае, когда есть один критерий, имеющий естественное количественное измерение. Когда же имеется несколько критериев, возникает задача их сопоставления между собой, взвешивания. Кроме того, в большом количестве задач присутствуют качественные критерии (например, оценка деловых качеств руководителя; оценка проектов с позиций экологии, обороноспособности, социальной и др.). Поэтому использование количественной оценки не всегда возможно.

Для решения задач классификации, ранжирования и количественной оценки используются соответственно номинальные, порядковые и количественные шкалы (рис. XIV. 3).

Рис.XIV.3. Типы шкал: номинальная (*а*); порядковая (*б*); количественная (*в*)



Информативность шкал возрастает по мере перехода от номинальной шкалы к порядковой и далее к количественной. На номинальной шкале выделены только классы решений, в которые попадают различные альтернативы (рис. XIV. 3 *а*).

В порядковых шкалах (рис.XIV.3*б*) альтернативы уже располагаются по порядку (рангу), по мере их предпочтения, т.е. альтернатива 5 предпочтительнее всех альтернатив, 7 – всех исключая 5, и т.д.

В количественных шкалах в дополнение к порядку дается также количественная оценка альтернатив. На рис.XIV.3*в* рассматривается задача максимизации критерия K (прибыль, тыс. руб.).

Как ранее было показано для разрешения неопределенности, возникающей при наличии нескольких критериев, используется информация ЛПР. Проанализируем основные операции, которые выполняет ЛПР в процессе принятия решений. Выделяют три группы операций:
A – с критериями;
$Б$ – с альтернативами, имеющими оценки по нескольким критериям;
B – с переменными.

Основные операции с критериями следующие:

A_1 – операция назначения весов критериев.

Эта операция используется для составления комплексного критерия, количественной оценки полезности альтернатив. Надежность этой операции невысока, действия ЛПР весьма субъективны.

A_2 – операция упорядочения критериев по важности, обычно выполняется ЛПР весьма надежно и устойчиво.

A_3 – операция назначения вероятности достижения определенного значения критерия. Эта операция еще более субъективна, чем A_1.

A_4 – операция сравнения полезности двух значений одного критерия. Эта операция весьма проста при использовании как количественных, так и порядковых шкал.

A_5 – операция сравнения полезности оценок двух критериев (либо их разности) при постоянных значениях других критериев. Эта весьма распространенная операция служит для установления предпочтительности одного из значений и установления компромисса между двумя критериями; выполняется ЛПР достаточно уверенно и надежно.

A_6 – операция количественного определения значения (или изменения значения) одного критерия, соответствующего количественному значению (или его изменению) другого критерия. Эта операция существенно более сложная, чем A_5.

A_7 – операция определения допустимого значения по одному из критериев. Эта операция выполняется достаточно надежно и используется при переводе критерия в ограничение.

A_8 – операция выделения критериев, по которым: значение целесообразно улучшить, решения могут быть ухудшены и значение должно остаться не хуже достигнутого уровня. По существу эта операция является совокупностью операции A_4 и A_5 и выполняется ЛПР достаточно уверенно и просто.

A_9 – выделение критериев, значение по которым необходимо улучшить в первую очередь. Эта операция весьма проста.

Основные операции с альтернативами, оцениваемыми по нескольким критериям, следующие:

$Б_1$ – операция обоснования "принципа" сравнения альтернатив, заключающаяся в обосновании ЛПР вида зависимости интегральной оценки альтернативы от оценок по многим критериям. Постулируя зависимость, ЛПР должен уметь предвидеть последствия своего решения, что весьма сложно. Эта операция близка операции A_1 и также выполняется весьма ненадежно.

$Б_2$ – оценка вероятностей выполнения решения в целом и достижения критериев, характеризующих данную альтернативу. Эта операция более сложна, чем операция A_3.

$Б_3$ – операция сопоставления двух альтернатив в целом и выделение лучшей (худшей) из них. Сложность этой весьма распространенной операции зависит от числа критериев, характеризующих альтернативу. Уже при трех критериях ЛПР испытывает затруднения, особенно при оценке сложных альтернатив, когда необходимо привлечение экспертов.

$Б_4$ – операция определения лучшей альтернативы из группы. Эта операция может рассматриваться как последовательное использование относительно более простых операций $Б_3$ и является сложной.

$Б_5$ – операция определения идеальной альтернативы, по близости к которой выявляется лучшая. Эта операция весьма сложна, в ряде случаев выполняется формально, когда за идеальную альтернативу условно принимается та, которая обеспечивает наилучшие значения по всем критериям (метод "идеальной точки").

В ряде задач при непрерывном или дискретном изменении переменных требуется выделять переменные, значения которых следует уменьшить (увеличить) исходя из каких-либо соображений. Такая необходимость возникает, например, в задачах управления развитием горных работ, когда по технологическим соображениям требуется форсированное развитие (или консервация) определенных участков карьерного (шахтного) поля. Этот тип операции (B_1) встречается относительно нечасто и не должен вызывать серьезных трудностей.

Методы многокритериальной оптимизации, реализующие перечисленные операции, выполняемые ЛПР, зависят от следующих основных факторов:

наличия или отсутствия объективной математической модели задачи;

вида требуемого решения;

новизны решаемой задачи;

сложности проблемы;

размерности задачи.

При возможности построения объективной математической модели задачи в целом при наличии нескольких критериев возникает многокритериальная задача математического программирования. В противном случае задача принятия решения сводится к задаче выбора, решаемой ЛПР самостоятельно (так называемый целостный выбор) или с помощью экспертов (критериально-экспертный выбор), если проблема сложна и ЛПР не может охватить самостоятельно всех ее аспектов.

Если проблема или условия, в которых она решается, новы для ЛПР, то он не может использовать имеющийся опыт и целиком полагается на знания и интуицию. При повторяющихся ситуациях ЛПР использует накопленный у него опыт решения задачи, его действия более уверенны и объективны; для таких задач могут быть выработаны и типовые правила решения.

Размерность задачи определяется количествами имеющихся альтернатив и используемых критериев оптимальности. Следует указать, что возможности человека не позволяют устойчиво и корректно решать задачи с количеством критериев более 5–7. Возможности человека по хранению и переработке информации должны обязательно учитываться при построении методов многокритериальной оптимизации. Методы решения задач выбора при наличии нескольких критериев можно условно разделить на 4 группы.

I. Методы, позволяющие свести многокритериальную задачу к одноцелевой, на основе использования функции ценности (полезности) альтернатив.

II. Теоретико-множественные методы, использующие функции выбора, описываемые бинарным отношением.

III. Методы последовательного формирования, анализа и улучшения решения.

IV. Комбинированные.

В методах I и III групп используются количественные шкалы, II группы – порядковые и номинальные шкалы, IV – в дополнение к количественным применяют порядковые, а иногда и номинальные шкалы.

Если методы I и II групп при относительно небольшой размерности задачи могут использоваться без ЭВМ, то основным средством реализации методов III и IV групп являются человеко-машинные (ЧМ) процедуры, называемые также интерактивными, диалоговыми. В человеко-машинных процедурах осуществляется непосредственное взаимодействие ЛПР с ЭВМ. ЛПР определяет первоначальные требования к задаче, ограничения, желательное соотношение критериев. Получив решения и реальные значения критериев, он уточняет свою информацию, изменяет требования и т.д. В результате такой итерационной процедуры ЭВМ вырабатывает все более лучшее решение. Процесс завершается, когда выданное ЭВМ решение приемлемо для ЛПР или он убеждается в нецелесообразности попыток дальнейшего улучшения решения при данных условиях.

§ 4. КОЛИЧЕСТВЕННЫЕ МЕТОДЫ РЕШЕНИЯ МНОГОКРИТЕРИАЛЬНЫХ ЗАДАЧ

Количественные методы основаны на введении единственного комплексного критерия оптимальности (функции ценности), позволяющего однозначно оценивать сравниваемые альтернативы. В этом случае проблема векторной оптимизации сводится к скалярной за счет свертки векторного критерия, т.е. задача сводится к одноцелевой.

При построении функции ценности наиболее часто используют аддитивный критерий, реже мультипликативный.

Для соизмерения различных целей устанавливают относительные "веса" критериев, показывающие взаимоотношения между их шкалами, т.е. их важность. Чаще всего используется линейная функция ценности, которая вычисляется следующим образом

$$U = \sum_{i=1}^{N} w_i K_i = w_1 K_1 + \ldots + w_i K_i + \ldots + w_N K_N , \quad \text{(XIV. 1)}$$

где i – индекс критерия оптимальности, $i = 1, N$; w_i – вес (важность) i-го критерия. Обычно используются относительные веса, т.е. $\sum_{i=1}^{N} w_i = 1$; K_i – значение i-го критерия, если K_i – целые числа, метод является балльным.

При мультпликативном критерии агрегированная оценка вычисляется следующим образом

$$U = \prod_{i=1}^{N} w_i K_i . \quad \text{(XIV. 2)}$$

В функциях ценности учитывается знак критерия. Например, если критерий учитывает отрицательные свойства продукции, то он входит в функцию ценности со знаком "минус". В мультипликативных критериях при решении задач максимизации в числителе отражаются критерии, значение которых желательно увеличить, а в знаменателе те, значение которых следует стремиться уменьшить. Например, типичный мультипликативный критерий – показатель рентабельности (отношение прибыли к фондам).

При решении стохастических задач можно учитывать вероятности достижения определенных целей (значений критерия) – P_i. Тогда ценность альтернативы определяется по формуле

$$U = \sum_{i=1}^{N} P_i w_i K_i . \quad \text{(XIV. 3)}$$

Различают три группы методов построения функций ценности: прямые, аксиоматические, компенсационные.

В прямых методах зависимость общей ценности (полезности) альтернатив от их значений по отдельным критериям известна или постулируется непосредственно ЛПР заранее. Известна (постулируется), как форма зависимости (чаще всего линейная), так и ее параметры (веса отдельных критериев).

При этом "вес" критерия может быть определен на основании специальных исследований (например, экономико-статистических), прошлого опыта или определен ЛПР по каким-либо другим соображениям.

Например, основные показатели любого проекта – удельные эксплуатационные (C) и капитальные (K) затраты. В настоящее время в качестве критерия оценки эффективности технических решений приняты приведенные затраты $C + E_н K$, т.е. $E_н$ по существу представляет "весовой" коэффициент (коэффициент приведения капитальных затрат к эксплуатационным) и определяется на основании специальных исследований.

При оценке месторождений комплексных полезных ископаемых, обосновании кондиций и решении различных вопросов управления качеством используется показатель приведенного содержания полезного компонента ($a_{пр}$), т.е. все полезные компоненты ($j = \overline{1, L}$) приводятся к одному – главному. При этом в качестве "весовых" коэффициентов используют отношения цен полезных компонентов ($Ц_j$) к цене главного компонента ($Ц_{гл}$), т.е.

$$a_{пр} = \sum_{j=1}^{L} \frac{Ц_j}{Ц_{гл}} a_j .$$

Часто "весовые" коэффициенты различных технологических критериев оптимальности (например, коэффициент вскрыши, уровень концентрации горных работ, показатели качества и их стабильности

и т.д.) определяют на основании прошлого опыта (имеющихся статистических материалов) методами математической статистики, устанавливая их влияние на результирующие экономические показатели работы предприятия. Следует, однако, иметь в виду, что статистика отражает лишь прошедший уровень: "весовые" коэффициенты, определенные на основании статистических исследований, не отражают конкретной реальной ситуации, а также предпочтения критериев при решении задачи в сложившихся условиях.

Учитывая ограниченность аналитических и статистических методов, "весовые" коэффициенты чаще всего устанавливают экспертным путем, непосредственно ЛПР. В случае неудовлетворенности полученным решением ЛПР может изменить "весовые" коэффициенты.

Рассмотрим определение "весовых" коэффициентов для задачи сопоставления четырех вариантов плана горных работ по трем технологическим критериям, условия которой приведены § 1.

Предположим, что на основании статистических исследований установлены следующие "весовые" коэффициенты $w_1 = 0{,}5$; $w_2 = 0{,}3$; $w_3 = 0{,}2$. Тогда функции ценности для четырех альтернатив соответственно составят (следует учесть, что увеличение коэффициента вскрыши $K_в$ и содержания вредного компонента отрицательно влияют на ценность альтернатив, поэтому они учитываются со знаком "минус")

$$U_1 = -0{,}5 \cdot 3{,}7 + 0{,}3 \cdot 2{,}29 - 0{,}2 \cdot 1{,}64 = -1{,}491,$$
$$U_2 = -0{,}5 \cdot 4{,}0 + 0{,}3 \cdot 2{,}46 - 0{,}2 \cdot 1{,}64 = -1{,}594,$$
$$U_3 = -0{,}5 \cdot 4{,}1 + 0{,}3 \cdot 2{,}35 - 0{,}2 \cdot 1{,}52 = -1{,}685,$$
$$U_4 = -0{,}5 \cdot 3{,}9 + 0{,}3 \cdot 2{,}38 - 0{,}2 \cdot 1{,}54 = -1{,}616.$$

Как видно, наивысшую ценность имеет альтернатива 1. Отрицательные значения ценности не должны приниматься во внимание, ибо нас интересует только сравнительная ценность альтернатив. Наличие отрицательных значений еще раз показывает условность функции ценности, отсутствие у нее в большинстве случаев реального экономического и технологического смысла.

