

НА:БЕОГРАДОВ:АМ:ГОЛУБЦОВ:А:А:СМНДЕН:УЗ

Н.А.БЕСПАЛОВ, А.И.ГОЛУБЦОВ, А.А.СИНДЕЕВ

**ЭКОНОМИКО-
МАТЕМАТИЧЕСКИЕ
МЕТОДЫ
В ТОПОГРАФО-
ГЕОДЕЗИЧЕСКОМ
ПРОИЗВОДСТВЕ**

Н. А. БЕСПАЛОВ, А. И. ГОЛУБЦОВ, А. А. СИНДЕЕВ

ЭКОНОМИКО- МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ТОПОГРАФО- ГЕОДЕЗИЧЕСКОМ ПРОИЗВОДСТВЕ

Под редакцией д-ра техн. наук
проф. Н. А. Беспалова



МОСКВА „НЕДРА“ 1983

Беспалов Н. А., Голубцов А. И., Синдеев А. А.
Экономико-математические методы в топографо-геодезическом производстве. М., Недра, 1983. с. 320.

Даны понятия об основных экономико-математических моделях, которые могут быть использованы в технико-экономическом планировании топографо-геодезических работ. Описаны соответствующие методы расчета, обоснованы целесообразные экономические критерии оптимального планирования.

Рассмотрены основные вопросы построения и организации АСУ, их особенности применительно к геодезической отрасли.

Для инженерно-технических работников геодезических организаций, занимающихся планированием и проектированием топографо-геодезических работ. Может быть полезна студентам, аспирантам и преподавателям геодезических специальностей вузов.

Табл. 78, ил. 83, список лит. — 31 назв.

Рецензент — канд. техн. наук *Р. П. Каширникова* (ЦНИИГАиК)

Николай Александрович БЕСПАЛОВ
Александр Иванович ГОЛУБЦОВ
Арнольд Александрович СИНДЕЕВ

**Экономико-математические методы
в топографо-геодезическом производстве**

Редактор издательства *Э. Н. Чумаченко*
Переплет художника *В. Ю. Новикова*
Художественный редактор *Е. Л. Юрковская*
Технический редактор *О. Н. Ласточкина*
Корректор *В. И. Ионкина*

ИБ № 2633

Сдано в набор 17.08.82. Подписано в печать 02.12.82. Т-20899.
Формат 60×90^{1/16}. Бумага кн.-журнальн. Гарнитура литературная. Печать высокая. Усл.-печ. 20,0. Усл. кр.-отт. 20,0. Уч.-изд. 22,0. Тираж 3800 экз. Заказ 1833/7224—15. Цена 1 р. 40 к.

Ордена «Знак Почета» издательство «Недра», 103633, Москва, К-12, Третьяковский проезд., 1/19

Ленинградская типография № 4 ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. в.Егении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 191126, Ленинград, Социалистическая ул., 14.

ПРЕДИСЛОВИЕ

Методы количественной оценки влияния производственных факторов на результирующие показатели деятельности отдельных подразделений топографо-геодезического производства и математические методы и приемы, позволяющие найти наилучшие или близкие к наилучшим решения в управлении топографо-геодезическим производством и в его организации — это одни из основных вопросов деятельности любого предприятия, требующих постоянного внимания и совершенствования.

Успех выполнения заданий одиннадцатой пятилетки, в том числе и внедрения комплексной народнохозяйственной программы совершенствования управления, в конечном счете зависит от людей, от того, насколько они подготовлены, правильно расставлены, участвуют в процессах управления и производства.

Отметим, что экономико-математические методы — это условное название научных дисциплин на стыке экономики с математикой и кибернетикой, впервые (в этом смысле) введенное В. С. Немчиновым в начале 60-х годов и ныне широко применяемое в СССР. В зарубежной литературе общепринятого названия для этой теоретической и прикладной области применения математики и кибернетики в экономике нет.

Обычно в экономико-математические методы включаются следующие группы научных дисциплин и направлений, хотя предмет их исследования и границы между ними в литературе толкуются по-разному [19]: 1) принципы экономико-математических методов; 2) математическая статистика; 3) математическая экономия и эконометрия; 4) методы принятия оптимальных решений, включая исследования операций в экономике; 5) экономическая кибернетика; 6) методы экспериментального изучения экономических явлений.

Интенсивное развитие математических методов и их внедрение в практику управления вызвано в основном двумя причинами: 1) материальное производство в настоящее время достигло такого уровня сложности, что управление обычными методами стало далеко не эффективным, 2) к настоящему времени созданы мощные ЭВМ, позволяющие поставить технику вычислений в принципиально новые условия. Современные математические методы и ЭВМ позволяют решать задачи, которые без них решать было бы невозможно, позволяют проектировать, создавать и эксплуатировать автоматизированные системы управления предприятиями и отраслями.

С конца 40-х — начала 50-х годов развернулось обсуждение теоретических принципов и практических схем планирования хозяйства и управления производством в условиях научно-технического прогресса. Эти дискуссии не завершились и в настоящее время.

Одни специалисты утверждают, что современные подходы к планированию и управлению непременно требуют построения формальной модели системы. Другие считают, что формальное описание достаточно сложных систем невозможно (а некоторые добавляют — и не нужно). Действительно, если формализацию сводить к математическому описанию системы, то оно, как правило, сложно и громоздко, не всегда возможно в привычных терминах и понятиях и, может быть, не всегда к нему следует стремиться.

Однако в настоящее время уже нельзя говорить о реалистическом подходе к задачам планирования и управления, если после предварительных этапов осознания и постановки проблемы в содержательных терминах не выбраны параметры управления и по крайней мере качественно не оценены цели (показатели качества решения) и ограничения задачи. Говоря о формальной модели выбора хозяйственного решения, можно не требовать явного, а тем более аналитического задания условий и целей задачи. Настаивая на необходимости формализации процессов планирования хозяйственной деятельности и управления производством, необходимо знать лишь некоторую априорную информацию о свойствах класса задач, к которому принадлежит исследуемая проблема, и иметь четкую запись задачи, из которой следовало бы, что задано и что требуется найти. Кроме того, должны быть указаны приемы (аналитические, алгоритмические, экспериментальные или основанные на экспертных оценках), позволяющие установить, удовлетворяет ли любое выбранное решение условиям задачи, и оценить его качество (гл. II).

Условные экстремальные модели — модели математического, главным образом линейного, программирования используются в практике решения экономических задач уже четверть века. Анализ отечественного и зарубежного опыта показал, что применение математических методов повысило эффективность хозяйственных решений по сравнению с решениями, подготовленными опытными специалистами с помощью традиционных методов, в среднем на 5—15 %. Однако следует признать, что 15—25 лет тому назад специалисты по математическому программированию ожидали большего эффекта от использования в планировании и управлении новых методов и вычислительных машин.

Можно указать целый ряд причин, объясняющих, почему не вполне оправдались надежды энтузиастов экономико-математических методов. Остановимся только на двух важных факторах. Первый фактор связан с соответствием математической модели реальной ситуации, второй определяется характером взаимодействия экономистов и математиков.

Большая часть известных экономико-математических описаний процессов планирования и управления представляет собой детер-

минированные модели математического программирования. Детерминированные модели описывают замкнутые системы, связями которых с внешним миром можно пренебречь. Далеко не во всех реальных ситуациях можно анализировать поведение системы, не учитывая влияния непредвиденных изменений в окружающей среде.

Следует полагать, что эффективность экономико-математических методов существенно возрастет при переходе от детерминированных моделей планирования к стохастическим (гл. III и XI). Стохастические модели лучше, чем детерминированные, приспособлены к учету последовательности поступления и использования информации и допустимой очередности выбора и корректирования решений.

В создании современных сложных систем, в частности автоматизированных систем управления и планирования, участвуют большие коллективы специалистов различного профиля. Недостаточная эффективность разработок и, главным образом, внедрения многих идей, планов и проектов определяется в значительной мере нерациональной организацией взаимодействия экономистов и инженеров, с одной стороны, и математиков и вычислителей — с другой.

Эффективное планирование производства и управления хозяйственной деятельностью может быть обеспечено лишь при естественном разделении обязанностей и ответственности между экономистами, инженерами и математиками.

Мы пришли к следующим выводам.

Ответственные хозяйственные решения должны быть в некотором смысле экстремальными или близкими к ним. Отступление от этого принципа обычно связано с излишними затратами, а также снижает качество плана и управления.

Планирование производства и управления отраслями народного хозяйства происходит при неполной информации об условиях реализации плана и управления. Исследование и учет стохастических факторов являются непременным условием рационального выбора хозяйственных решений.

Важно отметить, что методологической основой применения математических методов и моделей для решения экономических проблем является марксистско-ленинское учение о производстве.

Математические методы в экономике позволяют полнее использовать преимущества плановой системы хозяйства. Они дают возможность разрабатывать планы на глубокой научной основе и отражать в них всю сложность и многообразие зависимостей между действующими переменными факторами.

При плановой системе хозяйства применение математических методов и моделей не ограничивается никакими социально-экономическими рамками. В условиях капитализма их используют для решения отдельных конкретных производственных задач, однако за пределами предприятий они применяться не могут. Конкуренция, борьба за рынки сбыта, стремление монополий хранить в тайне производственные новшества, анархия, чередование подъемов и спа-

дов производства исключают возможность использования математических методов в управлении экономикой капиталистического общества.

В социалистическом обществе математические методы служат средством совершенствования планирования и управления экономикой на всех ее уровнях от предприятия до всего народного хозяйства.

Огромный вклад в развитие этой отрасли знаний внесли советские ученые В. С. Немчинов, Л. В. Конторович, Л. Б. Юдин, А. Л. Лурье, А. Г. Аганбегян, Н. П. Федоренко, Б. В. Гнеденко и др.

Важно отметить, что в экономико-математических методах применяются различные разделы математики, математической статистики и математической логики; большую роль в машинном решении экономико-математических задач играют вычислительная математика, теория алгоритмов и другие смежные дисциплины.

Если в технику математика вошла непосредственно через цикл общенаучных и общетехнических дисциплин, главным образом как средство их изучения, то в экономику, организацию и управление производством математика входит непосредственно, сама являясь источником основных идей и необходимого аппарата для научного построения этих дисциплин.

Проникновение математики в экономические науки и перестройка последних лишь частично основана на использовании классического аппарата математики: математического анализа, алгебры и теории вероятностей, хотя требования к знаниям экономистов и в этих областях математики повышаются. Большое значение приобретают, например, такие разделы, как линейная алгебра, функции многих переменных, разностные уравнения и т. д.

Сильно видоизменяется в связи с этим современная экономическая литература. В ней нередко можно найти ссылки на топологию и функциональный анализ, использование абстрактных аксиоматических теорий, сложных вероятностных моделей, физического моделирования и других самых современных средств теоретической и прикладной математики.

Математические построения, возникающие в современной экономической теории, в большей или меньшей степени отличаются от соответствующих построений в физических науках. Поэтому будет неудивительно, если попытки изучения свойств экономических систем путем формального применения того аналитического аппарата, который был удобным в физике, окажутся малопродуктивными. Глубокое проникновение в сущность экономических закономерностей может быть обеспечено лишь с учетом специфических требований к математическому аппарату, предъявляемых самой природой экономических систем. Такова общая точка зрения специалистов, работающих в области математической экономики.

Начало применения экономико-математических методов в топографо-геодезическом производстве относится к 1963 г. В последующие годы в геодезических публикациях («Труды МИИГАиК», «Труды МИИЗ», «Труды НИИГАиК») появился целый ряд статей, в которых

анализируются разнообразные планово-организационные, проектные и другие задачи с позиций математического программирования.

С 1967 г. в отделе экономики ЦНИИГАиКа проводят исследования возможностей применения экономико-математических методов в управленческой и планово-организационной деятельности аэрогеодезических предприятий по следующим направлениям: 1) по составлению технических проектов, 2) по планированию работы предприятий, 3) по управлению и организации работ в предприятии и его подразделениях, 4) по оптимизации технологических процессов.

Применение экономико-математических методов как в топографо-геодезическом производстве, так и в любой другой отрасли приводит к положительным результатам лишь в тех случаях, когда правильно учитываются экономико-организационные особенности объекта исследований, основные факторы и условия его развития.

Конечной целью топографо-геодезического производства является создание карт и планов на территорию суши, внутренних водоемов и континентального шельфа для потребителей — различных отраслей народного хозяйства. Руководство государственными топографо-геодезическими работами и государственный геодезический контроль в стране осуществляется общесоюзным органом — Главным управлением геодезии и картографии при Совете Министров СССР.

Основными производственными подразделениями в топографо-геодезическом производстве являются аэрогеодезические объединения, предприятия, Союзмаркштрест и специализированные организации по инженерным съемкам (АрмГИИГИС и УкрГИИГИС). Управление топографо-геодезическим производством в системе ГУГК построено с соблюдением производственно-территориального принципа. Территориальный принцип управления выражается в закреплении за каждым предприятием определенной территории страны. Экономико-организационные особенности топографо-геодезического производства обусловлены его технологией, большим разнообразием физико-географических условий объектов труда, специфическим характером самого труда, а также содержанием и назначением конечных результатов работ. Так, например, технологией работ предусматривается очередность выполнения процессов в полевых и камеральных (цеховых) условиях. При этом выполнение полевых работ связано с влиянием окружающей среды и по продолжительности ограничено полевым сезоном. В зависимости от географического положения объекта работ продолжительность полевого сезона может колебаться в пределах от 3 до 10 мес., что обуславливает различную эффективность использования средств труда и отражается на параметрах моделей производственной мощности предприятия.

На эффективность работ влияют и физико-географические особенности объекта. В зависимости от физико-географических условий для выполнения одного и того же процесса приходится оптимально выбирать технические средства и технологию работ.

Изменение структуры и географического размещения основных видов топографо-геодезических работ существенно влияет на выбор объективного показателя объектов производства, правильно отражающего динамику развития отрасли, уровень производительности труда, фондоотдачу и другие показатели эффективности.

Особенности, вызванные специфическим характером самого труда, проявляются в том, что топографо-геодезическое производство является трудоемким. Кроме того, труд работников отрасли является высококвалифицированным и в основном требует специалистов с высшим образованием. Таким образом, повышение эффективности топографо-геодезического производства в первую очередь связано с экономией живого труда ИТР, что и определяет в ряде случаев выбор критерия оптимальности в экономико-математических моделях.

Наконец, особенность топографо-геодезического производства заключается в специфическом характере конечных работ. Потребность отраслей народного хозяйства в любых видах топографо-геодезических работ должна задаваться не только указанием их вида (процесса), но и географическим положением объекта. Последнее обстоятельство очень важно учитывать при разработке экономико-математических моделей многовариантной производственной задачи, так как дополнительные ограничения сужают возможности ее варьирования.

Гл. I—VI, IX, XI, XII написаны д-ром техн. наук проф. Н. А. Беспаловым; канд. техн. наук доц. А. И. Голубцов написал гл. VII, VIII и X; гл. XIII написана канд. техн. наук А. А. Синдеевым.

Глава I

ЭЛЕМЕНТЫ СИСТЕМНОГО АНАЛИЗА

§ 1. Понятие системы, ее структуры, виды систем

В настоящее время в технической и экономической литературе употребляются термины: системный подход, системный анализ, теория систем, системное исследование, системная ориентация, системное движение. Все эти термины являются понятием одного порядка, тесно связанным с понятием системы.

В кибернетике под с и с т е м о й понимают множество предметов m , на котором реализуется заранее данное отношение R с фиксированными свойствами P [19].

В системном анализе используется понятие системы как единства взаимосвязанных элементов, совместно действующих для достижения общей цели [22].

Под системой понимают также наличие множества объектов (элементов) с набором связей между ними и между их свойствами, т. е. все состоящее из связанных друг с другом частей называется системой. При этом предполагается, что объекты (части) функционируют во времени «как единое целое» — каждый объект, подсистема, ячейка работают ради единой цели, стоящей перед системой в целом [27]. Системы, которые в течение определенного интервала времени переходят из одного состояния в другое, или функционируют, называются д и н а м и ч е с к и м и.

Рассмотрим теперь входящие в определение системы понятия: объекты, связи, свойства.

П о д о б ъ е к т а м и, и л и э л е м е н т а м и, системы понимают части, или компоненты, системы. Они могут быть одинаковыми и различными. Например, элементами системы предприятия могут быть его цехи, рабочие места и т. д. За объекты принимаются и математические переменные, уравнения, правила и законы, технологические процессы, производственные подразделения, приборы и т. д.

П о д с в о й с т в о м понимают качества параметров объектов, изменяющихся в результате действия системы. Свойства дают возможность описывать системы количественно, выражая их в единицах, имеющих определенную размерность.

П о д с в я з я м и понимают все то, что соединяется в системном процессе в целое. Связи существуют между всеми системными элементами и объектами, между системами и подсистемами. Различают три вида связей между элементами системы: механические (обмен усилиями между элементами), трофические (обмен энергией)

и сигнальные (обмен сигналами, информацией). Связи классифицируются на: а) связи первого порядка (функционально взаимно-необходимые); б) связи второго порядка (дополнительные связи, улучшающие действие системы, но не являющиеся функционально необходимыми); в) связи третьего порядка (излишние или противоречивые).

Элементы и связи определяют состояние системы. Динамическая система может иметь множество возможных состояний; последовательно переходя из одного состояния в другое. Однако данная система всегда находится только в одном из состояний и готова вступить в следующее состояние. Под состоянием системы понимают точно определенное условие или свойство, которое может быть опознано, если повторится снова. Смена состояний системы определяет линию ее поведения.

Число элементов систем, связей и состояний имеет математическую зависимость между собой. Максимально возможное число связей V между n элементами составляет $V = n(n - 1)$. Максимально возможное число состояний системы H определяется как $H = 2^{n(n-1)}$ или $H = 2^V$. Максимально возможное число комбинаций из связей элементов системы определяется как $K = 2^n$.

При анализе необходимо учитывать, что между переменными (элементами, объектами системы) существует связь, если они накладывают ограничения на поведение (изменение, функционирование) друг друга. Если такие ограничения отсутствуют, то связь также отсутствует. Связи исследуются корреляционным анализом (см. гл. XII), анализом неопределенности (см. гл. XI) и другими способами.

Исследователь проблемы, сам принимает решение — какие связи существенны, а какие тривиальны, т. е. вопрос о тривиальности оказывается связанным с личными интересами исследователя и задачами, которые стоят перед ним.

Система и внешняя среда. Для системного анализа важное значение имеет внешняя среда. Под внешней средой понимают совокупность всех объектов, изменение свойств которых влияет на систему, а также тех объектов, чьи свойства меняются в результате изменения поведения системы. Внешняя среда может быть отделена от системы границей. Граница может быть фиксируема в физическом и логическом смысле. Частным случаем границы является совокупность входов в систему, посредством которых внешняя среда влияет на систему, а также совокупность выходов из системы, через которые система воздействует на внешнюю среду. Таким образом, между системой и средой существует тесная взаимосвязь и взаимозависимость, заключающиеся в том, что система проявляет и формирует свои свойства в процессе взаимодействия со средой, являясь при этом его ведущим активным компонентом.

Системные объекты (части системы). Основными частями системы являются вход, процесс (операция), выход, цель, обратная

связь и ограничения. Рассмотрим эти понятия, пользуясь рис. 1.

Вход системы является комплексным понятием. Во-первых, он — субстанция, которая поступает в систему и подвергается определенным преобразованиям (сырье, материалы, энергия, информация), изменяется при протекании данного процесса. Компонентами входа являются рабочий вход (то, что «обрабатывается») и процессор (то, что «обрабатывает»).

Во-вторых, вход — внешняя окружающая среда, или совокупность факторов и явлений, воздействующих на систему (природные условия, внешнеполитическая обстановка и т. д.). Эти факторы можно разбить на две группы: а) случайные факторы, характеризующиеся законами распределения, неизвестными законами или действующие без всяких законов; б) факторы, находящиеся в распоряжении системы, являющейся внешней и активно, разумно действующей по отношению к рассматриваемой системе (например, факторы, находящиеся в распоряжении предприятия, используемые им в отношении экспедиции). В-третьих, к понятию входа системы относятся и установленные способы функционирования элементов системы, например инструкции, положения, приказы, определяющие порядок, правила, ограничительные условия, цели действия системы. Состояние входов называется импульсами.

Второй частью системы является процесс, или операция (см. гл. III). Процесс переводит вход в выход, и в то же время он — внутренняя структура системы, т. е. ее каналы, по которым проходят вещество, энергия, информация, поступающие в систему через входы и процессы или операции их преобразования. Для экономических систем процессом являются процессы воспроизводства материальных, трудовых и финансовых ресурсов.

Третья часть системы — выход, являющийся продуктом, или результатом, ее деятельности. Выход связывает систему с внешней средой. Отсутствие выхода делает невозможным существование системы. Всякий выход некоторого процесса системы является в то же время входом определенного процесса, а всякий вход системы является выходом этой или другой системы. Состояние выходов называется реакциями. Исследование того, какие импульсы вызывают реакции, позволяет в итоге раскрыть действие системы. Таким образом, реакции имеют характер зависимых переменных, а импульсы — независимых переменных, т. е. реакция является функцией импульсов.

Обратная связь представляет собой канал связи между выходом и входом системы, функционирующий либо прямо, либо через

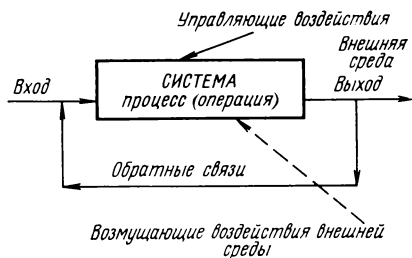


Рис. 1. Структура системы

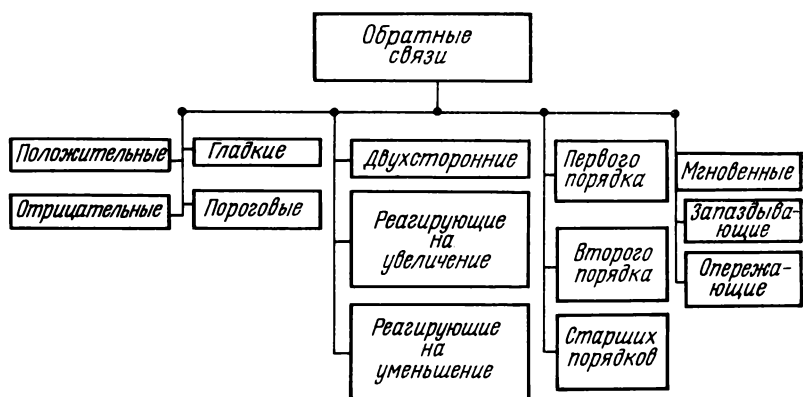


Рис. 2. Классификация обратных связей

другие элементы системы (например, через орган управления). Обратная связь передает данные о функционировании управляемой системы с ее выхода в управляющий орган, где она сравнивается с данными, задающими содержание и объем работы, например с планом. Если фактическое выполнение будет меньше плана, то управляющий орган системы вырабатывает меры по ликвидации этого несоответствия. Такое несоответствие между желаемым и существующим выходом (и тем самым между желаемой и существующей системой) называется *проблемой* (см. гл. II). Проблема может заключаться в предотвращении уменьшения выхода или же в его увеличении. За обнаружением проблемы следуют прогнозирование ее развития и оценка актуальности ее решения, позволяющая определить необходимость ее решения.

Приведем (рис. 2) следующую классификацию обратных связей [9].

Системы создаются и развиваются для достижения определенной *цели*. Под целью понимают конечное состояние, к которому стремится система в силу своей структурной организации.

В управляемых системах часть входа обособляется в управляющее воздействие, под которым понимается передача различных директивных распоряжений, приказов, указаний и тому подобной информации.

Наконец, каждая система испытывает на себе влияние *внешней среды*.

Иерархия систем — расположение частей, или элементов, целого в порядке от высшего к низшему. При построении иерархии систем сначала определяют первичную систему, которая не допускает дальнейшего расчленения на составные части (элементы) без потери основного качества всей системы. Так, например, применительно к производству первичной системой является производ-

ственный рабочий, имеющий определенные навыки, квалификацию, опыт, располагающий необходимыми орудиями труда, с помощью которых он превращает предметы труда в готовые продукты труда.

Системой второго порядка являются бригады, объединяющие двух и более рабочих; системой третьего порядка служит отдельный производственный участок; четвертого порядка — цех; пятого — предприятие и т. д.

Иерархическое построение систем в экономике позволяет успешно решать многие практические вопросы, связанные с совершенствованием управления в отрасли.

Классификация систем. Различают материальные и абстрактные системы. Материальные системы представляют собой объективную реальность, существующую в природе или обществе вне нашего сознания и независимо от него, развивающуюся по имманентным законам. К материальным системам, являющимся не отдельно взятой вещью или предметом, а совокупностью материальных объектов (элементов), относятся системы неорганической природы (физические естественные и физические искусственные, геологические, химические и т. д.) и живые системы, куда входят как простейшие биологические системы, так и очень сложные биологические объекты типа организма, вида, экосистемы. Особый класс материальных живых систем образуют социальные системы, многообразные по своим типам и формам (начиная от простейших социальных объединений и кончая социально-экономической структурой общества).

Абстрактная система является отражением материальной системы в человеческом сознании, продуктом человеческого мышления. Если материальная система состоит из определенных объектов или элементов (например, машина состоит из агрегатов, узлов и деталей, предприятие — из оборудования, машин, предметов труда, рабочих кадров, готовой продукции и т. п.), то абстрактная система состоит из символов свойств этих объектов. Поэтому абстрактные системы можно разделить на множества различных типов: понятия, гипотезы, идеи, теории, смену научных теорий, формулы, модели, планы и т. п.

Между материальными и абстрактными системами существует следующая связь: одна и та же материальная система может быть описана или изучена с помощью некоторого множества различных абстрактных систем. Например, предприятие может быть рассматриваемо как техническая система (совокупность применяемых на нем машин и оборудования), как технологическая система (совокупность применяемых технологических процессов), как система совместной организации труда (совокупность трудовых процессов и отношений), как экономическая система (совокупность производственных отношений), как система техники безопасности (совокупность средств и методов, обеспечивающих безопасную работу и т. д.). Различие в построении систем относительно одного и того же объекта определяется положением или точкой зрения наблюдателя, строящего систему.

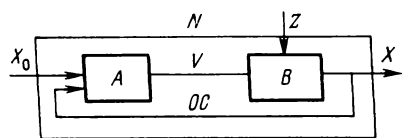


Рис. 3. Схематическое изображение управляемой системы

Равным образом, одна и та же абстрактная система (теория, гипотеза, формула, модель) может использоваться для изучения различных материальных систем, естественных, искусственных или социально-экономических.

Системы могут классифицироваться по ряду признаков: по природе образования, по степени связи с внешней средой, по зависимости от времени, по обусловленности действия, по месту в иерархии систем, по степени сложности и т. д.

Управляемая система. В кибернетике и системном анализе при изучении сложных систем приходится иметь дело с управляемыми системами (рис. 3). В составе управляемой системы N выделены два блока (две подсистемы): управляющая A и управляемая B , которые связаны между собой по меньшей мере двумя каналами связи.

Цель функционирования системы N , характеризующая ее вход X_0 , поступает на управляющую подсистему A , которая вырабатывает управляющие воздействия (командную информацию) V , поступающие на управляемую подсистему B . И управляющая и управляемая подсистемы должны быть динамичными, т. е. допускать возможность появления нескольких (многих) состояний. Между ними должны существовать причинно-следственные (прямые или обратные) связи. Под влиянием управляющих воздействий управляемая система переходит в такое состояние, когда обеспечивается достижение поставленной цели, характеризуемое выходом X . Так как на управляемую систему влияют не только управляющие воздействия V , но и так называемые возмущающие воздействия Z , то результат процесса X может отклоняться от цели. Для ликвидации этих отклонений в управляющую подсистему A все время поступают сигналы от управляемой подсистемы B в порядке осведомительной информации OC . Последняя позволяет управляющей подсистеме A сравнивать результат процесса X с его целью X_0 и вырабатывать на основе этого сличения дополнительные управляющие воздействия на подсистему B для ликвидации отклонений и максимального приближения результата процесса X к его цели X_0 .

В более сложных схемах построения управляемой системы в составе управляющей подсистемы выделяются два блока: управляющий и измерительный органы. Равным образом и в составе управляемой подсистемы выделяются исполнительный орган и блок управляемого объекта. На последние два блока оказывают влияние возмущающие воздействия. Результат процесса с помощью осведомительной информации доводят до сравнительного блока, а затем до управляющего блока. Последний вырабатывает мероприятия по корректировке управляющего воздействия и передает его на

исполнительный орган, который уточняет свою программу с целью максимального сокращения разницы $X - X_0$.

Описание систем. В кибернетике и в системном анализе прибегают к специальному приему — описанию систем. Система может быть описана тремя способами: 1) морфологическим (анализ внутреннего устройства), 2) функциональным (анализ взаимоотношений со средой и между частями системы), 3) информационным (анализ степени неопределенности состояния и его изменения).

Для морфологического описания системы рекомендуется [9] применять следующую четверку:

$$S_M = \{\Sigma, V, \sigma, K\},$$

где $\Sigma = \{\Sigma_i\}$ — множество элементов и их свойств (под элементом понимается подсистема, внутрь которой морфологическое описание не проникает); $V = \{V_j\}$ — множество связей; σ — структура; K — композиция.

Все множества авторы считают конечными.

Как указывают В. В. Дружинин и Д. С. Конторов [9], функциональное описание необходимо для того, чтобы осознать важность системы, определить ее место, оценить отношение к другим системам, ибо без этого невозможно использовать систему, а также создать правильную ориентацию в отношении внешних связей системы, ее контакта с окружающим миром, направлений ее возможного изменения.

Функциональное описание системы включает оценку значимости системы в ее конкретной функции. При функциональном описании предполагают, что всякая система выполняет некоторые функции: а) просто существует, б) служит областью обитания другой системы; в) обслуживает систему более высокого порядка; г) является контрольной для некоторого класса систем; д) служит средством или исходным материалом для создания более совершенной системы и т. д.

Что касается информационного описания систем, то оно должно давать представление об организации системы, о зависимости морфологических и функциональных свойств системы от качества и количества внутренней (в самой себе и среде) и внешней (поступающей из среды) информации. В целом совокупность всех трех описаний системы позволяет отразить ее главные свойства.

Композиция и декомпозиция систем. При исследовании структуры систем (множества всех возможных отношений между подсистемами и элементами внутри системы) приходится прибегать к декомпозиции системы. Под декомпозицией понимают расчленение системы на ее подсистемы, проводимое по различным признакам. Декомпозиция приводит к образованию структуры блоков, групп, элементов сложной системы, которые затем моделируются в зависимости от постановки задачи, ее целей.

Объединение отдельных оптимальных блоков и получение оптимальной (или близкой к оптимальной) сложной системы носят назва-

ние композиции системы. Систему можно построить при помощи композиции, если заданы структурные отношения и состав элементов.

§ 2. Системный подход в современной науке

Системный подход — направление методологии специально-научного познания и социальной практики, в основе которого лежит исследование объектов как систем. Методологическая специфика системного подхода определяется тем, что он ориентирует исследование на раскрытие целостности объекта и обеспечивающих ее механизмов, на выявление многообразных типов связей сложного объекта и сведение их в единую теоретическую картину.

Стремление к целостному охвату объекта изучения, к системной организации знания, всегда свойственное научному познанию, выступает как проблема уже в античной философии и науке. Но вплоть до середины XIX в. объяснение феномена целостности либо ограничивалось уровнем конкретных предметов (типа живого организма), внутренняя целостность которых была совершенно очевидна и не требовала специальных доказательств, либо переносилась в сферу спекулятивных натурфилософских построений. Идея же системной организованности рассматривалась только применительно к знанию.

Задачи адекватного воспроизведения в знании сложных социальных и биологических объектов действительности впервые в научной форме были поставлены К. Марксом, Ф. Энгельсом и Ч. Дарвиным. В. И. Ленин в своих работах, продолжая развивать системный подход, указывал: «... Чтобы действительно знать предмет, надо охватить, изучить все его стороны, все связи и «опосредствования»*.

В литературе подчеркивается, что системный подход к изучению биологических, физических и экономических объектов есть не что иное, как развитие марксистской диалектики с использованием достижений науки в области математики, информационных процессов и электронно-вычислительной техники.

В XX в. системный подход занимает одно из ведущих мест в научном познании. Предпосылкой его возникновения в науке явился прежде всего переход к новому типу научных задач: в целом ряде отраслей науки центральное место начинают занимать проблемы организации и функционирования сложных объектов; познание начинает оперировать системами, границы и состав которых далеко не очевидны и требуют специального исследования в каждом отдельном случае. Во второй половине XX в. аналогичные по типу задачи возникают и в социальной практике: техника все более превращается в технику сложных систем, где многообразные технические и другие средства тесно связаны решением единой крупной задачи (например, космические проекты, человеко-машинные си-

* Ленин В. И. Полн. собр. соч., т. 42, с. 290.

стемы разного рода и др.); в социальном управлении вместо господствующих прежде локальных задач, отраслевых задач и принципов ведущую роль играют крупные комплексные проблемы, требующие тесного взаимодействия экономических, социальных и иных аспектов общественной жизни (например, проблемы создания современных производственных комплексов, развитие городов, мероприятия по охране природы).

Наряду с развитием системного подхода «вширь», т. е. распространением его принципов на новые среды научного знания и практики, с середины XX в. начинается систематическая разработка этих принципов в методологическом плане. Первоначально методологические исследования группировались вокруг задач построения общей теории систем (первая программа ее построения и сам термин предложены Л. Берталанфи). Однако развитие исследований в этом направлении показало, что совокупность проблем методологии системного исследования существенно превосходит рамки задач общей теории систем. Для обозначения этой более широкой сферы методологических проблем и применяют термин «системный подход», который с 70-х годов прочно вошел в научный обиход (в научной литературе разных стран для обозначения данного понятия используют и другие термины — «системный анализ», «системные методы», «системно-структурный подход», «общая теория систем»; при этом за понятием системного анализа в общей теории систем закреплено еще и специфическое, более узкое значение; поэтому термин «системный подход» следует считать более точным, к тому же он наиболее распространен в литературе на русском языке).

Системный подход не существует в виде строгой методологической концепции: он выполняет свои эвристические функции, оставаясь не очень жестко связанной совокупностью познаваемых принципов, основной смысл которых состоит в соответствующей ориентации конкретных исследований. Эта ориентация осуществляется двояко. Во-первых, содержательные принципы системного подхода позволяют фиксировать недостаточность старых, традиционных предметов изучения для постановки и решения новых задач. Во-вторых, понятия и принципы системного подхода существенно помогают строить новые предметы изучения, задавая структурные и типологические характеристики этих предметов и, таким образом, способствуя формированию конструктивных исследовательских программ.

Значение критической функции новых принципов познания было убедительно продемонстрировано еще К. Марксом, «Капитал» которого далеко не случайно носит подзаголовок «Критика политической экономии»: именно последовательная критика принципов классической политэкономии позволила раскрыть узость, недостаточность ее исходной содержательно-концептуальной базы и расчистить путь для построения нового предмета этой науки, адекватного задачам изучения целостного функционирования и развития капиталистической экономики. Решение аналогичных задач

выступает важным предварительным условием и при построении современных системных концепций. Например, переходу к конструированию современных технических систем и возникновению системотехники (которая выступила одной из важных конкретизаций системного подхода в области современной техники) предшествовали осознание и критика подхода, господствующего на прежних ступенях развития техники, когда «единицей» конструирования было отдельное техническое средство (машина, отдельное орудие и т. д.), а не целостная функция (например, технологический процесс), как это имеет место в настоящее время. Условием разработки эффективных мероприятий по защите окружающей среды явилась весьма последовательная критика прежнего подхода к развитию производства, игнорировавшего системную связь общества и природы. Таким образом, эта функция системного подхода носит конструктивный характер и связана прежде всего с обнаружением неполноты наличных предметов изучения, их несоответствия новым научным задачам, а также с выявлением недостаточности применяемых в той или иной области науки и практики принципов объяснения и способов поступления знаний.

Позитивная роль системного подхода может быть сведена к следующим основным моментам. Во-первых, понятие и принципы системного подхода выявляют более широкую познавательную реальность по сравнению с той, которая фиксировалась в переднем знании (например, оптимальный подход в экономическом управлении и планировании).

Во-вторых, системный подход содержит в себе новую по сравнению с предшествующими схему объяснения, в основе которой лежит поиск конкретных механизмов целостности объекта и выявление достаточно полной типологии его связей. Реализация этой функции обычно сопряжена с большими трудностями: для действительно эффективного исследования мало зафиксировать наличие в объекте разнотипных связей, необходимо представить это многообразие в операциональном виде, т. е. изобразить различные связи как логически однородные, допускающие непосредственное сравнение и сопоставление.

В-третьих, из важного для системного подхода тезиса о многообразии типов связей объекта следует, что сложный объект допускает не одно, а несколько расчленений. При этом критерием обоснованного выбора наиболее адекватного расчленения изучаемого объекта может служить то, насколько в результате удается построить операциональную «единицу» анализа, позволяющую фиксировать целостные свойства объекта, его структуру и динамику (например, как товар в экономическом учении К. Маркса).

Использование системного подхода и общей теории систем дает возможность решить многие вопросы организации экономики производства: 1) установить понятие «различие» (формы проявления, количественная и качественная характеристики); 2) определить операнды производственной системы, позволяющие дать оценку

системе с точки зрения ее содержания; 3) выделить операторы окружающей среды, их взаимозависимость, взаимодействие и влияние на вещественный субстрат и структуру внутренних отношений; 4) вскрыть существенные и несущественные переменные производственной системы; 5) выбрать способ описания канонического представления производственной системы на основе определения функций отдельных переменных; 6) определить состояние производственной системы; 7) разработать методику определения диапазона переменной, характеризующего отличие ее настоящего состояния от возможного; 8) найти преобразование переменной до возможного состояния с целью совершенствования системы и ее локальных частей; 9) установить периоды изменения условий для переменной и последовательность состояний, принимаемых системой во времени; 10) определить этапы изменения условий окружающей среды для отдельных переменных и осуществить прогнозирование возможных изменений состояния системы; 12) найти инвариантность производственной системы; 13) дать классификацию производственных систем по размерам в зависимости от числа различий; 14) установить понятие «равновесие» системы; 15) найти возможность расширения обратных связей в системе для качественного регулирования; 16) определить количество и качество возмущающей информации внутри системы и в окружающей среде; 17) установить управляемость системы; 18) вскрыть надежность системы и ее отдельных элементов; 19) установить меру творчества и ответственности участников производственного процесса; 20) определить уровень организации производственной системы; 21) найти параметруемость отдельных процессов.

Приведенный перечень имеет важное значение также для проведения системного анализа производственных систем.

§ 3. Основные логические элементы системного анализа

Системный анализ в узком смысле — совокупность методологических средств, используемых для подготовки и обоснования решений по сложным проблемам политического, военного, социального, экономического, научного, технического характера. В широком смысле термин «системный анализ» иногда (особенно в английской литературе) употребляется как синоним системного подхода.

Привлечение методов системного анализа для решения указанных проблем необходимо прежде всего потому, что в процессе принятия решений приходится осуществлять выбор в условиях неопределенности, которая обусловлена наличием факторов, не поддающихся строгой количественной оценке. Процедуры и методы системного анализа направлены именно на выдвижение альтернативных вариантов решения проблемы, выявление масштабов неопределенности по каждому из вариантов и сопоставление

вариантов по тем или иным критериям эффективности. Специалисты по системному анализу только готовят или рекомендуют варианты решения, принятие же решений остается в компетенции соответствующих должностных лиц (или органа).

«Задача системного анализа заключается в выяснении проблем, стоящих перед лицами, принимающими решения, так, чтобы руководителю стали очевидными все основные последствия принятия решений, и он мог бы учитывать их в своих действиях» [22].

Интенсивное расширение сферы использования системного анализа тесно связано с распространением программно-целевого метода управления, при котором специально для решения важной проблемы составляется программа, формируется организация (учреждение или сеть учреждений) и выделяются необходимые материальные ресурсы. Первой широкой программой такого рода явился план ГОЭЛРО, разработанный в 1920 г. на основе указаний В. И. Ленина.

В 1972 г. в г. Лаксенбурге, близ Вены, создан Международный институт прикладного системного анализа (ПАСА), в котором участвуют 12 стран (в том числе СССР и США); институт ведет работу по применению методов системного анализа преимущественно к решению проблем, требующих международного сотрудничества (например, охрана окружающей среды, освоение ресурсов Мирового океана, совместное использование пограничных водных бассейнов).

Основой системного анализа считают общую теорию систем и системный подход. Системный анализ, однако, заимствует у них лишь самые общие исходные представления и предпосылки. Его методологический статус весьма необычен: с одной стороны, системный анализ располагает детализированными методами и процедурами, почерпнутыми из современной науки и созданными специально для него, что ставит его в ряд с другими прикладными направлениями современной методологии, с другой — в развитии системного анализа отсутствует тенденция к оформлению его в строгую и законченную теорию. В системном анализе тесно переплетены элементы науки и практики. Поэтому далеко не всегда обоснование решений с помощью системного анализа связано с использованием строгих формализованных методов и процедур; допускаются суждения, основанные на личном опыте и интуиции, необходимо лишь, чтобы это обстоятельство было ясно осознано. Важнейшие принципы системного анализа сводятся к следующему: процесс принятия решений должен начинаться с выявления и четкого формулирования конечных целей; надо рассматривать всю проблему и выявлять все последствия и взаимосвязи каждого частного решения; необходимы анализ и выявление возможных альтернативных путей достижения цели; цели отдельного подразделения не должны противоречить целям всей программы.

Центральной процедурой в системном анализе является построение обобщенной модели (или моделей), отображающей все факторы и взаимосвязи реальной ситуации, которые могут про-

явиться в процессе осуществления решения. Полученная модель исследуется с целью выявления близости результата применения того или иного из альтернативных вариантов действий к желаемому, сравнительных затрат ресурсов по каждому варианту, степени чувствительности модели к различным нежелательным внешним воздействиям. Системный анализ опирается на ряд прикладных математических дисциплин и методов, широко используемых в современной деятельности управления: исследование операций, метод экспертных оценок, метод критического пути, теорию очередей и т. д. Техническая основа современного математического анализа — современные ЭВМ и информационные системы.

Методологические средства, применяемые при решении проблем с помощью системного анализа, определяются в зависимости от того, предусматривается ли единственная цель или некоторая совокупность целей, принимает ли решение одно лицо или несколько и т. д. Когда имеется одна достаточно четко выраженная цель, степень достижения которой можно оценить на основе одного критерия, используются методы математического программирования. Если степень достижения цели должна оцениваться на основе нескольких критериев, применяют аппарат теории полезности, с помощью которого проводится упорядочение критериев и определение важности каждого из них. Когда развитие событий определяется взаимодействием нескольких лиц или систем, из которых каждая преследует свои цели и принимает свои решения, используются методы теории игр [6, 10].

Несмотря на то что диапазон применяемых в системном анализе методов моделирования и решения проблем непрерывно расширяется, системный анализ по своему характеру не тождествен научному исследованию: он не связан с задачами получения научного знания в собственном смысле, но представляет собой лишь применение методов науки к решению практических проблем управления и преследует цель рационализации процесса принятия решений, не исключая из этого процесса неизбежных в нем субъективных моментов.

Системный анализ можно рассматривать как некоторый процесс, в результате осуществления которого путем последовательного приближения решаются стоящие перед исследователем задачи.

Заслуживают внимания группировка этапов и работ по системному анализу, приведенная в табл. 1, и перечень применяемых в процессе системного анализа методов исследования [29].

В табл. 1 указаны 12 этапов работы по системному анализу и перечень работ по каждому этапу, а также научные методы системного анализа, применяемые на каждом из этих этапов. Расшифровка этих методов приведена в табл. 2.

Важно отметить, что приведенный в табл. 2 перечень методов и приемов системного анализа неоднороден по своему составу (как по объему выполняемых операций, так и по степени их разработанности). Так, например, экономико-математическое моделирование представляет собой набор ряда методов, которые образуют целую

Последовательность этапов и работ по системному анализу

Этапы и работы	Применяемые методы (см. табл. 2)
<i>I. Анализ проблемы</i>	1, 3, 4, 7
Существует ли проблема?	
Точное формулирование проблемы	
Анализ логической структуры проблемы	
Развитие проблемы (в прошлом и будущем)	
Внешние связи проблемы (с другими проблемами)	
Принципиальная разрешимость проблемы	
<i>II. Определение системы</i>	5, 10
Спецификация задачи	
Определение позиции наблюдателя	
» объекта	
Выделение элементов (определение границ разбиения систем)	
Определение подсистем	
» среды	
<i>III. Анализ структуры системы</i>	3, 5, 6, 8, 10
Определение уровней иерархии (в больших системах)	
Определение аспектов и языков (в сложной системе)	
Определение процессов-функций (в динамической системе)	
Определение и спецификация процессов управления и каналов информации (в кибернетических системах)	
Спецификация подсистем	
» процессов (функций) текущей деятельности (рутинных) и развития (целевых)	
<i>IV. Формулирование общей цели и критерия системы</i>	2, 4, 7, 8, 10, 12
Определение целей-требований подсистемы	
» целей и ограничений среды	
Формулирование общей цели	
Определение критерия	
Декомпозиция целей и критериев по подсистемам	
Композиция общего критерия из критериев подсистем	
<i>V. Декомпозиция цели, выявление потребности в ресурсах и процессах</i>	4, 6, 11
Формулирование целей верхнего ранга	
» целей текущих процессов	
» целей эффективности	
» целей развития	
» внешних целей и ограничений	
Выявление потребностей в ресурсах и процессах	
<i>VI. Выявление ресурсов и процессов, композиция целей</i>	2, 4, 7
Оценка существующих технологии и мощностей	
Оценка современного состояния ресурсов	
» реализуемых и запланированных проектов	
Оценка возможностей взаимодействия с другими системами	
Оценка социальных факторов	
Композиция целей	

Этапы и работы	Применяемые методы (см. табл. 2)
<i>VII. Прогноз и анализ будущих условий</i> Анализ устойчивых тенденций развития системы Прогноз развития и изменения среды Предсказание появления новых факторов, оказывающих сильное влияние на развитие системы Анализ ресурсов будущего Комплексный анализ взаимодействия факторов будущего развития Анализ возможных сдвигов целей и критериев	1, 2, 4, 6, 7, 9, 11
<i>VIII. Оценка целей и средств</i> Вычисление оценок по критерию Оценка взаимозависимости целей » относительной важности целей » дефицитности и стоимости ресурсов Оценка влияния внешних факторов Вычисление комплексных расчетных оценок	2, 7, 8
<i>IX. Отбор вариантов</i> Анализ целей на совместимость и входимость Проверка целей на полноту Отсечение избыточных целей Планирование вариантов достижения отдельных целей Оценка и сравнение вариантов Совмещение комплекса взаимосвязанных вариантов	4, 5, 7, 8
<i>X. Диагноз существующей системы</i> Моделирование технологического и экономического процессов Расчет потенциальной и фактической мощности Анализ потерь мощности Выявление недостатков организации производства и управления Выявление и анализ мероприятий по совершенствованию организации	3, 5, 7, 10
<i>XI. Построение комплексной программы развития</i> Формулирование мероприятий, проектов и программ Определение очередности целей и мероприятий по их достижению Распределение сфер деятельности » » компетенции Разработка комплексного плана мероприятий в рамках ограничений по ресурсам во времени Распределение по ответственным организациям, руководителям и исполнителям	5, 6, 7, 11, 12
<i>XII. Проектирование организации для достижения целей</i> Назначение целей организации Формулирование функций организации Проектирование организационной структуры » информационных механизмов » режимов работы » механизмов материального и морального стимулирования	3, 4, 5, 6, 10

Научный инструментарий системного анализа

Инструменты системного анализа	Этапы, на которых они применяются
<i>А. Неформальные методы</i> 1. Метод сценариев 2. Метод экспертных оценок («Дельфи») 3. Диагностические методы <i>Б. Графические методы</i> 4. Метод деревьев целей 5. Матричные методы 6. Сетевые методы <i>В. Количественные методы</i> 7. Метод экономического анализа 8. Морфологические методы 9. Статистические методы <i>Г. Методы моделирования</i> 10. Кибернетические модели 11. Описательные модели 12. Нормативные операционные модели (оптимизационные, имитационные, игровые)	
	I, VII VI, VII, VIII I, III, X, XII I, IV, V, VII, IX, XII II, III, IX, X, XI, XII III, V, VII, XI, XII I, VI, VII, VIII, IX, X, XI III, IV, VIII, IX VII II, III, IV, X, XII VI, VII, XI IV, XI

научную дисциплину, а экспертные оценки — единичный метод, используемый для специфического круга задач; экономический анализ деятельности предприятий разрабатывается в течение ряда десятилетий, а диагностика системы находится в начальной стадии разработки. Поэтому степень детализации описания методов и приемов зависит от перечисленных обстоятельств.

§ 4. Системный подход к анализу основных технико-экономических показателей развития топографо-геодезического производства

Топографо-геодезическое производство можно рассматривать как систему, входящую в состав более крупной суперсистемы — народное хозяйство. Топографо-геодезическое производство, как система, имеет свои входы и выходы, осуществляет процесс переработки материально-вещественных ресурсов в готовую продукцию (материалы), решает проблемы по достижению определенных целей своего развития.

Остановимся на некоторых особенностях и свойствах топографо-геодезического производства как системы. Как всякая система, топографо-геодезическое производство состоит из множества систем и подсистем со значительным числом взаимосвязанных и взаимодействующих элементов (объектов). Совокупность таких элементов (производственные объединения, предприятия, НИИ и т. д.) представляет собой определенную целостность, обладающую общими свойствами объектов по обеспечению целей производства.

Свойство делимости топографо-геодезического производства, как системы, проявляется в возможности выделения отдельных предприятий, которые, в свою очередь, делятся на экспедиции, комплексные экспедиции, партии, цехи, бригады и т. п. В составе каждого из выделенных объектов можно выделить элементы: трудовые ресурсы Т, основные производственные фонды Ф, основные материалы М, выпускаемые виды работ по себестоимости С и в натуре N и т. д.

Топографо-геодезическое производство, как система, обладает свойством автономности (изолированности): каждый ее объект или элемент можно рассматривать отдельно от остальных элементов. В то же время эта изолированность является относительной, так как все элементы через свои входы и выходы взаимосвязаны между собой и с внешней средой, т. е. с другими системами (отраслями народного хозяйства).

Идентификация (отделение) объекта может быть выполнена путем определения его характеристик с применением метода «черного ящика», построения экономико-математических моделей и разработки матриц, описывающих зависимость между входными и выходными переменными объектов и между переменными протекающих в них процессах.

Свойство разнообразия системы топографо-геодезического производства проявляется в том, что каждый ее элемент обладает собственным поведением и состоянием, отличным от поведения и состояния других элементов и системы в целом. Все входы и выходы объектов являются наблюдаемыми и контролируемыми.

Топографо-геодезическое производство обладает свойством неопределенности, так как нельзя заранее фиксировать все свойства и отношения ее элементов без проведения специального исследования.

Важным свойством системы топографо-геодезического производства является ее отображаемость с помощью различной статистической, бухгалтерской, оперативной информации и отчетности о работе. Вследствие этого свойства топографо-геодезического производства как экономической система обладает и свойством нетождественности отображения, ибо представление об ее состоянии, которое получает наблюдатель, всегда имеет какое-то отставание относительно состояния системы в данный момент.

Как уже отмечалось, топографо-геодезическое производство как самостоятельная экономическая система имеет все элементы: вход, выход, процесс, обратную связь, цель и т. п. Структура си-

стемы элементов, входящих в различные потоки входа, выхода и процесса, может быть представлена таким образом.

Вход X

Процесс Y

$X = \{T, \Phi, M, D, И, У, \dots\}$ $Y = \{ТЗ, \PhiЗ, З, М, ТУ, \dots\}$

Выход Z

$Z = \{B, N, И, Э, Т', \Phi^1, Д^1, \dots\}$

Множество показателей входа X включает: T — трудовые ресурсы; Φ — основные производственные фонды; M — основные материалы; D — ассигнования из госбюджета и другие денежные средства (кредиты, ассигнования на капитальные вложения и т. д.); $И$ — информация, включая планы, приказы, распоряжения вышестоящих органов, инструкции, нормативные акты, заявки от министерств и ведомств и др.; $У$ — услуги, получаемые от других организаций и систем.

Связи по входу преимущественно носят управляемый характер по отношению к вышестоящим системам.

Множество показателей процесса Y включает: $ТЗ$ — затраты труда (рабочего времени); $\PhiЗ$ — заработная плата работающих; $З$ — заделы по смежным процессам работ; $М$ — мощность полевого и камерального производства; $КМ$ — коэффициенты соответствия основных элементов производства по мощности; $ТУ$ — достигнутый технический уровень производства.

Множество потоков выхода Z включает: B — выпускаемые работы по видам; N — выпуск работ по видам в натуре; $И$ — поток информации, представляемой в вышестоящие органы, заказчикам, а также системам, подчиненным данной системе; $Э$ — народнохозяйственный эффект, получаемый в результате использования топографо-геодезических работ в народном хозяйстве; $Т'$ — поток выбывающих из системы работников; Φ' — поток убывающих основных фондов; $Д'$ — поток денежных средств.

Каждый из приведенных элементов системы топографо-геодезического производства, в свою очередь, является подмножеством. Например, T — трудовые ресурсы можно рассматривать как множество элементов: $T = \{T_о, T_в, T_{итр}, T_с, T_н, \dots\}$, т. е. численность основных работников, вспомогательных, инженерно-технических работников, служащих, научных работников и т. д.

Рассмотренные связи позволяют определить систему технико-экономических показателей топографо-геодезического производства. Связь между технико-экономическими показателями и характеристиками параметров сложной системы топографо-геодезического производства может быть представлена в виде уравнений статики и динамики, которые в общем случае являются математическим описанием процесса. Эти уравнения связывают между собой технико-экономические показатели, характеристики условий топографо-геодезического производства и технологические параметры и поэтому позволяют в некотором интервале времени судить о необходимых управляющих воздействиях, создавая основу для более

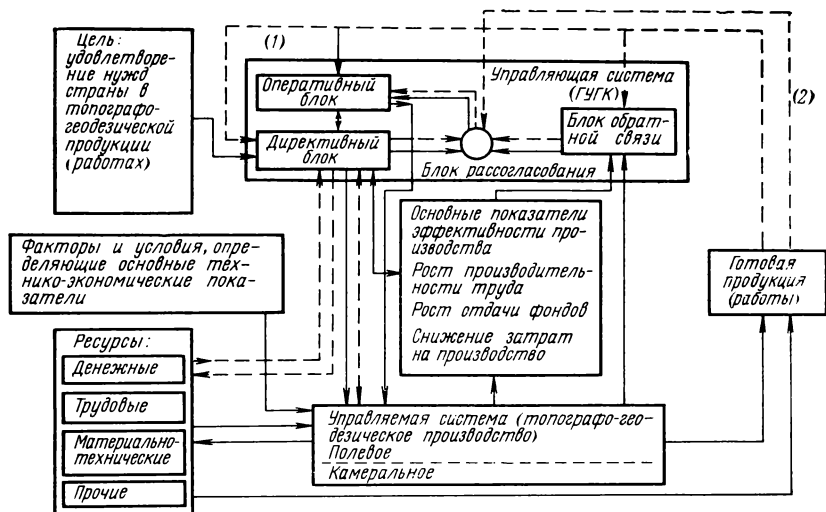


Рис. 4. Общая схема топографо-геодезического производства

действенного управления эффективностью как системой. Таким образом, возникает замкнутая, основанная на кибернетических принципах система эффективности топографо-геодезического производства, которая приводит требования централизованного планирования и управления в соответствие с принципом самостоятельности руководства производственным процессом на предприятии.

В свою очередь, система эффективности состоит из множества подсистем, поэтому случай, при котором выходные технико-экономические показатели не отвечают предъявляемым требованиям, не означает, что все подсистемы функционируют неудовлетворительно. Обычно не соответствуют условиям эффективного производства так называемые узкие места. Например, может возникнуть ситуация, когда на отдельных предприятиях при высокой производительности труда и низкой себестоимости существует низкая фондоотдача. Поэтому правильное рассмотрение взаимосвязанных элементов подсистем позволяет вскрыть причины отдельных недостатков деятельности, дать конкретные рекомендации по результатам технико-экономического анализа, выявить и мобилизовать резервы производства.

Отсюда следует, что системный подход к анализу технико-экономических показателей предполагает сопоставимость различных ресурсов, факторов и условий работ в соответствии с их способностью к достижению цели, т. е. максимального удовлетворения нужд страны в топографо-геодезических работах. Из рис. 4 видно, что вход системы, т. е. поступление ресурсов, является лишь частично регулируемым, образуя открытый конец системы. Другой

открытый конец — реализация целей системы (так называемое чистое потребление). Связь (1) — это обратная связь, по которой поступает информация о том, насколько топографо-геодезическое производство удовлетворяет заданной величине потребности. По каналу (2) задаются параметры выхода системы и осуществляется система регулирования производством. Таким образом, управление технико-экономическими показателями осуществляется передачей информации (1) и (2) в блок регулирования, а оттуда — в директивный блок и прямым обменом информации между директивным блоком и системой основных технико-экономических показателей [15].

Рассмотрение схемы системы управления топографо-геодезическим производством, ее функций и связей дает возможность сформулировать основные задачи и направления системного анализа основных технико-экономических показателей развития отрасли. Они сводятся к трем основным аспектам.

1. Анализ структуры системы, ее иерархии, состава подсистем, их элементов и объектов, их связей. Этот аспект анализа получил название морфологического, так как его главная цель состоит в получении морфологического описания системы топографо-геодезического производства и разработки управленческих решений по его улучшению.

2. Анализ отдельных функций топографо-геодезического производства. Цель этого направления сводится к обеспечению принятия управленческих решений по совершенствованию функции работы отрасли.

3. Анализ результатов работы топографо-геодезического производства. Он охватывает анализ показателей, характеризующих итоги работ отрасли, имеющиеся в ней резервы и факторы, способствующие лучшему ее функционированию. Это направление решает задачу как комплексной оценки всей системы показателей, характеризующих эффективность топографо-геодезического производства, так и отдельных показателей, отражающих отдельные стороны работы. Здесь определяется достигнутый уровень развития топографо-геодезического производства (в целом и по направлениям деятельности); исследуется динамика (темпы развития системы и ее подсистем, ход выполнения плана); анализируются затраты и эффективность, выявляются тенденции их развития; определяется влияние факторов и резервов, потерь и неиспользованных возможностей на развитие системы и т. д.

Принципы построения матричной модели показателей работы экономической системы

Подмножества показателей, определяющих вход, процесс и выход топографо-геодезического производства, взаимосвязаны (рис. 5).

Связи (1), (5), (9) можно рассматривать как петли графа, остальные — как ориентированные дуги, а показатели входа (X), процесса (Y) и выхода (Z) — как вершины графа.

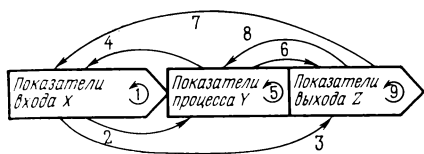


Рис. 5. Связи показателей входа, процесса и выхода

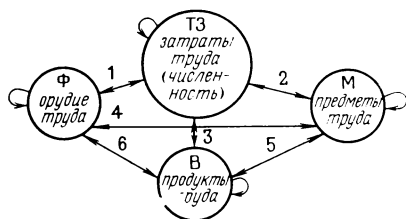


Рис. 6. Граф связей элементов

Связь между элементами входа, выхода и производственного процесса (процесса труда) можно представить в виде графа (рис. 6).

Из графа видно, что он охватывает четыре вершины (ТЗ, Ф, М, В), шесть ориентированных дуг, выражающих прямые и обратные им связи между элементами (вершинами) и четыре петли-дуги, начало и конец которых совпадают (связи ТЗ—ТЗ, Ф—Ф, М—М, В—В).

Как видно из рис. 6, рабочий своим трудом воздействует на орудия труда (связь 1) и с их помощью (связь 4) превращает предметы труда в готовые продукты труда (связь 5). В некоторых случаях воздействие рабочей силы на продукты труда совершается с помощью связей (1)—(6), например перевозка готовой продукции на электрокаре; связей (2)—(5) — ручная обработка предметов труда; связи (3) — упаковка продуктов труда вручную.

Связи между элементами (простыми моментами) процесса труда (производства) могут быть выражены с помощью пропорций, показывающих их отношение друг к другу.

Пользуясь принципом пропорциональности и соотнося различные показатели входа, процесса и выхода экономической системы друг с другом, можно построить матричную модель (табл. 3). В матричную модель для примера включены пять показателей, три из которых являются показателями входа (ТЗ, Ф и М), один — показатель процесса (С) и один — выхода (В). Все эти показатели можно условно назвать исходными.

Табл. 3 содержит в тексте ряд известных и ряд новых показателей, которые можно назвать показателями первого порядка.

Строки-векторы представляют собой показатели трудоемкости, фондоемкости, материалоемкости, себестоимости и выработки продукции (работ).

В свою очередь, каждый вектор-столбец матрицы включает затраты рабочей силы (рабочего времени), основных фондов, предметов труда, себестоимости, стоимости продукции (работ) на единицу каждого из перечисленных ресурсов.

Представленная в табл. 3 матрица обладает следующими свойствами:

1) каждому элементу, находящемуся выше главной диагонали (обозначенной цифрой 1), соответствует обратный ему элемент,

Т а б л и ц а 3

Матричная модель технико-экономических показателей топографо-геодезического производства

Шифр столбцов	Шифр строк	Показатели входа X			Показатели процесса Y	Показатели выхода Z
		Рабочая сила (численность) или затраты труда — рабочего времени	Орудия и средства труда (основные фонды)	Предметы труда (материальные затраты)		
	00	01	02	Oi	Oj	On
TЗ (чел.)	10	$TЗ_{TЗ} = 1$ (11)	$TЗ_{\Phi} = TЗ/\Phi$ (12)	$TЗ_M = TЗ/M$ (1i)	$TЗ_C = TЗ/C$ (1j)	$TЗ_B = TЗ/B$ (1n)
Ф (тыс. руб.)	20	$\Phi_{TЗ} = \Phi : TЗ$ (21)	$\Phi_{\Phi} = 1$ (22)	$\Phi_M = \Phi/M$ (2i)	$\Phi_C = \Phi/C$ (2j)	$\Phi_B = \Phi/B$ (2n)
M (тыс. руб.)	iO	$M_{TЗ} = M/TЗ$ (i1)	$M_{\Phi} = M/\Phi$ (i2)	$M_M = 1$ (ii)	$M_C = M/C$ (ij)	$M_B = M/B$ (in)
C (тыс. руб.)	jO	$C_{TЗ} = C/TЗ$ (j1)	$C_{\Phi} = C/\Phi$ (j2)	$C_M = C/M$ (ji)	$C_C = 1$ (jj)	$C_B = CB$ (jn)
B (тыс. руб.)	nO	$B_{TЗ} = B/TЗ$ (n1)	$B_{\Phi} = B/\Phi$ (n2)	$B_M = B/M$ (ni)	$B_C = B/C$ (nj)	$B_B = 1$ (nn)

находящийся на таком же расстоянии от главной диагонали, и наоборот. Эта зависимость может быть выражена следующим равенством: $A_{ij} = 1 : A_{ji}$ или $M_c = 1 : C_m$;

2) при повышении эффективности производства и при данном расположении показателей табл. 3 все показатели, находящиеся выше главной диагонали, имеют тенденцию к снижению, а ниже главной диагонали — к повышению;

3) каждый показатель, находящийся в тексте матрицы, является частным от деления какого-либо показателя подлежащего на показатель сказуемого $A_{2j} = A_{20} : A_{0j}$ или $\Phi_c = \Phi : C$;

4) каждый показатель подлежащего таблицы равен произведению любого показателя, помещенного в данной строке в тексте таблицы, на находящийся в этом же векторе-столбце показатель сказуемого $A_{i0} = A_{in} \cdot A_{0n}$ или $M = M_b \cdot B$; $C = M_{тз} \cdot TЗ$;

5) каждый показатель в сказуемом таблицы равен частному от деления любого показателя, стоящего в подлежащем таблицы, на показатель, находящийся в тексте таблицы на пересечении данных вектор-строки и вектор-столбца: $A_{0i} = A_{10} : A_{1i}$ или $M = TЗ : TЗ_m = \Phi : \Phi_m$.

Перечисленные закономерности дают возможность при системном анализе применять метод замены одних показателей другими. Например, показатель выработки на одного работника $B_{тз} = B : TЗ$ можно представить в виде следующих равенств: $B_{тз} = B_\phi \cdot \Phi_{тз} = B_m \cdot M_{тз} = B_c \cdot C_{тз}$ и т. д.: это означает, что показатель выработки продукции равен произведению фондоотдачи на фондоемкость труда и т. д. Указанный показатель можно выразить и как частное от деления одного показателя на другой: $B_{тз} = B_\phi : TЗ_\phi = B_m : M_{тз}$ и т. д.

Можно отметить, что вышеописанный метод анализа и система моделей показателей позволяют комплексно характеризовать в одной таблице результаты работы анализируемой системы не одним показателем, а их совокупностью — системой показателей.

Рассмотрим пример расчета матрицы основных показателей топографо-геодезического производства (табл. 4).

Исходная вектор-строка индексов изменения показателей Ψ , Φ , M и B имеет вид

$$I_B > I_\Phi > I_M > I_\Psi,$$

или в цифровом выражении

$$107,5 > 106,0 > 104,2 > 102,2.$$

Остальные неравенства, составленные по столбцам, будут иметь следующий порядок элементов:

$$I_{B_\Psi} > I_{\Phi_\Psi} > I_{M_\Psi} > 1, \text{ или } 105,2 > 103,7 > 102,0 > 100;$$

$$I_{B_\Phi} > 1 > I_{M_\Phi} > I_{\Psi_\Phi}, \text{ или } 101,4 > 100 > 98,3 > 96,4;$$

$$I_{B_M} > I_{\Phi_M} > 1 > I_{\Psi_M}, \text{ или } 103,2 > 101,7 > 100 > 98,0;$$

$$1 > I_{B_B} > I_{M_B} > I_{\Psi_B}, \text{ или } 100 > 98,6 > 96,9 > 95,0.$$

Как видно из приведенных неравенств, индексы развития топографо-геодезического производства позволяют оценить тип и формы

Матрица технико-экономических показателей работы
топографо-геодезического производства

	Численность Ч	Основные фонды Ф	Материалы М	Продукция В
Индекс	$I_{\text{ч}} = 102,2$	$I_{\text{ф}} = 106,0$	$I_{\text{м}} = 104,2$	$I_{\text{в}} = 107,5$
<i>Трудоемкость</i>				
Численность Ч $I_{\text{ч}} = 102,2$	1	$I_{\text{чф}} = 96,4$	$I_{\text{чм}} = 98,0$	$I_{\text{чв}} = 95,0$
<i>Фондоёмкость</i>				
Основные фонды Ф $I_{\text{ф}} = 106,0$	$103,7 = I_{\text{фч}}$	1	$101,7 = I_{\text{фм}}$	$98,6 = I_{\text{фв}}$
<i>Материалоемкость</i>				
Материалы М $I_{\text{м}} = 104,2$	$102,0 = I_{\text{мч}}$	$98,3 = I_{\text{мф}}$	1	$96,9 = I_{\text{мв}}$
<i>Выработка продукции</i>				
Продукция В $I_{\text{в}} = 107,5$	$105,2 = I_{\text{вч}}$	$101,4 = I_{\text{вф}}$	$103,2 = I_{\text{вм}}$	1
П р и м е ч а н и е. Даны условные значения показателей.				

расширенного воспроизводства. Из того обстоятельства, что индекс роста основных фондов опережает индекс роста численности основного производства, можно сделать вывод, что за рассматриваемый период существует интенсивное воспроизводство (при равенстве $I_{\text{ф}} = I_{\text{ч}}$ было бы экстенсивное воспроизводство). Индекс роста фондовооруженности больше индекса роста производительности труда $I_{\text{фч}} > I_{\text{вч}}$, а индекс $I_{\text{фч}}$ больше единицы ($I_{\text{фч}} > 1$), следовательно, в отрасли существует фондоемкая форма интенсивного производства. Она же является нематериалоемкой формой, что видно из зависимости $B_{\text{ч}} > M_{\text{ч}}$.

Рассмотренная матричная модель системы технико-экономических показателей топографо-геодезического производства может найти широкое применение при решении основных задач и направлений систематического анализа развития отрасли.

В заключение отметим, что если продолжить работу над матричной моделью второго, третьего порядков, то можно получить значительное число новых показателей, в которых будут объединяться не два-четыре элемента, а пять показателей и даже больше, если увеличить число исходных показателей. Все это дает методологию проектирования экономических показателей, основанную на характеристике связей между различными элементами экономических систем.

Новые показатели могут быть использованы при построении систем для принятия управляющих решений.

§ 5. Реализация системного подхода при исследовании вопросов комплексного управления качеством топографо-геодезических работ

Главное управление геодезии и картографии (ГУГК) при Совете Министров СССР вопросам повышения качества топографо-геодезических работ уделяет много внимания. Выполнению намеченных задач во многом способствуют научные разработки Л. А. Кашина [12, 13, 14], И. А. Кутузова [18], В. М. Сигалова [24], А. С. Земцева [11] и других авторов.

По современным понятиям, управление качеством продукции является составной частью управления процессом производства.

Комплексная система управления качеством геодезических и картографических работ (КС УКГКР) — это совокупность мероприятий, методов и средств, направленных на установление и поддержание необходимого уровня качества продукции при ее разработке, производстве и эксплуатации. Система управления качеством создается с учетом уже существующей структурной и функциональной схемы предприятия и его подразделений. Ее не рекомендуется рассматривать в отрыве от общей системы управления производством.

Все известные системы управления качеством можно отнести к классу сложных вероятностных систем. Поэтому в КС УКГКР процесс управления децентрализован. Комплексные системы разбиты на ряд подсистем, связанных между собой по горизонталям и уровням управления. В КС УКГКР, как и в любом контуре управления, эффективность функционирования системы определяется в первую очередь правильно выбранной целью.

Однако повышение качества работ имеет не только технические, но и экономические пределы, которые, как правило, лимитируют производство. Качество работ целесообразно повышать до тех пределов, когда изделие будет наиболее полно удовлетворять те потребности, для которых оно предназначено с минимумом затрат на производство и потребление, т. е. качество должно быть оптимальным. Оптимальный уровень качества определяется минимумом суммы затрат у изготовителя и потребителя продукта в результате применения последнего.

Кроме определения цели в комплексной системе управления качеством многое зависит от разработки и осуществления программ по достижению поставленной цели. К ним относятся мероприятия по совершенствованию технико-производственной базы, организации производства, улучшению структуры управления, повышению культуры производства, подготовке кадров и т. д.

Функционирование КС УКГКР, как и любой системы, невозможно без налаженной работы по сбору и анализу информации по качеству с целью ликвидации отклонений в управляемой системе. Поэтому значение контроля исполнения принимаемых решений в условиях КС УКГКР значительно возрастает.

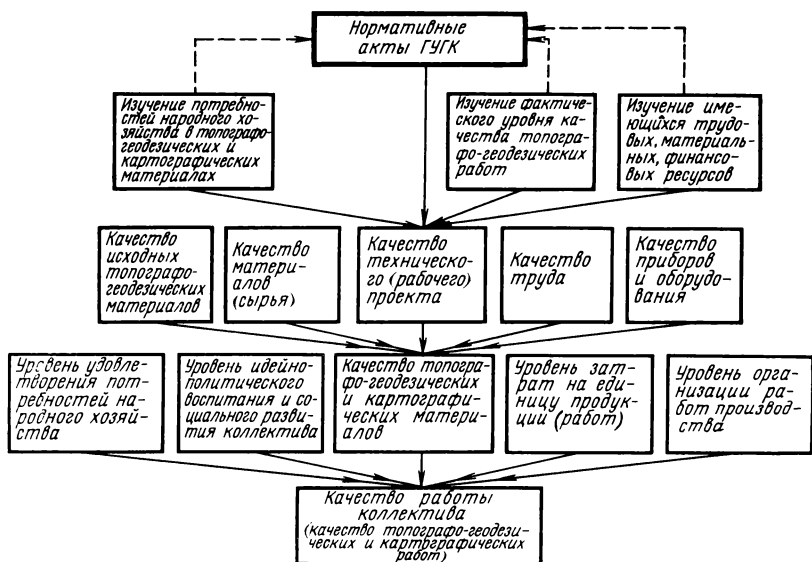


Рис. 7. Структура слагаемых качества топографо-геодезических работ

Функционирование КС УКГКР невозможно без системы объективных оценок качества. Достоверная оценка показателей качества требуется при прогнозировании возможных потребностей, планировании работ, развитии социалистического соревнования, стимулировании трудящихся, определении экономического эффекта от внедрения КС УКГКР. Поэтому в науке наметилось новое направление — квалиметрия, предметом изучения которой являются измерение и оценка качества.

Для внедрения в топографо-геодезическое производство результатов научных исследований и практических рекомендаций, разработанных применительно к промышленным предприятиям, требуется творческий подход, учитывающий специфику топографо-геодезического производства, местные условия каждого предприятия.

Качество топографо-геодезических работ — это совокупность свойств топографических, геодезических материалов, определяющих их пригодность в соответствии с назначением удовлетворять потребности народного хозяйства и обороны страны в своевременной, полной и достоверной информации о местности и ее предметах (рис. 7). Таким образом, в настоящей работе качество топографо-геодезических материалов отождествляется с качеством топографо-геодезических работ (ТГР).

Управление качеством топографо-геодезических работ в государственном масштабе призвано обеспечить Главное управление геодезии и картографии. Управление качеством топографо-геодези-

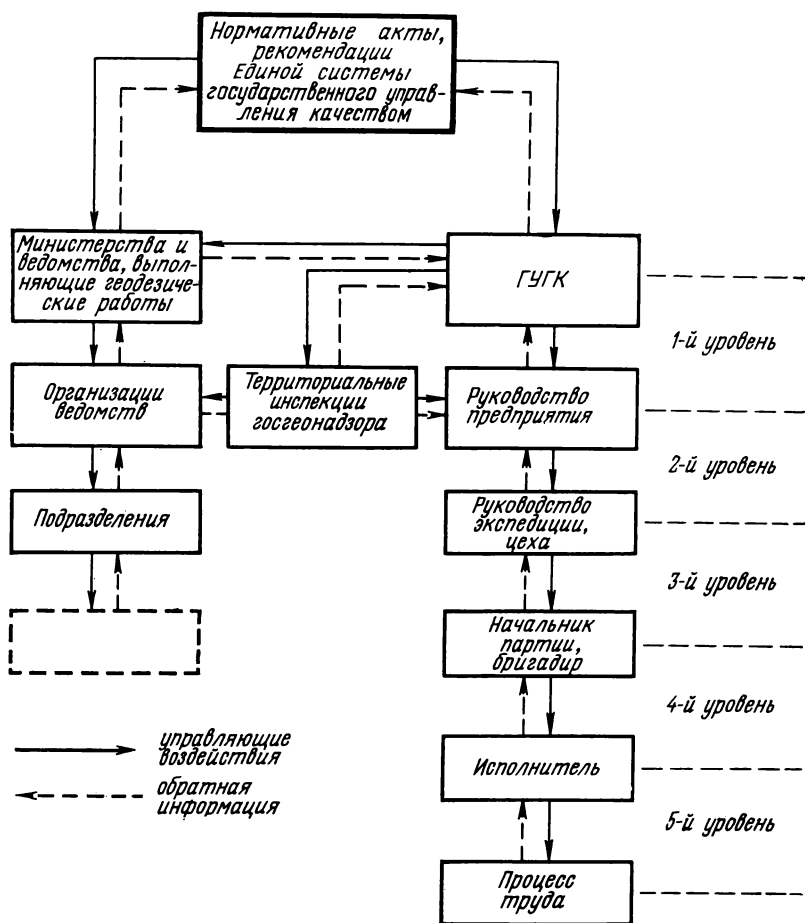


Рис. 8. Иерархия управления качеством топографо-геодезических работ

ческих работ должно осуществляться на всех уровнях управления топографо-геодезическим производством: межотраслевым — управлением государственного геодезического надзора с учетом рекомендаций научно-исследовательских институтов, отраслевым — управлением топографо-геодезической службы ГУГК, в предприятии — аппаратом управления предприятий, в территориальном отношении — территориальными инспекциями государственного геодезического надзора (рис. 8).

В задачу оценки уровня качества, по современным понятиям, входят: выбор номенклатуры оценочных показателей; разработка методов определения их численных значений; выбор базовых величин; определение фактически достигнутых значений показателей качества и их сопоставление с базовыми величинами; анализ и

обоснование рекомендаций для выработки управляющих воздействий.

Так, определив значения показателей качества P_i и их базовых значений $P_i^{\text{баз}}$, уровень качества K_i по i -му показателю можно оценить в общем виде с помощью функции

$$K_i = f(P_i, P_i^{\text{баз}}). \quad (1.1)$$

По формуле (1.1) получают дифференциальную оценку качества. Однако основная проблема заключается в получении не дифференциальных, а комплексных оценок. Последние можно получить двумя путями. Рассмотрим их. В случае, когда свойства топографо-геодезических материалов выражаются сравнительно небольшим числом показателей, из которых выделяются главные, отражающие основное назначение этих материалов, уровень качества можно обозначить комплексным оценочным числом K_0 , которое можно получить построением математической модели, отражающей внутреннюю взаимосвязь оценочного числа с дифференциальными показателями в виде

$$K_0 = A_1 P_1 + A_2 P_2 + \dots + A_n P_n + C.$$

или

$$K_0 = (A_1 P_1) (A_2 P_2) \dots (A_n P_n) C,$$

где $P_1, P_2, \dots, P_n$ — дифференциальные показатели, $A_1, A_2, \dots, A_n$ — коэффициенты дифференциальных показателей, C — постоянный коэффициент.

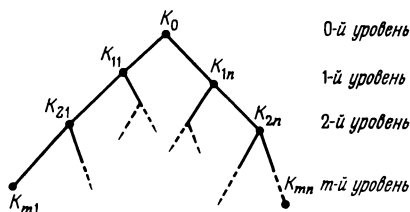


Рис. 9. Иерархия комплексных оценок качества

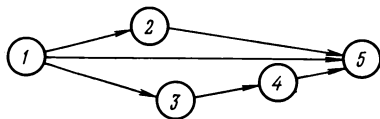


Рис. 10. Очередность и взаимосвязь системы оценки качества топографо-геодезических работ:

1 — оценка качества исполнения, 2 — оценка качества труда, 3 — пообъектная оценка качества ТГР, 4 — оценка качества совокупности работ, выполненных предприятием, 5 — комплексная оценка качества работы коллектива

В другом случае используют коэффициенты относительной важности показателей качества, когда зависимость (1.1) принимает вид

$$K_0 = f(P_i, P_i^{\text{баз}}, g_i, n), \quad (i = 1, 2, \dots, n), \quad (1.2)$$

где K_0 — комплексная оценка, g_i — коэффициент весомости показателей, n — число показателей.

Для того, чтобы охарактеризовать качество сложных топографо-геодезических материалов, представим на основе принципов системного анализа их комплексную оценку K_0 как функцию от комплексных групповых оценок $K_1, K_2, \dots, K_n$ (рис. 9).

Например, качество топографической съемки можно выразить комплексными оценками ее плано-высотного обоснования, съемки объектов ситуации, рельефа, внешнего оформления и т. п.

При построении иерархической структуры показателей качества топографо-геодезических работ необходимо опуститься до такого уровня рассмотрения, когда появится возможность однозначной оценки комплексных или единичных показателей по формуле (1.1).

Вид функциональной зависимости (1.1) применительно к конкретному показателю свойств оцениваемых материалов различен. В каждом отдельном случае его необходимо определять специальными исследованиями. Очень часто для простоты оценку по каждому показателю получают по формуле вида

$$K_i = \frac{P_i}{P_i^{\text{баз}}} \quad \text{или} \quad K_i = \frac{P_i^{\text{баз}}}{P_i}. \quad (1.3)$$

Затем комплексную оценку находят как среднее весовое значение из K_i . Чаще всего используют величины: среднюю арифметическую

$$K_0 = \sum_{i=1}^n q_i K_i \quad (1.4)$$

или геометрическую

$$K_0 = \prod_{i=1}^n K_i^{g_i}, \quad \text{где} \quad \sum_{i=1}^n g_i = 1, \quad 0 < g_i \leq 1. \quad (1.5)$$

Как дифференциальные, так и комплексные оценки качества получают при этом в виде безразмерных относительных величин (индексов).

А. А. Дразнюк [8] показал, что задачи совершенствования системы оценки качества в топографо-геодезическом производстве необходимо решать в очередности и взаимосвязи (рис. 10).

Важно отметить, что исходя из принципов системного подхода необходимо, чтобы различные системы оценок разрабатывались не параллельно, а были соподчинены. Показатели качества исполнения должны войти составной частью в комплексные оценки качества труда, пообъектную оценку качества ТГР и вместе с последними учитываться при оценке результатов деятельности коллективов подразделений и предприятия в целом.

В работе А. А. Дразнюка [8] приводятся результаты определения коэффициентов весомости g_i экспертным методом. Согласованность мнений экспертов относительно коэффициентов весомости определялась с помощью коэффициентов вариации. Полученные результаты свидетельствуют о достижении хорошей согласованности мнений экспертов уже в третьем туре их опроса.

Для выявления возможностей определения коэффициентов весо-

мости оценочных показателей по методу построения экономико-математической модели А. А. Дразнюком получено уравнение множественной регрессии:

$$y = +35,4 - 0,034 x_1 + 0,082 x_2 + 0,858 x_3 + 0,496 x_4, \quad (1.6)$$

где x_1 — процент перевыполнения норм выработки, x_2 — процент сдачи работ с первого предъявления, x_3 — процент работ, выполненных на отлично, x_4 — процент работ, сданных с личным штампом, y — затраты подразделений на 1 руб. сметной стоимости выполненных работ, коп. Однако логический анализ уравнения (1.6) показывает, что знаки и абсолютные величины некоторых параметров уравнения не соответствуют априорным экономическим представлениям о влиянии соответствующих факторов на себестоимость работ в подразделении. Следовательно, на современном этапе исследований не всегда можно определять коэффициенты весомости показателей методом построения экономико-математической модели. Дело все в том, что исследование вопросов комплексного управления качеством ТГР относится к разряду неструктуризованных проблем, для решения которых существуют свои методы, в том числе экспертный метод.

Глава II

УПРАВЛЕНИЕ, ПРИНЯТИЕ РЕШЕНИЙ

§ 6. Основные черты процесса управления

Повсюду в окружающем нас мире (природе, технике, человеческом обществе) протекают различные процессы, характер которых зависит от множества сопутствующих им условий и факторов. Изменяя условия протекания процессов, человек может влиять на их характер, изменять их, приспособлять к своим целям. Это вмешательство в естественный ход процесса, изменение естественного хода процесса и представляет собой сущность управления. Таким образом, можно сказать, что управление представляет собой такую организацию того или иного процесса, которая обеспечивает достижение определенных целей.

На рис. 11 представлена схема управления процессом (операцией), которая в общем случае представляет собой непрерывную процедуру в системе с обратной связью. На схеме можно отчетливо различить три этажа, характерные для любого процесса управления: 1) сбор и обработку информации с целью оценки сложившейся ситуации, 2) принятие решения о наиболее целесообразных действиях, 3) исполнение (передача) принятого решения.

Информация о ходе процесса поступает к управляющему органу. Здесь производится анализ информации, сравнение ее с тем, что желательно было бы иметь, принятие решения, оформление распоряжений и передача их. Управляемый процесс (операция) так или иначе реагирует на получение распоряжения, при этом на характер его функционирования оказывают влияние внешние условия. Информация об изменении процесса (операции) на распоряжение (управление) вновь поступает в систему управления, чем и осуществляется обратная связь.

Подача и обработка информации о ходе управляемого процесса и о внешних условиях (т. е. об обстановке) — важная часть управления. Если в систему управления будет поступать недостаточное количество информации, то качественное управление системой осуществить не удастся. Наоборот, если в систему управления будет поступать чрезмерно большое количество информации, то система может «захлебнуться», т. е. не сможет переработать всю поступающую информацию. Поэтому при организации системы управления большую роль играет определение необходимого минимума информации.

Не менее важно и рассмотрение времени управления, т. е. промежутков времени от начала сбора информации до выдачи распо-

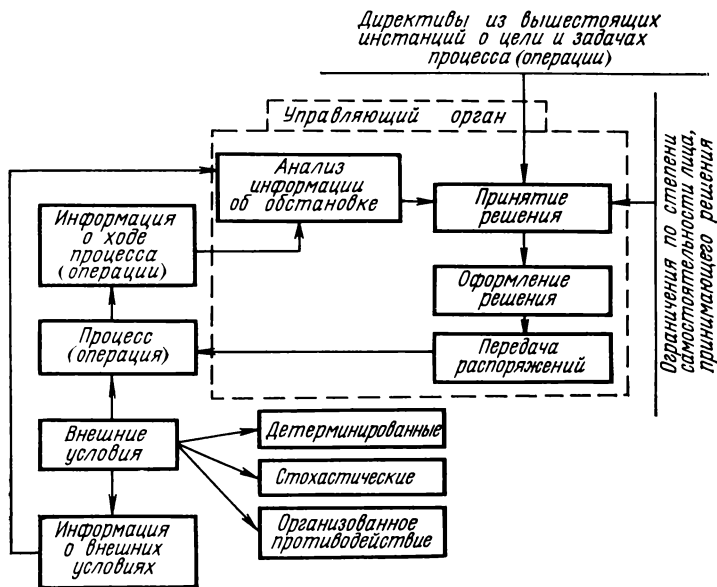


Рис. 11. Общая схема управления процессом (операцией)

ряжений в управляемый процесс (операцию), определяющее устойчивость системы, т. е. степень возможности ее нормального функционирования. Чем меньше времени на управление, тем устойчивее система, поэтому стремление к сокращению времени управления — основная тенденция в развитии систем управления, приведшая в конце концов к появлению автоматизированных систем управления (АСУ).

§ 7. Системный подход к анализу процесса управления

Процесс управления — это система управления в динамике, это процесс функционирования органов и процесс деятельности работников в соответствии с целями, на основе принципов и с использованием методов управления.

В общем виде процесс управления можно охарактеризовать с точки зрения содержания (что именно делается), организации (кем и в каком порядке) и технологии (как делается).

Содержание процесса управления в конечном счете задается той системой, которой управляют.

В содержании процесса управления можно выделить аспекты: технический, производственный, экономический и социальный.

Организационная характеристика процесса управления — это прежде всего характеристика участников процесса.

Этих участников, далее, необходимо ранжировать: ответственные исполнители, соисполнители.

Очень важна при организационной характеристике процесса управления временная последовательность его протекания, определяемая циклом управления.

И, наконец, еще одна характеристика процесса управления — технологическая. Она описывает этот процесс прежде всего как деятельность по принятию и реализации решений. Технологическая характеристика включает также анализ процесса управления как процесса преобразования информации. Она включает характеристику этапов сбора, преобразования и хранения информации.

Характер преобразования информации для технологии процесса управления имеет важное значение. Различают процессы формального и творческого преобразования информации. Творческие процессы — это те, в результате которых возникают новые, более качественные виды информации. Это не означает, что формальные преобразования не нужны. С применением математических моделей формальные преобразования информации позволяют получить, например, оптимальное решение. К тому же, хотя творческие процессы — главные, удельный вес творческих процессов обычно невелик и основной объем в процессе управления сводится к формализованным преобразованиям.

В состав технологической характеристики входит и анализ процесса управления в процедурном плане. В процедурную характеристику технологии процесса управления включаются действия, операции и процедуры, выполняемые работниками управления. В состав процедуры управления входят и проблемы организации деятельности работников управления. Речь идет о круге проблем научной организации работы руководителя. При этом речь идет не о содержании проблем, не о методах их решения по существу, а именно об общей технологии выработки и реализации решений.

Следующим элементом технологической характеристики являются такие важные общие методы принятия различных решений в управлении экономикой, как экономико-математические. Технология предполагает анализ роли этих методов, заключенных в них возможностей, видов, путей их эффективного применения на практике.

Содержательная, организационная, процедурная и технологическая характеристики применяются при анализе любого процесса управления. При его анализе, особенно при анализе деятельности кадров управления, большое значение имеет технологическая сторона процесса управления. Сложные этапы процесса управления поддаются только эвристическому анализу.

Организация и психология процесса управления связаны с использованием правовых средств для регламентации этого процесса.

Для системного анализа процесса управления важно деление этого процесса на макропроцесс — в масштабе всего народного

хозяйства, и на множество составляющих макропроцесс более локальных процессов — в отраслях, на территориях, внутри производственных организаций и учреждений.

§ 8. Своеобразие проблем, требующих управленческого решения

Причиной, вызывающей иногда крайне сложные проблемы, которые требуют управленческих решений, является диалектический характер самого процесса развития экономики, идущей через преодоление внутренних противоречий. Общественная собственность на средства производства, однородная социальная структура всех наций и национальностей, входящих в Советский Союз, исключают классовый антагонизм в нашем обществе.

Противоречия же как внутренние причины и источник всякого развития имеют место в развитом социалистическом обществе, и понимание этих противоречий во многом облегчает анализ ситуаций и проблем, возникающих в наши дни в управлении хозяйственной деятельностью и требующих соответствующих управленческих решений.

Огромное количество вопросов и проблем ставит научно-техническая революция, требующая в наших условиях использования всех преимуществ социалистического строя. Немало достаточно сложных вопросов и проблем, требующих управленческого решения, возникают между отраслевыми и территориальными органами управления.

Таким образом, мы можем констатировать следующие источники основных проблем и ситуаций, требующих управленческих решений.

1. Проблемы и вопросы, возникающие в связи с бурным ростом экономики, требованиями, выдвигаемыми научно-технической революцией.

2. Проблемы и вопросы, возникающие в результате нарушения требования объективных экономических законов развитого социализма.

3. Несоответствие организационных форм уровню развития и требованиям развитой экономической базы социализма.

4. Нарушение государственной и плановой дисциплины в процессе планирования, производства, материально-технического снабжения и поставок.

5. Недостаточная эффективность деятельности самой системы (предприятия, организация и др.).

6. Несоответствие целей, установленных для организации (предприятия), потребностям общества.

7. Непредвиденные изменения как вне, так и внутри системы.

В самом простейшем виде мы можем представить себе возникновение проблем, требующих управленческих решений по двум группам причин: неэффективной деятельности и неправильно установленных целей.

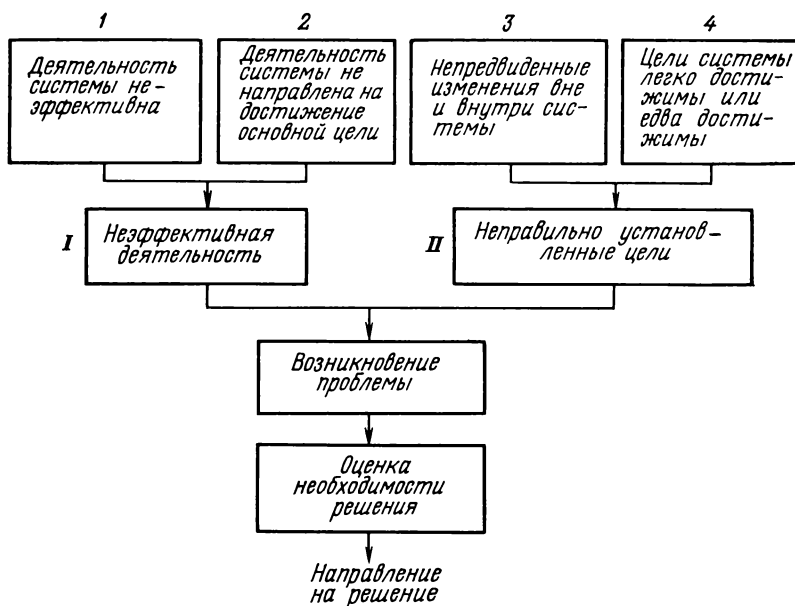


Рис. 12. Возникновение проблемы

Сказанное можно представить также в виде простейшей схемы (рис. 12), показывающей, как возникают проблемы, требующие управленческого решения.

§ 9. Системный подход к обоснованию решений

Важным при организации разработки и осуществления управленческого решения является общая методологическая сторона этого процесса. Поскольку вся социалистическая экономика и общество в целом представляют собой целеустремленную систему, основанную на плановом руководстве происходящими в ней экономическими и социальными процессами, то и каждая отдельная организация и предприятие в этой большой системе должны представлять собой систему.

На основе богатейшего опыта Советского государства в области управления можно сформулировать некоторые общие условия, необходимые для того, чтобы организацию, предприятие, отрасль, систему планирования, материальное обеспечение, сами процессы, происходящие в сфере управления, можно было бы рассматривать как системы (см. § 1). К необходимым условиям следует отнести:

1. Наличие цели, функции, а также измерителей деятельности.
2. Наличие устойчиво связанных между собой определенной структурой подсистем, имеющих свои измерители деятельности и функции, причем такие, что при улучшении показателей каждой

подсистемы происходит одновременно улучшение показателей деятельности системы в целом.

3. Возможность определить тех, кого обслуживает система — потребителей ее продукции, услуг и т. д., причем те, кого обслуживает система, должны определенным образом иметь возможность участвовать в установлении целей и измерителей деятельности системы, а также вносить предложения по их изменению.

4. Возможность определить организацию, которая выполняет функции разработки предложений по изменению показателей деятельности системы, а также ее подсистем таким образом, чтобы оптимизировать цели тех, чьи интересы система призвана обслуживать или удовлетворять.

5. Возможность определить организацию или звено как в самой системе, так и в вышестоящей системе, которая в зависимости от своих ресурсов, прав и ответственности принимает решения по внесению изменений в показатели деятельности подсистем.

