МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Пензенский государственный университет» (ПГУ)

Н. Ф. Добрынина, Д. В. Тарасов

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

Учебное пособие

Пенза
Издательство ПГУ
2017

УДК 519.8
Д57

Р е ц е н з е н т ы:

доктор технических наук, профессор,
заведующий кафедрой математики и математического моделирования
Пензенского государственного университета архитектуры и строительства
А. М. Данилов;

доктор физико-математических наук, профессор,
заведующий кафедрой математики и суперкомпьютерного моделирования
Пензенского государственного университета
Ю. Г. Смирнов

Добрынина, Н. Ф.
Д57 Математическое моделирование экономических процессов : учеб. пособие / Н. Ф. Добрынина, Д. В. Тарасов. – Пенза : Изд-во ПГУ, 2017. – 104 с.

ISBN 978-5-906975-02-7

Рассмотрены два раздела: модели линейного программирования и нелинейного программирования. Реализуется основная идея изучения курса высшей математики в применении к экономике – идея математического моделирования экономических процессов.

Издание написано в соответствии с требованиями государственных общеобразовательных стандартов в области математики.

Подготовлено на кафедре «Высшая и прикладная математика» Пензенского государственного университета и предназначено для обучающихся по направлению подготовки 01.03.04 «Математическое моделирование в экономике и технике», а также может быть использовано для других специальностей при изучении высшей математики и экономики.

УДК 519.8

ISBN 978-5-906975-02-7

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Исследование операций – научная дисциплина, занимающаяся разработкой и практическим применением методов наиболее эффективного управления различными организационными системами.

Управление любой системой реализуется как процесс, подчиняющийся определенным закономерностям. Их знание помогает определить условия, необходимые и достаточные для осуществления данного процесса. Цель исследования операций – *качественное обоснование принимаемых решений по организации управления.*

При решении конкретной задачи управления применение методов исследования операций предполагает:

- построение экономических и математических моделей для задач принятия решения в сложных ситуациях или в условии неопределенности;
- изучение взаимосвязей, определяющих принятие решений и установление критериев эффективности, позволяющих оценивать преимущество того или иного варианта действия.

Отметим основные понятия и определения исследования операций.

Операция – любое управляемое мероприятие, направленное на достижение цели. Результат операции зависит от способа ее проведения, организации, от выбора некоторых параметров.

Определенный выбор параметров называется *решением. Оптимальными* считаются те решения, которые по тем или иным соображениям предпочтительнее других. *Основной задачей исследования операций является предварительное количественное обоснование оптимальных решений.*

ОБЩИЕ МЕТОДИЧЕСКИЕ УКАЗАНИЯ

Модель и эффективность операций. Для применения количественных методов исследования требуется построить *математическую модель операции.* При построении модели операция упрощается, схематизируется и описывается с помощью математического аппарата. *Модель операции* – это достаточно точное описание операции с помощью математического аппарата (различного рода функций, уравнений, систем уравнений, неравенств и т.п.).

Эффективность операции – степень ее приспособленности к выполнению задачи. Эффективность операции количественно выражается в виде критерия эффективности – целевой функции.

Общая постановка задачи исследования операций. Важно усвоить методологию построения моделей задач исследования операций. Все факторы, входящие в описание операции, делятся на две группы:

- *постоянные факторы*, на которые мы влиять не можем, обозначаются $\alpha_1, \alpha_2, \ldots$;
- *зависимые факторы* (элементы решения) $x_1, x_2, \ldots$

Критерий эффективности, выражаемый некоторой функцией, называемой *целевой*, зависит от факторов обеих групп, поэтому целевую функцию Z можно записать в виде

$$Z = f(x_1, x_2, \ldots, \alpha_1, \alpha_2, \ldots).$$

Большой класс составляют *оптимизационные модели.* Такие задачи возникают при попытке оптимизировать планирование и управление сложными системами. Оптимизационную задачу можно сформулировать в следующем виде: *найти переменные* $x_1, x_2, \ldots, x_n$, *удовлетворяющие системе неравенств (уравнений)*

$$\varphi_i(x_1, x_2, \ldots, x_n) \le b_i, \quad i = 1, 2, \ldots, m, \qquad (0.1)$$

и обращающие в максимум (минимум) целевую функцию, т.е.

$$Z = f(x_1, x_2, \ldots, x_n) \to \max(\min). \qquad (0.2)$$

Точку n-мерного пространства будем обозначать $X = (x_1, x_2, \ldots, x_n)$, а оптимальное решение $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$.