При аксиоматических методах "весовые" коэффициенты функции ценности определяются на основе предпочтений ЛПР.

На основе "весовых" коэффициентов, первоначально установленных ЛПР или принятых по каким-либо соображениям, подсчитываются функции ценности альтернатив, которые предъявляются ЛПР. Предполагается, что между функцией ценности и предпочтениями ЛПР имеется соответствие. На основе анализа ЛПР устанавливает предпочтения между критериями или их комбинациями, которые приобретают форму равенств или уравнений типа $K_i \geqslant K_j$ или $K_i \geqslant \sum_{j=1}^{L} K_j$.

Если решение полученной системы неравенств и уравнений единственное, то оно дает новые "весовые" коэффициенты. Если же система имеет множество решений, то вводится целевая функция, соответствующая наибольшему приближению к ранее имеющемуся решению или к чебышевской точке. Например, $\sum_{i=1}^{N} (w_i - w_i^{ст}) \to \min$ или $\sum_{i=1}^{N} |w_i - w_i^{ст}| \to \min$, где $w_i^{ст}$ – ранее установленные "веса".

В этом случае "весовые" коэффициенты определяются на основе решения задачи математического (нелинейного) программирования. Недостаток аксиоматического метода – возможность получения несовместной системы неравенств, отсутствии явной связи между предпочтениями ЛПР и полученным решением.

Покажем применение аксиоматического метода на рассматриваемом примере. Предположим, что проанализировав функции ценности для всех альтернатив при исходных "весовых" коэффициентах (0,5; 0,3; 0,2), ЛПР после получения первого решения изменил ранжировку критериев и указал, что наибольший "вес" должен иметь второй критерий, далее – первый. Желание минимального изменения ранее установленных "весов" (0,5; 0,3; 0,2) приводит к следующей задаче математического программирования

$$(w_1 - 0{,}5)^2 + (w_2 - 0{,}3)^2 + (w_3 - 0{,}2)^2 \to \min$$

или $|w_1 - 0{,}5| + |w_2 - 0{,}3| + |w_3 - 0{,}2| \to \min$

при ограничениях

$$\left.\begin{aligned} w_2 - w_1 &> 0, \\ w_1 - w_3 &> 0, \\ w_1 + w_2 + w_3 &= 1. \end{aligned}\right\}$$

Решение данной задачи (с округлением до 0,01) дает следующие результаты $w_1 = 0{,}39$; $w_2 = 0{,}41$; $w_3 = 0{,}2$.

В этом случае ценность альтернатив составит

$$U_1 = -0{,}39 \cdot 3{,}7 + 0{,}41 \cdot 2{,}29 - 0{,}2 \cdot 1{,}64 = -0{,}83,$$
$$U_2 = -0{,}39 \cdot 4{,}0 + 0{,}41 \cdot 2{,}46 - 0{,}2 \cdot 1{,}66 = -0{,}87,$$
$$U_3 = -0{,}39 \cdot 4{,}1 + 0{,}41 \cdot 2{,}35 - 0{,}2 \cdot 1{,}52 = -0{,}94,$$
$$U_4 = -0{,}39 \cdot 3{,}9 + 0{,}41 \cdot 2{,}38 - 0{,}2 \cdot 1{,}54 = -0{,}85.$$

Лучшим опять является вариант 1.

Предположим, что проанализировав полученные оценки альтернатив ЛПР, желая еще больше познать значение второго критерия эффективности, устанавливают, что его "весовой" коэффициент должен быть равен сумме "весовых" коэффициентов первого и третьего коэффициентов. Кроме того, "весовой" коэффициент первого критерия должен быть больше, чем третьего. "Весовые" коэффициенты достаточно определить с точностью до десятых. Учитывая, что

сумма "весовых" коэффициентов равна единице, получаем следующую систему уравнений и неравенств

$$\left.\begin{aligned} &w_2 = w_1 + w_3, \\ &w_1 - w_3 > 0{,}1, \\ &w_1 + w_2 + w_3 = 1. \end{aligned}\right\}$$

Данная система имеет единственное решение $w_1 = 0{,}3$; $w_2 = 0{,}5$; $w_3 = 0{,}2$.

При таких коэффициентах функции ценности альтернатив примут следующие значения

$$U_1 = -0{,}3 \cdot 3{,}7 + 0{,}5 \cdot 2{,}29 - 0{,}2 \cdot 1{,}64 = -0{,}293,$$
$$U_2 = -0{,}3 \cdot 4{,}0 + 0{,}5 \cdot 2{,}46 - 0{,}2 \cdot 1{,}66 = -0{,}302,$$
$$U_3 = -0{,}3 \cdot 4{,}1 + 0{,}5 \cdot 2{,}35 - 0{,}2 \cdot 1{,}52 = -0{,}359,$$
$$U_4 = -0{,}3 \cdot 3{,}9 + 0{,}5 \cdot 2{,}38 - 0{,}2 \cdot 1{,}54 = -0{,}288.$$

Лучшим в этом случае является вариант 4 плана, что соответствует предпочтению ЛПР.

Методы компенсации основаны на идее компромисса между критериями, уравновешивания ценности их оценок.

Наиболее широко используются два метода перехода от сравнения качеств по различным критериям к сравнению альтернатив в целом: построение кривых безразличия и сравнение разностей оценок альтернатив по критериям.

Кривыми безразличия называются гиперповерхности уровня функции ценности U – т.е. множество точек в многокритериальном пространстве, для которых ценность постоянна ($U = \text{const}$). Семейство всех кривых безразличия (для отдельных значений функции ценности) дает карту безразличия. Термин "кривая безразличия" объясняется тем, что ценность любых двух альтернатив A и B, лежащих на одной кривой, одинакова, $U(A) = U(B)$. При двух критериях оптимальности кривая безразличия – кривая в двумерном пространстве, для трех критериев – поверхность, при N критериях – гиперповерхность в N-мерном пространстве.

При наличии в многокритериальном пространстве гиперповерхностей (кривых) безразличия сравнение многокритериальных альтернатив не представляет труда, поскольку исходные точки и соответствующие им кривые можно упорядочить по ценности.

Построение кривых безразличия осуществляется следующим образом (рис. XIV.4).

1. Берется исходная точка P_1 со значениями критериев K_1^1 и K_2^1 (нижний индекс – номер критерия, верхний – исходной точки) и рассчитывается для нее значение функции ценности U_1.

2. Выбирается приращение ΔK_2 по критерию K_2 (положительное или отрицательное) и определяется значение критерия в новой точке $P_2\, K_2^2 = K_2^1 + \Delta K_2$.

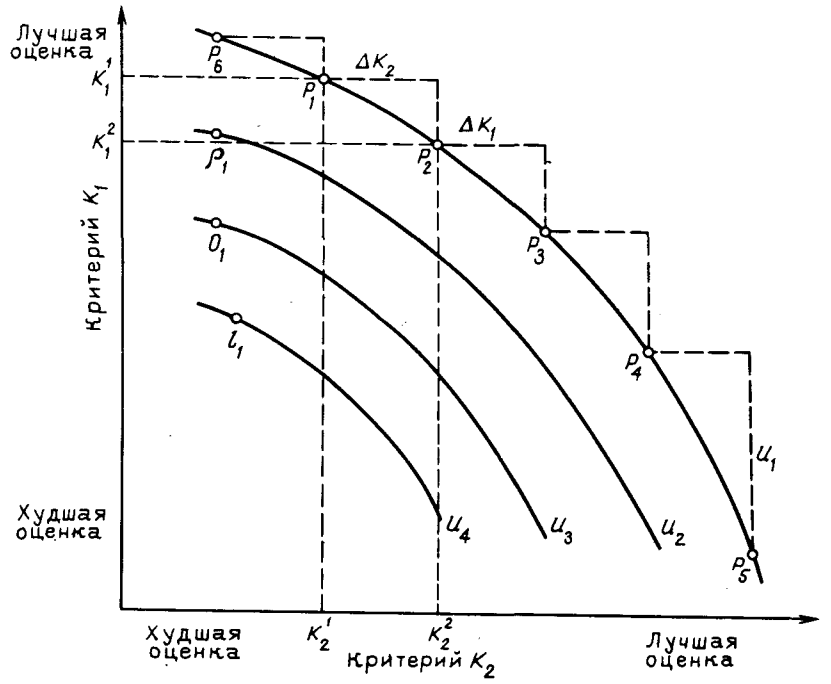


Рис. XIV.4. Кривые безразличия

3. Определяется значение K_1^2 по критерию K_1 таким образом, чтобы точка $P_2\,(K_1^2, K_2^2)$ была эквивалентна по ценности (полезности) точке $P_1\,(K_1^1, K_2^1)$.

4. Повторим последовательно операции 2 и 3, соединив полученные точки $P_1, P_2, P_3, P_4, P_5, P_6$, построим кривую безразличия.

5. Постепенно выбирая новые точки с другими значениями функции ценности (ρ_1, O_1, l_1) и используя процедуру 2–4, получаем карту безразличия.

Метод сравнения разностей оценок альтернатив основан на построении функции φ_1, определяющей влияние этой разности по критерию K_i на результат сравнения двух альтернатив. Альтернатива A считается предпочтительнее B, если

$$\sum_{i=1}^{N} \varphi_1\,(K_i^A - K_i^B) > 0.$$

Существенным недостатком методов количественной оценки многокритериальных альтернатив, основанных на построении функции ценности (полезности), является практически неограниченная возможность компенсации ухудшения качества по одному или ряду крите-

риев за счет его увеличения по другим, ибо одни и те же значения функции ценности получаются при различных комбинациях значений критериев. При этом может оказаться, что максимальное значение функции ценности имеет решение, противоречащее логике и здравому смыслу (например, при нулевых значениях отдельных критериев). Устранение указанного недостатка возможно путем введения нелинейных взаимоотношений между шкалами различных критериев (т.е. значение "весового" коэффициента зависит от величины критерия). Это приводит к нелинейным функциям ценности и усложняет решение проблемы в целом, особенно при большом количестве критериев.

Особую разновидность количественных методов решения многокритериальных задач представляют те, в которых ценность альтернатив определяется не на основе агрегирования оценок по отдельным критериям (построения функции ценности), а путем определения меры близости решения к некоему идеальному (так называемой точке идеала, или идеальной точке), т.е. решение должно обеспечивать наибольшее приближение к множеству одновременно недостижимых целей. В этом случае решается так называемая задача целевого программирования

$$d\,[K, b] \rightarrow \min, \qquad \text{(XIV.4)}$$

где K – векторная функция значения отдельных критериев $K_i (i = 1, N)$; b – N-мерный вектор, характеризующий идеальную точку, т.е. включающий желательные (идеальные) значения отдельных критериев; d – расстояние между K и b (т.е. между критериями, характеризующими полученное и идеальное решения), определенное на основании какой-либо метрики.

В задаче целевого программирования функционал (XIV.4) имеет содержательный смысл; он характеризует меру близости полученного решения к идеальному (т.е. расстояние между ними).

Параметры идеальной точки (т.е. желательные значения критериев) определяются ЛПР или вышестоящим органом. Часто в качестве координат идеальной точки берутся максимальные (или минимальные) значения отдельных критериев. В этом случае идеальная точка соответствует наилучшему достижению одновременно всех поставленных целей.

На рис. XIV.5 показана геометрическая интерпретация рассматриваемой задачи оценки четырех вариантов плана горных работ по трем критериям. Координатные оси соответствуют шкалам критериев K_B, α, β. В трехмерном пространстве показаны точки I–IV, соответствующие вариантам решения, и идеальная точка ($И$), координаты которой равны наилучшим значениям критериев. В трехмерном (по числу критериев) пространстве требуется определить, какая из точек I–IV расположена наиболее близко к точке $И$.