6. Наличие вышестоящей системы и иных объективных условий, обеспечивающих как само существование системы, так и то, что выработанные в соответствии с перечисленными выше п.п. 3, 4, 5 решения могут быть практически осуществлены.

Анализ практики хозяйственной деятельности показывает, что существуют по крайней мере три случая, когда вышеперечисленных условий оказывается недостаточно.

1. Случай, когда среда, в которой действует данная система (организация), объективно не соответствует достижению ее цели.

Развитие экономики и культуры всегда происходит в условиях ограниченных ресурсов, причем потребности, как правило, опережают возможности общества. Поэтому иногда правильные, на первый взгляд, решения и предложения не всегда можно осуществить в короткие сроки.

2. Тот, кто должен принимать решение о внесении изменений в существующую систему, будь то структура управления или показатели деятельности, не принимает этого решения. Имеется много случаев, когда решение признается в принципе правильным, но не принимается, а следовательно, и не осуществляется в результате отсутствия четкого разделения прав и обязанностей между соответствующими уровнями управления.

3. Любые изменения в организационной системе связаны с совместной деятельностью людей, трудового коллектива. Если первичный коллектив не принимал участия в разработке решения, он может быть не заинтересован в его проведении, проявить инертность.

Таким образом, приведенные выше дополнительные три обстоятельства имеют весьма существенное значение при системном анализе процессов, происходящих в среде управления, для выработки соответствующего решения.

На рис. 13 приведена логическая схема проверки достаточности условий для определения наличия системы.

Из всего сказанного можно сделать вывод, что для получения оптимальных управленческих решений мы прежде всего должны

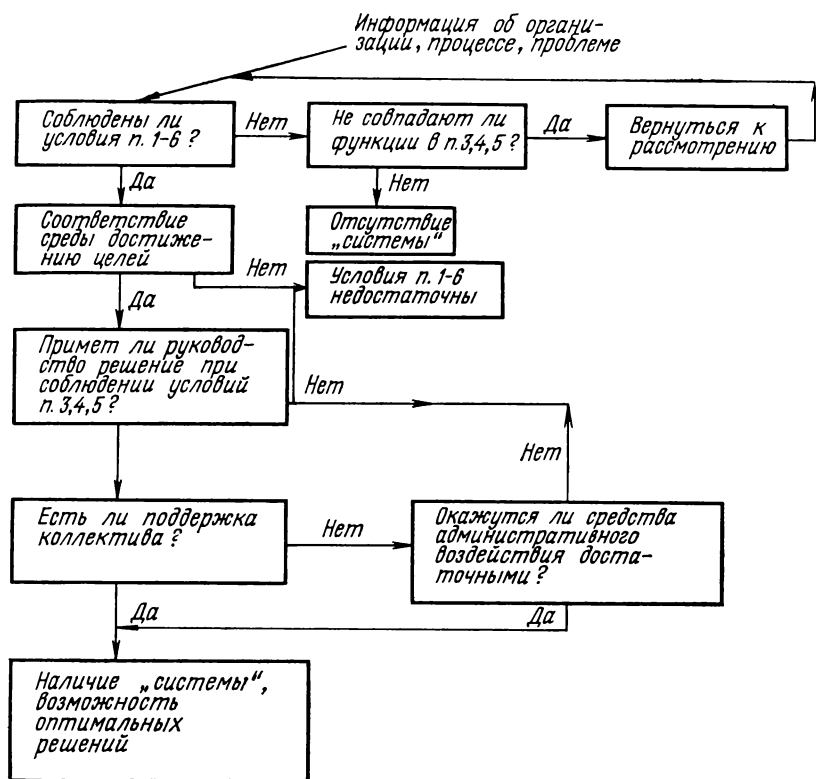


Рис. 13. Проверка достаточности условий для определения системы (системности)

исходить из общеметодологических требований марксистско-ленинской диалектики, требований объективной системы экономических законов развитого социализма, положений и принципов научного управления общественными процессами, разработанными в современном научном коммунизме, а также принципов системного подхода.

Выше было подчеркнуто, что все этапы принятия решений и реализации управленческого решения должны базироваться на методах, выработанных современной наукой. Это относится в известной мере к таким вопросам, как использование вероятностных оценок при выработке хозяйственного решения. В этих случаях, когда имеем дело с решениями, направленными на действия и результаты этих действий в будущем, всегда следует поискать соответствующий метод в одной из областей современной теории вероятностей.

Один из практических выводов этой теории логически подсказывает, что при выработке ответственного решения нужно предусматривать соответствующие действия (и резервы) на случай не-

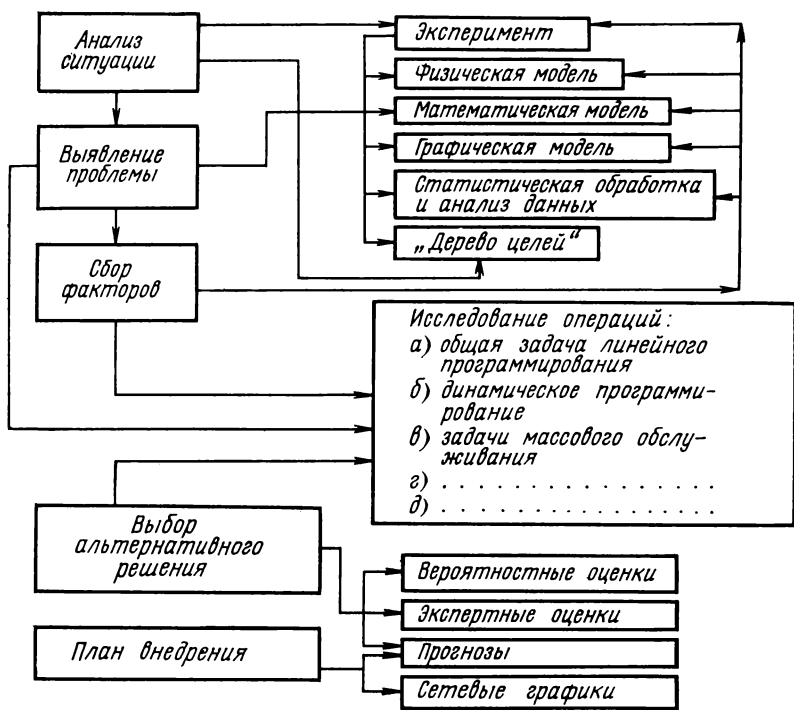


Рис. 14. Связь этапов процесса выработки управленческого решения с отдельными научными методами

предвиденных неудач и срывов в выполнении его на определенном этапе.

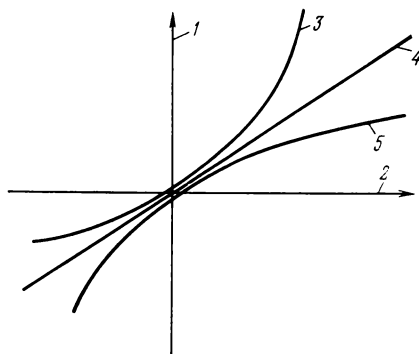
На рис. 14 предусмотрена связь этапов процесса выработки управленческого решения с отдельными методами.

В заключение следует коротко остановиться на теории полезности и применении ее результатов в процессе выработки решений. Дело в том, что при достижении тех или иных экономических результатов «психологические результаты» могут быть разными. Так, например, в обычной денежной игре проигрыш незначительной суммы хотя и описывается некоторым экономическим результатом, но «психологический результат» в этом случае практически равен нулю. Наоборот, проигрыш значительной суммы может вызвать намного больший «психологический результат», чем экономический.

Таким образом, экономический результат (в более общем случае материальный результат Θ) и «психологический результат» (в более общем случае полезность Π) не всегда будут пропорциональны. Соотношения между ними называют функцией полезности $\Pi = f(\Theta)$, которая определяется путем социологических исследо-

Рис. 15. Функция полезности:

1 — психологический эффект (полезность),
2 — экономический (материальный) эффект;
3 — «азартная» функция, 4 — объективная функция полезности, 5 — «осторожная» функция



ваний, опросов экспертов и другими эвристическими методами. Типичные виды этих функций для игровых ситуаций показаны на рис. 15.

Функция полезности должна учитываться при принятии решения. В ней находит отраже-

ние часть тех факторов, которые не включены в модель исследования.

Рассмотрев вопрос об общих принципах принятия решений, более детально остановимся на получении математической основы для их выработки, для чего в последующих главах рассмотрим общую методологию принятия решений [4, 26].

§ 10. Моделирование как метод научного познания

Моделирование в научных исследованиях стало применяться еще в глубокой древности и постепенно захватывало все новые области научных знаний: техническое конструирование, строительство и архитектуру, астрономию, классическую физику, химию и биологию и, наконец, общественные науки. Большие успехи и признание практически во всех отраслях современной науки принес методу моделирования XX в.

Методология моделирования долгое время развивалась независимо в недрах отдельных наук. Отсутствовала единая система понятий, единая терминология. Лишь постепенно стала осознаваться роль моделирования как универсального метода научного познания, как важной гносеологической категории. Теперь же приходится сталкиваться уже с трудностями совершенно другого рода. Понятия «модель» и «моделирование» стали настолько употребительными, что зачастую сложно разобраться, где же кончается «мир» моделей и что в процессах познания не является моделированием.

Термин «модель» широко используется в различных областях человеческой деятельности. Мы будем рассматривать только модели, применяемые в научных исследованиях.

М о д е л ь — это такой материально или мысленно представляемый объект, который в процессе исследования замещает объект-оригинал так, что его непосредственное изучение дает новые знания об объекте-оригинале.

Литература по методологии научных исследований содержит немало существенно различных определений понятия «модель». Приводимое определение ближе к определениям, содержащимся в работах [19, 28].

Под **м о д е л и р о в а н и е м** понимается процесс построения, изучения и применения моделей. Моделирование тесно связано с такими гносеологическими категориями, как абстракция, аналогия, гипотеза и др. Процесс моделирования обязательно включает и построение абстракций, и умозаключения по аналогии, и конструирование научных гипотез. Главная особенность моделирования в том, что это метод опосредственного познания с помощью объектов-заместителей. Модель выступает как своеобразный инструмент для познания, который исследователь ставит между собой и объектом и с помощью которого изучает интересующий его объект.

Необходимость использования метода моделирования определяется тем, что многие объекты (или проблемы, относящиеся к этим объектам) непосредственно исследовать или невозможно (объект недостижим, как, например, ядро Земли и глубины Вселенной), или они еще реально не существуют (будущее состояние экономики, будущие потребности общества и т. п.), или это исследование требует много времени и средств.

Все множество моделей делится на два больших класса: модели *материальные* (предметные) и *идеальные* (мысленные).

В *классе материальных (предметных)* моделей можно выделить две основные группы: модели **ф и з и ч е с к и е** и модели **п р е д м е т н о - м а т е м а т и ч е с к и е**.

Физическое моделирование особенно распространено в технических науках. Предметно-математические модели являются разновидностью математических моделей. Предметно-математическое моделирование основано на том, что характерные черты каких-либо процессов или явлений, принципиально различных по своей физической природе, могут выражаться одинаковыми математическими зависимостями.

Класс идеальных (мысленных) моделей объединяет довольно разнообразные модели, различающиеся прежде всего по степени формализации реальной действительности. В научном познании основным видом идеальных моделей являются **з н а к о в ы е** модели, использующие определенный формализованный язык. В свою очередь, важнейшим видом знаковых моделей являются **л о г и к о - м а т е м а т и ч е с к и е** модели, которые выражаются на языке математики и логики. Логико-математическая модель представляет собой определенную систему математических отношений и логических выражений (функций, уравнений, неравенств, алгоритмов и т. д.), отражающих существенные свойства исследуемого объекта.

Предметно-математическое и логико-математическое моделирования образуют математическое моделирование в широком смысле — метод исследования, основанный на аналогии процессов и явлений, различных по своей природе, но описываемых одинаково

выми математическими зависимостями. В современных научных исследованиях математическое моделирование является, безусловно, важнейшей формой моделирования, а в экономических исследованиях — доминирующей формой.

Роли двух разновидностей математических моделей неодинаковы. Предметно-математические модели служат средствами технической реализации логико-математических моделей и, следовательно, предполагают существование последних.

Известно немало примеров построения наглядных и эффективно действующих предметно-математических моделей для решения разнообразных экономических задач. Еще в XIX в. немецкий экономист-географ В. Лаунхардт предложил механическую модель для решения простейших задач размещения хозяйства; впоследствии эту модель обобщал А. Вебер и в историю экономико-географической науки данная модель вошла как «штандортный треугольник А. Вебера». Для выбора рациональных маршрутов смешанных перевозок «море — суша» предложена оптическая модель, использующая формулу преломления светового луча при переходе из одной прозрачной среды в другую. Для решения задач межотраслевого баланса применяются гидравлические модели, представляющие собой систему из заполняемых жидкостью резервуаров и сообщающихся сосудов. Большое число электрических и электронных моделей построено для изучения проблем регулирования экономического роста, циркуляции денежных потоков, товарного обращения и т. д.

Далее мы будем изучать только логико-математические модели (просто математические модели).

Математические модели экономических процессов и явлений более кратко можно называть экономико-математическими моделями. Для классификации этих моделей используют разные основания.

По целевому назначению экономико-математические модели делятся на теоретико-аналитические, используемые в исследованиях общих свойств и закономерностей экономических процессов, и прикладные, применяемые в решении конкретных экономических задач (модели экономического анализа, прогнозирования, управления и т. д.).

§ 11. Сущность процесса моделирования

Процесс моделирования включает три элемента: 1) субъект (исследователь), 2) объект исследования, 3) модель, опосредствующую отношения познающего субъекта и познаваемого объекта. Сущность процесса моделирования схематически изображена на рис. 16.

Пусть имеется некоторый объект A , который необходимо исследовать. Мы конструируем или находим в реальном мире другой объект B — модель объекта A . Этап построения модели (этап I)

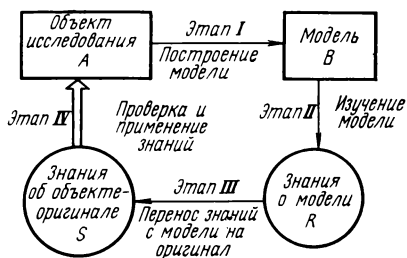


Рис. 16. Схема процесса моделирования

предполагает наличие некоторых знаний об объекте-оригинале. Познавательные возможности модели обуславливаются тем, что модель отображает (воспроизводит, имитирует) какие-либо существенные черты объекта-оригинала. Вопрос о необходимой и достаточной мере сходства оригинала и модели требует конкретного анализа. Очевидно, модель утрачивает свой смысл как в случае тождества с оригиналом (тогда она

перестает быть моделью), так и в случае чрезмерного во всех существенных отношениях отличия от оригинала.

Таким образом, изучение одних сторон моделируемого объекта осуществляется ценой отказа от исследования других сторон. Поэтому любая модель замещает оригинал лишь в строго ограниченном смысле. Из этого следует, что для одного объекта может быть построено несколько «специализированных» моделей, концентрирующих внимание на определенных сторонах исследуемого объекта или же характеризующих объект с разной степенью детализации.

На втором этапе процесса моделирования модель выступает как самостоятельный объект исследования. Одной из форм такого исследования является проведение «модельных» экспериментов, при которых сознательно изменяются условия функционирования модели и систематизируются данные об ее «поведении». Конечным результатом этого этапа является множество (совокупность) знаний о модели R .

На третьем этапе осуществляется перенос знания с модели на оригинал — формирование множества знаний S . Одновременно мы переходим с «языка» модели на «язык» оригинала. Этот процесс знаний проводится по определенным правилам. Знания о модели должны быть скорректированы с учетом тех свойств объекта-оригинала, которые не нашли отражения или были изменены при построении модели. Мы можем с достаточным основанием переносить какой-либо результат с модели на оригинал, если этот результат необходимо связан признаками сходства оригинала и модели. Если же определенный результат модельного исследования связан с отличием модели от оригинала (неадекватностью), то этот результат переносить неправомерно.

Четвертый этап — практическая проверка получаемых с помощью моделей знаний и их использование для создания обобщающей теории объекта, его преобразования или управления им. В итоге мы снова возвращаемся к проблематике реального объекта.

Для понимания сущности моделирования важно не упускать из виду, что моделирование — не единственный источник знаний

об объекте. Процесс моделирования «погружен» в более общий процесс познания. Это обстоятельство учитывается не только на этапе построения модели, но и на завершающей стадии, когда происходит объединение и обобщение результатов исследования, получаемых на основе многообразных средств познания.

§ 12. Этапы экономико-математического моделирования

Процесс моделирования носит циклический характер, и в каждом цикле выделяется несколько этапов. Основные этапы моделирования уже рассматривались (см. рис. 16). Но, поскольку в каждой области знаний применение метода моделирования имеет существенные особенности, целесообразно более внимательно проанализировать последовательность и содержание этапов экономико-математического моделирования.

1. **Постановка экономической проблемы и ее качественный анализ.** Главное здесь — четко сформулировать сущность проблемы, принимаемые допущения (предпосылки) и те вопросы, на которые требуется получить ответы. Этот этап включает выделение важнейших черт и свойств моделируемого объекта и абстрагирование от второстепенных; изучение структуры объекта и основных зависимостей, связывающих его элементы; формулирование гипотез (хотя бы предварительных), объясняющих поведение и развитие объекта.

2. **Построение математической модели.** Это этап формализации экономической проблемы (ситуации), выражения ее в виде конкретных математических зависимостей и отношений (функций, уравнений, неравенств и т. п.). Обычно сначала определяется основная конструкция (тип) математической модели и изучаются возможности ее применения, а затем уточняются детали этой конструкции (конкретный перечень переменных и параметров, форма связей). Таким образом, построение модели подразделяется, в свою очередь, на несколько стадий.

3. **Математический анализ модели.** Целью этого этапа является выяснение общих свойств модели (ее решений). Здесь применяются чисто математические приемы исследования. Наиболее важный момент — доказательство существования решений в сформулированной модели (теорема существования). При аналитическом исследовании модели выясняются такие вопросы, как, например, единственно ли решение, какие переменные (неизвестные) могут входить в решение, каковы будут соотношения между ними, в каких пределах в зависимости от каких исходных условий они изменяются, каковы тенденции их изменения (асимптотические свойства) и т. д. Преимущество аналитического исследования перед эмпирическим (численным) в том, что получаемые выводы сохраняют свою силу при различных конкретных значениях внешних и внутренних параметров модели.

4. Подготовка исходной информации. Это, пожалуй, наиболее трудоемкий этап моделирования, отнюдь не сводящийся к пассивному сбору исходных данных. Моделирование предъявляет жесткие требования к системе информации. В то же время реальные возможности получения информации ограничивают выбор моделей, предназначенных для практического использования. При этом принимается во внимание не только принципиальная возможность подготовки информации (за определенный срок), но и затраты на подготовку соответствующих информационных массивов. Эти затраты не должны превышать эффект от использования дополнительной информации.

В процессе подготовки информации широко используются методы теории вероятностей, теоретической и математической статистики (организация выборочных обследований, оценка достоверности данных, определение вероятных значений параметров и т. п.). При системном экономико-математическом моделировании исходная информация, используемая в одних моделях, является результатом функционирования других моделей.

5. Численное решение. Этот этап включает разработку алгоритмов (вычислительных схем) для численного решения задачи, составление программ на ЭВМ и непосредственное проведение расчетов. Трудности этого этапа обусловлены прежде всего большой размерностью экономических задач, необходимостью обработки значительных массивов информации.

Обычно расчеты по экономико-математической модели носят многовариантный характер. Благодаря высокому быстродействию современных ЭВМ удается проводить многочисленные «модельные» эксперименты, изучая «поведение» модели при различных измерениях некоторых условий. Исследование, проводимое численными методами, может существенно дополнить результаты аналитического исследования, а для многих моделей оно является единственно осуществимым. Класс экономических задач, которые можно решать численными методами, значительно шире, чем класс задач, доступных аналитическому исследованию.

6. Анализ численных результатов и их применение. На этом заключительном этапе цикла встает труднейший вопрос о правильности и полноте результатов моделирования, о степени практической применимости последних.

Пока еще не создано строгой теории проверки правильности экономико-математических моделей. Математические методы проверки могут выявлять некорректные построения моделей (доказывается неразрешимость модели или не подтверждаются принятые статистические гипотезы) и тем самым сужать класс потенциально правильных моделей.

Взаимосвязи этапов. На рис. 17 изображены связи между этапами цикла экономико-математического моделирования. Первые пять этапов более дифференцированно характеризуют процесс экономико-математического моделирования, чем общая схема моделирования, изображенная на рис. 16: этапы 1 и 2 соответ-

ствуют этапу I общей схемы, а этапы 3, 4, 5 — этапу II общей схемы. Наоборот, заключительный этап 6 включает этапы III и IV общей схемы (перенос знаний о модели на объект-оригинал, проверку и применение этих знаний).

Уже на этапе построения модели может выясниться, что постановка задачи противоречива или приводит к слишком сложной математической модели. В соответствии с этим исходная постановка задачи корректируется. Далее математический анализ модели (этап 3) может показать, что наибольшая модификация постановки задачи или ее формализации дает интересный аналитический результат.

Наиболее часто необходимость возврата к предшествующим этапам моделирования возникает при подготовке исходной информации (этап 4). Может обнаружиться, что необходимая информация отсутствует или же затраты на ее подготовку слишком велики. Тогда приходится возвращаться к постановке задачи и ее формализации, изменять их так, чтобы приспособиться к имеющейся информации.

Особо следует подчеркнуть, что выбор формальной модели реальной действительности и сложность получения алгоритма на основе модели решения проблемы существенно зависят от структурированности проблемы.

Структура любой проблемы определяется пятью основными логическими элементами: 1) целью или рядом целей, достижение которых будет означать, что проблема решена; 2) альтернативными средствами, т. е. курсами действия, с помощью которых может быть достигнута цель; 3) затратами ресурсов, требуемых для каждого курса действий; 4) моделью или моделями, в которых с помощью некоторого формального языка (в том числе математики, формальной логики, обычно словесного, графического или машинного описания и т. п.) отображаются связи между целями, альтернативами и затратами; 5) критерием, с помощью которого сопоставляются в каждом конкретном случае цели и затраты и отыскивается наиболее предпочтительное решение. Степень структуризации проблемы определяется тем, насколько хорошо выделены и осознаны указанные пять логических элементов проблемы. Именно от этого зависит возможность применения того или иного метода для ее решения. С этой точки зрения проблемы могут быть разделены на четыре типа: а) стандартные; б) хорошо структурированные; в) слабоструктурированные; г) неструктурированные. В соответствии с этим могут быть использованы четыре класса методов решения проблем: 1) стан-

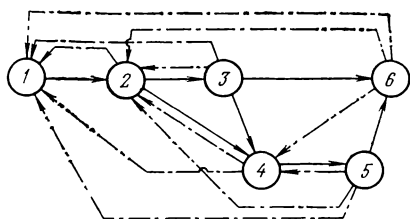


Рис. 17. Связи этапов экономико-математического моделирования (сплошной линией показаны последовательные связи этапов, прерывистой линией с точками — возвратные связи этапов (корректировка))

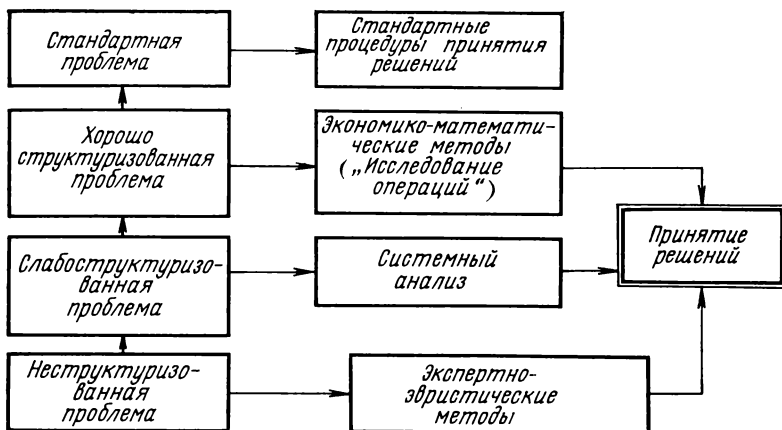


Рис. 18. Классификация спектра проблем и методов их решения

дартные процедуры и правила расчета решений; 2) экономико-математические методы поиска оптимальных решений; 3) системный анализ для построения рациональных хозяйственных альтернатив; 4) экспертно-эвристические методы принятия хозяйственных решений (рис. 18).

К стандартным проблемам можно отнести, например, расчет потребности в оборудовании, в материалах, в рабочей силе исходя из заданной производственной программы.

К хорошо структурированным проблемам можно отнести вопросы развития и реконструкции предприятий, выбор вариантов наилучшей загрузки производственных мощностей и поиск оптимальных режимов технологии производства.

К слабоструктурированным проблемам относятся разработки долгосрочных курсов действий: создание новых производственных комплексов, совершенствование организации управления предприятиями и многие другие.

Неструктуризованные проблемы отличаются значительной неопределенностью и неформализуемостью как самих целей деятельности, так и возможностью курсов действий. Сюда можно отнести формирование долгосрочных планов научно-исследовательских работ и опытно-конструкторских разработок, построение планов социального развития коллектива и т. п.

В современной литературе решения стандартных и хорошо структуризованных проблем часто называют программируемыми, а слабоструктуризованных и неструктуризованных проблем — непрограммируемыми.

Формальные постановки всевозможных модификаций хорошо структуризованных проблем вместе со специальными методами нахождения оптимальных вариантов объединяются в класс задач

математического программирования. Именно об этих задачах (моделях и методах решения) пойдет речь в дальнейшем.

Важно подчеркнуть, что отнесение той или иной проблемы к одному из названных четырех классов не носит постоянного характера. В процессе все более глубокого изучения, осмысливания и анализа проблемы она из неструктуризованной может превратиться в слабоструктуризованную, затем в хорошо структуризованную, а в ряде случаев и в стандартную. Соответственно одна и та же проблема может постепенно описываться все более точными методами, т. е. может осуществляться как бы постепенная формализация проблемы. Поэтому правильнее говорить не о классификации проблем, а о подходах к их описанию. При этом, пользуясь терминологией В. В. Налимова [28], удобно выделить три подхода к отображению объектов при системных исследованиях.

1. Представить анализируемый объект, если это удастся, в виде хорошо организованной системы, т. е. определить все взаимосвязи, зависимости между компонентами, их свойствами и целями, которые учитываются при решении задач. В этом случае задачи выбора целей и средств не разделяются. Цель представляется в виде выражения (как правило, аналитической зависимости), связывающего ее со средствами, и решение задачи заключается в применении достаточно хорошо разработанного математического аппарата (вариационного исчисления, математического программирования), который является основой теории оптимизации и исследования операций, применяющихся при решении хорошо структуризованных проблем (см. рис. 18).

2. Представить объект в виде плохо организованной, или диффузной, системы и применить для его исследования статистические методы, которые позволяют, не устанавливая всех связей между компонентами и целями системы, на основе выборочного исследования получать статистические закономерности и распространять их с какой-то вероятностью на поведение системы в целом, иными словами, пытаются заменить дефицит информации, необходимой для принятия решений по совершенствованию рассматриваемого объекта или деятельности, исследованием не всего объекта, а выборки компонентов или связей и использовать полученные при выборочном исследовании закономерности для совершенствования объекта (деятельности). Цель в этом случае формулируется содержательно.

3. Представить объект в виде самоорганизующейся, или развивающейся, системы. Это практически основной подход, который может применяться при анализе процессов в организованных объектах на уровне отрасли, объединения, крупного предприятия. Идею этого метода можно сформулировать следующим образом: с помощью принятой знаковой системы фиксировать известные на данный момент компоненты и связи между ними и, комбинируя их с применением установленных правил, получать новые, неизвестные ранее взаимоотношения, зависимости, которые могут быть положены в основу принимаемых решений (см. § 4).

Такой подход позволяет накапливать информацию об объекте, фиксируя новые компоненты в связи (правила их взаимодействия) и применяя их. При этом информация может поступать от различных специалистов и накапливаться во времени по мере ее выявления. «Механизм» накопления информации о компонентах, целях и их взаимосвязях должен постоянно сохраняться в системе и «включаться» при накоплении изменений в ней.

При применении этого подхода задачи определения целей и выбора средств, как правило, разделяются: цели обычно представляются в виде иерархических структур, фиксируемых в виде многоступенчатых списков, матриц, графов типа «дерева» («дерева целей»). Примером такого представления является структура плана, включающая взаимосвязанные задачи, решение которых с точки зрения руководства организации обеспечивает выполнение явно или неявно сформулированной глобальной (общей) цели организации на данном этапе ее развития.

Исследования процессов планирования показывают, что качество плана, скорость его разработки и корректировки в соответствии с реально изменяющимися условиями развития отрасли или организации существенно повышаются при применении специальных приемов и методов формирования и анализа структуры плана.

§ 13. Информационное обеспечение процесса управления

В настоящее время существуют различные подходы к пониманию информации. Большинство из них построено либо на математизации понятия информации, выделении ее количественных характеристик, либо на возведении категории «информация» в ранг общеполитической категории, отражающей, как и категория материи, всеобщее свойство мира.

Различное понимание информации во многом объясняется углубляющейся специализацией наук, каждая из которых имеет свой объект и предмет исследования, свои цели развития. Это накладывает заметный отпечаток на трактовку тех или иных категорий, используемых в целом ряде наук.

Понятие информации в социально-экономических дисциплинах связывают с отражением в сознании человека окружающего его мира. Информация характеризует одну из важнейших сторон деятельности человека — познание мира. На это обращал внимание и основатель кибернетики Н. Винер. Информация в таком ее понимании выступает необходимым условием трудовой деятельности человека, его развития. В социально-экономической системе она является средством коммуникации между людьми. Этим и определяется ее роль в управлении, осуществление которого было бы невозможно без связей, коммуникаций.

Играя важную роль в деятельности человека, информация имеет большое значение в управлении производством. С точки зрения исследования процесса управления информацией можно определить

как совокупность сведений о состоянии управляемой, управляющей систем и внешней среды. В таком понимании информация выступает как основа процесса управления. Все стадии и этапы процесса управления связаны с переработкой информации. Без информации невозможно определить цели управления, оценить ситуацию, сформулировать проблему, принять решение и проконтролировать его выполнение.

Организационно-практическая деятельность руководителя и аппарата управления также во многом носит информационный характер, ибо она включает получение информации для принятия решений и передачу информации о принятых решениях.

Иногда роль информации в процессе управления абсолютизируют, а сам процесс управления рассматривают только как процесс движения информации или же как процесс ее преобразования из исходной в командную, результирующую.

Процесс управления нельзя полностью сводить к преобразованию информации. Это содержание гораздо богаче. Воздействие на человека — это не столько акт передачи ему информации, сколько самый сложный социально-психологический процесс взаимодействия людей, включающий такие факторы, которые не относятся к информации, — интересы, долг, дисциплина, воля, авторитет, мораль и др. В осуществлении этих актов информация выступает как средство воздействия. Иногда создается видимость, что передача информации — это и есть воздействие. Но это лишь внешняя сторона процесса управления. Одна и та же информация по-разному действует на различных людей. И это потому, что не только информация оказывает действие в процессе управления, а человек воздействует на человека, по-разному используя при этом информацию как основу и средство воздействия в зависимости от того, кто или что является объектом воздействия.

Роль информации неодинакова для различных этапов или стадий процесса управления. На одних стадиях управления важен объем информации, ее виды, возможности получения новой, дополнительной информации, виды последней, возможности ее обработки.

Так, при определении цели управления важен объем информации, ее новизна, ценность, полнота и т. д. Чем больше сведений о современных научных достижениях использовано при разработке цели управления, тем объективней формулируется цель, тем значительней она для процесса управления.

При оценке ситуации наибольшее значение имеют виды информации, определяющие возможности комплексного и системного подхода к этой оценке, а в дальнейшем — к разработке управленческого решения.

На стадии разработки управленческого решения и аналитической деятельности большую роль играют возможности обработки отобранной для полученного решения информации, которые зависят от формы ее представления.

При организационно-практической деятельности по реализации решения наибольшую роль играют процессы передачи информации,

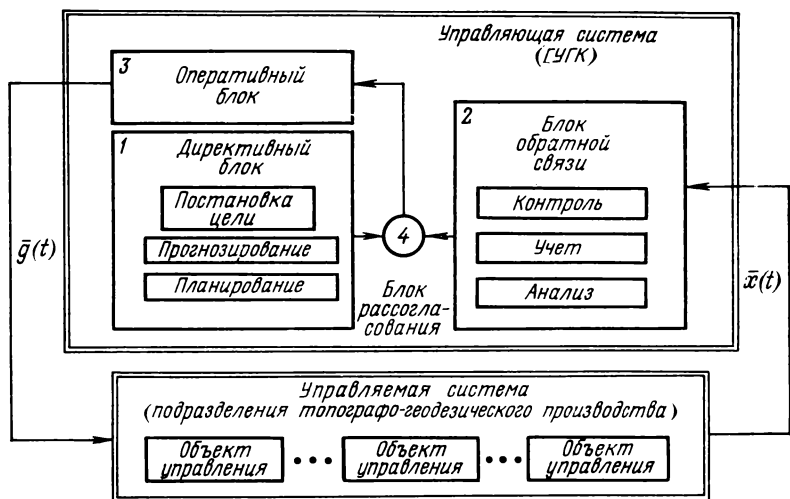


Рис. 19. Основные функции управления топографо-геодезическим производством как системы

скорость передачи, своевременность, полнота ее поступления и пр.

Роль информации для процесса управления двояка. С одной стороны, постоянное увеличение объема информации по мере развития производства создает трудности в управлении. Управляющей системе приходится перерабатывать все больше и больше информации. Установлено, что зависимость объема информации от объема производства является нелинейной, причем при экстенсивном развитии производства более высокими темпами растет объем информации, ибо в этом случае увеличивается число возможных производственных и хозяйственных связей.

Но, с другой стороны, по мере увеличения объема информации, особенно научно-технической, увеличиваются возможности совершенствования управления. Познание новых законов и закономерностей тем и ценно, что оно открывает новые возможности управления процессами развития, обуздания стихийных сил природы.

Таким образом, рост объема информации, используемой в управлении, нельзя оценивать односторонне. Он содержит в себе диалектическое единство трудностей управления и возможностей их успешного преодоления.

На рис. 19 показаны основные функции управления топографо-геодезическим производством как системы. Эти функции обобщены в трех блоках: директивном, обратной связи и оперативном. Выработка управляющих решений сосредоточена в оперативном и директивном блоках на основе данных блока рассогласования. Управляемая система характеризуется вектором состояния $\bar{x}(t)$. Вектор $\bar{g}(t)$ определяет динамическую характеристику объекта управления [15].

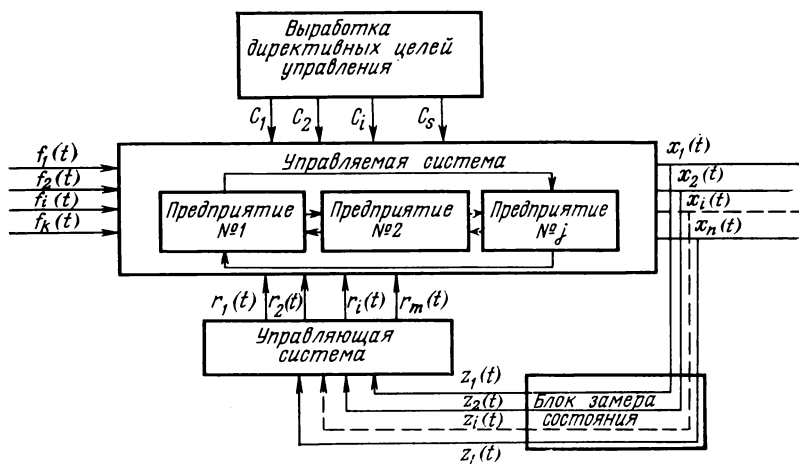


Рис. 20. Информационные связи управляемой системы

Информационные связи системы управления можно изобразить в виде, представленном на рис. 20.

Назначение системы управления заключается в том, чтобы система всегда достигала внешних целей управления ($C_1, C_2, \dots, C_s$). На рис. 20 приняты следующие обозначения: $x_1(t), x_2(t), \dots, x_n(t)$ — переменные состояния управляемой системы. Для рассматриваемой системы «отрасль ∇ предприятие» — это выходные показатели, характеризующие работу системы управления. В нашем случае это основные технико-экономические показатели, отражающие эффективность топографо-геодезического производства; $f_1(t), f_2(t), \dots, f_k(t)$ — внешние возмущения, создаваемые окружающей средой. Такими возмущениями могут быть влияние погодных условий, организационные неполадки и пр.;

$r_1(t), r_2(t), \dots, r_m(t)$ — управляющие переменные, такие как распределение денежных, трудовых и других ресурсов;

$z_1(t), z_2(t), \dots, z_l(t)$ — наблюдаемые переменные о состоянии системы, т. е. отчетные показатели. В частном случае, при безынерционном замере показатели $z_i(t) = x_i(t)$.

Указанные переменные можно рассматривать как компоненты многовекторных функций:

$$\bar{X}(t) = \begin{vmatrix} x_1(t) \\ x_2(t) \\ \dots \\ \dots \\ \dots \\ x_n(t) \end{vmatrix} ; \quad \bar{R}(t) = \begin{vmatrix} r_1(t) \\ r_2(t) \\ \dots \\ \dots \\ \dots \\ r_m(t) \end{vmatrix} ;$$

$$\bar{F}(t) = \begin{vmatrix} f_1(t) \\ f_2(t) \\ \dots \\ \dots \\ f_k(t) \end{vmatrix}; \quad \bar{Z}(t) = \begin{vmatrix} z_1(t) \\ z_2(t) \\ \dots \\ \dots \\ z_l(t) \end{vmatrix}.$$

Назовем векторы $\bar{X}(t)$, $\bar{R}(t)$, $\bar{F}(t)$ и $\bar{Z}(t)$ соответственно векторами состояния, управления, возмущения, наблюдения. Задачей анализа технико-экономических показателей $x_1(t)$, $x_2(t)$, $\dots$, $\dots$, $x_n(t)$ является определение параметров векторов $\bar{F}(t)$, $\bar{R}(t)$, при которых вектор $\bar{X}(t)$ принимает оптимальные значения.

В любой момент времени t состояние управляемой системы является функцией начального состояния $x_i(t_0)$ и векторов $r_i(t_0)$, $f_i(t_0)$. За начальное состояние могут быть приняты показатели предшествующего года в начале планового периода при текущем планировании, при перспективном планировании и сравнительном анализе — показатели первого года пятилетки. Таким образом, состояние управляемой системы можно выразить функцией

$$\bar{X}(t) = \bar{X} \{ \bar{X}(t_0); \bar{R}(t_0, t); \bar{F}(t_0, t) \}. \quad (\text{II.1})$$

Уравнение (II.1) эквивалентно системе n зависимостей вида

$$\bar{X}_i(t) = \bar{X}_i \{ \bar{X}(t_0); \bar{R}(t_0, t); \bar{F}(t_0, t) \}, \quad (\text{II.2})$$

где $i = 1, 2, \dots, n$; n — число технико-экономических показателей.

Такое расчленение одного уравнения на несколько равнозначно расчленению управляемой системы (см. рис. 20) на взаимосвязанные n элементов (подсистем), т. е. возможно применение известного в теории систем принципа декомпозиции.

§ 14. Информационная модель оперативного управления аэрогеодезическими предприятиями (АГП)

Как известно, процесс оперативного управления неразрывно связан с передачей, приемом, преобразованием и обработкой больших информационных массивов, описывающих свойства отдельных управляемых объектов. Каждый объект можно характеризовать соответствующим уровнем иерархии в системе управления: предприятие, экспедиция (цех, склад), партия (бригада); а также своим шифром: экспедиции, цеха, партии, бригады. Информационные массивы, описывающие свойства каждого такого объекта управления, кроме того можно группировать по своим информационным совокупностям: первичным данным (A_i), нормативно-справочной информации (B_i), результирующим промежуточным данным (C_i),

месячным (квартальным) отчетно-плановым данным (D_i), результатам оперативного контроля и анализа (E_i), годовым отчетно-плановым и статистическим данным (G_i), где i — шифры массивов. При решении любой из задач управления приходится учитывать связи как между различными объектами, так и отдельными информационными совокупностями. Характер и смысл таких связей могут быть самыми разнообразными, однако при всем их многообразии оказывается удобным представлять их в виде ориентированного графа (рис. 21), составленного в процессе создания информационной модели оперативного управления предприятием [23].

В ориентированном графе (см. рис. 21) множество вершин соответствует только одному объекту управления, информационный процесс которого охватывает взаимосвязи всех информационных совокупностей. В этом случае полное описание этих объектов и связей между ними можно формализовать в виде некоторого числа дискретных функций, определенных на графе, каждое значение которых соответствует одной ветви графа. Под каждой такой функцией следует понимать отдельный показатель или отдельную строку массива, или массив в целом. Тогда решение задач оперативного управления сводится к вычислению значений этих функций. Например,

$$F(G_1) = F[f_1(D_1), f_2(D_2), f_3(D_3)];$$

где

$$\left. \begin{aligned} f_1(D_1) &= f[\varphi(C_1), B_1]; \\ f_2(D_2) &= f[\varphi(C_1), \varphi(C_2)] \text{ и т. д.} \\ \varphi(C_1) &= \varphi[B_1, B_2, A_1, A_2]; \\ \varphi(C_2) &= \varphi[B_2, B_3, A_1, A_2] \text{ и т. д.} \end{aligned} \right\} \quad (II.3)$$

В формулах (II.3) F , f , φ — функции, определенные на графе. Получение любой из функций (II.3) требует выполнения полного цикла обработки данных: поиска, идентификаций, математической обработки, записи и контроля каждого показателя. Число полных циклов обработки данных (S) для функции $F(G_1)$, под которой подразумевается массив данных с числом показателей в нем N , можно подсчитать по формуле

$$S = N \left[\sum_{i=1}^n E_i \cdot \sum_{i=1}^m D_i \cdot \sum_{i=1}^l C_i \cdot \sum_{i=1}^k (A_i B_i) \right], \quad (II.4)$$

где m , n , l — число связей между вершинами графа. Если пред-

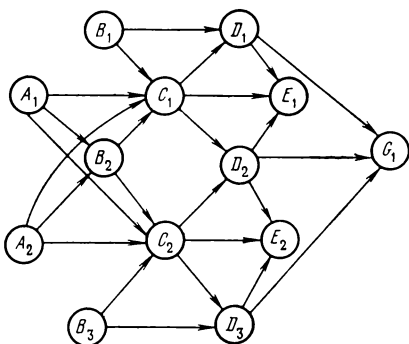


Рис. 21. Фрагмент дерева-графа информационной модели

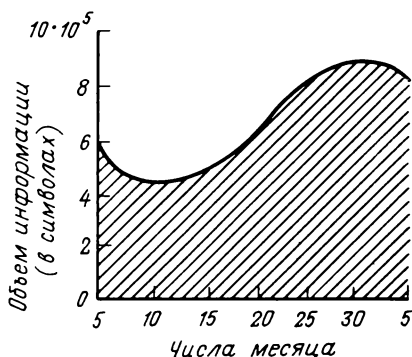


Рис. 22. График интенсивности потоков первичной информации в течение месяца в условиях автоматизированной обработки

положить, что в среднем значения этих величин будут: $N = 10\,000$, $n = 4$, $m = 2$, $l = 4$, то $S = 320\,000$ циклов. Из этого следует, что при большом числе вершин (показателей, строк, массивов) и функциональных зависимостей между ними, обеспечение реальности выполнения расчетов и доступа к ранее полученной информации становится очень сложной проблемой, решение которой целиком и полностью будет зависеть от рациональной организации информационных процессов.

Вся информация, принадлежащая к одной из совокупностей

$A_i, B_i, C_i, D_i, E_i, G_i$, группируется в один четырехмерный массив $M = \|S_{ij}^{lk}\|$, где i, j, l, k — порядковые номера строки, столбца, объекта управления и иерархического уровня управления соответственно.

В зависимости от типа решаемых задач структура каждого информационного массива может меняться путем сокращения его параметров l и k , при $i \geq 1$, $j \geq 2 = \text{const}$ для всех строк одного массива.

Основное достоинство описанной информационной модели — простота поиска нужной информации, так как структура каждого массива по существу идентична соответствующему документу, а каждый показатель характеризуется четырьмя координатами: i, j, l, k .

В «директиве-запросе» указываются или их точные значения, или сортировочные признаки требуемого показателя в такой же последовательности.

Ниже приведена интенсивность потоков оперативной информации в условиях автоматизированной обработки (рис. 22).

На основе описанных принципов решения задач оперативного управления топографо-геодезическим производством в предприятии № 2 решены следующие задачи: 1) расчет потребности в материальных ресурсах и составление сводного проекта плана материально-технического снабжения предприятия, 2) учет движения материально-технических ценностей и контроль за сохранением минимально необходимых запасов, 3) учет кадров в предприятии и анализ причин их текучести [23].

Глава III

ВВЕДЕНИЕ В ТЕОРИЮ ПРИНЯТИЯ РЕШЕНИЙ

§ 15. Развитие процесса принятия и реализации решения

Процесс принятия решения не нов — им интересовались еще жрецы и мудрецы глубокой древности, светила античных времен, им интересуются и современные ученые. За всю историю своего существования люди пользовались различными способами принятия решения. Поэтому неудивительно, что с древних времен до настоящего времени проблемы решения входят в число актуальных проблем науки. При этом большое внимание обращается на решение задач или проблем, весьма неопределенных, т. е. которые носят творческий характер.

Информационные источники, дошедшие до нас с древнейших времен, свидетельствуют, что уже тогда людей волновал вопрос, каким образом происходит принятие решения, особенно творческого, каков рациональный путь к выбору правильного поступка при новых, ранее не встречавшихся обстоятельствах, каким образом люди создают новое и полезное. Например, известно пособие по принятию решения, написанное около четырех тысяч лет назад в Китае И. Чингом. Это пособие явилось оригинальным руководством в творческой работе: оно, по мнению некоторых зарубежных специалистов (Ф. Д. Баррет и др.), не потеряло своего значения до наших дней и может быть рекомендовано современным хозяйственным руководителям и политическим деятелям (речь идет об управлении капиталистическим производством).

Несомненно, в те далекие времена научные данные о природе творческой деятельности, творческих решений были настолько скудны, что они не могли дать людям правильного объяснения этим процессам. Не находя правильного ответа на вопросы, люди объясняли их мистически. Так, древние греки и другие народы античного мира объясняли как силы природы, так и творческие способности их божественным происхождением. При таком объяснении основным орудием управления процессом принятия решения в неопределенных обстоятельствах или процессами творчества были молитвы музам — богиням творчества.

Однако и в те далекие времена люди искали более рациональные пути улучшения процесса принятия решения, особенно для таких случаев, когда было неизвестно, как изменятся обстоятельства. В ту пору предсказаниями будущего славился греческий город Дельфы со своим оракулом.

Вкратце суть этих предсказаний такова. Сначала с обстоятельствами дела досконально знакомились дельфийские жрецы. Затем они наблюдали за прорицательницей (Пифией), одурманенной выходящими из земли газами. Пифия выкрикивала отдельные слова и бессвязные фразы, вызывавшие у жрецов ассоциации, на основе которых они предсказывали будущее. Предсказание передавалось заинтересованному лицу или обнародовалось только после тщательного его обсуждения на совете*.

Несмотря на многовековую историю изучения процесса решения и его принятия, термин «принятие решения» появился в научной литературе в 30-е годы нашего столетия — в творческих работах по управлению производством в капиталистических странах для характеристики процессов децентрализации управления.

Современный этап развития науки управления характеризуется интенсивным поиском новых идей, подходов, методов и средств, способных повысить эффективность управления сложными социально-экономическими системами в условиях все возрастающего динамизма процессов, усложнения связей в системах и все более жесткого ограничения ресурсов. Одним из перспективных подходов является рассмотрение проблем управления с позиций принятия решений. Этот подход, связанный с использованием таких категорий, как «решение», «процесс принятия решений», «система принятия решений», в литературе называют концепцией принятия решений, теорией принятия решений, школой принятия решений [26, 16].

Актуальность и практическая ценность выделения и изучения проблем принятия решений в процессе управления определяются следующими причинами.

Во-первых, принятие решений занимает центральное место в процессе управления. Как известно, принятие решений наряду с прогнозированием, планированием, оценкой обстановки, исполнением решений, контролем и учетом является функцией управления. Центральная, важнейшая роль принятия решений определяется тем, что все другие функции управления направлены на формирование или реализацию решений. Кроме того, любую функцию управления технологически можно представить в виде последовательности решений. Например, при прогнозировании и планировании принимаются решения, связанные с выбором методов и средств, организацией работ, оценкой достоверности информации, выбором наиболее достоверного варианта прогноза и наилучшего варианта плана. Аналогичную цепочку решений можно построить и при рассмотрении других функций управления. Таким образом, функция принятия решений является с методологической и технологической точек зрения более общей, чем другие функции управления, поэтому в литературе иногда управление рассматривается как процесс принятия решений.

Во-вторых, принятие решений — это личная функция руководителя. Для руководителя любого ранга принятие решений яв-

* Бешелев С. Д., Гурвич Ф. Г. Экспертные оценки. М., Наука, 1973.

ляется основной задачей, которую он обязан решать в процессе управления. Поэтому знание методов, технологии и средств решения этой задачи является необходимым элементом квалификации руководителя.

В-третьих, современной моделью функционирования организационных систем является система принятия решений. Эта модель согласует положительные стороны двух предшествующих теоретических моделей, имеющих структурно-описательный характер: механической модели организации как «полностью рациональной» системы; естественной модели организации как «живой» социальной системы.

Система принятия решений является третьей моделью организационных систем, в которой как первичный элемент рассматривается «решение». Основное развитие эта модель получила на рубеже 60-х годов. В данной модели рассматриваются рациональные принципы механической модели с учетом социальной и психологической специфики естественной модели. В решении объединяются объективные факторы информационного анализа проблем, проводимого на основе логического мышления, математических методов и ЭВМ, и субъективные психологические факторы лица, принимающего решение. Поэтому система принятия решений позволяет осуществить системный подход к исследованию сложных организационных систем, включающий социально-технологическую форму реализации процессов управления.

В-четвертых, подход, ориентированный на принятие решений, создает прочную базу для дальнейшего совершенствования автоматизированных систем информационного обеспечения и управления. Эти системы должны развиваться от автоматизации трудоемких рутинных учетно-расчетных задач до логико-аналитических задач формирования и обоснования вариантов решений. Рассмотрение организационной системы принятия решений, выделение в ней центров принятия решений, формирование информационных потоков, соответствующей структуры и их динамической перестройки в зависимости от решаемых проблем являются перспективным направлением совершенствования управления сложными социально-экономическими системами.

Изложенное показывает, что изучение проблем управления с единой методологической позиции принятия решений позволяет практически осуществить комплексный, системный подход к анализу функционирования организационных систем, технологии их управления, применению современных методов и технических средств.

В настоящее время существует достаточно большое число современных научных дисциплин, посвященных проблеме принятия решений. К ним можно отнести математическое программирование, теорию игр, теорию статистических решений, теорию оптимального автоматического управления. К ним следует причислить ряд новейших прикладных дисциплин, даже названия которых были неизвестны лет сорок тому назад: исследование операций, экономическую кибернетику и др. (рис. 23):

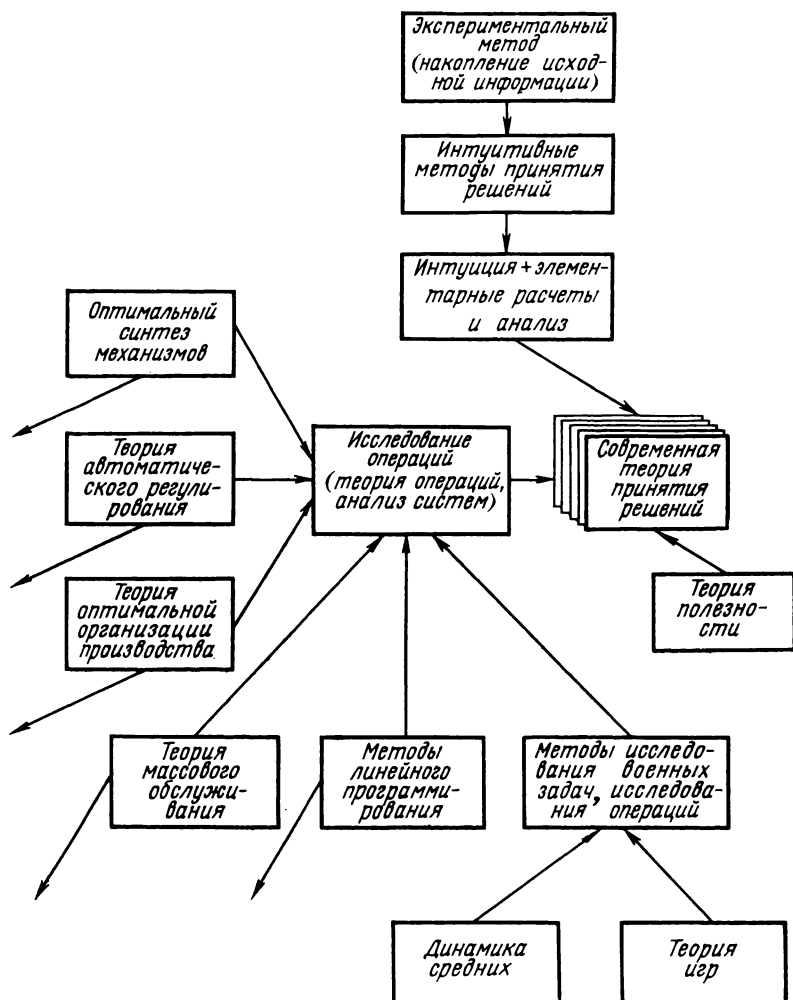


Рис. 23. Схема истории развития исследования операций и принятия решений

Все эти дисциплины занимаются рассмотрением одной и той же основной проблемы — научного анализа ряда возможных способов действия с целью нахождения такого из них, который в данных условиях был бы наилучшим. Иными словами, они занимаются рассмотрением проблем принятия оптимальных решений, но применительно к объектам управления различной природы и в разных условиях их существования. В этом смысле их можно считать составными частями единой научной дисциплины, для обозначения

которой в настоящее время все чаще применяется термин «теория принятия решений» (ТПР).

Ситуацию, в которой происходит принятие решений, в общем случае характеризуют следующие основные черты.

1. *Наличие цели (целей).* Необходимость принятия решения диктуется наличием некоторой цели, которую нужно достичь, например выполнить плановое задание, выбрать тип прибора, назначить план перевозок и т. д. Если же цель не поставлена, то не возникает и необходимость принимать какое-либо решение.

2. *Наличие альтернативных линий поведения.* Решения принимаются в условиях, когда существует более одного способа достижения цели, или иначе несколько альтернатив достижения цели. С различными альтернативами могут быть связаны различные затраты и разные вероятности достижения цели. Эти затраты и вероятности не всегда могут быть определены. Поэтому часто принятие решений сопряжено с неясностью и неопределенностью. Если же существует лишь одна линия поведения, то выбора нет и, следовательно, решение принимать не требуется: оно очевидно.

3. *Наличие ограничивающих факторов.* Решения обычно принимаются в условиях действия большого числа факторов, ограничивающих возможность выбора способов действий. Эти факторы иначе называются *дисциплинирующими условиями*. Ограничивающие факторы, подлежащие рассмотрению, можно разбить на три основные группы: экономические, механические и специальные. Под *экономическими* факторами понимают факторы, связанные с ресурсами: время, денежные средства, трудовые ресурсы, производственные возможности и т. п.

Техническими факторам обычно относят факторы, которые непосредственно связаны с инженерным анализом и выработкой требований к техническим характеристикам объектов: габаритам, массе, прочности, надежности и т. п.

Наконец, *социальные* факторы, в том числе и чисто человеческие, выражают требования не только политической или социальной целесообразности осуществления той или иной альтернативы, но и этики.

Все перечисленные факторы накладывают ограничения на возможности достижения поставленной цели. Очевидно, что отсутствие ограничений существенно упрощает задачу принятия решения.

Таким образом, задача принятия решения возникает в том и только в том случае, когда существует цель, которую нужно достичь, когда возможны различные способы ее достижения и существуют факторы, органичивающие возможности достижения цели.

Выяснение всех трех указанных элементов задачи принятия решений должно обязательно предшествовать ее непосредственному решению.

§ 16. Основные понятия теории принятия решений

Задача принятия решений направлена на определение наилучшего (оптимального) или приемлемого способа действий для достижения одной или нескольких целей.

Под целью понимается в широком смысле идеальное представление желаемого состояния или результата деятельности. Если фактическое состояние не соответствует желаемому состоянию, то имеет место *пр о б л е м н а я с и т у а ц и я*, или проблема, выработка плана устранения которой и составляет сущность задачи принятия решений.

Конечным результатом задачи принятия решений является *р е ш е н и е*. Решение есть предписание к действию. С точки зрения содержания решением может быть курс действия, способ действия, план работы, вариант проекта и т. д.

1. Одним из основных понятий теории принятия решений является понятие *о п е р а ц и и*. Под словом операция следует понимать организованную деятельность в любой области жизни, объединенную единым замыслом, направленную на достижение определенной цели, имеющую характер повторяемости, т. е. многократности. В данном определении подчеркиваются две особенности операции: ее целевая направленность и повторяемость. Именно отсюда возникает возможность проводить исследования, касающиеся количественных сторон операции, общими научными путями с использованием методов теории вероятностей, статистики, математики и данных различных наук — физики, химии, биологии, экономики и др.

Примеры различных операций: а) производственная деятельность отрасли, выпускающей некоторую народнохозяйственную продукцию; б) отражение воздушного налета средствами системы противовоздушной обороны; в) запуск группы искусственных спутников Земли для создания космической системы связи; 2) совокупность мероприятий, направленных на повышение надежности некоторого технического устройства и т. д.

Изучение операций может проводиться как путем исследования оригинала самой операции, так и путем исследования *м о д е л и* операции. Однако основным методом исследования операций, особенно крупного масштаба, является метод моделирования (см. § 10). Дело в том, что постановка специальных экспериментов с целью выбора оптимальных решений, особенно в операциях крупного масштаба, настолько сложна и дорога, что она просто нерациональна. Значительно проще и экономичнее изучить модель процесса и в процессе изучения выбрать лучшее решение.

Итак, основной метод изучения операций крупного масштаба — изучение моделей операции, главным образом моделей математических.

2. Второе важное понятие теории принятия решений — *о п е р и р у ю щ а я с т о р о н а*. Совокупность как лиц, которые стремятся в данной операции к достижению некоторой цели, так и

технических устройств с помощью которых она достигается, называется оперирующей стороной.

В операции могут участвовать одна или несколько оперирующих сторон, преследующих различные, несовпадающие цели. Несовпадение целей оперирующих сторон создает конфликтную ситуацию. Подобные операции называются многосторонними или конфликтными. Так, во втором, приведенном выше примере операции исход операции зависит от деятельности двух сторон, преследующих противоположные цели: нападающей стороны, совершающей воздушный налет, и обороняющейся, отражающей налет.

Наряду с оперирующими сторонами в операции могут участвовать арбитры и природные силы, поведение которых не подчинено стремлению к достижению цели операции.

Для достижения цели оперирующая сторона должна располагать некоторым запасом активных средств (ресурсов), используя или расходуя которые она может добиваться достижения цели. В качестве ресурсов в зависимости от сущности операции могут выступать: станки, запасы сырья, рабочая сила, денежные средства, средства противодействия в системе противовоздушной или противоракетной обороны и т. п.

Операция является управляемым мероприятием. Оперирующая сторона управляет операцией, выбирая те или иные способы использования ресурсов — способ действий. В качестве синонимов термина «способ действия» часто используют следующие термины: альтернатива, стратегия, управление, решение.

Возможности оперирующей стороны по управлению операцией всегда ограничены рядом естественных причин. Этот факт проявляется в наличии о г р а н и ч е н и й — дисциплинирующих условий — на выбор способов действий оперирующей стороны — стратегий. Стратегии, удовлетворяющие наложенным ограничениям, называются в с е в о з м о ж н ы м и или д о п у с т и м ы м и (в смысле наложенных ограничений). Понятие «допустимые стратегии» является относительным: класс допустимых стратегий определяется наложенными ограничениями и изменяется, если изменяются ограничения.

3. Реализация той или иной допустимой стратегии оперирующей стороны обычно приводит к различным исходам операции. Чтобы сравнивать между собой качество различных стратегий, нужно иметь возможность оценивать соответствующие исходы операций. Исход операции оценивается с помощью некоторых к р и т е р и е в к а ч е с т в а (критериев эффективности или критериев оптимальности). Критерий оптимальности является математическим выражением цели операции (математической моделью цели операции), позволяющим количественно определить (оценить) степень достижения этой цели. Стратегия, наилучшая в смысле выбранного критерия оптимальности, т. е. доставляющая ему требуемое экстремальное (максимальное или минимальное) значение, называется оптимальной стратегией (синонимами этого термина яв-

ляются термины «оптимальное решение», «оптимальное управление» и т. п.).

Следует всегда иметь в виду, что понятие «оптимальная стратегия» является не абсолютным, а относительным, как и понятие «допустимая стратегия». Не существует оптимальной стратегии вообще, всякая оптимальная стратегия является наилучшей лишь в некотором узком, совершенно конкретном смысле, определенном критерием оптимальности. Одна и та же стратегия, оптимальная в смысле одного критерия, может оказаться далеко не оптимальной и даже очень плохой по другому критерию.

Поскольку значение критерия оптимальности в любой операции зависит от каких-либо величин, описывающих свойства операции, используемые ресурсы и т. д., то критерий оптимальности часто называют также критерияльной или целевой функцией или функцией эффективности.

4. Следующее важное понятие исследования операций — и с л е д о в а т е л ь о п е р а ц и и. В составе оперирующей стороны специально выделяется и занимает особое место исследователь операции, или операционист. Он принадлежит к оперирующей стороне и должен преследовать ту же цель, что и оперирующая сторона. Однако он не принимает окончательных решений по выбору способов действий, а лишь помогает в этом оперирующей стороне, предоставляя ей количественные основания для принятия решений. Иными словами, исследователь операции имеет право не решающего, а лишь совещательного голоса. Естественно, что поэтому он не должен нести ответственности за принятые решения и последствия от реализации предпринятых действий.

Суть работы исследователя операции состоит в детальном изучении сущности и специфики решаемой проблемы, определении всего набора допустимых стратегий, оценке их качества, сравнении их между собой и определении оптимальной стратегии. Исследование операции завершается рекомендациями по выбору оптимальной стратегии. Само же принятие решения, т. е. окончательный выбор стратегии и ее реализация, выходит за рамки исследования и относится к компетенции ответственного лица — руководителя операции.

Обратимся к рассмотрению состава исследований, формирующих процесс принятия решений. Процессы принятия решений, реализуемые в самых различных сферах деятельности, имеют очень много общего, поэтому желательно создать некоторую универсальную, «типовую» схему, устанавливающую наиболее целесообразный набор и последовательность действий, проводимых при исследовании операций.

В настоящее время еще не выработана единая «технология», однако определенная тенденция в этом направлении имеется. В работах многих авторов по исследованию операций, системному анализу, управлению производством содержатся рекомендации по формированию состава и последовательности исследований в процессе принятия решений. На основании их анализа и обобщения

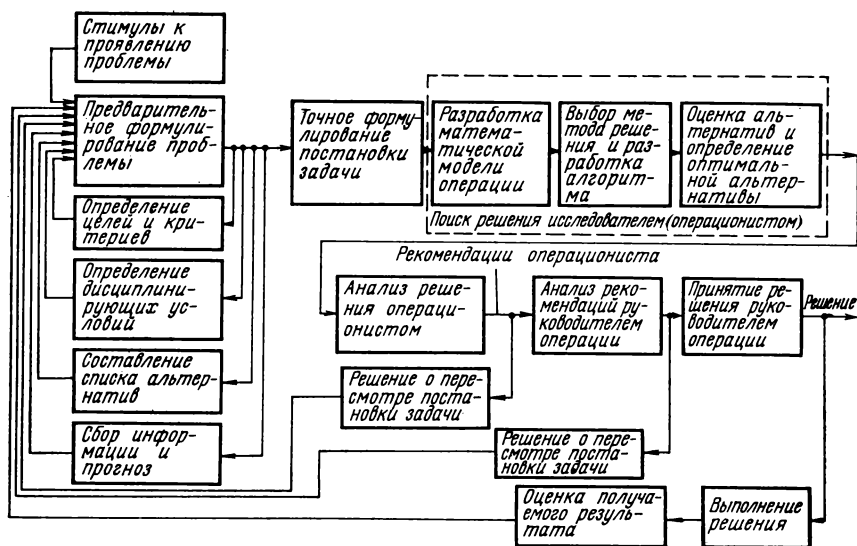


Рис. 24. Структурная схема процесса принятия решения

можно предложить следующий состав «типового» процесса принятия решений: 1) предварительное формулирование проблемы; 2) определение целей операции и выбор соответствующих критериев оптимальности; 3) выявление и формулирование дисциплинирующих условий; 4) составление возможно более полного списка альтернатив и предварительный их анализ с целью отбрасывания явно неэффективных; 5) сбор необходимой информации и прогнозирование изменений параметров операции в будущем; 6) точное формулирование постановки задачи; 7) разработку математической модели операции, позволяющей оценивать эффективность каждой альтернативы; 8) анализ и выбор метода решения задачи и разработка алгоритма решения; 9) оценку альтернатив и определение наиболее эффективных; 10) принятие решения ответственным руководителем; 11) выполнение решения и оценку результатов.

Процесс принятия решений является сложной итеративной циклической процедурой. Действительно, результат практически любого этапа исследований может повлиять на постановку задачи и привести к ее изменению. В частности, даже практическое опробование принятого решения, если оно дает нежелательный результат, также является стимулом к пересмотру постановки задачи и поиску новых решений. Структурная схема процесса исследования операций представлена на рис. 24. Указанные особенности процесса отображены с помощью обратных связей.

Отметим, что постановка задачи принятия решений индивидуальным лицом, принимающим решение (ЛПР), компактно может быть представлена в следующей символической форме:

$$(S_0, T, R \parallel S, \Pi, \Gamma, Y, f, K),$$

т. е. в условиях исходной проблемной ситуации S_0 , располагаемого времени T и ресурсов R необходимо доопределить ситуацию S_0 множеством ситуаций S , сформулировать множество целей C , ограничений Γ , решений Y , произвести оценку предпочтения решений f и найти оптимальное или приемлемое решение Y из множества решений, руководствуясь сформулированным критерием выбора K .

Ранее было отмечено, что в процессе принятия решений осуществляется преобразование информации. На различных этапах процесса принятия решений происходит количественная и качественная переработка исходной информации с целью получения новой информации, концентрированное выражение которой находит свое отражение в решении. Переработка информации происходит в виде последовательных фаз, для каждой из которых характерен свой уровень определенности решения задачи. Различают три фазы: 1) структуризацию, 2) характеризацию, 3) оптимизацию.

Структуризация заключается в выделении элементов решаемой задачи и установлении отношений между ними. В результате структуризации образуется логически упорядоченная система, позволяющая определить необходимую информацию и распределить ее по элементам структуры. Результат структуризации отображается в виде схем, таблиц или формальной символической записи. Примерами структуризации являются построение дерева целей, формирование множества решений, способов их реализации, установление взаимосвязанной системы гипотез о возможных событиях и т. п. (см. § 14).

Характеризация заключается в определении системы характеристик, параметров и показателей, количественно описывающих решаемую задачу. Например, для дерева целей могут быть определены приоритеты, или относительные коэффициенты важности целей: для множества решений — предпочтения (полезности), для гипотез о возможных событиях — вероятности их свершения и т. д. Проведение характеризации приводит к более полному и точному описанию решаемой задачи по сравнению с фазой структуризации и подготавливает исходные данные для выполнения оптимизации.

Оптимизация представляет собой наилучшее решение задачи. Именно на этой фазе вся имеющаяся информация преобразуется в конечную форму, описывающую решение и показывающую, как зависят характеристики решения от исходных данных. Проведение оптимизации приводит к полной определенности решения задачи. Для решения задач в условиях неопределенности не всегда возможно проведение фазы оптимизации в строго формальном виде. Во многих случаях лицо, принимающее решение, осуществляет оптимизацию в неявном виде, опираясь на некоторые общие принципы и свои предположения.

Итак, принятие решений невозможно без информации. Поэтому процесс принятия решений требует информационного обеспечения. Это обеспечение предполагает использование как имеющейся ин-

формации (априорной информации), так и получение, добывание с помощью экспериментов (опытов) новой послеопытной (апостериорной) информации.

Обобщенной характеристикой решения является э ф ф е к т и в н о с т ь р е ш е н и я. Эта характеристика включает эффект решения, определяющий степень достижения целей, и стоимость решения — совокупность затрат ресурсов для принятия и реализации решения. Таким образом, эффективность решения — это степень достижения целей, отнесенная к затратам на их достижение. Решение тем эффективнее, чем больше степень достижения целей и ниже стоимость затрат.

С и с т е м а п р и н я т и я р е ш е н и й — организованная совокупность людей, методов, технических средств, информации и технологии принятия решений для достижения поставленных целей [6, 10].

§ 17. Классификация задач принятия решений

В настоящее время не существует общепринятой универсальной классификации схемы задач принятия решений (ЗПР). Можно выделить следующие отдельные, важные классификационные признаки.

1. Число целей операции, преследуемых одной оперирующей стороной.

2. Наличие или отсутствие зависимости критерия оптимальности и дисциплинирующих условий от времени.

3. Наличие случайных и неопределенных факторов, влияющих на исход операции. Этот признак назван «определенность — риск — неопределенность».

По первому классификационному признаку ЗПР делятся на два больших класса: *одноцелевые*, или *однокритериальные* (скалярные), и *многоцелевые*, или *многокритериальные*, ЗПР.

Отметим, что многокритериальные задачи — задачи, которые состоят в поиске лучшего решения, удовлетворяющего нескольким несводимым друг с другом критериям. Для решения таких задач требуется их формализация, которая неизбежно связывается с экспертными оценками как самих критериев, так и взаимоотношений между ними.

Следует заметить также, что термин «многокритериальные задачи» нельзя смешивать с термином «задачи векторной оптимизации»: в последнем случае речь идет о сопоставлении однородных критериев разных участников (такие задачи решаются с помощью критерия оптимальности, по Парето). Нельзя также оба эти термина смешивать с термином «многоэкстремальные задачи», для которых характерны не разные критерии, а наличие у целевой функции не только глобального, но и локальных экстремумов.

По второму классификационному признаку ЗПР делятся на два больших класса: *статические* и *динамические*.

В статических ЗПР критериальная функция и функция ограничений не зависят от времени.

В статической экономико-математической модели значения всех переменных относятся к одному моменту времени. Статический подход к изучению экономики — это изучение состояния экономики на данный момент (статики) в отличие от динамического подхода, при котором экономика исследуется в развитии и движении. Динамические задачи сложнее статических. Модель является динамической, если как минимум одна ее переменная относится к периоду времени, отличному от времени, к которому отнесены другие переменные.

Динамические задачи отличают две характерные особенности:

а) в качестве критерия оптимальности в динамических ЗПР выступает обычно не функция, как в статических ЗПР, а функционал, зависящий от функции времени, описывающий поведение некоторых динамических объектов, участвующих в операции;

б) в числе дисциплинирующих условий в динамических ЗПР обычно присутствуют так называемые дифференциальные связи. Они представляют собой дифференциальные уравнения, описывающие поведение динамических объектов, участвующих в операции.

Для математического описания динамических моделей используются системы линейных дифференциальных уравнений (в моделях с непрерывным временем), разностные уравнения (в моделях с дискретным временем), а также системы обыкновенных алгебраических уравнений.

В качестве примера динамической ЗПР можно привести задачу вывода космического летательного аппарата в заданную точку пространства с заданной точностью и за заданное время с минимальным расходом топлива. В настоящее время динамические ЗПР еще не получили широкого применения в экономических исследованиях и поэтому нами в дальнейшем не рассматриваются.

По третьему классификационному признаку — «определенность — риск — неопределенность» — ЗПР делятся на три больших класса.

1. *Принятие решений при определенности* (детерминированные ЗПР). Такие ЗПР характеризуются однозначной, детерминированной связью между принятым решением и его исходом. Это наиболее простой и наиболее изученный случай принятия решений, когда относительно каждой стратегии оперирующей стороны заранее, до проведения операции, известно, что она неизменно приводит к некоторому конкретному результату. В детерминированных ЗПР критерий оптимальности и дисциплинирующие условия зависят только от стратегий оперирующей стороны и фиксированных детерминированных факторов, т. е. факторов, полностью известных оперирующей стороне.

2. *Принятие решений при риске* (стохастические ЗПР). В этом случае каждая стратегия оперирующей стороны может привести к одному из множества возможных исходов, причем каждый исход имеет определенную вероятность

появления. Предполагается, что принимающему решение эти вероятности заранее, до определения операции, полностью известны (во всяком случае они могут быть определены с любой требуемой для целей исследования степенью точности). В стохастических ЗПР критерий оптимальности зависит кроме стратегий оперирующей стороны и детерминированных факторов также от фиксированных стохастических факторов, т. е. от случайных факторов, законы распределения которых известны оперирующей стороне. Статистические характеристики (законы распределения, математические ожидания, дисперсии и т. п.) стохастических факторов, а также значения детерминированных факторов являются той исходной информацией, которая может быть использована исследователем операции при определении оптимальной стратегии.

Первое из приведенных названий рассматриваемого класса ЗПР — принятие решений при риске — связано со следующими обстоятельствами. Несмотря на то что все случайные явления и процессы, сопровождающие операцию и влияющие на ее исход, хорошо изучены и все их необходимые статистические характеристики полностью известны, исход каждой конкретной реализации операции заранее (до ее проведения) неизвестен, случаен. В этом смысле оперирующая сторона всегда рискует (в большей или меньшей степени) получить не тот результат, на который она ориентируется, выбирая свою оптимальную стратегию в расчете на осредненные, статистические характеристики случайных факторов.

3. Принятие решений в условиях неопределенности. В данных ЗПР критерий оптимальности зависит кроме стратегий оперирующей стороны и фиксированных факторов также от неопределенных факторов, неподвластных оперирующей стороне и неизвестных ей в момент принятия решения (или известных с недостаточной для принятия решения точностью). В результате влияния неопределенных факторов каждая стратегия оперирующей стороны оказывается связанной с множеством возможных исходов, вероятности которых либо неизвестны оперирующей стороне (или известны с недостаточной для принятия решения точностью), либо вовсе не имеют смысла. Первое соответствует неопределенным факторам стохастической природы (т. е. недостаточно изученным стохастическим факторам, относительно которых отсутствует необходимая статистическая информация), второе — неопределенным факторам нестохастической природы.

Детерминированные ЗПР и ЗПР в условиях неопределенности можно считать предельными случаями ЗПР (т. е. полное знание и полное незнание). ЗПР, в которых имеется элемент риска, занимают некоторое промежуточное положение. Очевидно, что любой предельный случай всегда представляет собой большую или меньшую идеализацию реальной ситуации.

Классификацию ЗПР завершим указанием на математический аппарат, применяемый при решении ЗПР того или другого класса. Однокритериальные статические детерминированные ЗПР в своей общей постановке полностью совпадают с общей постановкой

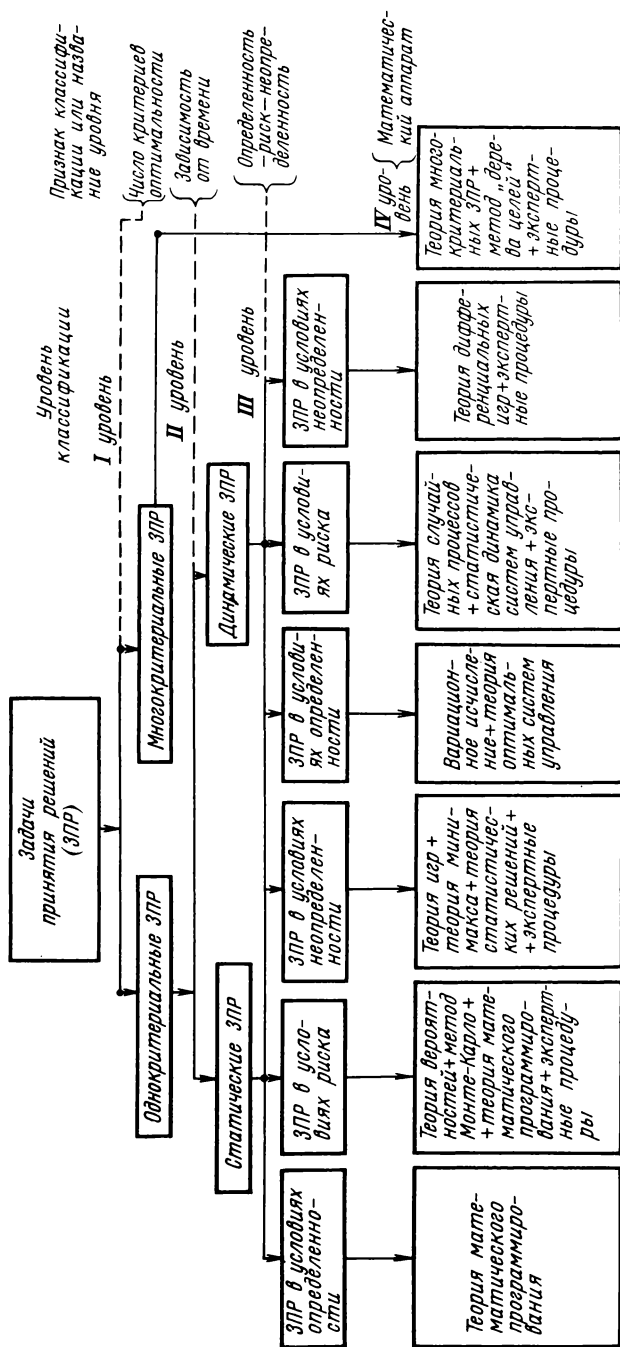


Рис. 25. Классификационное «дерево» ЗПР и методы решения

задачи математического программирования (МП).

Однокритериальные статические ЗПР в условиях риска решаются с использованием методов теории вероятностей и математического программирования. При моделировании задач этого класса находит широкое применение метод статистических испытаний (метод Монте-Карло).

При решении однокритериальных статических ЗПР в условиях неопределенности применяют ряд математических дисциплин: теорию игр, теорию минимакса, теорию статистических решений. Выбор той или иной теории диктуется природой неопределенных факторов.

Наименее разработанным классом ЗПР являются многокритериальные ЗПР [4, 10], хотя именно они должны иметь наибольшее применение в экономических исследованиях. Действительно, практически любое экономическое исследование столь сложно, что при его изучении и формализации в виде некоторой операции бывает очень трудно, а подчас и невозможно, ограничиться одной целью операции и соответственно одним критерием оптимальности. В настоящее время теория многокритериальных ЗПР находится в стадии становления. Наибольшие успехи достигнуты здесь в отношении статических детерминированных задач.

На рис. 25 представлено классификационное «дерево» ЗПР, соответствующее выделенным выше классификационным признакам. Здесь подробно выделена лишь «ветвь», связанная с однокритериальными ЗПР.

Очевидно, что любая реальная ЗПР может удовлетворять одновременно нескольким из перечисленных выше классификационных признаков, т. е. представлять собой комбинацию из рассмотренных классов ЗПР.

§ 18. Общая постановка однокритериальной задачи принятия решений

Имеется некоторая операция, т. е. управляемое мероприятие, на исход которого оперирующая сторона может в какой-то мере влиять. Эффективность этого управления характеризуется некоторым критерием оптимальности F , допускающим количественное представление. Критерий оптимальности может быть задан либо в виде функции (с помощью одного из способов задания функции), либо в более сложных случаях в виде функционала или может иметь лишь алгоритмическое задание.

Величина критерия оптимальности зависит от ряда факторов, которые можно разбить на две группы: 1) контролируемые (управляемые) факторы, выбор которых находится в распоряжении оперирующей стороны. Каждый конкретный выбор значений контролируемых факторов представляет собой стратегию оперирующей стороны; 2) неконтролируемые (неуправляемые) факторы, на которые оперирующая сторона влиять не может. В состав неконтроли-

руемых факторов может входить и время, если в операции участвуют динамические объекты, изменяющие свои свойства и поведение во времени.

Неконтролируемые факторы в зависимости от информированности о них исследователя операции можно разбить на три группы:

1) детерминированные факторы — неслучайные фиксированные факторы, или неслучайные величины, значения которых полностью известны оперирующей стороне до проведения операции;

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

3) неопределенные факторы, для каждого из которых известна только область возможных знаний фактора или область, внутри которой находится закон распределения, если фактор случаен. В последнем случае имеет место неопределенный закон распределения случайного фактора. Значения неопределенных факторов неизвестны оперирующей стороне в момент принятия решения о выборе оптимальной стратегии.

В соответствии с выделенными факторами критерий оптимальности F можно представить в виде зависимости

$$F = F(X_1, X_2, \dots, X_l, A_1, A_2, \dots, A_p, Y_1, Y_2, \dots, Y_q, Z_1, Z_2, \dots, Z_r, t), \quad (\text{III.1})$$

где $X_1, \dots, X_l$ — контролируемые факторы; $A_1, \dots, A_p$ — неконтролируемые детерминированные факторы; $Y_1, \dots, Y_q$ — неконтролируемые стохастические факторы; $Z_1, \dots, Z_r$ — неконтролируемые неопределенные факторы.

Величины X, A, Y, Z в общем случае могут быть массивами любой размерности: скалярами, векторами, матрицами и т. д.

Величины контролируемых (управляемых) факторов обычно ограничены рядом естественных причин, например ограниченностью располагаемых для операции ресурсов. Математически эти ограничения записываются в виде дисциплинирующих условий:

$$g_i = g_i(X_1, \dots, X_l, A_1, \dots, A_p, Y_1, \dots, Y_q, Z_1, \dots, Z_r, t) \{ \leq, =, \geq \} b_i, i \in \overline{1, m}, \quad (\text{III.2})$$

где $A_1, \dots, A_{p_i}$ — неконтролируемые детерминированные факторы; $Y_1, \dots, Y_{q_i}$ — неконтролируемые стохастические факторы; $Z_1, \dots, Z_{r_i}$ — неконтролируемые неопределенные факторы. Условное обозначение $\{ \leq, =, \geq \}$ есть краткая запись того факта, что в каждом i -м ограничении из совокупности m ограничений имеет место только один из знаков $\leq, =, \geq$, но разные ограничения могут иметь разные знаки.

Условия (III.2) определяют области $\Omega_{X_1}, \dots, \Omega_{X_l}$ пространства, внутри которых расположены возможные (допустимые) значения контролируемых факторов $X_1, X_2, \dots, X_l$.

Аналогичным образом могут быть ограничены и области возможных значений неконтролируемых факторов. В ЗПР обычно предполагается, что эти области известны оперирующей стороне.

Поскольку критерий оптимальности есть количественная мера степени достижения цели операции, то математически цель операции выражается в стремлении к максимально возможному увеличению (или уменьшению) значения критерия F , что можно записать в виде

$$F \rightarrow \max \text{ (или } \min \text{)}. \quad (\text{III.3})$$

Средством достижения этой цели является соответствующий выбор оперирующей стороной управлений $X_1, X_2, \dots, X_l$ из областей $\Omega_{X_1}, \Omega_{X_2}, \dots, \Omega_{X_l}$ их допустимых значений.

Таким образом, перед лицом, ответственным за принятие решения, стоит задача, которую можно сформулировать следующим образом: при заданных значениях и характеристиках фиксированных неконтролируемых факторов $A_1, \dots, A_p, Y_1, \dots, Y_q$ с учетом неопределенных факторов $Z_1, \dots, Z_l$ найти оптимальные значения $\bar{X}_1, \dots, \bar{X}_l$ управлений $X_1, \dots, X_l$ из областей $\Omega_{X_1}, \dots, \Omega_{X_l}$ их допустимых значений, которые по возможности обращали бы в максимум (минимум) критерий оптимальности F . Данная формулировка с некоторыми изменениями заимствована из работы Е. С. Вентцель [6].

Приведенная общая формулировка задачи принятия решения не является строго математической, о чем свидетельствует оговорка «по возможности». Эта нестрогость обусловлена прежде всего наличием неконтролируемых неопределенных, а также стохастических факторов.

§ 19. Принятие решений в условиях определенности. Общая постановка однокритериальной статической детерминированной ЗПР

Задаче принятия решений рассматриваемого класса в отличие от общей задачи, описанной в § 18, можно дать строгую математическую постановку. Она может быть представлена в следующем виде.

Допустим, имеет место некоторая операция, т. е. управляемое мероприятие, исход которого зависит от стратегий оперирующей стороны и некоторых неслучайных фиксированных факторов, полностью известных оперирующей стороне. Стратегии (управления) оперирующей стороны могут быть представлены в виде значений n -мерного вектора $X = (x_1, \dots, x_j, \dots, x_n)$, на компоненты которого наложен ряд ограничений, обусловленных физической и экономической сущностью задачи, имеющих вид

$$g_i = g_i(A_i, X) \{ \leq, =, \geq \} b_i; \quad (\text{III.4})$$

$$i \in \overline{1, m}; \quad m \{ <, =, > \} n,$$

где A_i — некоторый массив фиксированных неслучайных параметров.

Условия (III.4) определяют область Ω_X допустимых значений стратегии X .

Эффективность управления характеризуется некоторым численным критерием оптимальности вида

$$F = F(X, C), \quad (\text{III.5})$$

где C — массив фиксированных неслучайных параметров.

Массивы C и A_i в формулах (III.4) и (III.5) характеризуют свойства объектов, участвующих в операции, и условия протекания операции. Они предполагаются известными.

Цель оперирующей стороны состоит в максимизации значения критерия оптимальности F , что условно можно представить в виде

$$F = F(X, C) \rightarrow \max. \quad (\text{III.6})$$

Средством достижения этой цели является соответствующий выбор вектора управления X из области Ω_X . Следовательно, перед лицом, ответственным за принятие решения, стоит следующая за-

дача: требуется найти такое значение $\bar{X} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$ вектора управления $X = (x_1, x_2, \dots, x_n)$ из области Ω_X его допустимых значений, которое максимизирует значение критерия оптимальности F , а также найти значение $\bar{F}$ этого максимума; иначе, требуется найти значения $\bar{X}$ вектора X и $\bar{F}$ функции F , удовлетворяющие условию

$$F = F(\bar{X}, C) = \max F(X, C) \quad (\bar{X} \in \Omega_X), \quad (\text{III.7})$$

где область Ω_X представляется условием (III.4).

В формуле (III.7) символы $\bar{F}$ и $\bar{X}$ обозначают максимально достижимое в условиях (III.4) решаемой задачи значение критерия оптимальности F и соответствующее ему оптимальное значение вектора управления X (оптимальную стратегию). Значения $\bar{F}$ и $\bar{X}$ представляют собой решение задачи.

Совокупность выражений (III.4), (III.5), (III.7) есть общий вид математической модели однокритериальной статической детерминированной ЗПР.

§ 20. Пример обоснования решения однокритериальной статической детерминированной задачи принятия решений

Любое исследование, направленное на решение ЗПР, должно содержать как минимум шесть пунктов.

Мы рассмотрим пример решения одной простой задачи рассматриваемого класса — задачи оптимального планирования загрузки производственного оборудования. Пример строится по указанной в § 12 схеме (за исключением п.п. 4,6).

1. *Постановка экономической проблемы.* Рассмотрим операцию оптимального планирования загрузки оборудования производственного участка, выпускающего продукцию одного вида.

Производство продукции связано с затратами ряда исходных ресурсов — трудовых, материальных, энергетических, запасы которых для рассматриваемого планового периода ограничены заданными величинами. Для производства конечной продукции участок располагает несколькими заранее отработанными технологическими способами. Каждый технологический способ характеризуется своей производительностью, т. е. количеством продукции, производимой в единицу времени, а также соответствующими удельными расходами всех видов ресурсов, т. е. расходами в единицу времени. Технологические способы независимы между собой в том смысле, что они могут быть использованы одновременно.

Руководитель участка заинтересован в максимизации суммарного за время планового периода выпуска продукции, возможного при имеющихся ресурсах.

Средством достижения этой цели является выбор соответствующего плана загрузки технологических способов, представляющего собой совокупность времен работы каждого технологического способа в течение рассматриваемого планового периода. Следовательно, для достижения поставленной цели плановому подразделению предприятия требуется решить задачу: найти оптимальный в условиях располагаемых ресурсов план загрузки технологических способов, максимизирующий общий выпуск продукции за плановый период, а также величину максимально возможного выпуска.

2. Построение математической модели. Формулируем условия задачи. Число имеющихся технологических способов производства обозначим n и каждому способу присвоим порядковый номер j , где $j \in \overline{1, n}$. Общее количество типов ресурсов, потребных для производства конечной продукции, обозначим через m и каждому типу ресурсов присвоим порядковый номер i , где $i \in \overline{1, m}$.

Запасы ресурсов, которыми располагает производственный участок, представим в виде m -мерного вектора $\mathbf{B} = (b_i)$, где b_i — величина запаса i -го типа ресурсов.

Характеристику совокупности n технологических способов производства представим в виде n -мерного вектора производительности $\mathbf{C} = (c_j)$ и матрицы удельных расходов $\mathbf{A}$ размера mn :

$$\mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & & a_{mn} \end{vmatrix} = |a_{ij}|, \quad (\text{II.8})$$

где c_j — производительность j -го способа, т. е. количество продукции, производимой по j -му технологическому способу в единицу времени; a_{ij} — расход i -го ресурса за единицу времени использования j -го технологического способа.

Векторы $\mathbf{B}$, $\mathbf{C}$ и матрица $\mathbf{A}$ представляют собой совокупность параметров, характеризующих свойства объекта исследования про-

изводственного участка. Предполагается, что эти параметры являются неслучайными величинами, полностью известными руководителю участка и плановым органам.

Искомый план функционирования участка может быть представлен в виде n -мерного вектора $X = (x_j)$, компонента которого x_j есть время (в пределах рассматриваемого планового периода), в течение которого участок выпускает продукцию с использованием j -го технологического способа.

Вектор X представляет собой вектор управления. Компоненты его связаны рядом соотношений вида

$$\sum_{j=1}^n a_{ij}x_j \leq b_i; \quad i \in \overline{1, m}; \quad (III.9)$$

$$0 \leq x_j \leq T; \quad j \in \overline{1, n}, \quad (III.10)$$

где T — длительность планового периода.

Соотношения (III.9) обусловлены ограниченностью располагаемых ресурсов, а условия (III.10) отражают тот факт, что компоненты плана X не могут быть отрицательными в силу их физической сущности и не должны превышать длительность планового периода.

Условия (III.9) (III.10) определяют область Ω_X допустимых планов.

Эффективность функционирования производственного участка в рассматриваемой задаче оценивается общим количеством F продукции, произведенной всей совокупностью технологических способов за плановый период. Очевидно, что величина критерия оптимальности F является функцией плана функционирования участка X и вектора производительности C , а именно:

$$F = F(X, C) = \sum_{j=1}^n c_j x_j. \quad (III.11)$$

Цель руководителя участка математически выражается в стремлении максимизировать значение критерия оптимальности, т. е. $F \rightarrow \max$, за счет соответствующего выбора плана X из области Ω_X его допустимых значений. Отсюда следует задача: требуется найти оптимальный план $\bar{X} = (\bar{x}_j)$ и максимальный выпуск $\bar{F}$, удовлетворяющие условию

$$\bar{F} = F(\bar{X}, C) = \max_{X \in \Omega_X} \left| \sum_{j=1}^n c_j x_j \right|, \quad (III.12)$$

где Ω_X определяется условиями (III.9), (III.10).

Итак, дана критериальная функция

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

где $X = (x_j)$, $C = (c_j)$ — n -мерные векторы.

Вектор C задан, вектор $\bar{X}$ является искомым. Требуется найти значения $\bar{F}$ функции F и $\bar{X}$ вектора X , удовлетворяющие условию

$$\bar{F} = F(\bar{X}, C) = \max_{X \in \Omega_X} \left| \sum_{j=1}^n c_j x_j \right|,$$

где область Ω_X определяется условиями

$$\left. \begin{array}{l} \sum_{j=1}^n a_{ij} x_j \leq b_i; \quad i \in \overline{1, m} \\ 0 \leq x_j \leq T; \quad j \in \overline{1, n} \end{array} \right\} \quad (\text{III.13})$$

где матрица $A = \|a_{ij}\|$ размера $m \times n$ и величина T заданы.

Совокупность выражений (III.11)—(III.13) есть математическая модель задачи.

3. *Математический анализ модели и выбор метода решения.* Рассматриваемая задача относится к классу однокритериальных статических детерминированных ЗПР, поскольку 1) она имеет единственный критерий оптимальности и 2) ее критериальная функция и функция ограничений не зависят от времени, случайных и неопределенных факторов. Поэтому для ее решения могут быть использованы методы математического программирования.

Как следует из анализа математической постановки (III.11)—(III.13) рассматриваемой задачи оптимального планирования, она представляет собой типичную задачу линейного программирования. Поэтому для ее решения может быть применен один из методов линейного программирования, в частности универсальный метод линейного программирования — симплекс-метод.

§ 21. Обзор методов решения однокритериальных статических детерминированных задач принятия решений

При определении возможных методов решения любой ЗПР следует исходить из анализа математической модели задачи. Как видно из рассмотрения математической модели (III.4), (III.5), (III.7) общей постановки однокритериальной статической детерминированной ЗПР, она полностью совпадает с общей постановкой задачи математического программирования (МП). Поэтому весь богатый арсенал методов, разработанных для решения задач МП, может и должен быть использован для решения ЗПР рассматриваемого класса.