Важное значение в экономическом анализе имеет *классическая задача потребления.*

Пусть имеется n видов товаров и услуг, количества которых $x_1, x_2, \ldots, x_n$ по ценам соответственно $p_1, p_2, \ldots, p_n$ за единицу. Суммарная стоимость этих товаров и услуг составляет $\sum_{i=1}^{n} p_i x_i$.

Уровень потребления Z может быть выражен некоторой функцией $Z = f(x_1, x_2, \ldots, x_n)$, называемой *функцией полезности. Задача потребления состоит в том, чтобы найти такой набор товаров и услуг* $x_1, x_2, \ldots, x_n$ *при данной величине доходов* I*, чтобы обеспечить максимальный уровень потребления, т.е.*

$$Z = f(x_1, x_2, \ldots, x_n) \to \max \tag{0.3}$$

при условии

$$\sum_{i=1}^{n} p_i x_i \le I, \tag{0.4}$$

$$x_i \ge 0, \; i = 1, 2, \ldots, n. \tag{0.5}$$

Решения этой задачи, зависящие от цен $p_1, p_2, \ldots, p_n$ и величины дохода I, называются *функциями спроса.*

Очевидно, что рассмотренная задача потребления (0.3)–(0.5) является частным случаем более общей задачи (0.1), (0.2) на определение экстремума функции n переменных при некоторых ограничениях, т.е. задачей на *условный экстремум.*

В тех случаях, когда функции f и φ_i в задаче (0.1), (0.2) хотя бы дважды дифференцируемы, можно применять классические методы оптимизации. Однако применение этих методов для функций многих переменных трудоемко. Классические методы совсем не работают, если множество допустимых значений аргумента дискретно или функция Z задана таблично. В этих случаях для решения задачи (0.1), (0.2) применяются методы *математического программирования.*

Если критерий эффективности $Z = f(x_1, x_2, \ldots, x_n)$ представляет линейную функцию, а функции $\varphi_i(x_1, x_2, \ldots, x_n)$ в системе ограничений (0.1) также линейны, то такая задача является задачей *линейного программирования.* Если, исходя из смысла задачи, ее решения должны быть целыми числами, то эта задача *целочисленного линейного программирования.* Если критерий эффективности или система ограничений (или оба вместе) задаются нелинейными функциями, то получаем задачу *нелинейного программирования.* В частности, если указанные функции обладают свойствами выпуклости, то задача является задачей *выпуклого программирования.*

Если в задаче математического программирования имеется переменная времени и критерий эффективности содержит уравнения, описывающие протекание операций во времени, то такая задача является задачей *динамического программирования.*

Если критерий эффективности (0.2) и система ограничений (0.1) задаются функциями вида $cx_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$,то имеем задачу *геометрического программирования.* Если функции f и (или) φ_i зависят от параметров, то получаем задачу *параметрического программирования*; если эти функции носят случайный характер – задачу *стохастического программирования* .Если точный оптимум найти алгоритмическим путем невозможно из-за большого числа вариантов решения, то прибегают к методам *эвристического программирования.* Оно позволяет существенно сократить число вариантов и найти достаточно хорошее решение.

По своей содержательной постановке множество других, типичных задач исследования операций может быть разбито на ряд классов.

Задачи сетевого планирования и управления рассматривают соотношения между сроками окончания комплекса операций (работ) и моментами начала всех операций комплекса. Эти задачи состоят в нахождении минимальных продолжительностей комплекса операций, оптимального соотношения величин стоимости и сроков их выполнения.

Задачи массового обслуживания посвящены изучению и анализу систем обслуживания с очередями заявок или требований и состоят в определении показателей эффективности работы систем, в определении числа каналов обслуживания, времени обслуживания.