При определении расстояния d может быть также учтена относительная важность критериев, в этом случае в целевую функцию вво-

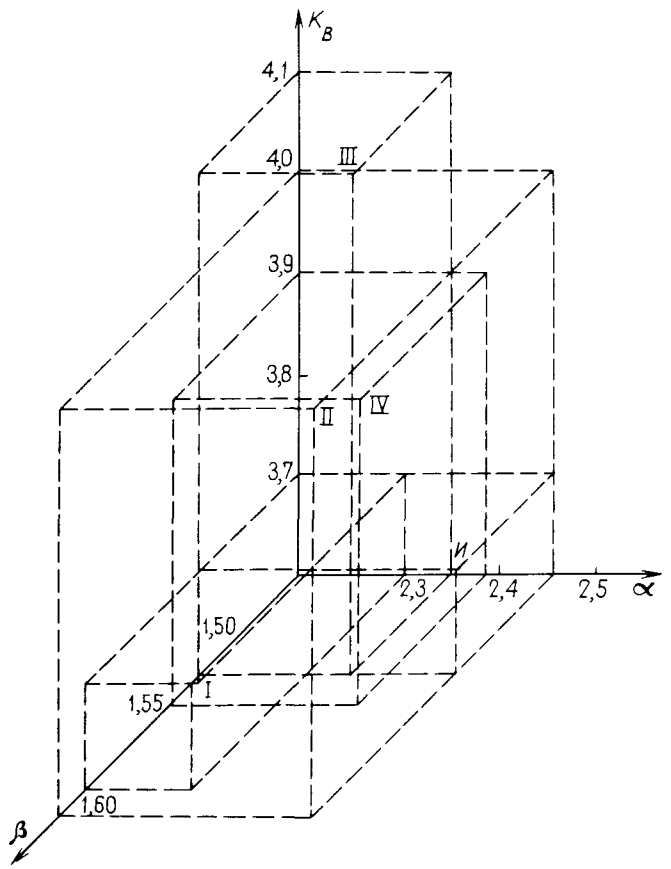


Рис. XIV.5. Геометрическая интерпретация задачи оценки четырех вариантов плана горных работ по трем критериям

дятся "весовые" коэффициенты $w_i (w_i = 1$, когда "веса" не вводятся). Введение "весового" коэффициента соответствует изменению масштаба шкал критериев. Следует отметить, что "весовые" коэффициенты в данном случае по смыслу несколько отличаются от "весовых" коэффициентов функции ценности U. Если в функциях ценности "весовые" коэффициенты соответствуют важности критериев, то в задаче целевого программирования они показывают значение расстояний, т.е. отклонений критерия от желательного значения (идеала).

Данный метод используется при различных метриках, функциях расстояния. В общем случае

$$d(K, b) = \left[\sum_{i=1}^{N} (K_i - b_i)^S \right]^{1/S} \rightarrow \min, \qquad \text{(XIV.5)}$$

где b_i – желательное значение i-го критерия; S – характеристика метрики.

При $S = 2$ вычисляется обычное эвклидово расстояние $d(K, b) = \sqrt{\sum_{i=1}^{N} w_i (K_i - b_i)^2}$, обычно при этом решают задачи квадратичного программирования

$$\sum_{i=1}^{N} w_i \Delta K_i^2 \to \min, \qquad \text{(XIV.6)}$$

где ΔK_i – разность идеального и реального значений по i-му критерию.

При $S = 1$ задача сводится к минимизации суммарного модуля относительных (с учетом "веса") отклонений.

При $S = \infty$ имеем равномерную метрику и задача целевого программирования сводится к минимизации максимального относительного отклонения, т.е. $\max w_i |\Delta K_i| \to \min$ – имеем минимаксную целевую функцию.

Перед решением задачи обычно производится нормирование значений критериев

$$K_i^{\text{н}} = \frac{K_i - K_i^{\min}}{K_i^{\max} - K_i^{\min}}, \qquad \text{(XIV. 7)}$$

где $K_i^{\max}$, $K_i^{\min}$ – соответственно максимальное и минимальное значения i-го критерия.

Знаменатель данной формулы, по существу, представляет длину шкалы, а числитель – положение критерия на ней.

Если в качестве идеальной точки используется точка с наилучшими значениями критерия, то нормированные значения ее координат равны 1 – для критериев, которые надо увеличивать, и 0 – для критериев, которые надо уменьшать.

Когда количество альтернатив ограничено и их параметры известны, задача целевого программирования может решаться без привлечения ЭВМ. Если же исходное множество альтернатив задано какой-либо математической моделью, то метод целевого программирования обычно используется в составе человеко-машинных процедур. Одним из диалоговых воплощений задач целевого программирования является траекторный подход, описываемый в § 6.

Проиллюстрируем использование методов целевого программирования для рассматриваемой задачи оценки четырех вариантов плана горных работ по трем критериям.

Прежде всего пронормируем значения критериев. Например, нормированные значения коэффициентов вскрыши соответственно по вариантам составят:

Таблица XIV. 1

Критерий	Варианты				
	I	II	III	IV	*И*
K_B, м³/т	0	0,75	1	0,5	0
α, %	0	1	0,35	0,53	1
β, %	0,86	1,0	0	0,14	0

Таблица XIV. 2

№ п/п	Вид критерия (метрика)	Варианты			
		I	II	III	IV
1	$\sqrt{\sum w_i (\Delta K_i)^2}$	0,81	0,61	0,71	0,47
2	$\sum_{1}^{3} w_i \lvert\Delta K_i\rvert$	0,67	0,425	0,625	0,415
3	$\min \max w_i \lvert\Delta K_i\rvert$	0,5	0,225	0,32	0,235

для I $\frac{3,7 - 3,7}{4,1 - 3,7} = 0$, для II $\frac{4,0 - 3,7}{4,1 - 3,7} = 0,75$ и т.д.

Здесь 3,7 и 4,0 значения критерия для вариантов I и II, а 4,1 и 3,7 соответственно максимальное и минимальное значения.

Значения нормированных критериев, в том числе и для идеальной точки (вариант *И*), приведены в табл. XIV. 1.

Предположим, что "весовые" коэффициенты расстояний для различных критериев соответственно равны 0,3; 0,5; 0,2. Значения расстояний при различных метриках приведены в табл. XIV. 2.

Как видно из данных табл. XIV. 2, при использовании эвклидовой метрики и при $S = 1$ (модульный критерий) лучшим является вариант IV, а при использовании равномерной метрики ($S = \infty$) (минимаксный критерий) вариант II.

§ 5. ТЕОРЕТИКО-МНОЖЕСТВЕННЫЕ МЕТОДЫ

Основная идея этой группы методов – использование бинарных соотношений между вариантами, определяющих условия превосходства ($>$) одной альтернативы над другой, их равнозначность ($=$) или несравнимость. Совокупность бинарных отношений между альтернативами может быть отображена ориентированным графом $G(U, V)$, где U – множество вершин графа, а V – множество дуг графа. При

этом дуги графа являются ориентированными и направлены от вершины, отображающей вариант с лучшей оценкой (предпочтительной), к вершине с более низкой оценкой. Если варианты эквивалентны, то между вершинами им соответствующими, проходят две разнонаправленных дуги.

Покажем принцип построения графа предпочтений. Предположим, требуется упорядочить альтернативы и выбрать лучшие из них в задаче определения места строительства горного предприятия. Сопоставляются восемь вариантов (I–VIII) по четырем основным критериям: K_1 – годовая производительность труда по готовой продукции (металлу), т/чел.; K_2 – рентабельность, %; K_3 – коэффициент извлечения металла из недр; K_4 – обеспеченность района кадрами. Причем, три первых критерия – количественные, а четвертый – качественный и имеет три градации (неудовлетворительную, удовлетворительную и хорошую).

Характеристики восьми сопоставляемых вариантов проектов (отличающихся технологией добычи и обогащения, мощностью по горной массе, структурой комплексной механизации и т.д.) по четырем рассматриваемым критериям приведены в табл. XIV. 3.

Построим графы предпочтений для каждого из критериев, соответственно G_1, G_2, G_3, G_4.

В графе G_1 стрелка направлена от варианта II к I, так как вариант II по первому критерию (производительность труда) предпочтительнее I (111 > 110, т.е. $K_1^{II} > K_1^{I}$, верхний индекс отражает номер варианта, а нижний – критерия), между II и IV вариантами две разнонаправленные стрелки, так как по первому критерию варианты равноценны ($K_1^{II} = K_1^{IV} = 111$).

Графы показаны на рис. XIV. 6 и дают четкое отношение (полный квазипорядок) между альтернативами по каждому из критериев. Если каждой вершине графа приписать значение соответствующего критерия, то получим всю имеющуюся информацию о задаче. С учетом этой

Таблица XIV. 3

Вариант	Критерии			
	K_1	K_2	K_3	K_4
I	110	8,30	82	Хорошая
II	111	7.96	80	Неудовлетворительная
III	109	8,32	80	Удовлетворительная
IV	111	8.28	79	Удовлетворительная
V	115	8,46	80	Хорошая
VI	110	8.30	81	Неудовлетворительная
VII	119	8.44	78	Хорошая
VIII	112	8.05	79	Удовлетворительная

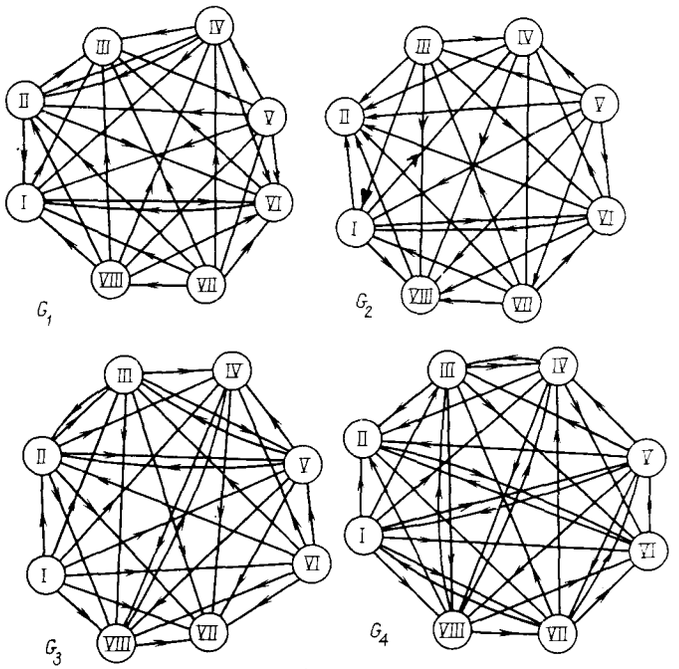


Рис. XIV. 6. Графы взаимосвязи альтернатив G по различным критериям K: G_1 – по K_1; G_2 – по K_2; G_3 – по K_3; G_4 – по K_4

информации необходимо на основе синтеза оценок по всем критериям получить новый граф, показывающий отношения между всеми имеющимися альтернативами. Предпочтения между вариантами легко выявить, если оценки одного варианта лучше оценок другого варианта по всем критериям (или по некоторым критериям, по крайней мере, не хуже), т.е. один из вариантов доминирует над другим.

В рассматриваемом примере вариант I лучше VI (по двум первым критериям оценки равны, а по третьему и четвертому оценки варианта I превосходят оценки VI), вариант V доминирует над II, III, IV, VIII, т.е. можем построить граф с дугами между I–VI, V–II, III, IV и VIII вариантами (рис. XIV. 7).

Полученный граф G, как и обычно, не является полным. Проблема состоит в том, чтобы на основании прямых предпочтений ЛПР или принципов, сформулированных им, дополнить граф G, чтобы он находился как можно больше в соответствии с исходными графами по отдельным критериям.

Существует несколько подходов к решению этой задачи. Сравнение двух альтернатив может производиться непосредственно ЛПР или при помощи аналитического метода. Ряд методов основывается на

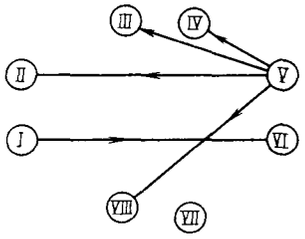


Рис.XIV.7. Обобщенный граф взаимосвязи альтернатив

групповых экспертных оценках и будет рассматриваться ниже (задача ранжирования многокритериальных объектов).

В самых простых методах ЛПР непосредственно сопоставляет два варианта и отвечает на вопрос, какой из них лучше. При этом возможен ответ об эквивалентности вариантов или даже несравнимости. Если появляются противоречия в ответах ЛПР, то часть альтернатив рассматривается им повторно. В результате предпочтений, указанных ЛПР, строится ориентированный граф, не имеющий контуров (циклов) при отсутствии противоречий в ответах ЛПР.

Другим методом, где сравнение альтернатив осуществляется непосредственно ЛПР, является "Запрос", в котором используется более корректный способ получения информации от ЛПР, включающий замкнутую процедуру проверки ответов ЛПР на непротиворечивость. В этом методе рассматриваются все пары, что приводит к увеличению избыточной информации и дает возможность проверки ее на непротиворечивость. ЛПР первоначально предъявляются альтернативы у так называемых опорных ситуаций (отсюда и название метода – замкнутые процедуры у опорных ситуаций – "запрос"). Опорными являются контрастные ситуации, имеющие только лучшие или только худшие оценки. ЛПР сопоставляет две альтернативы, у которых различаются два критерия, а остальные близки к опорным ситуациям.

Замкнутая процедура и проверка на непротиворечивость возникают из-за избытка информации. Например, если получено, что альтернатива A лучше B, а B лучше C, то можно сделать вывод, что A лучше C. В методе же сравнение A и C производится также ЛПР. Полученное от него сообщение проверяется на непротиворечивость с информацией, полученной ранее. При проявлении противоречий предпочтения ЛПР уточняются и при необходимости корректируются. Эффективность метода зависит от количества критериев и градации шкал оценок. При 7–8 критериях (или менее) и 3–4 градациях на шкале критериев может быть получена непротиворечивая информация. При этом, если при сравнении ряда альтернатив не удается устранить противоречия, то они считаются независимыми.