Математическое программирование представляет собой быстро развивающийся раздел современной прикладной математики. Содержание математического программирования составляют теория и методы решения задач о нахождении экстремумов функции в ограниченной области допустимых значений переменных, определяемой системой ограничений (равенств и неравенств). Термин «математическое программирование» получил широкое распространение в математической и инженерно-экономической литературе, однако его

нельзя считать достаточно удачным. Этот термин не имеет ничего общего с программированием для цифровых вычислительных машин и потому не является достаточно содержательным для определения существа соответствующей математической дисциплины. Исторически этот термин ведет свое происхождение из первых публикаций на тему о нахождении решения неклассических экстремальных задач, где решения задачи называли планом или программой.

Общая задача МП может быть сформулирована следующим образом [29]: требуется найти значения n переменных $x_1, x_2, \dots, x_n$, которые удовлетворяют системе из m соотношений:

$$g_i(x_1, x_2, \dots, x_n) \{ \leq, =, \geq \} b_i; \quad i \in \overline{1, m}, \quad (\text{III.14})$$

где b_i — заданные константы, g_i — заданные функции, и максимизируют функцию

$$F = F(x_1, x_2, \dots, x_n). \quad (\text{III.15})$$

Соотношения (III.14) в задачах МП называются **ограничениями**. Совокупность этих ограничений образует область Ω_X допустимых значений переменных $x_1, x_2, \dots, x_n$. Очевидно, что эта область не должна быть пуста. В противном случае исчезла бы свобода (хотя и ограниченная) в выборе значений переменных, которая позволяет сравнивать значения функции F в разных точках и определять искомый максимум и которая составляет принципиальную особенность задач МП.

Для краткости изложения в дальнейшем будем пользоваться следующей условной формой записи задачи МП в постановке (III.14), (III.15):

$$\left\{ \begin{array}{l} F(X) \rightarrow \max; \\ g_i(X) \{ \leq, =, \geq \} b_i; \quad i \in \overline{1, m}, \end{array} \right\} \quad (\text{III.16})$$

где $X = (x_1, x_2, \dots, x_n)$ — n -мерный вектор искомых переменных.

К задаче МП, представленной формулами (III.14), (III.15), можно отнести следующие общие замечания.

1). В каждом отдельном ограничении из системы (III.14) сохраняется только один из знаков $\leq, =, \geq$, причем разные ограничения могут иметь, конечно, разные знаки.

2). Величины n и m (n — число переменных в задаче, m — число ограничений) в общей задаче МП между собой не связаны и могут принимать любые значения при условии $n \geq 1, m \geq 0$, причем m может быть меньше, больше или равно n , т. е. $m \{ <, =, > \} n$. Исключение составляет случай, когда все m ограничений (III.14) являются строгими равенствами; в этом случае величины n и m должны быть связаны соотношением $0 \leq m \leq n$, т. е. число уравнений связи должно быть не больше числа переменных. В частном случае m может быть и нулем, так что общая постановка (III.14), (III.15) включает и случай, когда все ограничения (III.14) отсутствуют.

3). В задачах МП различают и обычно по-разному оформляют два вида ограничений: областные и функциональные. Каждое областное ограничение относится только к одной переменной, ограничивая область ее изменения. Областное ограничение может иметь, например, следующий вид:

$$x_j \geq b_j. \quad (\text{III.17})$$

Областное ограничение (III.17) представляет собой простейший частный случай ограничения (III.14), когда функция $g_i(X)$ имеет простейший вид $g_i(X) = x_j$.

В задачах МП встречаются и так называемые двусторонние областные ограничения:

$$b_j^{\min} \leq x_j \leq b_j^{\max}. \quad (\text{III.18})$$

Одно двустороннее ограничение равносильно двум односторонним ограничениям, имеющим вид, подобный выражению (III.17)

$$\left. \begin{aligned} x_j &\geq b_j^{\min}; \\ x_j &\leq b_j^{\max}. \end{aligned} \right\} \quad (\text{III.19})$$

4). Частным случаем областных ограничений являются так называемые ограничения на знак переменных. Эти ограничения выражают требования, состоящие в том, что некоторые или все переменные задачи должны удовлетворять условию неотрицательности. Условия неотрицательности имеют вид

$$x_j \geq 0; \quad j \in \overline{1, p}; \quad p \leq n. \quad (\text{III.20})$$

Требование (III.20) возникло в сфере приложений задач МП к решению практических ЗПР и отражает тот факт, что в большинстве практических ЗПР искомые переменные не могут принимать отрицательных значений вследствие их физической или экономической сути.

5). В отдельных задачах МП в качестве ограничений дополнительно к требованию (III.20) может фигурировать условие, по которому некоторые или все переменные могут принимать лишь некоторые дискретные значения, например только целочисленные. Это условие может иметь в частном случае вид

$$x_j = 0, \Delta_j, 2\Delta_j, 3\Delta_j, \dots, j \in \overline{1, k}; k \leq n, \quad (\text{III.21})$$

где Δ_j — некоторая константа.

Условия, подобные (III.21), можно назвать требованиями дискретности или ограничениями на непрерывность переменных. Если в частном случае в условии (III.21) $\Delta_j = 1$, то оно превращается в требование целочисленности переменных:

$$x_j = 0, 1, 2, 3, \dots, j \in \overline{1, k}; k \leq n. \quad (\text{III.22})$$

Условимся в дальнейшем изложении обозначение вида (III.14) оставлять лишь для собственно функциональных ограничений. Наличие же областных ограничений, требований неотрицательно-

сти и дискретности будем отмечать особо с помощью обозначений вида (III.17)—(III.22).

В настоящее время не существует единого метода, одинаково пригодного для решения всех задач МП. В зависимости от вида функций $F(X)$ и $g_i(X)$ среди задач МП выделяют частные типы задач, для решения которых разработаны специальные методы.

Обратимся к классификации задач МП и обзору методов их решения. В основу существующей классификации задач МП положены различия в характере критериальной функции $F(X)$ и функций ограничения $g_i(X)$.

По классификационной схеме задачи МП прежде всего подразделяются на два больших класса: классические и неклассические. Основным признаком такого деления выступает дифференцируемость критериальной функции и функциональных ограничений.

1. *Классические задачи* МП (классические задачи оптимизации). К ним относятся задачи, которые удовлетворяют совокупности следующих признаков: а) непрерывности критериальной функции $F(X)$ и функциональных ограничений $g_i(X)$ и наличия у них непрерывных частных производных по крайней мере второго порядка; б) отсутствия среди функциональных ограничений неравенств, что влечет за собой требование $m \leq n$; в) отсутствия областных ограничений и требований неотрицательности; г) отсутствия требований дискретности переменных. Если хотя бы одно из этих требований не удовлетворяется, задача не может быть классифицирована как классическая задача МП. Следует отметить, что из перечисленных требований принципиальными являются лишь первое и последнее. Если же не выполняются другие требования, то за счет введения вспомогательных переменных и, следовательно, увеличения размерности задачи эти требования всегда могут быть удовлетворены. (Под размерностью задачи здесь понимается число переменных.) Увеличение размерности задачи обычно усложняет процесс ее решения.

Классические задачи МП, в свою очередь, подразделяются на два подкласса по признаку отсутствия или наличия ограничений: задачи отыскания безусловного экстремума и задачи отыскания условного экстремума с ограничениями — равенствами.

2. *Неклассические задачи* МП (неклассические задачи оптимизации). В таких задачах на значения вектора управления $X = (x_j)$, кроме функциональных ограничений вида

$$g_i(X) \{ \leq, =, \geq \} \bar{b}_i, \quad i \in \overline{1, m},$$

где $g_i(X)$ — функция, обычно накладываются ограничения на знак, выражающие требование неотрицательности всех или некоторых компонент вектора X :

$$x_j \geq 0, \quad j \in \overline{1, p}, \quad p \leq n, \quad [\text{см. (III.20)}]$$

В дальнейшем для простоты будем считать, что требование (III.20) предъявляется ко всем n компонентам вектора X , т. е.

$p = n$. Если в частном случае это не так, то всегда за счет введения дополнительных переменных можно перейти к случаю $p = n$.

Общая постановка неклассической непрерывной (т. е. не имеющей требований дискретности) задачи МП обычно записывается в виде

$$\left. \begin{aligned} F(X) &\rightarrow \max(\min); \\ g_i(X) &\geq b_i; \quad i \in \overline{1, m}; \\ x_j &\geq 0; \quad j \in \overline{1, n}. \end{aligned} \right\} \quad (\text{III.23})$$

Заметим, что многие авторы публикаций по математическому программированию задачи МП отождествляют лишь с неклассическими задачами (III.23), выделяя классические задачи в особый класс задач оптимизации.

Неклассические задачи МП обычно подразделяют на два подкласса: специальные и неспециальные. К специальным задачам МП относят такие, для которых вследствие каких-то специфических особенностей их критериальных функций и функциональных ограничений разработаны специальные методы решения, например задача линейного программирования. Специфическая особенность этой задачи — линейность критериальной функции и функциональных ограничений.

Совокупность специальных методов в работах по математическому программированию получила и другое наименование — *прямые методы* поиска экстремума. В этом названии нашел отражение тот факт, что траектория поиска экстремума при пользовании этими методами проходит не непосредственно в направлении экстремума, а, образно говоря, какими-то окольными путями с учетом специфических особенностей критериальной функции и функциональных ограничений. Например, при решении задачи линейного программирования симплекс-методом траектория поиска начинается в некоторой начальной опорной точке на границе области допустимых значений вектора управления и затем переходит из одной опорной точки в другую в направлении возрастания критериальной функции.

Неспециальные задачи МП решают с использованием неспециальных методов поиска экстремума. Неспециальные методы имеют и другое название — *прямые методы* поиска экстремума. Эти методы являются универсальными в том смысле, что при их разработке учитывались не какие-либо специфические особенности конкретной задачи оптимизации, а лишь самые общие соображения, связанные с задачей отыскания экстремума. Прямые методы применимы для решения как специальных, так и неспециальных задач МП.

Разделение задач МП на специальные и неспециальные является условным, отражающим сегодняшнее состояние развития МП. С ростом достижений МП класс специальных задач МП будет непрерывно расширяться. Таким образом, список специальных задач МП

представляет собой своеобразный реестр теоретических достижений МП.

Перечислим основные типы специальных неклассических задач МП.

1. Задачи линейного программирования (ЛП). Они характеризуются тем, что функции $F(X)$ и $g_i(X)$ являются линейными по X .

Общая постановка задачи ЛП может быть представлена в виде

$$\left. \begin{aligned} F(X) &= \sum_{j=1}^n c_j x_j \rightarrow \max; \\ g_i(X) &= \sum_{j=1}^n a_{ij} x_j = \bar{b}_i; \quad i \in \overline{1, m}; \\ x_j &\geq 0; \quad j \in \overline{1, n}, \end{aligned} \right\} \quad (\text{III.24})$$

где b_i, c_j, a_{ij} — заданные постоянные величины.

Все остальные задачи МП, несводимые к постановке (III.24), являются нелинейными задачами МП.

Задачи ЛП представляют собой хорошо разработанный класс задач МП [4, 29, 31]. При пользовании линейными постановками ЗПР всегда, однако, следует иметь в виду, что линейное, а тем более детерминированное описание ЗПР обычно представляет собой довольно грубое приближение реальной задачи; более детальный анализ задачи часто позволяет обнаружить нелинейные и стохастические явления (под нелинейным явлением понимается такое, в котором отсутствует прямая пропорциональность между причиной и результатом).

2. Задачи квадратичного программирования. Они характеризуются квадратичной зависимостью критериальной функции $F(X)$ и линейной зависимостью функциональных ограничений $g_i(X)$ от X .

Общая постановка задачи квадратичного программирования имеет вид

$$\left. \begin{aligned} F(X) &= \sum_{j=1}^n c_j x_j + \sum_{j=1}^n \sum_{k=1}^n d_{jk} x_j x_k \rightarrow \max; \\ g_i(X) &= \sum_{j=1}^n a_{ij} x_j \geq b_i; \quad i \in \overline{1, m}; \\ x_j &\geq 0; \quad j \in \overline{1, n}, \end{aligned} \right\} \quad (\text{III.25})$$

где b_i, c_j, a_{ij}, d_{jk} — заданные постоянные величины.

Для решения задач этого типа разработаны специальные методы, базирующиеся на теории Куна—Таккера.

3. Задачи выпуклого программирования. В этих задачах критериальная функция $F(X)$ и функции ограничений $g_i(X)$ относятся к классу выпуклых функций.

Общая постановка задачи выпуклого программирования по форме совпадает с общей постановкой (III.23) неклассической за-

дачи МП. Для решения задач выпуклого программирования успешно применяются методы возможных направлений [10].

Отметим, что задачи линейного и квадратичного программирования представляют собой частные случаи задачи выпуклого программирования, поскольку линейные и квадратичные функции принадлежат к классу выпуклых функций. Следовательно, методы, разработанные для решения задач выпуклого программирования, применимы также для решения задач линейного и квадратичного программирования, но не наоборот.

4. Задачи с сепарабельными критериальными функциями и линейными ограничениями. Сепарабельной критериальной функцией называется такая функция n переменных, которая может быть представлена в виде суммы или произведения n функций одной переменной, т. е.

$$F(X) = F(x_1, \dots, x_n) = \sum_{j=1}^n f_j(x_j) = f_1(x_1) + f_2(x_2) + \dots + f_n(x_n), \quad (\text{III.26})$$

или

$$F(X) = f(x_1, \dots, x_n) = \prod_{j=1}^n f_j(x_j) = f_1(x_1) \cdot f_2(x_2) \cdot \dots \cdot f_n(x_n). \quad (\text{III.27})$$

В случае (III.26) функция F называется аддитивной, а в случае (III.27) — мультипликативной сепарабельной функцией. В формуле (III.27) символ $\prod_{j=1}^n$ означает произведение n функций.

Постановка задачи имеет вид

$$\left. \begin{aligned} F(X) &= \left\{ \sum_{j=1}^n f_j(x_j) \quad \text{или} \quad \prod_{j=1}^n f_j(x_j) \right\} \rightarrow \max; \\ g_i(X) &= \sum_{j=1}^n a_{ij}x_j \geq b_i; \quad i \in \overline{1, m}; \\ x_j &\geq 0; \quad j \in \overline{1, n}. \end{aligned} \right\} \quad (\text{III.28})$$

Для решения подобных задач может быть использован метод динамического программирования Беллмана [4, 2].

5. Задачи дискретного программирования. Сюда может быть отнесена любая задача МП, в которой есть дополнительное ограничение на вектор управления X , состоящее в требовании, чтобы все (полностью дискретные задачи) или некоторые (частично дискретные задачи) компоненты вектора X принимали только дискретные не отрицательные значения, т. е.

$$\left. \begin{aligned} x_k &= 0, \Delta_k, 2\Delta_k, 3\Delta_k, \dots; \\ k &\in \overline{1, p}; \quad p \leq n; \\ x_l &\geq 0; \quad l \in \overline{(p+1), n}, \end{aligned} \right\} \quad (\text{III.29})$$

где n — размерность вектора X , Δ_k — некоторый дискрет (положительная величина).

Частным случаем задач дискретного программирования являются задачи целочисленного программирования. В них $\Delta_k = 1$.

Глава IV

КЛАССИЧЕСКИЕ ЗАДАЧИ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

§ 22. Задачи отыскания безусловного экстремума

В этих задачах отсутствуют ограничения на область допустимых значений вектора переменных $X = (x_1, \dots, x_n)$, т. е. отсутствуют функции $g_i(X)$. Следовательно, область $\Omega(X)$ совпадает со всем n -мерным пространством переменных.

Таким образом, запись задачи математического программирования (III.16) для случая отыскания безусловного экстремума имеет вид

$$F(X) \rightarrow \max. \quad (\text{IV.1})$$

Задача (IV.1) является классической и рассматривается во всех курсах анализа. Теория решения таких задач получила развитие еще в трудах Эйлера, Лагранжа, Бернулли, Лейбница. Она не потеряла своего значения и в настоящее время, несмотря на то что с тех пор разработаны более общие методы, включающие классические как частный случай.

В дальнейшем будет показано, что решение задачи безусловной оптимизации (IV.1) классическим методом состоит из следующих этапов: 1) определения так называемых стационарных точек, 2) анализа тем или иным способом стационарных точек с целью выявления всех максимумов функции $F(X)$, 3) сравнения между собой максимальных значений функции $F(X)$ с целью определения ее глобального максимума.

Здесь мы последовательно рассмотрим задачу безусловной максимизации функции одной переменной ($X = E_1$), задачу нахождения безусловного экстремума функции нескольких переменных ($X = E_n$) и, наконец, задачу на условный (относительный) экстремум, т. е. задачу максимизации функции многих переменных при наличии ограничений типа равенств, когда x — множество решений уравнения

$$g(\bar{x}) = \bar{b},$$

где $g(\bar{x})$ — m -мерная вектор-функция, $m < n$.

1. Предварительные рассмотрения

Прежде всего сформулируем понятие р е ш е н и я задачи (IV.1).

Вектор переменных $\bar{x}^*$ является точкой г л о б а л ь н о г о

максимума (или решением), если он принадлежит к допустимому множеству X и функция принимает на этом векторе значение не меньше, чем в любой другой допустимой точке:

$$\bar{x}^* \in X \text{ и } F(\bar{x}^*) \geq F(\bar{x}) \text{ для всех } \bar{x} \in X. \quad (\text{IV.2})$$

Будем считать глобальный максимум строгим (сильным), если значение функции при $\bar{x} = \bar{x}^*$ строго больше любого другого значения функции на допустимом множестве, т. е.

$$F(\bar{x}^*) > F(\bar{x}) \text{ для всех } \bar{x} \in X, \bar{x} \neq \bar{x}^*. \quad (\text{IV.3})$$

Очевидно, что строгий глобальный максимум всегда единствен.

Вектор переменных $\bar{x}^*$ есть точка локального максимума, если он принадлежит к допустимому множеству X и на нем достигается значение функции, большее или равное значениям функции в некоторой малой окрестности этого вектора:

$$\bar{x}^* \in X \text{ и } F(\bar{x}^*) \geq F(\bar{x}) \text{ для всех } \bar{x} \in X \cap N_\varepsilon(\bar{x}^*), \quad (\text{IV.4})$$

где $N_\varepsilon \bar{x}^*$ — ε -окрестность вектора $\bar{x}^*$, в данном случае множество точек $\bar{x}$, удовлетворяющих условию

$$|\bar{x} - \bar{x}^*| = \sqrt{\sum_{j=1}^n (x_j - x_j^*)^2} < \varepsilon$$

для сколь угодно малого положительного числа ε .

Локальный максимум в $\bar{x}^*$ является строгим, если значение функции в точке $\bar{x}^*$ наибольшее, достигаемое функцией в некоторой малой окрестности:

$$F(\bar{x}^*) > F(\bar{x}) \text{ для всех } \bar{x} \in \bar{X} \cap N_\varepsilon(\bar{x}^*); \bar{x} \neq \bar{x}^*. \quad (\text{IV.5})$$

Очевидно, глобальный максимум является локальным; обратное утверждение неверно, так как могут существовать другие локальные максимумы, на которых функция принимает большие значения.

Все определения для минимума функции получаются заменой в выражениях (IV.1), (IV.2), (IV.5) знака $<$ на обратный $>$.

Теперь рассмотрим вопрос о существовании решения задачи (IV.1). Для этого рассмотрим некоторые варианты задания множества X . Естественно считать, что это множество непусто и не состоит из единственной точки. Только в этом случае задача максимизации (минимизации) содержательна, так как есть возможность выбора. Если множество X содержит конечное число точек, то решение задачи (IV.1) существует: можно перебрать все точки из X и выбрать среди них точку x^* , удовлетворяющую условию (IV.2). В случае, когда множество X содержит бесконечное число точек, задача максимизации (минимизации) $F(\bar{x})$ на X может не иметь решения.

Рассмотрим несколько примеров для E_1 .

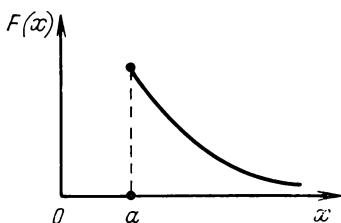


Рис. 26.

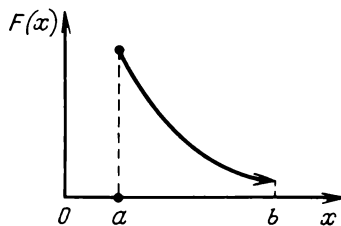


Рис. 27.

Пример IV.1. Пусть

$$X = \{x : x \geq a\}$$

(рис. 26), а функция $F(x)$ монотонно убывает (например, $F(x) = e^{-x}$). Очевидно, что здесь точки x^* , удовлетворяющей условию

$$F(x^*) \leq F(x)$$

не существует, так как допустимое множество не ограничено.

Пример IV.2. Пусть множество X задано в виде

$$X = \{x : a \leq x < b\},$$

т. е. не замкнуто: точка $x = b$ не содержится во множестве X (на рис. 27 это отмечено дужкой в точке b). Если функция $F(x)$ при $x \rightarrow b$ монотонно убывает, нижняя грань ее на множестве X не достигается (на рис. 27 это показано стрелкой) и точки x^* , удовлетворяющей условию

$$F(x^*) \leq F(x)$$

не существует.

Рассмотренные примеры свидетельствуют, что в случаях, когда множество X не замкнуто, задача может не иметь решения.

Сформулируем теперь теорему Вейерштрасса, выделяющую широкий класс задач максимизации (минимизации), заведомо имеющих решение.

Теорема Вейерштрасса (IV.1). Пусть допустимое множество X является компактным (т. е. ограниченным и замкнутым) и непустым. Тогда непрерывная функция $F[x]$, определенная на этом множестве, достигает глобального максимума на внутренней или граничной точке множества X .

Геометрическая иллюстрация теоремы Вейерштрасса в одномерном случае приведена на рис. 28. В этом случае вектор переменных сводится к действительному числу x . Допустимое множество X представляет собой интервал на горизонтальной оси x , а множество значений функции $F(x)$ — на вертикальной оси (включая граничные точки интервалов). В рассмотренных случаях внутреннее решение x^* является точкой глобального, но не строгого максимума, а граничное решение x^{**} — точкой строгого глобального максимума.

Однако надо отметить, что условия теоремы Вейерштрасса являются достаточными, но не необходимыми. Например, задача максимизации функции $F(x) = x^3$ при условии $0 < x \leq 1$ (или $x \leq 1$)

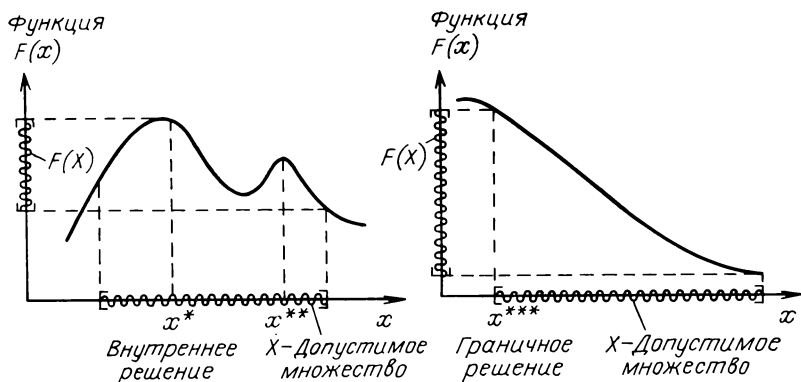


Рис. 28. Внутреннее и граничное решение в одномерном случае

имеет решение в точке $x = 1$, хотя допустимое множество не является компактным.

После того, как сформулирована задача максимизации функции и обсуждены вопросы, связанные с ее решением, необходимо указать признаки, с помощью которых можно найти точки x^* , являющиеся решением задачи (IV.1), или проверить, доставляет ли некоторая найденная точка $x \in X$ максимум функции $F(x)$. Эти признаки называются *необходимыми и достаточными условиями экстремума*.

2. Необходимые условия первого порядка

Выясним условия, которые должны выполняться в точках локального экстремума функции. Рассмотрим те случаи, когда множество X представляет вещественную ось.

Рассуждения сохраняют силу и для задач, в которых множество X не совпадает с E_1 , но открыто, т. е. состоит только из внутренних точек, либо экстремум достигается в его внутренней точке. Изучение случаев, когда экстремум реализуется на границе множества X , требует, как это будет описано ниже, *специальных методов*.

При выводе условий экстремума будем предполагать, что функция $F(x)$ имеет в окрестности исследуемой точки x^* непрерывные производные до второго порядка включительно.

Теорема (IV.2). *Для того, чтобы функция $F(x)$, определяемая на вещественной оси, имела безусловный локальный экстремум в точке x^* , необходимо, чтобы выполнялось условие*

$$\frac{dF}{dx}(x^*) = 0. \quad (\text{IV.6})$$

(Доказательство имеется во всех курсах анализа.) Точки, в которых выполняется условие (IV.6), называются *стационарными*, в них касательная горизонтальна.

Следует иметь в виду, что условие (IV.6) выделяет стационарные точки, но не определяет их характера: оно одинаково для точек максимума, минимума и перегиба. Чтобы избежать сравнения значений функции $F(x)$ во всех стационарных точках с целью определения глобального максимума (или локальных максимумов), желательно найти некоторые дополнительные условия, которые выполняются только в точках максимума (минимума).

3. Необходимые условия второго порядка

Пусть функция $F(x)$ удовлетворяет перечисленным выше условиям. Тогда справедлива теорема (IV.3). Для того, чтобы функция $F(x)$ имела в стационарной точке x^* безусловный локальный максимум (минимум), необходимо, чтобы ее вторая производная была неположительна (неотрицательна), т. е.

$$\frac{d^2F}{dx^2}(x^*) \leq 0 \quad \left(\frac{d^2F}{dx^2}(x^*) \geq 0 \right). \quad (\text{IV.7})$$

Условия (IV.6) и (IV.7) называются соответственно условиями первого и второго порядка, необходимыми для существования локального максимума (минимума) в x^* . Однако нетрудно показать, что совокупность условий (IV.6), (IV.7) не является достаточным условием максимума (минимума).

Пример IV.3. Пусть $F(x) = x^3$. Тогда точка $x^* = 0$ удовлетворяет необходимым условиям минимума второго порядка:

$$\begin{aligned} \frac{dF}{dx} \Big|_{x=0} &= 3x^2 \Big|_{x=0} = 0; \\ \frac{d^2F}{dx^2} \Big|_{x=0} &= 6x \Big|_{x=0} = 0. \end{aligned}$$

При этом она, однако, является точкой перегиба, а не минимума функции $F(x) = x^3$.

4. Достаточные условия

Если функция $F(x)$ дифференцируема достаточное число раз, то можно построить аналогичные формулам (IV.6), (IV.7) необходимые условия любого порядка. Налагая все более и более жесткие ограничения на выбор экстремальных точек, они тем не менее не дают окончательного ответа об их характере. Поэтому нужно иметь еще и достаточные условия экстремума.

Теорема (IV.4). Для того, чтобы функция $F(x)$ имела в стационарной точке x^* безусловный локальный максимум (минимум), достаточно, чтобы ее вторая производная была в точке x^* отрицательна (положительна):

$$\frac{d^2F}{dx^2}(x^*) < 0 \quad \left(\frac{d^2F}{dx^2}(x^*) > 0 \right). \quad (\text{IV.8})$$

Таким образом, достаточные условия наличия в x^* строгого локального максимума состоят в том, что первая производная

в этой точке обращается в нуль, а вторая производная строго меньше нуля. Иначе говоря, если выполнены условия

$$\left. \begin{aligned} \frac{dF}{dx}(x^*) &= 0; \\ \frac{d^2F}{dx^2}(x^*) &< 0, \end{aligned} \right\} \quad (IV.9)$$

то x^* есть точка строгого локального максимума:

$$F(x^*) > F(x^* + \Delta x).$$

На рис. 29 ход решения представлен в геометрической форме. Угловой коэффициент касательной к кривой $F(x)$ в точке x^* равен нулю, причем величина углового коэффициента убывает, так что точка x^* удовлетворяет условиям (IV.9) и, следовательно, является точкой строгого локального максимума. Эти условия выполнены и в точке x^{****} , которая также является точкой строгого локального максимума. В точках x^{**} и x^{***} выполняется условие первого порядка: угловой коэффициент касательной равен нулю, но условие второго порядка не выполнено, так как угловой коэффициент касательной в точке x^{**} возрастает, а в x^{***} остается постоянным. Точка x^{**} есть точка строгого локального минимума, а точка x^{***} , в которой первая и вторая производные обращаются в нуль, представляет собой один из случаев точки перегиба. Из примера с точкой x^{***} ясно, что хотя условие первого порядка (IV.6) и условие второго порядка (IV.7) являются необходимыми, однако выполнения одних только этих условий еще недостаточно для существования максимума.

Следовательно, полученные нами условия экстремума не определяют в этом случае характера стационарной точки. Очевидно, необходимо рассмотреть производные более высоких порядков.

Теорема (IV.5). Пусть функция $F(x)$, определенная на множестве $X = E_1$, имеет непрерывные производные до k -го порядка включительно, причем в некоторой точке x^* , т. е. $F'(x^*) = F''(x^*) = \dots = F^{k-1}(x^*) = 0$; $F^k(x^*) \neq 0$, тогда, если k — четное число, функция $F(x)$ имеет в точке x^* локальный максимум при $F^k(x^*) < 0$ и локальный минимум при $F^k(x^*) > 0$. Если k нечетное, то $F(x)$ не имеет в точке x^* ни максимума, ни минимума.

Эта теорема выражает достаточное условие общего вида.

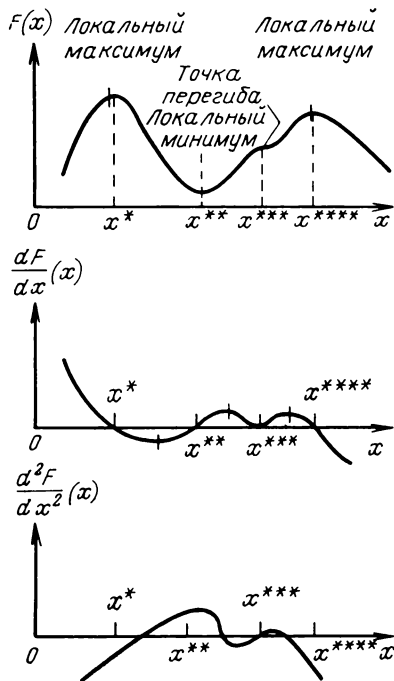


Рис. 29. Максимизация при отсутствии ограничений для случая одной независимой переменной

§ 23. Функция многих переменных

1. *Необходимые условия первого порядка.* Снова рассмотрим задачу безусловной максимизации, но будем считать, что $F(\bar{x})$ — скалярная функция векторного аргумента размерности n , т. е. $X = E_n$.

Т е о р е м а (IV.6). *Для того, чтобы в точке x^* функция $F(x_1, x_2, \dots, x_n)$ имела безусловный локальный экстремум, необходимо, чтобы все ее частные производные обращались в точке x^* в нуль:*

$$\left. \frac{\partial F}{\partial x^i} \right|_{x=x^*} = 0; \quad i = 1, 2, \dots, n. \quad (\text{IV.10})$$

Условие стационарности (IV.10) часто записывают в одной из следующих эквивалентных форм:

$$\text{grad } F(\bar{x}^*) = \nabla F(\bar{x}^*) = 0; \quad (\text{IV.11})$$

$$F'(\bar{x}^*) = 0. \quad (\text{IV.12})$$

Заметим, что необходимое условие экстремума (IV.10) эквивалентно нулю дифференциала функции $F(\bar{x})$ в точке x^*

$$dF(\bar{x}^*) = 0. \quad (\text{IV.13})$$

Условия (IV.10) образуют систему n уравнений для определения n компонент вектора x^* . Эти уравнения могут иметь различную природу и допускать любое число решений, в частности не иметь ни одного. Как и выше, точки $\bar{x}^*$, являющиеся решением системы уравнений (IV.10), называются стационарными, а условие (IV.10) — необходимым условием экстремума первого порядка.

2. *Необходимые условия второго порядка. Достаточные условия.* После того как решение $\bar{x}^*$ системы уравнений (IV.10) будет найдено, необходимо еще определить характер стационарной точки $\bar{x}^*$. Для этого нужно исследовать поведение функции $F(\bar{x})$ в окрестности стационарной точки $\bar{x}^*$. Воспользуемся разложением функции $F(\bar{x})$ в ряд Тейлора, предполагая ее дважды непрерывно дифференцируемой по всем переменным $x_1, x_2, \dots, x_n$. Ряд Тейлора для функции $y = F(\bar{x})$, где y — скаляр, $\bar{x}$ — вектор, при разложении относительно точки $\bar{x}^*$ с точностью до членов второго порядка малости может быть представлен

$$\begin{aligned} F(\bar{x}) = F(\bar{x}^*) + \sum_{j=1}^n \frac{\partial F}{\partial x_j}(\bar{x}^*) (x_j - x_j^*) + \\ + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 F}{\partial x_i \partial x_j}(\bar{x}^*) (x_i - x_i^*) (x_j - x_j^*). \end{aligned} \quad (\text{IV.14})$$

Так как $\bar{x}^*$ — стационарная точка, то согласно формуле (IV.12) в ней градиент равен нулевому вектору. Тогда в ε -окрестности по-

ведение функции $F(\bar{x})$ описывается следующим рядом Тейлора:

$$F(\bar{x}) = F(\bar{x}^*) + \frac{1}{2}(\bar{x} - \bar{x}^*)^T H_F(\bar{x}^*)(\bar{x} - \bar{x}^*). \quad (\text{IV.15})$$

Здесь и далее $(\bar{x} - \bar{x}^*)^T$ — вектор-столбец.

В выражении (IV.15) H_F обозначена матрица производных второго порядка:

$$\frac{\partial^2 F}{\partial \bar{x}^2}(\bar{x}) = \left\{ \begin{array}{cccc} \frac{\partial^2 F}{\partial x_1^2}(\bar{x}) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(\bar{x}) & \dots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(\bar{x}) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(\bar{x}) & \frac{\partial^2 F}{\partial x_2^2}(\bar{x}) & \dots & \frac{\partial^2 F}{\partial x_2 \partial x_n}(\bar{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1}(\bar{x}) & \frac{\partial^2 F}{\partial x_n \partial x_2}(\bar{x}) & \dots & \frac{\partial^2 F}{\partial x_n^2}(\bar{x}) \end{array} \right\}, \quad (\text{IV.16})$$

которая называется матрицей Гессе. Следовательно, характер стационарной точки $\bar{x}^*$ функции $F(\bar{x})$ связан со знакоопределенностью матрицы Гессе.

Напомним, что квадратичная форма $Q_A(\bar{x})$ называется неотрицательно определенной в точке $\bar{x}^*$, если

$$Q_A(\bar{x}) \geq 0, \quad (\text{IV.17})$$

и положительно определенной, если

$$Q_A(\bar{x}) > 0 \quad (\text{IV.18})$$

для любых векторов $\bar{x}$.

Соответственно симметричная матрица вторых производных $F''(\bar{x}^*)$ или матрица Гессе называется неотрицательно определенной в точке $\bar{x}^*$, если выполнено условие (IV.17), и положительно определенной, если выполнено условие (IV.18). Неположительно определенным и отрицательно определенным квадратичным формам и матрицам соответствуют противоположные знаки в неравенствах (IV.17), (IV.18). Таким образом с учетом разложения (IV.14) приходим к следующей формулировке условий второго порядка экстремальности функции $F(x_1, x_2, \dots, x_n)$.

Т е о р е м а (IV.7). Для того, чтобы дважды непрерывно дифференцируемая функция n переменных $F(\bar{x})$ имела в стационарной точке $\bar{x}^*$ безусловный локальный максимум (минимум), необходимо, чтобы матрица ее вторых производных была неположительно (неотрицательно) определенной, и достаточно, чтобы она была отрицательно (положительно) определенной.

Проверка знакоопределенности матриц может быть осуществлена, например, с помощью критерия Сильвестра. Согласно этому критерию необходимым и достаточным условием положительно определенной квадратичной матрицы $(\bar{x}, \bar{A}\bar{x})$, где $\bar{A} = \{a_{ij}\}$ — симметричная $n \times n$ матрица, является выполнение n неравенств:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \quad \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} > 0.$$

Необходимым и достаточным условием отрицательно определенной квадратичной формы $(\bar{x}, \bar{A}\bar{x})$ является выполнение цепочки следующих n неравенств:

$$(-1)^n a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \quad (-1)^n \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} > 0.$$

Пример IV.4. Исследуем на экстремум функцию

$$F(x_1, x_2, x_3) = x_1^3 + x_2^2 + x_3^2 + x_2 x_3 - 3x_1 + 6x_2 + 2. \quad (\text{IV.19})$$

Найдем сначала стационарные точки из решения системы (IV.10], которая в данном случае принимает вид

$$\frac{\partial F}{\partial x_1} = 3x_1^2 - 3 = 0;$$

$$\frac{\partial F}{\partial x_2} = 2x_2 + x_3 + 6 = 0;$$

$$\frac{\partial F}{\partial x_3} = 2x_3 + x_2 = 0.$$

Корни этой системы $x_1 = \pm 1$, $x_2 = -4$, $x_3 = 2$, т. е. имеем две стационарные точки:

$$x_1^* = \begin{pmatrix} -1 \\ -4 \\ 2 \end{pmatrix}; \quad x_2^* = \begin{pmatrix} +1 \\ -4 \\ 2 \end{pmatrix}.$$

Составим гессиан функции:

$$H_F = \left\| \frac{\partial^2 F}{\partial x_i \partial x_j} \right\| = \begin{pmatrix} 6x_1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Для первой стационарной точки x_1^* гессиан

$$H_F = \begin{pmatrix} -6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix},$$

согласно критерию Сильвестра, не является ни положительно, ни отрицательно определенной матрицей, так как в этом случае

$$D_1 = -6 < 0; \quad D_2 = \begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix} = -12 < 0;$$

$$D_3 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = -18 < 0.$$

Для второй стационарной точки x_2^* гессиан

$$H_F = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

является определенно положительной матрицей, так как

$$D_1 = 6 > 0, \quad D_2 = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} = 12 > 0; \quad D_3 = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = 18 > 0.$$

Поэтому стационарная точка x_2^* является точкой относительного минимума функции $F(x_1, x_2, x_3)$.

Итак, условия первого порядка для существования локального максимума функции двух переменных при отсутствии ограничений, т. е. когда требуется найти

$$\max_{x_1, x_2} F(x_1, x_2), \quad (\text{IV.20})$$

состоят в том, чтобы точка $\bar{x}^* = (x_1^*, x_2^*)^T$ была стационарной точкой:

$$\frac{\partial F}{\partial x_1}(\bar{x}^*) = 0; \quad \frac{\partial F}{\partial x_2}(\bar{x}^*) = 0. \quad (\text{IV.21})$$

Необходимые условия второго порядка в данном случае заключаются в том, что матрица Гессе отрицательно определена или отрицательно полуопределена, а это эквивалентно выполнению следующих условий, налагаемых на главные миноры матрицы:

$$\frac{\partial^2 F}{\partial x_1^2}(\bar{x}^*) \leq 0, \quad (\text{IV.22})$$

$$\begin{vmatrix} \frac{\partial^2 F}{\partial x_1^2}(\bar{x}^*) & \frac{\partial^2 F}{\partial x_1 \partial x_2}(\bar{x}^*) \\ \frac{\partial^2 F}{\partial x_2 \partial x_1}(\bar{x}^*) & \frac{\partial^2 F}{\partial x_2^2}(\bar{x}^*) \end{vmatrix} \geq 0. \quad (\text{IV.23})$$

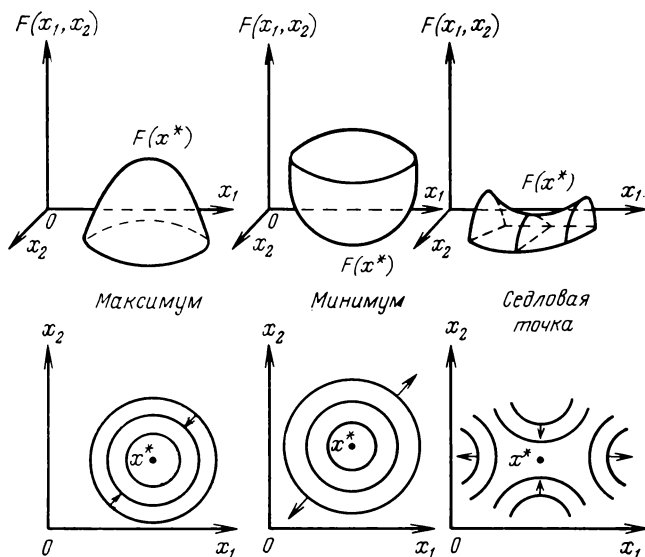


Рис. 30. Максимум, минимум и седловая точка в задаче оптимизации функции двух переменных при отсутствии ограничений

Завершая анализ данного случая, укажем, что если в стационарной точке

$$\frac{\partial^2 F}{\partial x_1^2}(\bar{x}^*) \geq 0 \quad (\text{IV.24})$$

и если выполнено условие (IV.23), то $\bar{x}^*$ будет точкой локального минимума, если же условие (IV.23) не выполнено, то $\bar{x}^*$ — седловая точка. Эти три варианта проиллюстрированы на рис. 30. В верхней части диаграммы непосредственно изображены все три типа стационарных точек, а в нижней части показаны линии уровня и направления скорейшего роста. Отметим, что в изображенной на рис. 30 седловой точке достигается минимум по переменной x_1 и максимум по переменной x_2 .

§ 24. Условный экстремум. Метод множителей Лагранжа

Условным экстремумом функции называется ее экстремум на множестве значений переменных, стесненном рядом ограничений, причем эти ограничения могут иметь вид равенств или неравенств. Для экономики более характерны ограничения вида неравенств. Тем не менее мы начинаем рассмотрение с задач, в которых ограничения имеют вид равенств, во-первых, потому что это

сведено с помощью необходимых условий к решению системы уравнений (IV.10), в результате чего определяются стационарные точки функции $F(\bar{x})$. Оказывается, что аналогичное сведение возможно и для задачи отыскания экстремумов функции $F(\bar{x})$ при наличии ограничений типа равенств

$$g_i(x) = 0; \quad i = 1, 2, \dots \quad (\text{IV.30})$$

Условия (IV.30) принято называть еще уравнениями связи.

Прежде всего уточним, что мы будем понимать под решением задачи на условный экстремум.

Допустимая точка $\bar{x}^*$ [точка, удовлетворяющая условиям (IV.30)] доставляет условный локальный максимум функции $F(\bar{x})$, если можно указать такое число $\varepsilon > 0$, при котором для всех $\bar{x}$, удовлетворяющих уравнениям (IV.30) и условию $\|\bar{x} - \bar{x}^*\| < \varepsilon$, имеет место неравенство

$$F(\bar{x}) \leq F(\bar{x}^*).$$

Определения строго условного локального максимума, а также все определения для условного минимума будут аналогичны приведенным в § 22.

Рассмотрим случай, когда уравнения ограничений (IV.30) могут быть разрешены относительно части переменных. Будем предполагать, что функции $g_i(x)$, $i = 1, \dots, m$ имеют в окрестности рассматриваемой допустимой точки $\bar{x}_j^*$ непрерывные частные производные по всем аргументам до второго порядка включительно и, кроме того, ранг матрицы Якоби для функций $g_i(x)$, $i = 1, \dots, m$, рассматриваемый в точке $\bar{x}_j^*$, равен m . Предположим, что отличен от нуля определитель (якобиан), составленный из частных производных по первым m аргументам, т. е.

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_m} \end{vmatrix} \neq 0. \quad (\text{IV.31})$$

Тогда, по теореме о неявных функциях, в некоторой окрестности точки $\bar{x}_j^*$ система уравнений (IV.30) разрешима относительно $x_1, \dots, x_m$, т. е. представима в виде

$$x_j = \varphi_j(x_{m+1}, \dots, x_n); \quad j = 1, 2, \dots, m, \quad (\text{IV.32})$$

где $\varphi_j(x_{m+1}, \dots, x_n)$ — непрерывно дифференцируемые в рассматриваемой окрестности функции. Переменные $x_{m+1}, \dots, x_n$ можно назвать «независимыми», в отличие от «зависимых» — $x_1, \dots, x_m$. Подставив выражение (IV.32) в функцию $F(\bar{x})$, полу-

чим задачу отыскания безусловного экстремума функции $n - m$ переменных:

$$F \varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), \\ (x_{m+1}, \dots, x_n) = F(x_{m+1}, \dots, x_n).$$

Однако провести исключение части компонент вектора $\bar{x}$ обычно бывает трудно и даже невозможно. Поэтому мы используем другой путь определения $\bar{x}^*$, который не предполагает наличия явных выражений типа (IV.32), хотя существенно использует условие (IV.31).

2. Метод множителей Лагранжа. Необходимые условия первого порядка. Этому методу мы придаем особое значение, поскольку он будет неоднократно использоваться в качестве основного подхода к решению почти всех видов задач оптимизации; кроме того, с его помощью можно получить ценную информацию о том, в какой степени оптимальное значение целевой функции чувствительно к изменениям констант ограничений. Последнее обстоятельство позволяет придавать множителям Лагранжа важный экономический смысл в задачах рациональной экономической деятельности.

Итак, ставится задача: найти

$$\max_{\bar{x}} F(\bar{x}) \quad (IV.33)$$

при условии

$$\bar{g}(\bar{x}) = \bar{b}. \quad (IV.34)$$

Предположим, что задача имеет локальное решение в $\bar{x}^*$ и что функции ограничений удовлетворяют условию Якоби, т. е. в точке $\bar{x}^*$ ранг ρ матрицы Якоби — матрицы частных производных функций ограничений — совпадает с числом строк матрицы

$$\rho \left(\frac{\partial \bar{g}}{\partial \bar{x}}(\bar{x}^*) \right) = \rho \left\{ \begin{array}{cccc} \frac{\partial g_1}{\partial x_1}(\bar{x}^*) & \frac{\partial g_1}{\partial x_2}(\bar{x}^*) & \dots & \frac{\partial g_1}{\partial x_n}(\bar{x}^*) \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_m}{\partial x_1}(\bar{x}^*) & \frac{\partial g_m}{\partial x_2}(\bar{x}^*) & \dots & \frac{\partial g_m}{\partial x_n}(\bar{x}^*) \end{array} \right\} = m. \quad (IV.35)$$

Отметим, что в приведенной выше задаче с одной степенью свободы матрица Якоби есть вектор-строка. Ранг этой матрицы равен единице тогда и только тогда, когда хотя бы одна из частных производных функций ограничений не обращается в нуль. При необходимости можно перенумеровать переменные таким образом, чтобы определитель матрицы, составленной из последних столбцов матрицы Якоби, не был бы равен нулю и чтобы вектор переменных можно было представить в виде $\bar{x} = (\bar{x}^1, \bar{x}^2)^T$, где вектор $\bar{x}^1$ состоит из $n - m$ переменных, а вектор $\bar{x}^2$ — из m переменных. Поскольку

условие Якоби выполнено, то, по теореме о неявной функции, в окрестности $\bar{x}^*$ можно разрешить систему ограничений, представив $\bar{x}^2$ в виде функции $\bar{x}^1$:

$$\bar{x}^2 = \bar{h}(\bar{x}^1), \quad (\text{IV.36})$$

где $\bar{h}$ — вектор-столбец, состоящий из m функций. Теперь задачу (IV.33) можно свести к задаче оптимизации при отсутствии ограничений

$$\max_{\bar{x}^1} \bar{H}(\bar{x}^1) = F(\bar{x}^1, \bar{h}(\bar{x}^1)). \quad (\text{IV.37})$$

Необходимое условие для существования локального максимума состоит в том, что

$$\frac{\partial \bar{H}}{\partial \bar{x}^1} = \frac{\partial F}{\partial \bar{x}^1} + \frac{\partial F}{\partial \bar{x}^2} \frac{\partial \bar{h}}{\partial \bar{x}^1} = 0, \quad (\text{IV.37, а})$$

где $\frac{\partial \bar{H}}{\partial \bar{x}^1}$ представляет собой вектор-строку, состоящую из $n - m$

элементов, а $\frac{\partial \bar{h}}{\partial \bar{x}^1}$ — матрицу размерности $m \times (n - m)$. Так как ограничения можно записать в виде тождества

$$\bar{g}(\bar{x}^1, \bar{h}(\bar{x}^1)) \equiv \bar{b},$$

то после дифференцирования получим

$$\frac{\partial \bar{g}}{\partial \bar{x}^1} + \frac{\partial \bar{g}}{\partial \bar{x}^2} \frac{\partial \bar{h}}{\partial \bar{x}^1} = 0. \quad (\text{IV.38})$$

Матрица $\frac{\partial \bar{g}}{\partial \bar{x}^2}$ размерности $m \times m$ невырожденная, так как выполняется условие Якоби. Следовательно,

$$\frac{\partial \bar{h}}{\partial \bar{x}^1} = - \left(\frac{\partial \bar{g}}{\partial \bar{x}^2} \right)^{-1} \left(\frac{\partial \bar{g}}{\partial \bar{x}^1} \right) \quad (\text{IV.39})$$

и условия (IV.37, а) можно записать следующим образом:

$$\frac{\partial F}{\partial \bar{x}^1} - \left(\frac{\partial F}{\partial \bar{x}^2} \right) \left(\frac{\partial \bar{g}}{\partial \bar{x}^2} \right)^{-1} \left(\frac{\partial \bar{g}}{\partial \bar{x}^1} \right) = 0. \quad (\text{IV.40})$$

Очевидно также, что

$$\frac{\partial F}{\partial \bar{x}^2} - \left(\frac{\partial F}{\partial \bar{x}^2} \right) \left(\frac{\partial \bar{g}}{\partial \bar{x}^2} \right)^{-1} \left(\frac{\partial \bar{g}}{\partial \bar{x}^2} \right) = 0. \quad (\text{IV.41})$$

Полагая теперь

$$\lambda = \left(\frac{\partial F}{\partial \bar{x}^2} \right) \left(\frac{\partial \bar{g}}{\partial \bar{x}^2} \right)^{-1} = (\lambda_1, \lambda_2, \dots, \lambda_m), \quad (\text{IV.42})$$

можно представить необходимые условия (IV.40) и (IV.41) в форме

$$\frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} = 0. \quad (\text{IV.43})$$

Переменные $\bar{\lambda}$ в формуле (IV.43) называются множителями Лагранжа.

Важно отметить, что эти необходимые условия наряду с исходными ограничениями можно получить путем дифференцирования по $\bar{x}$ и $\bar{\lambda}$ функции

$$L(\bar{x}, \bar{\lambda}) = F(\bar{x}) + \bar{\lambda}(\bar{b} - \bar{g}(\bar{x})), \quad (\text{IV.44})$$

которая называется функцией Лагранжа (лагранжианом) и представляет собой сумму целевой функции и скалярного произведения вектора множителей Лагранжа и вектора разности между постоянными ограничениями и функциями ограничений [29, 10].

Перечислим теперь основные этапы решения, по методу Лагранжа, классической задачи математического программирования, т. е. следующей задачи: найти $\max_{\bar{x}} F(\bar{x})$ при условии, что $\bar{g}(\bar{x}) = \bar{b}$.

1. Вводится вектор-строка из m новых переменных $\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)$ — множителей Лагранжа.

2. Определяется функция Лагранжа (IV.44)

$$L(\bar{x}, \bar{\lambda}) = F(\bar{x}) + \bar{\lambda}(\bar{b} - \bar{g}(\bar{x})), \quad (\text{IV.45})$$

или в развернутом виде

$$L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) = F(x_1, x_2, \dots, x_n) + \sum_{i=1}^m \lambda_i (b_i - g_i(x_1, x_2, \dots, x_n)). \quad (\text{IV.46})$$

3. Отыскивается точка $(\bar{x}^*, \bar{\lambda}^*)$, в которой все частные производные первого порядка функции Лагранжа обращаются в нуль:

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{x}}(\bar{x}^*, \bar{\lambda}^*) &= \frac{\partial F}{\partial \bar{x}}(\bar{x}^*) - \bar{\lambda}^* \frac{\partial \bar{g}}{\partial \bar{x}}(\bar{x}^*) = 0; \\ \frac{\partial L}{\partial \bar{\lambda}}(\bar{x}^*, \bar{\lambda}^*) &= \bar{b} - \bar{g}(\bar{x}^*) = 0. \end{aligned} \right\} \quad (\text{IV.47})$$

Первые n из полученных соотношений, совпадающие с условиями (IV.43), показывают, что градиент целевой функции должен равняться вектору множителей Лагранжа, умноженному на матрицу Якоби для функций ограничений, т. е.

$$\frac{\partial F}{\partial \bar{x}}(\bar{x}^*) = \bar{\lambda}^* \frac{\partial \bar{g}}{\partial \bar{x}}(\bar{x}^*), \quad (\text{IV.48})$$

или в развернутом виде

$$\frac{\partial F}{\partial x_j}(x_1^*, x_2^*, \dots, x_n^*) = \sum_{i=1}^m \lambda_i^* \frac{\partial g_i}{\partial x_j}(x_1^*, x_2^*, \dots, x_n^*);$$

$$j = 1, 2, \dots, n. \quad (\text{IV.49})$$

Остальные m условий представляют собой просто систему ограничений

$$\bar{g}(\bar{x}^*) = \bar{b}. \quad (\text{IV.50})$$

4. Решают совместно $m + n$ уравнений (IV.47), в результате чего получают значения $m + n$ неизвестных: переменных $\bar{x}^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ и множителей Лагранжа — $\bar{\lambda}^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_m^*)$.

Система уравнений (IV.47) представляет собой необходимые условия первого порядка в задаче на условный экстремум, а ее решение $x_1^*, x_2^*, \dots, x_n^*$ принято называть условно-стационарными точками. Как и в случае задач на безусловный экстремум, необходимые условия первого порядка не определяют характера условно-стационарной точки. Для выяснения этого вопроса следует привлечь производные более высоких порядков функций $\bar{F}(\bar{x})$ и $\bar{g}(\bar{x})$.

В заключение отметим, что если выполнены некоторые достаточные условия, указанные ниже, то значения переменных $\bar{x}^*$ дают локальное решение классической задачи математического программирования. Этот вывод достаточно очевиден, если учесть тот факт, что $\bar{x}^*$ удовлетворяет ограничениям и максимизирует функцию Лагранжа, которая в точке $(\bar{x}^*, \bar{\lambda}^*)$ просто совпадает со значением целевой функции:

$$L(\bar{x}^*, \bar{\lambda}^*) = F(\bar{x}^*). \quad (\text{IV.51})$$

3. *Необходимые условия второго порядка. Достаточные условия.* Необходимые условия экстремума второго порядка в задаче на условный экстремум и достаточные условия однозначно связаны со знакоопределенностью квадратичной формы

$$\frac{\partial^2 L}{\partial \bar{x}^2} = \left\{ \begin{array}{cccc} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \dots & \frac{\partial^2 L}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \dots & \dots & \frac{\partial^2 L}{\partial x_n^2} \end{array} \right\}, \quad (\text{IV.52})$$

т. е. матрицей Гессе, составленной из вторых частных производных лагранжиана по переменным, при условии, что

$$d\bar{g} = \frac{\partial \bar{g}}{\partial \bar{x}}(\bar{x}^*) d\bar{x} = 0, \quad (\text{IV.53})$$

так как смещения Δx из точки $\bar{x}^*$ не должны нарушать условий ограничений.

По аналогии с теоремой (IV.7) заключаем, что для того, чтобы условно-стационарная точка $\bar{x}^*$ реализовала локальный условный максимум, необходимо, чтобы квадратичная форма (IV.52) была неположительно определена, и достаточно, чтобы она была отрицательно определенной (соответственно неотрицательной и положительно определенной для минимума).

Пример IV.5. Пусть однородная продукция может производиться на предприятиях $\Pi_1, \Pi_2, \dots, \Pi_n$. Прибыль, получаемая j -м предприятием, представляется в виде следующей нелинейной функции:

$$F_j(x_j) = (a_j - k_j x_j) x_j, \quad (IV.54)$$

где x_j — количество единиц произведенной продукции, $a_j > 0$, $k_j > 0$ — постоянные коэффициенты.

Известны также производственные мощности M_j предприятий Π_j . Требуется распределить заказ по производству продукции в количестве N единиц между предприятиями так, чтобы получить наибольшую прибыль.

Математическая модель задачи будет следующей.

Требуется найти максимум целевой функции

$$F(\bar{x}) = \sum_{j=1}^n (a_j - k_j x_j) x_j \quad (IV.55)$$

при ограничениях

$$\sum_{j=1}^n x_j = N; \quad 0 \leq x_j \leq M_j. \quad (IV.56)$$

Конкретизируем задачу. Пусть $n = 5$, $N = 1000$. Параметры a_j , k_j и M_j заданы в табл. 5.

Т а б л и ц а 5

Параметры	Предприятия				
	Π_1	Π_2	Π_3	Π_4	Π_5
a_j	20	18	22	19	21
k_j	0,020	0,015	0,022	0,018	0,016
M_j	250	260	240	270	200

Определим тип целевой функции (IV.55). Вторая производная по x_j от каждого слагаемого этой функции

$$[(a_j - k_j x_j) x_j]'' = -2k_j < 0.$$

Следовательно, функция (IV.55) является строго вогнутой функцией, как сумма строго вогнутых функций, во всем n -мерном пространстве. Эта функция имеет единственный максимум на любом выпуклом множестве точек $\bar{x} = (x_1, x_2, \dots, x_n)$, в частности на выпуклом множестве, определенном линейными ограничениями (IV.56).

На первом этапе решим задачу (IV.55)—(IV.56) при отсутствии ограничений $0 \leq x_j \leq M_j$. Точка максимума $F(\bar{x})$ должна определяться по методу множителей Лагранжа как результат системы уравнений (IV.47), т. е.

$$\frac{\partial F(\bar{x})}{\partial x_j} = \bar{\lambda} \frac{\partial g(\bar{x})}{\partial x_j}; \quad \bar{g}(\bar{x}) = \bar{b}; \quad (j = 1, 2, \dots, n).$$

При условиях нашей задачи получаем

$$\left. \begin{aligned} a_j - 2k_j x_j &= \lambda; & (j = 1, 2, 3, 4, 5); \\ x_1 + x_2 + x_3 + x_4 + x_5 &= N. \end{aligned} \right\} \quad (\text{IV.57})$$

Находим

$$x_j = \frac{a_j - \lambda}{2k_j}, \quad (j = 1, 2, 3, 4, 5). \quad (\text{IV.58})$$

И, подставляя его в последнее уравнение системы (IV.57), получаем:

$$\sum_{j=1}^5 \frac{a_j}{2k_j} - \lambda \sum_{j=1}^5 \frac{1}{k_j} = 2N,$$

из последнего соотношения находим выражение множителя Лагранжа:

$$\lambda = \frac{\sum_{j=1}^5 a_j / k_j - 2N}{\sum_{j=1}^5 1/k_j}. \quad (\text{IV.59})$$

Затем, подставляя значение λ (IV.59) в выражение (IV.58), получаем все значения x_j , при которых функция (IV.55) имеет максимум.

После вычислений по формулам (IV.59) и (IV.58) получаем $\lambda = 12,734$; $x_1 = 181,6$; $x_2 = 175,5$; $x_3 = 210,6$; $x_4 = 174,1$; $x_5 = 258,3$.

Все значения x_j получились положительными, но значение $x_5 = 258,3$ получилось больше мощности предприятия $\Pi_5 M_5 = 200$. Зафиксируем значение $x_5 = 200$ и найдем максимум функции (IV.55) при условии $x_1 + x_2 + x_3 + x_4 = 800$ точно так же, как и в первом этапе, по формулам (IV.59) и (IV.58) при $N = 800$ и $n = 4$.

В результате расчета получаем на втором этапе: $\lambda = 12,198$; $x_1 = 195,0$; $x_2 = 193,4$; $x_3 = 222,8$; $x_4 = 188,8$.

Полученные значения неизвестных удовлетворяют условиям задачи (IV.55), (IV.56). Следовательно, мы имеем оптимальное распределение (с округлением до единицы) заказа между предприятиями: $x_1 = 195$; $x_2 = 193$; $x_3 = 223$; $x_4 = 189$; $x_5 = 200$. При этом максимальный доход составляет $\max F(x) = 16\,375$.

4. *Геометрическая интерпретация $m + n$ условий первого порядка* (IV.47). В системе (IV.47)

$$\frac{\partial L}{\partial \bar{\lambda}} (\bar{x}^* \bar{\lambda}^*) = \bar{b} - \bar{g}(\bar{x}^*) = 0$$

есть не что иное, как несколько иначе записанные ограничения (IV.34) исходной задачи; они означают требование, что искомое решение $\bar{x}^*$ должно принадлежать допустимому множеству « x »-сов.

Обратимся теперь к первому уравнению системы (IV.47). Если обозначить градиент функции $F(\bar{x})$ как $\text{grad } F(\bar{x})$, то уравнения системы (IV.47) можно переписать в векторном виде так

$$\text{grad } F(\bar{x}) = \bar{\lambda}^* \text{grad } \bar{g}(\bar{x}^*). \quad (\text{IV.60})$$

Известно, что градиент функции есть нормаль к гиперповерхности равного уровня этой функции, то последнее векторное соотношение (IV.60) можно интерпретировать следующим образом. Геометрически система уравнений (IV.47) означает требование, что точка $\bar{x}^*$ допустима и в этой точке градиент целевой функции является линейной комбинацией нормалей к гиперповерхностям $\bar{g}(\bar{x}) = \bar{b}$.

Можно показать, что это требование эквивалентно требованию, что касательная гиперплоскость в точке $\bar{x}^*$ к поверхности равного уровня функции $F(\bar{x})$ должна быть одновременно касательной гиперплоскостью к допустимому множеству, т. е. множеству «х»-сов, удовлетворяющих одновременно всем условиям (IV.34). Это последнее требование — простое следствие того факта, что в точке $\bar{x}^*$ условного экстремума гиперповерхность $F(\bar{x}) = F(\bar{x}^*)$ не должна пересекать допустимое множество, а может только его касаться.

Допустим, что значение функции $F(\bar{x})$ увеличивается в направлении, обозначенном на рис. 31 стрелкой, тогда точка $\bar{x}^*$ есть точка максимума $F(\bar{x})$ на допустимом множестве.

Для того, чтобы точка $\bar{x}^*$ была точкой максимума, линия $F(\bar{x}) = F(\bar{x}^*)$ равного уровня функции $F(\bar{x})$ не должна пересекать допустимое множество «х»-сов, лежащее на кривой $\bar{g}(\bar{x}) = \bar{b}$, так что вся кривая $\bar{g}(\bar{x}) = \bar{b}$ должна лежать по одну сторону от кривой $F(\bar{x}) = F(\bar{x}^*)$.

Заметим, что в случае, изображенном на рис. 31, совпадение касательных к кривой $\bar{g}(\bar{x}) = \bar{b}$ и к одной из линий равного уровня функции $F(\bar{x})$ в общей для этих двух кривых точке имеет место не только в точке $\bar{x}^*$, но и в точках $\bar{x}^{**}$, $\bar{x}^{***}$; первая из этих точек соответствует условному минимуму функции $F(\bar{x})$ при условии $\bar{g}(\bar{x}) = \bar{b}$, а вторая точка вообще не является условным экстремумом функции $F(\bar{x})$. Вместе с тем точки $\bar{x}^{**}$ и $\bar{x}^{***}$ будут решением системы уравнений (IV.43), (IV.50) наряду с точкой $\bar{x}^*$. Этот пример, таким образом, иллюстрирует тот факт, что уравнения (IV.43), (IV.50) выражают лишь необходимые условия условного экстремума.

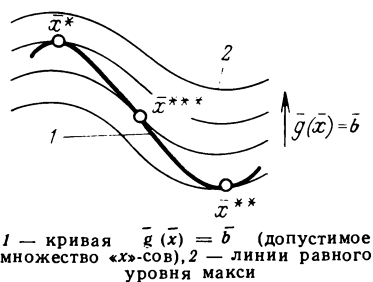


Рис. 31

§ 25. Интерпретация множителей Лагранжа

Решение системы уравнений (IV.47), выражающих условия первого порядка, содержат кроме вектора локального оптимума $\bar{x}^*$ еще и вектор множителей Лагранжа $\bar{\lambda}^*$. Если выполнено условие Якоби,

то существует единственный вектор $\bar{\lambda}^*$, соответствующий локальному решению $\bar{x}^*$. Знание значений множителей Лагранжа не является излишним; напротив, с их помощью можно получить ценную информацию о задаче. Широкое практическое использование метода множителей Лагранжа во многом объясняется именно последним обстоятельством. Множители Лагранжа, соответствующие решению задачи, измеряют чувствительность оптимального значения целевой функции $F^* = F(\bar{x}^*)$ к изменениям констант ограничений $\bar{b}$:

$$\bar{\lambda}^* = \frac{\partial F^*}{\partial \bar{b}}, \quad (\text{IV.61})$$

т. е.

$$\lambda_i^* = \frac{\partial F^*}{\partial b_i}; \quad i = 1, 2, \dots, m. \quad (\text{IV.62})$$

Для доказательства соотношений (IV.61) необходимо сначала показать, что если величины b_i рассматриваются как переменные, то переменные x_j и λ_i ($j = 1, \dots, n$; $i = 1, \dots, m$) можно представить в виде функций b_i . Рассмотрим с этой целью условия первого порядка (IV.47), которые можно записать в виде системы $m + n$ уравнений с $2m + n$ неизвестными $(\bar{b}, \bar{\lambda}, \bar{x})$, если считать b_i ($i = 1, 2, \dots, m$) переменными величинами:

$$\left. \begin{aligned} \psi^1(\bar{b}, \bar{\lambda}, \bar{x}) &\equiv \bar{b} - \bar{g}(\bar{x}) = 0; \\ \psi^2(\bar{b}, \bar{\lambda}, \bar{x}) &\equiv \frac{\partial F}{\partial \bar{x}}(\bar{x}) - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}}(\bar{x}) = 0. \end{aligned} \right\} \quad (\text{IV.63})$$

Запишем матрицу Якоби этой системы уравнений в виде

$$\begin{vmatrix} I & 0 & -\left(\frac{\partial \bar{g}}{\partial \bar{x}}\right) \\ 0 & -\left(\frac{\partial \bar{g}}{\partial \bar{x}}\right)^T & \frac{\partial^2 L}{\partial \bar{x}^2} \end{vmatrix}, \quad (\text{IV.64})$$

где I — единичная матрица порядка m . Следовательно, по теореме о неявной функции, решая систему уравнений, составленную из $m + n$ условий первого порядка, можно представить переменные и множители Лагранжа в виде функций постоянных ограничений $\bar{b}$:

$$\left. \begin{aligned} \bar{\lambda} &= \bar{\lambda}(\bar{b}); \\ \bar{x} &= \bar{x}(\bar{b}). \end{aligned} \right\} \quad (\text{IV.65})$$

Рассмотрим теперь функцию Лагранжа как функцию, зависящую от постоянных ограничений:

$$L(\bar{b}) = F(\bar{x}(\bar{b})) + \bar{\lambda}(\bar{b})[\bar{b} - \bar{g}(\bar{x}(\bar{b}))]. \quad (\text{IV.66})$$

Дифференцирование по $\bar{b}$
даёт

$$\begin{aligned}\frac{\partial L}{\partial \bar{b}} &= \frac{\partial F}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \bar{b}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \bar{b}} + \\ &+ (\bar{b} - g(\bar{x}))^T \frac{\partial \bar{\lambda}^T}{\partial \bar{b}} + \bar{\lambda} = \\ &= \left(\frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} \right) \left(\frac{\partial \bar{x}}{\partial \bar{b}} \right) +\end{aligned}$$

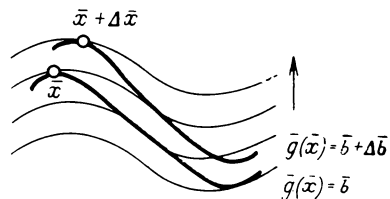


Рис. 32

$$+ (\bar{b} - \bar{g}(\bar{x}))^T \frac{\partial \bar{\lambda}^T}{\partial \bar{b}} + \bar{\lambda}. \quad (\text{IV.67})$$

Первые два члена этого выражения в точке $(\bar{x}^*, \bar{\lambda}^*)$ обращаются в нуль из-за условий первого порядка (IV.47), так что dL/db равно вектору множителей Лагранжа. Однако значение функции Лагранжа в точке $(\bar{x}^*, \bar{\lambda}^*)$ есть оптимальное значение целевой функции (IV.51). Следовательно,

$$\frac{\partial L}{\partial \bar{b}}(\bar{x}^*, \bar{\lambda}^*) = \frac{\partial F^*}{\partial \bar{b}} = \bar{\lambda}^*, \quad (\text{IV.68})$$

что и утверждалось ранее. Таким образом, помимо того, что метод множителей Лагранжа даёт решение классической задачи максимизации, он позволяет также проанализировать, насколько оптимальное значение целевой функции чувствительно к изменениям констант ограничений. Например, если какой-то множитель Лагранжа равен нулю, то малые изменения соответствующей константы ограничений не окажут никакого влияния на оптимальное значение целевой функции. Следовательно, множитель Лагранжа может служить показателем ценности соответствующего ресурса для экономической системы, т. е. показателем того, насколько в первом приближении возрастет (или упадет) значение целевой функции, если j -й ресурс или j -е задание возрастет на единицу.

Так, изменение j -го ограничения на величину Δb_j приводит к изменению допустимого множества « $\bar{x}$ »-сов [поскольку меняется одно из уравнений в условиях (IV.34)] и, следовательно, к новому решению $\bar{x} + \Delta \bar{x}$ задачи (IV.33), (IV.34); в результате изменится значение максимизируемой функции $F(\bar{x})$ (рис. 32). Показателем ценности j -го ресурса с точки зрения максимизации целевой функции $F(\bar{x})$ естественно считать величину $\frac{\partial L(\bar{b})}{\partial b_j}$, поскольку она показывает, как изменится оптимальное значение целевой функции при изменении j -го ограничения.

§ 26. Седловая точка функции Лагранжа. Минимаксная трактовка задачи на условный экстремум функции

Пусть функция $F(\bar{x})$ достигает в точке $\bar{x}^*$ относительного минимума при условии $\bar{g}(\bar{x}) = \bar{b}$.

Вектор множителей Лагранжа, соответствующий $\bar{x}^*$, обозначим $\bar{\lambda}^*$. В соответствии с необходимыми условиями величины $\bar{x}^*$, $\bar{\lambda}^*$ являются решениями системы уравнений

$$\begin{aligned}\frac{\partial L(\bar{x}, \bar{\lambda})}{\partial \bar{x}} &= \frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \left(\frac{\partial \bar{g}}{\partial \bar{x}} \right) = 0; \\ \frac{\partial L(\bar{x}, \bar{\lambda})}{\partial \bar{\lambda}} &= \bar{b} - \bar{g}(\bar{x}) = 0,\end{aligned}$$

где

$$L(\bar{x}, \bar{\lambda}) = F(\bar{x}) + \bar{\lambda}(\bar{b} - \bar{g}(\bar{x})).$$

Предположим теперь, что в ε -окрестности точки $\bar{\lambda}^*$ функция $L(\bar{x}, \bar{\lambda})$ имеет единственный для каждого $\bar{\lambda}^*$ безусловный относительный минимум по $\bar{x}$ (это предположение часто выполняется). Необходимым условием такого безусловного относительного минимума является, очевидно, выполнение первого уравнения (IV.47) для любых $\bar{\lambda}$ из ε -окрестности. Обозначим решение этого уравнения системы (IV.47) относительно $\bar{x}$ как $\hat{\bar{x}}$. Очевидно,

$$\hat{\bar{x}} = \hat{\bar{x}}(\bar{\lambda}); \quad \min_{\bar{x}} L(\bar{x}, \bar{\lambda}) = L(\hat{\bar{x}}, \bar{\lambda}) = h(\bar{\lambda}). \quad (\text{IV.69})$$

При $\bar{\lambda} = \bar{\lambda}^*$ имеем

$$\hat{\bar{x}} = \bar{x}^*; \quad h(\bar{\lambda}^*) = L(\bar{x}^*, \bar{\lambda}^*) = F(\bar{x}^*). \quad (\text{IV.70})$$

Так как $h(\bar{\lambda})$ безусловный минимум $L(\bar{x}, \bar{\lambda})$ по $\bar{x}$, а $h(\bar{\lambda}^*)$ — условный, то, очевидно, имеет место неравенство

$$h(\bar{\lambda}) \leq h(\bar{\lambda}^*), \quad (\text{IV.71})$$

т. е. $h(\bar{\lambda})$ имеет в точке $\bar{\lambda}^*$ относительный максимум

$$h(\bar{\lambda}^*) = \max_{\bar{\lambda}} h(\bar{\lambda}). \quad (\text{IV.72})$$

В соответствии с условиями (IV.70), (IV.69), (IV.72) получаем

$$h(\bar{\lambda}^*) = F(\bar{x}^*) = L(\bar{x}^*, \bar{\lambda}^*) = \max_{\bar{\lambda}} \min_{\bar{x}} L(\bar{x}, \bar{\lambda}). \quad (\text{IV.73})$$

Таким образом, решение задачи на условный экстремум функции $F(\bar{x})$ при ограничении $\bar{g}(\bar{x}) = \bar{b}$ свелось к решению задачи на минимакс функции Лагранжа $L(\bar{x}, \bar{\lambda})$.

Соотношение (IV.73) можно представить иначе. Так как, с одной стороны, учитывая выражения (IV.69)

$$F(\bar{x}^*) = L(\bar{x}^*, \bar{\lambda}^*) \leq F(\bar{x}, \bar{\lambda}^*), \quad (\text{IV.74})$$

а с другой —

$$L(\bar{x}^*, \bar{\lambda}) = F(\bar{x}^*), \quad (\text{IV.75})$$

имеем

$$L(\bar{x}^*, \bar{\lambda}) = L(\bar{x}^*, \bar{\lambda}^*) \leq L(\bar{x}, \bar{\lambda}^*). \quad (\text{IV.76})$$

Говорят, что функция $L(\bar{x}, \bar{\lambda})$ имеет седловую точку в точке $\bar{x}^*, \bar{\lambda}^*$, если для любых $\bar{x}$ и $\bar{\lambda}$ из ε -окрестности

$$L(\bar{x}^*, \bar{\lambda}) = L(\bar{x}^*, \bar{\lambda}^*) \leq L(\bar{x}, \bar{\lambda}^*). \quad (\text{IV.77})$$

При сделанных выше предположениях функция Лагранжа в точке $\bar{x}^*, \bar{\lambda}^*$ имеет, таким образом, седловую точку, которая является вырожденной, так как слева имеет место строгое равенство.

Пример IV.6. Рассмотрим задачу минимизации функции при условии

$$\begin{aligned} z = F(x_1, x_2) &= x_1^2 + x_2^2; \\ x_1 + x_2 &= 1. \end{aligned}$$

Составляя функцию Лагранжа

$$L(x_1, x_2, \lambda) = x_1^2 + x_2^2 + \lambda(1 - x_1 - x_2)$$

и приравнивая ее частные производные к нулю, получаем систему уравнений для определения x_1^*, x_2^*, λ^* :

$$\frac{\partial L}{\partial x_1} = 2x_1 - \lambda = 0;$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - \lambda = 0;$$

$$\frac{\partial L}{\partial \lambda} = 1 - x_2 - x_1 = 0.$$

Отсюда получаем $\lambda^* = 1$, $x_1^* = 1/2$, $x_2^* = 1/2$.

Это же решение можно получить и из условия $\max \min L(\bar{x}, \bar{\lambda})$.

Действительно, минимизируя $L(\bar{x}, \bar{\lambda})$, находим из условий:

$$\frac{\partial L}{\partial x_1} = 0; \quad \frac{\partial L}{\partial x_2} = 0;$$

$$\hat{x}_1 = \frac{\lambda}{2}; \quad \hat{x}_2 = \frac{\lambda}{2}.$$

Следовательно, минимальное по x значение $L(x, \lambda)$ будет

$$h(\lambda) = \min_{\bar{x}} F(\bar{x}, \bar{\lambda}) = \frac{\lambda^2}{2} + \lambda(1 - \lambda) = \lambda - \frac{\lambda^2}{2}.$$

Максимизируя $h(\lambda)$ по λ , находим из условия:

$$\frac{\partial h}{\partial \lambda} = 1 - \lambda = 0; \quad \lambda^* = 1.$$

Поэтому

$$x_1^* = \hat{x}_1(\lambda^*) = \frac{1}{2}; \quad x_2^* = \hat{x}_2(\lambda^*) = \frac{1}{2}.$$

Итак, в данном случае

$$L(\bar{x}^*, \bar{\lambda}^*) = \max_{\bar{\lambda}} \min_{\bar{x}} L(\bar{x}, \bar{\lambda}).$$

В предыдущих параграфах были выведены необходимые и достаточные условия, которым должна удовлетворять функция в экстремальных точках. При этом были рассмотрены те случаи, когда выбор значений аргумента $\bar{x}$ либо вообще не стеснен никакими ограничениями, либо подчиняется ограничениям типа равенств. Как уже неоднократно подчеркивалось, полученные необходимые условия экстремума не только позволяют проверить, доставляет ли некоторая точка максимум или минимум изучаемой функции, но и служат кроме того, инструментом для эффективного отыскания экстремумов.

Глава V

НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

§ 27. Общие понятия

Задача нелинейного программирования — это задача выбора таких неотрицательных значений некоторых переменных, подчиненных системе ограничений в форме неравенств, при которых достигается максимум или минимум данной функции. В обозначениях, принятых в гл. IV, задача нелинейного программирования на максимум состоит в следующем. Требуется найти

$$\max_{\bar{x}} F(\bar{x}), \quad (V.1)$$

при условии

$$\bar{g}(\bar{x}) \leq \bar{b}; \quad \bar{x} \geq 0 \quad (V.2)$$

или в развернутом виде найти

$$\max_{x_1, x_2, \dots, x_n} F(x_1, x_2, \dots, x_n) \quad (V.3)$$

при условии

$$\left. \begin{aligned} g_1(x_1, x_2, \dots, x_n) &\leq b_1; & x_1 &\geq 0; \\ g_2(x_1, x_2, \dots, x_n) &\leq b_2; & x_2 &\geq 0; \\ \dots &\dots & \dots & \dots \\ g_m(x_1, x_2, \dots, x_n) &= b_m; & x_n &\geq 0. \end{aligned} \right\} \quad (V.4)$$

Величины $x_1, x_2, x_3, \dots, x_n$ — составляющие n -мерного вектор-столбца $\bar{x}$ представляют собой переменные. Функция $F(\cdot)$ — это целевая функция, а функция ограничений $g_1(\cdot), g_2(\cdot), \dots, g_m(\cdot)$ — составляющие m -мерного вектор-столбца $\bar{g}(\cdot)$. Вектор $\bar{b}$ содержит константы ограничений $b_1, b_2, \dots, b_m$. Предположим, что m и n — конечные числа; заданы $m+1$ непрерывно дифференцируемых и не содержащих случайных элементов функций $F(\bar{x}), \bar{g}(\bar{x})$. Вектор $\bar{b}$ состоит из заданных вещественных чисел; $\bar{x}$ — это вектор с вещественными компонентами, удовлетворяющими $m+n$ ограничениям (V.1).

Отметим, что задачи, в которых $F(\bar{x}), \bar{g}(\bar{x})$ или $\bar{b}$ включают случайные элементы, относятся к классу задач стохастического программирования.

Раздел нелинейного программирования, который называется выпуклым программированием, представляет собой особый вид нелинейных экстремальных задач с выпуклыми

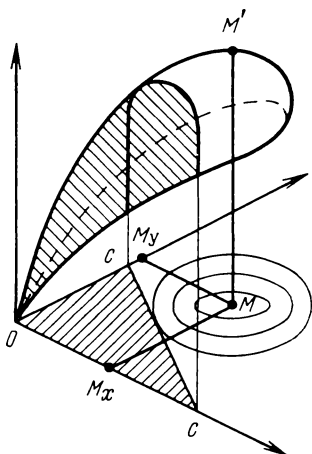


Рис. 33. Задачи вогнутого программирования

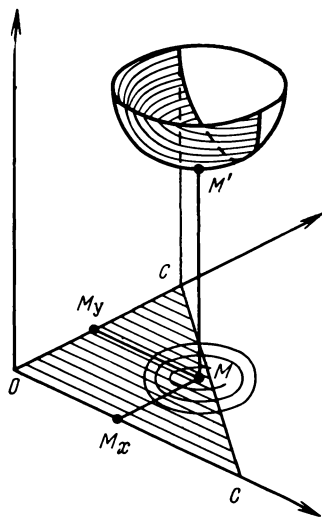


Рис. 34. Задачи выпуклого программирования

функциями. Выпуклым этот вид нелинейного программирования называется потому, что он имеет дело с выпуклыми целевыми функциями (они минимизируются) и выпуклыми системами ограничений.

Общая задача выпуклого программирования состоит в отыскании такого вектора $\bar{x}$ (т. е. такой точки выпуклого допустимого множества), который доставляет минимум выпуклой функции $F(\bar{x})$ или максимум вогнутой функции $F(\bar{x})$ (рис. 33, 34). Для второго случая (выпуклая область допустимых значений и максимум вогнутой функции) ряд авторов предпочитают термин «вогнутое программирование». Выпуклость (вогнутость) важна тем, что она гарантирует нахождение оптимального решения задачи, так как здесь соответственно локальные и глобальный экстремумы обязательно совпадают. Это положение основывается на второй основной теореме математического программирования (первая теорема — теорема Вейерштрасса), называемой «локально-глобальной» теоремой, дающей достаточные условия, при которых локальный экстремум является глобальным.

Поясним сказанное на примере.

Пример V. 1. Найти глобальный максимум функции

$$F = 2x_1 - x_1^2 + x_2$$

на множестве решений системы ограничений

$$x_2 \leq 3;$$

$$3x_1 + 2x_2 \leq 12;$$

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

Решение. Линиями уровня данной функции F являются параболы (рис. 35). Функцию F можно рассматривать как сумму двух вогнутых функций $f_1(x_1) = 2x_1 - x_1^2$ и $f_2(x_2) = x_2$. Поэтому F — вогнутая функция. Следовательно, локальный максимум функции F будет глобальным. $F_{\max} = 4$ и достигается в точке $D(1, 3)$.

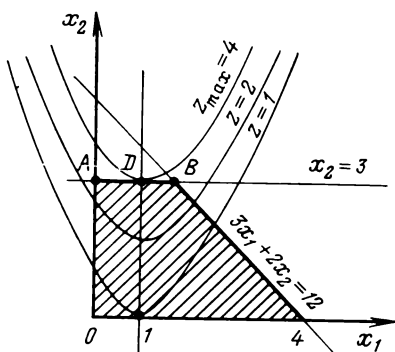


Рис. 35. Геометрическая интерпретация «локально-глобальной» теоремы

В постановках задач нелинейного программирования встречаются следующие случаи.

1. Задачи нелинейного программирования с линейной целевой функцией и нелинейной системой ограничений.

Пример V.2. На множестве решений системы неравенств

$$\begin{aligned} x^2 + y^2 &\leq 36; \\ x &\geq 0; \\ y &\geq 0 \end{aligned}$$

найти глобальные экстремумы функции

$$F = 2x + y.$$

Решение. На рис. 36, а множество допустимых решений заштриховано. Это множество выпукло. Линиями уровня функции $F = 2x + y$ являются параллельные прямые с угловым коэффициентом $K = -2$. Очевидно, что глобальный минимум достигается в точке $O(0; 0)$, а глобальный максимум — в точке A касания прямой уровня и окружности $x^2 + y^2 = 36$. Найдем координаты точки A . Для этого достаточно составить уравнение прямой l и решить систему, состоящую из уравнения прямой и уравнения окружности. Заметим, что прямая l перпендикулярна к линии уровня, а, следовательно, ее угловой коэффициент K_1 равен $\frac{1}{2}$ ($K_1 K = -1$). Прямая l проходит

через точку O и имеет угловой коэффициент $K_1 = \frac{1}{2}$. Поэтому ее уравнение

таково: $y = \frac{1}{2}x$. Решая систему

$$\begin{aligned} x^2 + y^2 &= 36; \\ y &= \frac{1}{2}x, \end{aligned}$$

получаем:

$$x = \frac{12\sqrt{5}}{5}; \quad y = \frac{6\sqrt{5}}{5}.$$

Итак, глобальный минимум, равный 0, достигается в точке $O(0; 0)$, а глобальный максимум, равный $6\sqrt{5}$, — в точке $A(\frac{12\sqrt{5}}{5}, \frac{6\sqrt{5}}{5})$. Локальных экстремумов, отличных от глобальных, функция не достигает.

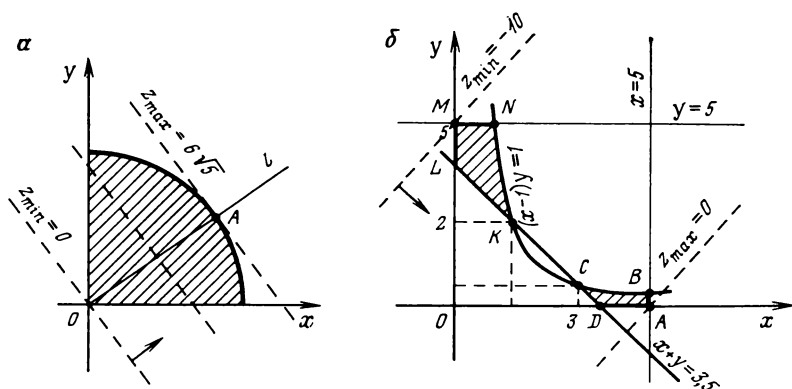


Рис. 36. Задачи нелинейного программирования с линейной целевой функцией и нелинейной системой ограничений:

а — пример V.2; б — пример V.3.

Пример V.3. Найти глобальные экстремумы функции

$$F = x - y - 5$$

на множество решений системы неравенств

$$(x - 1)y \leq 1;$$

$$x + y \geq 3,5,$$

$$0 \leq x \leq 5;$$

$$0 \leq y \leq 5.$$

Решение. Множество допустимых решений состоит из двух отдельных частей, каждая из которых выпукла (рис. 36, б). Вычислим значение целевой функции $F = x - y - 5$ в точках A, B, C, D, K, N, M, L , предварительно получив их координаты: $A(5; 0)$; $B(5; 1/4)$; $C(3; 1/2)$; $D(3,5; 0)$; $K(1 1/2; 2)$; $L(0; 3,5)$; $N(1 1/5; 5)$; $M(0; 5)$. $F_B = -1/4$; $F_C = -2,5$; $F_D = -1,5$; $F_K = -5,5$; $F_L = -8,5$; $F_N = -8,8$; $F_A = 0$; $F_M = -10$.

Итак, глобальный максимум достигается в точке $A(5; 0)$ и равен 0, а глобальный минимум — в точке $M(0; 5)$ и равен -10 .

Нетрудно видеть, что в точке C функция достигает локального минимума, равного $-2,5$, который отличен от глобального (значение целевой функции в точке C меньше, чем значение ее в соседних вершинах B и D). Аналогично в точке K достигается локальный максимум, отличный от глобального.

2. Задачи нелинейного программирования с линейной системой ограничений, но нелинейной целевой функцией. Множество допустимых решений таких задач всегда выпукло, так как линейные ограничения образуют выпуклый многогранник в n -мерном пространстве. Однако при нелинейной целевой функции оптимальное решение не обязательно находится в вершине этого многогранника.

Пример V.4. Определить наибольшее значение функции

$$F = \sqrt{x^2 + y^2}$$

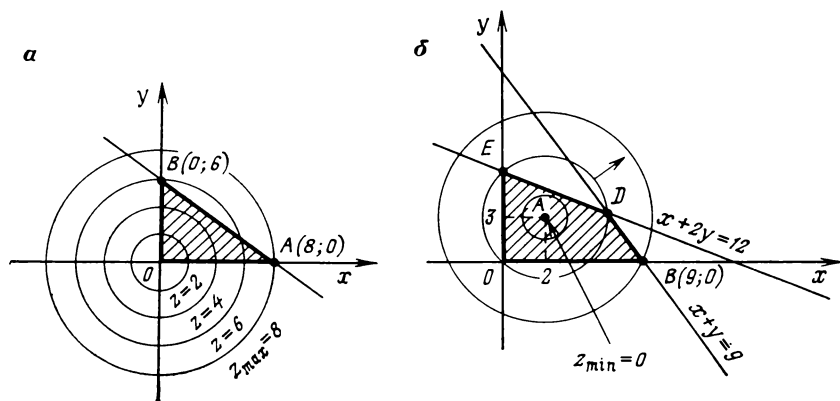


Рис. 37. Задачи нелинейного программирования с линейной системой ограничений, но нелинейной целевой функцией:

a — пример V.4; *б* — пример V.5

при условии

$$\left. \begin{aligned} 3x + 4y - 24 &\leq 0; \\ x &\geq 0; \\ y &\geq 0. \end{aligned} \right\}$$

Решение. Множество допустимых решений заштриховано на рис. 37, *a*. Если целевой функции придавать фиксированные значения c , то будем получать окружности с центром в начале координат и радиусом c^2 . Пусть $c = 1, 2, \dots$. Начертим ряд окружностей (линии уровня целевой функции). Из рис. 37, *a* видно, что функция $F = \sqrt{x^2 + y^2}$ достигает наибольшего значения, равного 8, в точке $A(8; 0)$. $F_{\max} = 8$.

Пример V.5. Найти глобальные экстремумы функции

$$F = (x - 2)^2 + (y - 3)^2$$

на множестве решений системы

$$\left. \begin{aligned} x + 2y &\leq 12; \\ x + y &\leq 9; \\ x &\geq 0; \\ y &\geq 0. \end{aligned} \right\}$$

Решение. Построим многоугольник допустимых планов и несколько линий уровня (рис. 37, *б*). Линии уровня $F = c$ представляют собой окружности с центром в точке $A(2; 3)$ и радиусом $r = \sqrt{c}$. Из рис. 37, *б* видно, что $F_{\min} = 0$ достигается в точке $A(2; 3)$, а $F_{\max}$ — в точке $B(9; 0)$. Таким образом, получаем, что $F_{\min} = 0$, $F_{\max} = (9-2)^2 + (0-3)^2 = 58$.

3. Задачи нелинейного программирования с нелинейной целевой функцией и нелинейной системой ограничений.

Пример V.6. На множестве решений системы ограничений

$$\left. \begin{aligned} x^2 + y^2 &= 36; \\ x &\geq 0; \quad y &\geq 0 \end{aligned} \right\}$$

найти глобальные экстремумы функции

$$F = (x - 3)^2 + (y - 2)^2.$$

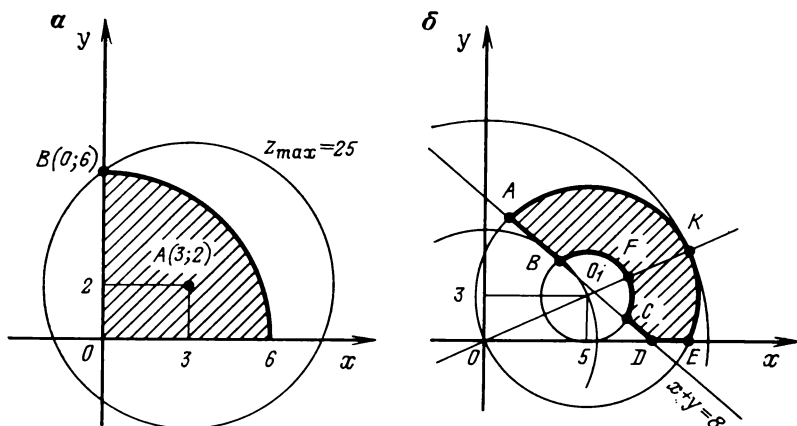


Рис. 38. Задачи нелинейного программирования с нелинейной целевой функцией и нелинейной системой ограничений:

a — пример V.6; *б* — пример V.7

Р е ш е н и е. Линиями уровня функции $F = (x - 3)^2 + (y - 2)^2$ являются окружности с центром в точке $A(3; 2)$ (рис. 38, *a*). Из рисунка видно, что глобальный минимум функции F достигается в точке $A(3; 2)$, а глобальный максимум — в точке $B(0; 6)$.

$$F_{\min} = 0; \quad F_{\max} = 25.$$

Пример V.7. Найти глобальные экстремумы функции

$$F = x^2 + y^2$$

на множестве системы ограничений

$$\left. \begin{aligned} (x - 5)^2 + (y - 3)^2 &\geq 9; \\ (x - 5)^2 + (y - 3)^2 &\leq 36; \\ x + y &\geq 8; \\ x &\geq 0; \\ y &\geq 0. \end{aligned} \right\}$$

Р е ш е н и е. На рис. 38, *б* множество допустимых решений заштриховано. Как видно из рисунка, оно не является выпуклым. Очевидно, что наименьшее значение функции F достигается в точке B , а наибольшее — в точке K (точка касания окружности $(x - 5)^2 + (y - 3)^2 = 36$ и линии уровня). Заметим, что точке E — точка локального максимума, так как значение функции F в ней больше, чем значения в соседних вершинах B и C . Аналогично точка C является точкой локального минимума.

Рассмотренные задачи позволяют установить ряд особенностей нелинейных задач, которые делают их более трудными для решения, чем линейные задачи. Однако модели нелинейного программирования позволяют учитывать нелинейный характер зависимости, например, общих затрат, выпуска производства и т. п. от плана $x = (x_1, x_2, \dots, x_n)$.

§ 28. Задача нелинейного программирования при ограничениях неотрицательности

Если на переменные налагаются только условия неотрицательности ($m = 0$), то основная задача (V.1) сводится к выбору неотрицательных значений переменных, максимизирующих функцию, т. е. к задаче вида: найти

$$\max_{\bar{x}} F(\bar{x}) \quad (\text{V.5})$$

при условии, что $\bar{x} \geq 0$.