Задачи управления запасами состоят в отыскании оптимальных значений уровня запасов (точки заказа) и размера заказа.

Задачи распределения ресурсов возникают при определенном наборе операций, которые необходимо выполнять при ограниченных наличных ресурсах и требуется найти оптимальные распределения ресурсов между операциями.

Задачи ремонта и замены оборудования возникают в связи с износом и старением оборудования и необходимостью его замены с течением времени. Задачи сводятся к определению оптимальных сроков, числа профилактических ремонтов и проверок, а также моментов замены оборудования модернизированным.

Задачи составления расписания (календарного планирования) состоят в определении оптимальной очередности выполнения операций на различных видах оборудования.

Задачи планировки и размещения состоят в определении оптимального числа и места размещения новых объектов с учетом их взаимодействия с существующими объектами и между собой.

Задачи выбора маршрута, или *сетевые задачи*, встречаются при исследовании разнообразных задач на транспорте и состоят в определении наиболее экономичных маршрутов.

Среди моделей исследования операций особо выделяются модели принятия оптимальных решений в конфликтных ситуациях. Они изучаются *теорией игр.* В конфликтных ситуациях сталкиваются интересы двух (или более) сторон, преследующих разные цели. В задачах теории игр необходимо выработать рекомендации по разумному поведению участников конфликта, определить их оптимальные стратегии.

В большинстве случаев успех операции оценивается не по одному, а по нескольким критериям, один из которых следует максимизировать, другие – минимизировать. Так возникли *многокритериальные задачи исследования операций.* Для того, чтобы из множества критериев выбрать целевую функцию, необходимо установить *приоритет* критериев. Обозначим $f_1(x), f_2(x), \ldots, f_n(x)$ набор целевых функций, расположенный в порядке убывания приоритетов. В зависимости от определенных условий возможны два варианта:

– в качестве целевой функции выбирается критерий $f_1(x)$, обладающий наиболее высоким приоритетом;

– рассматривается комбинация

$$f(x) = \omega_1 f_1(x) + \omega_2 f_2(x) + \ldots + \omega_n f_n(x), \qquad (0.6)$$

где $\omega_1, \omega_2, \ldots, \omega_n$ – некоторые коэффициенты (веса).

В создание современного математического аппарата и развитие многих направлений исследования операций большой вклад внесли российские ученые Л. В. Канторович, Н. П. Бусленко, Е. С. Вентцель, Н. Н. Моисеев, Д. Б. Юдин и многие другие.

Значительный вклад в формирование и развитие исследования операций внесли зарубежные ученые Р. Акоф, Р. Беллман, Г. Данциг, Г. Кун, Дж. Нейман, Т. Саати, Р. Черчмен, А. Кофман и др.

РАЗДЕЛ 1

МОДЕЛИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ И ЕГО ПРИЛОЖЕНИЯ

Глава 1

ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

1.1. Экономико-математическая модель

Моделью будем называть *условный образ какого-либо объекта, приближенно воссоздающий этот объект с помощью некоторого языка.* В экономико-математических моделях таким объектом является экономический процесс (например, использование ресурсов, распределение изделий между различными типами оборудования и т.п.), а языком – классические или специально разработанные методы.

Экономико-математическая модель – *математическое описание исследуемого экономического процесса или объекта.* Эта модель выражает закономерности экономического процесса в абстрактном виде с помощью математических соотношений. Использование математического моделирования в экономике позволяет углубить количественный экономический анализ, расширить область экономической информации, интенсифицировать экономические расчеты.

Можно выделить три основных этапа проведения экономико-математического моделирования. На первом этапе ставятся цели и задачи исследования, проводится качественное описание объекта в виде экономической модели. На втором этапе формируется математическая модель изучаемого объекта, осуществляется выбор методов исследования, проводится программирование модели на ЭВМ, подготавливается исходная информация. Далее проверяются пригодность машинной модели на основании правильности получаемых с ее помощью результатов и оценка их устойчивости. На третьем этапе экономико-математического моделирования осуществляются анализ математической модели, проведение машинных расчетов, обработка и анализ полученных результатов.