Аналитический метод выявления предпочтений используется в методах порогов сравнимости, включающих ряд модификаций метода "Электра", разработанного французским ученым Б. Руа.

В этих методах при сравнении двух вариантов A и B все критерии делятся на две группы: по которым вариант A превосходит B и по которым вариант A хуже B.

При этом по тем критериям, где вариант A превосходит B, подсчитывается индекс согласия (p), а по второй группе критериев индекс несогласия (q). Для сопоставления вариантов приняты (установлены) пороговые значения критериев p и q, соответственно, $p_{гр}$ и $q_{гр}$. Считается, что вариант A превосходит B, если индекс согласия $p > p_{гр}$ (т.е. вариант A превосходит B по большинству критериев с учетом их относительного "веса"), а индекс несогласия $q \leqslant q_{гр}$ (т.е. разница между остальными критериями незначительна).

Индекс согласия вычисляется по формуле

$$C_{AB} = \sum_{A} \pi_A , \qquad \text{(XIV.8)}$$

где π_A – относительный "вес" критериев, по которым вариант A превосходит B.

Как видно из формулы (XIV.8), индекс согласия изменяется от 0 до 1 и равен 1, когда вариант A по всем критериям превосходит вариант B. При замене какого-либо одного критерия их группой с тем же относительным "весом" значение индекса согласия не меняется.

При расчете индекса согласия для критериев в порядковых шкалах ЛПР устанавливаются оценки в баллах.

Индекс согласия вычисляется для всех пар альтернатив, в результате чего составляется матрица индексов согласия ($\|C_{ij}\|$, $i = \overline{1, I}$, $j = \overline{1, J}$ – изучаемые альтернативы).

Рассчитаем индексы согласия и составим их матрицу, для задачи сопоставления восьми альтернатив по четырем критериям, условия которой приведены в табл. XIV.3. Предположим, что ЛПР указан "вес" критериев $K_1 = 5$; $K_2 = 8$; $K_3 = 4$; $K_4 = 2$, тогда относительный "вес" критериев составит соответственно $\pi_1 = 0{,}3\left(\frac{6}{8+4+6+2}\right)$, $\pi_2 = 0{,}4$; $\pi_3 = 0{,}2$; $\pi_4 = 0{,}1$. Для качественного второго критерия установлены следующие количественные оценки: хорошая – 4; удовлетворительная – 3; неудовлетворительная – 2. Тогда индекс согласия между первой и второй альтернативами определится следующим образом. Вариант I превосходит II по второму, третьему и четвертому критериям, следовательно $C_{12} = \pi_2 + \pi_3 + \pi_4 = 0{,}4 + 0{,}2 + 0{,}1 = 0{,}7$.

Матрица индексов согласия имеет вид

$$\|C_{ij}\| = \begin{bmatrix} – & 0,3 & 0,4 & 0,4 & 0,7 & 0.0 & 0,7 & 0,3 \\ 0,7 & – & 0,5 & 0,5 & 0,8 & 0,6 & 0,8 & 0,8 \\ 0,6 & 0,3 & – & 0,3 & 0,8 & 0,5 & 0,8 & 0,3 \\ 0,6 & 0,2 & 0,6 & – & 1,0 & 0,6 & 0,8 & 0,3 \\ 0,2 & 0,0 & 0,5 & 0,0 & – & 0,2 & 0,4 & 0,0 \\ 0,3 & 0,3 & 0,5 & 0,4 & 0,8 & – & 0,8 & 0,4 \\ 0,2 & 0,2 & 0,2 & 0,2 & 0,6 & 0,2 & – & 0,2 \\ 0,7 & 0,2 & 0,6 & 0,4 & 1,0 & 0,6 & 0,8 & – \end{bmatrix}.$$

Так как при расчете индексов согласия равенство критериев не учитывается, то для доминирующих вариантов индекс согласия не всегда равен 1, доминирование одного варианта над другим можно определить по нулевому значению индекса согласия. Индекс согласия между I и VI вариантами – 0,3, но наличие нулевого индекса согласия между VI и I вариантами свидетельствует о доминировании варианта I над VI. Действительно, по первому и второму критерию они равнозначны, а по третьему и четвертому – вариант I предпочтительнее VI.

Наряду с индексом согласия, рассчитывается индекс несогласия, который характеризует несоответствие гипотезы о превосходстве варианта A над B. Индекс несогласия служит для того, чтобы исключить варианты с большой разницей значений критериев вариантов A и B, по которым B предпочтительнее A. Индекс несогласия рассчитывается как относительная разница интервальных значений критерия с учетом их "веса". Индекс несогласия рассчитывается по всем критериям, по которым вариант A не превосходит B и максимальный из них берется в качестве индекса несогласия между этими вариантами, т.е.

$$d_{AB} = \max \pi_B \frac{K_i^B - K_i^A}{d_l}, \qquad \text{(XIV.9)}$$

где K_l^A, K_l^B – оценка вариантов A и B по l-му критерию; π_B – относительный "вес" критериев, по которым вариант B превосходит A; d_l – максимальная разница оценок по шкале l-го критерия ($d_l = \max \Delta K_l$), т.е. длина шкалы. Индекс несогласия изменяется от 0 до 1 и равен 0, когда вариант A по всем критериям превосходит B.

Рассчитаем индексы несогласия и составим их матрицу для рассматриваемой задачи. Индекс несогласия между III и VI вариантами подсчитывается по первому и третьему критерию, где вариант IV предпочтительнее III.

Индекс несогласия по первому критерию

$$d_{36}^1 = \pi_1 \frac{K_1^6 - K_1^3}{d_1} = 0,3 \frac{110 - 109}{119 - 109} = 0,03,$$

где 0,3 – относительный "вес" первого критерия; 110, 109 – соответственно оценки вариантов VI и III по первому критерию; 119, 109 – максимальная и минимальная оценки по варианту I.

Индекс несогласия между вариантами III и VI по третьему критерию $d_{36}^3 = 0,2 \frac{81 - 80}{82 - 78} = 0,05$.

Максимальный индекс из двух рассчитанных значений (0,3; 0,05) равен 0,05, который и принимается за индекс несогласия между III и VI вариантами.

Матрица индексов несогласия имеет вид

$$\|d_{ij}\| = \begin{bmatrix} – & 0,27 & 0,10 & 0,15 & 0,10 & 0,10 & 0,20 & 0,15 \\ 0,03 & – & 0,06 & 0,05 & 0,00 & 0,03 & 0,10 & 0,05 \\ 0,02 & 0,27 & – & 0,05 & 0,00 & 0,05 & 0,10 & 0,27 \\ 0,03 & 0,26 & 0,06 & – & 0,00 & 0,05 & 0,05 & 0,23 \\ 0,16 & 0,40 & 0,18 & 0,14 & – & 0,15 & 0,10 & 0,33 \\ 0,00 & 0,27 & 0,05 & 0,10 & 0,05 & – & 0,15 & 0,20 \\ 0,27 & 0,38 & 0,30 & 0,24 & 0,12 & 0,27 & – & 0,31 \\ 0,06 & 0,07 & 0,09 & 0,03 & 0,00 & 0,06 & 0,06 & – \end{bmatrix}.$$

При большом числе критериев индекс несогласия, полученный по формуле (XIV.9), не отражает полностью всей картины о соотношении вариантов A и B. Одна и та же величина индекса несогласия по другим критериям близка к нулю или максимальному значению. В связи с этим в ряде случаев целесообразно брать не максимальный индекс, а какой-либо другой (например, 2-й или 3-й) из всего ряда индексов несогласия, расположенных в убывающем порядке. Кроме того, можно использовать также сумму индексов несогласия. В этом случае, естественно, граничное значение индекса несогласия должно быть принято бо́льшим. Построив матрицы индексов согласия и несогласия, можно выявить отношения превосходства между альтернативами. Как ранее указывалось, альтернатива A считается лучшей по сравнению с B, если $C_{AB} \geqslant p_{\text{гр}}$, а $d_{AB} \leqslant q_{\text{гр}}$.

Для выявления превосходства полученные матрицы индексов согласия и несогласия несколько видоизменяют, оставляя в них только те элементы, которые соответствуют пороговым значениям, устанавливаемым ЛПР. Допустим, что для рассматриваемого примера граничная величина индекса согласия 0,7, а индекса несогласия 0,05. Тогда матрицы индексов примут вид

$$C_{ij} = \begin{bmatrix} – & – & – & – & 0,7 & – & 0,7 & – \\ 0,7 & – & – & – & 0,8 & – & 0,8 & 0,8 \\ – & – & – & – & 0,8 & – & 0,8 & 0,8 \\ – & – & – & – & 1,0 & – & 0,8 & – \\ – & – & – & – & – & – & – & – \\ – & – & – & – & 0,8 & – & 0,8 & – \\ – & – & – & – & – & – & – & – \\ 0,7 & – & – & – & 1,0 & – & 0,8 & – \end{bmatrix}.$$

$$d_{ij} = \begin{bmatrix} - & - & - & - & - & - & - & - \\ 0{,}03 & - & - & 0{,}05 & 0 & 0{,}03 & - & 0{,}05 \\ 0{,}02 & - & - & 0{,}05 & 0 & 0{,}05 & - & - \\ 0{,}03 & - & - & - & 0 & 0{,}05 & 0{.}05 & - \\ - & - & - & - & - & - & - & - \\ 0 & - & 0{,}05 & - & 0{,}05 & - & - & - \\ - & - & - & - & - & - & - & - \\ - & - & - & 0{,}03 & 0 & - & - & - \end{bmatrix}$$

Прочерки в матрицах соответствуют элементам, индексы которых не соответствуют пороговому значению.

По полученным матрицам строим (или дополняем имеющийся по отношениям доминирования) граф, показывающий отношение превосходства между вариантами. Анализ матриц показывает следующие отношения превосходства между вариантами: I > II ($C_{12} = 0{,}7 = p_{\text{гр}} = 0{,}7$; $d_{12} = 0{,}03 < q_{\text{гр}} = 0{,}05$); V > II; V > III; V > IV; V > VI; V > VIII; VII > IV; VIII > 2 (знак > читается как предпочтительнее чем или лучше чем). В то же время, хотя по индексу согласия, например, вариант VII удовлетворяет условиям превосходства над вариантами I, II, III, VI, VIII, по установленному значению порога несогласия наличие этого отношения не подтверждается, так как соответственные индексы несогласия 0,20; 0,10; 0,10; 0,15; 0,06 превышают $q_{\text{гр}} = 0{,}05$. Граф, соответствующий выявленным отношениям превосходства альтернатив, показан на рис. XIV. 8.

Следующий этап – анализ графа для выделения на его основе лучших вариантов. Для этого, если известно, что альтернатива A превосходит B, то последняя может быть исключена из рассмотрения. Таким образом, из рассмотрения могут быть исключены все вершины, которые характеризуют варианты, по отношению к которым есть лучшие. Иначе говоря, все множество вариантов разбивается на два подмножества Ω_1 и Ω_2, в одном из них (Ω_1) содержатся варианты лучшие, а в другом Ω_2 – исключаемые из рассмотрения. Подмножество Ω_1 обладает свойствами внешней и внутренней устойчивости. Свойство внешней устойчивости означает, что для любого из исключенных вариантов в подмножестве Ω_1 остается хотя бы один вариант, который превосходит исключенный, а свойство внутренней устойчивости означает, что никакой из оставляемых в подмножестве Ω_1 вариантов не имеет превосходства перед другим вариантом. Оставшееся подмножество вариантов Ω_A называется ядром графа.

Граф может иметь несколько ядер, конечный ориентированный граф без циклов (контуров) имеет одно ядро. Если в графе обнаруживаются циклы (замкнутые отношения превосходства, например, $A > B$; $B > C$; $C > A$), то это означает равноценность альтернатив, которые входят в контур. Альтернативы, входящие в контур, могут быть заменены одна другой, а если требуется, то дополнительно уточняют

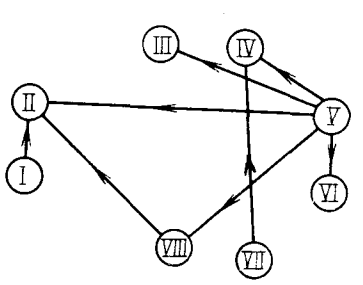


Рис.XIV.8. Граф, соответствующий выявленным отношениям превосходства альтернатив (при $p_{\text{ср}} \geqslant 0{,}7$, $q_{\text{гр}} \leqslant 0{,}5$)

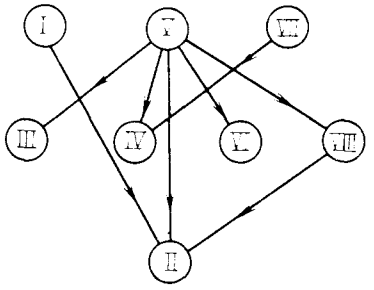


Рис.XIV.9. Упорядоченный граф (при $p_{\text{ср}} \geqslant 0{,}7$, $q_{\text{ср}} \leqslant 0{,}5$)

отношения между ними. Исключение контура осуществляется одновременно с построением ядра. Для этого последовательно исключают вершины графа, в которые входят дуги, показывающие отношение превосходства. Упорядочение графа и выделение ядра могут быть также осуществлены методом вычеркивания дуг, используемом при нумерации вершин сетевого графика (см. гл X).