Предположим, что в $\bar{x}^*$ существует локальный максимум задачи (V.5). Тогда в точках, близких к $\bar{x}^*$,

$$F(\bar{x}^*) \geq F(\bar{x}^* + h\Delta\bar{x}), \quad (\text{V.6})$$

где $\Delta\bar{x}$ определяет направление в E_n , а h — любое произвольное малое положительное число. Пусть $F(\bar{x})$ — дважды непрерывно дифференцируемая функция, тогда функцию в правой части (V.6) можно разложить по формуле Тейлора в окрестности $\bar{x}^*$:

$$\begin{aligned} F(\bar{x}^* + h\Delta\bar{x}) &= F(\bar{x}^*) + h \frac{\partial F}{\partial \bar{x}}(\bar{x}^*) \Delta\bar{x} + \\ &+ \frac{1}{2!} h^2 (\Delta\bar{x})^T \frac{\partial^2 F}{\partial \bar{x}^2}(\bar{x}^* + \theta h\Delta\bar{x}) \Delta\bar{x}, \end{aligned} \quad (\text{V.7})$$

где $0 < \theta < 1$.

Сопоставляя последние два неравенства, приходим к основному неравенству:

$$h \frac{DF}{d\bar{x}}(\bar{x}^*) \Delta\bar{x} + \frac{1}{2} h^2 (\Delta\bar{x})^T \frac{\partial^2 F}{\partial \bar{x}^2}(\bar{x}^* + \theta h\Delta\bar{x}) \Delta\bar{x} \leq 0, \quad (\text{V.8})$$

являющемуся необходимым условием существования локального максимума в $\bar{x}^*$. Если $\bar{x}^* > 0$, т. е. $\bar{x}^*$ — внутренняя точка, то основное неравенство выполняется для всех направлений $\Delta\bar{x}$. Поэтому, как и в классических задачах математического программирования, условием первого порядка является обращение в нуль первых производных. Допустим теперь, что одна из переменных принимает граничное значение $x_j^* = 0$. Пусть приращения всех остальных переменных равны 0. Тогда, поскольку при $x_j^* = 0$ единственным допустимым направлением является такое, при котором $\Delta x_j \geq 0$, из основного неравенства следует

$$\frac{\partial F}{\partial x_j}(\bar{x}^*) \Delta x_j \leq 0. \quad (\text{V.9})$$

Следовательно, условие первого порядка таково:

$$\frac{\partial F}{\partial x_j}(\bar{x}^*) \leq 0, \quad \text{если } x_j^* = 0. \quad (\text{V.10})$$

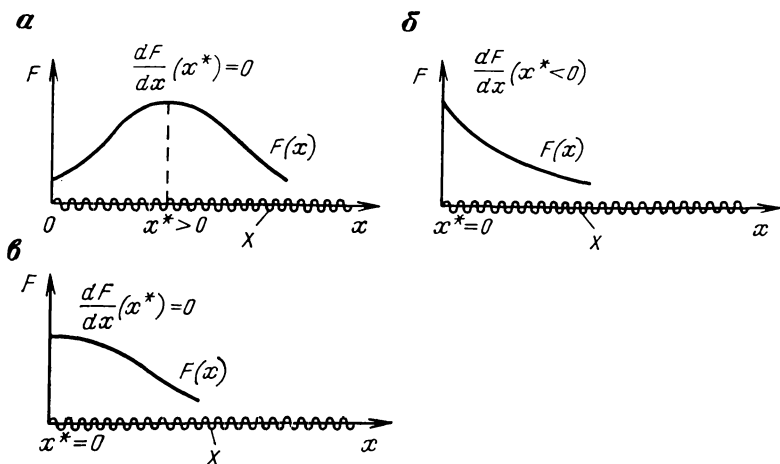


Рис. 39. Три возможных решения задачи максимизации функции одной неотрицательной переменной:

a — внутреннее решение (внутренний максимум). Угловой коэффициент касательной равен нулю; *б* — граничное решение (максимум на границе). Угловой коэффициент касательной меньше нуля; *в* — граничное решение (максимум на границе). Угловой коэффициент касательной равен нулю

Итак, в точке внутреннего максимума ($x_j^* > 0$) производная по x_j обращается в нуль, а если максимум достигается на границе ($x_j^* = 0$), то первая производная меньше или равна нулю. Так как либо производная равна 0 (во внутренней точке), либо соответствующая переменная принимает нулевое значение (в граничной точке), то их произведение в точке максимума всегда равно нулю:

$$\frac{\partial F}{\partial x_j}(\bar{x}^*) x_j^* = 0. \quad (V.11)$$

Сумма этих произведений также равна нулю

$$\frac{\partial F}{\partial \bar{x}}(\bar{x}^*) \bar{x}^* = \sum_{j=1}^n \frac{\partial F}{\partial x_j}(\bar{x}^*) x_j^* = 0. \quad (V.12)$$

Условие обращения в нуль суммы указанных произведений на самом деле требует, чтобы каждое слагаемое равнялось нулю (т. е. соотношение (V.11) должно выполняться для всех j). Это следует из того, что все переменные неотрицательны, а частные производные первого порядка не превышают нуля. Таким образом, локальный максимум в $\bar{x}^*$ определяется с помощью $2n + 1$ условий первого порядка:

$$\frac{\partial F}{\partial \bar{x}}(\bar{x}^*) \leq 0; \quad \frac{\partial F}{\partial x_j}(\bar{x}^*) \bar{x}^* = 0; \quad \bar{x}^* \geq 0. \quad (V.13)$$

Из этих условий следует приведенный выше результат: в точке максимума частная производная либо равна нулю, если соответ-

ствующая переменная положительна, либо меньше или равна нулю, если соответствующая переменная обращается в нуль

$$\left. \begin{aligned} \frac{\partial F}{\partial x_j}(\bar{x}^*) &= 0, & \text{если } x_j^* > 0 \\ \frac{\partial F}{\partial x_j}(\bar{x}^*) &\leq 0, & \text{если } x_j^* = 0 \end{aligned} \right\} \quad j = 1, 2, \dots, n. \quad (\text{V.14})$$

На рис. 39 проиллюстрированы три возможных решения задачи в одномерном случае.

§ 29. Условия Куна — Таккера

Используя результаты предыдущего параграфа, можно проанализировать общую задачу нелинейного программирования: найти

$$\max_{\bar{x}} F(\bar{x})$$

при условии

$$\bar{g}(\bar{x}) \leq b; \quad x \geq 0. \quad (\text{V.15})$$

Для того, чтобы преобразовать ограничения неравенства в ограничения в форме равенств, введем вектор, состоящий из m вспомогательных (свободных) переменных:

$$\bar{S} \equiv \bar{b} - \bar{g}(\bar{x}) = (S_1, S_2, \dots, S_m)^T. \quad (\text{V.16})$$

Теперь задача состоит в отыскании

$$\max_{\bar{x}, \bar{S}} F(\bar{x})$$

при условии

$$\bar{g}(\bar{x}) + S = \bar{b}; \quad \bar{x} \geq 0; \quad \bar{S} \geq 0. \quad (\text{V.17})$$

причем неотрицательность вспомогательных переменных обеспечивает выполнение ограничений неравенств. Если бы выражения (V.17) не включали $m + n$ ограничений неравенств, то наша задача являлась бы классической задачей математического программирования.

Функция Лагранжа такой задачи имела бы вид

$$L' = F(\bar{x}) + \bar{\lambda}(\bar{b} - \bar{g}(\bar{x}) - \bar{S}), \quad (\text{V.18})$$

где $\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)$ — вектор множителей Лагранжа (как и в предыдущей главе).

Необходимые условия первого порядка состояли бы в том, что все первые частные производные функции L' по $\bar{x}$, $\bar{\lambda}$ и $\bar{S}$ должны обращаться в нуль. Но так как x и $\bar{S}$ неотрицательны, то соответствующие условия на значениях первых производных по этим $m + n$ переменным заменяются на условия, которые можно получить по формулам (V.13). В соответствии с этим условия первого

порядка для существования локального максимума задачи (V.17) состоят в следующем:

$$\left. \begin{aligned} \frac{\partial L'}{\partial \bar{x}} &= \frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} \leq 0; \\ \frac{\partial L'}{\partial \bar{x}} \bar{x} &= \left(\frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} \right) \bar{x} = 0; \\ \bar{x} &\geq 0, \\ \frac{\partial L'}{\partial \bar{\lambda}} &= \bar{b} - \bar{g}(\bar{x}) - \bar{S} = 0; \\ \frac{\partial L'}{\partial \bar{S}} &= -\bar{\lambda} \leq 0; \\ \frac{\partial L'}{\partial \bar{S}} \bar{S} &= -\bar{\lambda} \bar{S} = 0; \\ \bar{S} &\geq 0, \end{aligned} \right\} \quad (\text{V.19})$$

где все переменные, функции и производные вычисляются при $\bar{x}^*$, $\bar{\lambda}^*$ и $\bar{S}^*$. Заменяя вектор вспомогательных переменных $\bar{S}$ на $\bar{b} - \bar{g}(\bar{x})$, приходим к условиям Куна — Таккера:

$$\left. \begin{aligned} \left(\frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} \right) &\leq 0; \\ \bar{x} &\geq 0. \\ \left(\frac{\partial F}{\partial \bar{x}} - \bar{\lambda} \frac{\partial \bar{g}}{\partial \bar{x}} \right) \bar{x} &= 0; \\ (b - g(x)) &\geq 0; \\ \bar{\lambda} &= 0. \\ \bar{\lambda}(\bar{b} - \bar{g}(\bar{x})) &= 0; \end{aligned} \right\} \quad (\text{V.20})$$

Те же условия можно получить, если определить функцию Лагранжа для исходной задачи (V.15) как сумму целевой функции и скалярного произведения векторов множителей Лагранжа и разности между константами ограничений:

$$L = L(\bar{x}, \bar{\lambda}) = F(\bar{x}) + \bar{\lambda}(\bar{b} - \bar{g}(\bar{x})),$$

или в развернутом виде

$$\begin{aligned} L(x_1, x_2, \dots, x_n; \lambda_1, \lambda_2, \dots, \lambda_m) &= F(x_1, x_2, \dots, x_n) + \\ &+ \sum_{i=1}^m \lambda_i (b_i - g_i(x_1, x_2, \dots, x_n)). \end{aligned} \quad (\text{V.21})$$

Условия Куна — Таккера записываются теперь следующим образом:

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{x}} (\bar{x}^*, \bar{\lambda}^*) &= \frac{\partial F}{\partial \bar{x}} (\bar{x}^*) - \bar{\lambda}^* \frac{\partial \bar{g}}{\partial \bar{x}} (\bar{x}^*) \leq 0; \\ \frac{\partial L}{\partial \bar{x}} (\bar{x}^*, \bar{\lambda}^*) \bar{x}^* &= \left(\frac{\partial F}{\partial \bar{x}} (\bar{x}^*) - \bar{\lambda}^* \frac{\partial \bar{g}}{\partial \bar{x}} (\bar{x}^*) \right) \bar{x}^* = 0, \\ \bar{x}^* &\geq 0. \\ \frac{\partial L}{\partial \bar{\lambda}} (\bar{x}^*, \bar{\lambda}^*) &= \bar{b} - \bar{g} (\bar{x}^*) \geq 0, \\ \bar{\lambda}^* \frac{\partial L}{\partial \bar{\lambda}} (\bar{x}^*, \bar{\lambda}^*) &= \bar{\lambda}^* (\bar{b} - \bar{g} (\bar{x}^*)) = 0, \\ \bar{\lambda}^* &\geq 0. \end{aligned} \right\} \quad (\text{V.22})$$

Эти условия являются необходимыми и достаточными для существования (строгого) локального максимума, если целевая функция (строго) вогнутая, а функции ограничений выпуклые и если, кроме того, выполнено некоторое условие регулярности, налагаемое на ограничения. Условия Куна — Таккера можно записать в развернутом виде как $2m + 2n + 2$ соотношений:

$$\frac{\partial L}{\partial x_j} = \frac{\partial F}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} \leq 0; \quad 1, 2, \dots, n; \quad (\text{V.23})$$

$$\sum_{j=1}^n \frac{\partial L}{\partial x_j} x_j = \sum_{j=1}^n \left(\frac{\partial F}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} \right) x_j = 0; \quad (\text{V.24})$$

$$x_j \geq 0; \quad j = 1, 2, \dots, n. \quad (\text{V.25})$$

$$\frac{\partial L}{\partial \lambda_i} = b_i - g_i(\cdot) \geq 0; \quad i = 1, 2, \dots, m; \quad (\text{V.26})$$

$$\sum_{i=1}^m \lambda_i \frac{\partial L}{\partial \lambda_i} = \sum_{i=1}^m \lambda_i (b_i - g_i(\cdot)) = 0; \quad (\text{V.27})$$

$$\lambda_i \geq 0; \quad i = 1, 2, \dots, m, \quad (\text{V.28})$$

где все переменные, функции и производные вычисляются по $(\bar{x}^*, \bar{\lambda}^*)$.

Для того чтобы понять важность этих условий, отметим прежде всего, что соотношения (V.25) и (V.26) выражают соответственно ограничения неотрицательности и ограничения неравенства исходной задачи нелинейного программирования. Отметим, что вслед-

ствие неравенств (V.23) и (V.25) каждое из слагаемых суммы (V.24) должно равняться нулю, так что или

$$\frac{\partial F}{\partial x_j} - \sum_{i=1}^m \lambda_i \left(\frac{dg_i}{dx_j} \right) = 0, \quad \text{или} \quad x_j = 0, \quad (\text{V.29})$$

или эти равенства выполняются одновременно, $j = 1, 2, \dots, n$. Иначе говоря, или каждое из соотношений между производными выполняется как равенство, или переменная равна нулю, или обе эти возможности осуществляются одновременно. Таким образом,

$$\left. \begin{aligned} \frac{\partial F}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} &\leq 0, \text{ но} \\ \frac{\partial F}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} &= 0, \text{ если } x_j^* > 0, \\ x_j^* &\geq 0, \text{ но } x_j^* = 0, \text{ если } \frac{\partial F}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} < 0. \end{aligned} \right\} j = 1, 2, \dots, n \quad (\text{V.30})$$

Каждое из слагаемых суммы (V.27) также равно нулю благодаря неравенствам (V.26) и (V.28), так что или

$$\lambda_i = 0, \quad \text{или} \quad g_i(\bar{x}^*) = b_i \quad (\text{V.31})$$

или же эти неравенства выполняются одновременно, $i = 1, 2, \dots, m$.

Иначе говоря, или множитель Лагранжа равен нулю, или соответствующее ограничение неравенства выполняется как равенство, или и то и другое выполняется одновременно. Итак,

$$\left. \begin{aligned} g_i(\bar{x}^*) &\leq b_i, \text{ но } g_i(\bar{x}^*) = b_i, \text{ если } \lambda_i^* > 0; \\ \lambda_i^* &\geq 0, \text{ но } \lambda_i^* = 0, \text{ если } g_i(\bar{x}^*) < b_i. \end{aligned} \right\} \quad (\text{V.32})$$

Условия (V.30) и (V.32), представляющие собой иной способ формулирования условия Куна—Таккера, называют условиями дополняющей нежесткости. Совершенно ясно, что

$$L(\bar{x}^*, \bar{\lambda}^*) = F(\bar{x}^*) + \bar{\lambda}(\bar{b} - g(\bar{x}^*)) = F(\bar{x}^*), \quad (\text{V.33})$$

т. е. значение функции Лагранжа в точке решения совпадает с оптимальным значением целевой функции, так как согласно формуле (V.27)

$$\bar{\lambda}^*(\bar{b} - \bar{g}(\bar{x}^*)) = 0.$$

§ 30. Теорема Куна — Таккера

Подход Куна — Таккера к общей задаче нелинейного программирования: найти

$$\max_{\bar{x}} F(\bar{x}) \quad (V.34)$$

при условии $\bar{g}(\bar{x}) \leq \bar{b}$, $\bar{x} \geq 0$, изложенный в предыдущем параграфе, состоит в том, что вводится вектор-строка множителей Лагранжа $\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)$, число которых равно числу ограничений неравенств и определяется функциями Лагранжа

$$L(\bar{x}, \bar{\lambda}) = F(\bar{x}) + \bar{\lambda}(\bar{b} - \bar{g}(\bar{x})). \quad (V.35)$$

В этом случае условия Куна — Таккера записываются так:

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{x}}(\bar{x}^*, \bar{\lambda}^*) &\leq 0; & \frac{\partial L}{\partial \bar{\lambda}}(\bar{x}^*, \bar{\lambda}^*) &\geq 0; \\ \frac{\partial L}{\partial \bar{x}}(\bar{x}^*, \bar{\lambda}^*) \bar{x}^* &= 0; & \bar{\lambda}^* \frac{\partial L}{\partial \bar{\lambda}}(\bar{x}^*, \bar{\lambda}^*) &= 0. \\ \bar{x}^* &\geq 0; & \bar{\lambda}^* &\geq 0. \end{aligned} \right\} \quad (V.36)$$

Из сопоставления этих неравенств с условиями существования максимума (V.13) видно, что $(\bar{x}^*, \bar{\lambda}^*)$ является седловой точкой функции Лагранжа, т. е. точкой, максимизирующей функцию по совокупности всех неотрицательных переменных $\bar{x}$ и минимизирующей ее по совокупности всех неотрицательных множителей Лагранжа $\bar{\lambda}$:

$$L(\bar{x}, \bar{\lambda}^*) \leq L(\bar{x}^*, \bar{\lambda}^*) \leq L(\bar{x}^*, \bar{\lambda}) \quad (V.37)$$

при всех $\bar{x} \geq 0$; $\bar{\lambda} \geq 0$.

Задача нахождения неотрицательных векторов $(\bar{x}^*, \bar{\lambda}^*)$, удовлетворяющих соотношению (V.37), называется задачей с седловой точкой (см. § 26). Тот факт, что $(\bar{x}, \bar{\lambda}^*)$ представляет собой седловую точку функции $L(\bar{x}, \bar{\lambda})$ и что два набора условий (V.36) симметричны, естественным образом приводит к двойственным задачам:

$$\begin{aligned} \max_{\bar{x}} L(\bar{x}, \bar{\lambda}) \text{ при условии } \frac{\partial L}{\partial \bar{\lambda}} &\geq 0; \quad \bar{x} \geq 0; \\ \min_{\bar{\lambda}} L(\bar{x}, \bar{\lambda}) \text{ при условии } \frac{\partial L}{\partial \bar{x}} &\leq 0; \quad \bar{\lambda} \geq 0. \end{aligned}$$

Этот подход оказался исключительно плодотворным в теории линейного программирования, однако в теории нелинейного программирования он не настолько удачен.

По теореме Куна — Таккера, $\bar{x}$ является решением задачи нелинейного программирования, если $(\bar{x}^*, \bar{\lambda}^*)$ — решение задачи о седловой точке, а $\bar{x}^*$ служит решением задачи нелинейного про-

граммирования лишь в том случае, если существует такой вектор $\bar{\lambda}^*$, что $(\bar{x}^*, \bar{\lambda}^*)$ является решением задачи о седловой точке.

Рассмотрим теперь задачу о седловой точке при дополнительном предположении о дифференцируемости функций $F(\bar{x})$ и $\bar{g}(\bar{x})$, которое не использовалось при доказательстве. Первая часть задачи о седловой точке заключается в отыскании максимума по неотрицательным переменным $\bar{x}$. Применяя выражения (V.13), получаем следующие условия:

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{x}}(\bar{x}^*, \bar{\lambda}^*) &\leq 0; \\ \frac{\partial L}{\partial \bar{x}}(\bar{x}^*, \bar{\lambda}^*) \bar{x}^* &= 0; \\ \bar{x}^* &\geq 0. \end{aligned} \right\} \quad (\text{V.38})$$

Для решения второй задачи о седловой точке — минимизации $L(\bar{x}^*, \bar{\lambda})$ по неотрицательным множителям Лагранжа $\bar{\lambda}$ требуется выполнение условий

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{\lambda}}(\bar{x}^*, \bar{\lambda}^*) &\geq 0; \\ \bar{\lambda}^* \frac{\partial L}{\partial \bar{\lambda}}(\bar{x}^*, \bar{\lambda}^*) &= 0; \quad \bar{\lambda}^* \geq 0. \end{aligned} \right\} \quad (\text{V.39})$$

Полученные соотношения представляют собой условия Куна — Таккера [29].

Сравним теперь утверждение рассматриваемой теоремы Куна — Таккера с утверждением аналогичной теоремы (условия первого порядка) (IV.47), касающейся решения классической задачи на условный экстремум.

Во-первых, в отличие от условий (IV.47) теорема Куна — Таккера устанавливает не только необходимые, но и достаточные условия максимума [условия (V.36)].

Во-вторых, вместо равенства нулю в точке $(x^*, \bar{\lambda}^*)$ всех частных производных функций Лагранжа, утверждаемого условиями (IV.47), настоящая теорема допускает возможность неравенств (V.38) и (V.39). Хотя в общем случае неравенство есть более слабое требование, нежели равенство, в данном случае этого сказать нельзя, поскольку условия (V.38) и (V.39) четко оговаривают, в каких случаях частные производные функции Лагранжа могут оказаться отличными от нуля.

В-третьих, настоящая теорема утверждает неотрицательность множителей Лагранжа, в то время как в «классическом случае» решения заранее знаки множителей Лагранжа оговорены быть не могут.

Таким образом, мы видим, что утверждение теоремы Куна — Таккера является более сильным, чем условия (IV.47).

Рассмотрим теперь характерный пример задачи нелинейного программирования.

Пример V.8. Найти

$$\max_{x_1, x_2} F(x_1, x_2) = 10x_1 + 20x_2 + x_1x_2 - 2x_1^2 - 2x_2^2$$

при условии

$$\begin{aligned} x_1 + x_2 &\leq 5; \\ x_1 &\geq 0; \quad x_2 &\geq 0. \end{aligned}$$

Решение. Сформулируем функцию Лагранжа

$$F(x_1, x_2, \lambda) = 10x_1 + 20x_2 + x_1x_2 - 2x_1^2 - 2x_2^2 - \lambda(x_1 + x_2 - 5). \quad (\text{V.40})$$

и условия Куна — Таккера

$$\left. \begin{aligned} \frac{\partial F}{\partial x_1} &= 10 + x_2 - 4x_1 - \lambda; \\ \frac{\partial F}{\partial x_2} &= 20 + x_1 - 4x_2 - \lambda; \\ \frac{\partial F}{\partial \lambda} &= -x_1 - x_2 + 5. \end{aligned} \right\} \quad (\text{V.41})$$

Неравенства (V.23) — (V.28) запишутся теперь следующим образом:

$$\left. \begin{aligned} \text{а) } 10 + x_2 - 4x_1 - \lambda &\leq 0; \\ \text{б) } 20 + x_1 - 4x_2 - \lambda &\leq 0; \\ \text{в) } x_1(10 + x_2 - 4x_1 - \lambda) &= 0; \\ \text{г) } x_2(20 + x_1 - 4x_2 - \lambda) &= 0; \\ \text{д) } x_1 + x_2 - 5 &\leq 0; \\ \text{е) } \lambda(x_1 + x_2 - 5) &= 0; \\ \text{ж) } x_1 &\geq 0; \quad x_2 \geq 0; \quad \lambda \geq 0. \end{aligned} \right\} \quad (\text{V.42})$$

Из рис. 40 ясно, что $x_1^* > 0$, $x_2^* > 0$. Поэтому

$$10 + x_2 - 4x_1 - \lambda = 0; \quad 20 + x_1 - 4x_2 - \lambda = 0.$$

Откуда

$$x_1 = 4 - 1/3\lambda; \quad x_2 = 6 - 1/3\lambda. \quad (\text{V.43})$$

Из формулы (V.43) следует, что $\lambda > 0$, так как точка $(x_1 = 4, x_2 = 6)$ лежит вне допустимой области. Тогда из соотношения (V.42, е) имеем

$$x_1 + x_2 - 5 = 0. \quad (\text{V.44})$$

Подставляя в уравнение (V.44) значения x_1 и x_2 из уравнения (V.43), получаем $10 - 2/3\lambda - 5 = 0$; $\lambda = 7,5$. Отсюда с использованием формул (V.43)

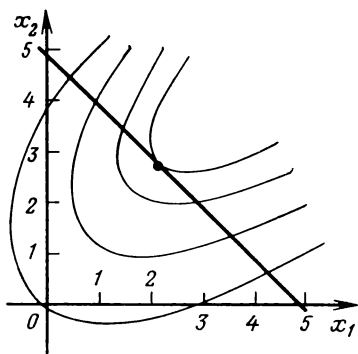


Рис. 40. Задача нелинейного программирования

$x_1^* = 1,5$; $x_2^* = 3,5$. Таким образом, в точке $(x_1^*, x_2^*) = (1,5; 3,5)$ функция

$$F(x_1, x_2) = 10x_1 + 20x_2 + x_1x_2 - 2x_1^2 - 2x_2^2$$

достигает максимума в пределах допустимой области, ограниченной неравенствами: $x_1 + x_2 \leq 5$; $x_1 \geq 0$; $x_2 \geq 0$.

Отметим, что условия Куна — Таккера дают полную характеристику решения, однако они не содержат конструктивного метода отыскания решения. Например, условия (V.42) описывают свойства решения задачи (V.40), но не указывают, как найти это решение.

§ 31. Геометрический способ решения задач нелинейного программирования

Предположим, что для изготовления продукции P_1 и P_2 требуется использование трех видов ресурсов R_1 , R_2 , R_3 . Располагаемое количество ресурсов и нормы их расхода на изготовление единицы каждого вида продукции известны (табл. 6).

Т а б л и ц а 6

Ресурсы	Количество ресурсов	Нормы расхода ресурсов на единицу продукции	
		P_1	P_2
R_1	80	$3 - 0,045x_1$	$4 - 0,071x_2$
R_2	125	1	5
R_3	60	2	1

Расход ресурсов R_1 на единицу продукции вида P_1 и P_2 непостоянный и уменьшается с увеличением соответствующего количества продукции x_1 и x_2 .

Прибыль, получаемая предприятием от реализации единицы продукции P_1 и P_2 , составляет соответственно 4 и 5 денежных единиц. В задаче требуется составить такой план выпуска продукции видов P_1 и P_2 , при котором прибыль предприятия от реализации всей продукции оказалась бы максимальной.

Требуется найти неотрицательные числа x_1 и x_2 , максимизирующие целевую функцию (функцию прибыли):

$$F(x_1, x_2) = 4x_1 + 5x_2 \quad (\text{V.45})$$

и удовлетворяющие условиям:

$$\left. \begin{aligned} (3 - 0,045x_1)x_1 + (4 - 0,071x_2)x_2 &\leq 80; \\ x_1 + 5x_2 &\leq 125; \\ 2x_1 + x_2 &\leq 60. \end{aligned} \right\} \quad (\text{V.46})$$

Эта задача нелинейная, так как одно из ограничений нелинейно. Если представить ограничения (V.46) в виде $g_i(x_1, x_2) \leq 0$ ($i = 1, 2, 3$), то функции $g_i(x_1, x_2)$ будут иметь вид:

$$\left. \begin{aligned} g_1(x_1, x_2) &= (3 - 0,045x_1)x_1 + (4 - 0,071x_2)x_2 - 80 \leq 0; \\ g_2(x_1, x_2) &= x_1 + 5x_2 - 125 \leq 0; \\ g_3(x_1, x_2) &= 2x_1 + x_2 - 60 \leq 0. \end{aligned} \right\} \quad (\text{V.47})$$

Вычислим градиенты всех функций:

$$\left. \begin{aligned} \nabla F(x_1, x_2) &= \left(\frac{\partial f}{\partial x_1}; \frac{\partial f}{\partial x_2} \right) = (4; 5); \\ \nabla g_1(x_1, x_2) &= \left(\frac{\partial g_1}{\partial x_1}; \frac{\partial g_1}{\partial x_2} \right) = (3 - 0,090x_1; 4 - 0,142x_2); \\ \nabla g_2(x_1, x_2) &= \left(\frac{\partial g_2}{\partial x_1}; \frac{\partial g_2}{\partial x_2} \right) = (1; 5); \\ \nabla g_3(x_1, x_2) &= \left(\frac{\partial g_3}{\partial x_1}; \frac{\partial g_3}{\partial x_2} \right) = (2; 1). \end{aligned} \right\} \quad (\text{V.48})$$

Вследствие неотрицательности $x_1 \geq 0$ и $x_2 \geq 0$ допустимая область должна быть расположена в первой четверти координатной плоскости x_1 и x_2 .

Для построения контура допустимой области следует приравнять к нулю функции $g_1(x_1, x_2)$ (V.47) и построить соответствующие нулевые линии уровня этих функций. Нулевые линии уровня $g_2(x_1, x_2)$ и $g_3(x_1, x_2)$ являются прямыми, способ построения их элементарен и поэтому не требует пояснений.

Нулевая линия уровня $g_1(x_1, x_2) = 0$ — кривая. Для построения ее достаточно вычислить координаты трех-четырех точек этой кривой и провести с помощью лекала эту линию (рис. 41). В произвольных точках этих линий следует построить соответствующие градиенты и потом заштриховать допустимую область, которая вполне определится (см. рис. 41).

Далее, задавая произвольным значением линейной целевой (V.45) функции, например равным 100, строим соответствующую прямую уровня

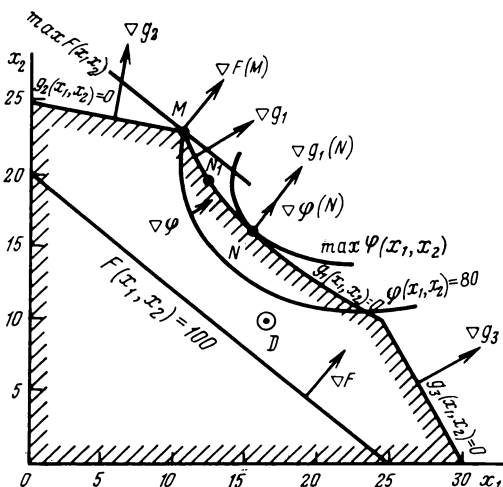


Рис. 41. Градиентный метод решения задачи нелинейного программирования

(см. рис. 41). Затем по градиенту ∇F определяем направление возрастания этой функции и строим допустимую линию уровня, соответствующую максимуму функции (V.45), которая проходит через точку M пересечения граничных линий $g_2(x_1, x_2) = 0$; $g_1(x_1, x_2) = 0$. Эта линия уровня не рассчитывается, а строится параллельным переносом линии уровня $4x_1 + 5x_2 = 100$ в направлении вектора ∇F , так как линии уровня всякой линейной функции параллельны. Координаты точки M , которая является пересечением кривой линии $g_1(x_1, x_2) = 0$ с прямой $g_2(x_1, x_2) = 0$, есть оптимальный план задачи.

Для вычисления координат оптимальной точки M следует решить нелинейную систему уравнений

$$\left. \begin{aligned} (3 - 0,045x_1)x_1 + (4 - 0,071x_2)x_2 &= 80; \\ x_1 + 5x_2 &= 125. \end{aligned} \right\} \quad (\text{V.49})$$

Неотрицательное решение этой системы: $x_1 = 10$; $x_2 = 23$. Таким образом, для обеспечения максимальной прибыли необходимо производить 10 единиц продукции вида P_1 и 23 единицы продукции вида P_2 . При этом максимальная прибыль составит

$$\max F(M) = 4 \times 10 + 5 \times 23 = 155 \text{ денежных единиц.}$$

Теперь изменим условия задачи. Представим себе, что целевая функция нелинейна и имеет вид

$$\varphi = (x_1, x_2) = (4 - 0,100x_1)x_1 + (5 - 0,126x_2)x_2. \quad (\text{V.50})$$

В этом случае прибыль на единицу продукции уменьшается с увеличением ее выпуска. Положим, что такой случай вызван, например, увеличением дополнительных непроизводительных затрат, связанных с дальнейшим увеличением выпуска продукции (расходы по хранению, местной транспортировке и т. д.). Задавая произвольным значением (V.50), например числом 80, строим соответствующую линию уровня этой функции (см. рис. 41). Вычисляем градиент функции (V.50):

$$\nabla \varphi(x_1, x_2) = \left(\frac{\partial \varphi}{\partial x_1}, \frac{\partial \varphi}{\partial x_2} \right) = (4 - 0,200x_1; 5 - 0,252x_2). \quad (\text{V.51})$$

Построение $\nabla \varphi$ в произвольной точке кривой $\varphi(x_1, x_2) = 80$ показывает, что целевая функция в допустимой области монотонно возрастает. Из рисунка ясно видно, что допустимая линия уровня, соответствующая максимуму этой функции, должна соприкоснуться с граничной линией $g_1(x_1, x_2) = 0$ в некоторой точке N , в которой градиенты $\nabla \varphi(N)$ и $\nabla g_1(N)$ имеют одинаковое направление. Следовательно, координаты точки N должны удовлетворять нелинейной системе уравнений, которая для данного примера будет иметь вид:

$$\left. \begin{aligned} 4 - 0,200x_1 &= \lambda(3 - 0,090x_1); \\ 5 - 0,252x_2 &= \lambda(4 - 0,142x_2); \\ (3 - 0,045x_1)x_1 + (4 - 0,071x_2)x_2 &= 80. \end{aligned} \right\} \quad (\text{V.52})$$

Неотрицательное решение этой системы $x_1 = 12,8$; $x_2 = 17,9$ является оптимальным планом выпуска продукции P_1 и P_2 с нелинейной целевой функцией (V.51), при этом максимальная прибыль составляет

$$\max \varphi(x_1, x_2) = \varphi(12,8; 17,9) \approx 84 \text{ денежных единицы.}$$

Несколько изменим условия задачи. Представим себе, что целевая функция имеет вид

$$\psi(x_1, x_2) = (4 - 0,125x_1)x_1 + (5 - 0,250x_2)x_2. \quad (\text{V.53})$$

Эта функция отличается от функции (V.50) тем, что прибыль на единицу продукции уменьшается с увеличением количества ее быстрее, чем в предыдущем случае. Вычислим градиент этой функции

$$\nabla \psi(x_1, x_2) = \left(\frac{\partial \psi}{\partial x_1}, \frac{\partial \psi}{\partial x_2} \right) = (4 - 0,025x_1; 5 - 0,50x_2); \quad (\text{V.54})$$

Приравняем составляющие градиента к нулю

$$4 - 0,25x_1 = 0;$$

$$5 - 0,50x_2 = 0.$$

Решением этой системы является внутренняя точка D допустимой области (см. рис. 41) с координатами $x_1 = 16$, $x_2 = 10$. Эта точка и будет оптимальным планом выпуска продукции видов P_1 и P_2 , при этом

$$\max \psi(x_1, x_2) = \psi(16; 10) = 57 \text{ денежных единиц.}$$

Отметим, что в реальной жизни связи между взаимозависимыми факторами носят нелинейный характер. Повышение дохода населения в два раза чаще всего не приводит к двукратному росту спроса на продукты потребления, увеличение пробега транспорта в два раза не приводит к двукратному увеличению затрат, в век научно-технического прогресса увеличение затрат ресурсов в n раз приводит к большему, чем в n раз, повышению выпуска продукции.

Поэтому большой практический интерес при моделировании процессов управления представляют методы и модели нелинейного программирования. Причиной же пока недостаточно широкого использования моделей нелинейного программирования (более адекватных реальной действительности) является, прежде всего, неразработанность самих моделей и математических методов их анализа [29].

Глава VI

ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

Задачи линейного программирования были первыми подробно изученными задачами оптимизации при наличии ограничений типа неравенств. Термин «линейное программирование» возник в результате неточного перевода английского «linear programming». Одно из значений слова «programming» — составление планов, планирование. Следовательно, правильным переводом этого термина было бы не линейное программирование, а линейное планирование или планирование на основе линейных соотношений, что более точно отражает содержание дисциплины.

§ 32. О постановках задач линейного программирования

1. Основные определения

Мы рассмотрим задачу максимизации линейной функции

$$F(\bar{x}) = c_1x_1 + c_2x_2 + \dots + c_nx_n = (\bar{c}\bar{x}). \quad (\text{VI.1})$$

Эта функция неограниченна, и поэтому искать ее максимум, не налагая никаких ограничений на область изменения вектора $\bar{x}$, бессмысленно. Интерес представляет задача максимизации $F(\bar{x})$ при условии, что $\bar{x}$ принадлежит некоторому множеству. Те из задач, в которых область изменения $\bar{x}$ — многогранник, т. е. выбор значений вектора $\bar{x}$ стеснен условиями типа линейных равенств и неравенств, составляют предмет линейного программирования.

В общем случае задача линейного программирования формулируется следующим образом: найти величины $x_1, x_2, \dots, x_n$, доставляющих максимум (минимум) линейной функции,

$$F(\bar{x}) = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \text{ext}$$

на множестве значений $x_1, x_2, \dots, x_n$, удовлетворяющих ограничениям, в числе которых могут присутствовать только равенства и неравенства вида

$$\left. \begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &\leq b_i; \\ a_{k1}x_1 + \dots + a_{kn}x_n &= b_k; \\ a_{l1}x_1 + \dots + a_{ln}x_n &\geq b_l. \end{aligned} \right\} \quad (\text{VI.2})$$

Среди ограничений задач линейного программирования часто встречаются условия неотрицательности всех или части переменных: $x_j \geq 0, j = 1, \dots, p$.

Хотя формально эти условия являются частным случаем представленных выше условий общего вида, на практике при построении алгоритмов их обычно выделяют в особую группу. Целевую функцию $F(\bar{x}) = (\bar{c}, \bar{x})$ принято называть критерием задачи линейного программирования (иногда ее называют также функционалом задачи линейного программирования).

Следует отметить, что предпосылка линейности, когда в реальной экономике подавляющее большинство зависимостей носит более сложный характер (например, нелинейный), есть определенное упрощение. В некоторых случаях оно достаточно реалистично, в других же выводы, получаемые с помощью решения задач линейного программирования (в том числе и на ЭВМ), оказываются весьма несовершенными.

2. Примеры прикладных задач линейного программирования

А. Задача о выборе оптимального плана [3]. Допустим, что экспедиция выполняет геодезические работы (рекогносцировку с постройкой) знаков и измерение углов на пунктах триангуляции 2 и 3 классов на трех объектах R_1, R_2, R_3 . Экспедиции заданы план работ по триангуляции (Т пунктов), их общая сметная (С) и плановая (Р) стоимость на текущий год и фонд зарплаты (Ф). В период подготовки к полевому сезону экспедиция должна решить вопрос о выборе объемов работ по постройке знаков и измерению углов на пунктах триангуляции 2 и 3 классов на каждом из трех объектов при условии, что для соблюдения, допустим, трехгодичного цикла работ на каждом объекте необходимо выполнить объем работ по триангуляции не менее t_4, t_5, t_6 (пунктов) соответственно.

Обозначим искомые объемы работ по рекогносцировке с постройкой знаков на объекте R_1 как x_1 , на объекте R_2 как x_2 и на объекте R_3 как x_3 , а объемы работ по измерению углов на пунктах триангуляции 2 и 3 классов — соответственно x_4, x_5 и x_6 . Следовательно, искомый план работ $X = (x_1, x_2, x_3, x_4, x_5, x_6)$.

Для формулировки ограничения по сметной стоимости работ должны быть известны значения сметной стоимости постройки знаков и наблюдений на каждом объекте. Обозначим их $a_{11}, a_{12}, \dots, a_{16}$. Тогда ограничение по сметной стоимости будет иметь вид

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + a_{15}x_5 + a_{16}x_6 = C$$

или

$$\sum_{j=1}^6 a_{1j}x_j = C. \quad (\text{VI.3})$$

Плановая стоимость искомого плана работ должна быть не выше заданной стоимости P . Если известны плановые себестоимости работ ($a_{21}, a_{22}, \dots, a_{26}$), то ограничение по плановой стоимости запишется в виде неравенства

$$\sum_{j=1}^6 a_{2j}x_j \leq P. \quad (\text{VI.4})$$

Аналогично запишется ограничение по фонду зарплаты

$$\sum_{j=1}^6 a_{3j}x_j \leq \Phi, \quad (\text{VI.5})$$

где a_{3j} ($j = 1, 2, \dots, 6$) — удельная зарплата, т. е. затраты по зарплате на постройку одного знака (a_{31}, a_{32}, a_{33}) и наблюдения на одном пункте (a_{34}, a_{35}, a_{36}).

Но ограничения (VI.3)—(VI.5) еще не в полной мере отражают условия поставленной задачи. Необходимо еще учесть заданный экспедиции объем работ по триангуляции. При этом следует применить ограничения снизу

$$x_4 + x_5 + x_6 \geq T \quad \text{или} \quad \sum_{j=4}^6 x_j \geq T, \quad (\text{VI.6})$$

так как экспедиция может перевыполнять заданные ей номенклатурные показатели. Кроме того, следует также учесть ограничения по обязательным объемам работ на каждом объекте. Это будут также ограничения снизу

$$x_4 \geq t_4; \quad x_5 \geq t_5; \quad x_6 \geq t_6. \quad (\text{VI.7})$$

Следует иметь в виду, что постройка знаков и наблюдение на пунктах — зависимые между собой процессы работ. Следовательно, в математическую модель задачи следует включить ограничения по заделам.

Допустим, минимальный задел по постройке знаков равен 30 %. Тогда ограничения по заделам будут иметь вид

$$x_1 \geq 1,3x_4; \quad x_2 \geq 1,3x_5; \quad x_3 \geq 1,3x_6 \quad (\text{VI.8})$$

или

$$x_1 - 1,3x_4 \geq 0; \quad x_2 - 1,3x_5 \geq 0; \quad x_3 - 1,3x_6 \geq 0.$$

После составления ограничений следует их логически проанализировать, установить, как ограничение влияет на оптимальный план, является ли оно необходимым, не противоречит ли условиям задачи. Кроме этого следует проанализировать, насколько полно составленная система ограничений отражает условия экономической задачи. Допустим, что в результате такого анализа установлено, что число строительных бригад в экспедиции не должно превышать Б (бригад), число бригад наблюдателей ничем не ограничивается. В этом случае нужно определить среднюю норму времени в бригадо-месяцах на постройку одного знака на каждом объекте a_{41}, a_{42}, a_{43} . Если обозначим П — продолжительность полевого периода в месяцах, то дополнительное ограничение по числу строительных бригад будет иметь вид

$$a_{41}x_1 + a_{42}x_2 + a_{43}x_3 \leq \text{БП} \quad (\text{VI.9})$$

или

$$\sum_{j=1}^3 a_{4j}x_j \leq \text{БП}.$$

Следующий этап решения задачи — установление критерия оптимальности для выбора из всех допустимых планов, описываемых системой ограничений (VI.3)—(VI.9), единственного оптимального плана. В рассматриваемом примере можно использовать различные критерии оптимальности: 1) выполнение заданного объема работ с минимальной плановой стоимостью; 2) максимальное перевыполнение заданного объема работ по триангуляции, т. е. максимальный выпуск продукции; 3) наименьшее число бригад при выполнении заданных показателей плана и т. д. В каждом конкретном случае выбирается наиболее подходящий критерий оптимальности либо находится оптимальный план по каждому из возможных критериев оптимальности и затем из них выбирается компромиссное решение.

Приведенным критериям оптимальности соответствуют следующие целевые функции:

минимум плановой стоимости работ

$$F_1 = \sum_{j=1}^6 a_{2j}x_j \rightarrow \min \quad (\text{VI.10})$$

максимум выпуска продукции

$$F_2 = \sum_{j=4}^6 x_j \rightarrow \max, \quad (\text{VI.11})$$

минимум числа бригад

$$F_3 = \sum_{j=1}^6 a_{4j}x_j \rightarrow \min. \quad (\text{VI.12})$$

где a_{44} , a_{45} , a_{46} — средние нормы времени в бригадо-месяцах на измерение углов на одном пункте триангуляции соответственно на объектах R_1 , R_2 , R_3 .

Итак, для составления плана работ экспедиции установлены система ограничений (VI.3)—(VI.9) и целевая функция (VI.10)—(VI.11) или (VI.12). Теперь можно записать экономико-математическую модель задачи с любым из приведенных критериев оптимальности. Например, с критерием оптимальности на минимум числа бригад экономико-математическую модель записывают следующим образом: найти план $X = (x_1, x_2, \dots, x_6)$, обращающий в минимум целевую функцию $F_3 = \sum_{j=1}^6 a_{4j}x_j$ при ограничениях

1) по сметной стоимости

$$\sum_{j=1}^6 a_{2j}x_j = C;$$

2) по плановой стоимости

$$\sum_{j=1}^6 a_{2j}x_j \leq P;$$

3) по фонду зарплаты

$$\sum_{j=1}^6 a_{3j}x_j \leq \Phi;$$

4) по заданному объему работ по измерениям углов на пунктах триангуляции 2 и 3 классов

$$\sum_{j=4}^6 x_j \geq T;$$

5) по обязательным объемам работ на каждом объекте

$$x_4 \geq t_4; \quad x_5 \geq t_5; \quad x_6 \geq t_6;$$

6) по заделам

$$x_1 - 1,3x_4 \geq 0;$$

$$x_2 - 1,3x_5 \geq 0;$$

$$x_3 - 1,3x_6 \geq 0;$$

7) по числу строительных бригад

$$\sum_{j=1}^3 a_{4j}x_j \leq \text{БП};$$

8) по неотрицательности искоемых переменных

$$x_j \geq 0.$$

Аналогично записывается экономико-математическая модель с критерием оптимальности на минимум плановой стоимости, целевой функцией которой служит выражение (VI.10), но в этой модели отпадает необходимость в ограничении по плановой стоимости.

В случае решения задачи на максимум выпуска продукции целевой функцией служит выражение (VI.12), система ограничений также остается прежней, за исключением ограничения по заданному объему работ, которое оказывается лишним, так как в оптимальном решении определяется максимальный объем пунктов по триангуляционным работам.

Б. Транспортная задача. Такая задача является одной из самых распространенных задач линейного программирования специального вида. Предположим, что имеется m складов, где хранится некоторый продукт в количестве $a_1, \dots, a_m$ и n пунктов реализации этого продукта, потребности которых равны $b_1, \dots, b_n$. Требуется найти наиболее экономичный способ перевозки продукта со складов к потребителям, если затраты на перевозку единицы продукции с i -го склада в j -й пункт потребления равны c_{ij} . Обозначим x_{ij} количество продукта, перевозимого с i -го склада в j -й пункт. Тогда задача минимизации транспортных расходов формулируется так:
найти

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \rightarrow \min,$$

ри ограничениях

$$\sum_{j=1}^n x_{ij} = a_i; \quad \sum_{i=1}^m x_{ij} = b_j; \quad x_{ij} \geq 0; \\ (i = 1, \dots, m); \quad (j = 1, \dots, n).$$

В. Задача о рационе. Предположим, что нужно выбрать суточный рацион питания, имея в распоряжении n продуктов, содержащих m питательных веществ.

Пусть a_{ij} — количество j -го питательного вещества в единице i -го продукта, b_j — суточная потребность организма в j -м питательном веществе, c_i — цена единицы i -го продукта. Если обозначить суточное потребление i -го продукта через x_i , стоимость рациона будет равна $c_1x_1 + c_2x_2 + \dots + c_nx_n$. Естественно поставить задачу минимизации при условии полного удовлетворения потребности организма в питательных веществах:

[illegible]

Потребности i -го продукта должны быть неотрицательны и $x_i \geq 0$
 $i = 1, 2, \dots, n$.

3. Стандартные формы представления задач линейного программирования

Существует немало вариантов формулировок задач линейного программирования, но наиболее употребительны две формы их записи. И к той и к другой форме можно привести любую задачу с помощью простых приемов. Некоторые из таких приемов перечислены ниже.

Переход от задачи минимизации к задаче максимизации. Для этого нужно изменить знак критеральной функции, т. е. задача минимизации функции

$$F(\bar{x}) = (\bar{c}, \bar{x})$$

Эквивалентна задача на поиск максимума функции

$$F(\bar{x}) = -(\bar{c}, \bar{x}).$$

Переход к эквивалентной системе неравенства. Меняя знаки свободного члена и коэффициентов в ограничении-неравенстве, можно поменять знак этого неравенства на обратный. Например, ограничение

$$d_1x_1 + d_2x_2 + \dots + d_nx_n \geq d$$

МОЖНО ЗАМЕНИТЬ СЛЕДУЮЩИМ:

$$(-d_1)x_1 + \dots + (-d_n)x_n \leq -d.$$

Переход от ограничения - неравенства к равенству. Любое ограничение в форме неравенства вве-

В матричной форме эта задача записывается так:

$$\max (\bar{c} \bar{x}); \quad (\text{VI.16})$$

$$A \bar{x} = \bar{b}; \quad (\text{VI.17})$$

$$\bar{x} \geq 0. \quad (\text{VI.18})$$

Здесь $\bar{A} = \|a_{ij}\|$ есть $(m \times n)$ -матрица условий, ее столбцы $\{a_{1j}, \dots, a_{mj}\}^T = 1, 2, \dots, n$ называются векторами условий, вектор $\bar{b} = \{b_1, \dots, b_m\}^T$ — вектором правых частей, а $(c_1, \dots, \dots, c_n)$ — вектором коэффициентов линейной формы. Далее всегда будем считать, что ранг матрицы $\bar{A}$ равен числу ее строк. Если это не так, а задача тем не менее разрешима, некоторые из уравнений (VI.14) должны быть линейными комбинациями остальных. Они «лишние» и их следовало бы отбросить.

Задача линейного программирования с однотипными условиями. Эта задача формулируется так: найти

$$\max (c_1x_1 + \dots + c_nx_n) \quad (\text{VI.19})$$

при ограничениях

[illegible]

или в матричной форме

$$\max (\bar{c} \bar{x}); \quad \bar{A} \bar{x} \leq \bar{b}.$$

Эту запись иногда называют сопряженной канонической формой.

§ 33. Геометрическая интерпретация задач линейного программирования

Двум основным формам задач линейного программирования (см. § 32) соответствуют два вида геометрических представлений этих задач. Одно из них реализуется в пространстве переменных, другое — в пространстве условий.

Рассмотрим геометрическую интерпретацию задачи с однотипными условиями в пространстве переменных. Геометрическая интерпретация линейного программирования в пространстве условий изложена в работе [31].

Итак, дана линейная форма (целевая функция)

$$F = c_1 x_1 + \dots + c_n x_n \quad (\text{VI.21})$$

и задана система $m > n$ линейных неравенств (ограничений)

$$a_{i1}x_1 + \dots + a_{in}x_n \leq b_i; \quad (i = 1, \dots, m), \quad (\text{VI.22})$$

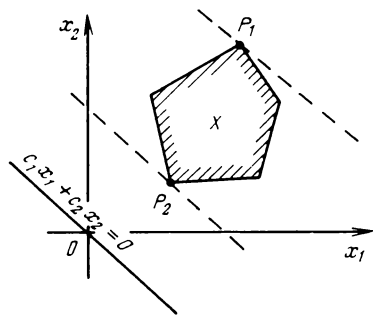


Рис. 42. Задача имеет два решения min и max

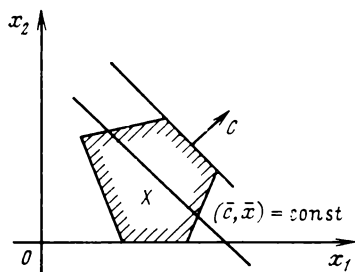


Рис. 43. Задача имеет множество решений

которую перепишем в виде

$$y_i \equiv a_{i1}(-x_1) + \dots + a_{in}(-x_n) + b_i \geq 0; \quad (i = 1, \dots, m). \quad (\text{VI.23})$$

Требуется определить максимум (минимум) формы (VI.21) при выполнении ограничений (VI.22) или (VI.23). Другими словами, среди решений системы (VI.22) надо отыскать такое, для которого форма (VI.21) принимает наибольшее (наименьшее) значение.

Каждый набор чисел $x_1, x_2, \dots, x_n$ будем рассматривать как координаты точки $\bar{x}(x_1, \dots, x_n)$ евклидова n -мерного пространства E_n . Тогда каждое уравнение

$$y_i \equiv a_{i1}(-x_1) + \dots + a_{in}(-x_n) + b_i = 0$$

можно трактовать как уравнение плоскости в этом пространстве. Каждое же неравенство $y_i \geq 0$ системы (VI.23) определяет некоторое полупространство пространства E_n , состоящее из точек $\bar{x}(x_1, \dots, x_n)$, расположенных по «одну сторону» от плоскости $y_i = 0$ и на самой этой плоскости. Система неравенств (VI.23), как пересечение полупространств определяет выпуклый многогранник (точнее, многогранное множество) X , любая точка которого (если он не пустой) принадлежит всем полупространствам $y_i \geq 0$ ($i = 1, \dots, m$), т. е. координаты точки удовлетворяют всем неравенствам (VI.23) и эта точка является одним из решений системы (VI.23).

Для каждой точки $\bar{x}' = (x'_1, \dots, x'_n)$ ее уклонение

$$F(x') = c_1x'_1 + \dots + c_nx'_n$$

от плоскости $c_1x_1 + \dots + c_nx_n = 0$ (т. е. число, получаемое при подстановке в уравнение плоскости координат данной точки x' вместо текущих координат) означает взятое с определенным знаком уравненное расстояние от точки x' до плоскости

$$F = c_1x_1 + \dots + c_nx_n = 0,$$

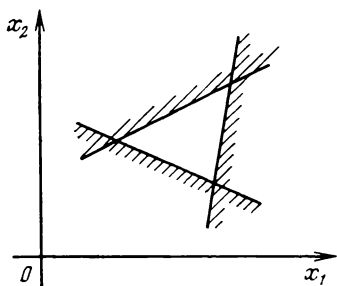


Рис. 44. Задача не имеет решений

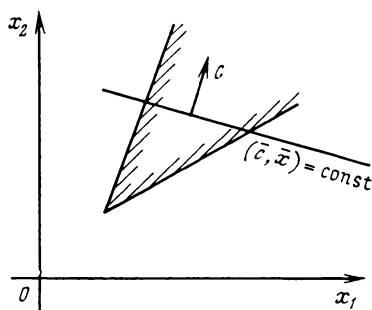


Рис. 45. Задача имеет решение $\min$

где уравненное расстояние — это расстояние, помноженное на величину (вес)

$$\sqrt{\sum_{j=1}^n c_j^2}.$$

Таким образом, геометрический смысл задачи линейного программирования заключается в отыскании такой точки многогранника X , которая наиболее (наименее) уклонена от плоскости.

В двумерном случае, изображенном на рис. 42, каждое уравнение

$$y_i \equiv a_{i1}(-x_1) + a_{i2}(-x_2) + b_i = 0$$

представляет собой прямую, а полупространство $y_i \geq 0$ — определенную полуплоскость (она отмечена штриховкой). Многогранник же X — выпуклый многоугольник.

Ясно, что решением задачи линейного программирования будет какая-то вершина многогранника X . На рис. 42 наибольшее значение F достигается в вершине P_1 , а наименьшее — в вершине P_2 , причем эти решения единственны. Задача может иметь бесчисленное множество решений, расположенных на некоторой грани (ребре), параллельной плоскости $c_1x_1 + \dots + c_nx_n = \text{const}$ (рис. 43), и может не иметь ни одного решения, если X — пустое множество (ограничения противоречивы) (рис. 44). Если допустимое множество задачи линейного программирования не ограничено, ответ на вопрос о существовании ее решения зависит от того, ограничена сверху на этом множестве целевая функция или нет. Если ограничена — задача разрешима, причем возможны те же ситуации, которые описаны выше. Если нет — решение в обычном понимании отсутствует (рис. 45).

Из приведенной геометрической интерпретации задач линейного программирования видно, что обычные методы классического математического анализа для отыскания наибольшего (наименьшего) значения функции неприменимы к рассматриваемой задаче.

Классические методы, сводя задачу к отысканию множества точек, «подозрительных на экстремум», и к сравнению значений функций в этих точках, становятся малопригодными, если число таких точек велико.

Линейная же форма (VI.21), определенная на многограннике X , заданном неравенствами (VI.22), достигает своего наибольшего (наименьшего) значения в некоторой вершине этого многогранника, так что множеством точек, «подозрительных на экстремум», является множество всех вершин многогранника X , число которых обычно бывает огромным. Основным методом решения общей задачи линейного программирования, позволяющим преодолеть эти трудности, является симплекс-метод.

§ 34. Симплекс-метод решения задач линейного программирования

Методы линейного программирования представляют собой последовательности однообразных по процедуре выполнения итераций, приводящих через конечное число шагов или в пределе к оптимальному плану задачи. Исходя из некоторого набора чисел $z^0 = (z_1^0, \dots, z_S^0)$ на итерации v , определяют вектор $z^v = (z_1^v, \dots, z_S^v)$, в котором смысле более близкий к решению задачи, чем $z^{v-1} = (z_1^{v-1}, \dots, z_S^{v-1})$. В одних методах $z^v = x^v = (x_1^v, \dots, x_n^v)$ — опорный план прямой задачи; в других $z^v = \lambda^v = (\lambda_1^v, \dots, \lambda_m^v)$ — опорный план двойственной задачи. Имеются также методы, в которых z^v не совпадает ни с x^v , ни с λ^v [4, 29].

Методы линейного программирования делятся на два широких класса — конечные и итеративные. Конечные методы позволяют получить точное решение задачи за конечное число шагов. Последовательность определяют итеративными методами, если задача сходится при $v \rightarrow \infty$ к оптимальному плану $\bar{x}^*$ задачи линейного программирования. Итеративные методы позволяют получить лишь приближенное решение задачи. При этом качество решения существенно зависит от числа проведенных итераций.

Конечные методы линейного программирования, в свою очередь, делятся на три класса в зависимости от того, используются ли для достижения оптимального плана прямая задача, двойственная задача или обе задачи двойственной пары одновременно.

Основным конечным методом является симплекс-метод, или метод последовательного улучшения плана. Термин «симплекс-метод» не отражает соответствующей вычислительной процедуры. Термин «метод последовательного улучшения плана» соответствует определяемой им процедуре решения задачи линейного программирования.

1. *Жордановы исключения.* Основу вычислительной схемы симплекс-метода для решения задачи линейного программирования составляют жордановы исключения, представляющие

собой важный аппарат линейной алгебры, позволяющей четко проводить обычно громоздкие вычисления.

Пусть рассматривается система

$$y_i = a_{i1}x_1 + \dots + a_{in}x_n, \quad (i = 1, \dots, m) \quad (\text{VI.24})$$

из m линейных форм с n независимыми переменными $x_1, \dots, x_n$. Запишем ее в виде табл. 7.

Т а б л и ц а 7

	x_1	...	x_s	...	x_n
$y_1 =$	a_{11}	...	a_{1s}	...	a_{1n}
$y_r =$	a_{r1}	...	a_{rs}	...	a_{rn}
$y_m =$	a_{m1}	...	a_{ms}	...	a_{mn}

Сверху расположены независимые переменные $x_1, \dots, x_n$ (координатные плоскости $x_1 = 0, \dots, x_n = 0$), а слева — зависимые переменные $y_1, \dots, y_m$ (плоскости $y_i \equiv a_{i1}x_1 + \dots + a_{in}x_n = 0$; $i = 1, \dots, m$).

Под шагом обыкновенного жорданова исключения, производимым над табл. 5 с разрешающим элементом $a_{rs} \neq 0$ с r -й разрешающей строкой и s -м разрешающим столбцом, будем понимать преобразование табл. 7 в новую по следующим пяти правилам: 1) разрешающий элемент a_{rs} заменяется единицей; 2) остальные элементы разрешающего s -го столбца остаются без изменений; 3) остальные элементы разрешающей r -й строки меняют лишь свои знаки; 4) элементы a'_{ij} , не принадлежащие разрешающей строке или столбцу, вычисляются по формуле

$$a'_{ij} = a_{ij}a_{rs} - a_{is}a_{rj}, \quad (i \neq r; \quad j \neq s);$$

5) все элементы новой таблицы делят на разрешающий элемент a_{rs} .

Таким образом, преобразованная таблица имеет вид (табл. 8).

Т а б л и ц а 8

	x_1	...	y_r	...	x_n
$y_1 =$	a'_{11}	...	a_{1s}	...	a'_{1n}
...	...	...	...	...	...
$x_s =$	$-a_{r1}$	...	1	...	$-a_{rn}$
...	...	...	...	...	...
$y_m =$	a'_{m1}	...	a_{ms}	...	a'_{mn}

: a_{rs} .

Если в уравнении

$$y_r = a_{r1}x_1 + \dots + a_{rs}x_s + \dots + a_{rn}x_n$$

системы (VI.24) выразить x_s через $x_1, \dots, y_n$ через x_n , подставить во все остальные уравнения и полученную новую систему записать в виде таблицы, то легко проверить, что приходим к табл. 8. Таким образом, шаг жорданова исключения над табл. 7 с разрешающим элементом a_{rs} эквивалентен перемене ролями между зависимой переменной y_r и независимой переменной x_s .

Геометрический смысл шага жорданова исключения заключается в замене старой координатной плоскости $x_s = 0$ новой $y_r = 0$ (вообще не ортогональной остальным координатным плоскостям).

Таким образом, жордановы исключения позволяют от случайно взятой декартовой системы координатных плоскостей перейти к новой, более интересной для той или другой задачи, координатной системе.

Запишем ее в виде таблицы

$$\begin{array}{c} y = \\ y_2 = \end{array} \begin{array}{|ccc|} \hline x_1 & x_2 & x_3 \\ \hline -2 & 3 & -4 \\ 1 & 2 & \boxed{2} \\ \hline \end{array}$$

Пример VI.1. Пусть задана система

$$y_1 = -2x_1 + 3x_2 - 4x_3;$$

$$y_2 = x_1 + 2x_2 + 2x_3.$$

и проведем один шаг жорданова исключения с разрешающей 2-й строкой и 3-м столбцом, т. е. поменяем ролями x_3 с y_2 . Получим

$$\begin{array}{c} y_1 = \\ x_3 = \end{array} \begin{array}{|ccc|} \hline x_1 & x_2 & y_2 \\ \hline 0 & 14 & -4 \\ -1 & -2 & 1 \\ \hline \end{array} : 2 \quad \text{т. е.} \quad \begin{array}{c} y_1 = \\ x_3 = \end{array} \begin{array}{|ccc|} \hline x_1 & x_2 & y_2 \\ \hline 0 & 7 & -2 \\ -1/2 & -1 & 1/2 \\ \hline \end{array}$$

В симплекс-методе удобнее пользоваться модифицированными жордановыми исключениями.

Система (VI.24) переписывается в виде

$$y_i \equiv -a_{i1}(-x_1) - \dots - a_{in}(-x_n) \quad (i = 1, \dots, m), \quad (\text{VI.25})$$

так что табл. 7 принимает вид (табл. 9).

Таблица 9

$$\begin{array}{c} -x_1 \dots \dots \dots -x_n \\ y_1 = \\ y_r = \\ y_m = \end{array} \begin{array}{|cccccc|} \hline \alpha_{11} & \dots & \dots & \alpha_{1s} & \dots & \alpha_{1n} \\ \alpha_{r1} & \dots & \dots & \alpha_{rs} & \dots & \alpha_{rn} \\ \alpha_{m1} & \dots & \dots & \alpha_{ms} & \dots & \alpha_{mn} \\ \hline \end{array}$$

где $\alpha_{ij} = -a_{ij}$ ($i = 1, \dots, m$; $j = 1, \dots, n$).

Шаг модифицированного жорданова исключения над табл. 9 проводится по тем же правилам, что и обыкновенное жорданово исключение, только правила 2) и 3) меняются ролями, а именно: 2) остальные элементы r -й разрешающей строки остаются без изменения; 3) остальные элементы s -го разрешающего столбца меняют только свои знаки.

Пример VI.2. Систему

$$y_1 = -2x_1 + 3x_2,$$

$$y_2 = x_1 - x_2,$$

$$y_3 = 3x_1 + 2x_2$$

запишем в виде таблицы

	$-x_1$	$-x_2$
$y_1 =$	2	-3
$y_2 =$	-1	1
$y_3 =$	-3	-2

и проведем один шаг модифицированного жорданова исключения с разрешающими первой строкой и первым столбцом. Получим

	$-y_1$	$-x_2$
$x_1 =$	$1/2$	$-3/2$
$y_2 =$	$1/2$	$-1/2$
$y_3 =$	$3/2$	$-13/2$

2. *Опорное решение**. Вернемся к задаче линейного программирования. В симплекс-методе предполагается, что все x_i неотрицательны. Форму (VI.21) и условия (VI.23) запишем в виде табл. 10.

Таблица 10

	$-x_1$	$\dots$	$-x_s$	$\dots$	$-x_n$	1
$y_1 =$	a_{11}	$\dots$	a_{1s}	$\dots$	a_{1n}	b_1
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$y_r =$	a_{r1}	$\dots$	a_{rs}	$\dots$	a_{rn}	b_r
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$y_m =$	a_{m1}	$\dots$	a_{ms}	$\dots$	a_{mn}	b_m
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$F =$	$-c_1$	$\dots$	$-c_s$	$\dots$	$-c_n$	0

* В некоторых руководствах по линейному программированию вместо термина «опорное решение» используются термины «базисное решение» или «допустимое решение».

Если же система ограничений не содержит условий неотрицательности всех x_j , т. е. некоторые из них (или все) являются свободными (могут принимать и отрицательные значения), то с помощью соответствующего числа последовательных шагов модифицированного жорданова исключения перебрасываем все свободные x_j налево таблицы и выписываем их выражения отдельно (исключаем все свободные x_j), так как они понадобятся лишь после получения решения задачи для выражения его в старых координатах. При этом в качестве разрешающего элемента берется любой отличный от нуля элемент a_{ij} (табл. 10).

Для удобства записи будем считать, что все x_j свободны и что полученная таблица имеет вид (табл. 11)

Т а б л и ц а 11

	$-y_1 \dots \dots \dots -y_n$	1
$x_1 =$	$\beta_{11} \dots \dots \dots \beta_{1n}$	β_1
$\dots$	$\dots \dots \dots \dots \dots \dots$	$\dots$
$x_n =$	$\beta_{n1} \dots \dots \dots \beta_{nn}$	β_n
$y_{n+1} =$	$\beta_{n+1,1} \dots \dots \dots \beta_{n+1,n}$	β_{n+1}
$\dots$	$\dots \dots \dots \dots \dots \dots$	$\dots$
$y_m =$	$\beta_{m1} \dots \dots \dots \beta_{mn}$	β_m
$F =$	$q_1 \dots \dots \dots q_n$	Q

Выписываем отдельно выражения для замененных $x_1, \dots, x_n$:

$$\begin{aligned}
 x_1 &= -\beta_{11}y_1 - \dots - \beta_{1n}y_n + \beta_1; \\
 &\dots \dots \dots \dots \dots \dots \\
 x_n &= -\beta_{n1} - \dots - \beta_{nn} + \beta_n
 \end{aligned}$$

и продолжаем работать лишь с оставшейся частью табл. 12,

Т а б л и ц а 12

	$-y_1 \dots \dots \dots -y_s \dots \dots \dots -y_n$	1
$y_{n+1} =$	$\beta_{n+1,1} \dots \dots \dots \beta_{n+1,s} \dots \dots \dots \beta_{n+1,n}$	β_{n+1}
$\dots$	$\dots \dots \dots \dots \dots \dots$	$\dots$
$y_r =$	$\beta_{r1} \dots \dots \dots \beta_{rs} \dots \dots \dots \beta_{rn}$	β_r
$\dots$	$\dots \dots \dots \dots \dots \dots$	$\dots$
$y_m =$	$\beta_{m1} \dots \dots \dots \beta_{ms} \dots \dots \dots \beta_{mn}$	β_m
$F =$	$q_1 \dots \dots \dots q_s \dots \dots \dots q_n$	Q

в которой переменные $y_1, \dots, y_n$ неотрицательны.

Итак, задача линейного программирования записана в виде табл. 12, где все переменные неотрицательны. Следует различать два случая: 1) все свободные члены неотрицательны; $\beta_i \geq 0$ ($i = n + 1, \dots, m$); 2) среди β_i есть отрицательные, например $\beta_r < 0$.

В первом случае, приняв в табл. 12 $y_1 = \dots = y_n = 0$, найдем $y_{n+1} = \beta_{n+1} \geq 0, \dots, y_m = \beta_m \geq 0$, т. е. получим некоторое решение системы (VI.23). Это будет вершина многогранника X , определяемая пересечением новых координатных плоскостей $y_i = 0$ ($i = 1, \dots, n$), т. е. новое начало координат. Решение системы (VI.23), соответствующее вершине многогранника X , будем называть опорным. Таким образом, в первом случае сразу получаем опорное решение и можем переходить к отысканию оптимального решения.

Во втором случае, когда, например, $\beta_r < 0$, значения $y_1 = \dots = y_n = 0$ тоже определяют точку пересечения n плоскостей, т. е. вершину, но расположенную вне многогранника X , так как в ней $y_r = \beta_r < 0$ [система (VI.23) не удовлетворяется] и плоскость $y_r = 0$ отделяет точку $y_1 = \dots = y_n = 0$ от многогранника X .

Симплекс-метод дает правило перехода от данной вершины $y_1 = \dots = y_n = 0$ к такой соседней, которая «ближе» расположена к многограннику X . Для осуществления такого перехода проводим шаг модифицированного жорданова исключения, выбирая разрешающий элемент следующим образом.

1. Выбираем строку с отрицательным заключительным членом (пусть, например, $\beta_r < 0$). Если среди коэффициентов этой строки нет отрицательных, то система (VI.23) несовместна.

2. Если среди коэффициентов рассматриваемой строки есть отрицательные, берем какой-нибудь из них (пусть $\beta_{rs} < 0$) и s -й столбец, содержащий этот коэффициент, берем в качестве разрешающего.

3. Вычисляем все неотрицательные отношения $\beta_i/\beta_{is} \geq 0$ свободных членов к отличным от нуля коэффициентам разрешающего столбца, находим среди них наименьшее, пусть β_{i_0}/β_{i_0s} , и строку i_0 -ю берем в качестве разрешающей, так что β_{i_0s} — разрешающий элемент.

Пример VI.3. Найти какое-нибудь опорное решение системы неравенств:

$$y_1 = -x_1 + 3x_2 - 3 \geq 0,$$

$$y_2 = -3x_1 - x_2 + 12 \geq 0,$$

$$y_3 = 3x_1 - 2x_2 + 6 \geq 0,$$

$$y_4 = x_1 + 2x_2 - 3 \geq 0$$

$$x_2 \geq 0.$$

Составим таблицу

	$-x_1$	$-x_2$	1
$y_1 =$	1	-3	-3
$y_2 =$	3	1	12
$y_3 =$	-3	2	6
$y_4 =$	-1	-2	-3

и исключим лишь координату x_1 (так как $x_2 \geq 0$), получим

	$-y_4$	$-x_2$	1
$y_1 =$	1	-5	-6
$y_2 =$	3	-5	3
$y_3 =$	-3	8	15

$$x_1 = y_4 - 2x_2 + 3.$$

В первой строке есть отрицательный свободный член -6 и отрицательный коэффициент -5 (во втором столбце). Поэтому выбираем второй столбец в качестве разрешающего и составляем неотрицательные отношения свободных членов к соответствующим коэффициентам второго столбца: $-6/-5$; $3/-5$; $15/8$.

Меньшее из них — первое, поэтому коэффициент -5 будет разрешающим. После шага жорданова исключения получим таблицу

	$-y_4$	$-y_1$	1
$x_2 =$	$-1/5$	$-1/5$	$6/5$
$y_2 =$	2	-1	9
$y_3 =$	$-7/5$	$8/5$	$27/5$

где все свободные члены неотрицательны и можно найти опорное решение системы.

Полагаем $y_4 = y_1 = 0$. Тогда $x_2 = 6/5$; $x_1 = -12/5 + 3 = 3/5$ — одно из решений заданной системы.

Дадим теперь геометрическую интерпретацию приведенного выше правила (рис. 46). Стремясь попасть в какую-нибудь вершину, принадлежащую многограннику $\bar{X}$, следует из полученной вершины $y_1 = \dots y_n = 0$ двигаться по направлению к X , т. е. к отделяющей плоскости $y_r = 0$. Двигаемся всегда по некоторому «верхнему» ребру, образованному пересечением $n-1$ координатных плоскостей, например

$$y_1 = \dots = y_{s-1} = y_{s+1} = \dots = y_n = 0.$$

В параметрическом виде уравнение этого ребра можно записать так:

$$y_1 = \dots = y_{s-1} = 0; \quad y_s = t > 0; \quad y_{s+1} = \dots = y_n = 0. \quad (\text{VI.26})$$

Уклонение любой точки ребра (VI.32) от отделяющей плоскости $y_r = 0$ выразится формулой

$$y_r = -\beta_{rs}t + \beta_r,$$

и так как мы стремимся продвинуться к этой плоскости, то ребро должно быть выбрано так, чтобы

$$|-\beta_{rs}t + \beta_r| < |\beta_r|,$$

т. е.

$$-(-\beta_{rs}t + \beta_r) < -\beta_r$$

$$\text{или } \beta_{rs}t < 0.$$

Но поскольку $t > 0$, то $\beta_{rs} < 0$. Этим объясняется правило выбора ведущего столбца.

Выбрав ребро, двигаемся по нему до встречи с одной из плоскостей $y_i = 0$ ($i > n$). В точке встречи (пересечения) с этой плоскостью, пусть $y_{i_0} = 0$, получим

$$0 = y_{i_0} = -\beta_{i_0s} + \beta_{i_0}.$$

Следовательно, $\beta_{i_0s} \neq 0$ и

$$t = \frac{\beta_{i_0}}{\beta_{i_0s}} \leq \frac{\beta_i}{\beta_{is}},$$

потому что речь идет о первой встреченной плоскости. Этим объясняется правило выбора ведущей строки.

Чтобы исчерпать вопрос улучшения опорного решения, остановимся на следующем положении: не может ли случиться, что в процессе улучшения нарушится опорность решения, т. е. что при переходе, по правилу улучшения, к следующей таблице в заключительном столбце появятся отрицательные элементы?

Ответ на эти вопросы содержится в следующей теореме.

Т е о р е м а (VI.1). При переходе от таблицы к таблице, по правилу улучшения, опорное решение остается опорным.

Допустим, что мы, по правилу улучшения, переходим от табл. 13 к табл. 14 с ведущим элементом $\beta_{rs} > 0$.

При этом элементы заключительного столбца β_i переходят в элементы β'_i заключительного столбца новой табл. 14. Надо показать, что если все $\beta_i \geq 0$ ($i = 1, 2, \dots, m$), то и все $\beta'_i \geq 0$ ($i = 1, 2, \dots, m$). Возьмем сначала элемент ведущей строки β'_r ; он будет неотрицательным, так как

$$\beta'_r = \frac{\beta_r}{\beta_{rs}} \geq 0.$$

Здесь числитель — неотрицательный, а знаменатель — положительный.

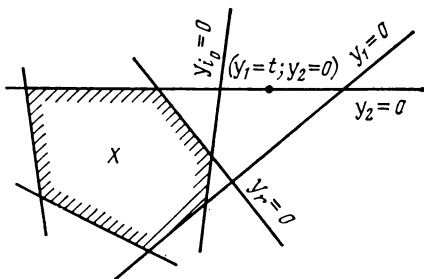


Рис. 46. Геометрическая интерпретация выбора ведущего столбца и ведущей строки

Таблица 13

$-x_s$	1
..... β_{1s}	β_1
..... β_{2s}	β_2
..... $\vdots$	$\vdots$
..... $\vdots$	$\vdots$
..... $\boxed{\beta_{rs}}$	β_r
..... $\vdots$	$\vdots$
..... $\vdots$	$\vdots$
..... β_{ms}	β_m

Таблица 14

1
..... β'_1
..... β'_2
..... $\vdots$
..... $\vdots$
..... β'_r
..... $\vdots$
..... $\vdots$
..... β'_m

Возьмем теперь какой-нибудь другой элемент, он получается по правилу, изложенному в п. 1 § 34, т. е.

$$\beta'_i = \beta_i - \frac{\beta_{is}\beta_r}{\beta_{rs}}.$$

При этом возможны два случая:

$$1) \beta_{is} \leq 0,$$

так как $\beta_r/\beta_{rs} \geq 0$; $\beta_i \geq 0$, то $\beta'_i \geq \beta_i \geq 0$;

$$2) \beta_{is} > 0;$$

тогда представим β'_i иначе

$$\beta'_i = \beta_{is} \left(\frac{\beta_i}{\beta_{is}} - \frac{\beta_r}{\beta_{rs}} \right),$$

Так как $\beta_r/\beta_{rs} = \min \beta_r/\beta_{is}$, то число в скобке неотрицательно, а множитель $\beta_{is} > 0$, значит, и в этом случае $\beta'_i \geq 0$.

Таким образом, теорема доказана.

Отметим, что последовательные переходы от одного опорного плана к следующему происходят до тех пор, пока будет получено решение задачи или установлена неограниченность ее линейной формы. Поскольку число опорных планов, совпадающее с числом вершин многогранного множества условий задачи, конечно, то один из этих случаев обязательно наступит через конечное число шагов.

3. *Оптимальное решение.* Допустим, что получено некоторое опорное решение системы (VI.23) в табл. 15, где все свободные переменные (β_i) положительны.

Теперь можно приступить к отысканию оптимального решения, т. е. к отысканию того опорного решения (вершины многогранника X), которое максимизирует линейную форму:

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \max.$$

	$-y_1$	...	$-y_s$	...	$-y_n$	1
$y_{n+1} =$	$\beta_{n+1,1}$	...	$\beta_{n+1,s}$	...	$\beta_{n+1,n}$	β_{n+1}
...	...	...	...	...	...	...
$y_r =$	β_{r1}	...	β_{rs}	...	β_{rn}	β_r
...	...	...	...	...	...	...
$y_m =$	β_{m1}	...	β_{ms}	...	β_{mn}	β_m
$F =$	q_1	...	q_s	...	q_n	Q

Шаг симплекс-метода для отыскания оптимального решения означает движение по такому ребру, выходящему из полученного опорного решения (начала координат), вдоль которого целевая функция возрастает (рис. 47). Двигаться следует до достижения новой (соседней) вершины, в которой целевая функция имеет, очевидно, большее значение.

При движении, например, по ребру

$$y_1 = \dots = y_{s-1} = 0; \quad y_s = t > 0; \quad y_{s+1} = \dots = y_n = 0 \quad (\text{VI.27})$$

получим

$$F = -q_s t + Q$$

и для возрастания F необходимо, чтобы

$$-q_s t + Q > Q$$

Q — значение F в вершине $y_1 = \dots = y_n = 0$). Отсюда — $q_s t > 0$, т. е. $q_s < 0$. Следовательно, если в F -строке нет отрицательных коэффициентов, то не существует ребра, при движении вдоль которого форма F увеличивается, т. е. вершина (начало координат) $y_1 = \dots = y_n = 0$ — искомое оптимальное решение и $\max F = Q$.

Если же в F -строке есть отрицательный коэффициент, например $q_s < 0$, то в качестве ведущего берем s -й столбец, т. е. из исходной вершины двигаемся вдоль ребра (VI.36) до соседней вершины или, что то же самое, до первой встречи с некоторой плоскостью $y_{i_0} = 0$.

В точке встречи получим

$$0 = y_{i_0} = -\beta_{i_0 s} t + \beta_{i_0}$$

и

$$t = \frac{\beta_{i_0}}{\beta_{i_0 s}} = \min_{\beta_{i s} > 0} \frac{\beta_{i_0}}{\beta_{i s}}.$$

Теперь можно сформулировать следующее правило выбора разрешающего коэффициента при отыскании оптимального решения.

В качестве разрешающего берем столбец, содержащий отрица-

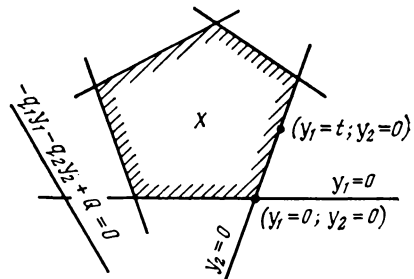


Рис. 47. Шаг симплекс-метода при отыскании оптимального решения

тельный коэффициент F -строки (например, s -й столбец); отбираем все положительные коэффициенты этого столбца (если такие имеются), делим на них соответствующие свободные члены, сравниваем полученные отношения и в качестве разрешающего берем тот из коэффициентов, для которого отношение имеет наименьшее значение (если их окажется больше одного, то берем любой из них).

После шага жорданова исключения с разрешающим коэффициентом, выбранным по только что сформулированному правилу, знак q_s изменится на противоположный, а новый коэффициент q_s окажется уже положительным (но при этом в F -строке могут появиться новые отрицательные коэффициенты). Продолжим так процесс до тех пор, пока в F -строке все коэффициенты окажутся неотрицательными и оптимальное решение получено, либо придем к случаю, когда в F -строке окажется отрицательным коэффициент, но в столбце, содержащем его, все коэффициенты — неположительны. В этом случае форма F не ограничена сверху.

Если оптимальное решение не достигнуто, а в F -строке имеется несколько отрицательных коэффициентов q_j , то в качестве разрешающего столбца берут обычно тот, в котором стоит коэффициент, наибольший по абсолютной величине:

$$|q_s| = \max |q_j|; \quad q_j < 0.$$

Основанием для этого служит следующее.

Пусть разрешающим будет коэффициент β_{rs} . В результате шага жорданова исключения свободным членом в F -строке будет

$$Q_s = \frac{Q\beta_{rs} - q_s\beta_r}{\beta_{rs}} = Q + \left(-q_s \frac{\beta_r}{\beta_{rs}} \right).$$

Поскольку $(-q_s) > 0$ и $\beta_r/\beta_{rs} > 0$, выражение в скобке в этом уравнении всегда будет положительно. А так как значение целевой функции всегда равно свободному члену Q , то выражение в скобке представляет собой дополнение к Q , который получается в результате сделанного шага.

Очевидно, чем большее приращение будет получать целевая функция на каждом шаге, тем меньше потребуется шагов.

На основании вышесказанного сформулируем две теоремы.

Т е о р е м а (VI.2) (признак оптимальности решения). *Если на каком-то шаге жорданова исключения оказалось, что все коэффициенты заключительного столбца и все коэффициенты заключительной строки неотрицательны (≥ 0), то это означает, что достигнут максимум целевой функции; при этом оптимальное опорное решение получится, если положить верхние переменные равными нулю, а боковые переменные — равными соответствующим коэффициентам заключительного столбца. Кроме того, заключительный угловой коэффициент Q представляет собой максимальное значение в целевой функции ($\max F$).*

Т е о р е м а (VI.3). *Если на каком-то шаге жорданова исключения в столбце над отрицательным коэффициентом заключительной строки все коэффициенты оказались неположительными, то целевая*

функция может принимать любые большие значения ($\max F = \infty$).

Действительно, если в s -м столбце $q_s < 0$ и $\beta_{is} \leq 0$ ($i = 1, 2, \dots, m$), то при любом большом значении $x_s > 0$ все значения $y_i = \beta_{is}(-x_s) + \beta_i \geq 0$, так что ограничения (VI.24) не нарушаются, значение же целевой функции $F = q_s(-x_s) + Q \rightarrow \infty$ при $x_s \rightarrow \infty$. Таким образом, если на каком-то шаге жордановых исключений оказалась такая ситуация, как в условии теоремы (VI.3), то вычисления обрываются, так как целевая функция не ограничена. Одновременно это свидетельствует о том, что область ограничений является неограниченной (не многогранником).

Пример VI.4. Максимизировать линейную форму

$$F = 2x_1 + 3x_2$$

при выполнении ограничений

$$y_1 = -x_1 + 2x_2 + 6 \geq 0,$$

$$y_2 = -x_1 - x_2 + 14 \geq 0,$$

$$y_3 = 2x_1 - 2x_2 + 5 \geq 0,$$

$$y_4 = 2x_1 - x_2 + 3 \geq 0.$$

Запишем в виде таблицы

	$-x_1$	$-x_2$	1
$y_1 =$	1	-2	6
$y_2 =$	1	1	14
$y_3 =$	-2	2	5
$y_4 =$	-2	1	3
$F =$	-2	-3	0

Исключив координаты x_1 и x_2 , получим таблицу:

	$-x_1$	$-x_2$	1
$y_3 =$	$\boxed{4/3}$	$2/3$	$67/3$
$y_4 =$	1	1	23
$F =$	$-1/3$	$7/3$	$92/3$

и выражение для исключенных координат

$$x_1 = -y_1 + 2x_2 + 6;$$

$$x_2 = 1/3y_1 - 1/3y_2 + 8/3.$$

Мы получили после исключения x_1 и x_2 свободные члены неотрицательные, следовательно, $y_1 = y_2 = 0$ дает опорное решение заданной системы ограничений.

Переходим к отысканию оптимального решения. В F -строке содержится отрицательный элемент $-1/3$. Столбец под $-y_1$, содержащий этот коэффи-

циент, выбирают в качестве разрешающего. Для выбора разрешающей строки составляем отношение свободных членов к соответствующим положительным коэффициентам ведущего столбца:

$$\frac{67}{3} : \frac{4}{3}; \quad 23 : 1.$$

Наименьшее из этих отношений $67/4$. Поэтому $4/3$ берем в качестве разрешающего. После шага жорданова исключения получим таблицу

	$-y$	y_2	1
$y_1 =$			$67/4$
$y_4 =$			$25/4$
$F =$	$1/4$	$5/2$	$145/4$

Теперь в F -строке нет больше отрицательных коэффициентов. Задача решена.

Из последней таблицы находим $\max F = \frac{145}{4}$; он достигается в точке $y_3 = y_2 = 0$. При этом $y_1 = 67/4$, $y_4 = 25/4$. Подставляя эти выражения для x_1 и x_2 , получаем $x_1 = 23/4$, $x_2 = 33/4$.

4. *Случай вырождения.* По правилу улучшения опорного решения разрешающей строкой следует взять ту, в которой достигается минимум отношения $\min \beta_i/\beta_{is}$ ($\beta_{is} > 0$). Если этот минимум достигается только в одной строке, то разрешающая строка определяется однозначно. Если же минимум достигается одновременно в нескольких строках, то в качестве разрешающей можно взять любую из них и выбор разрешающей строки уже не является однозначным. Именно при такой ситуации очередной шаг жорданова исключения приведет к вырождению, т. е. в заключительном столбце появятся нули.

Допустим, минимум соответствующих отношений достигается в p -й и r -й строках, так что

$$\min \frac{\beta_i}{\beta_{is}} = \frac{\beta_p}{\beta_{ps}} = \frac{\beta_r}{\beta_{rs}} \quad (\beta_{is} > 0).$$

Тогда, если мы выберем в качестве разрешающей p -ю строку, то в новой таблице вместо β_r появится нуль, так как

$$\beta'_r = \beta_r - \frac{\beta_{rs}\beta_p}{\beta_{ps}} = \beta_{rs} \left(\frac{\beta_r}{\beta_{rs}} - \frac{\beta_p}{\beta_{ps}} \right) = 0.$$

Если минимум соответствующих отношений достигается в k -строках, то в дальнейшем появится $k - 1$ нулей.

Более подробно этот вопрос рассмотрен в работах [7, 4].

§ 35. Двойственные задачи линейного программирования

В ходе развития математического программирования выяснилось, что при исследовании многих экстремальных задач (выявление признаков решения, установление тех или иных оценок искомого экстремального значения, изучение вопросов существования

и единственности решения, а также при разработке эффективных численных методов) оказывается полезным наряду с интересующей нас экстремальной задачей рассматривать другую, в известном смысле двойственную экстремальную задачу. Это привело к созданию общего принципа двойственности, позволяющего с единой точки зрения исследовать широкий круг задач максимального программирования.

Необходимо отметить, что наряду с принципом двойственности в математическом программировании важную роль играет другой универсальный подход к исследованию задач, который условно может быть назван принципом возможных направлений. При этом результаты, полученные с помощью одного из указанных принципов, как правило, удается переформулировать в терминах второго принципа.

Неявное использование наряду с исследуемой экстремальной задачей другой, двойственной к ней, можно проследить у П. Л. Чебышева, Х. Валле-Пуссена, С. Н. Бернштейна и других ученых в работах, посвященных задаче наилучшего приближения в той или иной форме заданной функции обобщенными полиномами (линейными комбинациями фиксированных функций). В явном виде задача, двойственная к этой задаче (точнее, к эквивалентной ей экстремальной задаче), была сформулирована впервые в работе М. Г. Крейна. Последняя систематически использовалась затем при исследовании различных экстремальных и вариационных проблем в теории функций и функциональном анализе.

Все же основное влияние на формирование принципа двойственности оказали традиционные разделы математического программирования. Идеи двойственности играли важную роль в первых работах фон Неймана по теории игр и математическим моделям экономики и Л. В. Канторовича по линейному и нелинейному программированию [20].

В настоящее время теорию двойственности можно считать в основном завершенной. Помимо детально разработанных схем, пригодных для исследования отдельных важных классов задач, имеется ряд достаточно универсальных схем принципа двойственности, позволяющих с единых позиций исследовать весьма широкий круг задач математического программирования. Во всех этих схемах важную роль играют теоремы отделимости выпуклых множеств. При этом в большинстве схем для облегчения их практического использования теоремы отделимости (как правило, классическая теорема Эйдельгайта) закладываются в саму схему. В частности, это относится к общей схеме, основанной на идеях Минковского — Фенхеля, которая интенсивно разрабатывается в последние годы [1, 29].

В теории линейного программирования также важную роль играет то обстоятельство, что каждой задаче линейного программирования соответствует некоторая двойственная задача. Если исходная задача линейного программирования, называемая прямой задачей, является задачей на максимум:

найти

$$\max_{\bar{x}} F = \bar{c}\bar{x} \quad (\text{VI.28})$$

при условии

$$\bar{A}\bar{x} \leq \bar{b}; \quad \bar{x} \geq 0.$$

то двойственная задача представляет собой задачу на минимум: найти

$$\min G = \bar{\lambda} \bar{b} \quad (\text{VI.29})$$

при условии, что $\bar{\lambda}\bar{A} \geq \bar{c}$; $\bar{\lambda} \geq 0$.

Здесь $\bar{\lambda}$ есть вектор-строка:

$$\bar{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m) \quad (\text{VI.30})$$

Запишем двойственную задачу линейного программирования в развернутом виде:

найти

$$\min_{\lambda_1, \lambda_2, \dots, \lambda_m} G(\lambda_1, \lambda_2, \dots, \lambda_m) = b_1\lambda_1 + b_2\lambda_2 + \dots + b_m\lambda_m$$

при умови

[illegible]

Сопоставим задачи (VI.20) и (VI.31). С одной стороны, общим для этих задач является то, что в каждой из них отыскивается экстремум линейной функции при условии, что переменные неотрицательны и удовлетворяют системе линейных неравенств. В обеих задачах используются одни и те же параметры, а именно: матрица $\bar{A}$, вектор-столбец $\bar{b}$ и вектор-строка $\bar{c}$. Число ограничений неравенств в этих задачах равно $m + n$; каждая из задач имеет геометрическую интерпретацию.

С другой стороны, в **прямой задаче** определяются значения n переменных — компонент **вектор-столбца** x , а в двойственной задаче — m переменных, являющихся компонентами вектор-строки $\bar{\lambda}$. В исходной задаче определяется **максимум**, а в двойственной — **минимум**: знаки неравенств в **этих задачах** различны: константы ограничений одной из задач являются коэффициентами целевой функции другой. Если применить к двойственной задаче те же преобразования, какие были сделаны для прямой задачи, то мы вновь приходим к исходной задаче, т. е. задача, двойственная к двойственной задаче, представляет собой исходную задачу.

Представим двойственную задачу линейного программирования в виде табл. 16.

	x_1	x_2	x_n	-1	
λ_1	a_{11}	a_{12}	$\dots$	a_{1n}	$b_1 \leq 0$
λ_2	a_{21}	a_{22}	$\dots$	a_{2n}	$b_2 \leq 0$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
λ_m	a_{m1}	a_{m2}	$\dots$	a_{mn}	$b_m \leq 0$
-1	c_1	c_2	$\dots$	c_n	$0 = F(\max)$
	≥ 0	≥ 0	≥ 0	$= G(\min)$	

Прочитаем эту таблицу **слева направо**, умножая элементы, стоящие внутри клетки, на **соответствующие** переменные из верхней строки и складывая эти **произведения**. В результате получим задачу максимизации. Читая **сверху вниз** и **умножая** коэффициенты, стоящие внутри клетки, на **соответствующие** переменные из левого столбца и складывая эти **произведения**, приходим к задаче минимизации. Нулевой коэффициент в **нижнем** правом углу таблицы можно в принципе заменить **любой другой** константой, вычитаемой из обеих целевых функций.

Оказывается, решив **одну из задач** двойственной пары, мы автоматически получим и **решение второй задачи**. Это устанавливается теоремами, которые изложены ниже.

Приведем теперь **интерпретацию двойственных задач** в терминах модели планирования **производства однородного продукта**.

Решение **прямой задачи (VI.28)** определяет план производства $\bar{x} = (x_1, x_2, \dots, x_n)$, который позволяет получить максимальное количество продукта при имеющихся в наличии ресурсах $\bar{b}$ ($b_1, \dots, b_m$) и производственных факторов.

Задаче (VI.28) ставится в **соответствие** двойственная задача вида (VI.29). В терминах задачи (VI.28) ей может быть дано следующее экономическое истолкование. Решение двойственной задачи определяет оценки $(\lambda_1, \dots, \lambda_m)$ **производственных факторов** — меры полезности единицы каждого фактора для данного производства при технологических способах и ограниченных ресурсах, обусловленных постановкой **прямой задачи**. Оценки производственных факторов отвечают **идеализированной замкнутой модели** производства, в которой связи с **внешним миром** строго фиксированы условиями задачи (VI.28). Оценки факторов являются, таким образом, внутренними оценками в **условиях** данного производства при выбранном наборе **технологических способов** и фиксированных ограничениях на **производственные факторы**. Изменение условий производства, в частности **изменение запасов** используемых факторов или появление **новой технологии**, приводит к необходимости переоценки производственных факторов.