Процедура экономико-математического моделирования заменяет дорогостоящие и трудоемкие натуральные эксперименты. Достаточно быстро и дешево производится сравнение многочисленных вариантов планов и управленческих решений. В результате отбираются наиболее оптимальные варианты.

Рассмотрим примеры экономико-математических моделей.

1.2. Примеры задач линейного программирования

1. Задача об использовании ресурсов (задача планирования производства).

1.1. Для изготовления двух видов продукции P_1 и P_2 используют четыре вида ресурсов S_1, S_2, S_3 и S_4. Запасы ресурсов, число единиц ресурсов, затрачиваемых на изготовление единицы продукции, приведены в табл. 1.1 (цифры условные).

Таблица 1.1

Вид ресурса	Запас ресурса	Число единиц ресурсов, затрачиваемых на изготовление единицы продукции	
		P_1	P_2
S_1	18	1	3
S_2	16	2	1
S_3	5	–	1
S_1	21	3	–

Прибыль, получаемая от единицы продукции P_1 и P_2 – соответственно 2 и 3 руб.

Необходимо составить такой план производства продукции, при котором прибыль от ее реализации будет максимальной.

Решение. Составим экономико-математическую модель задачи.

Обозначим x_1, x_2 – число единиц продукции соответственно P_1 и P_2, запланированных к производству. Для их изготовления потребуется $(1x_1 + 3x_2)$ единиц ресурса S_1, $(2x_1 + 1x_2)$ единиц ресурса S_2, $(1x_2)$ единиц ресурса S_3 и $(3x_1)$ единиц ресурса S_4. Поскольку потребление ресурсов S_1, S_2, S_3 и S_4 не должно превышать их запасов, соответственно 18, 16, 5 и 21 единицу, то связь между потреблением ресурсов и их запасами выражается системой неравенств:

$$\begin{cases} x_1 + 3x_2 \le 18, \\ 2x_1 + x_2 \le 16, \\ x_2 \le 5, \\ 3x_1 \le 21. \end{cases} \qquad (1.1)$$

По смыслу задачи переменные

$$x_1 \ge 0, \quad x_2 \ge 0. \qquad (1.2)$$

Суммарная прибыль F составит $2x_1$ руб. от реализации продукции P_1 и $3x_2$ руб. – от реализации продукции P_2, т.е.

$$F = 2x_1 + 3x_2. \qquad (1.3)$$

Экономико-математическую модель задачи можно сформулировать так: *найти такой план выпуска продукции* $X = (x_1, x_2)$, *удовлетворяющий системе (1.1) и условию (1.2), при котором функция (1.3) принимает максимальное значение.*

Задачу можно легко обобщить на случай выпуска n видов продукции с использованием m видов ресурсов.

Обозначим $x_j (j = 1, 2, \ldots, n)$ – число единиц продукции P_j, запланированной к производству; $b_i (i = 1, 2, \ldots, m)$ – запас ресурса S_i, a_{ij} – число единиц ресурса S_i, затрачиваемого на изготовление единицы продукции P_j (числа a_{ij} называют *технологическими коэффициентами*); c_j – прибыль от реализации единицы продукции P_j.

Экономико-математическая модель задачи об использовании ресурсов в общей постановке имеет вид: *найти такой план* $X = (x_1, x_2, \ldots, x_n)$ *выпуска продукции, удовлетворяющей системе*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m \end{cases} \qquad (1.4)$$

и условию

$$x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0, \qquad (1.5)$$

при котором функция

$$F = c_1x_1 + c_2x_2 + \ldots + c_nx_n \qquad (1.6)$$

принимает максимальное значение.

2. Задача составления рациона (задача о диете, задача о смесях).

1.2. Имеется два вида корма 1 и 2, содержащие питательные вещества (витамины) S_1, S_2 и S_3. Содержание числа единиц питательных веществ в 1кг каждого вида корма и необходимый минимум питательных веществ приведены в табл. 1.2.

Таблица 1.2

Питательное вещество (витамин)	Необходимый минимум питательных веществ	Число единиц питательных веществ в 1 кг корма	
		1	2
S_1	9	3	1
S_2	8	1	2
S_3	12	1	6

Стоимость 1кг кормов 1 и 2 равны соответственно 4 и 6 руб.