Выделенное ядро всегда содержит лучшую альтернативу. Выполнение описанных процедур для графа рассматриваемой задачи (см. рис. XIV. 8) приводит к упорядоченному графу (рис. XIV. 9), из которого видно, что его ядро составляет альтернативы I, V, VII.

Альтернативы, попадаемые в ядро, являются несравнимыми с точки зрения установленных порогов *согласия и несогласия*; для выяснения отношений между ними нужна дополнительная информация. Если ядро содержит один элемент и задача состояла в том, чтобы найти наилучший вариант из всего множества допустимых Ω, то она решена. Если в ядре содержится несколько вершин или нужно не только выделить наилучший вариант, но произвести их полное упорядочение (ранжировку), описанный метод может быть применен последовательно.

Для этого изменяется порог сравнимости, т.е. уменьшается индекс согласия и увеличивается индекс несогласия. Ядро, соответствующее новому порогу сравнимости, в общем случае содержит меньшее количество вершин – несравнимых элементов. Процесс изменения порога и получения ядер графа с уменьшающимся количеством вершин продолжается до тех пор, пока в ядре не останется одна вершина (или их требуемое количество).

Последовательно выделяя лучшую вершину или строя упорядоченный граф, можно получить полное ранжирование рассмотренных альтернатив.

Таким образом, хотя метод "Электра" в основном применяется для выделения лучшего варианта (или группы лучших), при его пос-

ледовательном использовании может быть решена задача ранжирования, т.е. упорядочены варианты по качеству.

Продолжим решение рассматриваемой задачи. При индексе согласия 0,7 и несогласия 0,05 в ядро графа попало три элемента. Изменим порог сравнимости, уменьшив индекс согласия до 0,6 и увеличив индекс несогласия до 0,1. Построим матрицы индексов согласия и несогласия, включив в них элементы, удовлетворяющие установленным требованиям ($p_{гр} \geqslant 0{,}6$; $q_{гр} \leqslant 0{,}1$)

$$C_{ij} = \begin{bmatrix} – & – & – & – & 0{,}7 & – & 0{,}7 & – \\ 0{,}7 & – & – & – & 0{,}8 & 0{,}6 & 0{,}8 & 0{,}8 \\ 0{,}6 & – & – & – & 0{,}8 & – & 0{,}8 & – \\ 0{,}6 & – & 0{,}6 & – & 1{,}0 & 0{,}6 & 0{,}8 & – \\ – & – & – & – & – & – & – & – \\ – & – & – & – & 0{,}8 & – & 0{,}8 & – \\ – & – & – & – & 0{,}6 & – & – & – \\ 0{,}7 & – & 0{,}6 & – & 1{,}0 & 0{,}6 & 0{,}8 & – \end{bmatrix}.$$

$$d_{ij} = \begin{bmatrix} – & – & 0.10 & – & 0{,}10 & 0{,}10 & – & – \\ 0{,}03 & – & 0{,}06 & 0{,}05 & 0 & 0{,}03 & 0{,}10 & 0{,}05 \\ 0{,}002 & – & & 0{,}05 & 0 & 0{,}05 & 0{,}10 & – \\ 0{,}03 & – & 0.06 & – & 0 & 0{,}05 & 0{,}05 & – \\ – & – & – & – & – & – & 0{,}10 & – \\ 0 & – & 0{,}05 & 0{,}10 & 0{,}05 & – & – & – \\ – & – & – & – & – & – & – & – \\ 0{,}06 & 0{,}07 & 0{,}09 & 0{,}03 & 0 & 0{,}06 & 0{,}06 & – \end{bmatrix}.$$

По полученным матрицам строим граф предпочтения (рис. XIV. 10). Как видно из рис. XIV. 10 снижение порога сравнимости увеличило связность (количество ребер графа), дополнительно выявлены отношения между вершинами I–III, I–IV, I–V, V–VI, II–VII, III–VII, III–VIII, IV–VI, VI–VIII, VII–VIII.

Описанным выше методом проведем упорядочение вершин графа и выделим его ядро. Новое упорядоченное изображение графа показано на рис. XIV. 11. Как видно из рис. XIV. 11, ядро графа содержит две вершины V и VII, следовательно, для выделения лучшего варианта, необходимо их сопоставить детальнее. Кроме того, для полного упорядочения необходимо выявить отношения между вершинами I–VII (если V > VII), III–VI, IV–VIII (если VIII > IV) и VII–VI (если I > VII, а VI > III).

Сопоставляя варианты V и VII, заметим, что V превосходит VII по двум критериям (K_2, K_3), по одному равен (K_4) и хуже по критерию (K_1), причем индекс согласия между вариантами 0,6; а индекс несогласия 0,15. То есть можно допустить, что вариант V лучше VII (вообще наилучший). Сопоставим I и VII варианты: вариант VII превосходит I по первому и второму критериям, равнозначен по четвер-

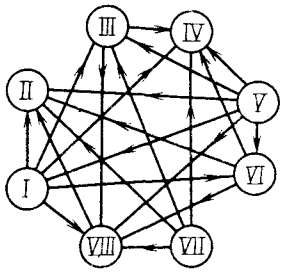


Рис. XIV. 10. Граф предпочтений альтернатив (при изменении порога сравнимости $p_{гр} \geqslant 0{,}6$; $q_{гр} \leqslant 0{,}1$)

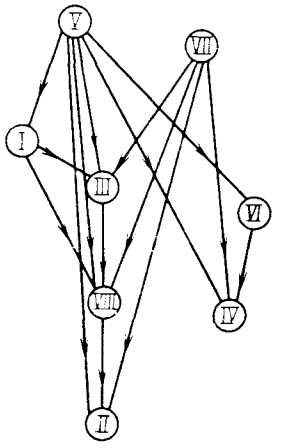


Рис. XIV. 11. Упорядоченный граф предпочтений (при $p_{гр} \geqslant 0{,}6$; $q_{гр} \leqslant 0{,}1$)

тому, и лишь уступает по третьему. Индексы согласия и несогласия составляют соответственно 0,7 и 0,2. Подняв предел индекса несогласия, можно допустить, что вариант VII превосходит I. Аналогичный, но более подробный анализ показывает практическую равноценность вариантов III и VI, некоторую предпочтительность варианта IV над VIII (он существенно лучше по критерию K_2, равноценен по K_3 и K_4 и лишь несколько уступает по K_1).

В соответствии с изложенным установлена следующая последовательность альтернатив по их предпочтительности: V > VII > I > III=VI > > IV > VIII > II.

Для сопоставления результатов и принятия окончательного решения построим функцию ценности и определим эвклидово расстояние до идеальной точки, координаты которой равны наилучшим значениям отдельных критериев (119; 8,46; 82; 4).

Так как шкалы критериев резко отличаются, функцию ценности определим для нормированных значений критериев.

Подсчитанные нормированные значения критериев сведены в табл. XIV. 4.

Значения функции ценности (XIV.1) и эвклидова расстояния (XIV.6), подсчитанные с "весовыми" коэффициентами 0,4; 0,3; 0,2; 0,1 и использованными в методе "Электра", приведены ниже.

Вариант	Функция ценности	Эвклидово расстояние
I	0,544	0,595
II	0,180	0,812
III	0,366	0,710
IV	0,218	0,817
V	0,740	0,336
VI	0,394	0,680
VII	0,788	0,445
VIII	0,286	0,752

Таблица XIV.4

Вариант	Критерии			
	K_1	K_2	K_3	K_4
I	0,1	0,68	1,0	1,0
II	0,2	0	0,5	0,0
III	0	0,72	0,5	0,5
IV	0	0,56	0,25	0
V	0,6	1,0	0,5	1,0
VI	0,1	0,68	0,75	0
VII	1,0	0,96	0	1,0
VIII	0,3	0,22	0,25	0,5

Сопоставим решения, полученные различными методами. В табл. XIV.5 приведена ранжировка вариантов, т.е. они расположены в порядке их предпочтительности.

Таблица XIV.5

Метод решения	Ранги							
	1	2	3	4	5	6	7	8
"Электра"	V	VII	I	VI	III	IV	VIII	II
Агрегация критериев (функция ценности)	VII	V	I	VI	III	VIII	IV	II
Целевое программирование (Эвклидово расстояние)	V	VII	I	VI	III	VIII	IV	II

Анализ полученных ранжировок показывает, что в принципе они совпадают. При необходимости можно дополнительно проанализировать варианты V и VII (ибо по функции ценности вариант VII превосходит V) и варианты IV и VIII, где решения, полученные методом "Электра", не совпадают с решением, полученным другими методами.

§ 6. ЧЕЛОВЕКО-МАШИННЫЕ ПРОЦЕДУРЫ ПРИНЯТИЯ РЕШЕНИЙ

В последние годы для решения многокритериальных задач математического программирования применяется подход, основанный на идее выявления предпочтений одновременно с исследованием допустимого множества действий для отыскания эффективных решений. Средством реализации этого подхода являются человеко-машинные (ЧМ) процедуры.

ЧМ-процедура принятия решений представляет собой циклический процесс сотрудничества человека и ЭВМ. Цикл состоит из фазы анализа и принятия решений (постановка задачи), выполняемой человеком, и фазы оптимизации (поиск решения и вычисление его характеристик), реализуемой машиной. В процессе взаимодействия человек проясняет характерные черты задачи, выявляет или уточняет свои предпочтения и в результате сообщает дополнительную информацию, благодаря которой машина вырабатывает все более совершенные решения. Процесс заканчивается, когда ЭВМ вырабатывает приемлемое для ЛПР решение или когда ЛПР убеждается в нецелесообразности дальнейших попыток получить лучшее решение при данной модели.

ЧМ-процедура состоит из чередующихся фаз анализа и оптимизации. Фаза может включать несколько шагов.

Фаза оптимизации: ЭВМ определяет новую оценку "весов" и (или) формирует новую область допустимых решений, используя полученную от ЛПР на предыдущем шаге информацию; вычисляет соответствующее новым данным решение (вместе с характеризующим его вектором); вырабатывает вспомогательную информацию.

Фаза анализа: ЛПР оценивает предъявленное решение (или несколько решений) и определяет, является ли оно приемлемым. Если да, то процедура окончена, в противном случае анализирует вспомогательную информацию; сообщает машине дополнительную информацию, с помощью которой можно вычислить новое решение.

ЧМ-процедуры, созданные для решения описанных выше задач, отличаются друг от друга содержанием и способами выполнения каждого из указанных шагов.

В качестве отправной точки при разработке этих процедур обычно выступает либо задача определения оптимального компромисса между частными целями, либо специфический способ поиска компромисса – включение в этот процесс человека. В совокупности диалоговых алгоритмов соответственно выделяются два класса процедур.

К 1-му классу относятся процедуры, организующие ЧМ-диалог с целью описания допустимого плана, оптимизирующего генеральную цель рассматриваемой системы. В основе процедур – математическая модель предпочтений и согласованный с ней оптимизационный алгоритм, обеспечивающий получение оптимального решения. Локальная информация о предпочтениях, необходимая для его выполнения, может быть получена у ЛПР, принимающего участие в диалоге. Эти процедуры называют нормативными.

Ко 2-му классу относятся процедуры, в основу которых положены алгоритмы дескриптивной информации, преследующие цель определения приемлемого решения, они непосредственно учитывают особенности поведения человека и поэтому значительно более удобны для ЛПР.

Диалоговые алгоритмы нормативного типа опираются на гипотезу о возможности описания предпочтений ЛПР целевой функцией $U(U_1, \ldots, U_m)$ (в явном виде, естественно, не заданной). Принятая гипотеза с формальной точки зрения позволяет представить задачу выбора решений в виде $U(\varphi(x)) \to \max$; $x \in X$. Поскольку оптимум заведомо является эффективным планом, задача может быть сформулирована следующим образом

$$U(u) \to \max, \; u \in E\,(\varphi(x)).$$

В основе ЧМ-процедур лежит экономико-математическая модель $P(a)$ вида

$$P(a): \varphi_j\,(x) \geqslant \Phi_j\,(a, Z), \; j = \overline{1, m}, \; x \in X,$$

являющаяся результатом параметрической скаляризации задачи (x, φ). Карта границы Парето $M = (A, \sigma)$, порожденная этой моделью, позволяет перенести поиск решения на множество параметров A и, в свою очередь, преобразовать задачу к виду

$$V(a) = U(\sigma(a)) \to \max, \; a \in A.$$

Выбирая тот или иной алгоритм решения этой задачи, получим формальную основу для организации ЧМ-диалога. Модель распределяет функции в этом диалоге следующим образом: экономико-математическая модель со скалярным критерием $P(a)$ используется для вычисления эффективных планов и перемещения по границе Парето путем корректировки параметров $a \in A$, а ЛПР направляет это перемещение, руководствуясь неформализованными представлениями о генеральной цели.