Оценки λ_i измеряются в единицах ценности выпускаемого продукта. Ценность продукта определяется условиями, внешними по отношению к данному производству. При рационально выбранных оценках λ_i производственных факторов оценка всех затрат определяется технологическим способом производства j ; она не может быть меньше ценности c_j произведенного по этой технологии продукта. Всякое производство требует соответствующих затрат. Это естественное соображение и фиксируется условиями (VI.29) двойственной задачи.

Величина $\min G = \bar{\lambda} \bar{b}$ решения двойственной задачи представляет собой суммарную оценку производственных факторов, которыми располагает предприятие. Рациональная система оценок λ_i соответствует минимальной суммарной оценке производственных факторов.

Метод множителей Лагранжа позволяет проанализировать основные свойства двойственных задач, поскольку двойственные переменные можно считать множителями Лагранжа для прямой задачи. Рассмотрим задачу максимизации: найти

$$\max_{\bar{x}} F = \bar{c} \bar{x} \text{ при условии } \bar{A} \bar{x} \leq \bar{b}, \bar{x} \geq 0 \quad (\text{VI.32})$$

как прямую задачу линейного программирования. Пусть функция Лагранжа определена следующим образом:

$$L(\bar{x}, \bar{\lambda}) = \bar{c} \bar{x} + \bar{\lambda} (\bar{b} - \bar{A} \bar{x}) = \bar{c} \bar{x} + \bar{\lambda} \bar{b} - \bar{\lambda} \bar{A} \bar{x}. \quad (\text{VI.33})$$

Тогда, согласно теореме Куна—Таккера, $\bar{x}^*$ будет решением задачи (VI.32), если существует такая вектор-строка $\bar{\lambda}^*$, что в $(\bar{x}^*, \bar{\lambda}^*)$ выполняются условия Куна — Таккера, имеющие вид

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{x}} &= \bar{c} - \bar{\lambda} \bar{A} \leq 0; \\ \frac{\partial L}{\partial \bar{x}} \bar{x} &= (\bar{c} - \bar{\lambda} \bar{A}) \bar{x} = 0; \\ \bar{x} &\geq 0. \\ \frac{\partial L}{\partial \bar{\lambda}} &= \bar{b} - \bar{A} \bar{x} \geq 0. \\ \bar{\lambda} \frac{\partial L}{\partial \bar{\lambda}} &= \bar{\lambda} (\bar{b} - \bar{A} \bar{x}) = 0; \\ \bar{\lambda} &= 0. \end{aligned} \right\} \quad (\text{VI.34})$$

Пусть теперь в качестве прямой задачи рассматривается задача минимизации: найти

$$\min_{\bar{\lambda}} G = \bar{\lambda} \bar{b} \text{ при условии } \bar{\lambda} \bar{A} \geq \bar{c}; \bar{\lambda} \geq 0. \quad (\text{VI.35})$$

Определим функцию Лагранжа как

$$L(\bar{x}, \bar{\lambda}) = \bar{\lambda} \bar{b} + (\bar{c} - \bar{\lambda} \bar{A}) \bar{x} = \bar{\lambda} \bar{b} + \bar{c} \bar{x} - \bar{\lambda} \bar{A} \bar{x}. \quad (\text{VI.36})$$

Тогда из теоремы Куна — Таккера следует, что $\bar{\lambda}^*$ будет решением, если существует такой вектор-столбец $\bar{x}^*$, что в $(\bar{x}^*, \bar{\lambda}^*)$ выполняются условия Куна — Таккера, имеющие вид

$$\left. \begin{aligned} \frac{\partial L}{\partial \bar{\lambda}} &= \bar{b} - \bar{A} \bar{x} \geq 0; \\ \bar{\lambda} \frac{\partial L}{\partial \bar{\lambda}} &= \bar{\lambda} (\bar{b} - \bar{A} \bar{x}) = 0; \\ \bar{\lambda} &\geq 0. \\ \frac{\partial L}{\partial \bar{x}} &= (\bar{c} - \bar{\lambda} \bar{A}) \leq 0; \\ \frac{\partial L}{\partial \bar{x}} \bar{x} &= (\bar{c} - \bar{\lambda} \bar{A}) \bar{x} = 0; \\ \bar{x} &\geq 0. \end{aligned} \right\} \quad (\text{VI.37})$$

Таким образом, функции Лагранжа и условия Куна — Таккера в обеих задачах совершенно одни и те же. На приведенных выше условиях базируются основные теоремы двойственности линейного программирования.

Т е о р е м а (VI.4). *Для того чтобы задача линейного программирования имела решение, необходимо и достаточно, чтобы допустимые множества как прямой, так и двойственной задачи были не пусты. Это основная теорема существования.*

Для доказательства того, что в обеих задачах существуют допустимые векторы, рассмотрим ограничения-неравенства в двойственных задачах

$$\bar{A} \bar{x} \leq \bar{b}, \quad (\text{VI.38})$$

$$\bar{\lambda} \bar{A} \geq \bar{c}. \quad (\text{VI.39})$$

Так как вектор $\bar{\lambda}$ неотрицателен, то

$$\bar{\lambda} \bar{A} \bar{x} \leq \bar{\lambda} \bar{b} = \bar{G}(\bar{\lambda}), \quad (\text{VI.40})$$

и так как вектор $\bar{x}$ неотрицателен, то

$$F(\bar{x}) = \bar{c} \bar{x} \leq \bar{\lambda} \bar{A} \bar{x}. \quad (\text{VI.41})$$

Следовательно, если $\bar{x}$ и $\bar{\lambda}$ — допустимые векторы, то

$$F(\bar{x}) \leq G(\bar{\lambda}), \quad (\text{VI.42})$$

т. е. значения целевой функции в задаче максимизации не превосходят значений целевой функции в двойственной задаче минимизации.

Это — фундаментальное неравенство теории двойственности в линейном программировании.

Экономический смысл теоремы в применении к задаче производственного планирования состоит в том, что, каков бы ни был производственный план и каков бы ни был план оценок производственных ресурсов, общая стоимость последних не будет оценена ниже, чем доход от продукции, которую можно получить по любому из производственных планов. Ясно, то такой позиции будет придерживаться предприятие, если ему предложат оценить свой запас производственных ресурсов.

Предположим, что в каждой из этих задач существуют допустимые векторы $\bar{x}^0$ и $\bar{\lambda}^0$. Допустимое множество прямой задачи содержит вектор $\bar{x}^0$ и, следовательно, не пусто. Целевая функция в прямой задаче ограничена, так как для любого допустимого вектора

$$F(\bar{x}) \leq G(\bar{\lambda}^0). \quad (\text{VI.43})$$

Следовательно, прямая задача имеет решение. Если при любом допустимом векторе $\bar{\lambda}$

$$F(\bar{x}^0) \leq G(\bar{\lambda}), \quad (\text{VI.44})$$

то целевая функция в двойственной задаче ограничена снизу. А так как допустимое множество в этой задаче содержит вектор, то и двойственная задача имеет решение.

Покажем теперь, что если решение задачи линейного программирования существует, то допустимые множества как прямой, так и двойственной задачи не пусты.

Пусть $\bar{x}^*$ является решением задачи максимизации. Тогда очевидно, что прямая задача имеет хотя бы один допустимый вектор, а именно, вектор $\bar{x}^*$. Известно, что целевая функция вогнута, а функции ограничений выпуклы (линейная функция является одновременно и выпуклой и вогнутой) и что выполняются условия регулярности ограничений. Поэтому, если $\bar{x}^*$ является решением задачи максимизации, то, согласно теореме Куна — Таккера, существует вектор $\bar{\lambda}^*$, удовлетворяющий условиям (VI.34). Следовательно,

$$\bar{\lambda}^* \bar{A} \geq \bar{c}; \quad \bar{\lambda}^* \geq 0, \quad (\text{VI.45})$$

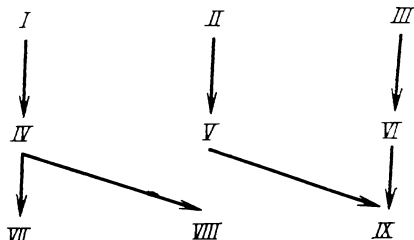
так что $\bar{\lambda}^*$ есть допустимый вектор. Теорема доказана.

Итак, решение существует тогда — и только тогда, когда каждая из двойственных задач имеет допустимые векторы. Если допустимое множество одной из задач пусто, то либо допустимое множество двойственной задачи пусто, либо ее целевая функция не ограничена. Вообще любые две двойственные задачи либо имеют допустимые векторы и, следовательно, по теореме существования, обладают решениями, либо лишь одна из двойственных задач имеет допустимый вектор, причем ее целевая функция не ограничена, либо ни одна из двойственных задач не имеет допустимых векторов (рис. 48).

Т е о р е м а (VI.5). Первая теорема двойственности: *некоторый допустимый вектор тогда — и только тогда, является решением задачи линейного программирования, когда существует такой допу-*

Рис. 48. Альтернативные варианты, возникающие при решении задач линейного программирования:

I — существуют допустимые векторы и для прямой и для двойственной задачи; II — допустимый вектор существует в прямой задаче, двойственная задача не обладает допустимым вектором; III — ни прямая, ни двойственная задачи не имеют допустимых векторов; IV — существует решение прямой задачи; V — целевая функция прямой задачи не ограничена; VI — допустимое множество прямой задачи пусто; VII — единственное решение; VIII — не единственное решение; IX — нет решений



стимый вектор двойственной задачи, при которых значения целевых функций обеих задач на этих векторах равны.

Докажем сначала, что если $\bar{x}^*$ — решение задачи максимизации, то существует такой $\bar{\lambda}^*$ — допустимый вектор двойственной задачи, что значения целевых функций на этих векторах равны. Рассмотрим условия Куна — Таккера (VI.43). Если вектор $\bar{\lambda}^*$ удовлетворяет этим условиям, то, согласно (VI.54),

$$\bar{\lambda}^* \bar{A} \geq \bar{c}; \quad \bar{\lambda}^* \geq 0, \quad (\text{VI.46})$$

т. е. этот вектор будет допустимым. Кроме того, из равенств

$$\left. \begin{aligned} (\bar{c} - \bar{\lambda}^* \bar{A}) \bar{x}^* &= 0; \\ \bar{\lambda}^* (\bar{b} - \bar{A} \bar{x}^*) &= 0 \end{aligned} \right\} \quad (\text{VI.47})$$

следует, что

$$F(\bar{x}^*) = \bar{c} \bar{x}^* = \bar{\lambda}^* \bar{A} \bar{x}^* = \bar{\lambda}^* \bar{b} = G(\bar{\lambda}^*), \quad (\text{VI.48})$$

т. е. значения целевых функций совпадают. Для того чтобы доказать эту часть теоремы в случае, когда прямой задачей является задача минимизации, достаточно провести аналогичные рассуждения, используя условия (VI.37).

Таким образом, критерий оптимальности решений пары двойственных задач состоит в том, что допустимые решения $\bar{x}^*$ и $\bar{\lambda}^*$ пары двойственных задач оптимальны тогда — и только тогда, когда соответствующие значения целевых функций равны, т. е. когда

$$F(\bar{x}^*) = G(\bar{\lambda}^*). \quad (\text{VI.49})$$

Установим экономический смысл первой теоремы двойственности. Теорема означает, что оптимальный план производства можно составить лишь в том случае, если удастся привести в соответствие всем производственным факторам рациональные оценки. И, наоборот, производственные факторы допускают естественную оценку лишь при наличии оптимального плана производства. При этом оценка продукта, полученного при реализации оптимального плана, совпадает с суммарной оценкой имеющихся ресурсов (производственных факторов). При любом другом плане оценка производственного продукта будет ниже оценки затрат. Такое производство убыточно: его возможности не используются полностью.

Решение двойственной задачи — оптимальные оценки ресурсов — называют объективно обусловленными оценками. Они были введены Л. В. Канторовичем. Это оценки продуктов, ресурсов, работ, вытекающие из условий решаемой оптимизационной задачи. Их называют также двойственными оценками, разрешающими множителями, множителями Лагранжа и целым рядом других терминов.

Вернемся к задаче распределения ресурсов, рассмотренной в § 32 п. 2, и зададимся вопросом, какова с точки зрения предприятия ценность имеющихся в его распоряжении ресурсов. При решении этого вопроса будем иметь в виду, что ресурсы, которые предприятие не может полностью использовать, имеют для него очень низкую ценность в том смысле, что предприятие не будет согласно нести даже небольшие расходы на увеличение запасов этих ресурсов. Так, дорогое оборудование, не используемое в технологическом процессе, имеет для предприятия нулевую ценность. Наибольшую ценность будут, очевидно, иметь те ресурсы, которые в наибольшей степени обеспечивают выпуск продукции, а следовательно, и доход предприятия, и на увеличение запасов которых предприятие согласно нести значительные расходы.

Поэтому можно считать, что каждый вид ресурсов обладает некоторой «теневой ценой» [19], определяющей ценность данного вида ресурсов для предприятия с точки зрения дохода от реализации выпускаемой продукции и зависящей от наличного запаса этого вида ресурсов и потребности в нем для выпуска продукции.

Если предприятие по каким-либо причинам ограничивается одним технологическим процессом, требующим больших затрат некоторого вида ресурсов, запасы которого ограничены, то теневая цена этого ресурса будет велика. Однако установленные в соответствии с этим технологическим процессом теневые цены не будут наилучшими, так как введение других технологических процессов позволит более рационально использовать все запасы ресурсов. Следовательно, существуют оптимальные теневые цены, которые соответствуют максимальному доходу предприятия, т. е. оптимальному распределению ресурсов. Оптимальные теневые цены и есть объективно обусловленные оценки [19]. Если исходить из объективно обусловленных оценок, то в оптимальном плане производства будут приняты интенсивности использования технологических способов, обеспечивающие максимально возможный при заданных ограничениях на ресурсы выпуск продукта.

Значение объективно обусловленных оценок состоит в том, что:

- 1) они показывают, насколько изменится значение критерия оптимальности в соответствующей прямой задаче линейного программирования при приращении данного вида ресурсов на единицу. Оценки выступают, следовательно, как мера дефицитности ресурсов и продукции, как мера влияния ограничений на функционал;

- 2) их можно использовать, далее, как инструмент определения эффективности отдельных технологических способов с позиций общего оптимума;

3) это инструмент балансирования суммарных затрат и результатов;

4) они являются также показателями взаимозаменяемости ресурсов относительно заданного критерия.

Т е о р е м а (VI.6). *Вторая теорема двойственности: для того чтобы допустимые векторы $\bar{x}^*$, $\bar{\lambda}^*$ являлись решениями двойственных задач, необходимо и достаточно, чтобы они удовлетворяли условиям дополнительной нежесткости*

$$\left. \begin{aligned} (\bar{c} - \bar{\lambda}^* \bar{A}) \bar{x}^* &= 0; \\ \bar{\lambda}^* (\bar{b} - \bar{A} \bar{x}^*) &= 0. \end{aligned} \right\} \quad (\text{VI.50})$$

Необходимость следует непосредственно из условий Куна — Таккера, а достаточность — из первой теоремы двойственности. Действительно, предположим, что $\bar{x}^*$, $\bar{\lambda}^*$ — допустимые векторы. Тогда из условий (VI.50) вытекает

$$F(\bar{x}^*) = \bar{c} \bar{x}^* = \bar{\lambda}^* \bar{A} \bar{x}^* = \bar{\lambda}^* \bar{b} = G(\bar{\lambda}^*). \quad (\text{VI.51})$$

Следовательно, $\bar{x}^*$ и $\bar{\lambda}^*$ являются искомыми решениями.

Таким образом, другой критерий оптимальности устанавливается второй теоремой двойственности, т. е. условиями дополнительной нежесткости (VI.50).

Выпишем условия дополняющей нежесткости в развернутом виде. Они приводят к равенствам

$$\left. \begin{aligned} \left(c_j - \sum_{i=1}^m a_{ij} \lambda_i^* \right) x_j^* &= 0; \quad j = 1, 2, \dots, n; \\ \lambda_i^* \left(b_i - \sum_{j=1}^n a_{ij} x_j^* \right) &= 0; \quad i = 1, 2, \dots, m. \end{aligned} \right\} \quad (\text{VI.52})$$

Сопоставив эти выражения с ограничениями в прямой и в двойственной задачах, получим:

$$\left. \begin{aligned} \lambda_i^* \geq 0, \text{ но } \lambda_i^* = 0, \text{ если } \sum_{j=1}^n a_{ij} x_j^* < b_i; \\ \sum_{j=1}^n a_{ij} x_j^* \leq b_i, \text{ но } \sum_{j=1}^n a_{ij} x_j^* = b_i, \text{ если } \lambda_i^* = 0; \\ x_j^* \geq 0, \text{ но } x_j^* = 0, \text{ если } \sum_{i=1}^m a_{ij} \lambda_i^* > c_j; \\ \sum_{i=1}^m a_{ij} \lambda_i^* \geq c_j, \text{ но } \sum_{i=1}^m a_{ij} \lambda_i^* = c_j, \text{ если } x_j^* > 0; \end{aligned} \right\} \begin{aligned} & i = 1, 2, \dots, m, \\ & j = 1, 2, \dots, n. \end{aligned} \quad (\text{VI.53})$$

Следовательно, если в оптимальной точке некоторое ограничение выполняется как строгое неравенство, то соответствующая двойственная переменная равна нулю, а если в этой точке некоторая переменная принимает положительное значение, то соответствующее ограничение-неравенство в двойственной задаче выполняется как равенство. Из решения двойственной задачи видно, какие переменные прямой задачи обращаются в оптимальной точке в нуль и какие из ограничений-неравенств в этой точке выполняются как равенства.

Объясним экономический смысл второй теоремы двойственности на двойственных парах условий (VI.53) с номерами i . Если в оптимальном плане производственный фактор i используется не пол-

ностью $\left(\sum_{j=1}^n a_{ij}x_j < b_i\right)$, естественно принять его оценку $\lambda_i = 0$.

Запасы фактора i превышают потребности в нем, поэтому он не представляет ценности для производства. Некоторое сокращение ресурса i не уменьшит возможностей производства. Однако, как только излишек производственного фактора i будет изъят, он приобретет определенную ценность для производства. Дальнейшее сокращение его запасов ведет к уменьшению выпуска продукции.

Интерпретируем теперь вторую теорему двойственности на двойственных парах условий (VI.53) с номерами j . Технологический способ производства с номером j используется в оптимальном плане ($x_j > 0$) только в том случае, если оценка соответствующего ему

выпуска продукта совпадает с оценкой затрат $\left(c_j = \sum_{i=1}^m a_{ij}\lambda_i\right)$.

В случае, когда при некотором технологическом способе расходы превышают доходы, использование этого способа лишено смысла ($x_i = 0$).

Если исходить из объективно обусловленных оценок, то в оптимальном плане будет соблюден принцип рентабельности — будут приняты только те технологические способы производства, при которых расходы не превысят доходы.

Приведем несколько примеров решения двойственных задач.

Пример VI.5. Найти минимум целевой функции

$$F = 0,5x_1 + 0,6x_2 + 0,4x_3 + 0,2x_4 + 0,3x_5 \quad (\text{VI.54})$$

при условиях

$$\left. \begin{aligned} x_1 \geq 0; \quad x_2 \geq 0; \quad x_3 \geq 0; \quad x_4 \geq 0; \quad x_5 \geq 0; \\ x_3 + x_4 \geq 500; \\ \left. \begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &\geq 1000; \\ 3x_1 \quad \quad \quad + x_5 &\geq 200; \\ x_2 + 2x_4 + x_5 &\geq 400. \end{aligned} \right\} \end{aligned} \right\} \quad (\text{VI.55})$$

После решения задачи симплекс-методом получено следующее оптимальное решение:

$$x_1 = \frac{200}{3}; \quad x_2 = 0; \quad x_3 = \frac{1000}{3}; \quad x_4 = 200; \quad x_5 = 0. \quad (\text{VI.56})$$

Совокупность неизвестных (VI.56) удовлетворяет ограничениям (VI.55). При этом второе, третье и четвертое ограничения являются равенствами, а первое ограничение — строгим неравенством $1600/3 > 500$. Значение целевой функции для допустимого решения (VI.56) $F = 620 : 3$.

Составим условие двойственной задачи. Найти максимум целевой функции

$$G = 500\lambda_1 + 1000\lambda_2 + 200\lambda_3 + 400\lambda_4 \quad (\text{VI.57})$$

при условиях

$$\left. \begin{aligned} \lambda_1 &\geq 0; \lambda_2 \geq 0; \lambda_3 \geq 0; \lambda_4 \geq 0; \\ 2\lambda_2 + 3\lambda_3 &\leq 0,5; \\ \lambda_2 + \lambda_4 &\leq 0,6; \\ \lambda_1 + 2\lambda_2 &\leq 0,4; \\ \lambda_1 + \lambda_2 + 2\lambda_4 &\leq 0,2; \\ \lambda_3 + \lambda_4 &\leq 0,3. \end{aligned} \right\} \quad (\text{VI.58})$$

(Переменная λ_1 соответствует первому ограничению неравенства системы VI.55), которое при оптимальном решении (VI.56) является строгим неравенством. Следовательно, по второй теореме двойственности, должно быть $\lambda_1 = 0$. По той же теореме, первое, третье и четвертое ограничения в системе (VI.58) должны быть равенствами, так как соответствующие им переменные x_1 , x_3 и x_4 положительны: $x_1 = 200/3 > 0$; $x_3 = 100/3 > 0$; $x_4 = 200 > 0$. Таким образом, учитывая, что $\lambda_1 = 0$, имеем:

$$2\lambda_2 + 3\lambda_3 = 0,5;$$

$$2\lambda_2 = 0,4;$$

$$\lambda_2 + 2\lambda_4 = 0,2.$$

Откуда $\lambda_2 = 1/5$; $\lambda_3 = 1/30$; $\lambda_4 = 0$.

Подставим найденные значения λ_2 , λ_3 и λ_4 во второе и пятое ограничения (VI.67):

$$0,2 + 0 < 0,6;$$

$$1/30 + 0 < 0,3,$$

которые выполняются как строгие неравенства и поэтому соответствующие этим ограничениям переменные x_2 и x_5 должны равняться нулю. (Они, действительно, равны нулю.)

Таким образом, допустимое решение (VI.56) и допустимое решение двойственной задачи $\lambda_1 = 0$; $\lambda_2 = 1/5$; $\lambda_3 = 1/30$; $\lambda_4 = 0$ удовлетворяют всем условиям второй теоремы двойственности и поэтому являются оптимальными. Значит, задачи (VI.54), (VI.55) решены правильно. Можно еще раз убедиться в правильности решения этой задачи, применив первую теорему двойственности, по которой должно быть совпадение значений целевых функций на допустимых решениях прямой (VI.54) и сопряженной (VI.57) задач.

Действительно, $F = 620/3$ и

$$G = 500,0 + 1000 \cdot 1/5 + 200 \cdot 1/30 + 400 \cdot 0 = 620/3.$$

Пример VI.6 [3]. Предприятие производит продукцию трех видов — А, Б, В. Уровень выпуска продукции лимитируется имеющимися в распоряжении предприятия ресурсами — сырьем, материалами и оборудованием двух различных групп. Исходные данные приведены в табл. 17.

Составим неравенства, отражающие требование — не превосходить имеющийся объем ресурсов по каждому их виду. Получим следующую систему ограничений:

$$5x_1 + 7x_2 + 4x_3 \leq 24;$$

$$10x_1 + 5x_2 + 20x_3 \leq 80;$$

$$5x_1 + 2x_2 + x_3 \leq 10;$$

$$2x_1 + x_2 + x_3 \leq 6.$$

Ресурсы		Норма затрат ресурсов на единицу продукции		
вид	объем	А	Б	В
Сырье, кг	24	5	7	4
Материалы, кг	80	10	5	20
Оборудование I группы, единиц	10	5	2	1
Оборудование II группы, единиц	6	2	1	1
Прибыль на единицу продукции, руб.		18	12	8
Неизвестные, искомая программа выпуска продукции		x_1	x_2	x_3

При соблюдении этих условий необходимо максимизировать общую сумму прибыли, т. е. величину

$$F = 18x_1 + 12x_2 + 8x_3 \rightarrow \max.$$

Оптимальное решение данной задачи: $x_1 = 1$; $x_2 = 1$; $x_3 = 3$; $F = 54$ руб.

В двойственной задаче требуется минимизировать

$$G = 24\lambda_1 + 80\lambda_2 + 10\lambda_3 + 6\lambda_4 \rightarrow \min$$

при ограничениях

$$5\lambda_1 + 10\lambda_2 + 5\lambda_3 + 2\lambda_4 \geq 18,$$

$$7\lambda_1 + 5\lambda_2 + 2\lambda_3 + \lambda_4 \geq 12,$$

$$4\lambda_1 + 20\lambda_2 + \lambda_3 + \lambda_4 \geq 8.$$

Оптимальным решением двойственной задачи является совокупность двойственных оценок соответствующих ресурсов:

сырья	$\lambda_1 = 0,834$
материалов	$\lambda_2 = 0$
единицы оборудования I группы	$\lambda_3 = 1,5$
единицы оборудования II группы	$\lambda_4 = 3,166$

Величина двойственной оценки того или иного вида ресурсов показывает, насколько возросло бы значение целевой функции, если бы объем данного вида ресурсов увеличился на единицу. Так, в нашем примере, если бы фонд времени работы оборудования I группы увеличился на единицу (1 прибор-ч), то прибыль возросла бы на 1,5 руб. Соответственно при увеличении фонда времени работы оборудования II группы на единицу — общая прибыль возрастает на 3,166 руб., т. е. в нашем примере увеличение оборудования II группы более выгодно.

Двойственная оценка материалов (λ_2) равна нулю. Это говорит о том, что материалы в оптимальном плане используются не полностью, поэтому дальнейшее их увеличение не повлияет на оптимальный план выпуска продукции и сумму прибыли.

На основании объективно обусловленных оценок можно определять своеобразные «нормы заменяемости ресурсов». Так, в нашем примере норма заменяемости ресурсов оборудования I и II групп составляет $3,166 : 1,5 = 2,1$. Здесь речь идет не об абсолютной заменяемости ресурсов, а об относительной, т. е. заменяемости с точки

зрения конечного эффекта производства и лишь в конкретных условиях данной задачи.

Известно, что не все задачи линейного программирования можно решить с помощью графического метода, поэтому для решения некоторых из них можно использовать графический анализ двойственной задачи.

Пример VI.7. Решить задачу линейного программирования с помощью графического анализа двойственной задачи:

$$\begin{aligned} 3x_1 + x_2 - 4x_3 - x_4 &\leq -3; \\ -2x_1 - 4x_2 - x_3 + x_4 &\leq -3; \\ x_1, x_2, x_3, x_4 &\geq 0; \\ F = -4x_1 - 18x_2 - 30x_3 - 5x_4 &\rightarrow \max. \end{aligned}$$

Задача содержит четыре переменных, поэтому не может быть решена графически. Двойственная к ней задача содержит две переменные и может быть решена графически. Последняя имеет вид

$$\begin{aligned} 3\lambda_1 - 2\lambda_2 &\geq -4; & -4\lambda_1 - \lambda_2 &\geq -30; \\ \lambda_1 - 4\lambda_2 &\geq -18; & -\lambda_1 + \lambda_2 &\geq -5; \\ \lambda_1, \lambda_2 &\geq 0; & G = -3\lambda_1 - 3\lambda_2 &\rightarrow \min. \end{aligned}$$

Решая ее графически, получаем $\bar{\lambda}^* = (6, 6)$ и $G_{\min} = -36$. В соответствии с первой теоремой двойственности исходная задача имеет оптимальное решение и $F_{\max} = -36$. Исходя из $\bar{\lambda}^* = (6, 6)$ определим $\bar{x}^*$ с помощью второй теоремы двойственности. Для $\bar{x}^* = (x_1, x_2, x_3, x_4, x_5, x_6)$ должны удовлетворяться равенства

$$\begin{aligned} x_1(-4 - 3\lambda_1 + 2\lambda_2) &= 0; & x_3(-18 - \lambda_1 + 4\lambda_2) &= 0; \\ x_2(-30 + 4\lambda_1 + \lambda_2) &= 0; & x_4(-5 + \lambda_1 - \lambda_2) &= 0. \end{aligned}$$

Поскольку в первом и четвертом равенствах при $\lambda_1 = \lambda_2 = 6$ выражения в скобках не равны нулю, то, по второй теореме двойственности, соответствующие переменные исходной задачи равны нулю: $x_1 = 0$; $x_4 = 0$.

Рассматривая исходную задачу как двойственную, по второй теореме двойственности, получаем:

$$\begin{aligned} \lambda_1(-3 - 3x_1 - x_2 + 4x_3 + x_4) &= 0; \\ \lambda_2(-3 + 2x_1 + 4x_2 + x_3 - x_4) &= 0. \end{aligned}$$

Принимая во внимание, что $\lambda_1 = 6 > 0$ и $\lambda_2 = 6 > 0$, для $\bar{x}^*$ должны удовлетворяться равенства:

$$-3 - 3x_1 - x_2 + 4x_3 + x_4 = 0; \quad -3 + 2x_1 + 4x_2 + x_3 - x_4 = 0.$$

Подставляя $x_1 = 0$ и $x_4 = 0$, получим:

$$\begin{aligned} -3 - x_2 + 4x_3 &= 0; \\ -3 + 4x_2 + x_3 &= 0. \end{aligned}$$

Из этой системы уравнений находим $x_2 = 9/17$ и $x_3 = 15/17$, следовательно, $\bar{x}^* = (0, 9/17, 15/17, 0)$ и $F = -36$.

В заключение отметим следующее. Как было указано в § 34, конечные методы линейного программирования делятся на три класса в зависимости от достижения оптимального плана с помощью прямой задачи, двойственной задачи или задачи двойственной пары одновременно. Метод первого класса (симплекс-метод) для решения

задачи линейного программирования, когда задана прямая задача, описан в § 34.

Основным представителем второго класса конечных методов линейного программирования является двойственный метод. Он представляет собой нечто иное, как симплекс-метод, примененный к решению двойственной задачи. При этом вычислительная процедура формируется в терминах прямой задачи. С каждым шагом двойственного метода уточняется план двойственной задачи. Каждый из опорных планов двойственной задачи можно рассматривать как приближенную систему оценок условий прямой задачи. Это значит, что с каждым шагом двойственного метода уточняются оценки условий исходной задачи. Поэтому двойственный метод называют также методом последовательного уточнения оценок. Вектор z^v в двойственном методе представляет собой опорный план $\lambda^v = (\lambda_1^v, \dots, \lambda_m^v)$ двойственной задачи.

С каждым вектором λ^v в пространстве переменных исходной задачи связан некоторый вектор $x^v = (x_1^v, \dots, x_n^v)$, называемый псевдопланом задачи.

Основной представитель третьего класса конечных методов линейного программирования, в которых для достижения оптимального плана используются обе задачи двойственной пары, это метод последовательного сокращения невязок [4, 6, 10].

Глава VII

ТРАНСПОРТНАЯ ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Впервые транспортная задача была рассмотрена А. Н. Толстым в 1930 г. (случай для двух поставщиков), в 1939 г. Л. В. Канторовичем был разработан метод разрешающих множителей и указана возможность решения ее этим методом.

В последнее время в нашей стране уделяется большое внимание изучению и разработке вопросов применения транспортной задачи в планировании. Так, А. Л. Лурье и Ю. А. Олейник разработали метод решения транспортной задачи, получивший название метода приближения условно-оптимальными планами. А. Л. Брудно, развивая этот метод, предложил модификацию, названную методом дифференциальных рент [19, 4].

§ 36. Математическая постановка задачи

Перепишем уравнения транспортной задачи (§ 32 п. Б) в развернутом виде ($m = n = 3$). Найти, x_{ij} , если

$$F = c_{11}x_{11} + c_{12}x_{12} + \dots + c_{33}x_{33} \rightarrow \min \quad (\text{VII.1})$$

при условиях

$$\left. \begin{array}{llll} x_{11} + x_{12} + x_{13} & & & = a_1 \\ & x_{21} + x_{22} + x_{23} & & = a_2 \\ & & x_{31} + x_{32} + x_{33} & = a_3 \\ x_{11} & + x_{21} & + x_{31} & = b_1 \\ & x_{12} & + x_{22} & + x_{32} = b_2 \\ & & x_{13} & + x_{23} & + x_{33} = b_3 \end{array} \right\} \quad (\text{VII.2})$$
$$x_{11} \geq 0; x_{12} \geq 0; x_{13} \geq 0; x_{21} \geq 0; x_{22} \geq 0; x_{23} \geq 0; x_{31} \geq 0; x_{32} \geq 0; x_{33} \geq 0.$$

Отметим, что ограничения транспортной задачи имеют особенность: коэффициенты при неизвестных равны 0 или 1. Транспортная задача относится к числу специфических задач линейного программирования, для решения которых разработаны специальные методы, более простые, чем симплекс-метод решения общей задачи линейного программирования.

Имеется всего $m + n$ уравнений, где m — число пунктов отправления, n — число пунктов назначения. Количество неизвестных

x_{ij} ($i = 1, \dots, m, j = 1, \dots, n$) равно $m \times n$. Вследствие того, что

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

сумма первых m уравнений

$$\sum_{i=1}^m x_{ij} = a_i$$

равна сумме последних n уравнений

$$\sum_{j=1}^n x_{ij} = b_j.$$

Таким образом, число линейно независимых уравнений только $(m + n - 1)$, поэтому опорное решение должно иметь не более $(m + n - 1)$ отличных от нуля неизвестных.

План x_{ij} ($i = 1, \dots, m; j = 1, \dots, n$) транспортной задачи, имеющий не более $m + n - 1$ отличных от нуля величин x_{ij} , называется опорным планом. Если в опорном плане число отличных от нуля величин x_{ij} равно $m + n - 1$, план называется невырожденным, а если меньше, то план называется вырожденным.

Отметим, что на практике условие

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j. \quad (\text{VII.3})$$

может не выполняться, т. е. может быть, что

$$\sum_{i=1}^m a_i > \sum_{j=1}^n b_j \quad (\text{VII.4})$$

или

$$\sum_{i=1}^m a_i < \sum_{j=1}^n b_j. \quad (\text{VII.5})$$

Формулировка транспортной задачи в виде (VII.1), (VII.3) называется закрытой транспортной задачей. Формулировка транспортной задачи в виде (VII.4) или (VII.5) называется открытой транспортной задачей.

В случае (VII.4) вводят еще одного фиктивного потребителя B_{n+1} с потребностью b_{n+1} , который как бы потребляет весь избыток ресурсов, т. е.

$$b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j > 0.$$

В случае (VII.5) вводится фиктивный поставщик A_{m+1} с наличием ресурсов a_{m+1} , определяемых как

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

Величины $c_{i,n+1}$ и $c_{m+1,j}$ принимаются равными нулю. Следовательно, любую задачу транспортного типа можно привести к закрытому виду.

§ 37. Двойственная задача. Условия оптимальности

Мы отмечали, что с каждой задачей линейного программирования тесным образом связана двойственная ей задача. Рассмотрим двойственную задачу к транспортной задаче линейного программирования. Для этого поставим в соответствие каждому уравнению задачи двойственную переменную. Все уравнения транспортной задачи распадаются на две группы: уравнения, относящиеся к строкам транспортной задачи, и уравнения, относящиеся к столбцам (так, уравнение $x_{21} + x_{22} + x_{23} = a_2$ соответствует второй строке, а уравнение $x_{11} + x_{21} + x_{31} = b_1$ соответствует первому столбцу. Поэтому каждой строке и каждому столбцу транспортной задачи отвечает своя двойственная переменная. Двойственную переменную, соответствующую строке i ($i = 1, 2, \dots, m$), обозначим u_i , а двойственную переменную, соответствующую столбцу j ($j = 1, 2, \dots, n$), обозначим v_j . Вид ограничений двойственной задачи сначала разберем на примере с двумя пунктами отправления и тремя пунктами назначения:

		v_1	v_2	v_3	
		B_1	B_2	B_3	
u_1	A_1	c_{11}	c_{12}	c_{13}	a_1
u_2	A_2	c_{21}	c_{22}	c_{23}	a_2
		b_1	b_2	b_3	

т. е. имеется задача — минимизировать

$$c_{11}x_{11} + c_{12}x_{12} + c_{13}x_{13} + c_{21}x_{21} + c_{22}x_{22} + c_{23}x_{23}$$

при условиях

$$\left. \begin{array}{l} u_1 \left\{ \begin{array}{l} x_{11} + x_{12} + x_{13} \\ x_{21} + x_{22} + x_{23} \end{array} \right. = a_1 \\ u_2 \left\{ \begin{array}{l} x_{21} + x_{22} + x_{23} \\ x_{11} + x_{12} + x_{13} \end{array} \right. = a_2 \end{array} \right\} \quad (\text{VII.6})$$

$$\left. \begin{array}{l} v_1 \left\{ \begin{array}{l} x_{11} + x_{21} \\ x_{12} + x_{22} \\ x_{13} + x_{23} \end{array} \right. = b_1 \\ v_2 \left\{ \begin{array}{l} x_{12} + x_{22} \\ x_{13} + x_{23} \end{array} \right. = b_2 \\ v_3 \left\{ \begin{array}{l} x_{13} + x_{23} \end{array} \right. = b_3 \end{array} \right\} \quad (\text{VII.7})$$

$$x_{11} \geq 0; x_{12} \geq 0; x_{13} \geq 0; x_{21} \geq 0; x_{22} \geq 0; x_{23} \geq 0.$$

Здесь первому уравнению, или первой строке транспортной таблицы, соответствует двойственная переменная u_1 , второму уравнению, или второй строке транспортной таблицы, соответствует двойственная переменная u_2 , третьему уравнению, или первому столбцу

транспортной таблицы — переменная v_1 , четвертому уравнению, или второму столбцу таблицы, соответствует переменная v_2 и, наконец, пятому уравнению, или третьему столбцу транспортной таблицы, — двойственная переменная v_3 . Составим ограничения двойственной задачи. Первому столбцу первой строки системы уравнений (V II.6), (V II.7) соответствует неравенство

$$u_1 + v_1 \leq c_{11}, \quad (\text{VII.8})$$

второму столбцу первой строки системы соответствует неравенство

$$u_1 + v_2 \leq c_{12}, \quad (\text{VII.9})$$

третьему столбцу первой строки соответствует неравенство

$$u_1 + v_3 \leq c_{13} \quad (\text{VII.10})$$

и т. д. Последнему столбцу (третьему столбцу второй строки) соответствует неравенство

$$u_2 + v_3 \leq c_{23}. \quad (\text{VII.11})$$

Требуется максимизировать

$$a_1 u_1 + a_2 u_2 + b_1 v_1 + b_2 v_2 + b_3 v_3. \quad (\text{VII.12})$$

На знак переменных u_1, u_2, v_1, v_2, v_3 ограничений не наложено. Заметим, что есть некоторая неоднозначность в формулировке двойственной задачи, связанная со следующим обстоятельством. Условия прямой задачи не изменятся, если правую и левую части любого уравнения умножить на -1 . Умножим на -1 обе части первых двух уравнений. Тогда вместо неравенства (V II.8) получим неравенство

$$v_1 - u_1 \leq c_{11}. \quad (\text{VII.13})$$

Вместо неравенств (V II.9)—(V II.11) будем иметь неравенства

$$\left. \begin{aligned} v_2 - u_1 &\leq c_{12}; \\ v_3 - u_1 &\leq c_{13}; \\ v_3 - u_2 &\leq c_{23}. \end{aligned} \right\} \quad (\text{VII.14})$$

Максимизация выражения (V II.12) заменяется максимизацией

$$b_1 v_1 + b_2 v_2 + b_3 v_3 - a_1 u_1 - a_2 u_2.$$

Это два эквивалентных варианта одной и той же двойственной задачи, причем общепринято формулировать задачу во втором варианте. Нетрудно убедиться, что и в общем случае при любых m и n имеют место неравенства, аналогичные неравенствам (V II.13), (V II.14). Итак, двойственную задачу можно сформулировать следующим образом. Для каждой строки $i = 1, 2, \dots, m$ транспортной таблицы и каждого столбца $j = 1, 2, \dots, n$ (для каждого пункта отправления $i = 1, \dots, m$ и пункта назначения $j = 1, \dots, n$) найти числа $u_1, \dots, u_m; v_1, \dots, v_n$, удовлетворяющие ограничениям

$$v_j - u_i \leq c_{ij}; \quad i = 1, \dots, m; \quad j = 1, \dots, n$$

и максимизирующие

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

Искомые числа называются *потенциалами*. Потенциалы $u_1, \dots, u_m; v_1, \dots, v_n$ позволяют ответить на следующий, очень важный вопрос. Предположим, что нам известен некоторый опорный план x_{ij}^* . Оптимален ли он? Справедливо следующее условие оптимальности, доказанное Л. В. Канторовичем.

Т е о р е м а (VI.1). План x_{ij}^* является оптимальным тогда — и только тогда, когда для каждой строчки $i = 1, \dots, m$ и каждого столбца $j = 1, \dots, n$ найдутся такие потенциалы $u_1^*, \dots, u_m^*; v_1^*, \dots, v_n^*$, при которых

$$v_j^* - u_i^* \leq c_{ij} \text{ при } x_{ij}^* = 0; \quad (\text{VII.15})$$

$$v_j^* - u_i^* = c_{ij} \text{ при } x_{ij}^* > 0. \quad (\text{VII.16})$$

Докажем эту теорему на основании теоремы (VI.5).

1. Пусть для плана x_{ij}^* справедливы условия (VII.15), (VII.16). Докажем, что план x_{ij}^* является оптимальным. Очевидно, числа $u_i^*, v_j^*, i = 1, \dots, m; j = 1, \dots, n$, удовлетворяющие условиям (VII.15), (VII.16), удовлетворяют ограничениям двойственной задачи. Пусть имеется любое решение $u_i, v_j (i = 1, \dots, m; j = 1, \dots, n)$ двойственной задачи, т. е.

$$v_j - u_i \leq c_{ij} \quad (\text{VII.17})$$

и произвольный план x_{ij} . Умножим обе части неравенства (VII.17) на x_{ij} и просуммируем по $i = 1, \dots, m$ и $j = 1, \dots, n$. Получим

$$\sum_{i=1}^m \sum_{j=1}^n (v_j - u_i) x_{ij} \leq \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}. \quad (\text{VII.18})$$

Преобразуем левую часть этого неравенства

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n (v_j - u_i) x_{ij} &= \sum_{i=1}^m \sum_{j=1}^n v_j x_{ij} - \\ &- \sum_{i=1}^m \sum_{j=1}^n u_i x_{ij} = \sum_{j=1}^n v_j \sum_{i=1}^m x_{ij} - \sum_{i=1}^m u_i \sum_{j=1}^n x_{ij} = \sum_{j=1}^n v_j b_j - \sum_{i=1}^m u_i a_i. \end{aligned}$$

Итак, из неравенства (VII.17) следует, что для любого решения u_i, v_j двойственной задачи имеем

$$\sum_{j=1}^n b_j v_j - \sum_{i=1}^m a_i u_i \leq \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij},$$

т. е. любое значение целевой функции двойственной задачи не превышает любого значения целевой функции прямой задачи (аналогичное утверждение имеет место и для общей задачи линейного про-

граммирования см. § 35, гл. VI). Поэтому, чтобы убедиться в оптимальности плана x_{ij}^* , достаточно показать, что

$$\sum_{j=1}^n b_j v_j^* - \sum_{i=1}^m a_i u_i^* = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^*. \quad (\text{VII.19})$$

Имеем

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* &= \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* - \left(\sum_{j=1}^n b_j v_j^* - \sum_{i=1}^m a_i u_i^* \right) + \\ &+ \left(\sum_{j=1}^n b_j v_j^* - \sum_{i=1}^m a_i u_i^* \right). \end{aligned} \quad (\text{VII.20})$$

Выражение

$$\begin{aligned} &\sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* - \left(\sum_{j=1}^n v_j^* b_j - \sum_{i=1}^m a_i u_i^* \right) = \\ &= \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* - \left(\sum_{i=1}^m \sum_{j=1}^n v_j^* x_{ij}^* - \sum_{i=1}^m \sum_{j=1}^n u_i^* x_{ij}^* \right) = \\ &= \sum_{i=1}^m \sum_{j=1}^n (c_{ij} - v_j^* + u_i^*) x_{ij}^*. \end{aligned} \quad (\text{VII.21})$$

Вследствие уравнений (V II.15) и (V II.16)

$$\sum_{i=1}^m \sum_{j=1}^n (c_{ij} - v_j^* + u_i^*) x_{ij}^* = 0,$$

поэтому из равенства (V II.20) следует нужное нам равенство (V II.19).

2. Предположим теперь, что план x_{ij}^* оптимальный, и докажем, что найдутся числа u_i^* , v_j^* ($i = 1, \dots, m$; $j = 1, \dots, n$), удовлетворяющие условиям (V II.15) и (V II.16). Из того, что существует оптимальный план прямой задачи, согласно теореме (V I.5), следует, что существует и решение двойственной задачи. Обозначим это решение $\bar{u}_i$, $\bar{v}_j$ ($i = 1, \dots, m$; $j = 1, \dots, n$) и докажем, что оно удовлетворяет условиям (V II.15) и (V II.16). Согласно теореме (VII.5), имеем

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* = \sum_{j=1}^n b_j \bar{v}_j - \sum_{i=1}^m a_i \bar{u}_i. \quad (\text{VII.22})$$

Предположим, что не для всех $\bar{u}_i$, $\bar{v}_j$ ($i = 1, \dots, m$; $j = 1, \dots, n$) справедливы условия (V II.15) и (V II.16), и покажем, что в этом случае не может иметь место равенство (V II.22). Имеем

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* &= \left[\sum_{i=1}^m \sum_{j=1}^n c_{ij} - x_{ij}^* - \left(\sum_{j=1}^n b_j \bar{v}_j - \sum_{i=1}^m a_i \bar{u}_i \right) \right] + \\ &+ \left(\sum_{j=1}^n b_j \bar{v}_j - \sum_{i=1}^m a_i \bar{u}_i \right). \end{aligned}$$

Произведя над выражением в квадратных скобках преобразования, аналогичные (VII.21), получим:

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* = \left[\sum_{i=1}^m \sum_{j=1}^n (c_{ij} - v_j + u_i) x_{ij}^* \right] + \left(\sum_{j=1}^n b_j \bar{v}_j - \sum_{i=1}^m a_i \bar{u}_i \right). \quad (\text{VII.23})$$

Так как $c_{ij} - \bar{v}_j + \bar{u}_i \geq 0$ и если хотя бы для некоторой пары $(i, j) = (s, t)$ не выполняются условия (VII.15) и (VII.16), т. е. $(c_{st} - \bar{v}_t + \bar{u}_s) x_{st}^* > 0$, то из уравнения (VII.23) следует, что равенство (VII.22), действительно, не может иметь места. Итак, теорема доказана.

§ 38. Способы построения опорного плана

Транспортные задачи решаются вручную и на ЭВМ, причем первым способом решают задачи сравнительно небольшого размера. При решении задач большое значение имеет задание исходного плана. От того, каким будет исходный план, зависит число итераций, а следовательно, и время, необходимое для получения оптимального решения. Особенно это важно при решении задач вручную. Рассмотрим несколько способов построения начального плана.

При решении задач на ЭВМ наиболее часто применяется метод северо-западного угла (табл. 18).

Т а б л и ц а 18

Откуда	Куда								Ресурсы
	B_1		B_2		B_3		B_4		
A_1	14	70		38		24		92	14
A_2	16	58	4	18		56		72	20
A_3		19	18	10	8	100		30	26
A_4		3		36	7	121	34	8	41
Потребности	30		22		15		34		

Т а б л и ц а 19

Откуда	Куда				Ресурсы
	B_1	B_2	B_3	B_4	
A_1	70	38	24	92	14
A_2	58	18	56	72	20
A_3	19	10	100	30	26
A_4	3	36	121	8	41
Потребности	30	22	15	34	

Построение исходного плана начинается с левого верхнего угла матрицы. Распределение ресурсов первого поставщика осуществляется так, что сначала удовлетворяются потребности первого потребителя, затем второго и т. д. до полного распределения ресурсов. После распределения ресурсов первого поставщика переходят ко второму и распределяют подобным же образом его ресурсы. Так же распределяют ресурсы остальных поставщиков. В результате получается характерная для данного метода ступенчатая фигура, начинающаяся в верхнем левом углу.

При построении исходного плана перевозок методом северо-западного угла не обращают внимания на стоимости, и план в общем случае будет далек от оптимального. Поэтому применение данного метода для построения начального опорного плана при решении транспортных задач вручную нецелесообразно, так как число итераций для получения оптимального плана будет больше, чем если бы начальный план был построен с учетом стоимостей. При решении задач на ЭВМ, учитывая их быстроедействие, можно идти на увеличение числа итераций, зато программа построения исходного плана будет проще, чем при учете стоимостей.

С п о с о б н а и м е н ь ш е й с т о и м о с т и. Построение начального плана способом наименьшей стоимости рассмотрим на примере (табл. 19).

В матрице стоимостей ищем минимальный элемент. В данном случае он находится в клетке (A_4, B_1) . В эту клетку помещаем максимально возможную перевозку. В данном случае 30 единиц. Первый столбец при дальнейшем просмотре исключается. В оставшейся части матрицы снова находим минимальный элемент. Теперь

Т а б л и ц а 20

Откуда	Куда								Ресурсы
	B_1		B_2		B_3		B_4		
A_1		70		38	*	24		92	14
					*	14			
A_2		58	*	18	1	56	19	72	20
A_3		19	*	10		100		30	26
			*	22			4		
A_4	*	3		36		121	*	8	41
	*	30					11		
Потребности	30		22		15		34		

он находится в клетке (A_4, B_4) . Помещаем в эту клетку оставшиеся ресурсы поставщика A_4 — 11 единиц. Строку 4 исключаем, как исчерпавшую свои ресурсы. В оставшейся части матрицы минимальный элемент находится в клетке (A_3B_2) . В эту клетку помещаем 22 единицы и исключаем столбец 2. В таком порядке процесс повторяется до окончательного распределения всего объема перевозок.

Начальный план, построенный способом наименьшей стоимости, ближе к оптимальному по сравнению с планом построенным методом северо-западного угла. Следовательно, для решения вручную потребуется меньше времени, но при построении оптимального плана, особенно в матрицах большого объема, трудность заключается в нахождении минимального элемента.

Способ двойного предпочтения. Этот способ очень удобен при решении транспортных задач вручную, он дает наилучшие результаты. В связи со сложной логикой построения начального плана способ двойного предпочтения не оправдывает затраты времени на программирование и машинное решение.

В матрице стоимости сначала находим минимальный элемент в каждой строке. Клетки, имеющие минимальные элементы, отмечаем знаком «*». Затем ищем минимальные стоимости в столбцах и те клетки, в которых они находятся, отмечаем также знаком «*». В клетки, имеющие два знака «*», помещаем максимально возможное количество единиц груза. В клетки с одним знаком «*» и в не отмеченные знаком «*», но с возможно меньшей стоимостью распределяем остальное количество груза. В табл. 20 показан план, построенный методом двойного предпочтения.

Т а б л и ц а 21

Откуда	Куда								Ресурсы	Столбец-разность
	B_1		B_2		B_3		B_4			
A_1	*	70	*	38	14	24	*	92	14	14, 14, к
A_2	*	58	20	18	*	56	*	72	20	38, к
A_3	23	19	2	10	1	100	*	30	26	19, 19, 19 11, 81, к
A_4	7	3	*	36	*	121	34	8	41	5, 5, 5, 5 119, к
Потребности	30		22		15		34			
Строка-разность	16		8		32		22			
	16		26		76		22			
	16		26		21		22			
	16		к		21		22			
	16				21		к			
	к				к					

В данном случае начальный план, построенный методом двойного предпочтения, совпал с начальным планом, составленным методом наименьшей стоимости. Однако такое совпадение не является обязательным.

С п о с о б Ф о г е л я. Построение начального плана по данному способу начинается с того, что ищутся разности в каждой строке и в каждом столбце матрицы между наименьшей стоимостью и стоимостью ближайшей к ней по величине (табл. 21).

Разности по строкам записывают справа, разности по столбцам — внизу. Из всех разностей как по строкам, так и по столбцам ищем максимальную. В данном примере она равна 38 и находится во второй строке.

В клетку A_2B_2 , где находится наименьшая стоимость (18) строки с наибольшей разностью, помещаем возможно больший объем перевозок. Он равен 20 единицам груза, т. е. всему наличию груза поставщика A_2 . Поскольку все ресурсы поставщика A_2 исчерпаны, строку A_2 исключаем при дальнейших расчетах, обозначив все клетки этой строки знаком «*». В столбце разностей напротив второй строки пишем к (конец).

После этого снова вычисляем разности по столбцам и строкам, исключая стоимости в занятых местах и в клетках со знаком «*», определяем максимальную разность в строке или столбце, ищем минимальный элемент в строке или столбце с максимальной разностью и в данную клетку направляем возможно большее количество груза. Таким образом все операции повторяются до полного распределения. На последнем этапе, когда остается разметить один или два груза, разности можно не вычислять, а направить груз в клетки с максимальными стоимостями оставшейся части матрицы. Подсчитаем теперь значения целевых функций, составленных по данным приведенных выше способов. Так, стоимость начального плана, построенного методом северо-западного угла, составляет 4079 единиц, методом наименьшей стоимости и методом двойного предпочтения — по 2268 единиц, методом Фогеля 1546 единиц. Стоимость оптимального плана, полученного методом потенциалов, равна 1494 единицам.

Можно убедиться, что для доведения начального плана, составленного методом Фогеля, до оптимального, полученного методом потенциалов, потребуется одна итерация; для доведения начального плана, составленного методом наименьшей стоимости, до оптимального тем же методом — две итерации; для доведения начального плана, составленного методом северо-западного угла, до оптимального потребуется еще больше итераций, а следовательно, и времени.

§ 39. Метод потенциалов

Определение оптимального решения транспортной задачи этим методом заключается в следующем: строится начальный невырожденный опорный план, например, способом северо-западного угла, далее начальный план за определенное число итераций доводится до оптимального, после каждой итерации план проверяется на оптимальность, если оптимальное решение получено — процесс прекращается, если же план неоптимальный, строится новый план, стоимость которого меньше предыдущего [4].

Рассмотрим подробнее метод потенциалов. Исходный план приведен в табл. 22.

	$v_1 = 40$	$v_2 = 15$	$v_3 = -33$	$v_4 = -167$	Ресурсы
$u_1 = 0$	16 40	+ 15	35	100	20
$u_2 = -36$	68	51	3	18	31
$u_3 = -183$	120	30	150	13 16	15
$u_4 = -247$	* 30	54	13	- 80	42
Потребности	16	25	12	55	

Клетки, где запланированы перевозки, в дальнейшем будем называть занятыми местами, клетки, где нет перевозок, — свободными местами. Для оценки оптимальности имеющегося плана определяем потенциалы строк и столбцов. Обозначим потенциалы строк u_i , столбцов — v_j , где i — порядковый номер строки, j — порядковый номер столбца. Для определения потенциалов можно составить систему уравнений. Именно из условия оптимальности известно, что если на пересечении i -й строки с j -м столбцом стоит занятая клетка, то

$$v_j - u_i = c_{ij}.$$

[см. уравнение (VII.16)].

Составим такие уравнения для всех занятых клеток. В нашем примере

$$\begin{aligned} v_1 - u_1 &= 40; & v_3 - u_3 &= 150; \\ v_2 - u_1 &= 15; & v_4 - u_3 &= 16; \\ v_2 - u_2 &= 51; & v_4 - u_4 &= 80. \\ v_3 - u_2 &= 3; \end{aligned}$$

Таким образом, имеем семь уравнений, из которых необходимо определить восемь неизвестных. Один из потенциалов можно взять произвольным. Обычно потенциал первой строки принимается равным нулю. Если $u_1 = 0$, тогда

$$\begin{aligned} v_1 - 0 &= 40; & v_1 &= 40; \\ v_2 - 0 &= 15; & v_2 &= 15. \end{aligned}$$

Далее аналогично

$$u_2 = v_2 - 51 = 15 - 51 = -36;$$

$$v_3 = 3 + u_2 = 3 + (-36) = -33;$$

$$u_3 = v_3 - 150 = -33 - 150 = -183;$$

$$v_4 = 16 + u_3 = 16 + (-183) = -167;$$

$$u_4 = v_4 - 80 = -167 - 80 = -247.$$

На следующем этапе проверяем план на оптимальность. Если найденные величины потенциалов таковы, что для всех свободных мест

$$\gamma_{ij} = u_i - v_j + c_{ij} \geq 0,$$

то план оптимальный [см. уравнение (VII.15)].

Обозначим свободные места в матрице буквами $a, b, c, d, e, f, g, h, i$ и проверим, удовлетворяют ли они условию оптимальности:

$$a - \gamma_{13} = 0 - (-33) + 35 = 68; \quad b - \gamma_{11} = 0 - (-167) + 100 = 267;$$

$$c - \gamma_{21} = -36 - 40 + 68 = -8; \quad d - \gamma_{24} = -36 - (-167) + 18 = 149;$$

$$e - \gamma_{31} = -183 - 40 + 120 = -103; \quad f - \gamma_{32} = -183 - 15 + 30 = -168;$$

$$g - \gamma_{41} = -247 - 40 + 30 = -257; \quad h - \gamma_{42} = -247 - 15 + 54 = -208;$$

$$i - \gamma_{43} = -247 - (-33) + 13 = 201.$$

Можно было бы все свободные клетки не проверять на оптимальность. Достаточно обнаружить одну свободную клетку, не удовлетворяющую условию оптимальности, чтобы сказать, что план неоптимальный и его можно улучшить.

Однако если мы выберем для улучшения плана свободную клетку, имеющую наибольшее нарушение оптимальности (наименьшее отрицательное значение v_{ij}), то получения оптимального плана можно достичь за меньшее число итераций. В данном случае клетка (4.1) имеет наибольшее нарушение условия оптимальности. Обозначим эту клетку знаком «*» и проведем улучшение плана. Для этого из свободного места, обозначенного знаком «*», начинаем двигаться по замкнутой линии, делая шаги только по занятым местам (кроме конечного), причем в каждом занятом месте делаем поворот только под прямым углом. Это движение продолжаем до тех пор, пока не придем в исходное свободное место. В нашем примере движение осуществляется по следующим клеткам: (4.1)—(1.1)—(1.2)—(2.2)—(3.3)—(3.4)—(4.4)—(4.1).

Далее по линии движения, считая от свободного места, представляем в занятых клетках поочередно знаки (—) и (+). Так как замкнутая линия имеет четное число шагов, то исходное свободное место будет иметь всегда (+). В клетках со знаком (—) выбираем наименьшее число, равное объему перевозок, и вычитаем его из

Т а б л и ц а 23

	$u_1 = 40$	$u_2 = 15$	$u_3 = -33$	$u_4 = 90$	Ресурсы
$u_1 = 0$	40 14 —	15 + 6	35	100	20
$u_2 = 36$	68	— 51 19	3 12	* 18 +	31
$u_3 = 74$	120	30	150	15 16	15
$u_4 = 10$	30 2 +	54	13	40 — 80	42
Потребности	16	25	12	55	

всех чисел, находящихся в клетках со знаком (—), а к числам, находящимся в клетках со знаком (+), прибавляем это минимальное положительное число. В данном случае это число находится в клетке (3.3) и равно 2. В результате объем перевозок в отрицательных клетках уменьшится на 2 единицы, а в положительных увеличится на 2 единицы, т. е. в данном случае в клетке (1.1) объем перевозок вместо 16 будет 14, в клетке (2.2) вместо 21 будет 19, в клетке (4.4) вместо 42 будет 40, клетка (3.3) станет в результате вычитания свободной, а клетка (4.1) вместо свободной станет занятой с величиной перевозки равной 2 единицам, в клетке (1.2) вместо 4 будет 6, в клетке (2.3) вместо 10 будет 12, в клетке (3.4) вместо 13 будет 15. Все изменения в плане отражены в табл. 23.

В результате изменения стоимость плана уменьшится на $257 \times 2 = 514$ единиц по сравнению с начальным и составит $5669 - 514 = 5155$ единиц.

Проверяем измененный план на оптимальность. Для этого снова вычисляем потенциалы данной матрицы и проверяем на оптимальность свободные клетки. Вычисление потенциалов довольно просто осуществляется непосредственно по таблице и нет необходимости выписывать уравнения. Значения потенциалов нового плана приведены в табл. 23:

$$\gamma_{13} = 0 - 33 + 35 = +2.$$

$$\gamma_{32} = 74 - 15 + 30 = 89;$$

$$\gamma_{14} = 0 - 90 + 100 = 10;$$

$$\gamma_{33} = 74 + 33 + 150 = 257;$$

$$\gamma_{21} = -36 - 40 + 68 = -8;$$

$$\gamma_{42} = 10 - 15 + 54 = 49;$$

$$\gamma_{24} = -36 - 90 + 18 = -108;$$

$$\gamma_{43} = 10 + 33 + 13 = 56.$$

$$\gamma_{31} = 74 - 40 + 120 = 154;$$

Имеются клетки (2.1), (2.4) с отрицательными значениями, следовательно, план неоптимальный и его можно улучшить. Улучшение

Таблица 24

	$u_1 = -68$	$u_2 = 15$	$u_3 = -33$	$u_4 = -18$	Ресурсы
$u_1 = 0$	40	15	35	100	20
$u_2 = -36$	68	51	3	18	31
$u_3 = -2$	120	30	150	16	15
$u_4 = -98$	30	54	13	80	42
Потребности	16	25	12	55	

плана производим в клетке (2.4) (табл. 24), имеющей нарушение условия оптимальности.

В результате второй итерации стоимость плана уменьшится на $108 \times 14 = 1512$ единиц по сравнению с предыдущим и составит $5155 - 1512 = 3643$ единицы. Проверяем план на оптимальность:

$$\begin{aligned}
 \gamma_{11} &= 0 + 68 + 40 = 108; & \gamma_{32} &= -2 - 15 + 30 = 13; \\
 \gamma_{13} &= 0 + 33 + 35 = 68; & \gamma_{33} &= -2 + 33 + 150 = 181; \\
 \gamma_{14} &= 0 + 18 + 100 = 118; & \gamma_{42} &= -98 - 15 + 54 = -59; \\
 \gamma_{21} &= -36 + 68 + 68 = 100; & \gamma_{43} &= -98 + 33 + 13 = -52. \\
 \gamma_{31} &= -2 + 68 + 120 = 186;
 \end{aligned}$$

Таблица 25

	$u_1 = -39$	$u_2 = 15$	$u_3 = -4$	$u_4 = 11$	Ресурсы
$u_1 = 0$	40	15	35	100	20
$u_2 = -7$	68	51	3	18	31
$u_3 = -5$	120	30	150	16	15
$u_4 = -69$	30	54	13	80	42
Потребности	16	25	12	55	

Улучшаем план в клетке (4.2) (см. табл. 24). Измененный план показан в табл. 25.

В результате третьей итерации стоимость плана уменьшится на $59 \times 5 = 295$ единиц по сравнению с предыдущим и составит $3643 - 295 = 3348$ единиц. Проверяем план на оптимальность:

$$\gamma_{11} = 0 + 39 + 40 = 79;$$

$$\gamma_{31} = -5 + 39 + 120 = 154;$$

$$\gamma_{13} = 0 + 4 + 35 = 39;$$

$$\gamma_{32} = -5 - 15 + 30 = 10;$$

$$\gamma_{14} = 0 - 11 + 100 = 89;$$

$$\gamma_{33} = -5 + 4 + 150 = 149;$$

$$\gamma_{21} = -7 + 39 + 68 = 100;$$

$$\gamma_{43} = -69 + 4 + 13 = -52.$$

$$\gamma_{22} = -7 - 15 + 51 = 29;$$

Улучшаем план в клетке (4.4) (см. табл. 25) и получаем новый план, показанный в табл. 26.

Т а б л и ц а 26

	$u_1 = -9$	$u_2 = -15$	$u_3 = -26$	$u_4 = 41$	Ресурсы
$u_1 = 0$	40	15	35	100	20
$u_2 = 23$	68	51	3	18	31
$u_3 = 25$	120	30	150	16	15
$u_4 = -39$	30	54	13	80	42
Потребности	16	25	12	55	

В результате изменения стоимость плана уменьшится на $52 \times 12 = 624$ единицы по сравнению с предыдущим и составит $3348 - 624 = 2724$ единицы.

Проверяем план на оптимальность:

$$\gamma_{11} = 0 + 9 + 40 = 49;$$

$$\gamma_{23} = 23 + 26 + 3 = 52;$$

$$\gamma_{13} = 0 + 26 + 35 = 61;$$

$$\gamma_{31} = 25 + 9 + 120 = 154;$$

$$\gamma_{14} = 0 - 41 + 100 = 59;$$

$$\gamma_{32} = 25 - 15 + 30 = 40;$$

$$\gamma_{21} = 23 + 9 + 68 = 100;$$

$$\gamma_{33} = 25 + 26 + 150 = 201.$$

$$\gamma_{22} = 23 - 15 + 51 = 59;$$

Все свободные места имеют положительные значения. План оптимальный. Решение окончено.

§ 40. Случай вырождения

При построении начального опорного плана или в процессе решения задачи методом потенциалов часто возникает так называемое вырождение. Как мы уже видели, для построения потенциалов путем решения соответствующих уравнений число занятых мест в матрице должно равняться $(m + n - 1)$, где m — число строк в матрице (исходных данных), n — число столбцов.

Если число занятых мест в матрице меньше $(m + n - 1)$, то имеет место случай вырождения. В табл. 27 приведен пример вырожденного плана. В данном случае начальный опорный план составлен методом северо-западного угла. Однако ввиду соответствия потребности первого потребителя ресурсам первого поставщика число занятых мест оказалось равным 5, в то время как оно должно быть равно $(m + n - 1) = 3 + 4 - 1 = 6$. Так как число занятых мест будет меньше $(m + n - 1)$, нельзя найти все значения потенциалов и, таким образом, нельзя исследовать свободные места, т. е. задачу решить невозможно.

Существует несколько способов устранения вырождения плана:

1) выбираем несколько клеток, дополняющих сумму занятых мест до величины $(m + n - 1)$, вводим в них малые объемы перевозок ϵ . В дальнейшем с этими клетками производят такие же действия, как и с другими занятыми местами;

2) для получения суммы занятых мест, равной $(m + n - 1)$, в некоторые свободные клетки вводим перевозки, равные нулю, и данные клетки считаем занятыми местами. Естественно, вводя нулевую перевозку, мы не изменим величину перевозок, следовательно, на оптимальное решение данный прием не повлияет.

Необходимо заметить, что при введении нулевой перевозки надо правильно выбрать клетку. Может оказаться, что на некоторой итерации при изменении плана занятое место с наименьшим значением в отрицательной клетке окажется равным нулю, а следовательно, план не изменится. В таком случае необходимо нулевую

Т а б л и ц а 27

	$u_1 = 34$	$u_2 = 25$	$u_3 = 22$	$u_4 = 3$	Ресурсы			
$u_1 = 0$	24	34	16	86	10	24		
$u_2 = -16$	0 -	50	28	41	8 +	38	18	36
$u_3 = -78$	*	3	13	100	-	81	10	44
Потребности	24	28	42	10				

	$u_1 = 34$	$u_2 = 16$	$u_3 = 13$	$u_4 = -6$	Ресурсы
$u_1 = 0$	— 34 24	16 0 +	86	10	24
$u_2 = -25$	50	— 41 28	8 + 38	18	36
$u_3 = -78$	3 +	13	— 100 34	81 10	44
Потребности	24	28	42	10	

перевозку поместить в другую клетку. Предположим, нулевую перевозку мы поместили в клетку (2.1) (см. табл. 27). Определяем потенциалы данной матрицы. Проверяем план на оптимальность:

$$u_3 - v_1 + c_{31} = -78 - 34 + 3 = -109,$$

свободное место имеет отрицательное значение, план можно улучшить. Обозначим клетку (3.1) знаком «+», строим замкнутую линию, начиная с этой клетки, шагая по занятым местам и делая повороты под прямым углом. При этом оказывается, что минимальное значение объема перевозки, находящееся в отрицательной клетке, равно нулю. План улучшить нельзя. Передвигаем нулевую перевозку, предположим, в клетку (1.2) (табл. 28).

Определяем потенциалы для этой матрицы. Проверяем свободные места на оптимальность по формуле

$$\gamma_{ij} = u_i - v_j + c_{ij} \geq 0.$$

Наименьшее отрицательное значение находится в клетке (3.1) и равно -109 . Для улучшения плана строим линию указанным выше способом. Минимальное значение объема перевозки в клетках со знаком «—» равно 24. Таким образом, стоимость плана уменьшится на $109 \times 24 = 2616$ единиц по сравнению с начальным. В результате изменения получим новый план (табл. 29).

Число занятых мест равно $(m + n - 1) = 6$. План невырожден. Решение можно продолжать обычным порядком. Иногда для устранения вырождения достаточно поменять местами некоторые столбцы или строки. Так, в нашем примере достаточно поменять первую и третью строки (табл. 30).

Т а б л и ц а 29

	v_1	v_2	v_3	v_4	Ресурсы
u_1	34	16	86	10	24
	24				
u_2	50	41	38	18	36
	4	32			
u_3	3	13	100	81	44
	24		10	10	
Потребности	24	28	42	10	

Т а б л и ц а 30

	v_1	v_2	v_3	v_4	Ресурсы
u_1	3	13	100	81	44
	24	20			
u_2	50	41	38	18	36
	8	28			
u_3	34	16	86	10	24
		14		10	
Потребности	24	28	42	10	

§ 41. Особенности дискретных экстремальных задач

Целочисленное программирование является одним из наиболее важных и наиболее трудных разделов математического программирования. Большое количество разнообразных задач планирования экономики, управления отраслью, организации, размещения и специализации производства формально сводятся к выбору лучших в каком-то смысле значений параметров из некоторой дискретной — конечной или счетной — совокупности заданных величин. Интерес к дискретным экстремальным задачам определяется не только широким кругом приложений, но и органической связью дискретных задач с другими разделами математики — комбинаторным анализом, математической логикой, теорией конечных групп и т. д.

Целочисленное программирование возникло из непрерывного. Поэтому на первом этапе развитие теории и методов решения дискретных задач находилось под явным воздействием идей непрерывного программирования. В дальнейшем появились новые «чисто дискретные» идеи, техника и методы исследования проблем дискретной оптимизации.

Задачи и методы, относящиеся к перечисленному кругу вопросов, различные авторы именуют по-разному. Наибольшее распространение получил термин «целочисленное программирование». Встречаются, однако, другие термины: «дискретное» программирование и реже «комбинаторное», или «диофантово», программирование.

В дискретном программировании рассматриваются главным образом задачи вида

$$\left. \begin{aligned} F(\bar{x}) = \bar{c}\bar{x} \rightarrow \text{exs } A; \\ \bar{a}\bar{x} \{ \leqslant . . . \geqslant \} \bar{b}; \end{aligned} \right\} \quad (\text{VIII.1})$$

$$x_k = 0, \Delta_k, 2\Delta_k, 3\Delta_k, \dots, k \in \overline{1, p}, p \leqslant n; x_l \geqslant 0; l \in \overline{(p+1), n}.$$

Сюда может быть отнесена любая задача математического программирования, в которой есть дополнительное ограничение на вектор $\bar{x}$, состоящее в требовании, чтобы все (полностью дискретные задачи) или некоторые (частично дискретные задачи) компоненты вектора $\bar{x}$ принимали только дискретные неотрицательные значения, т. е.

$$\left. \begin{aligned} x_k = 0, \Delta_k, 2\Delta_k, 3\Delta_k, \dots, \\ k \in \overline{1, p}, p \leqslant n, x_l \geqslant 0, l \in \overline{(p+1), n}, \end{aligned} \right\} \quad (\text{VIII.2})$$

где n — размерность вектора $\bar{x}$, Δ_k — некоторый дисконт (положительная величина).

Частным случаем задач дискретного программирования являются задачи целочисленного программирования. В них $\Delta_k = 1$. Подчеркнем, что требование дискретности хорошо сочетается с методом динамического программирования Беллмана, применяемым для решения задач с сепарабельными целевыми функциями.

В задачах целочисленного программирования область допустимых решений является невыпуклой и несвязной. Поэтому отыскание решения таких задач сопряжено со значительными трудностями. В частности, невозможно применение округления найденного решения до ближайшего целочисленного.

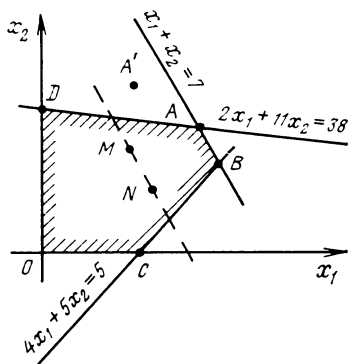


Рис. 49

Пример VIII. 1. Найти максимум функции

$$F(x_1, x_2) = x_1 + x_2$$

в области, определенной условиями

$$2x_1 + 11x_2 \leq 38;$$

$$x_1 + x_2 \leq 7;$$

$$4x_1 - 5x_2 \leq 5;$$

$$x_1 \geq 0; x_2 \geq 0; x_1, x_2 — \text{целые числа.}$$

Решим задачу геометрически (рис. 49). Область поиска экстремума — многоугольник $ODABC$, но, так как линия уровня целевой функции параллельна стороне AB многоугольника, экстремум достигается в вершинах $A(8/3, 13/3)$ и $B(23/9, 40/9)$, а также в любой точке отрезка AB и равен 7. Однако нас интересуют лишь точки с целочисленными координатами, следовательно, ни A , ни B не являются допустимым решением задачи. Округляя значения координат A , получим $A'(2,4)$. Но точка A' не принадлежит к области поиска. Можно показать, что целочисленный оптимум достигается в точках $N(3,2)$ и $M(2,3)$ и равен 5. Обе точки — внутри области поиска.

Приведенный пример показал, что для решения задач с требованием целочисленности необходимо рассмотреть особые методы оптимизации, и, кроме того, мы видим, что оптимальное решение задач целочисленного программирования не обязательно принадлежит к границе многогранника (многоугольника) условий, что было характерно для задач линейного программирования.

Можно отметить несколько основных классов задач, приводящих к моделям целочисленного программирования.

1. Задача с неделимостями. К таким задачам относятся многие задачи, связанные с планированием выпуска неделимых видов продуктов или с использованием неделимых ресурсов (автомашин и т. д.).

2. Экстремальные комбинированные задачи: задачи календарного планирования, многошаговые задачи принятия решений и др.
3. Задачи с разрывной целевой функцией.
4. Задачи на несвязных и невыпуклых областях.

§ 42. Задачи целочисленного программирования

Рассмотрим несколько примеров постановки таких задач.

1. Задачи с неделимостями

Транспортная задача (см. гл. VII) является одной из самых распространенных задач целочисленного программирования. Она имеет много модификаций и обобщений, поэтому иногда ее называют обобщенной транспортной задачей. В данной задаче целочисленность является обязательным условием.

Задача о ранце. Положим, что турист готовится к длительному переходу в горах. В рюкзаке он может нести груз, масса которого не более W кг. Этот груз может включать n видов предметов, каждый предмет типа j , массой ω_j кг, $j = 1, 2, \dots, n$. Для каждого вида предметов турист определяет его ценность E_j во время перехода. Сколько предметов каждого вида он должен положить в рюкзак, чтобы суммарная ценность снаряжения была максимальной?

Обозначим x_j — количество предметов в рюкзаке j -го вида. Тогда математическая модель задачи такова:

$$\sum_{j=1}^n E x_j \rightarrow \max \quad (\text{VIII.3})$$

при ограничениях

$$\sum_{j=1}^n \omega_j x_j \leq W; \quad x_j \text{ — целое; } x_j \geq 0; \quad j = 1, 2, \dots, n. \quad (\text{VIII.4})$$

Графическое решение задачи о ранце приведено в работе [3].

2. Экстремальные комбинаторные задачи

Задача о назначении. Имеется n работ и n кандидатов на выполнение каждой работы. Назначение i -го рабочего на j -ю работу вызывает затраты c_{ij} (или дает прибыль). Требуется определить наилучшее с точки зрения минимума суммарных затрат распределение рабочих. Комбинированная задача представляет собой перестановку $(p_1, p_2, \dots, p_n)$ чисел $(1, 2, \dots, n)$.

Каждое назначение одного i -го рабочего на p_i рабочее место описывается соответствием $i \rightarrow p$. Требуется определить перестановку $(p_1^*, p_2^*, p_3^*, \dots, p_n^*)$, при которой

$$\sum_{i=1}^n c_i p_i = \min. \quad (\text{VIII.5})$$

Значение суммы необходимо вычислить по всем перестановкам $(p_1, p_2, \dots, p_n)$ индексов $(1, 2, \dots, n)$, например при $(1, 2, \dots, n)$,

(2, 1, 3, ..., n). При больших n получается очень сложный вычислительный процесс: сильно возрастает число перестановок $[n! = 1, 2, \dots, (n-1)n]$. Посредством булевых переменных

$x_{ij} = \begin{cases} 1, & \text{если } i\text{-й кандидат назначается на } j\text{-ю работу;} \\ 0 & \text{— в противном случае.} \end{cases}$

Сведем эту задачу к задаче целочисленного программирования вида

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min; \quad (\text{VIII.6})$$

$$\sum_{j=1}^n x_{ij} = 1; \quad \sum_{i=1}^m x_{ij} = 1,$$

где x_{ij} — целые числа.

Пример назначения бригад на полевые работы приведен в работе [3].

Задача о коммивояжере («бродячем» торговце). Коммивояжер, выезжая из пункта A_0 , должен побывать по одному разу в пунктах $A_1, A_2, \dots, A_n$ и вернуться в пункт A_0 . Известны расстояния c_{ij} между любыми двумя пунктами A_i и A_j . Задача заключается в выборе маршрута, минимизирующего общую длину пути, пройденного коммивояжером. Чтобы записать эту комбинаторную задачу в терминах целочисленного программирования, введем булевы переменные

$x_{ij} = \begin{cases} 1, & \text{если коммивояжер переходит из пункта } A_i \text{ в пункт } A_j; \\ 0 & \text{— в противном случае} \end{cases}$

и целочисленные вспомогательные переменные $u_j, j = \overline{1, n}$. В этих терминах задача коммивояжера записывается в виде линейной целочисленной модели:

$$\sum_{i=0}^n \sum_{j=0}^n c_{ij} x_{ij} \rightarrow \min; \quad (\text{VIII.7})$$

$$\sum_{i=0, i \neq j}^n x_{ij} = 1; \quad j = \overline{0, n}; \quad \sum_{j=0, j \neq i}^n x_{ij} = 1; \quad i = \overline{0, n}; \quad (\text{VIII.8})$$

$$u_i - u_j + n x_{ij} \leq n - 1; \quad i, j = \overline{1, n}; \quad i \neq j; \quad (\text{VIII.9})$$

$$u_i \geq 0 - \text{целые}; \quad i = \overline{1, n}; \quad x_{ij} \in \{0, 1\}; \quad i, j = \overline{1, n}; \quad i \neq j. \quad (\text{VIII.10})$$

Условия (VIII.7) и (VIII.8) означают, что из каждого пункта коммивояжер только один раз выезжает и в каждый пункт только один раз приезжает. Таким образом, исключаются циклы на маршруте. Условия (VIII.9) гарантируют наличие определенных последовательностей посещения пунктов и не допускают расщепления замкнутого маршрута из n звеньев на несколько замкнутых маршрутов с меньшим числом звеньев.

Один из вариантов применения задачи коммивояжера в геодезической практике описан в работе [3].

Задача планирования. В кибернетике большинство задач оптимизации делится на две группы: распределения и расписания. В задачах первой группы (распределения ресурсов, транспортной) время, как правило, отсутствует. В задачах второй группы, которые часто возникают в автоматизированных системах управления, требуется ресурсы распределить по времени. Встречаются и смешанные задачи распределения и расписания, в которых требуется одновременно оптимальным образом распределить ресурсы по объектам и расписать это распределение по времени. Так, задачи о назначении и транспортная относятся к задачам распределения, так как в них отсутствует параметр время, или применительно к задачам целочисленного программирования календарь, под которым понимается упорядоченная последовательность целых (дискретных) значений времени: 1, 2, 3, . . . , T . Однако практически задачи планирования представляют собой сложную совокупность этих двух задач, к которым всегда добавляются условия соблюдения заданной последовательности операций и, как правило, целочисленность переменных.

3. Задачи с разрывной целевой функцией

Рассмотрим следующую задачу. Пусть x_j — объем выпускаемой продукции типа j ; c_j — прибыль за комплект продукции типа j . Предполагается, что прибыль поступает только за готовые комплекты, отвечающие целочисленным значениям x_j . Параметр a_{ij} означает количество ресурса вида i , необходимое для производства комплекта продукции типа j , а b_i — наличные запасы ресурса i . Цель — получить максимум прибыли при данных условиях. Формальная модель задачи имеет вид

$$\sum_{j=1}^n c_j [x_j] \rightarrow \max; \quad (\text{VIII.11})$$

$$\sum_{j=1}^n a_{ij} x_j \leq b_i; \quad i = \overline{1, m}; \quad (\text{VIII.12})$$

$$x_j \geq 0; \quad j = \overline{1, n}. \quad (\text{VIII.13})$$

Здесь $[x_j]$ — целая часть переменной x_j .

При $c_j \geq 0$ задачи (VIII.11)—(VIII.13) сводятся к частично целочисленной. Введем дополнительные целочисленные переменные $z_j \geq 0; j = \overline{1, n}$. Тогда исходная задача с разрывной целевой функцией эквивалентна следующей частично целочисленной задаче линейного программирования:

$$\begin{aligned} \sum_{j=1}^n c_j z_j &\rightarrow \max, \\ \sum_{j=1}^n a_{ij} x_j &\leq b_i; \quad i = \overline{1, m}; \\ x_j - 1 &\leq z_j \leq x_j; \quad j = \overline{1, n}; \\ z_j &\geq 0 - \text{целое}; \quad x_j \geq 0; \quad j = \overline{1, n}. \end{aligned}$$

Отметим, что к задачам с разрывной целевой функцией относятся также задачи с фиксированными доплатами [10], задачи о смесях [10] и др.

4. Задачи на несвязных и невыпуклых областях

Еще одним источником возникновения дискретности могут служить многоэкспериментальные задачи оптимизации. Каждую многоэкстремальную задачу можно рассматривать как задачу математического программирования с линейным целевым функционалом и с невыпуклой или несвязной областью определения. Любую такую область можно точно или приближенно представить в виде объединения конечного числа выпуклых областей или даже выпуклых многогранников. Ввод дополнительных переменных позволяет свести эту задачу к частично целочисленной задаче линейного программирования. Рассмотрим это на общем примере.

Пусть невыпуклая область M определения задачи представляет собой объединение двух выпуклых многогранников M_1 и M_2 . Им соответствуют системы ограничений

$$L'_i(x) = \sum_{j=1}^n a'_{ij}x_j \leq b'_i; \quad i = \overline{1, m_1} \quad \text{и} \quad L''_i(x) = \sum_{j=1}^n a''_{ij}x_j \leq b''_i; \quad i = \overline{1, m_2}.$$

Допустим, что известно число b , для которого справедливы неравенства

$$\max_{x \in M} L'_i(x) \leq b'_i + b; \quad i = \overline{1, m_1};$$

$$\max_{x \in M} L''_i(x) \leq b''_i + b; \quad i = \overline{1, m_2}.$$

Введем дополнительное переменное y , принимающее значение 0 или 1 (величина y называется булевой переменной). Тогда, как нетрудно видеть, многогранник M можно описать системой ограничений

$$\begin{aligned} L'_i(x) &\leq b'_i + by; \quad i = \overline{1, m_1}; \\ L''_i(x) &\leq b''_i + b(1 - y); \quad i = \overline{1, m_2}; \\ y &\in [0, 1]. \end{aligned}$$

Пример VIII.2. Найти

$$F(x_1, x_2) = c_1x_1 + c_2x_2 \rightarrow \max$$

при условии либо

$$3x_1 + x_2 \leq 3,$$

либо

$$x_1 + x_2 \leq 2.$$

Если оба ограничения должны выполняться одновременно, то в результате получается выпуклая область. Ограничение же типа «либо—либо» порождает заштрихованную невыпуклую область (рис. 50).

Запишем оба ограничения в таком виде:

$$\left. \begin{aligned} f(x_1, x_2) &\leq 0; \\ g(x_1, x_2) &\leq 0. \end{aligned} \right\} \quad (\text{VIII.14})$$

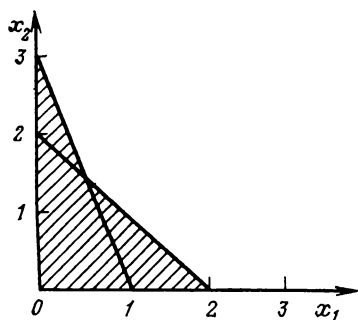


Рис. 50

Предположим, что B_f и B_g — верхние границы значений функций f , g на некоторой области S . В приведенном примере

$$f(x_1, x_2) = 3x_1 + x_2 - 3;$$

$$g(x_1, x_2) = x_1 + x_2 - 2.$$

Можно умозрительно заключить, что значения ни x_1 , ни x_2 не превзойдут, скажем, 5. Поэтому мы могли бы заведомо считать допустимой область

$$S = \{(x_1, x_2): 0 \leq x_1 \leq 5; 0 \leq x_2 \leq 5\}$$

и положить $B_f = 17$ и $B_g = 8$. Зная B_f и B_g (эти числа не обязательно должны быть точными верхними гранями), перепишем неравенства (VIII.14) в следующем виде:

$$f(x_1, x_2) - \delta B_f \leq 0; g(x_1, x_2) - (1 - \delta) B_g \leq 0; \delta = (0, 1). \quad (\text{VIII.15})$$

Новая переменная δ должна быть или 0, или 1. Если $\delta = 0$, действует первое ограничение, а второе выполняется автоматически для любой точки $(x_1, x_2) \in S$. Если же $\delta = 1$, то первое ограничение справедливо автоматически для всех $(x_1, x_2) \in S$ и реальным ограничением является второе.

Добавление третьей переменной δ преобразует невыпуклую двумерную допустимую область исходной задачи в выпуклую трехмерную область, задаваемую неравенствами:

$$f(x_1, x_2) - \delta B_f \leq 0; g(x_1, x_2) - (1 - \delta) B_g \leq 0; 0 \leq \delta \leq 1. \quad (\text{VIII.16})$$

Для того чтобы решить исходную задачу, достаточно наложить на δ требования целочисленности.

Задачи такого типа с дополнительными логическими условиями вида «либо — либо» будем называть **дихотомическими**.

Определение выгодности выполнения работы (задача типа «либо — либо»). Предположим, что x_j — объем определенной работы, которую можно принять к исполнению на данном промежутке времени либо отказаться от нее. При этом возможны две ситуации: или $x_j \geq L_j$, где L_j — наименьший из возможных (еще выгодный) объем принятой к исполнению работы, или $x_j = 0$, т. е. работа не принимается (невыгодна). Если ввести в задачу число M такой величины, что

$$\max \{x_j\} \leq M,$$

и булеву переменную y_i , причем $y_i = 0$, когда работа не принимается, или $y_i = 1$, когда работа принимается к исполнению, то в общем виде модель целочисленной задачи для выполнения k работ из m возможных формулируется таким образом.

Найти

$$\sum_{j=1}^n c_j x_j \rightarrow \max \quad (\text{VIII.17})$$

при ограничениях

$$\sum_{i=1}^m y_i = k; \quad \sum_{j=1}^n a_{ij}x_j \leq b_i y_i + M; \quad i = \overline{1, m}, \quad j = \overline{1, n}; \quad n \geq m;$$
$$x_j = 0, 1, 2 \dots \text{— целые числа.} \quad (\text{VIII.18})$$

Заметим, что к этому же классу задач относится задача типа «да — нет», которая рассматривается при распределении капиталовложений [4].

§ 43. Общая характеристика методов решения целочисленных задач

При переходе от оптимального нецелочисленного решения к оптимальному целочисленному обычно уменьшается значение функционала $F(\bar{x})$. Поэтому кажется заманчивым осуществить переход к целочисленному решению путем округления значений переменных в оптимальном плане непрерывной задачи без уменьшения значения функционала. Такой подход часто оправдывается тем, что получение целочисленного решения связано как с возрастанием вычислительных трудностей, так и с потерей оптимальности функционала. Так, некоторые авторы утверждают, что потери от нецелочисленности могут быть меньше, чем потери от достижения целочисленного результата, подразумевая под последним уменьшение значения функционала в целочисленном оптимальном плане. С таким утверждением согласиться нельзя, ибо выгода от лучшего значения функционала при нецелочисленном плане иллюзорна, так как в дальнейшем при реализации плана целочисленные факторы так или иначе проявят себя и фактический ущерб будет много больше того теоретического преимущества в величине функционала, которое достигается при нецелочисленном плане. Метод простого округления нельзя назвать методом целочисленного программирования и в математическом смысле. Простейший пример, приведенный в § 41, показывает, что при округлении получается план не оптимальный, а частный, не соответствующий поставленным в задаче условиям.

В настоящее время в целочисленном программировании существуют два вида методов решения целочисленных задач: 1) точные и 2) приближенные.

Точные методы целочисленного программирования. За двадцать лет, прошедших после опубликования первых работ по целочисленному программированию, появилось много различных алгоритмов, обеспечивающих точное решение целочисленных задач. Однако все они укладываются в ограниченное число вычислительных схем. Наибольшее число известных алгоритмов основаны на идеях: а) метода отсечения, б) метода ветвей и границ, в) различных модификаций динамического программирования.

Ниже будут изложены основные особенности и принципы построения этих методов. Вопросы подхода к целочисленному программированию, основанного на теоретико-групповых методах и на учете специфики конкретных классов дискретных задач, приведены в работах [30, 4].

Приближенные методы целочисленного программирования. Приведенное здесь деление приближенных методов на «традиционные» и «новые» основано главным образом на материалах монографии [30].

«Традиционные» приближенные методы можно разделить на два класса:

приближенные методы, порожденные известными точными методами;

методы, с самого начала ориентирующиеся на приближенное решение задачи.

Методы второго класса, в свою очередь, подразделяются на три группы: детерминированные методы, методы случайного поиска и методы, использующие специфику задачи. Имеются, конечно, и комбинаторные методы, основанные на идеях методов разных групп.

Среди «новых» подходов к построению приближенных методов целесообразно выделить следующие направления:

ϵ -подход, объединяющий методы приближенного решения задачи с относительной погрешностью, не превышающей заданной величины;

статистические эффективные приближенные методы, обеспечивающие в большинстве случаев малую погрешность;

методы, связанные с варьированием исходной модели — с выбором из многих возможных формализаций исследуемой содержательной модели наиболее приемлемой с точки зрения вычислений. Это главным образом эвристические методы.

§ 44. Точные методы целочисленного программирования

Естественно начать изучение точных методов целочисленного программирования с характеристики класса линейных задач, для которого требование целочисленности оптимального плана выполняется автоматически. Для этих задач любой конечный метод линейного программирования оказывается в то же время точным методом целочисленного программирования.

Рассмотрим задачу линейного программирования

$$\sum_{j=1}^n c_j x_j \rightarrow \min; \quad (\text{VIII.19})$$

$$\sum_{j=1}^n a_{ij} x_j = b_i; \quad \overline{i, m}, \quad (\text{VIII.20})$$

$$x_j \geq 0; \quad j = \overline{1, n}. \quad (\text{VIII.21})$$

Ограничения (VIII.20), (VIII.21) определяют выпуклое многогранное множество M . Будем считать элементы a_{ij} матрицы

$$\|a_{ij}\| = A$$

фиксированными целыми числами, а компоненты b_i вектора b — произвольными. Обозначим семейство многогранников, определяемых условиями задачи, $M(b)$.

Следующее утверждение выявляет класс матриц A , при которых все вершины многогранного множества $M(b)$ целочисленны.

Т е о р е м а (VIII.1). *Для целочисленности всех вершин любого из многогранных множеств $M(b)$ с произвольным t -мерным вектором ограничений b , имеющим целочисленные компоненты, необходимо и достаточно, чтобы любой минор порядка t матрицы A был равен 0 или ± 1 .*

Условием теоремы (VIII.1) удовлетворяет, в частности, матрица A классической транспортной задачи и некоторых других задач транспортного типа. Жесткость условий теоремы означает, что класс задач линейного программирования, у которых имеются целочисленные решения, крайне беден. Как правило, решив задачу, мы получим нецелочисленный оптимальный план. Поэтому возникает необходимость в специальных методах целочисленного программирования.

Метод отсекающих плоскостей (метод отсечения). Идея метода отсечения, насколько известно, принадлежит Дж. Данцигу [7]. Сначала предлагается пренебречь требованием целочисленности переменных и вместо целочисленной задачи решить соответствующую ей задачу линейного программирования. Если же на первом шаге получено нецелочисленное решение, то к исходным линейным ограничениям задачи добавляется новое линейное неравенство. Оно формируется на следующем основании: полученное нецелочисленное решение не должно удовлетворять новому неравенству, а любой целочисленный план (не только оптимальный) заведомо удовлетворяет ему.

Затем описанная процедура повторяется с новой задачей линейного программирования, система ограничений которой содержит на одно неравенство больше, чем исходная задача. Решение полученной задачи вновь проверяется на целочисленность, и, если возникает необходимость, к системе ограничений задачи вновь добавляется линейное неравенство. Если правило формирования дополнительных ограничений разработано удовлетворительно, то через некоторое число этапов процедура завершится. Будет найдено решение целочисленной задачи, или очередная система линейных неравенств окажется противоречивой. В последнем случае будет зафиксирована несовместимость условий исходной задачи.