Необходимо составить дневной рацион, имеющий минимальную стоимость, в котором содержание каждого вида питательных веществ было бы не менее установленного предела.

Решение. Составим экономико-математическую модель задачи.

Обозначим x_1, x_2 – количество кормов 1 и 2, входящих в дневной рацион. Тогда этот рацион будет включать $(3x_1 + 1x_2)$ единиц питательного вещества S_1, $(1x_1 + 2x_2)$ единиц вещества S_2 и $(1x_1 + 6x_2)$ единиц питательного вещества S_3. Так как содержание питательных веществ S_1, S_2 и S_3 в рационе должно быть не менее 9, 8 и 12 единиц, то получим систему неравенств

$$\begin{cases} 3x_1 + x_2 \geq 9, \\ x_1 + 2x_2 \geq 8, \\ x_1 + 6x_2 \geq 12. \end{cases} \qquad (1.7)$$

Кроме того, переменные

$$x_1 \geq 0, \quad x_2 \geq 0. \qquad (1.8)$$

Общая стоимость рациона составит, в рублях,

$$F = 4x_1 + 6x_2. \qquad (1.9)$$

Экономико-математическая модель задачи: *составить дневной рацион* $X=(x_1,x_2)$, *удовлетворяющий системе (1.7) и условию (1.8), при котором функция (1.9) принимает минимальное значение.*

Для формулировки задачи в общей постановке обозначим: $x_j\,(j=1,2,\ldots,n)$ – число единиц корма n-го вида; $b_i(i=1,2,\ldots,m)$ – необходимый минимум содержания в рационе питательного вещества S_i; a_{ij} – число единиц питательного вещества S_i в единице корма j-го вида; c_j – стоимость единицы корма j-го вида.

Экономико-математическая модель задачи примет вид: *найти такой рацион* $X=(x_1,x_2,\ldots,x_n)$, *удовлетворяющий системе*

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \ge b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \ge b_2, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \ge b_m \end{cases} \tag{1.10}$$

и условиям

$$x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0, \tag{1.11}$$

при котором функция

$$F=c_1x_1+c_2x_2+\ldots+c_nx_n \tag{1.12}$$

принимает максимальное значение.

3. Задача об использовании мощностей (задача о загрузке оборудования).

Предприятию задан план производства продукции по времени и по номенклатуре: требуется за время T выпустить $n_1,n_2,\ldots,n_k$ единиц продукции $P_1,P_2,\ldots,P_k$. Продукция производится на станках $S_1,S_2,\ldots,S_m$. Для каждого станка известны производительность a_{ij} (число единиц продукции P_j, которое можно произвести на станке S_i) и затраты b_{ij} на изготовление продукции P_j на станке S_i в единицу времени.

Необходимо составить такой план работы станков (так распределить выпуск продукции между станками), чтобы затраты на производство всей продукции были минимальными.

Составим экономико-математическую модель задачи.

Обозначим x_{ij} – время, в течение которого станок S_i будет занят изготовлением продукции P_j $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, k)$.

Время работы каждого станка ограничено и не превышает T, следовательно, справедливы неравенства

$$\begin{cases} x_{11} + x_{12} + \ldots + x_{1k} \le T, \\ x_{21} + x_{22} + \ldots + x_{2k} \le T, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ x_{m1} + x_{m2} + \ldots + x_{mk} \le T. \end{cases} \tag{1.13}$$

Для выполнения плана выпуска по номенклатуре необходимо, чтобы выполнялись следующие равенства:

$$\begin{cases} a_{11}x_{11} + a_{21}x_{21} + \ldots + a_{m1}x_{m1} = n_1, \\ a_{12}x_{12} + a_{22}x_{22} + \ldots + a_{m2}x_{m2} = n_2, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{1k}x_{1k} + a_{2k}x_{2k} + \ldots + a_{mk}x_{mk} = n_k. \end{cases} \tag{1.14}$$