Метод покоординатной оптимизации. Согласно этому методу для отыскания максимума функции нескольких переменных $V(a)$ вначале фиксируются все компоненты вектора a, кроме первой, по которой осуществляется одномерная оптимизация, затем значение первой компоненты a_1 "замораживается" на оптимальном уровне, освобождается следующая переменная a_2 и т.д.; после очередной оптимизации всех составляющих a алгоритм возвращается к первой из них и циклически возобновляет работу. Соответствующая ЧМ-процедура опирается на использование скаляризации ограничениями (метод ограничений), так что поочередной оптимизации подлежат компоненты вектора правых частей $a = (a_1, \ldots, a_{m-1})$.

Градиентный метод. Этот метод предусматривает движение из текущей точки в направлении $h(a^{(i)}) = \operatorname{grad} V(a^{(i)})$ градиента максимизируемой функции $V(a)$ в точке $a^{(i)}$. Последовательность итераций строится следующим образом:

$$a^{(i+1)} = a^{(i)} + \lambda^{(i)} h\,(a^{(i)}),$$

где $\lambda^{(i)}$ – величина шага в направлении градиента.

Такой порядок корректировки параметров оправдывается известным свойством градиента, согласно которому вектор grad $V(a)$ указывает направление быстрейшего роста предпочтений, индуцированных на множестве A картой M (предполагается, что используемая карта порождена ассортиментной скаляризацией или методом ограничений).

Градиент функции $V(a) = U(\sigma(a))$, представляющий собой композицию целевой функции U и параметризации σ, рассчитывается по правилам дифференцирования следующей функции, согласно которому

$$h_j\,(a) = \frac{dV}{da_j}\,(a) = \sum_{k=1}^{m} \frac{dU}{dU_k}\Big|_{\sigma(a)} \frac{d\sigma_k\,(a)}{da_j}.$$

Таким образом, для построения вектора $h(a)$ необходимо располагать:

градиентом целевой функции

$$\operatorname{grad} U\Big|_{\sigma(a)} = \left(\frac{dU}{dU_1}\,(\sigma(a)), \ldots, \frac{dU}{dU_m}\,(\sigma(a))\right);$$

частными производными $\dfrac{d\sigma_k}{da_j}$ компонент параметризующего отображение σ в точке a.

Информацию о градиенте целевой функции дает ЛПР. С этой целью определяют при взаимодействии с ЛПР предельные нормы замещения $\Delta_{ij}(\sigma(a))$ одного из критериев, например первого $(i = 1)$, всеми остальными $(j = 2, \ldots, m)$ в точке $\sigma(a)$. В качестве оценки направления $\operatorname{grad} U\Big|_{\sigma(a)}$ используют вектор $\eta = \eta(\sigma(a)) = (1, \Delta_{21}(\sigma(a)), \ldots, \Delta_{m1}(\sigma(a)))$.

Производные же $\dfrac{d\sigma_k}{da_j}\,(a)$ рассчитываются согласно формулам

$$\frac{d\sigma_i}{da_j}\,(a) = \frac{d\Phi_i}{da_j}\,(a, Z(a)) - \left[\sum_{k=1}^{m} P_k\,(a)\,\frac{d\Phi_k}{da_j}\,(a, Z(a))\right]$$

$$\frac{d\Phi_i}{dZ}\,(a, Z(a)); \qquad \text{(XIV. 10)}$$

$$\frac{d\sigma_i}{da_j}\,(a) = \begin{cases} \sigma_{ij} - p_i\,(a)\,\sigma_{im}, \; i = \overline{1, m}, \; j = \overline{1, (m-1)}, \\ Z(a)\,[\sigma_{ij} - a_i\,p_j\,(a)], \; i, \; j = \overline{1, m}. \end{cases} \qquad \text{(XIV. 11)}$$

Необходимые при этом двойственные оценки критериев $p_j(a)$ ($j = \overline{1,m}$) определяются при вычислении эффективного значения (a) путем решения скаляризованной задачи $P(a)$. Подстановка равенств (XIV.11) в (XIV.10) после замены $\frac{dU}{dU_k}$ на $\eta_k = \eta_k(\sigma(a))$ приводит к следующему выражению для компонент вектора h

$$h_i(a) = \begin{cases} \eta_i - p_i \eta_m, \ j = \overline{1,(m-1)} & \text{– метод ограничений,} \\ Z(a)\,[\eta_j - (\sum_{k=1}^{m} a_k \eta_k) p_j], \ j = \overline{1,m} & \text{– ассортиментная скаляризация.} \end{cases}$$

Организация поиска оптимума в диалоговом режиме зависит от того, каким образом определяется величина шага $\lambda^{(i)}$ в направлении $h(a^{(i)})$. Одна из имеющихся здесь возможностей – последовательность шагов $\{\lambda_i\}_{i=0}^{\infty}$ формируется заранее, причем подчиняется стандартным условиям

$$\lim_{i\to\infty} \lambda^{(i)} = 0; \ \sum_{i=0}^{\infty} \lambda^{(i)} = \infty .$$

Алгоритмы дескриптивной ориентации, преследующие цель определения приемлемого решения, непосредственно учитывают особенности поведения человека и потому значительно более удобны для ЛПР.

Рассмотрим процедуры дескриптивной ориентации, в основе которых лежит параметризация границы Парето методом ограничений. Один из простейших – **метод последовательных уступок.** Для этого все критерии, характеризующие различные цели, располагают в порядке убывающей важности: сначала основной критерий C, затем остальные K_1, K_2, K_3 и т.д. Находят решение, оптимизирующее критерий C. Назначают уступку ΔC, на которую можно пойти в этом показателе, чтобы улучшить решение относительно критерия K_1. На показатель C накладывают ограничение, чтобы он был не меньше $(C - \Delta C)$; при этом находят решение, оптимизирующее критерий K_1. Затем назначают уступку в критерии K_1, за счет чего оптимизируют критерий K_2 и т.д.

В методе последовательно устанавливаемых ограничений (метод STEM) каждая итерация состоит из вычислительного этапа и этапа принятия решений. Вычислительный этап заключается в построении "платежной матрицы" и расчете допустимого компромиссного решения, предлагаемого ЛПР на этапе принятия решений. Платежная матрица содержит в j-й строке значения целевых функций V_i для некоторого действия из A, обеспечивающего экстремум функции V_i.

По диагонали матрицы расположены оптимальные значения, представляющие "идеальное" решение, которое, вообще говоря, не содержится в A. Рассчитываемое на этом этапе компромиссное решение является "ближайшим" к идеальному в минимаксном смысле. При расчете расстояния от идеального решения вводятся "веса" их относительной важности. Значения этих "весов" зависят от информации об относительной важности целевых функций, но существенную роль при их определении играют значения целевых функций в платежной матрице. Далее, на этапе принятия решения идеальное и компромиссное решения предъявляются ЛПР. Если ЛПР рассматривает компромиссное решение как удовлетворительное, процедура окончена. В противном случае ЛПР должно пойти на ухудшение значений по какому-либо критерию с целью выигрыша по другим. Выбор такого критерия и величины приемлемого ухудшения значений по нему ЛПР выполняет с помощью анализа чувствительности, имеющего своей целью исследование поведения целевых функций в окрестности полученного при этом компромиссного решения. Затем осуществляется вычислительный этап следующей итерации. В соответствии с полученной информацией определяются новые "веса" и предлагается новое компромиссное решение.

Метод последовательной ликвидации "узких мест" состоит в предъявлении ЛПР двойственных оценок частных критериев; отношения этих оценок, будучи предельными коэффициентами компенсации, позволяют судить о том, как изменение одного из критериев может отразиться на остальных. Метод начинается с того, что на нулевой итерации ($i = 0$) ЛПР указывает минимально приемлемые значения частных критериев $\alpha_1^{(0)}, \ldots, \alpha_m^{(0)}$.

Предполагается, что вектор $\alpha^{(0)} = (\alpha_1^{(0)}, \ldots, \alpha_m^{(0)})$ является достижимым, так как существует такой допустимый план $x^{(0)} \in X$, что $\varphi(x^{(0)}) = \alpha^{(0)}$.

Среди компонент вектора $\alpha^{(i)} = (\alpha_1^{(i)}, \ldots, \alpha_m^{(i)})$ ЛПР выделяет критерий K, значение которого должно быть увеличено в первую очередь ("узкое место"). Прочие компоненты $\alpha^{(i)}$ рассматриваются как ограничения $a_j = \alpha_j^{(i)}$, $j = \overline{1,m}$, $j \neq K$. Решается задача оптимизации критерия K. Далее найденное значение векторного критерия $\tilde{\alpha} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_m)$ и вектор оценок критериев $p = (p_1, \ldots, P)$ предъявляются ЛПР. Если это решение удовлетворяет ЛПР, то он решает, можно ли сократить значение критерия K. Если да, то указывается величина сокращения и продолжается этот процесс (номер итерации увеличивается на единицу) и так до получения удовлетворительного решения. Если же для улучшения значения критерия K необходимо некоторые из ограничений A уменьшить, то это делается на основе анализа отношений оценок критериев $P_j/P_k (j \neq K)$ и после этого решается задача оптимизации критерия K. Процесс продолжается до получения решения, приемлемого для ЛПР.

Контрольные вопросы и задачи

1. В каких случаях используются методы многокритериальной оптимизации?
2. Какие вопросы изучаются в теории выбора и принятия решений?
3. В чем заключается решение задачи многокритериальной оптимизации? В чем трудность решения?
4. Укажите решения, оптимальные по Парето, по четырем критериям K_1, K_2, K_3, K_4, если первый и третий надо максимизировать, а второй и четвертый минимизировать. Характеристики альтернатив даны в табл. XIV. 6. Дополнительные варианты решения задачи определяются в табл. XIV. 7.
5. Почему при решении многокритериальных задач необходимо участие человека и что при этом достигается?
6. Объясните, в чем разница между ЛПР и экспертами?
7. В чем заключается задача принятия решений? Из каких этапов она состоит?
8. Какие методы используются для формирования альтернатив принятия решений?
9. Область применения и особенности экспертного метода формирования альтернатив.
10. Использование методов математического моделирования для формирования альтернатив.
11. Особенности и область применения методов морфологического анализа и поэтапного формирования альтернатив.
12. Методы построения процедур при поэтапном формировании альтернатив.
13. Назовите основные типы задач принятия решений.
14. Типы шкал, используемые в задачах принятия решений.
15. Основные операции, используемые ЛПР в процессе принятия решений.
16. Классификация методов многокритериальной оптимизации и области их использования.
17. Оцените альтернативы, характеристики которых по четырем критериям даны в табл. XIV. 6.

"Веса" критериев по вариантам приведены в табл. XIV. 8.

18. Определите ценность альтернатив из задачи 17, если вероятность достижения целей составляет: 1) 0,7; 0.8; 0,9; 0,8; 2) 0.8; 0.6; 0.8; 0.7; 3) 0.8; 0,7; 0.7; 0.9; 4) 0.9; 0,7; 0,6; 0.7.
19. Решите задачу 17 при мультипликативном критерии.
20. Задаться предпочтениями между критериями из задачи 17, установить "весовые" коэффициенты критериев и найти наилучшую альтернативу. Принять первоначальные "весовые" коэффициенты равными соответственно 0.2; 0,3; 0.1; 0.4.

Таблица XIV. 6

Критерий	Значения критериев по альтернативам									
	1	2	3	4	5	6	7	8	9	10
K_1	4	3	7	4	5	8	7	6	3	8
K_2	6	2	5	4	3	7	6	5	4	3
K_3	6	5	8	7	4	6	7	3	8	5
K_4	3	5	7	6	4	5	5	3	2	5

Таблица XIV. 7

Критерии	Альтернативы									
	1–6	2–7	3–8	4–9	5–10	2–9	3–10	1–10		
K_1, K_2	1	2	3	4	5	6	7	8		
K_1, K_3	9	10	11	12	13	14	15	16		
K_2, K_3	17	18	19	20	21	22	23	24		
K_1, K_2, K_3	25	26	27	28	29	30	31	32		
K_1, K_2, K_4	33	34	35	36	37	38	39	40		
K_1, K_3, K_4	41	42	43	44	45	46	47	48		
K_2, K_3, K_4	49	50	51	52	53	54	55	56		
K_1, K_2, K_3, K_4	57	58	59	60	61	62	63	64		

Таблица XIV. 8

Вариант	"Веса" критериев			
	K_1	K_2	K_3	K_4
1	0,25	0,25	0,25	0,25
2	0,1	0,2	0,3	0,4
3	0,1	0,3	0,3	0,3
4	0,1	0,4	0,2	0,3
5	0,2	0,2	0,3	0,3
6	0,2	0,1	0,3	0,4
7	0,2	0,3	0,2	0,3
8	0,2	0,3	0,3	0,2
9	0,2	0,3	0,1	0,4
10	0,2	0,4	0,1	0,3
11	0,2	0,4	0,3	0,1
12	0,2	0,1	0,2	0,5
13	0,2	0,1	0,4	0,1
14	0,2	0,2	0,1	0,5
15	0,2	0,1	0,2	0,4

21. Что такое кривая безразличия?