А л г о р и т м Г о м о р и. Первым, кому удалось удовлетворительно реализовать идею Дж. Данцига, был Р. Гомори. Ему принадлежит строго аргументированные алгоритмы полностью целочисленного и частично целочисленного линейного программирования [4].

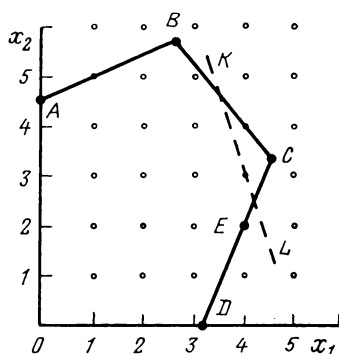


Рис. 51

гранника ограничений, содержащую вершину-оптимум (точку C) и проходящую через какую-то точку решетки (на рис. 51 через точку E).

Рассмотрим пример, при решении которого будет разъяснено правило составления дополнительных ограничений.

Пример VIII.3. Найти целочисленное решение задачи:

$$\begin{aligned} 3x_1 + 2x_2 &\leq 10; \\ x_1 + 4x_2 &\leq 11; \\ 3x_1 + 3x_2 + x_3 &\leq 13; \\ F = 4x_1 + 5x_2 + x_3 &\rightarrow \max, \\ x_1x_2x_3 &\geq 0. \end{aligned}$$

Решение. Составим жорданову таблицу, соответствующую данной задаче (табл. 31).

Т а б л и ц а 31

	$-x_1$	$-x_2$	$-x_3$	1
$y_1 =$	3	2	0	10
$y_2 =$	1	4	0	11
$y_3 =$	3	3	1	13
$F =$	-4	-5	-1	0

Т а б л и ц а 32

	$-x_1$	$-x_2$	$-x_3$	1
$y_1 =$	10/4	-2/4	0	18/4
$x_2 =$	1/4	1/4	0	11/4
$y_3 =$	9/4	-3/4	1	19/4
$F =$	-11/4	5/4	-1	55/4

Решение получается в результате следующих трех итераций — табл. 32, 33, 34 (разрешающие коэффициенты заключены в рамку).

Т а б л и ц а 33

	$-y_1$	$-y_2$	$-x_3$	1
$x_1 =$	4/10	-2/10	0	18/10
$x_2 =$	-1/10	3/10	0	23/10
$y_3 =$	-9/10	-3/10	1	7/10
$F =$	11/10	7/10	-1	187/10

Т а б л и ц а 34

	$-y_1$	$-y_2$	$-y_3$	1
$x_1 =$	4/10	-2/10	0	18/10
$x_2 =$	-1/10	3/10	0	23/10
$x_3 =$	-9/10	-3/10	1	7/10
$F =$	2/10	4/10	1	194/10

Оптимальное решение получаем из последней таблицы $x_1 = 18/10$; $x_2 = 23/10$; $x_3 = 7/10$; $F = 194/10$; оно — нецелочисленно.

Теперь вводят дополнительное ограничение:

1) выделяем какую-нибудь строку, в которой заключительный коэффициент (т. е. соответствующая координата полученного оптимального плана) — нецелое число; например, выделим в последней таблице третью строку, в которой заключительный коэффициент равен $7/10$;

2) для каждого коэффициента выделенной строки находят разность между ближайшим целым числом, содержащимся в нем, и им самим — эти разности служат соответствующими коэффициентами в дополнительном ограничении. В нашем примере будут следующие разности:

$$-1 - (9/10) = -1/10;$$

$$-1 - (3/10) = -7/10;$$

$$1 - 1 = 0;$$

$$0 - 7/10 = -7/10.$$

Эти числа являются коэффициентами дополнительного ограничения

$$S_3 = -1/10 (-y_1) - 7/10 (-y_2) + 0 (-y_3) - 7/10 \times 1;$$

3) вносим дополнительное ограничение в последнюю таблицу и снова решаем задачу оптимизации.

В результате получим табл. 35.

Т а б л и ц а 35

	$-y_1$	$-y_2$	$-y_3$	1
$x_1 =$	4/10	-2/10	0	18/10
$x_2 =$	-1/10	3/10	0	23/10
$x_3 =$	-9/10	-3/10	1	7/10
$S_3 =$	-1/10	-7/10	0	-7/10
$F =$	2/10	4/10	1	194/10

После внесения дополнительного ограничения полученный ранее план перестал быть оптимальным. Он даже не опорный, так как в заключительном столбце появился отрицательный коэффициент ($-7/10$). Чтобы получить опорное решение, надо проделать жордановы исключения, работая со строкой для S_3 . По соответствующему правилу (см. § 34), находим разрешающий коэффициент, который оказывается в той же строке для S_3 и, сделав один шаг жорданова исключения, приходим к табл. 36, из которой получаем оптимальный целочисленный план: $x_1 = 2$; $x_2 = 2$, $x_3 = 1$, $F_{\max} = 19$.

Т а б л и ц а 36

	$-y_1$	$-S_3$	$-y_3$	1
$x_1 =$				2
$x_2 =$				2
$x_3 =$				1
$y_2 =$				1
$F =$	1/7	4/7	1	19

Перейдем к обоснованию п р а в и л а Г о м о р и. Допустим, что при решении задачи оптимальное решение оказалось нецелочисленным, т. е. в жордановой таблице на последнем шаге по крайней мере один из свободных членов, например b_i , оказался нецелым числом. Пусть i -я строка в этой последней таблице имеет вид (табл. 37).

Т а б л и ц а 37

	$-\xi_1$	$-\xi_2$	. . .	$-\xi_j$	. . .	$-\xi_n$	1
i	b_{i1}	b_{i2}		b_{ij}		b_{in}	b_i

Если обозначить $[b_{ij}]$ наибольшее целое число, содержащееся в коэффициенте b_{ij} , то это значит, что $[b_{ij}] \leq b_{ij} < [b_{ij}] + 1$. Если $[b_i]$ — наибольшее целое число, содержащееся в свободном члене b_i , то поскольку b_i — нецелое, $[b_i] < b_i < [b_i] + 1$ (b_i — существенно дробное число). В соответствии с правилом Гомори нужно образовать разности $[b_{ij}] - b_{ij}$ — они неположительны, а разность $[b_i] - b_i$ — существенно отрицательна. Эти разности обозначим соответственно через $-\beta_i$ и $-\beta_{ij}$, следовательно, $[b_{ij}] - b_{ij} = -\beta_{ij} \leq 0$; $[b_i] - b_i = -\beta_i < 0$.

Дополнительное ограничение, которое вводится, по правилу Гомори, будет

$$S_i = \sum_{j=1}^n (-\beta_{ij})(-\xi_j) - \beta_i. \quad (\text{VIII.22})$$

Значит, для оптимального решения, в котором $\xi_1 = \xi_2 = \dots = \xi_n = 0$ (верхние переменные), будем иметь

$$S_i = -\beta_i < 0,$$

т. е. оптимальная вершина многогранника ограничений и сам многогранник, для точек которого $S_i \geq 0$, расположены по разные стороны от гиперплоскости $S_i = 0$. Значит, эта гиперплоскость действительно отсекает нецелочисленную вершину-оптимум.

Можно также показать, что эта гиперплоскость проходит через некоторую точку целочисленной решетки.

Метод ветвей и границ. Другой общий подход к точному решению задач целочисленного программирования дается схемой методов ветвей и границ (или схемой методов направленного перебора). Методы ветвей и границ используются для решения целочисленных задач, обладающих конечным (возможно, очень большим) числом допустимых планов. Разные реализации методов ветвей и границ объединяет общая идея перехода от полного перебора к сокращенному (направленному) перебору. Сокращение перебора достигается за счет массового отсеивания неперспективных вариантов. Выявление множества вариантов, подлежащих отсеиванию на каждом шаге, представляет собой наиболее важный этап метода, определяющий его эффективность. Этот этап существенно использует особенности структуры ограничений задачи.

В основе различных модификаций методов ветвей и границ лежит несколько принципов, позволяющих во многих случаях существенно сократить перебор вариантов решений дискретных экстремальных задач: 1) вычисление нижней (при решении задач на минимум) границы оценки целевой функции задачи на области ее определения (или на некотором ее подмножестве); 2) последовательное разбиение (ветвление) области определения задачи, шаг за шагом сокращающее подмножества; 3) пересчет оценок; 4) установление признака оптимальности, фиксирующего выбор решения.

Алгоритмы метода ветвей и границ различаются стратегией ветвления и принципами оценивания. Стратегии ветвления определяются обычно нижними оценками, построенными на последнем шаге и, возможно, на предыдущих шагах. Простейший способ вычисления оценок сводится к решению задач линейного программирования, игнорирующих условия целочисленности. Усиление оценок позволяет ускорить вычислительный процесс. При решении практических задач стратегии ветвления и оценивания определяются обычно эвристическими соображениями, существенно использующими специфику задачи.

Впервые идеи метода ветвей и границ использовали А. Лэнд и Л. Дойг в работе по решению задачи линейного целочисленного

программирования. Однако своей популярностью метод ветвей и границ обязан главным образом алгоритму Дж. Литла, К. Мурти, Д. Суини и К. Кэрл, предложенному ими не для линейной целочисленной модели, а для задачи коммивояжера в ее комбинаторной формулировке.

Рассмотрим модель вида:
максимизировать

$$\sum_{j=1}^n c_j x_j \quad (\text{VIII.23})$$

при ограничениях

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad (i = 1, 2, \dots, m); \quad (\text{VIII.24})$$

$$x_j - \text{целые числа } (j = 1, 2, \dots, p); \quad p \leq n; \quad (\text{VIII.25})$$

$$x_j \geq 0 \quad (j = p + 1, \dots, n). \quad (\text{VIII.26})$$

Допустим, что для каждой целочисленной переменной можно задать верхнюю U_j и нижнюю Z_j границы, в пределах которых, безусловно, содержится ее оптимальное значение:

$$Z_j \leq x_j \leq U_j \quad (j = 1, 2, \dots, p). \quad (\text{VIII.27})$$

Обычно $Z_j = 0$, но это не обязательно. Идея состоит в том, что если рассмотреть любую переменную x_j и принять, что S есть некоторое число, где $Z_j \leq S \leq U_j - 1$, то оптимальный план задачи (VIII.23) — (VIII.26) будет удовлетворять одному из двух линейных ограничений:

либо

$$x_j \geq S + 1, \quad (\text{VIII.28})$$

либо

$$x_j \leq S. \quad (\text{VIII.29})$$

Если найти оптимальный план нашей задачи без учета условия целочисленности (VIII.25), то сразу можно будет поставить и решить еще две задачи линейного программирования, каждая из которых содержит условия (VIII.24), (VIII.25) и (VIII.26), а условие (VIII.27) для одной задачи записывается как (VIII.28), а для другой — как (VIII.29). В случае, когда каждая из этих задач имеет оптимальный план, удовлетворяющий условию целочисленности (VIII.25), то остается выбрать такой план, который доставляет большее значение целевой функции. Это и будет оптимальный план целочисленной задачи. Если одна (или обе) из этих задач не имеет оптимального плана, удовлетворяющего условию целочисленности, то потребуются дополнительные вычисления, о которых будет сказано ниже.

Как и все алгоритмы математического программирования, метод ветвей и границ является итеративным. Для этого метода важно, чтобы на каждой итерации (обозначим ее t) имелась нижняя граница оптимального значения целевой функции Z_0^t . Обычно характер

решаемой задачи позволяет установить эту границу для первой итерации хотя бы приближенно. Но если этого сделать нельзя, принимают даже $Z'_0 = -\infty$; важно, чтобы граница была. Затем составляют так называемый «список задач» линейного программирования, подлежащих решению. Множество задач этого списка отличаются друг от друга только ограничениями (VIII.27).

На первой итерации ($t = 1$) основной список содержит лишь одну задачу (VIII.23), (VIII.24) (VIII.26) и (VIII.27).

Алгоритм метода ветвей и границ состоит в следующем.

Шаг 1. Выбрать из основного списка одну задачу линейного программирования. Если основной список пуст, то вычисления прекращаются.

Шаг 2. Решить выбранную задачу. Если на этой k -й итерации оптимальное значение целевой функции $Z_0^k \geq Z_0^t$, то переходят к шагу 3. Если же окажется, что $Z_0^k \leq Z_0^t$, т. е. меньше или равно нижней ее границе, то принимают $Z_0^k = Z_0^t$ и возвращаются к шагу 1.

Шаг 3. В случае, когда полученный оптимальный план задачи линейного программирования удовлетворяет условию целочисленности, этот план фиксируют и возвращаются к шагу 1, принимая $Z_0^{k+1} = Z$. В противном случае переходят к шагу 4.

Шаг 4. Выбирают любую переменную x_j ($j = 1, 2, \dots, p$), не имеющую целого значения в полученном на k -й итерации оптимальном плане выбранной задачи линейного программирования. В основной список добавляют две задачи линейного программирования, идентичные задаче, выбранной на шаге 1, но первая из них удовлетворяет ограничению (VIII.28), а вторая — (VIII.29). Принимают $Z_0^{k+1} = Z_0^t$ и возвращаются к шагу 1.

Пример VIII.4. Максимизировать

$$3x_1 + 3x_2 + 13x_3$$

при ограничениях

$$\left. \begin{aligned} -3x_1 + 6x_2 + 7x_3 &\leq 8; \\ 6x_1 - 3x_2 + 7x_3 &\leq 8; \end{aligned} \right\} \text{ (Задача 1).}$$

$$x_1 \geq 0; \quad x_2 \geq 0; \quad x_3 \geq 0; \quad x_1, x_2, x_3 \text{ — целые числа.}$$

Допустим, что установлены (или заданы) такие границы на каждую переменную:

$$0 \leq x_j \leq 5 \quad (j = 1, 2, 3).$$

Решение. 1. Вносим в основной список задачу 1. Принимаем $Z'_0 = 0$. Выбираем эту задачу на шаге 1 и находим ее оптимальный план на шаге 2.

$$Z_0 = 16; \quad x_1 = x_2 = 2^{2/3}; \quad x_3 = 0.$$

План не является целочисленным; переходим к шагу 4 и выбираем x_1 . Это дает возможность внести в основной список еще две задачи:

задачу 2

$$3 \leq x_1 \leq 5; \quad 0 \leq x_2 \leq 5; \quad 0 \leq x_3 \leq 5$$

и задачу 3

$$0 \leq x_1 \leq 2; \quad 0 \leq x_2 \leq 5; \quad 0 \leq x_3 \leq 5.$$

2. Возвращаемся к шагу 1 при $Z_0^2 = Z_0^1 = 0$. Выбираем задачу 2. На шаге 2 обнаруживаем, что допустимого решения у задачи 2 нет. Поэтому примем $Z_0^3 = Z_0^2 = 0$ и вновь вернемся к шагу 1.

3. Выбираем теперь задачу 3 и находим ее оптимальный план:

$$Z_0 = 15^{4/7}; \quad x_1 = x_2 = 2; \quad x_3 = 1/2.$$

Этот план также не удовлетворяет требованию целочисленности. Поэтому переходим к шагу 4, где выберем x_4 и внесем в основной список еще две задачи: задачу 4

$$0 \leq x_1 \leq 2; \quad 0 \leq x_2 \leq 5; \quad 1 \leq x_3 \leq 5$$

и задачу 5

$$0 \leq x_1 \leq 2; \quad 0 \leq x_2 \leq 5; \quad 0 \leq x_3 \leq 0.$$

4. Выбираем задачу 4 при $Z_0^4 = 0$. Оптимальный план этой задачи

$$Z_0 = 15; \quad x_1 = x_2 = 1/3; \quad x_3 = 1.$$

Вновь переходим к шагу 4, выбираем x_2 и вносим в основной список две задачи:

задачу 6

$$0 \leq x_1 \leq 2; \quad 1 \leq x_2 \leq 5; \quad 1 \leq x_3 \leq 5$$

и задачу 7

$$0 \leq x_1 \leq 2; \quad 0 \leq x_2 \leq 0; \quad 1 \leq x_3 \leq 5.$$

5. $Z_0^5 = 0$. Выбираем задачу 6. После чего в списке остаются еще две задачи 5 и 7. Обнаруживаем, что задача 6 не имеет допустимого решения. Возвращаемся к шагу 1, приняв $Z_0^6 = 0$.

6. Выбираем задачу 7. Ее оптимальный план:

$$Z_0 = 14^{6/7}; \quad x_1 = x_2 = 0; \quad x_3 = 1^{1/7}.$$

Выбираем x_3 и получаем задачу 8

$$0 \leq x_1 \leq 2; \quad 0 \leq x_2 \leq 0, \quad 2 \leq x_3 \leq 5$$

и задачу 9

$$0 \leq x_1 \leq 2; \quad 0 \leq x_2 \leq 0; \quad 1 \leq x_3 \leq 1.$$

7. Выбрав задачу 8, убеждаемся, что допустимого решения она не дает.

8. Выбор задачи 9 дает план

$$Z_0 = 13; \quad x_1 = x_2 = 0; \quad x_3 = 1.$$

На шаге 3 фиксируется этот план и принимается $Z_0^3 = 13$.

9. В основном списке осталась только задача 5.

Решив ее, находим:

$$Z_0 = 13; \quad x_1 = 2; \quad x_2 = 2^{1/3}; \quad x_3 = 0.$$

10. На шаге 1 прекращены вычисления, так как в списке задач больше нет.

Оптимальным планом задачи целочисленного программирования является оптимальный план задачи 9, зафиксированный на восьмой итерации.

Отметим, что число итераций, необходимых для решения задачи, зависит от удачного выбора задачи, вычеркиваемой из основного списка на шаге 1, и от выбора переменной, по которой формируются дополнительные задачи на шаге 4.

Ход вычислений можно представить графически блок-схемой в виде дерева. Отсюда, видимо, и название метода: ветвей — итераций и границ — оценок наименьшего Z_0^t .

Метод динамического программирования в целочисленных задачах. Для точного решения задач целочисленного программирования с малым числом «существенных ограничений» (не более двух-трех) часто оказываются полезными принципы динамического программирования. Продемонстрируем возможности этого метода на примере одномерной задачи о рюкзаке. Эта задача имеет вид

$$\sum_{j=1}^n c_j x_j \rightarrow \max; \quad (\text{VIII.30})$$

$$\sum_{j=1}^n a_j x_j \leq b_i; \quad (\text{VIII.31})$$

$$x_j \in [0, 1]; \quad j = \overline{1, n}. \quad (\text{VIII.32})$$

Здесь c_j, a_j, b — целые числа.

Функция Беллмана $f_k(z)$ — функция целочисленного аргумента z для $k = \overline{1, n}$ — представляет собой оптимальное значение целевой функции задачи, в которой параметры n и b заменены k и z . Функция Беллмана удовлетворяет следующее рекуррентное соотношение (уравнение Беллмана):

$$f_k(z) = \max [f_{k-1}(z - a_k h) + c_k h],$$

где максимум берется по h при условии

$$0 \leq h \leq h(z, k) = \left[\frac{z}{a_k} \right].$$

Учитывая, что $f_1(z) = c_1 \left[\frac{z}{a_1} \right]$, можно последовательно найти из уравнения Беллмана значения всех функций $f_k(z)$ для $0 \leq z \leq b$; $k = \overline{1, n}$. Зная величины $f_k(z)$, легко последовательно вычислить составляющие искомого решения x^* задачи о рюкзаке. Выберем $x_n^* \in [0, 1]$ из условия

$$c_n x_n^* + f_{n-1}(b - a_n x_n^*) = f_n(b).$$

Зная x_n^* , вычисляем $x_{n-1}^* \in [0, 1]$ из соотношения

$$c_{n-1} x_{n-1}^* + f_{n-2}(b - a_n x_n^* - a_{n-1} x_{n-1}^*) = f_{n-1}(b - a_n x_n^*)$$

и т. д. Значение $x_2^* \in [0, 1]$ находим из условия

$$c_2 x_2^* + f_1(b - a_n x_n^* - \dots - a_2 x_2^*) = f_2(b - a_n x_n^* - \dots - a_3 x_3^*).$$

и наконец, $x_1^* \in [0, 1]$ вычисляются по формуле

$$c_1 x_1^* = f_1(b - a_n x_n^* - \dots - a_2 x_2^*).$$

Определенные таким образом значения составляющих x_j^* удовлетворяют условиям задачи о рюкзаке. Суммируя n предыдущих равенств, убеждаемся в том, что

$$\sum_{j=1}^n c_j x_j^* = f_n(b),$$

т. е. найденный указанным образом вектор x^* с булевыми составляющими x_j^* является действительно решением задачи о рюкзаке.

Операции, связанные с вычислением x_j^* по известным значениям $f_k(z)$, оказываются излишними, если при вычислении правой части уравнения Беллмана каждый раз запоминать значение h , на котором достигается максимум. Общее число операций (сложение, сравнение и деление), необходимых для вычисления всех $f_k(z)$; $k = \overline{2, n}$, равно $2b(n-1)$. Решение одномерной задачи о рюкзаке с помощью полного перебора вариантов (их число C_n^b при $b < n$) требует bC_n^b операций. Ясно, что при немалых n и b и $b \leq n$ динамическое программирование существенно упрощает решение задачи.

Динамическое программирование оказывается также эффективным методом решения ряда специальных целочисленных задач. В работе [2] принцип динамического программирования использован для формирования алгоритма решения целочисленной распределительной задачи.

Общая схема динамического программирования эффективно используется в тех целочисленных задачах, в которых удается выразить последовательность рекуррентно связанных между собой функций Беллмана через возможно меньшее число аргументов.

Глава IX

ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

§ 45. Сущность и основные свойства

Для решения некоторых типов задач нелинейного программирования может быть применено динамическое программирование.

Динамическое программирование является сравнительно новым разделом математического программирования. Оно возникло в 50-е годы в результате изучения задач, в которых были существенные изменения во времени. Точнее, возникновение его связано с исследованием некоторых типов многошаговых процессов управления, особенно многошаговых и стохастических процессов, возникающих в теории создания запасов.

В дальнейшем аппарат динамического программирования был значительно усовершенствован и развит для решения задач, не связанных с временными изменениями.

Термин «динамическое программирование» относится скорее к вычислительному методу, чем к особому типу задач нелинейного программирования.

Таким образом, под динамическим программированием следует понимать вычислительный метод, опирающийся на аппарат рекуррентных соотношений, разработанных в значительной степени американским ученым Р. Беллманом [2, 4].

Идея метода динамического программирования состоит в том, что отыскание максимума (или минимума) функции многих переменных заменяется многократным отысканием максимума (минимума) функции одного или небольшого числа переменных.

Сущность метода динамического программирования сводится к составлению функциональных уравнений, управляющих процессом, и дальнейшему решению этих уравнений посредством нестандартных вычислительных процедур.

Функциональным называется такое уравнение, которое выражает функциональную связь между множеством функций.

Составление функциональных уравнений проводится на основе принципа оптимальности Р. Беллмана, сущность которого заключается в следующем.

Оптимальное поведение обладает тем свойством, что, каковы бы ни были начальное состояние и начальное решение, последующие решения должны составлять оптимальное поведение относительно состояния, полученного в результате первоначального решения.

Принцип оптимальности имеет конструктивный характер и непосредственно указывает процедуру нахождения оптимального

решения. Математически он записывается выражением вида

$$f_{n-l}(S_l) = \text{optimum} [R_{l+1}(S_l, U_{l+1}) + (f_{n-(l+1)}S_{l+1})], \quad (l = \overline{0, n-1}), \quad (\text{IX.A})$$

где $U_l = (u_l^{(1)}, \dots, u_l^{(m)})$ — решение (управление), выбранное на l -м шаге, $S_l = (s_l^{(1)}, \dots, s_l^{(m)})$ — состояние системы на l -м шаге, R_l — непосредственный эффект, достигаемый на l -м шаге, f_{n-l} — оптимальное значение эффекта, достигаемого за $n - l$ шагов, n — число шагов (этапов).

Optimum в выражении (IX.A) означает максимум или минимум в зависимости от условия задачи.

Все вычисления, дающие возможность найти оптимальное значение эффекта, достигаемого за n шагов, $f_n(S_0)$, проводятся по формуле (IX.A), которая носит название основного функционального уравнения Беллмана или рекуррентного соотношения. Действительно, при вычислении очередного значения функции f_{n-l} используются значения функции $f_{n-(l+1)}$, полученные на предыдущем шаге, и непосредственное значение эффекта $R_{l+1}(S_l, U_{l+1})$, достигаемого в результате выбора решения U_{l+1} при заданном состоянии системы S_l . Процесс вычисления значений функции f_{n-l} ($l = \overline{0, n-1}$) осуществляется при естественном начальном условии $f_0(S_n) = 0$, которое означает, что за пределами конечного состояния системы эффект равен нулю.

Вычислительная схема. Оптимальное решение задачи методом динамического программирования находится на основе функционального уравнения (IX.A). Чтобы определить его, необходимо:

1) записать функциональное уравнение для последнего состояния процесса (ему соответствует $l = n - 1$):

$$f_1(S_{n-1}) = \text{optimum}_{U_n} [R_n(S_{n-1}, U_n) + f_0(S_n)];$$

2) найти $R_n(S_{n-1}, U_n)$ из дискретного набора его значений при некоторых фиксированных S_{n-1} и U_n из соответствующих допустимых областей; так как $f_0(S_n) = 0$, то

$$f_1(S_{n-1}) = \text{optimum}_{U_n} [R_n(S_{n-1}, U_n)].$$

В результате после первого шага известно решение U_n и соответствующее значение функции $f_1(S_{n-1})$;

3) уменьшить значение l на единицу и записать соответствующее функциональное уравнение. При $l = n - k$ ($k = \overline{2, n}$) оно имеет вид

$$f_k(S_{n-k}) = \text{optimum}_{U_{n-k+1}} [R_{n-k+1}(S_{n-k}, U_{n-k+1}) + f_{k-1}(S_{n-k+1})]; \quad (\text{IX.B})$$

4) найти условно-оптимальное решение на основе выражения (IX.B);

5) проверить, чему равно значение l . Если $l = 0$, расчет условно-оптимальных решений закончен, при этом найдено оптимальное

решение задачи для первого состояния процесса. Если $l \neq 0$, перейти к выполнению п. 3;

6) вычислить оптимальное решение задачи для каждого последующего шага процесса, двигаясь от конца расчетов к началу.

Пример IX. 1. Требуется перевести груз из города *A* в город *B* (см. рис. 51). Стоимость перевозки груза из города *s* в город *j* представлена над соответствующими дугами сети. Необходимо найти маршрут, связывающий города *A* и *B*, для которого суммарные затраты на перевозку груза были бы наименьшими.

Решение. На рис. 52 вершинам сети поставлены в соответствие города, а дугам — транспортные магистрали.

Разобьем все множество вершин (городов) на подмножества. В первое подмножество включим исходную вершину 1. Во второе — вершины, в которые входят дуги, выходящие из вершины 1. В третье — вершины, в которые входят дуги, выходящие из вершин второго подмножества. Таким образом, мы получим пять подмножеств: {1}, {2, 3, 4}, {5, 6, 7}, {8, 9}, {10}. Очевидно, что любой маршрут из города 1 в город 10 содержит ровно четыре дуги, каждая из которых связывает вершины, принадлежащие к соответствующим подмножествам. Следовательно, процесс решения задачи (нахождения оптимального маршрута) разбивается на четыре этапа. На первом этапе принимается решение, через какой город, принадлежащий ко второму подмножеству, везти груз из города 1. На втором этапе необходимо определить, через какой город третьего подмножества везти груз из некоторого города, принадлежащего ко второму подмножеству, и т. д.

Перенумеруем этапы от конечного пункта сети к начальному (см. рис. 51) и введем обозначения: n — номер шага ($n = 1, 2, 3, 4$); $f_n(s)$ — минимальные затраты на перевозку груза от города *s* до конечного города, если до конечного города осталось n шагов; $j_n(s)$ — номер города, через который нужно ехать из города *s*, чтобы достичь $f_n(s)$; c_{sj} — стоимость перевозки груза из города *s* в город *j*.

Здесь все обозначения несут важную смысловую нагрузку: f — целевая функция; s — состояние системы (номер города); индекс n несет динамическую информацию о том, что из города *s* до конечного города осталось n шагов.

Предположим, что груз доставлен в город 10, следовательно, число оставшихся шагов равно нулю ($n = 0$) и $f_n(s) = f_0(10) = 0$, так как из города 10 груз везти не надо.

Рассмотрим последний шаг ($n = 1$) и вычислим для него значение функции. Очевидно, что в город 10 груз может быть доставлен или из города 8 или 9. Вычислим затраты на перевозку для этих двух состояний:

$$f_1(8) = c_{8,10} + f_0(10) = 5 + 0 = 5, \quad s = 8, \quad j_1(8) = 10;$$

$$f_1(9) = c_{9,10} + f_0(10) = 3 + 0 = 3; \quad s = 9; \quad j_1(9) = 10.$$

Чтобы провести расчет для $n = 2$, выдвинем гипотезы о месте нахождения груза: 1-я гипотеза — груз находится в городе 5; 2-я гипотеза — груз находится в городе 6; 3-я гипотеза — груз находится в городе 7.

Из города 5 в город 10 можно провезти груз или через город 8, или через город 9. Поэтому оптимальный маршрут из города 5 найдется из выражения

$$f_2(5) = \min_j [c_{5,8} + f_1(8), c_{5,9} + f_1(9)] = \min(9 + 5; 8 + 3) = 11.$$

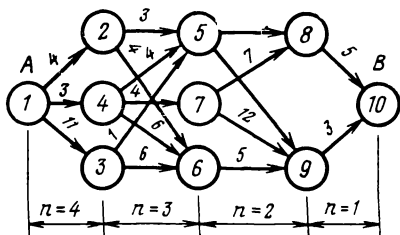


Рис. 52. Схема дорог, связывающих города *A* и *B*

Т а б л и ц а 38

s	j	$f_1 s$	$j_1(s)$
	10		
8	5 + 0	5	10
9	3 + 0	3	10

Т а б л и ц а 39

s	j		$f_2(s)$	$j_2(s)$
	8	9		
5	9+5	8+3	11	9
6		5+3	8	9
7	7+5	12+3	12	8

Здесь $s = 5$ и $j_2(5) = 9$, т. е. условно-оптимальный маршрут проходит через город 9.

Аналогично находим значения функций для $s = 6$ и $s = 7$:

$$f_2(6) = c_{8,9} + f_1(9) = 8;$$

$$f_2(7) = \min_j [c_{7,8} + f_1(8), c_{7,9} + f_1(9)] = 12.$$

Все вычисления удобно выполнять в таблицах. Расчеты первого [$n = 1$, $c_{sj} + f_0(j)$] и второго [$n = 2$, $c_{sj} + f_1(j)$] этапов помещены в табл. 38 и 39 соответственно.

Цифры в столбцах таблиц, находящиеся слева от двойной вертикальной черты, представляют собой сумму стоимости c_{sj} доставки груза из города s в город j и стоимости $f_{n-1}(j)$ доставки груза от города j до города B . В каждой строке выбирается наименьшая из этих сумм. Этим определяются условно-оптимальные затраты на доставку груза из города s в конечный город. Затраты (значение функции) обозначены $f_n(s)$ и записаны в первом столбце справа от вертикальной черты, а город, через который проходит условно-оптимальный маршрут, обозначен $j_n(s)$.

Рекуррентное соотношение для $n = 3$ имеет вид

$$f_3(s) = \min_{s,j} [c_{s,j} + f_2(j)].$$

Отметим, что для подсчета условно-оптимальных значений используется значение $f_2(j)$, полученное на предыдущем шаге, из табл. 39.

Вычисления для третьего шага [$n = 3$, $c_{s,j} + f_2(j)$] приведены в табл. 40. Здесь две клетки заштрихованы, поскольку из городов 2 и 3 нельзя попасть в город 7.

Вычисления для четвертого шага ($n = 4$, $c_{s,j} + f_3(j)$) приведены в табл. 41. Из табл. 41 видно, что минимальные затраты на перевозку груза

Т а б л и ц а 40

s	j			$f_3(s)$	$j_3(s)$
	5	6	7		
2	3+11	4+8		12	6
3	1+11	6+8		12	5
4	4+11	6+8	4+12	14	6

$f_4(1) = 16$ и оптимальный маршрут проходит через второй город, так как $j_4(1) = 2$. Далее из табл. 40 при $s = 2$ следует, что оптимальный маршрут проходит через город 6, так как $j_3(2) = 6$.

Т а б л и ц а 41

s	j			$f_4(s)$	$j_4(s)$
	2	3	4		
1	4 + 12	11 + 12	3 + 14	16	2

Продолжая рассмотрение таблиц, для $n = 2$ определяем, что оптимальный маршрут проходит через город 9 [$j_2(6) = 9$]. Наконец, из города 9 груз доставляется в конечный город 10 (место назначения). Таким образом двигаясь от последней таблицы к первой, мы определили оптимальный маршрут $\mu = (1-2-6-9-10)$, затраты на перевозку груза по которому составляют $f_4(1) = 4 + 4 + 5 + 3 = 16$.

§ 46. Задача по оптимизации распределения ресурсов

Рассмотрим одну из самых простых задач распределения ресурсов. Пусть имеется заданное исходное количество средств Z_0 , которое мы должны распределить между двумя предприятиями Π_1 и Π_2 в течение m лет.

Средства, вложенные в предприятие в начале какого-либо года, приносят за этот год определенный доход, зависящий от количества средств. Но эти средства, принося в течение года доход, частично тратятся, т. е. уменьшаются, и к концу года от них остается остаток, также зависящий от количества средств, вложенных в предприятие в начале года.

В табл. 42 в общем виде представлена форма исходных данных, приведены средства, вложенные в предприятие в начале i -го года и соответствующие им значения функций дохода и функций остатков. Предполагается, что новых средств не поступает и по истечении каждого года оставшиеся средства заново распределяются между предприятиями.

При таком предположении между величинами, приведенными в табл. 42, существуют следующие соотношения.

Суммарный доход за i -й год

$$W_i = f(x_i) + g(y_i). \quad (IX.1)$$

Т а б л и ц а 42

Предприятие	Количество, выделяемых средств	Функция	
		дохода	остатков
Π_1	x_i	$f(x_i)$	$\varphi(x_i)$
Π_2	y_i	$g(y_i)$	$\psi(y_i)$

Т а б л и ц а 43

Предприятие	Количество выделяемых средств на i -й год	Функция	
		дохода	остатков
Π_1	x_i	$(3-0,002x_i) x_i$	$0,6x_i$
Π_2	y_i	$(2-0,001y_i) y_i$	$0,8y_i$

Общее количество средств, вкладываемое в предприятие в начале i -го года

$$Z_{i-1} = x_i + y_i = \Phi(x_{i-1}) + \Psi(y_{i-1}) \quad (i = 1, 2, \dots, m). \quad (\text{IX.2})$$

При указанных условиях требуется так распределить средства между предприятиями по годам в течение m лет, чтобы суммарный доход за весь период (m лет) был наибольшим.

Рассмотрим, как может быть решена такая задача методом динамического программирования.

Процесс решения этой задачи будем развертывать в обратном по времени направлении. Сначала условно-оптимально спланируем распределение средств в последний m -й год, затем оптимально спланируем распределение средств в $(m - 1)$ год и т. д., пока не дойдем до 1-го года. Каждый год планирования будем называть этапом операции.

Первый этап. Функция суммарного дохода за m -й год будет

$$W_m = f(x_m) + g(y_m). \quad (\text{IX.3})$$

Максимальный суммарный доход за m -й год получим в результате решения задачи максимизации сепарабельной функции двух переменных x_m, y_m (IX.3) при линейном ограничении $x_m + y_m = Z_{m-1}$ и неотрицательности переменных x_m и y_m . Этот максимум будет являться функцией от Z_{m-1} :

$$\max W_m = W_m^*(Z_{m-1}). \quad (\text{IX.4})$$

Задаваясь различными значениями Z_{m-1} , мы можем каждый раз найти оптимальное решение (x_m^*, y_m^*) задачи максимизации функции (IX.3) при этом Z_{m-1} . Таким образом, мы можем получить либо аналитическое выражение функции (IX.4), либо таблицу ее значений, либо график этой функции. Каждому значению Z_{m-1} будет соответствовать условная оптимальная стратегия (x_m^*, y_m^*) . На этом заканчивается первый этап решения задачи.

Второй этап. На втором этапе требуется найти условную оптимальную стратегию (x_{m-1}^*, y_{m-1}^*) . Согласно принципу оптимальности Р. Беллмана, функция суммарного дохода за $(m - 1)$ -й и m -й годы должна слагаться из функций суммарного дохода за $(m - 1)$ -й год и максимального дохода за m -й год, т. е.

$$W_{m-1, m} = W_{m-1} + \max W_m = f(x_{m-1}) + g(y_{m-1}) + W_m^*(Z_{m-1}). \quad (\text{IX.5})$$

Согласно соотношению (IX.2)

$$Z_{m-1} = \Phi(x_{m-1}) + \Psi(y_{m-1}),$$

поэтому

$$W_{m-1, m} = W_{m-1, m}(x_{m-1}, y_{m-1}), \quad (\text{IX.6})$$

т. е. является функцией двух неотрицательных аргументов x_{m-1}, y_{m-1} , которые должны удовлетворять линейному ограничению

$$Z_{m-2} = x_{m-1} + y_{m-1}. \quad (\text{IX.7})$$

Таким образом, на втором этапе требуется максимизировать функцию (IX.6) при линейном ограничении (IX.7) и неотрицательности переменных x_{m-1} , y_{m-1} . Но это точно такая же задача, которая возникла на первом этапе. Для каждого Z_{m-2} могут быть найдены условная оптимальная стратегия (x_{m-1}^*, y_{m-1}^*) и соответствующий ей условный максимальный доход:

$$\max W_{m-1, m} = W_{m-1, m}^*(Z_{m-2}). \quad (\text{IX.8})$$

Т р е т ь и й э т а п. На третьем этапе аналогичным образом находится условная оптимальная стратегия (x_{m-2}^*, y_{m-2}^*) для предпоследнего года в результате максимизации суммарного дохода за последние 3 года:

$$W_{m-2, \dots, m} = W_{m-2} + \max W_{m-1, m} = f(x_{m-2}) + g(y_{m-2}) + \\ + W_{m-1, m}^*[\Phi(x_{m-2}) + \Psi(y_{m-2})] \quad (\text{IX.9})$$

при условии

$$Z_{m-3} = x_{m-2} + y_{m-2}; \quad x_{m-2} \geq 0; \quad y_{m-2} \geq 0. \quad (\text{IX.10})$$

Точно так же, продолжив процесс условной оптимизации, получим для любого $(m-i)$ -го года условную оптимальную стратегию (x_{m-i}^*, y_{m-i}^*) и соответствующий ей условный максимальный доход за все годы, начиная с данного.

$$W_{m-i, \dots, m}^*(Z_{m-i-1}) = \max W_{m-i, m}(x_{m-i}, y_{m-i}), \quad (\text{IX.11})$$

где

$$W_{m-i, \dots, m} = f(x_{m-i}) + g(y_{m-i}) + W_{m-i+1, m}^*[\Phi(x_{m-i}) + \Psi(y_{m-i})]. \quad (\text{IX.12})$$

Когда проведем условную оптимизацию всех годов, кроме первого, останется найти оптимальную стратегию, т. е. оптимальное распределение средств на первый год, и найти максимальный полный суммарный доход за все m лет. Планирование на первый год качественно отлично от остальных, так как здесь мы будем исходить из известного начального запаса средств Z_0 , в то время как при планировании последующих лет средства Z_{i-1} изменялись.

Для получения функций дохода за весь период надо в общей формуле (IX.12) положить $i = m - 1$, тогда

$$W_{1m} = f(x_1) + g(y_1) + W_{2m}^*[\Phi(x_1) + \Psi(y_1)]. \quad (\text{IX.13})$$

Для получения оптимальной стратегии для первого года следует максимизировать функцию (IX.13) двух аргументов x_1 , y_1 при условии

$$Z_0 = x_1 + y_1; \quad x_1 \geq 0; \quad y_1 \geq 0. \quad (\text{IX.14})$$

Оптимальное решение x_1^* , y_1^* этой задачи не будет условно-оптимальным, а будет просто оптимальной стратегией, т. е. опти-

мальным распределением средств в начале первого года между предприятиями Π_1 и Π_2 , и $W = \max W_{1m}$ будет максимальным доходом за весь период (m лет).

Зная оптимальное распределение средств в начале первого года, мы можем найти оптимальное распределение средств между предприятиями во все последующие годы. Для этого надо снова пройти все годы, но уже в обратном направлении — от начала к концу. Получив x_1, y_1 , мы можем вычислить:

$$Z_1 = \varphi(x_1) + \psi(y_1). \quad (\text{IX.15})$$

По ранее полученной зависимости оптимальной стратегии (x_2^*, y_2^*) от значений Z_1 находим оптимальное распределение средств (x_2, y_2) в начале второго года, по которому мы можем найти средства

$$Z_2 = \varphi(x_2) + \psi(y_2), \quad (\text{IX.16})$$

по которым найдем оптимальное распределение средств (x_3, y_3) в начале третьего года и т. д. вплоть до последнего года. Таким образом будет найдено окончательное решение задачи.

Мы рассмотрели общее решение задачи о распределении ресурсов между двумя объектами. Все проведенные рассуждения не изменятся, если мы в соотношениях (IX.1), (IX.2) будем брать не два, а n слагаемых, соответствующих n предприятиям:

$$W_i = \sum_{j=1}^n f_j(x_{ij}); \quad (\text{IX.17})$$

$$Z_{i-1} = \sum_{j=1}^n x_{ij} = \sum_{j=1}^n \varphi_j(x_{i-1, j}). \quad (\text{IX.18})$$

где $f_j(x)$ — функции дохода; $\varphi_j(x)$ — функции остатка для j -го предприятия; x_{ij} — количество средств, выделяемых j -му предприятию в начале i -го года.

Рассмотрим пример на условных данных.

Положим, что планируется работа двух предприятий Π_1 и Π_2 на 4-летний период. Функции дохода и функции остатков в начале i -го года представлены в табл. 43.

Требуется распределить ресурсы, исходная величина которых $Z_0 = 1000$ (с точностью до 0,1), между предприятиями Π_1 и Π_2 на каждый год планируемого периода, так чтобы получить максимальный суммарный доход за весь период.

Первый этап решения. Условная оптимальная стратегия (x_4^*, y_4^*) на последнем, четвертом году (количество средств, выделяемых предприятиям) находится как оптимальное решение задачи максимизации нелинейной функции дохода на четвертом году:

$$W_4(x_4, y_4) = (3 - 0,002 x_4) x_4 + (2 - 0,001 y_4) y_4 \quad (\text{IX.19})$$

при линейных ограничениях

$$x_4 + y_4 = Z_3, \quad x_4 \geq 0; \quad y_4 \geq 0, \quad (\text{IX.20})$$

где Z_3 — суммарный остаток средств по прошествии 3 лет, считающийся в этой задаче постоянным фиксированным числом.

Функция $W_4(x_4, y_4)$ строго вогнутая, поэтому существует единственная точка (x_4, y_4) , в которой эта функция достигает своего максимума. Если не учитывать условия неотрицательности переменных x_4 и y_4 , то точка максимума функции (IX.19) легко может быть найдена по методу множителей Лагранжа. Для этого надо составить функцию Лагранжа и приравнять все ее частные производные к нулю. В результате должна получиться система уравнений, из которой определяется точка максимума.

Составляем функцию Лагранжа:

$$F(x_4, y_4, \lambda) = (3 - 0,002 x_4) x_4 + (2 - 0,001 y_4) y_4 + \lambda (x_4 + y_4 - Z_3). \quad (\text{IX.21})$$

Вычисляем частные производные этой функции и приравниваем их к нулю:

$$\left. \begin{aligned} \frac{\partial F}{\partial x_4} &= 3 - 0,004 x_4 + \lambda = 0; \\ \frac{\partial F}{\partial y_4} &= 2 - 0,002 y_4 + \lambda = 0; \\ \frac{\partial F}{\partial \lambda} &= x_4 + y_4 - Z_3 = 0. \end{aligned} \right\} \quad (\text{IX.22})$$

Получилась система трех линейных уравнений с тремя неизвестными x_4 , y_4 и λ . Поскольку значения множителей Лагранжа λ нас не интересуют, мы можем их исключить почленным вычитанием второго уравнения из первого. В результате получаем систему двух линейных уравнений:

$$\left. \begin{aligned} 2 x_4 - y_4 &= 500; \\ x_4 + y_4 &= Z_3. \end{aligned} \right\} \quad (\text{IX.23})$$

Единственно решение этой системы

$$x_4 = \frac{1}{3} (Z_3 + 500); \quad y_4 = \frac{1}{3} (2 Z_3 - 500). \quad (\text{IX.24})$$

Это решение будет удовлетворять условию неотрицательности, если $Z_3 \geq 250$. Следовательно, выражение (IX.24) будет представлять собой условную оптимальную стратегию на четвертый год планирования при любом $Z_3 \geq 250$.

Подставляя выражения (IX.24) в формулу (IX.19) и приводя подобные члены с Z_3 и Z_3^2 , получаем условный максимальный доход за четвертый год:

$$\max W_4(Z_3) = \frac{1}{3} (250 + 7 Z_3 - 0,002 Z_3^2). \quad (\text{IX.25})$$

Второй этап решения. Перейдем к рассмотрению средств на третий год. Пусть мы подошли к нему с запасом средств Z_2 . Найдем условную оптимальную стратегию (x_3^*, y_3^*) на третий год и условный максимальный доход за два последних года.

По принципу оптимальности Р. Беллмана, функция дохода за

последние 2 года должна слагаться из функции дохода за третий год и максимальной функции дохода (IX.25) за четвертый год:

$$W_{3.4} = W_3(x_3, y_3) + \max W_4(Z_3) = (3 - 0,002 x_3) x_3 + \\ + (2 - 0,001 y_3) y_3 + \frac{1}{3} (250 + 7 Z_3 - 0,002 Z_3^2).$$

Но

$$Z_3 = 0,6 x_3 + 0,8 y_3 \quad (\text{IX.26})$$

и, следовательно,

$$W_{3.4} = (3 - 0,002 x_3) x_3 + (2 - 0,001 y_3) y_3 + \\ + \frac{1}{3} [250 + 7 (0,6 x_3 + 0,8 y_3) - 0,002 (0,6 x_3 + 0,8 y_3)^2]. \quad (\text{IX.27})$$

Условная оптимальная стратегия (x_3^*, y_3^*) найдется в результате максимизации функции (IX.27) при ограничениях

$$x_3 + y_3 = Z_2; \quad x_3 \geq 0; \quad y \geq 0. \quad (\text{IX.28})$$

Отвлекаясь от условия неотрицательности переменных и применяя метод Лагранжа, получаем систему линейных уравнений, которой должна удовлетворять условная оптимальная стратегия (x_3, y_3) на третий год

$$\left. \begin{aligned} -144 x_3 + 83 y_3 &= 15000; \\ x_3 + y_3 &= Z_2. \end{aligned} \right\} \quad (\text{IX.29})$$

Решение этой системы

$$\left. \begin{aligned} x_3 &= 0,366 Z_2 - 66,1; \\ y_3 &= 0,634 Z_2 + 66,1 \end{aligned} \right\} \quad (\text{IX.30})$$

представляет собой условную оптимальную стратегию на третий год при любом $Z_2 \geq 181$.

Подставляя выражения (IX.30) в формулу (IX.27) и приводя подобные члены с Z_2 и Z_2^2 , получаем условный максимальный доход за последние два года (третий и четвертый):

$$\max W_{3.4}(Z_2) = 34,8 + 4,062 Z_2 - 0,001022 Z_2^2. \quad (\text{IX.31})$$

Т р е т и й э т а п р е ш е н и я. Найдем распределение средств на второй год. Пусть мы подошли к нему с запасом средств Z_1 . Задача отыскания условной оптимальной стратегии (x_2^*, y_2^*) на второй год по своему решению ничем не отличается от решения аналогичной задачи на втором этапе. Оптимальная стратегия на второй год найдется как оптимальное решение задачи максимизации дохода за последние 3 года (второй, третий и четвертый)

$$W_{2.3.4}(x_2, y_2) = W_2(x_2, y_2) + \max W_{2.4}(Z_2) = (3 - 0,002 x_2) x_2 + \\ + (2 - 0,001 y_2) y_2 + 34,8 + 4,062 Z_2 - 0,001022 Z_2^2, \quad (\text{IX.32})$$

где

$$Z_2 = 0,6 x_2 + 0,8 y_2 \quad (\text{IX.33})$$

при ограничениях

$$x_2 + y_2 = Z_1; \quad x_2 \geq 0; \quad y_2 \geq 0. \quad (\text{IX.34})$$

Применение метода Лагранжа для решения этой задачи приводит к следующей системе уравнений:

$$\left. \begin{aligned} 37 x_2 - 23 y_2 &= 1876; \\ x_2 + y_2 &= Z_1. \end{aligned} \right\} \quad (\text{IX.35})$$

Условная оптимальная стратегия на второй год является решением этой системы:

$$\left. \begin{aligned} x_2 &= 0,383 Z_1 + 31,3; \\ y_2 &= 0,617 Z_1 - 31,3 \end{aligned} \right\} \quad (\text{IX.36})$$

при любом $Z \geq 51$.

Для отыскания условного максимального дохода за последние 3 года надо выражения (IX.36) подставить в формулу (IX.32) и привести подобные члены с Z_1 и Z_1^2 , в результате чего имеем

$$\max W_{2,3,4}(Z_1) = 37,7 + 5,321 Z_1 - 0,001209 Z_1^2. \quad (\text{IX.37})$$

Четвертый этап решения. На этом последнем этапе мы найдем не условную, а действительно оптимальную стратегию (x_1^*, y_1^*) планирования на первый год, так как мы будем теперь исходить не из предполагаемого, а из определенного наличия начальных средств $Z_0 = 1000$. Функция дохода за весь период по принципу оптимальности имеет выражение

$$\begin{aligned} W(x_1, y_1) &= W_1(x_1, y_1) + \max W_{2,3,4}(Z_1) = (3 - 0,002 x_1) x_1 + \\ &+ (2 - 0,001 y_1) y_1 + 37,7 + 5,321 Z_1 - 0,001209 Z_1^2, \end{aligned} \quad (\text{IX.38})$$

где

$$Z_1 = 0,6 x_1 + 0,8 y_1. \quad (\text{IX.39})$$

Чтобы определить оптимальное распределение начальных средств $Z_0 = 1000$, надо максимизировать функцию полного дохода (IX.38) при ограничениях

$$x_1 + y_1 = 1000; \quad x_1 \geq 0; \quad y_1 \geq 0. \quad (\text{IX.40})$$

Применяя метод Лагранжа, получаем следующую систему линейных уравнений:

$$\left. \begin{aligned} -19 x_1 + 12 y_1 &= 321; \\ x_1 + y_1 &= 1000. \end{aligned} \right\} \quad (\text{IX.41})$$

Эта система имеет единственное решение (с точностью до 0,1):

$$x_1 = 376,7; \quad y_1 = 623,3, \quad (\text{IX.42})$$

которое и представляет собой начальную оптимальную стратегию.

Подставляя значения (IX.42) в формулу (IX.38), получаем максимальный суммарный доход за весь четырехлетний период:

$$\max W(x_1, y_1) = 4963,1. \quad (\text{IX.43})$$

Таким образом, в начале четырехлетнего периода предприятию Π_1 должно быть выделено средств $x_1 = 376,7$, а предприятию Π_2 — остальные начальные средства $y_1 = 623,3$.

Для определения оптимальных стратегий в последующие годы надо пройти все этапы в обратном направлении.

Вычисляем по формуле (IX.39)

$$Z_1 = 0,6 \times 376,7 + 0,8 \times 623,3 = 724,7.$$

Но тогда мы можем вычислить по формулам (IX.36) значения x_2 и y_2 :

$$x_2 = 0,383 \times 724,7 + 31,3 = 308,9;$$

$$y_2 = 0,617 \times 724,7 - 31,3 = 415,8.$$

Значит, в начале второго года предприятию Π_1 выделяется средств 308,9, а предприятию Π_2 — 415,8.

Зная значения x_2 и y_2 , мы можем вычислить Z_2 по формуле (IX.33):

$$Z_2 = 0,6 \times 308,9 + 0,8 \times 415,8 = 518,0 > 181,$$

затем по формулам (IX.30) вычисляем значения x_3 и y_3

$$x_3 = 0,366 \times 518,0 - 66,1 = 123,5;$$

$$y_3 = 0,634 \times 518,0 + 66,1 = 394,5.$$

В начале третьего года предприятию Π_1 выделяется средств 123,5, а предприятию Π_2 — 394,5.

Вычисляем по формуле (IX.26) остаток средств после третьего года

$$Z_3 = 0,6 \times 123,5 + 0,8 \times 394,5 = 389,7 > 250$$

и по формулам (IX.24) определяем значения x_4 и y_4

$$x_4 = \frac{1}{3} (389,7 + 500) = 296,6;$$

$$y_4 = \frac{1}{3} (2 \times 389,7 - 500) = 93,3.$$

Итак, в последний год предприятию Π_1 должно быть выделено средств 296,6, а предприятию Π_2 — 93,3.

Полный результат решения этой задачи представим в виде табл. 44.

Т а б л и ц а 44

Год	Оптимальный план распределения ресурсов		Доход
	Π_1	Π_2	
1-й	376,7	623,3	1704,4
2-й	308,9	415,8	1394,6
3-й	123,5	394,5	973,4
4-й	296,6	93,3	891,8
	Суммарный доход		4964,2

При таком распределении суммарный доход получился на 1,1 больше ранее рассчитанного (IX.43). Это объясняется ошибками округлений при вычислениях.

Из табл. 44 видно, что относительная доля средств, выделяемых первому предприятию, меньше в течение первых 3 лет, в последний же год первому предприятию выделяется средств относительно больше, чем второму, в три с лишним раза.

§ 47. Сущность сетевых методов планирования и управления производством

Методы сетевого планирования и управления (СПУ) применяются для решения задач упорядочения и координации. Содержанием этих задач является планирование сложных комплексов работ и управление ими [16].

Комплексом работ называют некоторую совокупность последовательно и параллельно выполняемых работ, направленных на достижение намеченных целей. Комплексы работ, которые являются объектами СПУ, обладают следующими свойствами: 1) работы, входящие в комплекс, взаимно обуславливают друг друга. Иначе говоря, одни работы не могут быть начаты раньше, чем завершатся некоторые другие работы; 2) в определенных интервалах времени отдельные работы можно начинать и заканчивать независимо друг от друга (при наличии соответствующих ресурсов); 3) существуют одна или много целей функционирования комплекса работ.

Примерами таких комплексов работ (и соответственно областей применения СПУ) могут быть строительство или реконструкция крупных промышленных и гражданских объектов, капитальный ремонт сложного оборудования и машин, подготовка и освоение производства новых видов продукции, организация научно-исследовательских и опытно-конструкторских работ, материально-технического снабжения крупных предприятий, разработка техпромфинплана предприятия, планов НОТ и др.

Методы СПУ представляют собой систему, объединяющую расчетные методы, организационные мероприятия и приемы контроля, направленные на разработку и оптимизацию плана выполнения комплекса работ, оперативное управление им и систематический контроль за его ходом. Отличительной особенностью СПУ является представление плана работ в виде сетевых графиков (схем). Сетевой график — это графическая модель выполнения комплекса работ, отображающая взаимосвязь между отдельными работами, их параметры и последовательность выполнения.

Полная характеристика комплекса работ содержится в сетевой модели. В такой модели должна быть отражена вся информация о комплексе: цель его функционирования, состав и порядок выполнения работ, условия, при которых оно возможно, и т. д.

Благодаря системному подходу СПУ обеспечивает значительное улучшение качества планирования и, в частности, установление наиболее целесообразной последовательности и сроков выполне-

Признак	Модели системы СПУ
Состав и характер содержащейся в модели информации	Детерминированные и вероятностные
Число технологически независимых комплексов работ	Односетевые и многосетевые
Число самостоятельных результатов	Одноцелевые и многоцелевые
Состав параметров управления	С учетом времени, стоимости ресурсов

ния работ, рациональное распределение и использование материальных и трудовых ресурсов, возможность гибкого оперативного управления производством.

Сетевые методы используют для решения не только частных задач, но и систем СПУ. Система СПУ — это система организационного управления, реализующая функции планирования и управления комплексом работ на основе построения, анализа и оптимизации сетевых моделей. Основным методом решения плановых и управленческих задач в рамках системы СПУ является метод СПУ, что не исключает применения для решения отдельных задач и других экономико-математических методов [21].

В табл. 45 приведена характеристика систем СПУ по характеру моделей.

Математической основой сетевых графиков является теория графов, основанная на топологических принципах и теории множеств. Исторически первой формулой, относящейся к топологии, является формула Эйлера о выпуклом многограннике

$$\alpha_0 - \alpha_1 + \alpha_2 = 2,$$

где α_0 — число вершин выпуклого многогранника, α_1 — число ребер, α_2 — число граней.

Т о п о л о г и я — это область математики, изучающая такие свойства фигур, которые взаимно однозначны и непрерывны. Геодезическим примером топологического преобразования является случай, когда все грани многогранника треугольные и в результате проектирования их на плоскость получается триангуляция с прямолинейными или криволинейными сторонами.

Г р а ф как математическое понятие есть некоторое множество, состоящее из точек, соединенных прямыми или кривыми линиями. Другими словами, графом называют совокупность конечного числа точек, называемых вершинами графа, и линий, называемых дугами или ребрами графа, попарно соединяющих некоторые из этих вершин. Область ориентированного (направленного) графа называют с е т е в ы м г р а ф и к о м. Под ориентированным гра-

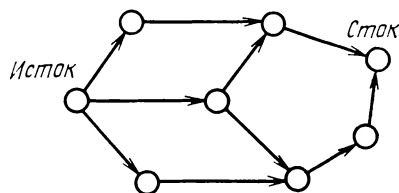


Рис. 53. Ориентированный граф

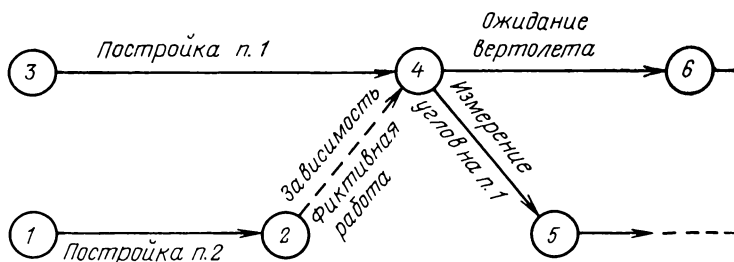


Рис. 54. Отражение зависимости (фиктивных работ) на сетевом графике (работа изображена сплошной линией, фиктивная работа — прерывистой)

фом понимают такой, у которого для всех его ребер (дуг) стрелками указаны направления (рис. 53). В последние годы вопросы топологии и теории графов стали применять при уравнивании геодезических сетей и решении некоторых инженерно-экономических задач. Например, пользуясь теорией графов, можно проводить оценку геометрических особенностей отдельных геодезических сетей, уравнивать аналитические сети с применением конформного преобразования, решать системы линейных уравнений, определять веса искомым неизвестных, решать транспортные и другие инженерно-экономические задачи методами математического программирования.

Теория графов развивалась совершенно независимо от сетевого планирования. Начало теории графов было положено Эйлером еще в 1736 г. Когда в 1958 г. были изобретены сетевые графики, оказалось, что они представляют собой направленные графы, т. е. такие, которые не содержат циклов, имеют вершину (исток), в которую не входит ни одна дуга, и вершину (сток), из которой не выходит ни одна дуга. В сетевом планировании стали использовать готовые результаты теории графов, а теория графов получила много задач, вытекающих из ее нового применения.

В расчетах по сетевому планированию мы не будем доказывать теорем и воспользуемся готовыми положениями теории графов, давая качественный анализ рассматриваемых задач.

Сетевое планирование обычно осуществляется по параметру время, т. е. минимизируются процессы по времени и реже по параметрам время — стоимость, время — ресурсы, время — стоимость — ресурсы. В методах СПУ вместо традиционной линейной формы дается более обозримая сетевая форма, которая моделирует систему и позволяет достаточно полно отражать ее состояние и различные текущие изменения.

Элементы сетевого графика. Сетевая модель содержит три важных элемента: работу, или операцию, событие и путь (критический путь). Р а б о т а — это прежде всего трудовой процесс, требующий затрат времени и ресурсов (операция 4,5). К работе относят также и ожидание, т. е. пассивный процесс, на который не затрачи-

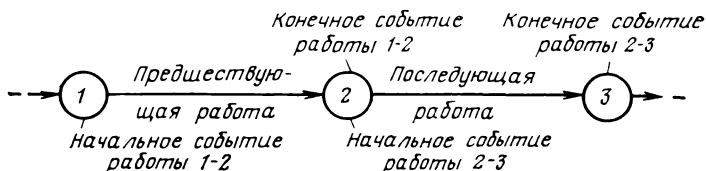


Рис. 55. Моменты производственного процесса

вается труд, но затрачивается время. Например, когда закончена постройка знака и бригада ожидает транспорт для переброски на другой пункт (операция 4,6 на рис. 54). Работой также называется и просто зависимость, другими словами логическая связь между двумя событиями. Такую работу называют *ф и к т и в н о й р а б о т о й*, так как на нее не требуется затрат ни времени, ни труда, ни материальных ресурсов. Например, измерение углов на пункте 1 не может быть начато раньше, чем будут построены пункты 1 и 2 (операция 2, 4 на рис. 54).

С о б ы т и е представляет собой итог какой-либо деятельности, промежуточный или окончательный результат выполнения одной или нескольких предшествующих работ. Например, событие 4 (см. рис. 54) означает, что постройка пункта 1 закончена. Событие не обладает длительностью, не требует никаких затрат, совершается мгновенно. На сетевом графике события обычно показываются кружками. На любом сетевом графике выделяют *и с х о д н о е* и *з а в е р ш а ю щ е е* события. Первое служит моментом начала производственного процесса, второе — моментом его окончания. Все остальные события на графике называются *п р о м е ж у т о ч н ы м и*. Событие называется *н а ч а л ь н ы м*, если оно служит началом какой-либо работы, и *к о н е ч н ы м*, если оно является окончанием какой-либо работы. Ясно, что одно и то же событие, за исключением исходного и завершающего, служит начальным событием для следующей за ним работы и конечным событием для предшествующей работы (рис. 55). Каждое событие имеет номер, который проставляется внутри кружка. Не должно быть двух событий с одинаковыми номерами.

Построение и анализ сетевой модели осуществляются в несколько этапов: 1) уяснение и изучение процесса, для которого составляется сетевая модель, подготовка исходных данных; 2) построение первичного сетевого графика; 3) расчет параметров сетевого графика; 4) анализ и оптимизация сетевого графика; 5) определение сроков выполнения всех работ, переход к календарному графику.

Рассмотрим пример [3]. Пусть в период подготовки к полевому сезону партии предстоит выполнить некоторые организационные мероприятия (табл. 46).

№ п. п.	Обозначение работ	Состав работ	Продолжительность работы (в днях)
1	А	Изучение и анализ заказ-наряда партии. Распределение объемов работ по бригадам	12
2	Б	Составление маршрутов передвижения бригад. Их утверждение	8
3	В	Укомплектование бригад оборудованием, материалами и снаряжением	7
4	Г	Написание и выдача технических предписаний бригадам	5
5	Д	Составление оргплана партии	10

Первый этап — уяснение и изучение процесса, для которого составляется сетевая модель, подготовка исходных данных. Организационные технические мероприятия, представленные в табл. 46, можно изобразить на ленточном графике Ганта (рис. 56). Для этого следует установить, в какой зависимости друг от друга находятся мероприятия А, Б, В, Г, Д. Предположим, что между этими мероприятиями существуют следующие связи: работа Б может быть начата спустя три дня после начала работы А и должна быть закончена на три дня раньше работы А, работа В может быть начата после выполнения половины работы Б, работу Г можно начинать сразу же после окончания работы А и необходимо закончить за три дня до конца работы Д, работа Д может быть начата только после окончания работы Б, работу В следует закончить за пять дней до окончания работы Д и работа А должна быть закончена за два дня до конца работы В.

Ленточный график (см. рис. 56) в совокупности с приведенными описаниями в принципе дает полное представление о порядке проведения организационных мероприятий. Но видно, что такое описание явно неудобно. Кроме того, проводить какой-либо анализ с ленточным графиком трудно. Гораздо проще представить ленточный график в виде сетевого графика (см. рис. 63, 64).

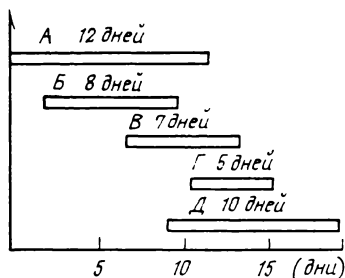


Рис. 56. Ленточный график Ганта

В сетевых графиках эти недостатки устраняются и, кроме того, появляются новые возможности в нахождении правильных решений. Однако и сетевые графики имеют недостатки, которые могут отсутствовать в ленточных графиках. Поэтому при применении сетевого планирования используют и различные ленточные диаграммы.

К первому этапу относится подготовка исходных

д а н н ы х. Каждая работа (операция) характеризуется продолжительностью, которая обозначается $t(i, j)$, где (i, j) — код работы (операции), i — код (номер) начального события, j — код (номер) конечного события работы.

При установлении продолжительности работ в сетевом графике существуют два подхода: детерминистический, когда для каждой работы указывается одна оценка времени (чаще всего это официальный норматив), и вероятностный, когда оценка времени выполнения работы таит некоторую (а иногда и значительную) долю неопределенности.

Однозначная оценка времени дает хорошие результаты в таких программах, где существуют нормы на отдельные производственные процессы или где имеется большой опыт реализации сходных работ. Даже в проектах, посвященных разработке новых технологий или технических средств, встречаются «стандартные» работы. В этих случаях детерминистический подход себя вполне оправдывает.

Однако мы можем встретить довольно большое число работ (особенно в научных исследованиях, опытно-конструкторских разработках и др.), когда неопределенность во времени играет существенную роль. Продолжительность выполнения топографо-геодезических работ зачастую носит случайный характер, поскольку она зависит от таких факторов, как погода, категория трудности и т. д. Если для работы нет норм или опыта осуществления, нельзя заранее дать точную оценку продолжительности работы и, следовательно, детерминистический подход к составлению графика непригоден. В этих случаях применяется вероятностный подход.

При вероятностном подходе используются некоторые, интуитивно подсказываемые, границы (верхняя и нижняя) для продолжительности каждой работы. Это так называемые пессимистическое (b_{ij}) и оптимистическое время (a_{ij}) выполнения работы (i, j) .

Пессимистическое время b_{ij} — максимальная продолжительность выполнения работы, когда предусматриваются все, даже маловероятные задержки, а оптимистическое a_{ij} — минимальная продолжительность выполнения данной работы, которая может иметь место при наиболее благоприятных условиях. Планируемую продолжительность работы, которую называют о ж и д а е м о й, вычисляют по формуле

$$t_{ij}^{(e)} = \frac{3a_{ij} + 2b_{ij}}{5},$$

которая вытекает из некоторых вероятностных соображений.

Иногда оказывается возможным установить и наиболее в е р о я т н о е время m_{ij} выполнения работы $(i \rightarrow j)$. В этих случаях ожидаемая продолжительность данной работы (снова из вероятностных соображений) будет

$$t_{ij}^{(e)} = \frac{a_{ij} + 4m_{ij} + b_{ij}}{6}.$$

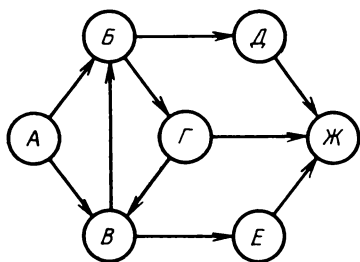


Рис. 57. Сетевой график

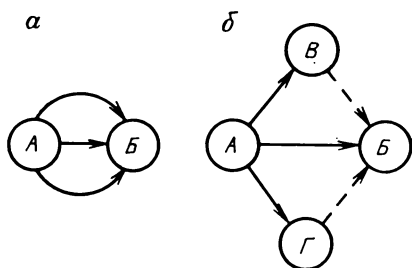


Рис. 58. Сетевой график составлен а — неправильно, б — правильно

Для оценки разброса используется величина дисперсии, которая вычисляется по формуле

$$\sigma_{ij}^2 = \left(\frac{b_{ij} - a_{ij}}{6} \right)^2.$$

Определив продолжительность работы $t_{ij}^{(e)}$ для всех видов работ, в дальнейшем сетевую модель составляют как в детерминированном случае.

В подготовку исходных данных также входит определение ресурсов (трудовых, материальных и стоимостных), необходимых для выполнения каждой операции. Одновременно уточняются сроки завершения как производственного процесса в целом, так и отдельных промежуточных операций.

В т о р о й э т а п — построение первичного сетевого графика. Существует несколько самых общих правил, относящихся к составлению и построению сетевого графика. Приведем основные из них.

1. В сетевом графике не должно быть замкнутых контуров (циклов), т. е. путей, которые начинаются и заканчиваются в одном и том же событии (рис. 57), если в сетевом графике обнаруживается контур, то это свидетельствует об ошибке, допущенной при построении. Действительно, для того чтобы свершилось, например, событие Г, требуется выполнить работу Б, Г, но, в свою очередь, для осуществления события Б — начала работы Б, Г — требуется выполнить работу В, Б, а значит, и работу ГВ, а выполнение последней возможно лишь после свершения события Г.

В небольших сетях замкнутые контуры обнаруживаются «на глаз». При анализе сетевого графика на ЭВМ для обнаружения и ликвидации контуров используются специальные программы.

2. В сетевом графике не должно быть работ, имеющих одинаковые коды, т. е. работ с одним и тем же предшествующим событием и одним и тем же последующим событием (рис. 58, а). Если $k (>1)$ работ исходят из одного и того же события и входят в одно и то же событие, то, для того чтобы каждую пару событий соединила не больше чем одна работа, необходимо ввести $k - 1$ дополнительных событий, а затем их следует соединить с последующим событием фиктивными работами (на рис. 58, б эти работы изображены прерывистой линией).

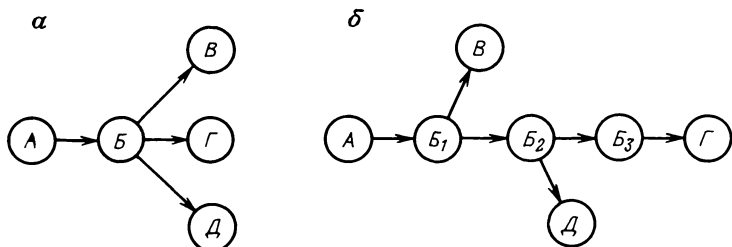


Рис. 59. Сетевой график составлен *а* — неправильно, *б* — правильно

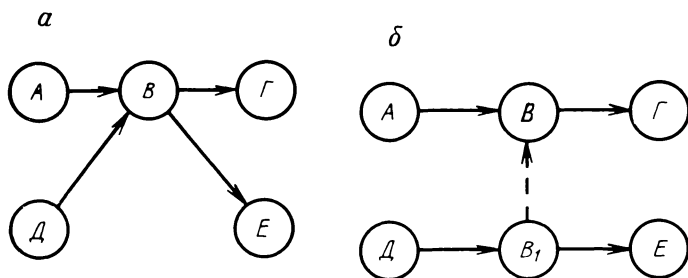


Рис. 60. Сетевой график составлен *а* — неправильно, *б* — правильно

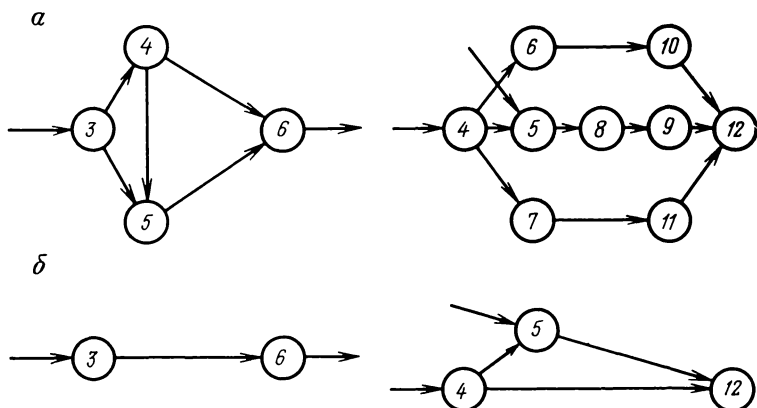


Рис. 61. Сетевой график: *а* — до укрупнения, *б* — после укрупнения

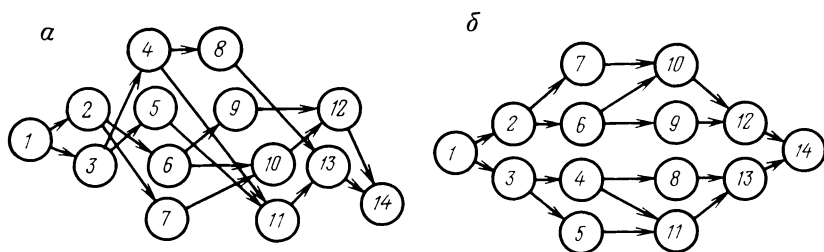


Рис. 62. Два варианта одного и того же сетевого графика:
а — неудачный, *б* — удачный

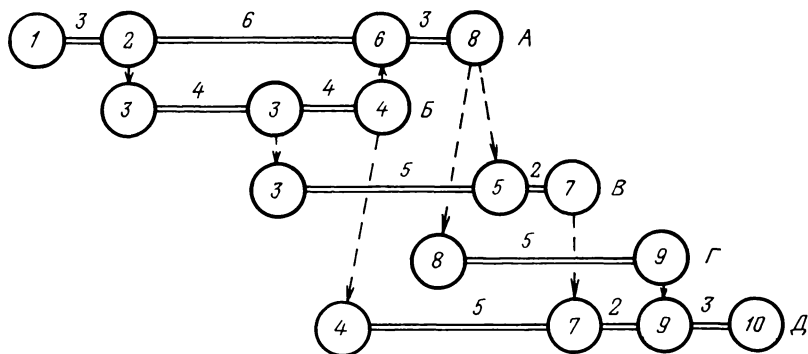


Рис. 63. График Ганта с указанием связей между операциями

3. Все работы в сетевом графике должны быть простыми. Это означает следующее: если какие-либо работы в сети (например, работы БВ, БД на рис. 59, а) могут быть начаты еще до полного окончания непосредственно предшествующей им работы АВ, то работа АВ должна быть разбита на несколько последовательных работ (АВ₁, Б₁ Б₂, Б₂Б₃ на рис. 59, б).

4. Если для выполнения одной из работ (например, работы ВГ на рис. 60, а) необходимо получение результатов всех работ, входящих в начальное для нее событие В, а для другой работы ВЕ необходимо окончание только одной из этих работ (например, ДВ), то для отражения этих обстоятельств следует ввести дополнительное событие В₁ и фиктивную работу В₁В, связывающую новое событие с прежним (рис. 60, б).

5. Если внутри большого сетевого графика можно выделить некоторую подсеть, планирование которой может быть произведено независимо от всей сети, то можно составить укрупненную сеть. Для этого следует заменить намеченную подсеть одной работой, продолжительность которой равна критическому пути подсети (рис. 61, б).

6. При построении сетевого графика полезно обращать внимание на взаимное расположение событий. На рис. 62, а приведен пример неудачного изображения сети: из-за неправильного взаимного расположения событий создается ложное впечатление о сложности комплекса. Но стоит переставить несколько событий, как сразу обнаруживаются две почти самостоятельные ветви (рис. 62, б).

Займемся теперь построением сетевого графика для нашего примера. Нанесем на ленточный график Ганта (см. рис. 56) пунктирными стрелками все связи, существующие между работами А, Б, В, Г, Д.

В результате получим новый график (рис. 63), от которого легко перейти к сетевому графику (рис. 64). Стрелки на графике не имеют масштаба, их величина выбирается такой, чтобы полученный сете-

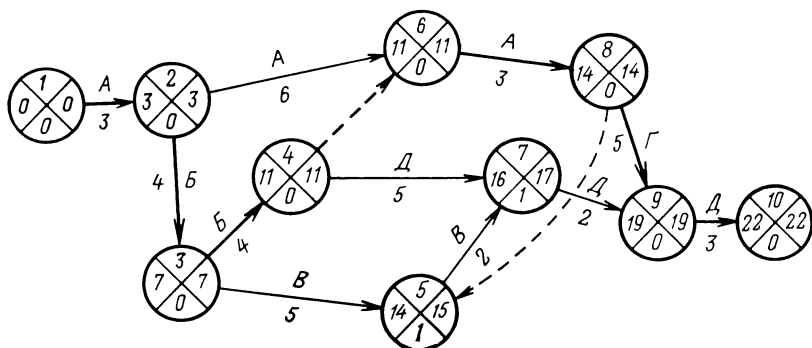


Рис. 64. Сетевой график выполнения работ, перечисленных в табл. 46.

В верхнем секторе кружка указан номер события, ранний срок наступления событий показан в левом секторе, поздний срок — в правом секторе; под каждой работой указана ее продолжительность. В нижнем секторе каждого кружка показан резерв времени i -го события R_i .

вой график был наглядным и удобочитаемым. Под каждой работой показана ее продолжительность.

Третий этап — расчет параметров сетевого графика. При расчете сетевого графика определяют следующие основные временные параметры сетей: ранние и поздние сроки наступления (свершения) событий, сроки начала и окончания работ и резервы времени, событий и работ.

Ранний срок наступления любого последующего события j можно определить по максимальной величине суммы раннего срока наступления предшествующего события i и продолжительности работы $t(i, j)$, т. е.

$$t_p(j) = \max_i [t_p(i) + t(i, j)]. \quad (X.1)$$

Используем эту формулу в нашем примере. Обычно ранний срок наступления исходного события принимают равным нулю, т. е. $t_p(1) = 0$.

Тогда ранний срок наступления второго события

$$t_p(2) = t_p(1) + t(1, 2) = 0 + 3 = 3.$$

Для остальных событий $t_p(j)$

$$t_p(3) = t_p(2) + t(2, 3) = 3 + 4 = 7;$$

$$t_p(4) = t_p(3) + t(3, 4) = 7 + 4 = 11;$$

$$t_p(6) = \max_{i=2, 4} [t_p(2) + t(2, 6), t_p(4) + t(4, 6)] = \max [3 + 6, 11 + 0] = 11;$$

$$t_p(8) = t_p(6) + t(6, 8) = 11 + 3 = 14;$$

$$t_p(5) = \max_{i=3, 8} [t_p(3) + t(3, 5), t_p(8) + t(8, 5)] = \max [7 + 5, 14 + 0] = 14;$$

$$t_p(7) = \max_{i=4, 5} [t_p(4) + t(4, 7), t_p(5) + t(5, 7)] = \max [11 + 5, 14 + 2] = 16;$$

$$t_p(9) = \max_{i=7, 8} [t_p(7) + t(7, 9), t_p(8) + t(8, 9)] = \max [16 + 5, 14 + 2] = 19;$$

$$t_p(10) = t_p(9) + t(9, 10) = 19 + 3 = 22.$$

Таким образом, алгоритм закончился определением раннего срока конечного события $t_p(n) = t_p(10)$, и тем самым найдено критическое время, так как $T_{кр} = t_p(n)$. Что касается критического пути (или критических путей), то он находится попутно.

Поздний срок наступления любого предшествующего события i легко определяется по минимальной величине разности между поздним сроком наступления последующего события j и продолжительностью работы, связывающей эти события (i, j) :

$$t_n(i) = \min_j [t_n(j) - t(i, j)]. \quad (X.2)$$

Принимая в нашем примере за поздний срок наступления завершающего события величину, равную 22 единицам, вычислим поздние сроки всех остальных событий по формуле (X.2):

$$t_n(9) = t_n(10) - t(9, 10) = 22 - 3 = 19;$$

$$t_n(7) = t_n(9) - t(7, 9) = 19 - 2 = 17;$$

$$t_n(5) = t_n(7) - t(5, 7) = 17 - 2 = 15;$$

$$t_n(8) = \min_{j=9, 5} [t_n(9) - t(8, 9), t_n(5) - t(5, 8)] = \min [19 - 5, 15 - 0] = 4;$$

$$t_n(6) = t_n(8) - t(6, 8) = 4 - 3 = 11;$$

$$t_n(4) = \min_{j=6, 7} [t_n(6) - t(4, 6), t_n(7) - t(4, 7)] = \min [11 - 0, 17 - 5] = 11;$$

$$t_n(3) = \min_{j=4, 5} [t_n(4) - t(3, 4), t_n(5) - t(3, 5)] = \min [11 - 4, 15 - 5] = 7;$$

$$t_n(2) = \min_{j=3, 6} [t_n(3) - t(2, 3), t_n(6) - t(2, 6)] = \min [11 - 6, 7 - 4] = 3;$$

$$t_n(1) = t_n(2) - t(1, 2) = 3 - 3 = 0.$$

Следует запомнить, что ранние сроки свершения событий находятся по формуле (X.1) «движением вперед» от начального события, тогда как поздние сроки находятся по формуле «движением вспять» от конечного события.

Резерв времени R_i наступления любого события равен разности между поздним и ранним сроком наступления этого события, т. е.

$$R(i) = t_n(i) - t_p(i). \quad (X.3)$$

Так, резерв времени события 5 для нашего примера

$$R(5) = t_n(5) - t_p(5) = 15 - 14 = 1 \text{ дню.}$$

Как показано на рис. 64, одни события имеют нулевой резерв времени, другие, например события 5 и 7, имеют отличный от нуля положительный резерв времени. Понятно, что если события имеют нулевой резерв времени, то задержка их наступления хотя бы на один день повлечет за собой удлинение времени выполнения всего процесса работ. В то же время задержка в наступлении события 5 на один день не изменит общей продолжительности работ. Поэтому события, имеющие наименьший (в нашем случае нулевой) резерв времени, называют критическими. Путь, проходящий через критические события, называется критическим. На рис. 64

критический путь изображен жирными линиями, он проходит через события 1, 2, 3, 4, 6, 8, 9, 10. Продолжительность критического пути $T_{кр}$ определяет время окончания производственного процесса. Это наибольшая сумма времени, требуемого на выполнение всех работ по одному или нескольким путям от исходного до завершающего события. В сетевом графике может быть не один, а несколько критических путей. Основное внимание руководства должно быть сосредоточено на работах критического пути, которые не имеют резервов. Одновременно с этим в поле зрения руководителя должны находиться подкритические пути, близкие по продолжительности к критическим, а также наименее напряженные пути.

Разность между продолжительностью критического пути $T_{кр}$ и продолжительностью $T(L_n)$ любого пути L_n называют резервом времени $R(L_n)$ этого пути:

$$R(L_n) = T_{кр} - T(L_n). \quad (X.4)$$

Зная $t_p(i)$ и $t_n(i)$ для всех событий сети, можно для любой работы (i, j) определить следующие параметры:
ранний срок начала работы

$$t_{рн}(i, j) = t_p(i), \quad (X.5)$$

поздний срок начала работы

$$t_{пн}(i, j) = t_n(j) - t_{\{i, j\}}, \quad (X.6)$$

полный резерв времени работы

$$R_n(i, j) = t_{пн}(j) - t_p(i) - t(i, j). \quad (X.7)$$

Четвертый этап — анализ и оптимизация сетевого графика. После вычисления всех временных параметров графика получают исходный сетевой план, который затем анализируют и корректируют в соответствии с заданными сроками и имеющимися возможностями его выполнения. При анализе сетевого графика в первую очередь обращают внимание на критические пути, от которых в основном зависит выполнение всего комплекса работ.

На основе полученных параметров сетевого графика проводят его анализ, оценивая при этом структуру сети, загрузку исполнителей, возможные изменения начала работ и другие факторы. В этих целях сначала анализируют работы, используя найденные размеры их резервов.

Оптимизацию можно проводить по следующим критериям: а) директивным срокам окончания всего производственного процесса в целом и отдельных его работ, б) материальным и трудовым ресурсам, в) стоимости.

Первый тип ограничений вызывает необходимость сокращения продолжительности критического пути. Это может быть выполнено несколькими путями. Первый путь — анализ и сокращение продолжительности работ, лежащих на критическом пути. Второй путь — изменение сети благодаря параллельному выполнению работ. И третий путь — изменение топологии сети путем частичного изменения технологии производственного процесса.

Анализ и оптимизацию сетевого графика по ресурсам сделаем на примере графика, изображенного на рис. 64. Для этой цели все события сетевого графика расположим на оси времени, т. е. в календарной сетке. Далее прочертим все работы в масштабе времени. Длина каждой стрелки — работы соответствует ее протяженности. Затем прочертим все связи, существующие между работами. В результате получим график, показанный на рис. 65, а. На нем под каждой операцией указано необходимое число ИТР для ее выполнения. Используя полученный график, нетрудно подсчитать число ИТР, занятых в каждый момент времени. График использования ИТР изображен на рис. 65, б.

Предположим, что нам задано ограничение по ресурсам, что число ИТР для выполнения работ (см. рис. 64) не должно превышать пять человек. Чтобы уменьшить число ИТР и сгладить график их использования, переместим (см. рис. 65, а) работы (5, 7) и (7, 9') на один день, а работу (3, 5') — на три дня. Такое перемещение показано на рис. 66, а. Снова строим график использования ИТР (рис. 66, б). Теперь видно, что для выполнения работ в нашем примере требуются пять ИТР вместо шести. Аналогично выполняются анализ и оптимизация использования любых ресурсов.

Отметим, что строгого алгоритма оптимизации сетевых графиков по всем показателям пока не найдено. Но при корректировке графика по одному какому-нибудь критерию (например, минимизации сроков или стоимости) используются методы линейного программирования или специальным алгоритмом Келли, или Фулкерсона, или другим алгоритмом [4].

Экономико-математическую модель задачи минимизации стоимости выполнения работ по сетевому графику после его анализа можно записать так:

$$F = \sum_{(i, j) \in n} c_{ij} x_{ij} \rightarrow \min;$$

$$t(i, j) \min \leq x_{ij} \leq t(i, j) \max \text{ для всех } x_{ij};$$

$$\sum_{(i, j) \in n} x_{ij} \leq T_{кр} \text{ для всех путей,}$$
(X.8)

где F — стоимость всего комплекса работ, (i, j) — код работы (операции), $c_{i, j}$ — стоимость операции за один день, n — множество, представляющее собой номера путей, x_{ij} — продолжительность операции в днях, $t_{ij} \min$ — минимальная возможная продолжительность операции, $t_{ij} \max$ — максимальная возможная продолжительность операции, $T_{кр}$ — продолжительность критического пути.

П я т ы й э т а п — определение сроков выполнения всех работ, переход к календарному графику. В результате оптимизации сетевого графика по времени, ресурсам и стоимости необходимо выбрать рациональное время начала каждой работы и вычертить затем календарный график подобно тому, как это сделано на рис. 65, а и 66, а.

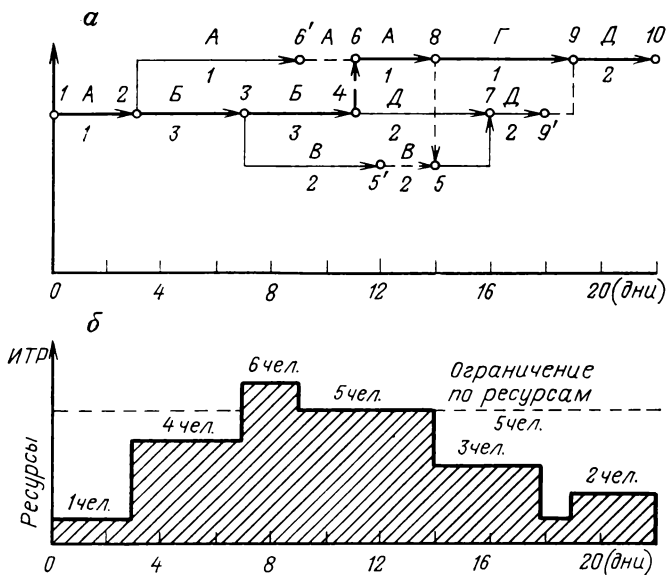


Рис. 65. Сетевой график в календаре. Анализ использования ресурсов

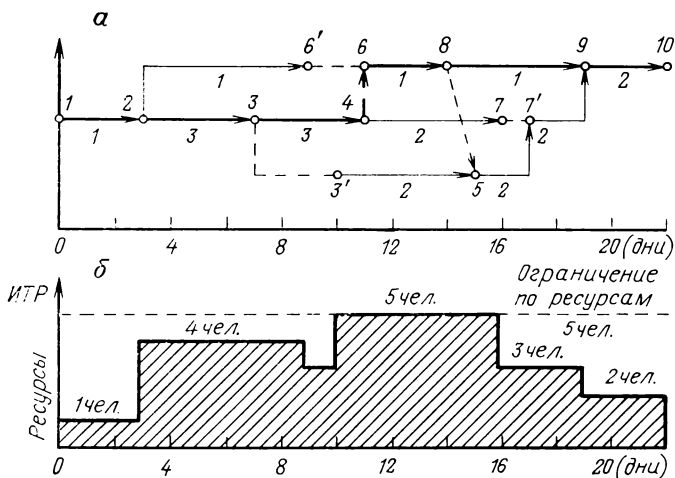


Рис. 66. Сглаживание ресурсов по сетевому графику

§ 48. Решение транспортной задачи на основе сетевых графиков

Некоторые задачи оптимизации сетевых графиков можно сформулировать так, что их можно решить методами линейного программирования. Так, транспортную задачу можно графически изобразить в виде сети на карте или схеме и решить методом потенциалов.

Рассмотрим элементарный пример решения транспортной задачи на основе сетевых графиков. Пусть имеются три пункта отправления: A_1 , A_2 и A_3 , в которых находятся соответственно $a_1 = 8$, $a_2 = 11$ и $a_3 = 10$ единиц груза, и три пункта назначения B_1 , B_2 и B_3 , куда надо доставить $b_1 = 6$, $b_2 = 9$ и $b_3 = 14$ единиц груза, при этом $\sum a_i = \sum b_j$. Эти данные и затраты c_{ij} на перевозку единицы груза от пункта A_i до пункта B_j напомним на рис. 67.

В кружках запишем мощности поставщиков, а в прямоугольниках — спрос потребителей и будем называть их вершинами сети (графа).

На линиях (дорогах, ребрах), соединяющих вершины, запишем затраты c_{ij} на перевозку единицы продукции от пунктов A_i до пунктов B_j . Такая совокупность вершин и линий, их соединяющих, есть граф, имеющий ребра с оценками.

Цель задачи — найти такие неотрицательные значения x_{ij} , при которых суммарная стоимость перевозок будет минимальна. В нашем случае требование это можно записать так:

$$Z = \sum_{i=1}^3 \sum_{j=1}^3 c_{ij} x_{ij} \rightarrow \min. \quad (\text{X.9})$$

Исходный план будем составлять, начиная с вершины A_3 , из которой надо вывозить продукцию $a_3 = 10$ единиц. Значение c_{3j}

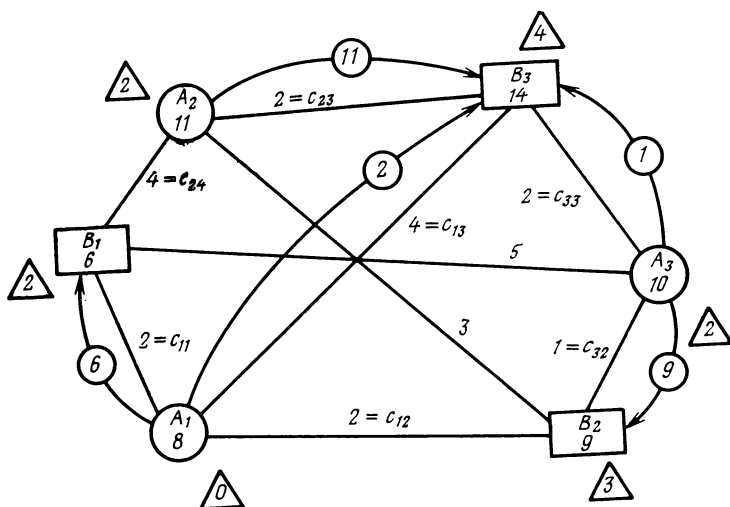


Рис. 67. График исходного плана

является наименьшим по линии A_3B_2 , поэтому в пункт B_2 направляем максимально возможное количество продукции $x_{32} = 9$, записывая ее в стрелке. В пункт B_3 направляем оставшуюся в пункте A_3 продукцию, т. е. полагаем $x_{33} = 1$, и также записываем ее в стрелке. Из вершины A_1 направляем груз в пункты B_1 и B_3 , полагая $x_{11} = 6$ и $x_{13} = 2$. Из вершины A_2 груз $a_2 = 11$ единицам остается направить в пункт B_3 , т. е. положить $x_{23} = 11$.

Мы получили базисный план поставок, который должен удовлетворять следующим требованиям: 1) все запасы должны быть распределены и все запросы удовлетворены; 2) к каждой вершине должна подходить или из нее выходить хотя бы одна стрелка; 3) общее число стрелок должно равняться $m + n - 1$, где m — число пунктов отправления, n — число пунктов получения; 4) стрелки не должны образовывать замкнутого контура, т. е. не должно быть циклов.

Полученное базисное решение проверяем на оптимальность при помощи потенциалов. Произвольно вершине A_1 присваиваем потенциал $\alpha_1 = 0$ и записываем его в треугольнике около этой вершины. Прибавляя к потенциалу вершины A_1 значения $c_{11} = 2$ и $c_{13} = 4$, получим потенциалы соответственно для вершин B_1 и B_3 , равные $\beta_1 = 2$ и $\beta_3 = 4$, которые записываем в треугольниках около соответствующих вершин.

К вершине B_3 подходят стрелки из пунктов A_2 и A_3 , поэтому для получения потенциала вершин A_2 и A_3 из потенциала вершины B_3 надо соответственно вычесть показатели $c_{23} = 2$ и $c_{33} = 2$. Получаем $\alpha_2 = 2$ и $\alpha_3 = 2$. Потенциал пункта B_2 вычисляем по формуле $\beta_2 = \alpha_3 + c_{32} = 2 + 1 = 3$; значение его записываем около вершины B_2 .

Проверка правильности вычисления потенциалов осуществляется путем двойного определения значения функции цели:

1) по формуле

$$Z_1 = \sum_{j=1}^n b_j \beta_j - \sum_{i=1}^m \alpha_i a_i = 12 + 27 + 56 - 0 - 22 - 20 = 53;$$

2) по формуле

$$\sum_{i,j} c_{ij} x_{ij} = 12 + 8 + 22 + 2 + 9 = 53.$$

По найденным потенциалам проверяем план на оптимальность путем вычисления характеристик $\sum_{ij} \epsilon_{ij}$ для всех ребер сети по формуле

$$\sum_{ij} \epsilon_{ij} = c_{ij} - (\beta_j - \alpha_i). \quad (X.10)$$

Очевидно, характеристики ребер, имеющих стрелки, должны равняться нулю. Характеристики для ребер, не имеющих стрелок, могут быть любыми: нулевыми, положительными и отрицательными.

Если хотя бы одно ребро имеет отрицательную характеристику, то план неоптимален. В нашем примере для ребра A_1B_2 имеем $\sum_{12} = c_{12} - (\beta_2 - \alpha_1) = 2 - (3 - 0) = -1$. Следовательно, состав-

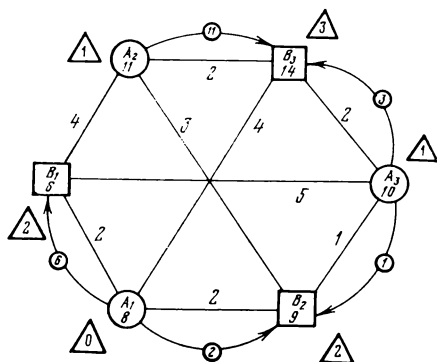


Рис. 68. График оптимального плана

внимание не принимается, но оно обязательно учитывается при перераспределении поставок.

Перед уточнением плана определяют направление новой стрелки в ребре, к которому построен контур (в нашем примере контур $A_1B_2A_3B_3$).

Новая стрелка в нашем примере имеет направление от пункта A_1 к пункту B_2 .

После этого просматривают все стрелки в контуре, направление которых противоположно направлению новой стрелки, и выбирают стрелку с наименьшей поставкой. В нашем примере это будет стрелка A_1B_3 с поставкой 2. Найденная величина поставки прибавляется ко всем поставкам, записанным в стрелках, имеющих одинаковое направление с новой стрелкой, и вычитается из поставок в стрелках противоположного направления. Полученные результаты записывают на новой сети (рис. 68). Поставки в стрелках, которые не входят в контур, остаются без изменения. Таким путем получают уточненный план, в котором общее число стрелок остается прежним и равным $m + n - 1 = 5$.

После этого проверяют план на оптимальность. Значение потенциала вершины A_1 можно положить равным 0 (можно взять любое значение), вычислить потенциалы всех остальных вершин и записать их в треугольниках около соответствующих вершин. Вычисленное двумя путями значение функции цели равно 51.

Для каждого ребра сети характеристики неотрицательны. Следовательно, план оптимален.

При решении транспортных задач на сети может встретиться много различных особенностей и затруднений, аналогичных тем, которые возникают при решении их табличными методами. Может встретиться вырождение, которое проявляется в том, что число стрелок окажется меньше $m + n - 1$. Вырождение преодолевается путем введения требуемого числа стрелок с нулевыми поставками так, чтобы можно было образовать допустимый контур.

В случае открытой модели вводят фиктивного потребителя или

вленный план неоптимален, и его надо улучшать, т. е. перераспределять поставки.

Для перераспределения поставок составляют замкнутый контур, состоящий из ребра без стрелки (характеристика которого отрицательна) и ребер со стрелками. Если план допустимый, то к каждому ребру без стрелок, имеющему отрицательную характеристику, можно построить контур и только один. При составлении контура направление стрелок во

поставщика с мощностью равной небалансу. При решении транспортных задач на сети можно так же, как и в случае матричных алгоритмов, использовать запрещенные поставки.

Для простой транспортной сети (без перевалочных пунктов) матричная поставка транспортной задачи эквивалентна сетевой ее поставке. Сетевой метод более нагляден и позволяет лучше учесть густоту движения и пропускную способность на отдельных участках сети.

Глава XI

ПРИНЯТИЕ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

§ 49. Классификация задач принятия решений в условиях неопределенности и обзор методов их решения

Задача принятия решений в условиях неопределенности определялась нами ранее как задача выбора оптимальной стратегии в операции, исход которой помимо стратегий оперирующей стороны и ряда фиксированных факторов (детерминированных или стохастических, или тех и других вместе) зависит также от некоторых неопределенных факторов, неподвластных оперирующей стороне и неизвестных ей в момент принятия решения. Вследствие влияния неопределенных факторов каждой конкретной стратегии оперирующей стороны соответствует не единственный, как в детерминированном случае, а множество возможных исходов. Конкретная реализация исхода операции для каждой стратегии оперирующей стороны определяется конкретной реализацией неопределенных факторов (а также, конечно, конкретной реализацией фиксированных стохастических факторов, если в рассматриваемой операции они входят в состав фиксированных факторов, влияющих на исход операции).

Подчеркнем здесь еще раз принципиальное различие между фиксированными стохастическими факторами и неопределенными факторами. И те и другие факторы приводят к разбросу в возможных исходах операции при многократной реализации одной и той же стратегии оперирующей стороны в аналогичных условиях проведения операции. В этом их сходство между собой и отличие от детерминированных факторов. Различие же между фиксированными стохастическими факторами и неопределенными факторами состоит в том, что относительно первых исследователь операции располагает всей полнотой статистической информации; этой информации достаточно для определения вероятностей появления возможных исходов операции и принятия решения о выборе оптимальной «средней» стратегии. В этом смысле фиксированные стохастические факторы являются факторами, полностью изученными исследователем операции. Относительно неопределенных факторов исследователь операции подобной информацией не располагает.

В дальнейшем изложении для простоты и определенности условимся рассматривать в основном ЗПР в условиях неопределенностей в упрощенном, «чистом» виде, т. е. «очищенными» от влияния фиксированных стохастических факторов. Это позволит нам уделить основное внимание изучению влияния неопределенных факторов.

Конечно, подобные ЗПР представляют собой определенную идеализацию реальных операций, исходы которых обычно зависят от всех видов факторов.

В отношении вероятностей различных реализаций исходов операции возможны два случая: 1) вероятности возможных исходов операции не имеют физического смысла; 2) вероятности возможных исходов операции имеют физический смысл, но либо вовсе не известны исследователю операции, либо известны с недостаточной для принятия решений точностью. Первое соответствует неопределенным факторам нестохастической природы, которые не могут быть описаны в терминах теории вероятностей; второе — факторам стохастической (вероятностной) природы, неопределенность в отношении которых обусловлена их недостаточной изученностью.

1. В соответствии с характером причины, вызывающей неопределенность нестохастической природы, можно выделить две группы факторов:

а) стратегические неопределенности — неопределенные факторы, появляющиеся вследствие участия в операции нескольких оперирующих сторон, т. е. разумных, активно действующих участников операции, преследующих различные (несовпадающие) цели;

б) концептуальные неопределенности — неопределенные факторы, сопутствующие принятию особо сложных решений, имеющих долговременные и далеко идущие последствия, и связанные с нечеткими представлениями оперирующей стороны о своих собственных целях и целях других участников операции.

II. Неопределенные факторы стохастической природы имеют и другое распространенное название — природные неопределенности. Природные неопределенности — это неопределенные факторы, появляющиеся из-за недостаточной изученности оперирующей стороной «природы», связанной с рассматриваемой операцией. Под словом «природа» в теории принятия решений понимают всю совокупность обстоятельств, в условиях которых приходится принимать решения. Сюда можно отнести неизвестные характеристики некоторых процессов, связанных с протеканием операции, или неизвестные свойства объектов, участвующих в операции, а также неизвестные внешние условия выполнения операции.

ЗПР в условиях действия неопределенных стратегических факторов можно назвать многосторонними ЗПР, так как исход соответствующих операций зависит от деятельности нескольких оперирующих сторон. Несовпадение целей оперирующих сторон многосторонней операции создает конфликт между ее участниками, иначе конфликтную ситуацию принятия решений. Отсюда идет еще одно название подобных ЗПР — ЗПР в условиях конфликтной ситуации, или конфликтные ЗПР.

Конфликтные ЗПР можно подразделить на одноуровневые и многоуровневые.

Одноуровневые конфликтные ЗПР могут быть как антагонистическими, так и неантагонистическими. Антагонистическая ЗПР связана с операцией, в которой сталкиваются интересы двух сторон, преследующих прямо противоположные цели. Очевидно, что многоуровневые ЗПР не являются антагонистическими: исследование ситуаций прямого неподчинения и противодействия нижестоящих руководителей вышестоящим не является предметом математической теории принятия решений.

Обратимся к обзору математического аппарата, употребляемого для решения различных ЗПР в условиях неопределенности.

Жизненная необходимость анализа конфликтных ситуаций и принятия решений в условиях неопределенности, создаваемой активными действиями нескольких участников операций, вызвала к жизни специальный математический аппарат — теорию игр и теорию минимакса (максимина). Эти дисциплины изучают одноуровневые конфликты. Иерархические ЗПР являются наименее изученными классом конфликтных ЗПР [20].

Теорию игр обычно определяют как математическую теорию конфликтных ситуаций [20]. Это достаточно общее определение теории игр привело к довольно распространенному в настоящее время заблуждению. Многие считают, что любая конфликтная ситуация может быть формализована в терминах теории игр, т. е. представлена в виде игры и решена методами теории игр. При этом забывают о том, что методы теории игр разработаны применительно к специфическим конфликтным ситуациям, таким, которые обладают свойством многократной (а в теории бесконечной) повторяемости. Задача теории игр — выработка рекомендаций по выбору рационального образа действий участников многократно повторяющегося конфликта.

В отличие от теории игр теория минимакса представляет собой математический аппарат, который может быть использован для решения задач принятия решений в конфликтных антагонистических ситуациях, реализуемых ограниченное число раз или всего один раз и ведущих к тяжелым и долговременным последствиям. Теория минимакса является молодой отраслью математики, находящейся в стадии становления.

Перейдем к рассмотрению методов решения ЗПР в условиях природных неопределенностей, называемых также часто ЗПР в условиях неизвестного состояния природы. Математическим аппаратом решения этих задач является теория статистических решений. Для теории статистических решений часто используют другие, более образные названия, например «теория статистических игр» или «теория игр с природой». Чтобы подчеркнуть отличие теории статистических игр от собственно теории игр, занимающейся конфликтными ситуациями, для наименования последней иногда используют наименование «теория стратегических игр»; прямо указывающее на наличие в операции нескольких активных участников, применяющих свои собственные стратегии.

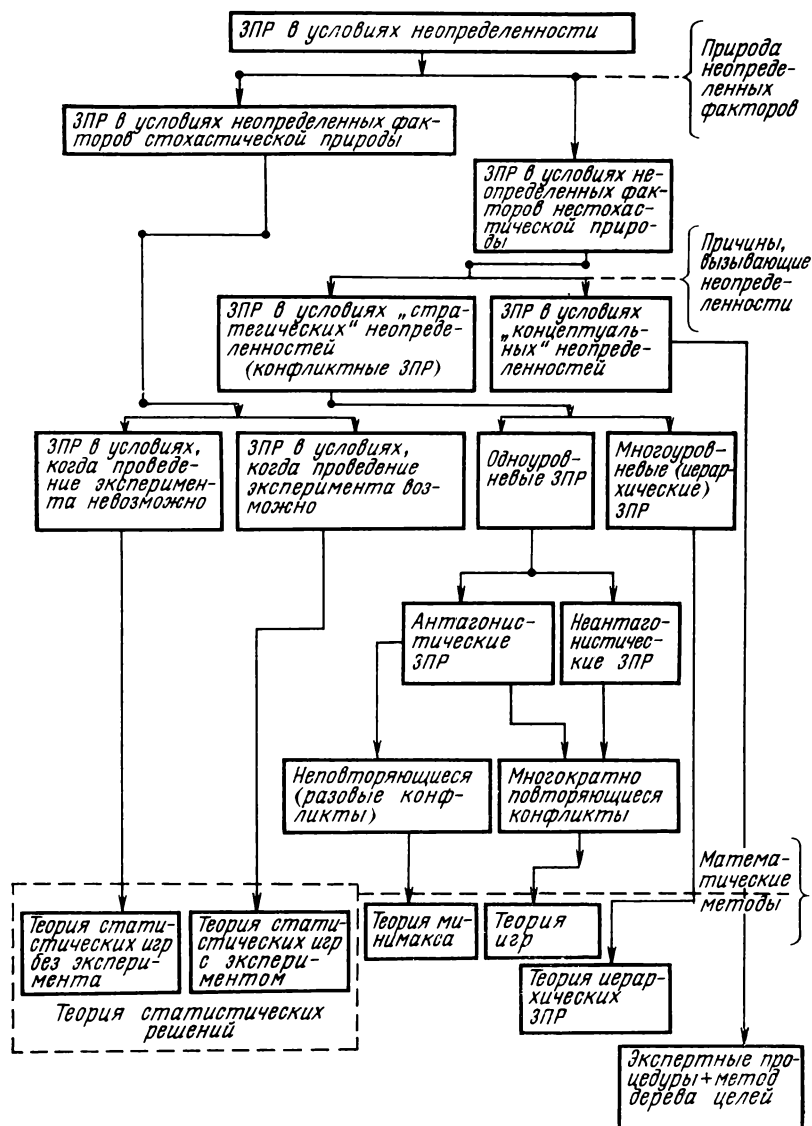


Рис. 69. Классификационная схема ЗПР в условиях неопределенности

В теории статистических решений обычно рассматриваются два случая: 1) операция такова, что не представляется возможным проведение каких-либо экспериментов с целью уточнения значений вероятностей возможных исходов операции; 2) проведение экспериментов возможно. Первый случай является наиболее сложным. Здесь решения принимаются на основании различных критериев

оптимальности, сформулированных в теории статистических решений. Выбор того или иного критерия является концептуальной проблемой. Соответствующие ЗПР часто называют статистическими играми без экспериментов. Во втором случае в связи с тем, что всякий эксперимент обладает стоимостью, возникает вопрос, сколько и каких экспериментов целесообразно предпринять для уточнения вероятностей возможных исходов. На эти вопросы отвечает теория планирования экспериментов, являющаяся основным содержанием теории статистических решений. Соответствующие ЗПР часто называют статистическими играми с экспериментами.

Подчеркнем здесь принципиальное отличие ЗПР в условиях конфликтной ситуации от других видов ЗПР в условиях неопределенности. С конфликтными ЗПР обычно связывают два предположения: 1) каждой оперирующей стороне многосторонней операции известны как цели, так и весь набор возможных стратегий всех других участников операции; 2) поскольку каждый участник многосторонней операции является разумным и активно действующим, то определяющим в его поведении является стремление к максимально возможному достижению собственных целей. Оба эти предположения определяют и одновременно упрощают подходы к принятию решений в конфликтных ЗПР. На них строятся, в частности, теория игр и теория минимакса.

В указанном смысле ЗПР в условиях «природных» неопределенностей являются принципиально более сложными. Здесь вторым участником операции будет «природа». Однако природу нельзя рассматривать как разумного, активно действующего участника операции; ей нельзя приписать никаких сознательно поставленных целей, к достижению которых она стремится. Тем самым неопределенность выступает здесь в более «тяжелой» форме. Однако природа развивается и «действует» в соответствии со своими объективно существующими законами. У человека есть возможность постепенно изучать эти законы, в частности, с помощью специальных экспериментов и тем самым снижать степень неопределенности. Это — ключ к решению ЗПР в условиях «природных» неопределенностей.

Проведенному здесь обзору ЗПР в условиях неопределенности и методов их решения соответствует классификационная схема («дерево») (рис. 69). Поскольку построенное классификационное «дерево» относится к классу неупорядоченных по уровням, то классификационные признаки (или названия уровней) в нем представлены лишь на отдельных уровнях.

§ 50. Принятие решений в условиях повторяющейся одноуровневой конфликтной ситуации. Элементы теории стратегических игр

Мы рассмотрим элементы теории игр как математического аппарата, который может быть использован для обоснования решений в условиях конфликтной ситуации.

Зарождение теории игр связано с именем известного американского математика Дж. фон Неймана, которого часто называют «отцом» теории игр. Еще в 1928 г. Дж. фон Нейман доказал основополагающую теорему теории игр — знаменитую теорему о минимаксе. Но только после публикации в 1944 г. первого фундаментального трактата по теории игр Дж. фон Неймана и О. Моргенштерна «Теория игр и экономическое поведение» [20], в котором впервые было дано систематическое, полное и строгое изложение теории игр, а также после появления быстродействующих ЭВМ, позволяющих решать громоздкие игровые задачи, эта теория стала бурно развиваться.

В настоящее время теория игр представляет собой емкую и развитую математическую дисциплину. Развита (в разной степени) теория антагонистических и неантагонистических (кооперативных) игр, конечных игр, бесконечных игр, позиционных игр. В последние годы помимо теории этих игр, которую можно уже считать классической, начала развиваться теория рефлексивных игр, связанная с именем советского математика Лефевра, а также теория дифференциальных игр. В рефлексивных играх существенным моментом являются задачи преследования одного управляемого объекта другим управляемым объектом; особенностью дифференциальных игр является учет динамики поведения объектов; их поведение описывается дифференциальными уравнениями.

Примеры конфликтных ситуаций весьма многообразны. Дж. фон Нейман и О. Моргенштерн в своей работе по теории игр ориентировались на анализ конфликтных ситуаций в вопросах капиталистической экономики, когда при наличии свободной конкуренции в роли борющихся сторон выступают торговые фирмы, промышленные предприятия, монополии и т. д. Однако конфликтные ситуации встречаются и во многих других областях человеческой деятельности. Любая ситуация, складывающаяся в ходе военных действий, принадлежит к конфликтным, так как здесь каждое решение должно приниматься с учетом сознательного противодействия разумного противника. К той же категории принадлежат и ситуации, возникающие при планировании любых военных действий. Интересными примерами конфликтных ситуаций являются спортивные состязания, арбитражные споры, аукционы, выборы в парламент при наличии нескольких кандидатов на одно место и т. д.

Основные понятия и методы теории игр мы изложим на примере простейшей и наиболее теоретически разработанной игры — конечной парной антагонистической игры. Данное изложение элементов теории игр в основном соответствует изложенному в литературе [6, 16].

1. *Основные понятия, используемые в теории игр двух лиц.* В игре участвуют два игрока A и B с определенными интересами. Сама игра — это ряд действий, или «ходов», игроков A и B , выполняемых по определенным правилам. Правила игры четко формулируются и должны включать описания возможных действий игроков, объем сведений, доступных каждой стороне, о поведении другой

стороны и результат (исход) игры, к которому приводит каждая совокупность ходов игроков.

Игра называется игрой с нулевой суммой (матричной игрой), если выигрыш игрока A в точности равен проигрышу игрока B (или наоборот).

Игроки A и B проводят ряд последовательных ходов, т. е. осуществляют ряд последовательных действий, предусмотренных правилами игры. Все ходы подразделяются на личные и случайные. Личным ходом называется выбор и осуществление самим игроком одного из предписанных правилами действий. Случайным ходом называется выбор и осуществление одного из предписанных правилами действий, которые проводит не сам игрок, а некоторый механизм случайного выбора (бросание монеты, вытаскивание номерков из барабана, использование таблиц случайных чисел). Для каждого случайного хода правила игры определяют распределение вероятностей возможных исходов.

Стратегия игрока — это совокупность правил, определяющих выбор варианта действий при каждом ходе в зависимости от ситуации, которая сложилась в ходе игры.

Теория игр дает указания игрокам при выборе ходов, т. е. рекомендует им лучшие стратегии.

Задание пары стратегий игроков A и B в игре двух лиц полностью определяет ее исход, т. е. выигрыш одного игрока и проигрыш другого. Игра называется конечной, если у каждого игрока имеется лишь конечное число стратегий. Если таких стратегий бесконечно много (пусть даже только у одного игрока), то игра называется бесконечной.

Целью теории игр является выработка оптимальных для игроков A и B стратегий, которые при многократном повторении игры обеспечивают игрокам максимально возможный средний выигрыш или минимально возможный средний проигрыш. Причем всегда при выборе стратегий предполагается, что если игрок A стремится максимально увеличить свой выигрыш, то игрок B делает все, чтобы помешать игроку A это сделать.

2. Платежная матрица. Цена игры. Принципы максимина и минимакса.

Перейдем непосредственно к описанию математического аппарата, используемого в теории игр. Рассмотрим конечную игру, в которой игрок A имеет m стратегий ($A_1, A_2, \dots, A_m$), а игрок B — n стратегий ($B_1, B_2, \dots, B_n$). Такая игра называется игрой $m \times n$. Если игроки A и B используют только личные ходы, то выбор стратегий A и B однозначно определяет исход игры a_{ij} , т. е. число, характеризующее выигрыш игрока A и проигрыш игрока B . Причем a_{ij} может быть и положительным, и отрицательным. Будем считать, что при $a_{ij} > 0$ игрок A выигрывает, а игрок B проигрывает величину a_{ij} . Если $a_{ij} < 0$, то, наоборот, выигрывает игрок B и проигрывает игрок A . В этом случае вместо проигрыша часто говорят об отрицательном выигрыше игрока A .

Если в игре используются случайные ходы, то выигрыш при

Стратегии A	Стратегии B			
	B_1	B_2	$\dots$	B_n
A_1	a_{11}	a_{12}	$\dots$	a_{1n}
A_2	a_{21}	a_{22}	$\dots$	a_{2n}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_m	a_{m1}	a_{m2}	$\dots$	a_{mn}

двух стратегиях A_i , B_j является случайным. В этом случае за оценку ожидаемого выигрыша берется его математическое ожидание.

Предположим, что нам известны все значения a_{ij} в игре $m \times n$. Эти значения удобно записать в виде таблицы платежей $m \times n$ матрицы (табл. 47), где строки соответствуют стратегиям A_i , а столбцы — стратегиям B_j .

Построение платежной матрицы не всегда просто, поскольку количества стратегий m и n могут оказаться очень большими. Однако в принципе любая конечная игра может быть приведена к матричному виду.

Рассмотрим простейший пример построения платежной матрицы. Разберем игру «поиск». Эта игра двух игроков A и B состоит в том, что игрок A прячется, а игрок B его ищет. В распоряжении игрока A имеются только два убежища, любое из которых он может использовать по своему усмотрению. Когда используется первое убежище, будем говорить, что игрок A выбрал стратегию A_1 . Когда выбрано второе убежище, то это значит, что выбрана стратегия A_2 . Аналогично игрок B может искать игрока A в первом или втором убежище, выбирая соответственно стратегии B_1 и B_2 . Договоримся об условиях игры: если игрок B находит игрока A , то последний платит ему 1 руб.; если B не находит A , то B сам должен уплатить игроку A 1 руб. Таким образом, игрок A может выиграть +1 руб. или —1 руб. (проиграть 1 руб.). Игрок B может при этих же условиях проиграть 1 руб. или —1 руб. (выиграть 1 руб.). Построим платежную матрицу для этой игры. Поскольку у каждого из игроков A и B имеется только по две стратегии, то это будет игра 2×2 , для которой платежная матрица имеет вид, представленный в табл. 48.

Допустим, что сначала игрок A принял стратегию A_1 , т. е. спрятался в первом убежище. Если игрок B принял стратегию B_1 , то он нашел игрока A , и последний проиграл 1 руб. (в табл. 48 — клетка на пересечении A_1 и B_1), т. е. B_1 выиграл у игрока A_1 1 руб. Предположим, что игра продолжается далее. Если при этом игрок A постоянно придержива-

Таблица 48

Стратегии A	Стратегии B	
	B_1	B_2
A_1	—1	+1
A_2	+1	—1

вается одной и той же стратегии, например A_1 , то игрок B , запомнив предыдущие ситуации, очень скоро усвоит это и начнет каждый раз выигрывать у игрока A по 1 руб. Аналогично игрок B будет выигрывать, если игрок A будет все время придерживаться стратегии A_2 или даже чередовать каким-либо закономерным образом стратегии A_1 и A_2 (например, через раз использовать A_1 или A_2). Здесь мы предполагаем, как уже было оговорено выше, что игроки всегда стараются выбрать для себя наиболее выгодные стратегии и при этом выборе используют весь имеющийся в их распоряжении опыт. Поэтому приходится признать тот факт, что, как бы ни старался игрок A скрыть закономерность использования им стратегий A_1 и A_2 , все равно через некоторое время игрок B разгадает эту закономерность и будет постоянно выигрывать.

Очевидно, единственным способом, гарантирующим игрока A от верного проигрыша, будет такой выбор стратегий A_1 и A_2 , когда отсутствует какая-либо закономерность в их чередовании, т. е. этот выбор осуществляется случайно, например путем бросания монеты; выпал герб — применяем стратегию A_1 , выпала цифра — применяем стратегию A_2 . В этом случае игрок B полностью лишается возможности подстраиваться под действия игрока A и ему остается только также организовать случайный выбор стратегий B_1 и B_2 .

Такой подход к игре, когда для выигрыша приходится использовать случайный выбор стратегии, называется использованием смешанных стратегий. Итак, смешанные стратегии получаются путем случайного чередования отдельных ч и с т ы х стратегий. При этом надо еще ввести вероятности использования чистых стратегий при образовании смешанных стратегий. В рассмотренной выше игре «поиск» мы неявно предполагали, что вероятности стратегий A_1 и A_2 одинаково равны 0,5. В более сложных примерах с большим числом чистых стратегий бывает необходимо определить эти вероятности, т. е. установить распределение вероятностей использования чистых стратегий в смешанной стратегии.

Предположим, что игра составлена таким образом, что существует ее решение в чистых стратегиях. Сначала определим наилучшую из стратегий игрока A , т. е. наилучшую из $A_1, A_2, \dots, A_m$ с учетом всех возможных ответов на нее игрока B (табл. 49).

Т а б л и ц а 49

Стратегии A	Стратегии B					
	B_1	B_2	$\dots$	B_j	$\dots$	B_n
A_1	a_{11}	a_{12}	$\dots$	a_{1j}	$\dots$	a_{1n}
A_2	a_{21}	a_{22}	$\dots$	a_{2j}	$\dots$	a_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_i	a_{i1}	a_{i2}	$\dots$	a_{ij}	$\dots$	a_{in}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_m	a_{m1}	a_{m2}	$\dots$	a_{mj}	$\dots$	a_{mn}

При этом мы должны рассчитывать, на то, что на любую стратегию A_i игрок B ответит стратегией B_j , для которой выигрыш игрока A окажется минимальным. Чтобы найти эту стратегию B_j , надо в строке платежной матрицы, соответствующей стратегии A_i (строке с номером i), найти минимальное из чисел a_{ij} . Обозначим его α_i , т. е. $\alpha_i = \min_j a_{ij}$, где минимум определяется путем перебора всех столбцов.

При изменении стратегий игрока A соответствующее каждой из этих стратегий число α_i тоже будет меняться. Естественно, что игроку A выгоднее всего остановиться на такой стратегии A_i , для которой значение α_i будет максимальным. Обозначим это, максимальное значение α , т. е. $\alpha = \max_i \alpha_i$, или, учитывая выражение для α_i , получаем $\alpha = \max_i \min_j a_{ij}$.

Величину α принято называть **нижней ценой** игры или максимальным выигрышем (сокращенно максимином). Стратегию игрока A , которой соответствует максимин α , назовем **максиминной стратегией**.

Если игрок A будет придерживаться максиминной стратегии, то ему при любом поведении игрока B гарантирован выигрыш: во всяком случае не меньше чем α . Поэтому величину α называют **нижней ценой** игры, т. е. это тот гарантированный минимум, который получит игрок A в данной игре.

Аналогично можно определить наилучшую из стратегий игрока B . Он стремится обратить выигрыш игрока A в минимум. Для этого игрок B старается для каждой своей стратегии B_j получить максимальное значение выигрыша при любой стратегии игрока A , т. е. он ищет значение β_j такое, что $\beta_j = \max_i a_{ij}$. Однако игрок B , не может рассчитывать на то, что игрок A позволит ему получить любой из выигрышей β_j . Единственно на что может рассчитывать игрок B , так это на то, что он получит выигрыш, не меньший чем величина β , определяемая выражением

$$\beta = \min_j \max_i a_{ij}.$$

Эта величина β называется **верхней ценой** игры при минимальном выигрыше (минимаксом). Соответствующая минимаксу стратегия игрока B называется **минимаксной стратегией**. Это наиболее осторожная стратегия игрока B , обеспечивающая ему в любом случае проигрыш не больше β и соответственно выигрыш игроку A также не больше β .

В теории игр принцип осторожности, рекомендуемый игрокам придерживаться максиминной стратегии, называется **принципом минимакса**. Он вытекает из предположения об осторожности игроков или иначе из желания разрешить конфликтную ситуацию наилучшим для всех участвующих в ней сторон образом.

Определим нижнюю и верхнюю цены игры, а также минимаксную и максиминную стратегии для разобранных выше двух примеров.

Т а б л и ц а 50

Стратегии A	Стратегии B		
	B_1	B_2	α_i
A_1	-1	+1	-1
A_2	+1	-1	-1
β_j	+1	+1	

Для игры «поиск» запишем платежную матрицу и все значения α_i и β_i в табл. 50.

Так как все величины α (α_1 и α_2) равны одному и тому же числу -1, а величины β (β_1 и β_2) равны +1, то нижняя и верхняя цены игры различны и равны: $\alpha = -1$; $\beta = +1$,

Любая стратегия игрока A является максиминной, любая стратегия игрока B является минимаксной. Из решения игры следует, что, придерживаясь любой из своих стратегий, игрок A проигрывает не более 1 руб., а игрок B также выигрывает не более 1 руб.

В случае, если $\alpha = \beta$, то минимаксные стратегии образуют так называемую седловую точку. Эти стратегии являются устойчивыми в том случае, что отклонение от этих стратегий невыгодно для обоих игроков.

3. *Решение игры в смешанных стратегиях. Основная теорема теории матричных игр.* Смешанные стратегии представляют собой математическую модель изменчивой и гибкой тактики игрока, при которой противостоящий ему игрок не может узнать заранее то положение, с которым ему придется встретиться в игре, ибо перед каждой партией производится случайный выбор одной из чистых стратегий с помощью некоторого механизма, производящего этот выбор с определенными и заранее заданными вероятностями. Покажем, как определить эти вероятности исходя из требования получения максимина, т. е. оптимального решения игры.

Рассмотрим, как и прежде, игру $m \times n$, когда у игрока A имеется m стратегий, а у игрока B — n стратегий. Введем обозначение для смешанных стратегий игроков A и B :

$$S_A = S_A(p_1, p_2, \dots, p_m);$$

$$S_B = S_B(q_1, q_2, \dots, q_n),$$

где $p_1, p_2, \dots, p_m$ — вероятности использования чистых стратегий $A_1, A_2, \dots, A_m$; $q_1, q_2, \dots, q_n$ — вероятности чистых стратегий $B_1, B_2, \dots, B_n$. Поскольку эти стратегии по условию игры полностью исчерпывают возможные ходы игроков A и B , то они образуют полную группу событий. Поэтому для вероятностей p_i справедливо условие

$$\sum_{i=1}^m p_i = 1.$$

Аналогичное условие имеет место для вероятностей q_j

$$\sum_{j=1}^n q_j = 1.$$

Очевидно, каждая чистая стратегия является частным случаем смешанной стратегии, когда все стратегии, кроме выбранной чистой

стратегии, имеют вероятности, равные нулю, а выбранная стратегия — вероятность, равную единице.

Оказывается, если допустить возможность использования смешанных стратегий, то для каждой конечной игры можно найти оптимальное решение, т. е. пару устойчивых оптимальных стратегий игроков.

В теории игр доказывается основная теорема матричных игр: *любая матричная игра имеет решение в смешанных стратегиях.*

Из этой теоремы следует, что каждая матричная игра имеет цену v , причем эта цена всегда лежит между нижней α и верхней β ценой игры:

$$\alpha \leq v \leq \beta.$$

Действительно, α есть минимальный гарантированный выигрыш, который может обеспечить себе игрок. Применяя смешанные стратегии, игрок A , во всяком случае, не уменьшит свой выигрыш, т. е. $v \geq \alpha$. Для игрока B аналогичные рассуждения приведут к выводу, что $v \leq \beta$. Поэтому в общем случае

$$\alpha \leq v \leq \beta.$$

Если в некоторой игре $m \times n$ найдено оптимальное решение, состоящее в использовании двух оптимальных смешанных стратегий:

$$S_A^*(p_1, p_2, \dots, p_m) \text{ и } S_B^*(q_1, q_2, \dots, q_n),$$

то чистые стратегии $A_1, A_2, \dots, A_m$ и $B_1, B_2, \dots, B_n$, для которых вероятности p_i и q_i оказались отличными от 0, называются **активными** стратегиями игроков A и B .

В теории игр доказывается теорема об активных стратегиях: *если один из игроков придерживается своей оптимальной смешанной стратегии, то его выигрыш не меньше цены игры, независимо от того, что делает другой игрок.*

Смысл этой теоремы состоит в том, что если один из игроков начнет последовательно придерживаться своей оптимальной смешанной стратегии, то противостоящий игрок уже не может изменить исход игры.

Рассмотрим пример использования смешанных стратегий для нахождения решения в игре «поиск». В этой игре каждый из игроков имеет по две чистые активные стратегии: A_1, A_2 и B_1, B_2 . Составим из них смешанные стратегии: $S_D = S_D(p_1, p_2)$ и $S_B = S_B(q_1, q_2)$, где p_1, p_2 и q_1, q_2 — вероятности использования чистых стратегий. Платежную матрицу игры можно записать в виде табл. 51.

Согласно теореме об активных стратегиях, можно утверждать, что при использовании игроком A стратегий A_1 и A_2 с вероятностями p_1 и p_2 игрок B может применять любую из своих чистых стратегий B_1 или B_2 , и при этом выигрыш игрока A всегда составит цену

Т а б л и ц а 51

Стратегии A	Стратегии B	
	B_1	B_2
A_1	a_{11}	a_{12}
A_2	a_{21}	a_{22}

игры. Это утверждение позволяет получить два уравнения:

$$a_{11}p_1 + a_{21}p_2 = v;$$

$$a_{12}p_1 + a_{22}p_2 = v.$$

Первое из них получено в предположении, что игрок B придерживается стратегии B_1 , второе — стратегии B_2 . Принимая во внимание условие $p_1 + p_2 = 1$, получим:

$$p_1 = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}};$$

$$p_2 = \frac{a_{11} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}},$$

$$v = \frac{a_{22}a_{11} - a_{12}a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}}.$$

Рассуждая аналогично для игрока B , получаем уравнения для вероятностей, определяющих его оптимальную смешанную стратегию:

$$a_{11}q_1 + a_{12}q_2 = v;$$

$$a_{21}q_1 + a_{22}q_2 = v,$$

откуда

$$q_1 = \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}}; \quad q_2 = 1 - q_1.$$

Для условий игры «поиск» значения выигрышей следующие: $a_{11} = -1$; $a_{12} = 1$; $a_{21} = 1$; $a_{22} = -1$. Подставляя эти значения в полученные формулы, находим: $p_1 = 0,5$; $p_2 = 0,5$; $q_1 = 0,5$; $q_2 = 0,5$. Оптимальными смешанными стратегиями игроков A и B будут: $S_A^*(0,5; 0,5)$ и $S_B^*(0,5; 0,5)$. Следовательно, оптимальные стратегии игроков состоят в случайном чередовании чистых стратегий с вероятностями 0,5, и при этом средний выигрыш будет равен нулю. Итак, точное решение полностью подтверждает тот интуитивный вывод об оптимальных стратегиях, который был получен ранее.

4. *Решение игры методами линейного программирования.* В общем случае при больших m и n решение игры $m \times n$ представляет сложную задачу. Покажем, что решение любой конечной игры может быть сведено к задаче линейного программирования.

Будем считать, что платежная матрица игры a_{ij} (см. табл. 49) нам известна. Найдем решение игры, т. е. две оптимальные смешанные стратегии игроков A и B :

$$S_A^*(p_1, p_2, \dots, p_m) \text{ и } S_B^*(q_1, q_2, \dots, q_n).$$

Найдем сначала оптимальную стратегию S_A^* . Эта стратегия должна обеспечивать игроку A выигрыш не меньше v при любом поведении игрока B и выигрыш, равный v , при оптимальном поведении игрока B .

Предположим, что цена игры $v > 0$. Действительно, для выполнения условия $v > 0$ достаточно, чтобы все элементы платежной матрицы a_{ij} были неотрицательными. Этого можно добиться в любой игре, прибавляя ко всем элементам a_{ij} одну и ту же достаточно большую величину M . При этом цена игры увеличится на M , а само решение (выбор стратегий) не изменится.

Предположим далее, что игрок A применяет свою оптимальную смешанную стратегию S_A^* (пока нам неизвестную), а игрок B — только свою любую чистую стратегию B_j . Тогда средний выигрыш игрока A составит

$$a_j = p_1 a_{1j} + p_2 a_{2j} + \dots + p_m a_{mj}.$$

Но любое из чисел a_j не может быть меньше v , согласно теоремам теории игр. Отсюда получаем в общем случае n неравенств:

$$p_1 a_{11} + p_2 a_{21} + \dots + p_m a_{m1} \geq v;$$

$$p_1 a_{12} + p_2 a_{22} + \dots + p_m a_{m2} \geq v;$$

$$\dots \dots \dots$$

$$p_1 a_{1n} + p_2 a_{2n} + \dots + p_m a_{mn} \geq v.$$

Введем новые обозначения:

$$x_1 = \frac{p_1}{v}; \quad x_2 = \frac{p_2}{v}, \dots, \quad x_m = \frac{p_m}{v}.$$

Тогда после деления всех неравенств на v получим

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{m1}x_m &\geq 1; \\ a_{12}x_1 + a_{22}x_2 + \dots + a_{m2}x_m &\geq 1; \\ \dots \dots \dots \\ a_{1n}x_1 + a_{2n}x_2 + \dots + a_{mn}x_m &\geq 1, \end{aligned} \right\} \quad (\text{XI.1})$$

где, по условию, все x_i — неотрицательные переменные, удовлетворяющие условию

$$x_1 + x_2 + \dots + x_n = \frac{1}{v}.$$

Игрок A стремится сделать свой гарантированный выигрыш v максимальным. Это значит, что величина $1/v$ должна принять минимальное значение. Таким образом, задача решения игры свелась к следующей задаче линейного программирования: определить такие неотрицательные значения переменных $x_1, x_2, \dots, x_m$, чтобы обратилась в минимум линейная функция

$$F_x = x_1 + x_2 + \dots + x_m \quad (\text{XI.2})$$

при условии, что выполняются ограничения (XI.1).

Рассуждая аналогично, т. е. последовательно фиксируя чистые стратегии A_i игрока A , оптимальную стратегию игрока B определим, решая следующую задачу линейного программирования: определить такие неотрицательные значения переменных $\lambda_1, \lambda_2, \dots, \lambda_n$, чтобы обратилась в максимум линейная функция

$$F_y = \lambda_1 + \lambda_2 + \dots + \lambda_n \quad (\text{XI.3})$$

при условии, что выполняются ограничения:

[illegible]

Здесь введены обозначения

$$\lambda_1 = q_1/v; \quad \lambda_2 = q_2/v, \quad \dots, \quad \lambda_n = q_n/v.$$

Вместо требования минимизации функции $F_x = 1/v$ теперь используется требование максимизации функции $F_y = 1/v$, так как игрок B в отличие от игрока A стремится минимизировать v (проиграть как можно меньше).

Итак, решение игры $m \times n$ свелось к решению двойственной пары задач линейного программирования. Согласно теоремам игр, это решение всегда существует.

§ 51. Принятие решений в условиях действия неопределенных факторов стохастической природы. Элементы теории статистических игр

Выше были рассмотрены вопросы, связанные с принятием решений одной из оперирующих сторон в условиях неопределенности, обусловленной не известными заранее способами действий других оперирующих сторон — разумных, активно действующих участников операции. Соответствующие неопределенные факторы были названы «стратегическими» неопределенностями. При выборе решения в этом случае мы исходили из предположения, что противник является разумным и принимает именно те действия, которые нам наименее выгодны.

Однако при решении задач принятия решений приходится встречаться не только с таким видом неопределенности. Часто неопределенность бывает обусловлена нашей недостаточной осведомленностью об условиях, в которых будет проводиться операция, и свойствах объектов, участвующих в ней. Например, могут быть заранее неизвестны: погода в некотором районе, объем перевозок и т. д. [6, 16].

Во всех такого рода случаях неопределенность зависит не от сознательно противодействующего нам противника, а от неизвестной нам объективной действительности, которую в теории принятия решений принято называть *природой*. «Природные» неопределенные факторы следует отнести к неопределенным факторам стохастической природы. Вопросами принятия решений в условиях неопределенности, обусловленной недостаточным знанием природы, занимается специальная прикладная математическая дисциплина, называемая теорией статистических решений.

Природа в теории статистических решений рассматривается как некоторая незаинтересованная инстанция, поведение которой неизвестно, но, во всяком случае, не содержит элемента враждебности и сознательного противодействия достижению нашей цели. Оперирующую сторону в соответствующих ситуациях часто называют статистиком, а сами ситуации — играми статистика с природой (иначе играми с природой, или статистическими играми).

В теории статистических решений разработан ряд принципов оптимальности подходов к принятию решений в условиях природных неопределенностей. Прежде чем обратиться к их рассмотрению, остановимся на вопросе предварительного анализа информации, связанной с решаемой задачей и содержащейся в матрице выигрышей.

Как и в стратегических играх, прежде чем приступить к решению игры с природой, целесообразно постараться упростить ее, уменьшив ее размерность путем отбрасывания дублирующих и явно невыгодных стратегий. Отличие здесь состоит в том, что отбрасывать явно невыгодные стратегии следует только за статистика. За природу этого делать нельзя, так как природа не выбирает свои стратегии (т. е. состояние), тем более так, чтобы как можно больше нам навредить.

После предварительного анализа матрицы во многих случаях бывает целесообразно перейти от матрицы выигрышей к матрице рисков, поскольку матрица выигрышей может вносить некоторые искажения в наши представления об относительной выгодности той или иной стратегии.

Рассмотрим такую ситуацию. Пусть игрок A имеет m различных стратегий $A_1, A_2, \dots, A_m$, относительно «природы» P известно, что она имеет различные состояния $P_1, P_2, \dots, P_n$. Так же, как и ранее, считается известным выигрыш a_{ij} игрока A при каждой паре стратегий игрока и природы, т. е. известна платежная матрица (табл. 52).

Требуется выбрать стратегию игрока A , обеспечивающую наилучший результат. Риском игрока A при использовании стратегий A_i в условиях природы P_j называется разность между выигрышем, который он получил бы, если бы знал P_j , и выигрышем, который он получает в тех же условиях, применяя стратегию A_i .

Обозначим r_{ij} риск игрока при его стратегии A_i в условиях P_j . Если бы A знал заранее состояние природы P_j , то он выбрал бы ту стратегию A_i , которой соответствует максимальный выигрыш в столбце P_j , т. е. он играл бы величину

$$\beta_j = \max_i a_{ij}.$$

Тогда, согласно определению, получаем

$$r_{ij} = \beta_j - a_{ij} \geq 0.$$

Т а б л и ц а 52

Стратегии	Состояние природы			
	P_1	P_2	$\dots$	P_n
A_1	a_{11}	a_{12}	$\dots$	a_{1n}
A_2	a_{21}	a_{22}	$\dots$	a_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_m	a_{m1}	a_{m2}	$\dots$	a_{mn}

Т а б л и ц а 53

Стратегии	Условия			
	Π_1	Π_2	Π_3	Π_4
A_1	2	6	2	9
A_2	3	8	4	3
A_3	4	6	6	2

Т а б л и ц а 54

Стратегия	Условия			
	Π_1	Π_2	Π_3	Π_4
A_1	2	2	4	0
A_2	1	0	2	6
A_3	0	2	0	7

Для каждого элемента платежной матрицы можно определить величину риска, т. е. перейти от платежной матрицы a_{ij} к матрице рисков r_{ij} . Эта матрица дает более наглядную картину неопределенной ситуации, чем исходная матрица a_{ij} .

Рассмотрим пример построения матрицы рисков. Пусть планируется операция в заранее неясных условиях Π . Относительно этих условий имеется набор предположений $\Pi_1, \Pi_2, \Pi_3, \Pi_4$, а игрок A имеет три чистые стратегии A_1, A_2, A_3 . Платежную матрицу запишем в виде табл. 53.

Перейдем к матрице рисков. Для этого находим в каждом столбце максимальные элементы a_{ij} . Для первого столбца этим элементом будет $\beta_1 = a_{31} = 4$, для второго $\beta_2 = a_{22} = 8$, для третьего $\beta_3 = a_{33} = 6$, для четвертого $\beta_4 = a_{14} = 9$. Затем из этих элементов β_j вычитаем все элементы a_{ij} и записываем их на соответствующих местах матрицы рисков (табл. 54).

При рассмотрении этой матрицы становятся яснее некоторые особенности данной игры. Так, в платежной матрице при стратегии A_2 по второй строке мы имеем $a_{21} = a_{24} = 3$. В действительности оказывается, что при состоянии конъюнктуры Π_4 мы можем выиграть самое большее 4 единицы и выбор стратегии A_2 достаточно хорош; если же состояние конъюнктуры оказалось бы Π_4 , то выбор стратегии A_2 будет плох из-за очень большого риска проиграть ($r_{24} = 6$). Поэтому выбор стратегии A_1 с меньшим максимальным риском ($r_{13} = 4$) может оказаться более предпочтительным.

Обратимся теперь к рассмотрению подходов к принятию решений в условиях неопределенности стохастической природы и применяемых при этом критериев оптимальности. Подходы к принятию решений здесь существенно зависят от возможности или невозможности проводить эксперименты с целью снизить степень неопределенности.

§ 52. Принятие решений в условиях стохастической неопределенности в случае, когда проведение экспериментов невозможно.

Статистические игры без эксперимента

Рассмотрим три случая, связанные с различной степенью неопределенности в отношении вероятностей возможных состояний природы: 1) имеется некоторая априорная информация о вероят-

ностях возможных состояний природы, заданных в виде вектора; 2) априорная информация отсутствует, но есть некоторые основания для выдвижения ряда гипотез относительно вероятностей возможных состояний природы; 3) информация о вероятностях возможных состояний природы полностью отсутствует.

1. Первый из перечисленных случаев является наиболее простым. В формально математическом плане он ничем не отличается от стохастической ЗПР. Однако в концептуальном плане эти задачи принципиально различны. Напомним, что стохастическая ЗПР была нами ранее квалифицирована как задача, при решении которой исследователь располагает достоверной информацией о вероятностях возможных состояний природы (эти вероятности полностью известны, неопределенность в их знании отсутствует). В рассматриваемых же случаях имеется лишь некоторая априорная, недостоверная информация о вероятностях возможных состояний природы, неопределенность связана именно с этой недостоверностью информации.

Поскольку исследователь не располагает возможностью провести эксперименты, чтобы получить дополнительную информацию о вероятностях возможных состояний природы, то ему не остается ничего другого, как применить те же приемы, что и при решении стохастической ЗПР, и, в частности, прием оптимизации в среднем.

Положим, что известны все априорные вероятности состояний природы $\Pi_1, \Pi_2, \Pi_3, \dots, \Pi_n$, т. е. известны величины $q_1, q_2, \dots, q_n$. Ясно, что должно выполняться условие $\sum_{j=1}^n q_j = 1$. Примем в качестве критерия эффективности, который мы будем стремиться обратить в максимум, среднее значение (математическое ожидание) выигрыша, которое определим с учетом вероятностей состояний природы q_j . Обозначим это среднее значение для каждой стратегии игрока A a_i . Тогда получим

$$\bar{a} = a_{i1}q_1 + a_{i2}q_2 + \dots + a_{in}q_n = \sum_{j=1}^n a_{ij}q_j.$$

В качестве оптимальной стратегии естественно выбрать ту из стратегий A_i , для которой величина $\bar{a}_i$ обращается в максимум:

$$a_i^* = \max_i \bar{a}_i,$$

где a_i^* — максимальный средний выигрыш игрока A при данных условиях игры.

При таком подходе задача принятия решения в условиях неопределенности превращается фактически в задачу принятия решения в условиях определенности (известны все выигрыши $\bar{a}_i$ для всех стратегий A_i), только принятое решение является оптимальным не в каждом отдельном случае, а в среднем, т. е. при многократном повторении одной и той же игры. Например, пусть материал хра-

нится под открытым небом. В этих условиях погода непосредственно влияет на качество материала. Предположим, что заранее неизвестные метеоусловия могут быть подразделены на состояния $\Pi_1, \Pi_2, \Pi_3, \Pi_4$ с вероятностями соответственно $q_1 = 0,1$; $q_2 = 0,2$; $q_3 = 0,5$; $q_4 = 0,2$ (например, вероятность того, что все дни месяца будут дождливыми, равна 0,1, ясными — 0,5, промежуточные состояния имеют вероятности по 0,2). Допустим, игрок A может выбрать три стратегии поведения: A_1, A_2, A_3 и получить разный доход от каждого решения. Представим эти доходы в виде табл. 55 с условными данными.

Т а б л и ц а 55

Стратегии	Условия				
	Π_1	Π_2	Π_3	Π_4	$\bar{a}_i$
A_1	1	4	5	9	5,2
A_2	3	8	4	3	4,5
A_3	4	6	6	2	5,0
g_j	0,1	0,2	0,5	0,2	

В последней строке этой таблицы записаны вероятности состояний g_j . В последнем столбце приведены вычисленные значения $\bar{a}_i$ средних ожидаемых выигрышей. Из этого столбца видно, что оптимальной стратегией игрока, дающей наибольший средний доход 5,2, является стратегия A_1 .

При выборе оптимальной стратегии в неизвестных условиях с известными вероятностями состояний природы можно пользоваться не только значениями среднего выигрыша $\bar{a}_i$, но и значениями среднего риска:

$$\bar{r}_i = \sum_{j=1}^n r_{ij} q_j,$$

который, естественно, следует стремиться сделать минимальным.

Покажем, что в общем случае стратегия, максимизирующая средний выигрыш $\bar{a}_i$, совпадает со стратегией, минимизирующей средний риск $\bar{r}_i$. Для этого вычислим сумму этих показателей:

$$\bar{a}_i + \bar{r}_i = \sum_{j=1}^n a_{ij} q_j + \sum_{j=1}^n (\beta_j - a_{ij}) q_j = \sum_{j=1}^n \beta_j q_j.$$

Эта сумма есть среднее значение (математическое ожидание) максимумов столбцов платежной матрицы. Для каждой матрицы эта сумма является постоянной величиной, т. е.

$$\sum_{j=1}^n \beta_j q_j = c; \quad \bar{r}_i = c - \bar{a}_i.$$

Очевидно, что $\bar{r}_i$ обращается в минимум, когда величина $\bar{a}_i$ стано-

вится максимальной. Напомним, что, по условию, всегда $r_i \geq 0$. Таким образом, стратегия A_i^* , лучшая относительно среднего выигрыша, оказывается также лучшей с точки зрения уменьшения среднего риска.

Если известны вероятности состояний природы P_j , то при решении игры с такой природой всегда можно обойтись только чистыми стратегиями, не применяя смешанных. В данном случае применение смешанных стратегий всегда менее выгодно, чем применение чистой стратегии A_i^* , соответствующей

$$a_i^* = \max_i \bar{a}_i.$$

Необходимые для решения задачи вероятности состояний могут быть определены из статистических данных, полученных при многократном воспроизведении (наблюдении) состояний природы.

2. Обратимся к рассмотрению второго из перечисленных выше случаев. Часто, приступая к выполнению операции, мы ничего не знаем о вероятностях состояний природы. Тогда не остается ничего другого, как оценить эти вероятности субъективно, выдвинув гипотезу о вероятностных состояниях. Например, если ни одно из состояний природы нельзя предпочесть другим, то, естественно, выдвинем гипотезу о том, что все состояния равновероятны, т. е. $g_1 = g_2 = \dots = g_n = \frac{1}{n}$. Эта гипотеза называется «принцип недостаточного основания Лапласа».

В другом случае мы имеем некоторое представление о том, какие события являются более вероятными, а какие — менее вероятными. Для такого ранжирования событий может быть привлечена группа квалифицированных экспертов. Метод опроса экспертов широко применяется при оценке неопределенной ситуации. Опыт применения такого метода показал, что из оценок экспертов с применением теории игр можно получить обоснованное решение.

3. Перейдем к рассмотрению третьего, наиболее сложного случая.

Можно сформулировать критерии принятия такого решения, не пользуясь понятием вероятностей состояний, основываясь на введенных выше в теории игр понятиях максимина и минимакса.

Максиминный критерий Вальда основан на выборе стратегии игрока A (в игре с природой Π), для которой минимальный выигрыш максимален. Этот критерий предписывает игроку A придерживаться стратегии, гарантирующей при любых условиях выигрыш не меньший, чем максимин соответствующей платежной матрицы игры:

$$\omega = \max_i \min_j a_{ij}.$$

В качестве примера применения критерия Вальда определим оптимальную стратегию игрока A в игре против природы Π с платежной матрицей, представленной в табл. 56 (см. разобранный выше пример).

Стратегии	Условия				
	П ₁	П ₂	П ₃	П ₄	α_i
A_1	1	4	5	9	1
A_2	3	8	4	3	3
A_3	4	6	6	2	2

В последнем столбце этой матрицы представлены значения

$$\alpha_i = \min_j a_{ij},$$

т. е. минимальные значения выигрыша игрока A , которые будут получены при использовании им его чистых стратегий A_i (A_1, A_2, A_3). Согласно критерию Вальда, игроку A следует выбрать стратегию A_2 , как обеспечивающую ему максимальный выигрыш:

$$\omega = \max_i \alpha_i = 3$$

в самых худших условиях игры.

Минимаксный критерий Сэвиджа рекомендует игроку A выбирать ту стратегию A_i , при которой наименьшей является величина риска r_{ij} в самой неблагоприятной ситуации в игре с природой:

$$S = \min_i \max_j r_{ij}.$$

Этот критерий обеспечивает принятие решения без большого риска в любых условиях.

Определим оптимальную с точки зрения критерия Сэвиджа стратегию игрока A в игре с природой Π в условиях разобранного выше примера. Пусть матрица рисков имеет вид табл. 57.

Т а б л и ц а 57

Стратегии	Условия				
	П ₁	П ₂	П ₃	П ₄	r_i
A_1	3	4	1	0	4
A_2	1	0	2	6	6
A_3	0	2	0	7	7

В последнем столбце этой матрицы представлены значения

$$r_i = \max_j r_{ij},$$

т. е. значения максимальных рисков игрока A , которые будут допущены им при выборе стратегий A_1, A_2, A_3 . Согласно критерию

Сэвиджа, игроку A следует выбрать стратегию A_1 , обеспечивающую ему при всех возможных условиях природы Π минимальный риск:

$$S = \min_i r_i = 4.$$

Критерий Вальда и Сэвиджа — это критерии крайнего пессимизма, только пессимизм в каждом из этих критериев понимается по-разному. При использовании критерия Сэвиджа пессимистическое предположение делается относительно величины риска.

К р и т е р и й Г у р в и ц а называют часто также критерием пессимизма — оптимизма. Этот критерий рекомендует при выборе решения руководствоваться не крайним пессимизмом и не крайним оптимизмом. Критерий Гурвица записывается в виде

$$H = \max_i [x \min_j a_{ij} + (1 - x) \max_j a_{ij}],$$

где x — коэффициент, выбираемый произвольно принимающим решение игроком ($0 \leq x \leq 1$).

При $x = 1$ критерий Гурвица превращается в критерий Вальда:

$$H = \omega = \max_i \min_j a_{ij}.$$

При $x = 0$ критерий Гурвица становится критерием крайнего оптимизма, рекомендуящим выбирать ту стратегию, для которой в наилучших условиях выигрыш максимален.

При $0 < x < 1$ получается средний подход между крайним пессимизмом и крайним оптимизмом. В зависимости от характера и оценки игроком ситуации коэффициент x выбирается или более близким к 1, или ближе к 0.

Выбор какого-либо одного критерия для принятия решения в большинстве случаев не является самым лучшим выходом. Поэтому анализ матрицы игры с природой с учетом разных критериев часто дает лучшее представление о ситуации, достоинствах и недостатках каждого решения, чем какой-либо отдельный критерий. Если окажется, что разные критерии приводят к одному и тому же решению, то тем лучше, ибо тогда можно смело принимать рекомендуемое решение.

§ 53. Принятие решений в условиях стохастической неопределенности с использованием экспериментов.

Статистические игры с экспериментами

При решении некоторых ЗПР в условиях природных неопределенностей исследователь располагает возможностью проводить различные эксперименты, позволяющие получить дополнительную информацию и тем самым снизить степень неопределенности в отношении знания действительного состояния природы. Соответствующие ЗПР часто называют задачами принятия статистических решений с проведением экспериментов или статистическими играми с экспериментами.

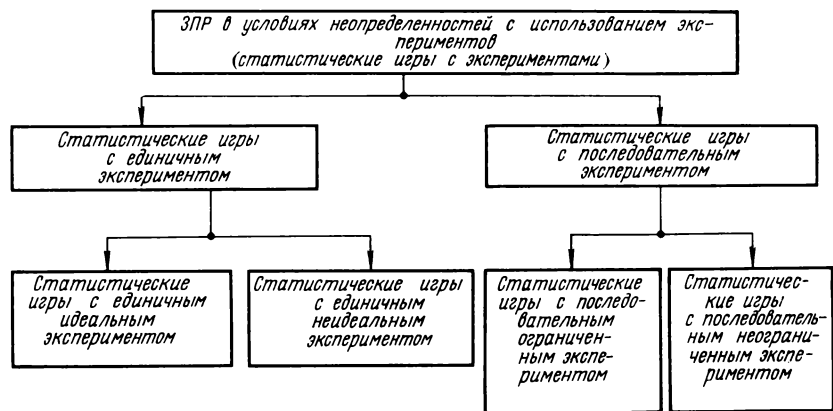


Рис. 70. Классификационная схема ЗПР в условиях стохастической неопределенности с использованием экспериментов

На рис. 70 приведена классификационная схема ЗПР в условиях стохастической неопределенности с использованием экспериментов.

Мы рассмотрим простейшую из задач принятия статистических решений с использованием единичного идеального эксперимента [4, 16].

Статистические игры с единичным идеальным экспериментом

Пусть требуется провести некоторую операцию в недостаточно выясненных условиях ее протекания или с использованием объектов, свойства которых недостаточно изучены. Опирующаяся сторона — статистик располагает возможностью выбора из m стратегий: $x_1, x_2, \dots, x_m$. Относительно неопределенных факторов — состояний природы — можно сделать n предложений $\Pi_1, \Pi_2, \dots, \Pi_n$. Вероятности возможных состояний природы оцениваются как $g_1, g_2, \dots, g_n$, где $\sum_{j=1}^n g_j = 1$, образующих n -мерный вектор $Q = (g_j)$ априорных вероятностей состояний природы. Вектор априорных вероятностей состояний природы известен. Известна также матрица выигрышей $\bar{A} = [a_{ij}]$ размера mn , где a_{ij} — выигрыш оперирующей стороны в случае реализации ею i -й стратегии x_i в условиях j -го состояния природы Π_j .

Для снижения степени неопределенности относительно действительного состояния природы статистик может провести единичный идеальный эксперимент, приводящий к полному выяснению действительного состояния природы. Стоимость проведения эксперимента известна и равна C (в тех же единицах, что и выигрыш a_{ij}).

Требуется определить, целесообразно ли проведение данного эксперимента и какая стратегия является оптимальной в случае проведения эксперимента и при отказе от него.

Чтобы сделать вывод о целесообразности или нецелесообразности проведения эксперимента, сравним наш максимально воз-

можный средний выигрыш без проведения эксперимента со средним выигрышем, который можно получить после проведения эксперимента.

Средний выигрыш a_i^{cp} , соответствующий стратегии x_i , определяется по формуле для математического ожидания:

$$a_i^{cp} = a_i^{cp}(x_i) = q_1 a_{i1} + q_2 a_{i2} + \dots + q_n a_{in} = \sum_{j=1}^n q_j a_{ij}. \quad (XI.5)$$

В качестве оптимального решения без проведения эксперимента исходя из байесовского принципа оптимальности следует выбрать ту стратегию $\bar{x}_t$ из возможных стратегий, для которой достигается максимальный средний выигрыш $\bar{a}^{cp}$, т. е. ту, которая максимизирует выражение (XI.5).

Максимальный средний выигрыш $\bar{a}^{cp}$ и номер t оптимальной стратегии $\bar{x}_t$ определяется из условия

$$\bar{a}^{cp} - a^{cp}(\bar{x}_t) = \max_{1 \leq i \leq m} a_i^{cp} = \max_{1 \leq i \leq m} \sum_{j=1}^n q_j a_{ij}. \quad (XI.6)$$

Максимальный средний выигрыш $\bar{a}^{cp}$ и есть тот выигрыш, на который мы можем ориентироваться при выборе оптимальной стратегии в рассматриваемой операции, если эксперимент не проводится.

Теперь положим, что мы провели эксперимент и выяснили, какое из состояний $\Pi_1, \Pi_2, \dots, \Pi_n$ является действительным состоянием природы. Если им оказалось состояние Π_j , то, очевидно, мы должны применить ту стратегию, для которой в состоянии Π_j выигрыш максимален, т. е. для которой достигается выигрыш

$$\beta_j = \max_{1 \leq i \leq m} a_{ij}. \quad (XI.7)$$

Выигрыш β_j (XI.7) будет получен, если эксперимент будет проведен. Однако мы должны заранее, до проведения эксперимента, решить, целесообразно ли проводить эксперимент. Нам неизвестно, какое из состояний природы имеет место на самом деле и каков будет наш действительный выигрыш после проведения эксперимента.

Чтобы оценить возможный выигрыш после проведения эксперимента, осредним выигрыш β_j с весами, равными априорным вероятностям q_j состояний природы:

$$\beta_{cp} = q_1 \beta_1 + q_2 \beta_2 + \dots + q_n \beta_n = \sum_{j=1}^n q_j \beta_j. \quad (XI.8)$$

С учетом стоимости C эксперимента наш средний выигрыш с применением идеального эксперимента

$$\beta_{cp}^{эк} = \beta_{cp} - C = \sum_{j=1}^n q_j \beta_j - C. \quad (XI.9)$$

Правило, определяющее целесообразность проведения эксперимента, можно сформулировать следующим образом: если выигрыш

$\beta_{\text{ср}}^{\text{эк}}$, определяемый формулой (XI.9), превышает выигрыш $\bar{a}^{\text{ср}}$ без эксперимента, определяемый выражением (XI.6), то эксперимент проводить целесообразно. С учетом формул (XI.6), (XI.7) и (XI.9) это правило можно сформулировать следующим образом: проведение эксперимента целесообразно, если справедливо условие

$$\max_{1 \leq i \leq m} \sum_{j=1}^n q_j a_{ij} < \sum_{j=1}^n q_j \beta_j - C. \quad (\text{XI.10})$$

Учитывая, что

$$-\max_i f_i = \min_i (-f_i),$$

приведем выражение (XI.10) к виду

$$C < -\max_{1 \leq i \leq m} \sum_{j=1}^n q_j a_{ij} + \sum_{j=1}^n q_j \beta_j = \min_{1 \leq i \leq m} \left[\sum_{j=1}^n q_j (\beta_j - a_{ij}) \right]. \quad (\text{XI.11})$$

Но величина $\beta_j - a_{ij}$ есть, по определению, риск r_{ij} , а сумма в выражении (XI.11) под знаком минимума представляет собой средний риск $r_i^{\text{ср}}$, соответствующий стратегии x_i , так как

$$Z_i^{\text{ср}} = \sum_{j=1}^n q_j (\beta_j - a_{ij}) = \sum_{j=1}^n q_j r_{ij}. \quad (\text{XI.12})$$

Минимальный средний риск

$$\bar{r}^{\text{ср}} = \min_{1 \leq i \leq m} r_i^{\text{ср}} = \min_{1 \leq i \leq m} \left[\sum_{j=1}^n q_j (\beta_j - a_{ij}) \right], \quad (\text{XI.13})$$

т. е. равен правой части выражения (XI.11).

Поэтому правило о целесообразности проведения эксперимента формулируется следующим образом: эксперимент проводить целесообразно, если затраты C на его осуществление меньше минимального среднего риска:

$$C < \bar{r}^{\text{ср}}. \quad (\text{XI.14})$$

После проведения эксперимента и выяснения действительного состояния природы в качестве оптимальной стратегии следует выбирать ту стратегию, которая максимизирует выигрыш при найденном действительном состоянии природы. В противном случае, если стоимость эксперимента превышает минимальный средний риск, следует воздержаться от эксперимента и в качестве оптимальной стратегии взять ту стратегию, для которой достигается максимальный средний выигрыш $\bar{a}^{\text{ср}}$ (или минимальный средний риск $\bar{r}^{\text{ср}}$).

Пример. На технологическую линию может поступать сырье разного качества. Из прошлого опыта известно, что в 60 % случаев поступает сырье с малым количеством примесей P_1 , в 40 % случаев сырье с большим количеством примесей P_2 .

На технологической линии предусмотрены три режима работы: x_1 , x_2 , x_3 . Прибыль предприятия от реализации продукции, производимой технологической линией, зависит от качества используемого сырья и режима работы технологической линии. Эта прибыль в расчете на один день работы приведена в виде табл. 58 (в условных единицах).

Т а б л и ц а 58

x_i	Π_j	
	Π_1	Π_2
x_1	5	1
x_2	4	2
x_3	2	3
q_j	0,6	0,4

Т а б л и ц а 59

x_i	Π_j		
	Π_1	Π_2	r_i^{cp}
x_1	0	2	0,8
x_2	1	1	1
x_3	3	0	1,8
q_j	0,6	0,4	

Требуется определить предельную стоимость эксперимента, который целесообразно проводить один раз в день с целью точного определения качества.

Р е ш е н и е. Перейдем к матрице рисков r_{ij} (табл. 59) и определим средний риск r_i^{cp} для каждого режима x_i , считая априорные вероятности q_1 и q_2 появления сырья Π_1 и Π_2 равными 0,6 и 0,4. Из анализа величин r_i^{cp} можно определить, что минимальный средний риск $\bar{r}^{cp} = 0,8$.

В ы в о д. Эксперимент следует проводить, если его стоимость не превосходит 0,8 единицы. В противном случае в качестве оптимального следует использовать режим x_1 , обеспечивающий минимальный средний риск $\bar{r}^{cp} = 0,8$ (или максимальную среднюю прибыль):

$$\bar{a}^{cp} = 5 \times 0,6 + 1 \times 0,4 = 3,4.$$

Глава XII

КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗ В ЭКОНОМИКЕ ТОПОГРАФО-ГЕОДЕЗИЧЕСКОГО ПРОИЗВОДСТВА

§ 54. Сущность корреляционного и регрессионного анализа

Отдельную группу среди экономико-математических задач составляют экономико-статистические задачи. Их цель — определение математических выраженных зависимостей между экономическими показателями, т. е. построение экономико-статистических моделей этих зависимостей с последующим их анализом. Модель позволяет судить о характере и силе связи между отдельными параметрами, дать их статистическую оценку и сделать выводы, представляющие практический интерес.

Экономико-статистические модели только математически формализуют изучаемые зависимости, являясь описательными, и в отличие от оптимизационных называются **д е с к р и п т и в н ы м и** [19, 5].

Решение дескриптивных задач можно разделить на следующие этапы: 1) постановку задачи (качественную формулировку модели); 2) разработку формализованной схемы (выбор функции и аргументов); 3) формирование исходных данных (статистической совокупности); 4) проверку связи между функцией и аргументами; 5) притятие гипотезы о форме связи; 6) построение модели и определение ее параметров; 7) проверку соответствия модели реальной действительности.

Постановка задачи предусматривает четкое определение области применения корреляционного и регрессионного анализа, формулировку цели задачи, выявление факторов, подлежащих изучению.

При выборе функции и аргумента целесообразно, основываясь на данных экономической информации, дать предварительную качественную оценку предполагаемым функции и аргументу, проанализировать относительную связь между ними. Если заведомо известно, что два фактора в высшей степени зависимы друг от друга (т. е. не являются достаточно независимыми), то нет смысла включать их в математическую модель. Менее существенный из них следует исключить из рассмотрения.

Центральное место в решении задачи занимают построение гипотезы о форме связи между функцией и аргументом и определение параметров модели. И на этом этапе весьма эффективными оказываются методы корреляционного и регрессионного анализа. Для определения корреляционной зависимости необходимо располагать

числовыми данными. Такие данные определяют путем статистических наблюдений или исследований. Полученный числовой материал группируют в ряды распределения по факториальному x и результативному y признакам (составляют корреляционную таблицу с двумя входами).

При наличии единственного факториального признака разработка гипотезы о форме связи его с функцией упрощается, поскольку данную зависимость можно показать в виде графика. Если же на результативный признак y воздействует множество факторов $x_1, x_2, x_3, \dots, x_m$, то представить такую зависимость графически не возможно. В этом случае рассматривают ряд гипотез о характере связи между y - и m -факторами и с помощью специальных критериев предполагаемые зависимости проверяют на достоверность.

Линейные модели просты и легко поддаются технико-экономической интерпретации. Поэтому их нередко используют для отображения зависимости между y и x непосредственно либо через другие нелинейные зависимости, которые несложно привести к линейному виду.

Запишем в общем виде уравнения регрессии для различных типов статистических зависимостей в задачах однофакторного анализа:

линейная:

$$y = b + ax;$$

параболическая:

$$y = ax^2 + bx + c;$$

гиперболическая

$$y = b + a/x;$$

степенная:

$$y = ax^\alpha;$$

показательная

$$y = ae^x.$$

При многофакторном анализе в качестве уравнений регрессии выступают следующие зависимости:

прямолинейная:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m;$$

параболическая:

$$y = a_0 + a_1x_1 + a_2x_2^2 + \dots + a_mx_m^m;$$

показательная

$$y = a_0a_1^{x_1}a_2^{x_2} \dots a_m^{x_m};$$

степенная:

$$y = a_0x_1^{a_1}x_2^{a_2} \dots x_m^{a_m}.$$

Линеанизировать нелинейные функции, т. е. свести их с линейным, можно разными приемами: логарифмированием, заменой нели-

Следовательно, всегда получают систему из $(m + 1)$ нормальных уравнений с $m + 1$ неизвестными (m — число факториальных признаков). Такую систему можно решать любым методом, применяемым для решения систем линейных уравнений. Вычисление целесообразно проводить на ЭВМ по стандартным программам.

Далее для установления тесноты (степени) связи между значением y и совокупным влиянием факторов определяют коэффициент множественной корреляции:

$$R = \sqrt{\sum_{j=1}^m \beta_j r_{0j}},$$

где β_j — нормированный коэффициент регрессии; $\beta_j = a_j \sigma_{x_j} / \sigma_y$; σ_y и σ_{x_j} — среднеквадратические отклонения соответственно результирующего и факториального признаков; r_{0j} — коэффициент парной корреляции между y и j -м фактором; $j = 1, 2, \dots, m$ (где m — число факторов):

$$r_{0j} = \frac{\sigma_{jk}}{\sqrt{\sigma_{x_j}} \sqrt{\sigma_y}}.$$

В заключение предполагаемую зависимость подвергают статистическому испытанию на надежность. Для этого по критериям Стьюдента или Фишера проверяют достоверность коэффициентов парной r_{0j} и множественной R корреляции. Например, для парной корреляции по критерию Стьюдента вычисляют величину

$$t_r = \frac{r_{0j} \sqrt{n-2}}{\sqrt{1-r_{0j}^2}}.$$

Зависимость считается близкой к достоверности, если найденная таким путем величина критерия t_r при заданном уровне значимости больше значения, определенного по специальным таблицам.

§ 55. Двухфакторная модель прогнозирования производительности труда

Полевому топографо-геодезическому производству свойственна неравномерность выполнения однотипных работ не только в различных районах, но и в пределах одного объекта в разные отчетные периоды. Поэтому в практике планирования имеется неопределенность в учете производительности труда, соответствующей реальной производственной ситуации, что приводит к неправильному определению годовых объемов работ во времени и фактической потребности во всех видах ресурсов. Еще большая неопределенность возникает при распределении организационно-ликвидационных расходов (10—15 % стоимости всех полевых работ) в течение года, что значительно влияет на качество календарного планирования в целом.

Как показали исследования В. И. Семенова [23], производительность труда P_{ij}^0 при стабилизированных нормах выработки

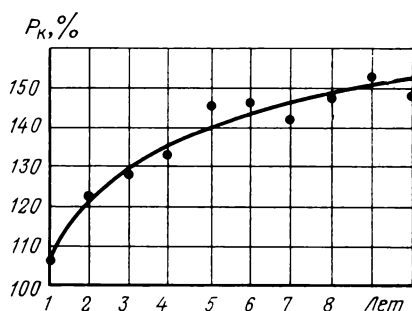


Рис. 71. График зависимости производительности труда от стажа работы исполнителя

зависит от двух основных факторов: уровня квалификации исполнителя или стажа работы k и отклонения продолжительности благоприятной погоды в данном районе или в определенный период j от среднего значения Δy_j^p т. е.

$$P_{ij}^p = P_k + \Delta p_j^p; \quad (XII.3)$$

$$P_k = f(k); \Delta p_j^p = f(\Delta y_j^p); \Delta y_j^p = y_j^p - \bar{y}_j, \quad (XII.4)$$

где y_j^p — продолжительность (часы) благоприятной погоды за период j в районе p в часах; $\bar{y}_j$ — средняя продолжительность благоприятной погоды для всего района работ; Δp_j^p — изменение производительности труда в зависимости от времени j и места p .

Влияние этих факторов на производительность труда изучалось раздельно. В табл. 60 приведены среднестатистические данные о производительности труда исполнителей по предприятию № 2 за 1971—1973 гг. на различных видах и процессах работ. На рис. 71 эти данные представлены в виде дискретных точек, наилучшая аппроксимирующая функция для которых будет

$$P_k = a + b \lg k, \quad (XII.5)$$

Т а б л и ц а 60

Стаж работы k_i	Средняя производительность труда $\bar{P}_i$	$\lg k_i$	$\bar{P}_i \lg k_i$	Прогноз P_k	$(P_k - \bar{P}_i)^2$
1	106	0	0	107	1
2	122	0,30103	36,726	121	1
3	129	0,47712	61,548	129	0
4	133	0,60206	80,074	135	4
5	145	0,69897	101,351	139	36
6	146	0,77815	113,610	143	9
7	144	0,84510	121,694	146	4
8	149	0,90309	134,560	149	0
9	152	0,95424	145,044	151	1
≥10	148	1,0000	148,000	153	25
	1374	6,55976	942,607	—	81

$$a = 107,2; b = 46; m = \sqrt{\frac{(P_k - \bar{P}_i)^2}{k - 1}} = \pm 3 \%. \quad (XII.6)$$

Для нахождения коэффициентов a и b составлены уравнения

$$\left. \begin{aligned} a_1 + b_1 \lg k_1 &= P_1; \\ a_2 + b_2 \lg k_2 &= P_2; \\ &\dots \dots \dots \\ a_i + b_i \lg k_i &= P_i; \\ &\dots \dots \dots \\ a_{10} + b_{10} \lg k_{10} &= P_{10}. \end{aligned} \right\} \quad (\text{XII.6})$$

из которых получены суммарные уравнения

$$a \sum_{i=1}^{10} k_i + b \sum_{i=1}^{10} \lg k_i = \sum_{i=1}^{10} P_i; \quad a \sum_{i=1}^{10} \lg k_i + b \sum_{i=1}^{10} (\lg k_i)^2 = \sum_{i=1}^{10} (P_i \lg k_i). \quad (\text{XII.7})$$

В результате вычислений по методу наименьших квадратов получилось при $k \geq 1$

$$P_k = 107,2 + 46 \lg k \pm 3. \quad (\text{XII.8})$$

Т а б л и ц а 61

Интервалы отклонений состояния погоды от средних значений Δy_i	Отклонения производительности труда от средних значений $(P - \bar{P}_i)$	Весовые коэффициенты n_i	Средние значения Δp_i	Коэффициенты функциональной зависимости c_i	$n_i c_i$	Прогноз Δp_i	$n_i (\Delta p_i - \Delta \bar{p})^2$
0,1	+1; +2; +4; 0; +4; -2	6	0,2	2,0	12,0	1,8	15,4
0,2	+4; -6; +10; +14; 0	5	4,4	22,0	110,0	3,6	3,2
0,3	+17; -7; -1	3	3,0	10,0	30,0	5,4	17,3
0,4	+1; -4; +8; +22	4	6,8	17,0	68,0	7,2	0,6
0,5	+8; +5; +15; +1; +10; +3; +2; +12	8	7,8	15,0	112,5	9,0	18,0
0,6	+8; +11; -1; -1; +8; +23	6	8,0	13,3	106,4	10,8	47,0
0,7	+15; +19; +12	3	15,3	21,9	65,7	12,6	21,9
0,8	+23; +4; +11; +20; +24; 0	6	13,7	17,1	234,3	14,4	2,9
0,9	+18; +12	2	15,0	16,7	33,4	16,2	2,9
1,0	+15; +6; +26	3	15,7	15,7	47,1	18,0	15,9
1,1	+26	1	26,0	23,6	23,6	19,8	38,4
1,2	+12; +32	2	22,0	18,3	36,6	21,6	0,3
		49		18,0	879,6		183,8

$$\Delta \bar{p}_i = \frac{[(P - \bar{P})_i]}{n_i}; \quad c_i = \frac{\Delta p_i}{\Delta y_i}; \quad c = \frac{[n_i c_i]}{[n_i]}; \quad \Delta p_i = c \Delta y_i;$$

$$m_c = \sqrt{\frac{[n_i (\Delta p_i - \Delta \bar{p})^2]}{[n_i]}} = \pm 1,9.$$

На рис. 71 приведены графики месячных изменений фактической производительности труда при производстве мензульной съемки со среднесуточной производительностью при благоприятной погоде (в часах) по данным гидрометеослужбы. Изменение производительности труда во времени Δp_i^p находится в прямой зависимости от состояния погоды Δy_j^p

$$\Delta p_j^p = c \Delta y_j^p, \quad (\text{XII.9})$$

где c — коэффициент зависимости производительности труда от продолжительности благоприятной погоды (табл. 61).

Тогда прогнозирующая функция (XII.3) примет окончательный вид

$$P_{ij}^p = 107,2 + 46 \lg k + 18 \Delta y_j^p \pm \varepsilon, \quad (\text{XII.10})$$

где ε — погрешность прогноза.

Средний стаж k исполнителей, которые будут выполнять работу, известен заранее; Δy_j^p для любого района и временного периода можно получить из многолетних статистических данных. Для определения допускаемой погрешности и стандартных отклонений состояния погоды от среднегодовых значений для различных временных интервалов были обработаны данные об осадках за 22 года. Результаты приведены в табл. 62. Для предприятия № 2 среднее значение $\bar{y} = 21$, а $(\Delta y_j^p) = \pm 2$. Тогда средняя квадратическая погрешность величины Δy_j^p получается: на месяц $\pm 0,88$, на квартал $\pm 0,52$, на полевой сезон $\pm 0,52$, на год $\pm 0,27$. Зная средние квадратические погрешности m_k , m_c , $m_{\Delta y}$ и дифференцируя формулу (XII.70), получаем абсолютные средние квадратические погрешности прогноза производительности труда:

Т а б л и ц а 62

Месяц	Доверительные интервалы единичных отклонений состояния погоды от среднегодовых значений, %			
	месяц	квартал	полугодие	год
Январь	4,9	—	—	—
Февраль	3,4	—	—	—
Март	3,8	2,1	—	—
Апрель	3,6	—	—	—
Май	4,9	—	—	—
Июнь	5,5	2,6	1,4	—
Июль	5,6	—	—	—
Август	5,1	—	—	—
Сентябрь	4,5	3,5	—	—
Октябрь	2,5	—	—	—
Ноябрь	2,9	—	—	—
Декабрь	4,0	1,7	2,3	1,3
Среднее	$\pm 4,2$	$\pm 2,5$	$\pm 1,8$	$\pm 1,3$

Рис. 72. Графики:

1 — среднесуточной продолжительности благоприятной погоды (в часах); 2 — фактической производительности труда при мензуральной съемке

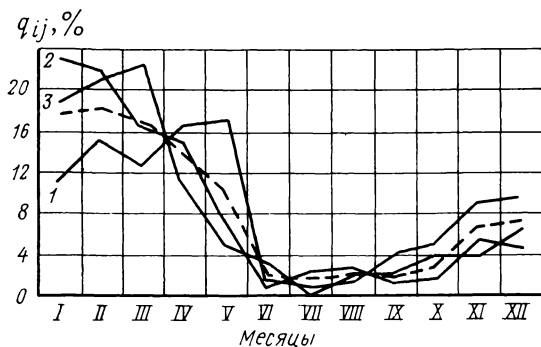
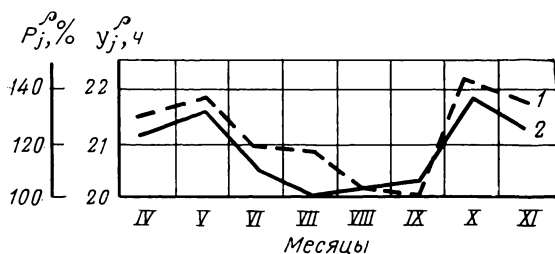


Рис. 73. Графики месячных оргликрасходов для:
1 — триангуляции, 2 — нивелирования, 3 — топографической съемки

$$m_p = \sqrt{m_k^2 + c^2 m_{\Delta y}^2 + \Delta y^2 m_c^2} = \pm \left\{ \begin{array}{l} 16,6\% \text{ — на месяц,} \\ 10,6\% \text{ — на квартал,} \\ 9,6\% \text{ — на полевой сезон,} \\ 6,8\% \text{ — на год.} \end{array} \right\} \quad (\text{XII.11})$$

Отсюда следует, что прогнозировать по предлагаемой методике производительность труда по объектам на год можно с точностью $\pm 6,8\%$, а на квартал с точностью $\pm 10,6\%$, что значительно выше, чем при планировании обычными методами. Это подтверждается данными апробирования прогнозирующей функции (XII.10), приведенными в табл. 63.

Графики зависимости производительности труда от стажа работы исполнителей и продолжительности благоприятной погоды для различных районов деятельности предприятия (см. рис. 71 и 72) используются в настоящее время при планировании. Для составления планов на ЭВМ эти данные представлены в качестве нормативно-справочной информации в виде временных рядов динамики производства по видам работ на все месяцы полевого сезона; в таком же виде представлено и распределение организационно-ликвидационных расходов (оргликрасходов) в течение года (рис. 73).

В общем виде объемно-календарный план производства работ экспедиции в натуральном выражении при условии постоянства

Т а б л и ц а 63

Объект	Местонахождение	Стаж работы (лет)	Вид работы	Производительность труда, %				Отклонения, %			
				полевой сезон	квартал			полевой сезон	квартал		
					II	III	IV		II	III	IV
1	11	6	Триангуляция:								
			план	114	114	114	—	+24	+21	+27	—
			отчет	138	135	141	—	—	—	—	—
			прогноз	136	132	139	—	+2	+3	+2	—
2	12	7	То же	113	113	113	—	+17	+14	+20	—
				130	127	133	—	—	—	—	—
				132	135	133	—	—2	—8	0	—
3	13	8	»	120	120	120	115	+30	+50	+24	+22
				150	170	144	137	—	—	—	—
				147	160	151	131	+13	+10	—7	+6
4	14	2	»	110	110	110	100	+12	+9	+20	+18
				122	119	130	118	—	—	—	—
				118	113	131	109	+3	+6	—1	+9
5	15	6	Нивелирование:								
			план	112	112	112	112	+42	+57	+38	+31
			отчет	154	169	150	143	—	—	—	—
			прогноз	153	155	153	150	+1	+14	—3	—7
6	16	3	То же	124	124	124	—	+2	—3	+6	—
				126	121	130	—	—	—	—	—
				124	124	124	—	+2	—3	+6	—
			Средняя погрешность:								
			план					+21	+25	+22	+24
								+3	+4	0	1

числа исполнителей и обязательного выполнения годовых заданий может быть представлен системой уравнений по каждому объекту:

[illegible]

где

$$v_{ij} = \frac{V_i P_{ij}^0 d_j}{\sum_{j=t_1}^{t_2} P_{ij}^0 d_j}$$

— месячные объемы работ в натуральном выражении; d_j — рабочие дни в текущем месяце j ; t_1 и t_2 — даты начала и конца работ, обу-

словленные длительностью полевого сезона или сроками выполнения работ; n — число видов и процессов работ на объекте.

Сметная стоимость работ на год S_i представлена в объемно-календарном плане работ подразделений двумя видами расходов: на производство A_i и организационно-ликвидационные мероприятия Q_i , т. е.

$$S_i = A_i + Q_i. \quad (\text{XII.13})$$

Тогда месячная сметная стоимость работ подсчитывается по формуле

$$S_{ij} = \frac{A_i v_{ij}}{V_i} + \frac{q_{ij} Q_i}{100}, \quad (\text{XII.14})$$

где q_{ij} — месячные оргликарисходы в процентах от годовой плановой суммы.

Таким образом с помощью разработанной методики удалось не только автоматизировать основные вычислительные процессы, связанные с планированием полевого производства, но и повысить точность и научную обоснованность плановых решений.

§ 56. Взаимосвязь качества труда исполнителей с производительностью их труда

Известно, что показатели производительности и качества труда зависят от возраста и стажа работы исполнителя. Согласно данным исследований В. И. Семенова (см. § 55), зависимость производительности труда P_k от стажа k работы исполнителя в топографо-геодезическом производстве выражается уравнением вида

$$P_k = 107,2 + 46 \lg k.$$

Приведенная формула получена на основе анализа работы исполнителей со стажем до 10 лет. А. А. Дразнюк [8] уточнил эту формулу на основе более обширных статистических данных, взяв в обработку показатели производительности труда исполнителей со стажем работы до 30 лет. Полученные первичные результаты (рис. 74) свидетельствуют, что групповые средние показатели производительности труда исполнителей растут до 20—24 лет стажа, затем медленно снижаются. На этом основании можно предположить, что между стажем работы и производительностью

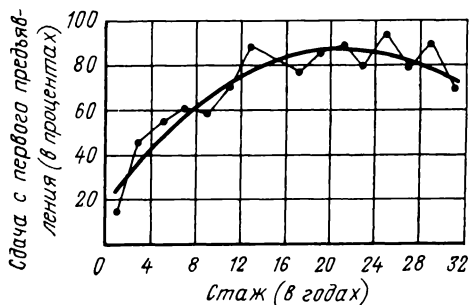


Рис. 74. Зависимость производительности труда от стажа работы в топографо-геодезическом производстве

Расчет коэффициентов нормальных уравнений для установок зависимости производительности труда от стажа работника

№ п. п.	Стаж (годы, x)	Выполне- ние норм (проценты, $\bar{y}$)	x^2	yx	x^3	yx^2	x^4	y_x	$(y - y_x)^2$	$(y - \bar{y})^2$	$(y_x - \bar{y}_x)^2$
1	2	3	4	5	6	7	8	9	10	11	12
1	1	117	1	117	1	117	1	121,7	22,09	1849,0	1482,2
2	3	131	9	393	27	1 179	81	131,9	0,81	841,0	789,6
3	5	150	25	750	125	3 750	645	141,0	81,00	100,0	361,0
4	7	152	49	1 064	343	7 448	2 401	149,1	8,41	64,0	118,8
5	9	148	81	1 332	729	11 988	6 561	156,0	64,00	144,0	16,0
6	11	171	121	1 881	1 331	20 691	14 641	161,9	82,81	121,0	3,6
7	13	171	169	2 223	2 197	28 899	28 561	166,6	19,36	121,0	43,6
8	15	151	225	2 265	3 375	33 975	50 625	170,3	368,64	81,0	106,1
9	17	178	289	3 026	4 913	51 442	83 521	172,8	27,04	324,0	163,8
10	19	176	361	3 344	6 859	63 536	130 321	174,3	2,89	256,0	204,5
11	21	181	441	3 801	9 261	79 821	194 481	174,6	29,16	441,0	213,2
12	23	149	529	3 427	12 167	78 821	279 841	173,9	620,01	121,0	193,2
13	25	185	625	4 625	15 625	115 625	390 625	172,0	169,00	625,0	144,0
14	27	181	729	4 887	19 683	131 949	531 441	169,1	141,61	441,0	82,8
15	29	167	841	4 843	24 389	140 447	707 281	165,0	4,00	49,00	25,0
16	31	152	961	4 712	29 791	146 072	923 521	159,9	62,41	64,0	0,0
Итого	256	2560	5456	42 690	130 816	915 760	3 344 528	2560,1	1703,24	5642,0	3947,4
Среднее		$\bar{y} = 160,0$	—	—	—	—	$\bar{y}_x = 160,0$	—	—	—	—

стью труда существует не полулогарифмическая (XII.8), а параболическая корреляционная зависимость.

Учитывая, что дисперсия первичных статистических результатов значительная, а высокой точности аппроксимации для наших целей не требуется, можно остановиться на данной стадии исследования на параболе второго порядка. Вычислим по методу наименьших квадратов коэффициенты уравнения регрессии

$$y_x = a_0 + a_1x + a_2x^2, \quad (\text{XII.15})$$

где x , y_x — факторный и результативный признаки соответственно стаж в годах и производительность труда в процентах.

Для вычисления коэффициентов уравнения (XII.15) имеем систему нормальных уравнений:

$$\left. \begin{aligned} a_0n + a_1 \sum x + a_2 \sum x^2 &= \sum y_x; \\ a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 &= \sum y_x x; \\ a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 &= \sum y_x x^2. \end{aligned} \right\} \quad (\text{XII.16})$$

Коэффициенты нормальных уравнений приведены в табл. 64. Система нормальных уравнений примет вид

$$\left. \begin{aligned} 16a_0 + 256a_1 + 5456a_2 &= 2560; \\ 256a_0 + 5450a_1 + 130\,816a_2 &= 42\,690; \\ 5456a_0 + 130\,816a_1 + 3\,344\,528a_2 &= 915\,716. \end{aligned} \right\} \quad (\text{XII.17})$$

Решение системы (XII.17) дает следующее уравнение регрессии:

$$y_x = 116,14 + 5,670x - 0,1374x^2. \quad (\text{XII.18})$$

Среднеквадратическое отклонение m_y фактических данных y от найденных y_x подсчитывается по формуле

$$m_y = \pm \sqrt{\frac{1}{n-l} \sum (y - y_x)^2},$$

где n — число группировок, взятых в обработку; l — число связей, накладываемых функцией y_x на выборку; для параболы второго порядка ($l = 3$). Тогда в рассмотренном примере $m_y = \pm 11,4 \%$.

Оценим силу найденной связи с помощью теоретического корреляционного отношения η_r (индекс корреляции):

$$\eta_r = \sqrt{\frac{\sigma_{y_x}^2}{\sigma_y^2}}, \quad (\text{XII.19})$$

где $\sigma_{y_x}^2$ — вариация производительности труда под влиянием фактора—стажа работы исполнителя

$$\sigma_{y_x}^2 = \frac{\sum (y_x - \bar{y}_x)^2}{n}; \quad (\text{XII.20})$$

σ_y^2 — вариации производительности труда под влиянием всех факторов

$$\sigma_y^2 = \frac{\sum (y - \bar{y})^2}{n}, \quad (\text{XII.21})$$

где y , y_x — соответственно фактическое и теоретическое значения результативного признака. Подставив выражения (XII.20) и XII.21) в формулу (XII.19), найдем:

$$\eta_T = \sqrt{\frac{\sum (y_x - \bar{y}_x)^2}{(y - \bar{y})^2}}. \quad (\text{XII.22})$$

Полученное в рассмотренном примере теоретическое корреляционное отношение $\eta_T = 0,836$ говорит о высокой связи между производительностью труда и стажем работы исполнителя.

Коэффициент детерминации D рассчитаем по формуле

$$D = \frac{\sigma_{y_x}^2}{\sigma_y^2}. \quad (\text{XII.23})$$

В нашем примере коэффициент детерминации ($D = 0,699$) свидетельствует, что изменение производительности труда на предприятии на 69,9 % обусловлено фактором роста стажа исполнителя и на 30,1 % — другими факторами.

Проанализируем таким же методом связь между качеством работы исполнителя и его стажем. Показанные на рис. 75 результаты первичного статистического наблюдения позволяют предположить, как и в предыдущем случае, наличие параболической корреляции зависимости.

В результате составления и решения системы нормальных уравнений по приведенной выше методике получено уравнение связи между качеством труда y_x и стажем работы x :

$$y_x = 20,00 + 6,403x - 0,14393x^2. \quad (\text{XII.24})$$

Как и ранее, получим следующие оценки силы найденной связи: $m_y = \pm 6,5 \%$; $\eta_T = 0,95$; $D = 0,89$, которые подтверждают наличие более тесной, чем в предыдущем примере, связи между качеством труда и стажем работы.

Таким же образом получим и уравнение связи между качеством труда y_x и производительностью труда исполнителя x . Из рис. 76 видно, что и здесь характер зависимости можно выразить параболой второго порядка.

Уравнение регрессии имеет вид

$$y_x = 87,10 + 1,53266x - 0,00353x^2 \quad (\text{XII.25})$$

и оценки силы связи будут $m_y = \pm 4,4 \%$; $\eta_T = 0,98$; $D = 0,95$.