22. Найти альтернативу из десяти возможных, наиболее близкую к идеальной точке, определяемой наилучшими значениями каждого из критериев. Характеристика альтернатив приведена в табл. XIV. 6.

Наилучшую альтернативу выбрать по наименьшему эвклидовому расстоянию, минимальному суммарному отклонению значения критерия от наилучшего и по минимизации максимального отклонения значения критерия от оптимального.

Сопоставьте полученные результаты.

23. Решите задачу из примера 22, если известны "весовые" значения критериев (см. табл. XIV. 8).

24. Решите задачу из примера 22, если идеальная точка характеризуется значениями критериев, приведенными в табл. XIV. 9.

25. Решить задачу из примера 24, если известны "весовые" значения критериев (см. табл. XIV. 8).

26. Какие метрики используются при определении альтернативы, наиболее близкой к идеальной?

27. Составить графы предпочтений для десяти альтернатив из задачи 4.

28. Построить обобщенный граф предпочтений для предшествующей задачи. Взяв на себя роль ЛПР, попарно сравните альтернативы и постройте ориентированный граф. При появлении противоречий в ответах постепенно устраняйте их.

29. Определите индексы согласия и несогласия десяти альтернатив, оцениваемых по четырем критериям (см. табл. XIV. 6), если "весовые" значения их даны в табл. XIV. 8. Составить матрицу индексов согласия и несогласия.

30. Преобразуйте матрицы индексов согласия и несогласия, если пороговые значения их составляют соответственно 0,80 и 0,03; 0,70 и 0,05; 0,60 и 0,01.

31. На основе полученных при решении задачи 30 матриц постройте граф предпочтений альтернатив и на основе его последовательных преобразований проранжируйте все альтернативы и выделите из них лучшую. При необходимости можно корректировать индекс согласия и несогласия.

Таблица XIV. 9

Вариант	"Веса" критериев			
	K_1	K_2	K_3	K_4
1	8	3	8	1
2	8	2	8	2
3	8	3	7	3
4	7	3	8	1
5	7	2	8	2
6	7	2	7	3
7	9	3	8	2
8	9	2	8	3
9	9	3	7	1
10	6	2	8	3
11	6	3	7	2
12	6	2	7	2
13	10	3	7	1
14	10	2	8	3
15	10	3	8	2

32. Сопоставьте решение многокритериальной задачи (см. табл. XIV. 6, 7) различными методами.

33. Структура человеко-машинных процедур принятия решений.

34. Какие алгоритмы оптимизации лежат в основе человеко-машинных процедур принятия решений.

35. В чем суть и какова процедура методов последовательных уступок, последовательно устанавливаемых ограничений и последовательной ликвидации "узких мест".

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. *Автоматизация* планирования горных работ на железорудных карьерах / Ю.П. Астафьев, А.С. Давидкович, Н.Д. Бевз и др. М., Недра, 1982.

2. *Акулич И.Л.* Математическое программирование в примерах и задачах. М., Высшая школа, 1986.

3. *Аленичев В.М.* Экономико-математическое моделирование горнотехнических задач на рудных карьерах. М., Недра, 1983.

4. *Банди Б.* Методы оптимизации. Вводный курс. М., Радио и связь, 1988.

5. *Бешелев С.Д., Гурвич Ф.Г.* Математико-статистические методы экспертных оценок. М., Статистика, 1980.

6. *Бронштейн И.Н., Семендяев К.А.* Справочник по математике для инженеров и учащихся вузов. М., Наука, 1986.

7. *Бугров Я.С., Никольский С.М.* Элементы линейной алгебры и аналитической геометрии. М., Наука, 1988.

8. *Васильев Ф.П.* Численные методы решения экстремальных задач. М., Наука, 1988.

9. *Вентцель Е.С.* Исследование операций: задачи, принципы, методология, М., Наука, 1980.

10. *Ганицкий В.И.* Организация производства на карьерах. М., Недра, 1983.

11. *Гилл Ф., Мюррей У., Райт М.* Практическая оптимизация. М., Мир, 1985.

12. *Гранберг А.Г.* Динамические модели народного хозяйства. М., Экономика, 1985.

13. *Гранберг А.Г.* Моделирование социалистической экономики. М., Экономика, 1988.

14. *Дегтярев Ю.И.* Исследование операций. М., Высшая школа, 1986.

15. *Емельянов С.В., Ларичев О.И.* Многокритериальные методы принятия решений. М., Знание, 1985.

16. *Еремин И.И.* Противоречивые модели оптимального планирования. М., Наука, 1988.

17. *Известия АН СССР.* Техническая кибернетика. 1987. № 1.

18. *Ириков В.А., Ларик В.Я., Самущенко Л.М.* Алгоритмы и программы решения прикладных многокритериальных задач. – Техническая кибернетика. 1986. № 1. с. 5–15.

19. *Каранчук В.П., Сваровский И.Н., Суздальницкий И.Д.* Основы применения ЭВМ. М., Радио и связь, 1988.

20. *Карманов В.Г.* Математическое программирование. М., Наука, 1986.

21. *Кашьян Р.Л., Рао А.Р.* Построение динамических стохастических моделей по экспериментальным данным. М., Наука, 1983.

22. *Кини Р.Л., Райфа Х.* Принятие решений при многих критериях: предпочтения и замещения. М., Радио и связь, 1981.

23. *Кузнецов Ю.Н., Кузубов В.И., Волощенко А.Б.* Математическое программирование. М., Высшая школа, 1980.

24. *Курносов А.М., Кудин И.Б.* Совершенствование методов математического программирования в горном деле. М., Наука, 1984.

25. *Липский В.* Комбинаторика для программистов. М., Мир, 1988.

26. *Лотов А.В.* Введение в экономико-математическое моделирование. М., Наука, 1984.

27. *Майника Э.* Алгоритмы оптимизации на сетях и графах. – М., Мир, 1981.

28. *Математические* методы и модели в планировании и управлении горным производством / А.Г. Протосеня, С.А. Кулиш, Е.И. Азбель и др. М., Недра, 1985.

29. *Муртаф Б.* Современное линейное программирование. М., Мир, 1984.

30. *Негруцкий Б.Ф.* Интенсификация монтажа оборудования угольных шахт. М., Недра, 1983.

31. *Пакеты* прикладных программ: Методы оптимизации. Под общ. ред. А.А. Самарского, М., Наука, 1984.

32. *Пакеты* прикладных программ: Программное обеспечение оптимизационных задач. Под общ. ред. А.А. Самарского. М., Наука, 1987.

33. *Потапов В.Д., Яризов А.Д.* Имитационное моделирование в горной промышленности. М., Высшая школа, 1981.

34. *Применение ЭВМ* и математических методов в горном деле. – Материалы XXVII Междунар. симпоз., Москва, 20–25 октября 1980 г. М., Недра, 1982.

35. *Программное* обеспечение микроЭВМ. Под общ. ред. В.Ф. Шаньгина. М., Высшая школа, 1988.

36. *Реклейтис Г., Рейвиндран А., Рэгсдел К.* Оптимизация в технике. Кн. 1. М., Мир, 1986.

37. *Ромакин М.И.* Оптимизация планирования производства: экономико-математические модели и методы. М., Финансы и статистика, 1981.

38. *Ромашкин И.П.* Перспективное планирование в угольной промышленности. М., Недра, 1982.

39. *Справочник* по математике для экономистов. Под ред. В.И. Ермакова. М., Высшая школа, 1987.

40. *Справочник* по теории вероятностей и математической статистике / В.С.Королюк, Н.И. Портенко, А.В. Скороход, А.Ф. Турбин. М., Наука, 1985.

41. *Таха Х.* Введение в исследование операций: В 2 кн. М., Мир, 1985.

42. *Теория* выбора и принятия решений / И.М. Макаров, Т.М. Виноградская, А.А. Рубчинский, В.Б. Соколов. М., Наука, 1982.

43. *Хохлюк В.И.* Параллельные алгоритмы целочисленной оптимизации. М., Радио и связь, 1987.

Абсолютная пропускная способность системы 318, 321, 323
Аддитивный алгоритм 182
Адекватность 16
Активная стратегия 377
Алгоритм:
  Лэнда и Дойга 176
  нумерации событий 275
  построения сетевой модели 264
  решения 38
Анализ чувствительности решения задач 116
Асинхронное статистическое моделирование 354

Базисная переменная 54
Бесконечные игры 370
Бинарное отношение 399
Булева переменная 152, 155

Вероятностная оценка 367
Верхняя цена игры 372
Ветвление при решении задач ЦЛП 177
Внешняя устойчивость 406
Внутренняя устойчивость 406
Временной анализ сетевой модели 281
Временные параметры 275, 260
Вспомогательные переменные 43
Выпуклое множество 49
Вырождение 75

Геометрическая интерпретация задачи линейного программирования 51, 74
Глобальный максимум 37
Градиентный метод 412
Граница целевой функции верхняя 177
– – – нижняя 178
Граф 239, 243
– состояний 313
Графическое решение оптимизационной задачи 31

Двойственная задача линейного программирования 110
– – параметрического программирования 118
Двойственность 110
Двойственные оценки 115
Двойственный симплекс-метод 124
Дескриптивная информация 411
Детерминированная оценка 367
Детерминированные параметры 9
Дискретная случайная величина 10
Дисциплина очереди 312
Длительность работы 277, 260

Жордановы исключения 55

Задача в условиях неопределенности 18
– – риска 18
– классификации объектов 386
– линейного целочисленного программирования (ЦЛП) 151
– математического программирования 29
– об оптимальном использовании ресурсов 89
– плане выпуска продукции 89
– о назначениях 81
– о расстановке оборудования 83
– о смеси 81, 94
– планирования добычных работ в режиме усреднения качества 91
– перевозки грузов горных предприятий 94
– работы группы горных предприятий 97
– ранжирования 386
– распределения 81
– с дихотомией 156
– транспортная 81, 94
Закон распределения случайной величины 11
Закрытая модель транспортной задачи 128
Замкнутая система обслуживания 329
Зондирование вершин 178

Игра с нулевой суммой 369
– с седловой точкой 373
Интенсивность обслуживания 314
– потока требований 313
Исследование операций 3
Исходное базисное решение 59
Исходное множество альтернатив 383

Календарный план (ленточно-сетевой график) 286, 260

Канонический вид задачи линейного программирования 42
Качественные параметры 12
– характеристики математической модели 16
Классификация методов математического программирования 38
Количественная оценка вариантов?
Компромиссный вариант 381
Конечные игры 370
Контур (цикл) 135
Конфликтные ситуации 369
Крайняя точка 50
Кривые безразличия 394
Критерий 360
– Байеса 363
– Вальда 364
– Гурвица 364
– Лапласа 363
– несогласия 403
– Пирсона χ^2 315
– принятия решений 366
– согласия 403
– Сэвиджа 364
– эффективности (оптимальности) 19, 22
Критический путь 260, 278

Лицо, принимающее решение (ЛПР) 282
Локальный максимум 382

Максимальная стратегия 372
Максимин 372
Математическая вариационная задача 19
– модель 9
Матрица выигрышей 361
– риска 361
Метод ветвей и границ 176, 153
– двойного предпочтения 130
– компенсации 394
– морфологического анализа 385
– наименьших стоимостей 130
– отсечений 164, 153
– покоординатной оптимизации 412
– последовательно устанавливаемых ограничений 414
– последовательной ликвидации "узких мест" 415
– последовательных уступок 414
– потенциалов 134
– поэтапного формирования 385
– решения 29
– северо-западного угла 130
Методы безусловной оптимизации (минимизации) локальные:
  градиентный 197
  Ньютона 199
  покоординатного спуска 196
  спряженных направлений 198
  сходимости 196
– – – глобальные 200
Методы условной оптимизации:
  возможных направлений 201
  линеаризации (Б.Н. Пшеничного) 202
  модифицированных функций Лагранжа 204
  Ньютона 202
  проекции градиента 202
  условного градиента 202
  штрафных функций 203
Многоканальная система массового обслуживания 322
Многокритериальная задача 23, 380
Множественные игры 369
Множество допустимых решений 8
Моделирование 8
Модель 8
Модифицированные жордановы исключения 58
Мультипликативный критерий 390

Надежность реализации программы 281
Неограниченная переменная 45
Неограниченное множество решений 47
Несовместимость системы ограничений 46
Неупорядоченная очередь 312
Неуправляемые переменные 21
Нижняя цена игры 272
Нормальный план 300
Нормативная процедура 411

Область допустимых значений: 194
  выпуклая 194
  многогранная 195
  невыпуклая 195
  разрывная 195
Область минимума 369
Обслуживание с приоритетом 312
Обслуживающая система 312
Обслуживающий аппарат 312
Ограничения выпуклые:
  линейные 201–202
  нелинейные 193
Одноканальная система массового обслуживания 319
Оперативное управление 261
Операция 3
Опорное решение 51,64
Опорное решение транспортной задачи 130

Оптимальная стратегия 370
Оптимальное решение 19
Оптимальность условная 219–224, 226
Оптимальные решения задачи линейного программирования 69
Оптимизация календарного плана:
по времени 290
по затратам 299
Оптимум (минимум, максимум) локальный 193
– глобальный 193, 194, 200
Ординарность потока 313
Основная задача линейного программирования 42
Основная теорема теории игр 374
Особенности двойственной задачи 112
Особенности решения транспортной задачи 128
Относительная пропускная способность системы 318, 321, 323
Ошибка исходной информации 367

Параметрическое программирование 117
Параметры модели 9
Парные игры 369
Пессимистический подход 364
Петля 259
Платежная матрица 361, 370
Поздний срок наступления события 276, 279, 260
Поиск условного экстремума функций многих переменных 29
Полностью целочисленная задача ЦЛП 152
Полный путь 278
Полный резерв времени выполнения работы 260, 276, 280
Полуцепь положительная 136
– отрицательная 136
Порог сравнимости 402
Поток без последствия 313
Принцип минимакса 372
Программа производства или исследований 257
Прокатная оценка 141
Пропускная способность системы 349
Прямая задача параметрического программирования 117
Пуассоновский поток 313
Путь 260

Работа (действительная, ожидание, эффективная) 259
Разомкнутые системы массового обслуживания:
с неограниченной длиной очереди 324
с ограниченной длиной очереди 324
с ожиданием 324
Разрешающий элемент 57
Ранжирование многокритериальных объектов 402
Ранний срок наступления события 278, 279, 260
Регулярный поток событий 312
Резерв времени наступления события 210, 276, 279
Решение по Парето 381
Решение транспортной задачи по критерию времени 144
Риск 360

Свойства множества решений 46, 48
Сводный резерв времени выполнения работы 260, 276, 280
Седловая точка 373
Сетевая постановка транспортной задачи 140
Сетевые методы планирования и управления 257
Сеть (сетевая модель) 235, 258
Сечение 165, 167
Симплекс-таблица 56
Система 6
Система массового обслуживания с потерями 314
Система ограничений 21
Система потенциалов 134
Сложность модели 17
Случайная величина 10
Смешанные стратегии 374
Событие 259
Составление оптимизационной математической модели 26
Состояние системы 6
Срочный план 300
Статистическая связь между параметрами системы 14
Статистическое моделирование 342
Стационарный поток 312
Стохастические параметры модели 9
Стратегия 361, 370
Стратегия поиска оптимума 34
Структура процесса 342
Структура системы 6
Структурный анализ 342
Структурный план 260
Сходимость алгоритма решения 38

Таблица работ 261
Таблица событий 280
Теорема двойственности 113
Теория выбора и принятия решений 380
Теория игр 360, 369
Теория статистических решений 360
Траектория динамической системы 216, 224, 245
Транспортная матрица 129
Транспортная сеть 140

Улучшение плана перевозок 135
Упорядоченная очередь 312
Управление в динамической системе 215, 232
Управляемые переменные 21
Устойчивость оптимального решения 116

Фиктивный поставщик 128
– потребитель 128
Форма записи задач линейного программирования:
развернутая 43
приведенная 43
матричная 43
Формула Эрланга 322
Функциональная связь между параметрами системы 13
Функция Беллмана, уравнение Беллмана 220, 227, 232, 247
Функция ценности 394

Характеристики системы массового обслуживания 313
Ход 369

Целевая функция 18
– – квадратичная 194, 196, 201
– – кусочно-линейная аппроксимация 206–212
– – нелинейная 193
– – сепарабельная 194

Частичное решение 183
Частично целочисленная задача ЦЛП 152
Чебышевская точка 393
Человеко-машинные процедуры 410
Чистая цена игры 373

Эвклидово расстояние 409
Эвристический метод решения задач ЦЛП 153
Экспертный метод 384
Этапы решения оптимизационной задачи 24
Этапы постановки задачи 24

Ядро графа 406

ОГЛАВЛЕНИЕ

Предисловие 3

Глава I. Введение в математическое моделирование 6
§ 1. Назначение и типы моделей 6
§ 2. Свойства производственных систем и виды математических моделей . . . 9
§ 3. Качественные характеристики математических моделей 16
§ 4. Оптимизационные модели, их структура 18
§ 5. Этапы решения оптимизационной задачи 24

Глава II. Методы и средства поиска оптимальных решений 29
§ 1. Особенности решения задач математического программирования 29
§ 2. Графическое решение оптимизационной задачи 31
§ 3. Стратегия поиска оптимума 34
§ 4. Краткая характеристика задач, методов и программных средств математического программирования 38

Глава III. Решение основной задачи линейного программирования 42
§ 1. Основная задача линейного программирования 42
§ 2. Свойства множества решений 46
§ 3. Стратегия поиска оптимального решения симплекс-методом 50
§ 4. Расчет нового базисного решения. Жордановы исключения 55
§ 5. Получение исходного базисного решения 59
§ 6. Поиск опорного решения 64
§ 7. Поиск оптимального решения 69
§ 8. Вырождение 75

Глава IV. Применение линейного программирования в задачах планирования и управления горным производством 81
§ 1. Условия применения и классификация линейных моделей 81
§ 2. Задачи о расстановке оборудования 83
§ 3. Задачи об оптимальном использовании ресурсов (оптимальном плане выпуска продукции) 89
§ 4. Планирование добычных работ в режиме усреднения качества 91
§ 5. Планирование перевозок грузов горных предприятий 94
§ 6. Модель задачи планирования работы группы горных предприятий (добывающих и перерабатывающих) 97
§ 7. Модели задач размещения 100

Глава V. Двойственность в линейном программировании и оценка результатов решения линейных задач 110
§ 1. Прямая и двойственная задачи линейного программирования 110
§ 2. Теоремы двойственности и их экономический смысл 113
§ 3. Анализ чувствительности решения линейных экономических задач 116
§ 4. Двойственный симплекс-метод 124

Глава VI. Транспортная задача линейного программирования 127
§ 1. Особенности транспортной задачи 127
§ 2. Построение опорного решения 130
§ 3. Условия и метод построения оптимального решения транспортной задачи 134
§ 4. Алгоритм решения транспортной задачи на сети 140
§ 5. Постановка и решение транспортной задачи по критерию времени 144

Глава VII. Целочисленное линейное программирование и его применение при решении задач планирования горного производства 151
§ 1. Особенности целочисленных линейных задач и методов их решения 151
§ 2. Использование булевых переменных при построении моделей целочисленных задач планирования 155
§ 3. Модель планирования размещения углеобогатительных фабрик 157
§ 4. Модель оперативного планирования расстановки самоходного оборудования по очистным блокам рудника 160
§ 5. Модель планирования развития шахтного фонда производственного объединения 162
§ 6. Модель задачи о раскрое 163
§ 7. Решение целочисленных задач методом отсечений 164
§ 8. Решение целочисленных задач методом ветвей и границ 176
§ 9. Частичный перебор в задачах с булевыми переменными 182

Глава VIII. Нелинейное программирование и его использование в планировании и управлении горным производством 193
§ 1. Общая характеристика, основные типы и особенности задач нелинейного программирования 193
§ 2. Методы решения задач безусловной оптимизации 196
§ 3. Прямые методы решения задач условной оптимизации 200
§ 4. Методы преобразования для решения задач условной оптимизации 203
§ 5. Приближенные методы решений нелинейных задач 206

Глава IX. Динамические оптимизационные модели планирования и управления горным производством 215
§ 1. Общая постановка и геометрическая интерпретация динамических задач оптимизации 215
§ 2. Принцип оптимальности, основное функциональное уравнение и порядок решения задач методом динамического программирования 217

§ 3. Задача об определении оптимальной траектории перемещения системы и ее решение методом динамического программирования 222
§ 4. Использование динамического программирования для решения статических задач распределения ресурсов 225
§ 5. Динамическая задача распределения ресурсов 231
§ 6. Задача отыскания кратчайших расстояний на сети 235
§ 7. Использование динамического программирования при оптимизации альтернативных графов 239
§ 8. Задачи дискретного оптимального управления 243
§ 9. Динамическое программирование и его локальные варианты в задачах дискретного оптимального управления 247
§ 10. Метод возможных направлений в задачах дискретного оптимального управления 251
§ 11. Методы, основанные на условиях оптимальности 254

Глава X. Сетевое планирование и управление реализацией программ 257
§ 1. Основные определения и этапы сетевого планирования и управления . . . 257
§ 2. Сетевое представление программ (сетевая модель) 261
§ 3. Расчет временных параметров сетевой модели 275
§ 4. Построение календарного плана реализации программ 286
§ 5. Оптимизация календарного плана по времени при ограниченных ресурсах 290
§ 6. Оптимизация календарного плана по затратам 299
§ 7. Управление процессом реализации программ 308

Глава XI. Аналитические модели систем массового обслуживания 312
§ 1. Системы массового обслуживания 312
§ 2. Порядок решения задач массового обслуживания 315
§ 3. Моделирование систем массового обслуживания с отказами 319
§ 4. Разомкнутые системы массового обслуживания с ожиданием 324
§ 5. Замкнутые системы массового обслуживания 329

Глава XII. Статистическое моделирование производственных процессов 341
§ 1. Задачи моделирования процессов и классификация типов взаимодействия машин и механизмов 341
§ 2. Моделирование непосредственного взаимодействия машин и механизмов 344
§ 3. Моделирование взаимодействия через склад 348
§ 4. Статистическое моделирование систем массового обслуживания 354

Глава XIII. Принятие решений в условиях неопределенности 360
§ 1. Элементы теории статистических решений 360
§ 2. Выбор критерия решений и определение рациональной точности исходной информации 366
§ 3. Основные понятия теории игр 369
§ 4. Методы решения парных игр 373

Глава XIV. Элементы теории выбора и принятия решений 379
§ 1. Задачи многокритериальной оптимизации 379
§ 2. Методы формирования альтернатив 383
§ 3. Общие сведения о методах многокритериальной оптимизации 386
§ 4. Количественные методы решения многокритериальных задач 390
§ 5. Теоретико-множественные методы 399
§ 6. Человеко-машинные процедуры принятия решений 410

Список рекомендуемой литературы 420
Алфавитно-предметный указатель 422

Резниченко С.С., Подольский М.П., Ашихмин А.А.

Р 34 Экономико-математические методы и моделирование в планировании и управлении горным производством: Учеб. для вузов. – М.: Недра, 1991. – 429 с.: ил.

ISBN 5-247-01000-0

Изложены принципы математического моделирования, описаны типы и модели задач организации, планирования и управления горным производством. Основное внимание уделено моделям и методам линейного и динамического программирования, сетевому планированию, теории массового обслуживания и статистического моделирования. Приведены примеры построения моделей типовых задач планирования и управления горным производством.

Для студентов вузов, обучающихся по специальности "Экономика и управление в отраслях горной промышленности и геологоразведке".

$\text{Р} \frac{2501010000 - 057}{043(01) - 91} 82-91$ ББК 65.050.9(2)

УЧЕБНОЕ ИЗДАНИЕ

Резниченко Семен Саулович
Подольский Михаил Петрович
Ашихмин Алексей Анатольевич

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИРОВАНИЕ В ПЛАНИРОВАНИИ И УПРАВЛЕНИИ ГОРНЫМ ПРОИЗВОДСТВОМ

Заведующий редакцией *Е.И. Кит*
Редактор издательства *Э.Я. Освальдт*
Технические редакторы *Л.А. Миронова, Н.С. Анашкина*
Корректоры *К.И. Савенкова, И.Н. Таранева*
Операторы *М.Г. Булковская, Ю.А. Савельева*

ИБ 7666

Подписано в печать с репродуцированного оригинал-макета 17.01.91. Формат $60 \times 90^1/_{16}$. Бум. офсетная № 2. Гарнитура Пресс-роман. Печать офсетная. Усл.-печ. л. 27,0. Усл. кр.-отт. 27,0. Уч.-изд. л. 28,35. Тираж 2600 экз. Зак. № 175. /1723-7. Цена 1 р. 40 к.
Набор выполнен на наборно-пишущей машине

Ордена "Знак Почета" издательство "Недра".
125047 Москва, Тверская застава, 3.

Тульская типография Государственного Комитета СССР по печати.
300600, г. Тула, Проспект Ленина, 109