Приведенные выше результаты свидетельствуют о том, что характеры зависимостей производительности и качества труда от стажа

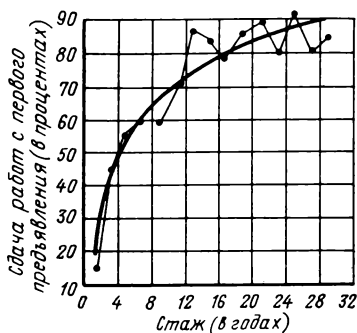


Рис. 75. Зависимость качества труда от стажа работы в топографо-геодезическом производстве

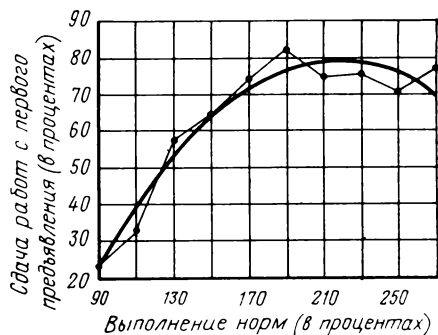


Рис. 76. Зависимость между ростом производительности труда исполнителей и качеством их труда

работы, а также качества труда от производительности совпадают. Как показатели производительности, так и качество труда у среднего исполнителя растут до 21 года трудового стажа в топографо-геодезическом производстве. После этого заметно незначительное снижение показателей производительности и качества труда в последующие 10 лет работы на 15 и 14 % соответственно. Также практически не снижаются показатели качества труда у исполнителей с высокими показателями производительности труда. Заметное (на 10 %) снижение показателей качества наблюдается лишь после того, когда исполнитель перевыполняет нормы выработки более чем в два раза.

§ 57. Минимум затрат как один из критериев определения оптимальных размеров топографо-геодезического подразделения

Р. П. Каширникова в своей работе [15] показала, что минимум затрат может быть использован как один из возможных критериев при определении оптимальных размеров подразделения.

Из табл. 65 видно, что общие затраты по себестоимости складываются из многих элементов, на которые влияют различные факторы. Так как большая доля затрат приходится на заработную плату, можно ожидать, что различный уровень себестоимости связан прежде всего с неодинаковой численностью работников на предприятии. Кроме того, затраты на 1 руб. сметной стоимости зависят от величины объемов производства. Следовательно, решение задачи сводится к отысканию корреляционной зависимости себестоимости (фактические затраты на 1 руб. сметной стоимости) от численности работников и объемов работ (сметная стоимость, тыс. руб.; и приведенные единицы топографической съемки масштаба 1 : 10 000, км²). Выбор аппроксимирующей функции строится на предпосылке, что до известного предела увеличение размеров предприятия (по вы-

Распределение фактической себестоимости топографо-геодезического производства ГУГК по элементам расхода

Годы	Расходы										Накладные	
	Основные											
	основная заработная плата производственного персонала	дополнительная заработная плата	начисление на заработную плату	полевое довольствие	материалы	производственный транспорт	амортизация	износ	прочие	содержание линейного аппарата в полевом и камеральном производстве	содержание центрального аппарата	
1973	36	3	2	4	5	29	3	2	1	12	3	
1974	34	3	2	4	5	24	4	1	7	12	4	
1975	34	3	2	4	5	30	4	1	1	12	4	
1976	33	3	2	4	5	30	5	1	1	12	4	
1977	36	3	2	4	5	24	5	1	1	13	6	

пуску продукции) способствует снижению себестоимости (в результате больших возможностей механизации, автоматизации, организации минимальных технологических циклов, снижения доли постоянной части накладных расходов и т. п.), но за ним следует ее увеличение.

Точность аппроксимации оценивалась на основе анализа дисперсий. Линейная функция исследовалась на интервале убывания или возрастания.

$$y = a + bx; \quad (\text{XII.26})$$

$$y = a + bx + cx^2; \quad (\text{XII.27})$$

$$y = ae^{b \cdot x}, \quad (\text{XII.28})$$

где y — зависимая переменная; x — независимая переменная; a , b , c — параметры уравнений. Для

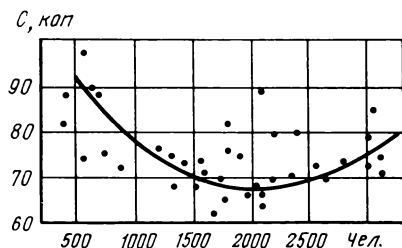


Рис. 77. Зависимость удельных затрат по численности работников топографо-геодезического производства

аппроксимации использовались линейная, параболическая и степенная функции (XII.26)—(XII.28) в интервалах, ограниченных размерами реально действующих предприятий. В табл. 66 приведены результаты расчетов на ЭВМ.

Анализ дисперсий свидетельствует о параболических формах связей для всех исследуемых зависимостей. Графическое изображение корреляционной зависимости между фактически-

Параметры уравнений по видам зависимостей и величины дисперсий σ^2

Вид зависимости, уравнение аппроксимации	Параметры			Дисперсия
	<i>a</i>	<i>b</i>	<i>c</i>	σ^2
Зависимость удельных затрат от численности работающих				
(XII.26)	0,819	$-0,392 \cdot 10^{-5}$	—	0,0767
(XII.27)	1,116	$-0,052 \cdot 10^{-2}$	$0,014 \cdot 10^{-5}$	0,0735
(XII.28)	0,796	15,732	—	0,0745
Зависимость удельных затрат от объемов работ в приведенной топо съемке				
(XII.26)	0,881	$-0,283 \cdot 10^{-5}$	—	0,0642
(XII.27)	1,094	$-0,243 \cdot 10^{-4}$	$0,003 \cdot 10^{-7}$	0,0620
(XII.28)	0,804	46,972	—	0,0742
Зависимость удельных затрат от объемов работ в сметной стоимости				
(XII.26)	0,818	$-0,102 \cdot 10^{-5}$	—	0,0750
(XII.27)	0,924	$-0,499 \cdot 10^{-4}$	$0,026 \cdot 10^{-7}$	0,0682
(XII.28)	0,793	70,460	—	0,0731

ми затратами на 1 руб. сметной стоимости и численностью, а также объемами работ в приведенной топографической съемке и сметной стоимости представлено на рис. 77, 78, 79.

Таким образом, выполненные исследования, расчеты и графический анализ полученных зависимостей позволяют сделать вывод, что по эффективности затрат оптимальными являются предприятия численностью около 2000 чел., выполняющие работы в приведенной топо съемке масштаба 1 : 10 000 порядка 35—40 тыс. км², в сметной стоимости на 6—8 млн. руб. Аналогичные исследования могут быть выполнены и для других показателей эффективности топографо-

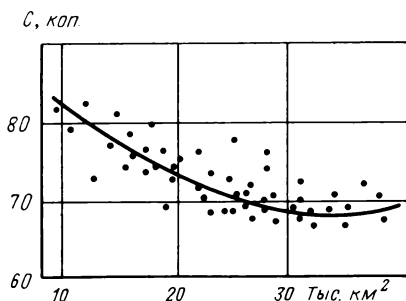


Рис. 78. Зависимость удельных затрат от объемов работ в приведенной топографической съемке масштаба 1 : 10 000

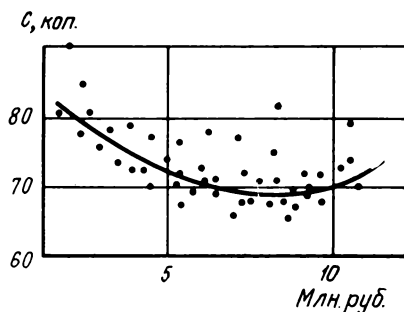


Рис. 79. Зависимость удельных затрат от объемов работ в сметной стоимости

геодезического производства — производительности труда, фондоотдачи и др.

Модель определения оптимальных размеров предприятий на основе критерия минимума затрат может быть использована и для других подразделений топографо-геодезического производства, например для определения оптимальных размеров экспедиций. Особый эффект дает ее применение в сочетании с методом группировок, позволяющий по различным признакам (район работ, мощность и т. д.) найти в образованных группировках оптимальные подразделения.

§ 58. Метод прогнозирования ожидаемого уровня производительности труда по многофакторной регрессионной модели

На основе статистических данных о производственно-хозяйственной деятельности отрасли за период 1961—1975 гг. (табл. 67) была изучена динамика производительности труда и отобраны основные факторы, обуславливающие изменение этого показателя [25].

Используя аппарат математической статистики и программы пошаговой множественной регрессии, на ЭВМ ЕС 1030 рассчитали параметры регрессионной модели для прогнозных оценок ожидаемого роста производительности труда.

На основе регрессионной модели были определены значения ожидаемого роста производительности труда на десятилетку

Таблица 67

Статистические данные о динамике производительности труда и основных факторов ее роста за период 1961—1975 гг. (% к 1961 г.)

Год	Показатели			
	производительность труда	фондовооруженность	уровень специализации	объем работ
1961	100	100	100	100
1962	109	108	104	109
1963	119	130	104	123
1964	132	135	99	139
1965	139	212	97	153
1966	148	252	102	157
1967	156	258	120	164
1968	165	265	147	171
1969	173	288	151	179
1970	188	305	155	195
1971	192	325	171	204
1972	197	350	199	215
1973	202	360	226	228
1974	206	400	226	236
1975	211	392	238	256

(1976—1980 гг.), которые были сопоставлены с аналогичными значениями, полученными по фактическим отчетным данным за период 1976—1978 гг.

В общем случае производительность труда может быть представлена зависимостью

$$P = f(x_1, x_2, \dots, x_n), \quad (\text{XII.29})$$

где $x_1, x_2, \dots, x_n$ — факторы-аргументы.

Для отбора факторов в уравнение регрессии необходимо наличие у них тесной связи с производительностью труда. При отборе руководствовались следующими требованиями:

факторы должны отражать особенности производства и оказывать существенное влияние на изменение уровня производительности труда;

при построении уравнений зависимости функции от факторов целесообразно в качестве переменных использовать показатели, имеющиеся в статистической отчетности;

все факторы должны быть количественно измеримы. Если фактор характеризуется качественным показателем, учет его влияния и включение в уравнение экономико-статистической зависимости требует обязательной количественной определенности. Это может быть сделано, например, с помощью балльной оценки;

в уравнение не включаются факторы, которые находятся в функциональной или тесной корреляционной зависимости друг от друга. Если они связаны тесно, то оценка влияния их на функцию теряет смысл. Кроме того, часто эти факторы характеризуют одну и ту же сторону изучаемого явления, т. е. дублируют друг друга;

при использовании регрессионной зависимости для прогнозирования рекомендуется включать в нее минимальное число факторов, наиболее существенных и характерных для основной тенденции развития моделируемого показателя. Дополнительное включение в модель новых факторов позволяет дать более полное и точное описание изучаемого явления. В то же время это ведет к усложнению зависимости и увеличению времени вычислений.

С учетом указанных требований на первом этапе в обработку были включены статистические данные в процентах к 1961 г. по следующим факторам-аргументам: x_1 — рост фондовооруженности на одного человека (по активной части основных фондов); x_2 — рост уровня специализации (удельный вес съемочных работ в общем объеме топографо-геодезических работ); x_3 — рост объемов топографо-геодезических работ (в трудовых затратах); x_4 — рост удельного веса ИТР (в общей численности промышленно-производственного персонала); x_5 — рост удельного веса камеральных процессов работ.

В результате ранжирования значимости факторов для регрессионной модели были выбраны следующие основные факторы, оказывающие существенное влияние на изменение уровня производительности труда: фондовооруженность, уровень специализации и объем топографо-геодезических работ.

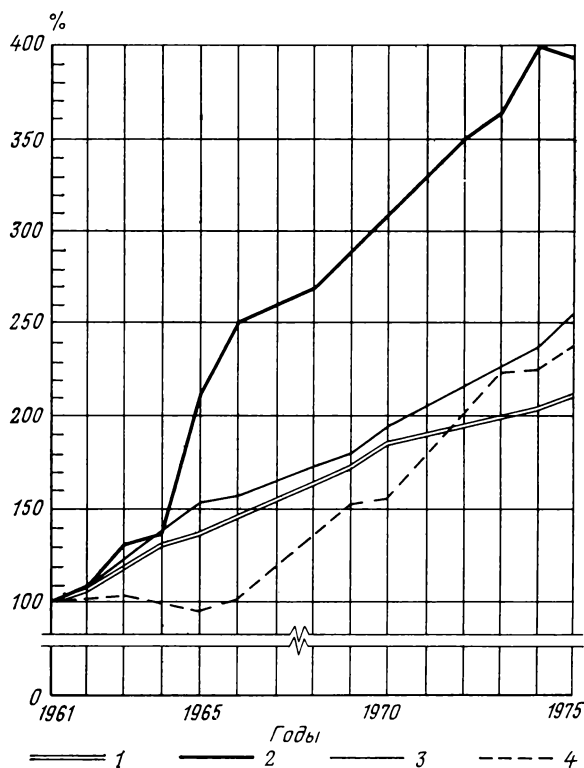


Рис. 80. Темпы роста производительности труда, фондовооруженности, уровня специализации и объема работ за 1961—1975 гг.

Парные коэффициенты корреляции $r_{p, x}$ между производительностью труда и выбранными x_1, x_2, x_3 по результатам расчета на ЭВМ получили следующие значения:

$$\left. \begin{aligned} r_{p, x_1} &= 0,990; \\ r_{p, x_2} &= 0,991; \\ r_{p, x_3} &= 0,985. \end{aligned} \right\} \quad (\text{XII.30})$$

Все значения парных коэффициентов корреляции выше табличных для 1%-ного уровня значимости. Это говорит о достаточно тесной связи между производительностью труда и выбранными факторами.

Распределение во времени значений производительности труда P и факторов-аргументов x_1, x_2, x_3 представлено на рис. 80. Мы видим, что производительность труда 1, фондовооруженность 2 и объем работ в трудовых затратах 3 растут достаточно равномерно.

Фактические их значения хорошо аппроксимируются уравнением прямой. Это особенно важно, поскольку параметры модели линейной множественной регрессии рассчитываются значительно проще, чем нелинейной. Уровень специализации 4 также изменяется во времени без каких-либо заметных скачков. Но, аппроксимируя этот показатель уравнением прямой, можно допустить ошибку, поскольку из графика видно, что закон изменения уровня специализации производства далек от линейного. Прибор аппроксимирующей функции по результатам статистической оценки временного ряда на ЭВМ показал, что для этого фактора больше всего подходит уравнение параболы.

Нахождение основных тенденций (трендов) или, другими словами, выравнивание рассматриваемых динамических рядов осуществлялось методом наименьших квадратов. Критерием подбора аппроксимирующей функции служило среднее квадратическое отклонение фактических значений показателя от рассчитанных по проверяемому уравнению. Окончательная же проверка гипотезы о правильности выбора вида тренда осуществлялась с помощью критерия серий. В результате расчетов на ЭВМ (табл. 68) получили уравнения трендов

$$\left. \begin{aligned} P &= 97,0 + 8,18t; \\ x_1 &= 79,2 + 22,43t; \\ x_2 &= 98,7 - 2,14t + 0,818t^2; \\ x_3 &= 91,6 + 10,45t. \end{aligned} \right\} \quad (\text{XII.31})$$

Статистические оценки динамических рядов по каждому фактору-аргументу могут быть полезны в том случае, если значения каждого фактора в планируемом периоде неизвестны из прямых расчетов и их придется определять по уравнениям регрессии (XII.31).

Т а б л и ц а 68

Результаты оценки аппроксимирующей функции
 $P = 97,0 + 8,18 t \ (\sigma^2 = 4,62)$

№ п. п.	t	P _{факт.}	P _{аппрокс}	Отклонение
1	1,0000	100,0000	105,21669	—5,21669
2	2,0000	109,0000	113,39526	—4,39526
3	3,0000	119,0000	121,57382	—2,57382
4	4,0000	132,0000	129,75240	2,24760
5	5,0000	139,0000	137,93097	1,06903
6	6,0000	148,0000	146,10953	1,89047
7	7,0000	156,0000	154,28810	1,41190
8	8,0000	165,0000	162,46667	2,53333
9	9,0000	173,0000	170,64523	2,35477
10	10,0000	188,0000	178,82381	9,17619
11	11,0000	192,0000	187,00237	9,99763
12	12,0000	197,0000	195,18094	1,81906
13	13,0000	202,0000	203,3591	—1,35951
14	14,0000	206,0000	211,53807	—5,53807
15	15,0000	211,0000	219,71664	—8,71664

Уравнение множественной регрессии, показывающее зависимость производительности труда от факторов-аргументов x_1, x_2, x_3 , рассчитано на ЭВМ ЕС 1030 по программам пошаговой множественной регрессии (табл. 69, 70) и имеет следующий вид:

$$P = 43,9 + 0,467x_3 + 0,157x_1 - 0,026x_2. \quad (\text{XII.32})$$

Т а б л и ц а 69

Корреляционная матрица

	x_1	x_2	x_3	y
x_1	1,00000	0,93604	0,97684	0,99116
x_2	0,93604	1,00000	0,89515	0,95222
x_3	0,97684	0,89515	1,00000	0,97588
y	0,99116	0,95222	0,97588	1,00000

Т а б л и ц а 70

Статистические характеристики уравнений регрессии

Характеристика	I шаг	II шаг	III шаг
Множественный коэффициент корреляции R	0,985	0,990	0,990
Множественный коэффициент корреляции, исправленный для степеней свободы, R_c	0,985	0,989	0,988
Дисперсионное отношение (критерий Фишера) $F_{\text{рас}}$	432,7	295,9	182,3
Остаточная дисперсия $\sigma_{\text{ост}}^2$	0,971	0,980	0,980
a_0	26,892	45,34	43,887
a_1	0,773	0,428	0,467
a_2		0,163	0,157
a_3			-0,026

Коэффициенты регрессии уравнения (XII.32) показывают степень влияния каждого фактора на анализируемый показатель при среднем уровне других факторов. С изменением каждого фактора на единицу (при среднем значении остальных факторов) зависимая переменная P изменяется на соответствующий коэффициент регрессии. В уравнении (XII.32) факторы расположены в порядке значимости их влияния на показатель.

Значимость полученного уравнения множественной регрессии определялась при помощи F -критерия:

$$F = \frac{\sigma^2}{\sigma_{\text{ост}}^2} = 182,3, \quad (\text{XII.33})$$

где σ^2 и $\sigma_{\text{ост}}^2$ — общая и остаточная дисперсии.

Поскольку рассчитанное значение F получилось значительно больше табличного ($F_{\text{табл}} = 6,54$), то можно сделать вывод, что данное уравнение значимо для 1 %-ного уровня значимости.

Т а б л и ц а 71

Значения факторов-аргументов			
Год	Индекс факторов-аргументов		
	x_1	x_2	x_3
1976	437	272	258
1977	460	294	268
1978	482	317	278
1979	505	341	289
1980	527	366	300

Множественное корреляционное отношение η , показывающее тесноту связи между производительностью труда и всеми факторами, оказалось равным 0,990. Значимость множественного корреляционного отношения проверялась при помощи t_η — критерия Стьюдента:

$$t_\eta = \frac{\eta}{S_\eta} = 150,5, \quad (\text{XII.34})$$

где η — множественное корреляционное отношение, S_η — погрешность множественного корреляционного отношения.

Рассчитанные значения t_η -критерия значительно больше табличного ($t_{\eta\text{табл.}} = 3,50$) для 1 %-ного уровня значимости.

Используя регрессионные модели (XII.31), (XII.32), определим ожидаемый уровень производительности труда на десятую пятилетку. Значения факторов-аргументов x_1, x_2, x_3 по годам пятилетки получим, подставляя в уравнение регрессии (XII.31) соответственно $t_{1976} = 16, t_{1977} = 17, \dots, t_{1980} = 20$. Результаты расчета приведены в табл. 70.

Используя регрессионную модель (XII.32) и данные табл. 71, подсчитаем ожидаемый уровень производительности труда по годам пятилетки в процентах к 1975 г. Результаты расчетов приведены в табл. 72, для сравнения в ней даны значения фактически достигнутого уровня производительности труда по отчетным данным за период 1976—1978 гг.

Сопоставляя результаты расчетов ожидаемого уровня производительности труда по модели (XII.32) и фактически достигнутые в 1976—1978 гг., можно сделать вывод, что расхождения в абсолютной величине ожидаемого уровня производительности труда в топографо-геодезическом производстве, расчет и анализ статистических характеристик, а также данные табл. 72 дают основание утверждать, что предложенное уравнение множественной регрессии (XII.32) позволяет осуществить прогнозирование уровня производительности труда.

С течением времени может оказаться, что соотношение темпов роста факторов-аргументов существенно изменится. Таким образом,

Т а б л и ц а 72

Ожидаемый уровень производительности труда

Год	По регрессионной модели		Фактически
	в % к 1961 г.	в % к 1975 г.	в % к 1975 г.
1975	214	100	100
1976	224	104	103
1977	232	108	107
1978	240	112	112
1979	248	116	—
1980	257	120	—

параметры регрессионной модели (XII.32) по мере накопления статистических данных необходимо уточнять.

Достоинство метода прогнозирования ожидаемого уровня производительности труда по многофакторной регрессионной модели по сравнению с методом поэлементного расчета [25] заключается не только в значительном упрощении технико-экономических расчетов, но и в том, что можно быстро обчислить различные варианты направлений капитальных вложений в развитие отрасли и уже в начальной стадии разработки перспективных планов выбрать наиболее эффективные из них.

§ 59. Методы учета влияния отраслевых факторов на показатели эффективности производства

Рассмотрим степень влияния на показатели эффективности топографо-геодезического производства объективных особенностей условий и факторов производства, которые не зависят от самих предприятий. Эти особенности и факторы условно можно разделить на две группы [15]:

I группа — объективно неуправляемые, к которым можно отнести местоположение предприятия, влияющее на длительность полевого периода, и погодные условия; стоимость материалов и транспортных средств; размер надбавок к заработной плате и др.;

II группа — условно неуправляемые: производственная и организационная структура предприятия; техническая вооруженность; организация производства. К этой же группе можно отнести и факторы, которые, по действующим положениям, нельзя изменять (например, объем работ в сметной стоимости или удельный вес отдельных видов работ).

Определение степени влияния обеих групп факторов может быть выполнено путем нахождения корреляционной зависимости и уравнения связи, характеризующего степень влияния этих групп на рассматриваемый показатель эффективности. В качестве примера показателей эффективности топографо-геодезического производства рассмотрим выработку в сметной стоимости y_1 на одного работника и себестоимость работ y_2 . Эти показатели взяты для исследования потому, что отклонения их значений для разных предприятий довольно велики. Для целей сравнительного анализа могут быть взяты любые технико-экономические показатели (например, выработка на одного работающего в приведенных условных единицах топосъемки, прибыль, фондоотдача и др.).

При расчете коэффициентов уравнений регрессии, учитывающих влияние на указанные показатели объективных условий и факторов топографо-геодезического производства, можно использовать количественные оценки по следующим факторам.

I группа: x_1 — длительность полевого периода, месяцы; x_2 — среднегодовая заработная плата на одного работника, руб.; x_3 — среднегодовая выработка в трудовых затратах, человеко-мес.

II группа: x_4 — соотношение полевого и камерального производства по трудовым затратам, %; x_5 — техническая вооруженность труда (величина активной части производственных основных фондов на одного работника), руб.; x_6 — удельный вес транспортных расходов в общем объеме расходов, %.

В качестве неуправляемых факторов учитывались также объемные показатели, регламентируемые централизованно: x_7 — объем работ в сметной стоимости, тыс. руб.; x_8 — удельный вес съёмочных работ в общем объеме производства, %.

Правильный подбор уравнения регрессии и его решения относится к наиболее ответственному этапу корреляционного анализа. Предварительный графический анализ корреляционных полей функции $y_1 = \varphi_1(x_1, x_2, \dots, x_n)$ для выработки в сметной стоимости на одного работника и $y_2 = \varphi_2(x_1, \dots, N)$ для себестоимости работ показал, что заданные функции линейные и могут быть аппроксимированы полиномом первой степени, имеющим вид

$$y = a_0 + \sum_{j=1}^N a_j x_j, \quad (\text{XII.35})$$

где a_0, a_j — коэффициенты регрессии; x_i — рассматриваемые факторы, $j = 1, 2, \dots, 8$ — число факторов.

При построении моделей зависимостей исследуемых показателей от принятых факторов применялась пошаговая множественная регрессия. Расчеты по программе пошаговой множественной регрессии выполнены в ЦНИИГАиКе по данным за 1975—1976 гг. На основе анализа матрицы коэффициентов парной корреляции из дальнейшей обработки были исключены факторы x_2 и x_7 [15].

С использованием исходной информации о количественных значениях факторов $x_1, x_3, x_4, x_5, x_6, x_8$ и показателей y_1, y_2 по каждому предприятию выполнены расчеты для различных комбинаций факторов при определении окончательных уравнений регрессии. Как указывалось, выбор последовательности наиболее значимых факторов на основании анализа квадратов отклонений осуществлялся программно, т. е. в процессе обработки на каждом шагу расчетов автоматически вводилась новая независимая переменная x_j , наиболее влияющая на зависимую переменную y . Уравнения регрессии для анализируемых функций на последнем шаге имеют вид

$$\left. \begin{aligned} y_1 &= -268,87 + 60,15x_6 + 15,25x_5 + 24,76x_3 - 16,50x_1 - \\ &\quad - 14,39x_8 - 9,26x_4; \\ y_2 &= 73,65 - 0,80x_6 + 0,13x_5 - 0,21x_3 - 0,20x_4 - \\ &\quad - 0,07x_1 + 0,08x_8. \end{aligned} \right\} \quad (\text{XII.36})$$

Независимые переменные расположены в порядке уменьшения влияния на зависимые переменные.

Подставляя фактические значения факторов $x_1, x_3, x_5, x_6, x_8, x_4$ го каждому предприятию в уравнения (XII.36), можно вычислить выработку в сметной стоимости на одного работающего и себестоимость на 1 руб. сметной стоимости для всех предприятий.

Однако результаты, полученные по уравнениям (XII.36), еще недостаточны для сравнительного анализа. Для сопоставления предприятий по уровню производительности труда и себестоимости значения в сопоставимый вид, т. е. рассчитать коэффициенты приведения к условиям производства работ, принятые за базовые, на основе количественной оценки влияния на величины $y_i = \varphi(x_{ij})$, $i = 1, 2, \dots, n$ комплекса неуправляемых факторов x_{ij} ($j = 1, 2, \dots, N$). Под коэффициентом приведения K_i понимается система чисел, позволяющая сравнить, насколько изменится величина y_i при переходе от условий, принятых за базовые y_0 , к i -м условиям производства:

$$K_i = \frac{y_i - y_0}{y_0}. \quad (\text{XII.37})$$

За базовые можно принять средние значения исследуемых величин, т. е. $y_0 = \bar{y}$. Подставляя в выражение (XII.37) значение y_i из формулы (XII.35) и учитывая, что базовые значения определяются как

$$\bar{y} = a_0 + \sum_{j=1}^N a_j \bar{x}_j, \quad (\text{XII.38})$$

коэффициенты приведения можно вычислить по формуле

$$K_i = \sum_{j=1}^N \frac{100a_j}{\bar{y}} (x_j - \bar{x}_j). \quad (\text{XII.39})$$

В уравнении (XII.39) отклонения фактических значений j -х факторов по предприятиям от одноименных базовых значений для удобства вычислений выражены в абсолютных величинах, а не в процентах. Исходными данными для расчета коэффициентов приведения к средним условиям топографо-геодезического производства для всех предприятий в анализируемый период служат уравнения регрессии функций y_1 и y_2 , определяемые зависимостями (XII.36). Подставляя значения коэффициентов уравнения (XII.37) в уравнение (XII.39), получаем коэффициенты приведения:

для выработки в сметной стоимости на одного работающего

$$K_i = 1,20x_6 + 30,48x_5 + 4,94x_3 - 3,29x_1 - 0,29x_8 - 0,18x_4, \quad (\text{XII.40})$$

для себестоимости на 1 руб. сметной стоимости

$$K_i = 1,05x_6 + 0,16x_5 - 0,26x_3 - 0,25x_4 + 0,09x_1 + 0,09x_8. \quad (\text{XII.41})$$

По уравнениям (XII.40) и (XII.41) рассчитаны коэффициенты приведения для ряда предприятий. Пример расчета сопоставимых значений выработки в сметной стоимости для предприятий, имеющих условные индексы A , B , B , и себестоимости для тех же предприятий приведен в табл. 73, 74.

Сопоставимые значения себестоимости по предприятиям

Показатель себестоимости	Предприятия		
	А	Б	В
Базовое значение, коп.	80	80	80
Фактическое значение, коп.	88	72	90
Коэффициент приведения K_i , %	—10,0	15,6	—12,4
Сопоставимые значения, коп.	80	83	79

Как видно из табл. 74, коэффициенты K_i для предприятий А, Б, В равны соответственно —30, 40 и —34 %. Это означает, что для приведения фактической выработки в сметной стоимости на одного работника к среднеотраслевому уровню необходимо изменить фактические значения указанных показателей для предприятий А, Б, В соответственно на —30, 40 —34 %. Эти же величины можно использовать при оценке предприятий по уровню производительности труда, а отклонения между ними объясняются только качеством работы предприятий, так как в расчете учтено влияние большинства объективных факторов, не зависящих от деятельности предприятия и поддающихся количественному определению. Степень различия полученных показателей характеризует уровень эффективности производства по производительности труда предприятий А, Б, В.

Этот метод сравнительного анализа достаточно точно отражает, существующее положение. Полученные результаты согласуются с экспертной оценкой специалистов, назвавших в числе передовых по выработке в сметной стоимости предприятие Б.

Таблица 74

Исходные данные расчета сопоставимых значений выработки на одного работника по предприятиям

Показатель	Базовое значение показателя	Предприятие								
		А			Б			В		
		фактическое значение	отклонение	поправка	фактическое значение	отклонение	поправка	фактическое значение	отклонение	поправка
y		6898			8855			7590		
x_6	22,44	38,00	—15,96	—18,70	11,00	11,44	13,75	31,00	—8,56	—1029
x_5	1,77	1,97	—0,20	—6,10	1,35	0,42	12,36	2,22	—0,45	—13,71
x_4	15,20	14,30	1,00	4,94	13,90	1,40	6,92	15,00	0,30	1,49
x_3	5,92	3,90	2,02	—6,65	6,20	—0,28	0,92	3,80	2,12	—6,97
x_2	68,68	47,00	21,68	—6,28	84,00	—15,32	4,44	49,00	19,68	—5,71
x_1	65,59	73,00	—7,41	1,43	73,00	—7,41	1,43	70,00	—4,41	0,79
K_i				—0,36						—24,41

Аналогичные расчеты выполнены для ряда предприятий ГУГК по показателям себестоимости, выраженной фактическими затратами на 1 руб. сметной стоимости. Из табл. 73 видно, что по этому показателю эффективнее работало предприятие В.

Коэффициенты приведения, рассчитанные в соответствии с изложенной схемой (см. табл. 73, 74), могут быть использованы при утверждении плановых заданий топографо-геодезическим подразделением по производительности труда, себестоимости работ и др.

Если после расчетов получается система коэффициентов вида (XII.37), то при распределении показателей по выработке в сметной стоимости на одного работающего между предприятиями топографо-геодезического производства необходимо соблюдение условия

$$\frac{\sum_{i=1}^n v_i}{\bar{y}} = \sum_{i=1}^n \frac{v_i}{y_i}, \quad (\text{XII.42})$$

где v_i — объем работ в сметной стоимости i -го подразделения.

Будем считать, что разность в уровнях между предприятиями определяется коэффициентами K_i . Найденные из формулы (XII.39) значения y_i подставляем в условие (XII.42) и получаем

$$\frac{\sum_{i=1}^n v}{\bar{y}} = \sum_{i=1}^n \frac{v_i}{y_0 \left(1 - \frac{K_i}{100}\right)}. \quad (\text{XII.43})$$

Так как $\sum_{i=1}^n v_i$, $\bar{y}$, v_i заданы, то из уравнения (XII.43) можно отыскать базовое значение y_0 .

На упрощенном примере можно показать порядок распределения плановых заданий по выработке в сметной стоимости между условными предприятиями А, Б, В. Пусть на плановый период группе из трех предприятий задается:

1) общий объем в сметной стоимости

$$\sum_{i=1}^3 v_i = 27\,450 \text{ тыс. руб.};$$

2) объемы работ в сметной стоимости по каждому предприятию $v_A = 11\,050$ тыс. руб., $v_B = 5700$ тыс. руб., $v_V = 10\,700$ тыс. руб.,

3) среднее значение выработки, которое должно быть достигнуто в плановом периоде $\bar{y} = 5766$ руб.;

4) коэффициенты приведения, полученные на плановый период, $K_A = 10$, $K_B = -32$, $K_V = 34$.

Общее задание по производительности труда может быть распределено между предприятиями А, Б, В следующим образом. Подставляя исходные данные в уравнение (XII.43), решаем его относи-

тельно y_0 и получаем $y_0 = 5548$ руб. Следующим этапом расчетов является определение индивидуальных заданий предприятиям по производительности труда. С учетом коэффициентов приведения на плановый период получаем $y_A = 6103$ руб., $y_B = 3773$ руб., $y_B = 7434$ руб. Фактическая выработка в сметной стоимости (в руб.) на одного работающего, достигнутая предприятиями за анализируемый год,

$$y_A^{\Phi} = 5760 \text{ руб.}, \quad y_B^{\Phi} = 3919 \text{ руб.}, \quad y_B^{\Phi} = 7620 \text{ руб.},$$

т. е. отклонения фактических значений от рассчитанных

$$\Delta_A = -343 \text{ руб.}, \quad \Delta_B = 146 \text{ руб.}, \quad \Delta_B = 186 \text{ руб.}$$

Таким образом, рассчитанные значения соответствуют фактически достигнутым результатам, отличаясь от них на $-5,9$; $3,7$ и $2,4$ %.

Выполненные исследования и полученные результаты показали, что изложенный метод расчета коэффициентов приведения может быть использован для учета и соизмерения объективных условий работы топографо-геодезических предприятий, а также для сравнительного анализа уровня эффективности топографо-геодезического производства. Этот метод может служить для устранения искажающего влияния стоимостных структурных сдвигов, дает возможность определять их действительное влияние на величину таких показателей эффективности производства, как себестоимость, производительность труда, прибыль, фондоотдача и др. Это, в свою очередь, необходимо для объективного обоснования мощностей подразделений, их размерных групп и совершенствования планирования.

Глава XIII

АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ УПРАВЛЕНИЯ В ТОПОГРАФО- ГЕОДЕЗИЧЕСКОМ ПРОИЗВОДСТВЕ

§ 60. Сущность и основные предпосылки создания автоматизированных систем управления на современном этапе

Быстрый рост масштабов топографо-геодезического производства, его динамичность, широкая автоматизация технологических процессов вызывают увеличение объемов учетной, перерабатываемой, хранимой и выдаваемой информации, влекут за собой увеличение штатов административно-управленческого персонала. За последние десять лет число топографо-геодезических предприятий значительно увеличилось. Развитие новых направлений топографо-геодезических работ, технических средств и методов их выполнения обусловило быстрый рост промежуточных технологических процессов и отчетных показателей в экспедициях и цехах предприятий. В свою очередь, перегрузка различных звеньев системы управления наряду с запаздыванием обработки информации приводит к значительным производственным потерям. При этом простое увеличение численности управленческого персонала, как правило, не решает возникшей проблемы улучшения качества управления производством. Таким образом, возникла объективная необходимость искать новые пути совершенствования управления топографо-геодезическим производством.

Один из самых эффективных путей повышения производительности и качества управленческого труда на современном этапе применение средств вычислительной техники и экономико-математических методов. Необходимой предпосылкой успешного применения указанных средств и методов в топографо-геодезическом производстве является заметный рост оснащенности предприятий и фабрик ГУГКа вычислительной техникой и средствами связи. В десятой пятилетке вычислительные центры предприятий и НИИ ГУГКа стали оснащаться ЭВМ третьего поколения. Эти машины получили название ЕС ЭВМ (единая система ЭВМ). В настоящее время вычислительные центры (ВЦ) предприятий оснащены ЕС ЭВМ с различной памятью и быстродействием. К этим ЭВМ разработан и серийно выпускается большой комплекс дополнительных и периферийных устройств. Важной предпосылкой к успешному внедрению ЭВМ и экономико-математических методов в сфере управления топографо-геодезическим производством являются также выполненные исследования по указанной проблеме в научных учреждениях и предприятиях ГУГКа.

Начало широкого развертывания работ по практическому применению экономико-математических методов и ЭВМ в задачах планирования и управления в нашей стране относится к 1957—1962 гг. К этому времени был накоплен известный опыт разового решения отдельных производственных задач, который показал, что кибернетические модели и средства таят в себе огромные возможности. Вместе с тем все более становилось ясным, что подлинной эффективности от применения экономико-математических методов и ЭВМ можно ожидать лишь при условии комплексного подхода к экономическому объекту, когда предметом исследования и реорганизации на новой основе становится вся система управления, а не отдельные управленческие задачи. Это означает, что должна быть построена совокупность взаимосвязанных моделей управления, описывающих различные функции и уровни управления. Таким образом, возникает задача создания комплексных методов и средств управления, где отдельные модели, алгоритмы и способы их организации объединены в стройную систему. Иначе говоря, возникает задача создания автоматизированных систем управления (АСУ).

Работа по созданию АСУ за последние годы приобретает все более широкий размах.

В девятой пятилетке в Главном управлении геодезии и картографии были начаты, в десятой пятилетке продолжены работы по созданию и развитию автоматизированных систем управления и обработки информации. В настоящее время разработана и сдана в промышленную эксплуатацию в объеме 1-й очереди отраслевая автоматизированная система управления ГУГК (ОАСУ ГУГК).

АСУ представляет собой организационно-технический комплекс, обеспечивающий выполнение функций управления производством на основе использования экономико-математических и современных средств обработки данных (ЭВМ, накопительные устройства, устройств регистрации, отображения и др.). Основная цель разработки АСУ — совершенствование управления на соответствующих уровнях управления народным хозяйством, обеспечение интенсификации производственно-хозяйственных и технологических процессов, наилучшее использование материальных и трудовых ресурсов, увеличение выпуска продукции, улучшение ее качества, повышение рентабельности и прибыли.

На ранних этапах использования ЭВМ и экономико-математических методов в управлении количество функций и задач в области технико-экономического планирования, выполняемых машинами, было незначительно и они были довольно четко разграничены. В АСУ эти функции и задачи представляют собой взаимосвязанную систему, причем человек освобождается от выполнения неквалифицированного труда и за ним остаются функции творческого характера. Например, АСУ создает исключительно благоприятные возможности для проведения аналитических и различного рода исследовательских работ, в которых творческая мысль человека играет первостепенную роль.

АСУ относится к так называемым человеко-машинным системам. Центральное место в них занимает человек, который активно участвует в регулировании производством, используя для этой цели вычислительную технику и другие средства. Вычислительная машина, как бы совершенна она ни была, играет, однако, вспомогательную роль при решении производственно-хозяйственных задач предприятия, объединения, отрасли. В этом заключается принципиальное отличие АСУ от автоматических систем управления технологическими процессами, в которых управление машинами (приборами) осуществляется программными устройствами.

В иерархии систем управления различают (снизу вверх) следующие группы:

автоматизированные системы управления объединениями, предприятиями (АСУП). Эти системы относятся к низовому звену в иерархии систем управления. В зависимости от области применения ЭВМ и видов обрабатываемой информации на этом уровне различают также автоматизированные системы обработки информации (АСОИ), автоматизированные информационно-поисковые системы (АИПС) и автоматизированные системы управления технологическими процессами (АСУТП);

отраслевые автоматизированные системы управления (ОАСУ). Эти системы образуют следующий (второй) уровень в иерархии управления. К ним относятся автоматизированные системы управления министерств и ведомств как общесоюзного, так и республиканского значения;

автоматизированные системы функциональных органов управления общегосударственного значения — это системы третьего уровня в иерархии АСУ. В Госплане СССР создана, например, автоматизированная система плановых расчетов (АСПР), в ЦСУ СССР функционирует автоматизированная система управления государственной статистики (АСГС); в Госснабе СССР действует автоматизированная система управления материально-техническим снабжением (АСУМТС). Существует целый ряд и других АСУ функциональных органов управления при государственных комитетах, Госбанке СССР, Министерстве финансов СССР;

общегосударственная автоматизированная система (ОГАС) сбора и обработки информации для учета, планирования и управления народным хозяйством на базе государственной сети вычислительных центров и единой автоматизированной системы связи страны. Главная задача ОГАС — обеспечить органы планирования и управления союзного, республиканского и территориального значений достоверной информацией, необходимой для решения задач учета, планирования и управления. За основу структуры ОГАС принята иерархическая структура, соответствующая системе государственных органов планирования и отраслевому принципу управления народным хозяйством. В ОГАС выделяются три звена: функциональное, отраслевое и административно-территориальное. Функциональное звено ОГАС образует автоматизированные системы функциональных органов управления общегосударственного

значения (АСПР, АСГС, АСУ МТС и др.). Отраслевое звено ОГАС — ОАСУ образует автоматизированные системы управления отраслевых союзных, союзно-республиканских министерств и ведомств, отраслевых органов управления местного подчинения. Административно-территориальное звено ОГАС образует автоматизированные системы управления республиками, краями, областями и другими, административно-территориальными единицами страны.

§ 61. Структура, составные части и основные принципы построения АСУ

Для каждого вида АСУ можно выделить присущие системе основные функции, состав и организационную структуру.

Функции ОАСУ ГУГК определяются задачами, стоящими перед Главным управлением геодезии и картографии по планированию и управлению топографо-геодезическим и картографическим производством в условиях постоянного роста масштабов работ и объемов перерабатываемой информации. В соответствии с техническим заданием на создание ОАСУ ГУГК основными функциями системы являются:

планирование топографо-геодезических и картографических работ, материально-технического снабжения, труда и заработной платы по предприятиям и организациям ГУГК; учет и контроль производственной деятельности, качества выполнения работ, использования материально-технических ресурсов и оборудования, выполнения планов по труду, состава и движения кадров, хозяйственной деятельности подразделений ГУГК;

анализ выполнения производственных планов, использования основных производственных фондов, трудовых показателей, движения кадров и основных показателей хозяйственной деятельности предприятий и организаций ГУГК;

информационно-справочное обслуживание центрального аппарата ГУГКа по его запросам.

Реализация этих функций ОАСУ ГУГК позволит сократить сроки и затраты на обработку, повысить достоверность и аналитичность технико-экономической информации, организовать решение принципиально новых задач и будет способствовать изменению характера труда работников центрального аппарата, которые будут освобождены от счетной работы и смогут больше внимание уделять вопросам анализа и принятия решений.

В состав ОАСУ ГУГКа входят шесть функциональных и четыре обеспечивающие подсистемы.

Под функциональной подсистемой понимается часть системы, выделяемая по общности функциональных признаков и объединяющая комплексы задач, административных и организационных методов, необходимых для реализации заданных функций управления. В составе ОАСУ ГУГКа выделены следующие функциональные подсистемы: «Технико-экономическое планиро-

вание», «Оперативное управление», «Планирование, учет и анализ труда и заработной платы», «Планирование, учет и анализ кадров», «Бухгалтерский учет».

Обеспечивающие подсистемы в зависимости от их назначения объединяют наборы необходимых элементов, обуславливающих возможность решения функциональных задач. В ОАСУ ГУГКа к ним относится информационное, математическое, техническое и программное обеспечение*.

Информационное обеспечение ОАСУ ГУГКа представляет собой совокупность показателей, документов, классификаторов, словарей, массивов информации, а также методов их организации, хранения и контроля, обеспечивающих взаимосвязанное решение комплексов задач ОАСУ и информационное взаимодействие с внешними автоматизированными системами.

Математическое обеспечение представляет собой совокупность математических методов, моделей и алгоритмов для решения задач и обработки информации с применением вычислительной техники в ОАСУ ГУГКу.

Техническое обеспечение ОАСУ ГУГКа включает комплекс технических средств (КТС), определяющий ее эффективное функционирование. Он представляет собой совокупность устройств (оргтехника, периферийная техника, ЭВМ, средства связи), предназначенных для реализации процессов сбора, передачи, обработки, хранения и отображения информации, а также связи с предприятиями и организациями ГУГКа.

Программное обеспечение — это совокупность программных средств, необходимых для реализации задач ОАСУ ГУГКа и обеспечивающих функционирование комплекса технических средств. Оно состоит из двух основных частей: системного программного обеспечения и конкретного программного обеспечения. Системное обеспечение объединяет программные средства, предназначенные для автоматизации прохождения задач через вычислительную систему, средства автоматизации программирования, а также совокупность тестовых программ, контролирующих состояние вычислительной системы. Конкретное программное обеспечение представляет собой совокупность программ, создаваемых с помощью системного программного обеспечения и предназначенных для решения конкретных задач функциональных подсистем ОАСУ ГУГКа.

В состав организационной структуры ОАСУ ГУГКа входят следующие звенья:

аппарат ГУГКа, оснащенный устройствами, связанными с Головным информационно-вычислительным центром (ГИВЦ);

ГИВЦ с комплексом технических средств, включая ЭВМ и средства узла связи;

* В состав обеспечивающей части АСУ министерств и ведомств могут также входить подсистемы кадрового и правового обеспечения.

информационные пункты предприятий и организаций ГУГК, связанные непосредственно с ГИВЦ.

Основными функциями ГИВЦ является: сбор, накопление и обработка поступающей от подразделений ГУГКа информации; формирование и выдача аппарату ГУГКа необходимых данных; передача предприятиям и организациям ГУГКа оперативных решений, принятых руководством.

Информационные пункты предназначены для подготовки исходящей информации и передачи ее в ГИВЦ по каналам связи для решения задач ОАСУ ГУГКа.

АСУ ТПП создается в неразрывной связи с ОАСУ ГУГКа, что позволяет обеспечить их единое функционирование и взаимодействие.

Функциональная часть АСУ ТПП состоит из пяти подсистем: «Технико-экономическое планирование», «Нормативное хозяйство», «Учет», «Оперативное управление», «Материально-техническое снабжение».

Кроме функциональной части АСУ ТПП, как и ОАСУ ГУГКа, имеет обеспечивающую часть, в которую включены подсистемы информационного, математического, технического и программного обеспечения.

Состав организационной структуры АСУ ТПП так же, как и ОАСУ ГУГКа, территориально имеет три уровня управления: экспедиции (цеха) — вычислительный центр — аппарат предприятия.

Одновременно с проектированием и внедрением АСУ ТПП в предприятиях и организациях ГУГКа созданы и развиваются АСОИ. Несмотря на более низкий уровень программного и технического обеспечения по сравнению с ОАСУ ГУГКа и АСУ ТПП эти системы сыграли большую роль в решении как экономико-организационных, так и технологических задач. Кроме того, АСОИ способствовали повышению эффективности использования электронно-вычислительной и счетно-перфорационной техники в системе ГУГКа.

Функциональная часть АСОИ состоит из двух основных подсистем: «Обработка технологической информации», «Обработка технико-экономической информации».

Во всех звеньях АСОИ создана единая организационная структура, включающая три группы: разработку АСОИ; приема, контроля и выпуска информации; выполнения расчетов и хранения информации.

По мере развития программного и технического обеспечения АСОИ предусматривается постепенный их перевод в АСУ ТПП. На базе подсистемы обработки технологической информации в дальнейшем будут создаваться самостоятельные автоматизированные системы управления технологическими процессами (АСУ ТП) на предприятиях и фабриках ГУГКа.

Существуют различные классификации принципов построения АСУ. С учетом экономико-организационных особенностей топогра-

фо-геодезического производства, на наш взгляд, можно выделить следующие принципы, реализованные при создании ОАСУ ГУГКа и АСУ ТПП: 1) комплексный, системный подход; 2) перспективный подход; 3) оптимальный подход; 4) интегрированную обработку данных; 5) постановку новых задач; 6) развитие системы.

Рассмотрим сущность перечисленных принципов проектирования АСУ.

АСУ проектируется с учетом всех требований, предъявленных к управлению. Она должна предусматривать выполнение всех функций управления. Для реализации этого условия в состав АСУ включаются соответствующие подсистемы. Соблюдение этого условия и является принципом комплексности. С ним неразрывно связан принцип системного подхода. Принцип системного подхода предусматривает разделение системы на элементы, изменяющиеся во времени, взаимосвязанные и взаимодействующие в процессе функционирования объекта управления. Системный подход предполагает анализ связей и взаимодействий между элементами системы в динамике, а также в перспективе.

Внедрение АСУ, как правило, должно сопровождаться крупными мероприятиями организационного характера: изменением форм документов, изменениями организационной и функциональной структуры органов управления, совершенствованием системы показателей, нормативной базы и др. Не всегда все вопросы могут быть решены в рамках самой разрабатываемой системы. Принцип системного подхода требует, чтобы все такие вопросы были выявлены и поставлены в соответствующих инстанциях.

Принцип перспективного подхода вытекает из необходимости учета возможных долговременных последствий научно-технического прогресса и структурных сдвигов в развитии экономики, которые приводят к перестройке хозяйственных связей, к необходимости изменения организационной структуры управления. Кроме того, соблюдение принципа перспективного подхода связано с ограниченностью материальных и трудовых ресурсов, необходимостью планирования перспективных результатов производственной деятельности.

Системный и перспективный подходы предполагают оптимизацию как развития, так и связей объекта управления. Оптимальный план развития топографо-геодезического производства — это план топографо-геодезических работ, максимально удовлетворяющих потребности министерств и ведомств, при ограничении в каждый данный момент времени производственных ресурсов. Этим подчеркивается необходимость соблюдения принципа оптимального подхода при создании АСУ. Соблюдение принципа оптимального подхода позволяет решать задачи планирования принципиально по-новому, обеспечивает рациональное распределение производственных ресурсов, размещение производства и выбор наиболее эффективных направлений развития предприятий и организаций ГУГКа.

Принцип интегрированной обработки данных заключается в том, что система должна строиться на основе единой информационной базы и предусматривать выполнение всех функций управления — планирования, учета, контроля, анализа и регулирования. Сущность такого подхода состоит в том, что на машинных носителях накапливается (и постепенно обновляется) информация, необходимая для решения не какой-то одной или нескольких задач, а всех задач управления. При этом обеспечивается однократный ввод значения любого показателя, характеризующего состояние системы и многократный вывод совокупности показателей в различных сочетаниях и формах документов.

Принцип постановки новых задач предполагает не просто перекладывать на ЭВМ традиционно сложившиеся методы и приемы управления, алгоритмы решения сложившихся в старой системе задач управления, а пересматривать эти методы в соответствии с новыми возможностями, которые открывают ЭВМ и экономико-математические методы. На практике это значит, что в функциональный состав АСУ включаются задачи оптимального планирования, прогнозирования, системного анализа и другие, которые из-за сложности алгоритмов, расчета и больших объемов информации в условиях традиционных методов и техники управления не решались.

Принцип развития системы заключается в том, что ее создание осуществляется в несколько этапов. Кроме того, необходимо обеспечить возможность включения появляющихся задач и совершенствования уже решаемых задач.

§ 62. Организационные основы проектирования и очередность создания АСУ

Проектирование ОАСУ ГУГКа, АСУ ТПП и АСОИ осуществляется в соответствии с руководящими методическими материалами по созданию АСУ и включает следующие этапы: 1) обследование объектов управления и разработку технического задания; 2) разработку технического проекта; 3) разработку рабочего проекта; 4) внедрение.

Объектами проектирования являются обеспечивающие и функциональные подсистемы, территориальные звенья и другие элементы АСУ.

Независимо от этапа разработки проектные материалы делятся на:

общесистемные проектные материалы, содержащие проектные решения, относящиеся к системе в целом (организационная и функциональная структура АСУ, общесистемные классификаторы и словари, общая схема постановки задач, системное программное обеспечение и др.);

конкретные проектные материалы (постановка задач, конкретные программы решения задач, специфические схемы обработки информации, инструкции и др.).

На этапе обследования объекта управления и разработки технического задания производится общее ознакомление с объемами входящей и исходящей информации, ее содержанием, динамикой и связями информационных потоков, организационной структурой управления, функциями и задачами аппарата управления. После определения комплекса задач АСУ проводится детальное обследование информации, необходимой для их решения. В результате такого обследования определяются состав, периодичность и объем информации для каждой проектируемой подсистемы и задачи АСУ. Материалы обследования служат основой подготовки технического задания. В техническом задании определяются: общие положения и цель создания АСУ; состав подсистем, функции и задачи, подлежащие автоматизации; требования, предъявляемые к информационной базе, математическому и программному обеспечению, комплексу технических средств; организация работ и исполнители; этапы разработки и сроки их выполнения; предварительная оценка эффективности внедрения.

Технический проект разрабатывается с целью определения основных проектных решений и содержит технический проект системы в целом и технические проекты функциональных подсистем.

В разделах технического проекта на систему в целом приводятся проектные решения по следующим вопросам: 1) назначению системы и ее функциональной и организационной структуре; 2) методам управления производством в условиях функционирования автоматизированной системы; 3) видам и содержанию информации, обрабатываемой с помощью комплекса технических средств (КТС) в вычислительном центре; 4) параметрам информационных потоков; 5) типам машинных носителей для каждого вида информации; 6) принципам построения и характеристикам классификаторов и словарей; 7) методам контроля информации; 8) составу информационного фонда АСУ; 9) принципам организации информационных массивов на машинных носителях; 10) порядку хранения информации, ее поиска и внесения изменений; 11) математическим методам решения задач; 12) составу программного обеспечения АСУ, включая средства автоматизации программирования, конкретные программы решения задач на ЭВМ, управление процессом решения всей совокупности задач системы; 13) структуре комплекса технических средств и его функциональных групп; 14) технологии обработки данных в системе; 15) оценке экономической эффективности внедрения системы; мероприятиям по подготовке АСУ к внедрению.

Разделы технических проектов функциональных подсистем характеризуют: назначение подсистем и целесообразность автоматизации их функций; функциональный состав подсистем и их организационные структуры; экономико-математическую постановку и информационное обеспечение отдельных комплексов задач, входящих в состав подсистемы. По каждой задаче определяют ее назначение, постановку, периодичность решения, информационные связи задачи, алгоритм решения, характеристики входной, хра-

нимой и выходной информации, а также сроки ее передачи и выдачи; образцы форм входных и выходных документов.

Цель рабочего проектирования АСУ — окончательная разработка материалов, необходимых для внедрения системы. Рабочий проект АСУ содержит сводный рабочий проект, в котором излагаются общесистемные проектные решения и рабочие проекты функциональных подсистем.

В общесистемных разделах рабочего проекта содержатся следующие материалы: уточненные проектные решения по организационному построению системы, ее информационному, программному и техническому обеспечению; классификаторы технико-экономической информации; общесистемные инструкции по организации информационной базы системы; альбомы форм входных и выходных унифицированных документов; план-график мероприятий по внедрению системы; уточненный расчет экономической эффективности АСУ.

В разделах технических проектов на отдельные функциональные подсистемы приведены: характеристики функциональных структур, информационного, математического и программного обеспечения подсистем; инструктивные материалы, регламентирующие порядок сбора, передачи, обработки, хранения и выдачи информации (инструкции по подготовке данных для передачи по каналам связи, инструкции о порядке приема информации в ВЦ, инструкции по перфорации, инструкции по эксплуатации программ и др.); словари технико-экономической информации; описание программ решения задач; мероприятия по подготовке подсистем к внедрению.

Внедрению автоматизированных систем управления предшествует опытная эксплуатация ее подсистем, осуществляемая на реальных информационных массивах и предусматривающая решение всех внедряемых задач и анализ результатов решения, проверку готовности обеспечивающих подсистем и надежности их функционирования.

По результатам апробации вносятся необходимые коррективы в элементы информационного и программного обеспечения и организации технологического процесса. Внедрение системы оформляется актом в установленном порядке.

Для организации и координации проектных работ составляется и утверждается координационный план работ по созданию АСУ; устанавливаются генеральный заказчик на систему в целом, заказчик на функциональные подсистемы и определяются их функции; устанавливается головной разработчик, т. е. организация, ответственная за проектирование системы в целом и обеспечивающая единое направление разработок подсистемы и элементов системы, а также определяющая разработчиков (организации-исполнители) отдельных подсистем и их функции.

В соответствии с принципом развития АСУ ее проектирование и внедрение осуществляется поочередно. Проектная очередь системы — это законченный комплекс внедренных подсистем и задач, выполняемых на определенной технической базе. Очереди АСУ раз-

личаются по объему обрабатываемой в автоматизированном режиме информации, составу подсистем и комплексов задач, типу применяемых ЭВМ, совершенству технологических процессов, степени интеграции и автоматизации обработки данных, уровню взаимодействия и совместимости с другими автоматизированными системами.

§ 63. Функциональная и обеспечивающая части отраслевой автоматизированной системы управления ГУГКа

Функциональная часть ОАСУ ГУГКа объединяет относительно взаимосвязанные функциональные подсистемы: 1) технико-экономическое планирование (ТЭП); 2) оперативное управление (ОУ); 3) управление материально-техническим снабжением (УМТС); 4) планирование, учет и анализ труда и заработной платы (ПУ и АТЗ); 5) планирование, учет и анализ кадров (ПУ и АК); 6) бухгалтерский учет (БУ).

Функциональные подсистемы решают задачи по планированию, учету и проводят анализ производственно-хозяйственной деятельности топографо-геодезических предприятий и фабрик ГУГКа. Комплексы задач по подсистемам, которые предусмотрены техническим заданием на создание второй очереди ОАСУ ГУГК, приведены в табл. 75. Эти комплексы определены на основе анализа важнейших функций управления топографо-геодезическим производством, выполняемых аппаратом ГУГКа, а также с учетом реальных возможностей привлечения к разработке системы специалистов и выделенных средств на создание технической базы ОАСУ ГУГКа.

Как видно из табл. 75, в состав второй очереди ОАСУ ГУГКа включены 39 комплексов задач (свыше 150 выходных форм). Так, например, создание подсистемы ТЭП предусматривает расчет проектов оптимальных вариантов пятилетних и годовых планов предприятий (фабрик), формирование аналитических таблиц, характеризующих использование основного оборудования, состояние заделов по основным технологическим процессам, исследования себестоимости работ по статьям и элементам затрат. В результате решения задач на ЭВМ аппарату ГУГКа предоставляются выходные сводные документы и аналитические таблицы, позволяющие рационально использовать дефицитные ресурсы, вскрыть резервы увеличения объемов и снижения себестоимости работ, повысить экономичность и рентабельность производства.

Обеспечивающая часть ОАСУ ГУГКа включает информационное, математическое, программное и техническое обеспечение.

Как видно из табл. 76, системы классификации и кодирования совместно с системой документации образуют язык информационного обеспечения АСУ, создаваемый на основе взаимосвязанной системы показателей и необходимый для обеспечения взаимодействия всех элементов АСУ как между собой, так и с внешними системами управления.

Комплексы задач по подсистемам

Код под-системы	Подсистема	Комплекс задач
01	Технико-экономическое планирование	1. Формирование и выдача аппарату ГУГК сведений о производственно-хозяйственной деятельности отрасли, топографо-геодезических предприятий и фабрик 2. Расчет основных показателей пятилетнего плана топографо-геодезических предприятий 3. Расчет основных технико-экономических показателей производственного плана топографо-геодезических предприятий 4. Расчет основных технико-экономических показателей производственного плана фабрик 5. Расчет производственной мощности топографо-геодезических предприятий 6. Расчет производственной мощности фабрик 7. Анализ себестоимости топографо-геодезических работ 8. Анализ себестоимости картографической продукции 9. Анализ себестоимости продукции приборостроения 10. Учет и анализ заделов топографо-геодезических работ 11. Анализ использования основного производственного оборудования на фабриках
02	Оперативное управление	1. Учет и контроль выполнения месячных, квартальных и годовых производственных планов топографо-геодезических предприятий 2. Учет и контроль выполнения месячных и квартальных производственных планов фабрик 3. Учет и контроль выполнения месячных и квартальных производственных планов предприятий приборостроения 4. Учет и контроль наличия и движения основного производственного оборудования по топографо-геодезическим предприятиям 5. Учет и контроль наличия и движения основного производственного оборудования по фабрикам 6. Учет и контроль наличия и движения основного производственного оборудования по предприятиям приборостроения 7. Составление оперативных сведений о выполнении поручений по приказам и директивным указаниям ГУГКа 8. Учет и контроль качества выполнения работ по топографо-геодезическим предприятиям 9. Составление сведений о состоянии труда и техники безопасности в топографо-геодезических предприятиях

Код под-системы	Подсистема	Комплекс задач
03	Управление материально-техническим снабжением	1. Расчет средних норм расходования материальных ресурсов 2. Расчет годовой потребности в основных материальных ресурсах на производственные нужды 3. Составление сводной статистической отчетности по формам № 1-СН, № 1-СН (ОСТ), № 4-СН 4. Учет и контроль использования дефицитных материалов по топографо-геодезическим предприятиям 5. Учет и контроль использования дефицитных материалов по фабрикам
04	Планирование, учет и анализ труда и заработной платы	1. Учет и контроль выполнения плана по труду и заработной плате по ГУГКу и его подразделениям 2. Анализ соотношения темпов роста производительности труда и заработной платы по ГУГКу и его подразделениям 3. Анализ использования фонда заработной платы по ГУГКу и его подразделениям 4. Анализ средней заработной платы по топографо-геодезическим предприятиям, фабрикам отрасли 5. Анализ численности персонала по ГУГКу и его подразделениям 6. Расчет показателей и плана по труду и заработной плате по топографо-геодезическим предприятиям 7. Расчет показателей плана по труду и заработной плате по фабрикам
05	Планирование, учет и анализ кадров	1. Составление сводных годовых отчетов по кадрам ГУГКа 2. Учет и анализ состава и распределения кадров в отрасли 3. Учет и анализ движения кадров 4. Учет и анализ деловых качеств работников отрасли 5. Определение дополнительной потребности организаций ГУГКа в специалистах
06	Бухгалтерский учет	1. Контроль бухгалтерских отчетов топографо-геодезических и промышленных предприятий 2. Составление сводной бухгалтерской периодической и годовой отчетности в разрезе видов деятельности (топографо-геодезическое производство, промышленные и бюджетные организации)

Под классификацией понимается упорядоченное распределение элементов соответствующего множества, выполненное по общности какого-либо признака. Такое множество представляет собой объект классификации, а его элементы — классифицируемые единицы. Понятие, обозначающее совокупность свойств, по которым различаются элементы в пределах множества, является признаком классификации.

Так, например, информация, используемая в ОАСУ ГУГКа, характеризуется следующими основными классификационными признаками: направлением движения, функциональным составом, степенью стабильности информации, периодом использования.

По направлению движения (по отношению к ГИВЦ) информация подразделяется на входную, внутреннюю (промежуточную) и выходную.

Информация по функциональному составу делится на нормативно-справочную, включая словарную информацию, плановую, оперативно-производственную, отчетную и статистическую.

По степени стабильности информация подразделяется на постоянную, переменную и условно-постоянную.

По периоду использования информацию можно разделить на годовую, полугодовую, квартальную, месячную и разовую (по запросам).

Упорядоченное расположение классифицируемых элементов на основе установленных связей и зависимостей между признаками составляет систему классификации. Эта система определяет комплекс классификаторов, для каждого из которых устанавливается сфера его действия.

В табл. 77 приведен перечень классификаторов технико-экономической информации ОАСУ ГУГКа.

На основе классификаторов технико-экономической информации разрабатываются словари на машинных носителях. Словари — это

Т а б л и ц а 76

Состав и основные функции информационного обеспечения АСУ

Основные элементы информационного обеспечения		Основные функции информационного обеспечения				
		Формирование данных	Корректировка и преобразование данных	Хранение и защита данных	Поиск и выдача данных	
					Для решения задач	для информационного обслуживания пользователей
Элементы языка информационного обеспечения	Система классификации информации					
	Система кодирования информации					
	Система документации					
	Массивы информации					

Классификаторы технико-экономической информации ОАСУ ГУГКа

Классификатор	Вид классификатора
Высшие классификационные группировки общесоюзного классификатора промышленной и сельскохозяйственной продукции (ВКГ ОКП)	Общесоюзный
Система обозначений единиц величин и счета, используемых в автоматизированных системах управления народным хозяйством для обработки технико-экономической информации (СОЕВС)	»
Общесоюзный классификатор профессий рабочих, должностей служащих и тарифных разрядов (ОКП ДТР)	»
Информация по кадрам	»
Раздел ОКП «Картографическая продукция»	»
Основные фонды	Отраслевой
Материальные ценности	»
Структурные подразделения ГУГКа	»
Виды и процессы топографо-геодезических работ	»
Основные технико-экономические показатели	»
Подсистемы ОАСУ	»
Задачи ОАСУ	»
Периоды времени	»
Управленческая документация	»
Характеристика директивных документов и указаний ГУГКа	»
Причины невыполнения плана	»
Информационные массивы	»
Виды информации	»

документ, который включает сокращенные наименования отдельных позиций классификаторов, предназначенные для печати в выходных формах.

При построении классификаторов используются две наиболее распространенные системы классификации — иерархическая и многоаспектная.

Под и е р а р х и ч е с к о й понимается система классификации, между классификационными группировками которой установлено отношение подчинения. Классифицируемое множество объектов делится по некоторому выбранному признаку (основание деления) на крупные группировки, затем каждая группировка в соответствии с выбранным основанием деления разбивается на ряд последующих группировок, которые, в свою очередь, делятся на более мелкие. Такая система обеспечивает построение классификации в виде графа, вершиной которого является объект классификации, включающей все множество, дугами — признак классификации, узлами — группировки соответствующего уровня деления (рис. 81).

Исходное множество A на первом уровне делится на классификационные группировки (классы) $A_1, A_2, \dots, A_n$, которые на втором уровне, в свою очередь, делятся соответственно на группировки (подклассы) $A_{11}, A_{12}, A_{13}, \dots, A_{1n}$ и на третьем — на группы $A_{111}, A_{112}, A_{113}, \dots, A_{nmk}$.

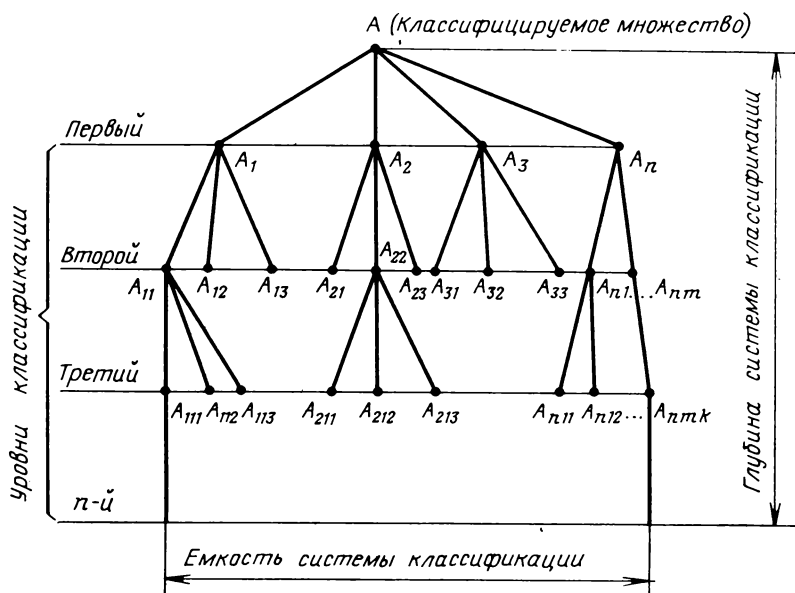


Рис. 81. Граф классифицируемого множества

Другой системой классификации является **многоаспектная**, которая строится по независимым признакам без подчинения их в рамках одной структуры и, следовательно, не имеет жесткой структуры и заранее построенных конечных группировок. Объединяемое в ней множество объектов делится многократно и независимо.

Разновидность многоаспектной системы классификации — **фасетный** способ описания и представления элементов классификации. При фасетной классификации задачи идентификации объектов и их классификации решаются отдельно. Независимые признаки получают самостоятельные коды и деление множества может проводиться по любому числу признаков. В результате запись приобретает гибкую блочную структуру.

С построением систем классификации тесно связан вопрос кодирования информации. В процессе кодирования объектам классификации и их группировкам по определенным правилам присваиваются цифровые, буквенные или буквенно-цифровые кодовые обозначения. В полном кодовом обозначении выделяются содержательная-адресная часть (классификационная), идентификационная (регистрационная) и контрольная части. В связи с этим код каждого объекта может быть построен как с помощью отдельной идентификационной и классификационной частей, так и идентификации объектов в пределах классификационных группировок.

Регистрационные системы отличаются методом кодирования: порядковым и серийно-порядковым.

Порядковый метод кодирования соответствует идентификации объектов в несистематизированном виде. При этом все множество объектов кодируется по порядку расположения.

Для серийно-порядкового метода характерна неглубокая классификация. При этом методе кодирования все множество по определенному признаку группируется в серии. Каждому элементу этого множества путем порядковой регистрации присваивается код. В каждой серии обычно предусматриваются резервные номера (на случай появления новых позиций), которые могут располагаться в любом месте серии.

Классификация системы отличается методом кодирования: последовательным и параллельным.

При последовательном методе значение показателя, записанного в виде цифры или буквы в определенном разряде кодового обозначения, зависит от значений предыдущих разрядов. В этом случае код последующей группировки образуется путем добавления кодов соответствующего количества разрядов к коду предыдущей (вышестоящей) группировки. Последовательный метод кодирования чаще всего используется в иерархической системе классификации, например в высших классификационных группировках общесоюзного классификатора промышленной и сельскохозяйственной продукции, в котором первые два показателя обозначают класс продукции, третий — подкласс, четвертый — группу, пятый — подгруппу, шестой — вид.

При параллельном методе кодирования значение характеристики объекта, выраженного каким-либо числом в определенных разрядах, не зависит от конкретного значения предыдущих разрядов кодовой комбинации. Такой метод используется в основном в фасетной системе классификации. Для обозначения каждого фасета отводится определенное число разрядов.

В качестве примера использования различных систем классификации и кодирования при проектировании обеспечивающей части ОАСУ ГУГКа в табл. 78 приведены основные характеристики некоторых классификаторов.

Т а б л и ц а 78

Классификатор	Значность кода	Метод кодирования	Метод кодирования свойств или показателей объектов	Система классификации
Картографическая продукция	10	Идентификация в пределах классификации	Комбинированная система (последовательное кодирование в сочетании с параллельным)	Комбинированная система (многоаспектная, иерархическая классификации в сочетании с фасетной)
Структурные подразделения ГУГКа	3	То же	Последовательное кодирование	Иерархическая классификация

Система документации включает в себя комплекс нормативно-методических материалов и взаимосвязанных документов, содержащих информацию, необходимую для решения функциональных задач АСУ.

При проектировании информационного обеспечения системы важное значение имеет унификация входных и выходных форм документов. Унификация форм документов заключается прежде всего в типизации состава показателей, выборе формата бланков, проектировании размеров строк и граф.

Унификация входных и выходных форм документов проводится на основе формуляров-образцов. Так, например, формы статистической отчетности состоят из трех частей, в пределах которых выделяется девять зон (рис. 82). Зоны 1—7 относятся к заголовочной части, 8 — к содержательной (предметной), а 9 — к оформительной.

Зона 1 отводится для кодов документа и отчитывающейся организации.

В зоне 2 записываются наименование и почтовый адрес получателя, а в зоне 3 — наименование, почтовый адрес, ведомственная подчиненность и отраслевая принадлежность отчитывающейся организации. В зоне 4 типографским способом печатается индекс формы документа.

В зоне 5 указываются способ представления документа и периодичность составления, а в зоне 6 — отчитывающиеся организации, в которые посылаются отчет и сроки представления. В зоне 7 указываются название формы отчета и период отчетности.

Зоны 1—7 размещаются только на лицевой стороне бланка документа.

Основной частью отчета является зона 8. Как правило, эта зона строится в табличной форме, показатели в которой размещаются в виде таблиц с графами по вертикали и строками по горизонтали, и содержит наименования, коды и числовые значения показателей. Структура таблицы должна предусматривать возможность контроля правильности заполнения, для чего вводятся контрольные суммы по строкам и графам. Зона 8 может располагаться на лицевой и обратной сторонах документа или на последующих его страницах.

Зона 9 является оформительской. Здесь указываются должности лиц, подписавших документ, их подписи и даты подписей.

2	1
3	4
	5
	6
7	
8	
9	

Рис. 82. Формуляр-образец для унификации форм статистической отчетности

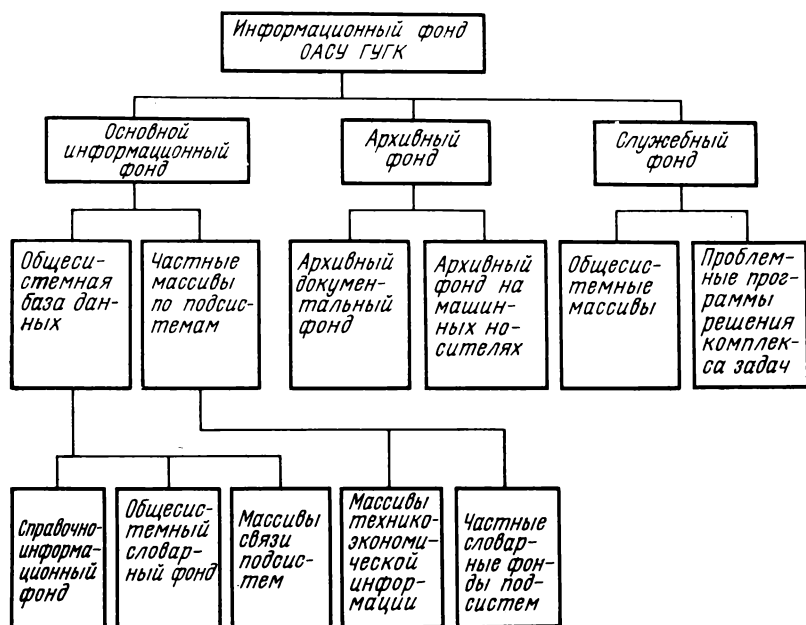


Рис. 83. Блок-схема построения информационного фонда ОАСУ ГУГКа

В состав системы документации ОАСУ входит 261 форма документов различных видов.

Функционирование АСУ связано с накоплением и хранением в вычислительном центре больших массивов информации: документальной информации и информации на машинных носителях. Совокупность информационных массивов (ИМ), хранящихся в памяти системы, представляет собой информационный фонд АСУ. Рассмотрим особенности построения информационного фонда на примере ОАСУ ГУГКа. Общая блок-схема его построения приведена на рис. 83.

В соответствии с назначением и способами обработки информации информационный фонд имеет три части: основной информационный, архивный и служебный фонды.

Основной информационный фонд включает в свой состав общесистемную базу данных и частные массивы по подсистемам.

Общесистемная база данных ОАСУ представляет собой совокупность справочно-информационного фонда (СИФ), общесистемного словарного фонда и массивов связи подсистем. Наличие указанных массивов позволяет обеспечить взаимоувязанное решение функциональных задач системы, а также выдачу аппарату ГУГКа справочной информации в оперативном режиме «запрос—ответ». В массивах справочно-информационного фонда хранятся основные показатели, характеризующие производственно-хозяйственную дея-

тельность ГУГКа и его подразделений за определенные периоды времени.

Массивы общесистемного словарного фонда содержат словари, используемые для распечатки текстов в выходных формах (табуляграммах).

В массивы связи подсистем входят промежуточные данные, которыми обмениваются подсистемы в процессе решения задач.

Частные массивы подсистемы содержат информацию, которая используется в решении задач одной подсистемы.

В состав архивного фонда входят: архив справочных и руководящих технических материалов; архив документальной информации; архив информации на машинных носителях.

Служебный фонд содержит информацию, которая предназначена для организации процесса переработки технико-экономической информации, циркулирующей в ОАСУ (программа решения задач, каталоги информационные массивов, стандартные программы, трансляторы и др.).

Перечисленные в табл. 76 функции информационного обеспечения АСУ позволяют сделать вывод о том, что основной задачей его является своевременное и полное обслуживание всех информационных запросов системы достоверными и точными данными.

При проектировании информационного обеспечения АСУ прежде всего определяют: состав информации, ее структуру, т. е. перечень отдельных единиц информации или информационных совокупностей (реквизиты, показатели, отчеты и т. д.); закономерность преобразования информации, т. е. правила построения показателей, документов, потоков информации, агрегирования и преобразования единиц от входа в систему и до выхода результатов; количественные оценки информации, т. е. характеристики движения потоков информации (объем, периодичность, маршруты движения документов, временные характеристики функционирования источников информации, сбора данных, продолжительности хранения и обновления данных); качественные оценки информации (полнота, достоверность, значимость, своевременность и т. д.); методики получения необходимой информации.

В состав обеспечивающей части АСУ входят также математическое, программное и техническое обеспечение.

Математическое обеспечение ОАСУ ГУГКа включает совокупность математических методов, моделей и алгоритмов для решения задач и обработки информации с применением ЭВМ. Среди основных математических методов, использованных для решения задач ОАСУ ГУГКа, можно выделить: линейное программирование, математическую логику и теории алгоритмизации, метод экспертных оценок, корреляционный, регрессионный и факторный анализы, имитационное моделирование, информационно-логические методы и методы прямого счета.

Программное обеспечение ОАСУ ГУГКа представляет собой совокупность программ и состоит из внутреннего и внешнего обеспечения.

Внутреннее программное обеспечение поставляется одновременно с комплексом технических средств и включает: операционную систему, рассчитанную на диски; трансляторы с алгоритмических языков; стандартные, управляющие, обслуживающие программы и блоки; комплекс программ к типовым задачам.

Внешнее программное обеспечение ОАСУ ГУГКа состоит из двух составных частей, общесистемной и программ решения конкретных задач.

Общесистемное программное обеспечение включает: универсальные системные программы ввода и обработки информации с перфоленты и перфокарт; пакеты прикладных программ для обработки словарной и выходной информации; универсальный пакет для решения задач линейного программирования; информационно-справочную систему.

Разработка внешнего программного обеспечения включает набор методов решения задач ОАСУ ГУГКа и составление программ.

Функционирование автоматизированных систем управления основано на использовании разнообразных средств вычислительной техники, обеспечивающих регистрацию, сбор, обработку и выдачу информации в процессе управления предприятиями, фабриками и отраслью в целом. Указанные вычислительные средства образуют техническое обеспечение или комплекс технических средств (КТС) любой АСУ.

КТС ОАСУ ГУГКа состоит из следующих функциональных групп: сбора, передачи и представления данных; подготовки и первичной обработки данных; средств ввода-вывода данных; средств обработки данных; средств оформления и размножения документов; средств связи.

В группу средств передачи данных входят вычислительные машины, телетайпы, телеграфные линии Министерства связи.

Группу средств подготовки данных образуют карточные и ленточные перфораторы.

Устройства ввода с перфокарт и перфолент объединяются в группу средств ввода данных. От устройств ввода с перфокарт и перфолент информация через арифметическое и оперативное запоминающее устройства записывается на магнитные ленты и диски.

В группу обработки данных входят электронно-вычислительные машины, точнее, их процессоры и внешние накопители.

§ 64. Опыт создания и перспективы развития ОАСУ ГУГКа

Работы по проектированию и внедрению первой очереди ОАСУ ГУГКа были завершены в 1976 г. В этих работах принимали участие специалисты ЦНИИГАиКа, НИИПГ и предприятий и аппарата ГУГКа.

В процессе функционирования системы аппарату ГУГКа регулярно выдаются месячные и квартальные отчеты о выполнении производственных планов топографо-геодезическими предприятиями, сводные отчеты о выполнении производственных планов по

ГУГКу, аналитические таблицы по себестоимости топографо-геодезических работ, сводные статистические отчеты по материально-техническому снабжению, сведения о наличии и движении основного оборудования по предприятиям, сводные данные о качественном и количественном составе кадров подразделений ГУГКа, учетно-аналитические таблицы, характеризующие деятельность отрасли в целом и ее организаций в области труда и заработной платы.

В результате решения задач на ЭВМ ежегодно для аппарата ГУГКа готовится информация по 46 формам, что позволяет улучшить качество принятых управленческих решений. В то же время в процессе функционирования первой очереди ОАСУ ГУГКа были вскрыты отдельные недостатки, которые обусловлены целым рядом объективных и субъективных факторов. Так, например, при проектировании первой очереди системы недостаточно использовался принцип системного подхода, не было обеспечено взаимодействие между ОАСУ ГУГКа и АСУП ТПП, что в значительной мере снижает эффективность их функционирования. В первой очереди ОАСУ ГУГК в основном решаются задачи учета и анализа производственной деятельности подразделений ГУГКа. Отсутствие укрупненных нормативов затрат ресурсов по видам работ не позволило решить задачи оптимального планирования топографо-геодезического производства. Большие затруднения в проектировании первой очереди системы возникли в связи с ориентацией ее программного обеспечения на ЭВМ второго поколения типа М-222.

В десятой пятилетке работы по созданию и развитию ОАСУ ГУГКа были продолжены. Для второй очереди системы характерно наличие отдельных информационных связей между подсистемами и комплексами задач, создание общесистемных информационных массивов. Значительно увеличен объем входной и выходной информации, прослеживается взаимосвязь между отдельными задачами ОАСУ ГУГКа и АСУ ТПП. Таким образом, при проектировании второй очереди ОАСУ ГУГКа осуществлен принцип системного подхода.

Вторая очередь ОАСУ ГУГКа ориентирована на новый комплекс технических средств на базе ЭВМ ЕС. Он характеризуется высокой производительностью ЭВМ, наличием дисковой памяти, совместимостью с основными моделями ЕС, развитым программным обеспечением.

В соответствии с принципом развития АСУ ее проектирование и внедрение осуществляется очередями. Главная цель развития ОАСУ ГУГКа — постоянное повышение ее экономической эффективности и научно-технического уровня. Последнее должно достигаться путем комплексной автоматизации управленческих функций, увеличения числа подсистем и комплексов задач, создания интегрированной системы обработки информации на базе банка данных, широкого применения экономико-математических моделей оптимизации и прогнозирования развития отрасли, создания мощных программных средств ОАСУ, совершенствования КТС, обеспечения четкого взаимодействия ОАСУ ГУГКа с системами верхнего и ниж-

него уровней, совершенствования технологии проектирования и эксплуатации системы. Таким образом, одним из главных условий дальнейшего развития ОАСУ ГУГКа должно быть комплексное решение всех вопросов, связанных с повышением эффективности и научно-технического уровня системы.

Изучение современных тенденций развития АСУ в нашей стране, а также определенный опыт создания первой и второй очередей ОАСУ ГУГКа позволили определить целый ряд взаимосвязанных направлений развития ОАСУ ГУГКа. Вся совокупность направлений развития системы объединена в три блока: развитие структуры ОАСУ и ее элементы (блок 1); обеспечение взаимодействия ОАСУ ГУГКа с системами верхнего и нижнего уровней (блок 2); совершенствование технологии создания ОАСУ (блок 3).

Основными задачами первого блока являются комплексная автоматизация управленческих функций, расширение и изменение характера автоматизации функций, переход от решения отдельных комплексов задач к системному, взаимосвязанному их решению.

Важность обеспечения взаимодействия ОАСУ ГУГКа с АСУ верхнего и нижнего уровней обуславливается тенденциями развития АСУ в нашей стране: созданием интегрированных автоматизированных систем управления отраслями, объединяющих АСУ и АСУП, и созданием общегосударственной автоматизированной системы (ОГАС). В связи с этим возникает задача обеспечения методологической, организационной, технической и информационной совместимости ОАСУ ГУГКа с АСУП и высшими звеньями ОГАС.

К основным задачам совершенствования технологии создания ОАСУ следует отнести: повышение качества проектирования, сокращение сроков разработки и внедрения ОАСУ, обеспечение аппарата ГУГКа своевременной полной и достоверной информацией при функционировании системы.

Существующая организационная структура, методы и формы деятельности аппарата управления оказывают большое влияние на проектирование ОАСУ. В то же время разработка ОАСУ способствует созданию объективных предпосылок для совершенствования организационной структуры и методов управления отраслью.

Важное значение при дальнейшем развитии ОАСУ приобретает широкое применение экономико-математических методов и технически обоснованной системы норм и нормативов затрат трудовых, материальных и финансовых ресурсов по видам и физико-географическим районам работ.

В заключение следует отметить, что рассмотренные основные направления дальнейшего развития ОАСУ ГУГКа несут в себе лишь принципиальную возможность повышения экономической эффективности и научно-технического уровня системы. Чтобы практически реализовать эту возможность, необходимо выполнить целевую комплексную программу по решению этой важной научно-технической проблемы.

СПИСОК ЛИТЕРАТУРЫ

1. *Акоф Р., Сасиени М.* Основы исследования операций. М., Мир, 1971.
2. *Беллман Р., Дрейфус С.* Прикладные задачи динамического программирования. М., Наука, 1965.
3. *Брыкин П. А., Кимельман С. А.* Математическое программирование в планировании геодезических и топографических работ. М., Недра, 1972.
4. *Вагнер Г.* Основы исследования операций. Т. I—III. М., Мир, 1972, 1973.
5. *Венецкий И. Г., Венецкая В. И.* Основные математико-статистические понятия и формулы в экономическом анализе. М., Статистика, 1974.
6. *Вентцель Е. С.* Исследование операций. М., Советское радио, 1972.
7. *Данциг Дж. Б.* Линейное программирование, его обобщения и применение. Пер. с англ. М., Прогресс, 1966.
8. *Дражнюк А. А.* Исследование вопросов комплексного управления качеством топографо-геодезических работ. Дисс. на соиск. уч. степени канд. техн. наук. М., МИИГАиК, 1980.
9. *Дружинин В. В., Конторов Д. С.* Проблемы систематологии. (Проблемы теории сложных систем). М., Советское радио, 1976.
10. *Зайченко Ю. П.* Исследование операций. Киев, Вища школа, 1975.
11. *Земцев А. С.* За высокое качество работ. — Геодезия и картография, 1974, № 12, с. 1—6.
12. *Кашин Л. А.* Совершенствовать управление качеством работ. — Геодезия и картография, 1978, № 3, с. 1—7.
13. *Кашин Л. А.* Качество топографо-геодезических и картографических работ на уровень требований задач десятой пятилетки. — Геодезия и картография, 1976, № 3, с. 1—9.
14. *Кашин Л. А.* Задачи по улучшению планирования и совершенствования хозяйственного механизма в отрасли. — Геодезия и картография, 1980, № 1, с. 1—8.
15. *Каширникова Р. П.* Анализ основных технико-экономических показателей развития топографо-геодезического производства. Дисс. на соиск. уч. степени канд. техн. наук. М., ЦНИИГАиК, 1979.
16. *Коршунов Ю. М.* Математические основы кибернетики. М., Энергия, 1972.
17. *Крынский Х. Э.* Математика для экономистов. М., Статистика, 1970.
18. *Кутузов И. А.* Полнее использовать резервы производства. — Геодезия и картография, 1979, № 4, с. 1—12.
19. *Математика и кибернетика в экономике: Словарь-справочник* — 2-е изд. перераб. и доп. М., Экономика, 1975.
20. *Фон Нейман Дж., Моргенштерн О.* Теория игр и экономическое поведение. М., Наука, 1970.
21. *Параубек Г. Э.* Сетевое планирование и управление. М., Экономика, 1967.
22. *Райзберг Б. А., Голубков Е. П., Пекарский Л. С.* Системный подход в перспективном планировании. М., Экономика, 1975.
23. *Семенов В. И.* Автоматизация процессов оперативного управления топографо-геодезическим предприятием. Дисс. на соиск. уч. степени канд. техн. наук. М., МИИГАиК, 1976.
24. *Сигалов В. М.* Опыт комплексного управления качеством продукции. — Геодезия и картография, 1976, № 5, с. 1—6.
25. *Синдеев А. А.* Моделирование и автоматизация разработки опти-

мальных вариантов развития топографо-геодезического производства. Дисс. на соиск. уч. степени канд. техн. наук. М., ЦНИИГАиК, 1980.

26. *Системный анализ и структуры управления*. Под общ. ред. проф. В. Г. Шорина, серия Наука и управление, кн. 8. М., Знание, 1975.

27. *Чернов Г., Мозес Л. Элементарная теория статистических решений*. Пер. с англ. М., Наука, 1968.

28. *Черняк Ю. И. Системный анализ в управлении экономикой*. М., Экономика, 1975.

29. *Хедли Дж. Нелинейное и динамическое программирование*. Пер. с англ. М., Мир, 1967.

30. *Ху Т. Целочисленное программирование и потоки в сетях*. Пер. с англ., М., Мир, 1974.

31. *Юдин Д. Б., Гольштейн Е. Г. Линейное программирование*. М., Наука, 1968.

ОГЛАВЛЕНИЕ

Предисловие	3
Глава I. Элементы системного анализа	9
§ 1. Понятие системы, ее структуры, виды систем	9
§ 2. Системный подход в современной науке	16
§ 3. Основные логические элементы системного анализа	19
§ 4. Системный подход к анализу основных технико-экономических показателей развития топографо-геодезического производства	24
§ 5. Реализация системного подхода при исследовании вопросов комплексного управления качеством топографо-геодезических работ	33
Глава II. Управление, принятие решений	39
§ 6. Основные черты процесса управления	39
§ 7. Системный подход к анализу процесса управления	40
§ 8. Своеобразие проблем, требующих управленческого решения	42
§ 9. Системный подход к обоснованию решений	43
§ 10. Моделирование как метод научного познания	47
§ 11. Сущность процесса моделирования	49
§ 12. Этапы экономико-математического моделирования	51
§ 13. Информационное обеспечение процесса управления	56
§ 14. Информационная модель оперативного управления аэрогеодезическими предприятиями (АГП)	60
Глава III. Введение в теорию принятия решений	63
§ 15. Развитие процесса принятия и реализации решения	63
§ 16. Основные понятия теории принятия решений	68
§ 17. Классификация задач принятия решений	73
§ 18. Общая постановка однокритериальной задачи принятия решений	77
§ 19. Принятие решений в условиях определенности. Общая постановка однокритериальной статической детерминированной задачи принятия решений	79
§ 20. Пример обоснования решения однокритериальной статической детерминированной задачи принятия решений	80
§ 21. Обзор методов решения однокритериальных статических детерминированных задач принятия решений	83
Глава IV. Классические задачи математического программирования	90
§ 22. Задачи отыскания безусловного экстремума	90
§ 23. Функция многих переменных	96
§ 24. Условный экстремум. Метод множителей Лагранжа	100
§ 25. Интерпретация множителей Лагранжа	109
§ 26. Седловая точка функции Лагранжа. Минимаксная трактовка задачи на условный экстремум функции	112
Глава V. Нелинейное программирование	115
§ 27. Общие понятия	115
§ 28. Задача нелинейного программирования при ограничениях неотрицательности	121
§ 29. Условия Куна — Таккера	123
§ 30. Теорема Куна — Таккера	127
§ 31. Геометрический способ решения задач нелинейного программирования	130
Глава VI. Линейное программирование	134
§ 32. О постановках задач линейного программирования	134
§ 33. Геометрическая интерпретация задач линейного программирования	141
§ 34. Симплекс-метод решения задач линейного программирования	144
§ 35. Двойственные задачи линейного программирования	156

Глава VII. Транспортная задача линейного программирования	171
§ 36. Математическая постановка задачи	171
§ 37. Двойственная задача. Условия оптимальности	173
§ 38. Способы построения опорного плана	177
§ 39. Метод потенциалов	181
§ 40. Случай вырождения	187
Глава VIII. Целочисленное программирование	190
§ 41. Особенности дискретных экстремальных задач	190
§ 42. Задачи целочисленного программирования	192
§ 43. Общая характеристика методов решения целочисленных задач	197
§ 44. Точные методы целочисленного программирования	198
Глава IX. Динамическое программирование	209
§ 45. Сущность и основные свойства	209
§ 46. Задача по оптимизации распределения ресурсов	213
Глава X. Сетевое планирование и управление	222
§ 47. Сущность сетевых методов планирования и управления производством	222
§ 48. Решение транспортной задачи на основе сетевых графиков	236
Глава XI. Принятие решений в условиях неопределенности	240
§ 49. Классификация задач принятия решений в условиях неопределенности и обзор методов их решения	240
§ 50. Принятие решений в условиях повторяющейся одноуровневой конфликтной ситуации. Элементы теории стратегических игр	244
§ 51. Принятие решений в условиях действия неопределенных факторов стохастической природы. Элементы теории статистических игр	254
§ 52. Принятие решений в условиях стохастической неопределенности в случае, когда проведение экспериментов невозможно. Статистические игры без эксперимента	256
§ 53. Принятие решений в условиях стохастической неопределенности с использованием экспериментов. Статистические игры с экспериментами	261
Глава XII. Корреляционный и регрессионный анализ в экономике топографо-геодезического производства	266
§ 54. Сущность корреляционного и регрессионного анализа	266
§ 55. Двухфакторная модель прогнозирования производительности труда	269
§ 56. Взаимосвязь качества труда исполнителей с производительностью их труда	275
§ 57. Минимум затрат как один из критериев определения оптимальных размеров топографо-геодезического подразделения	279
§ 58. Метод прогнозирования ожидаемого уровня производительности труда по многофакторной регрессионной модели	282
§ 59. Методы учета влияния отраслевых факторов на показатели эффективности производства	288
Глава XIII. Автоматизированные системы управления в топографо-геодезическом производстве	294
§ 60. Сущность и основные предпосылки создания автоматизированных систем управления на современном этапе	294
§ 61. Структура, составные части и основные принципы построения АСУ	297
§ 62. Организационные основы проектирования и очередность создания АСУ	301
§ 63. Функциональная и обеспечивающая части отраслевой автоматизированной системы управления ГУГКа	304
§ 64. Опыт создания и перспективы развития ОАСУ ГУГКа	314
Список литературы	317

НА.БЕОГРАДОВ.:АМ.ГОЛУБЦОВ.:А.А.СИНДЕНБ