Кроме того,

$$x_{ij} \ge 0 \; (i = 1, 2, \ldots, m, \; j = 1, 2, \ldots, k). \tag{1.15}$$

Затраты на производство всей продукции выразятся функцией

$$F = b_{11}x_{11} + b_{12}x_{12} + \ldots + b_{mk}x_{mk}. \tag{1.16}$$

Экономико-математическая модель задачи об использовании мощностей примет вид: *найти такое решение* $X = (x_{11}, x_{12}, \ldots, x_{mk})$, *удовлетворяющее системам (1.13) и (1.14) и условию (1.15), при котором функция (1.16) принимает минимальное значение.*

1.3. Общая задача линейного программирования

Рассмотренные примеры задач линейного программирования позволяют сформулировать ***общую задачу линейного программирования.***

Дана система m *линейных уравнений и неравенств с* n *переменными:*

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n \le b_k, \\
a_{k+1,1}x_1 + a_{k+1,2}x_2 + \ldots + a_{k+1,n}x_n = b_{k+1}, \\
a_{k+2,1}x_1 + a_{k+2,2}x_2 + \ldots + a_{k+2,n}x_n = b_{k+2}, \\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
\end{cases} \tag{1.17}
$$

и линейная функция

$$F = c_1x_1 + c_2x_2 + \ldots + c_nx_n. \tag{1.18}$$

Необходимо найти такое решение системы $X = (x_1, x_2, \ldots, x_j, \ldots, x_n)$,

где

$$x_j \ge 0 \; (j = 1, 2, \ldots, l; l \le n), \tag{1.19}$$

при котором линейная функция F *(1.18) принимает оптимальное значение.*

Система (1.17) называется *системой ограничений*, а функция F – *линейной функцией, линейной формой, целевой функцией, или функцией цели.*

Кратко общую задачу линейного программирования можно представить в виде

$$F = \sum_{j=1}^{n} c_j x_j \to \max(\to \min)$$

при ограничениях:

$$
\begin{cases}
\sum\limits_{j=1}^{n} a_{ij}x_j \le b_i \; (i = 1, 2, \ldots, k), \\
\sum\limits_{j=1}^{n} a_{ij}x_j = b_i \; (i = k+1, k+2, \ldots, m),
\end{cases}
$$

$$x_j \ge 0 \; (j = 1, 2, \ldots, l; \; l \le n).$$

Оптимальным решением (оптимальным планом) задачи линейного программирования называется решение $X = (x_1, x_2, \ldots, x_j, \ldots, x_n)$

системы ограничений (1.17), удовлетворяющее условию (1.20), при котором линейная функция (1.19) принимает оптимальное (максимальное или минимальное) значение.

Термины «решение» и «план» – синонимы. Первый используется, когда речь идет о формальной стороне задачи (ее математическом решении), второй – о содержательной стороне (экономической интерпретации).

Если все переменные неотрицательны $(x_j \geq 0)$, то такая задача линейного программирования называется *стандартной*; если система ограничений состоит из одних уравнений, то задача называется *канонической.* В некоторых работах по математическому программированию стандартную задачу называют *симметричной,* а каноническую – *основной.*

Любая задача линейного программирования может быть сведена к канонической, стандартной или общей. Рассмотрим вспомогательную теорему без доказательства.

Теорема 1.1. *Всякому решению* $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *неравенства*

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \ldots + \alpha_{in}x_n \leq b_i \tag{1.20}$$

соответствует определенное решение $(\alpha_1, \alpha_2, \ldots, \alpha_n; \alpha_{n+i})$ *уравнения*

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \ldots + \alpha_{in}x_n + x_{n+i} = b_i, \tag{1.21}$$

в котором

$$x_{n+i} \geq 0, \tag{1.22}$$

и, наоборот, каждому решению $(\alpha_1, \alpha_2, \ldots, \alpha_n; \alpha_{n+i})$ *уравнения (1.21) и неравенства (1.22) соответствует определенное решение* $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *неравенства (1.20).*

Используя эту теорему, представим в качестве примера стандартную задачу (1.4)–(1.6) в каноническом виде. Для этого в каждое из m неравенств системы ограничений (1.4) введем дополнительные неотрицательные переменные $x_{n+1}, x_{n+2}, \ldots, x_{n+m}$, и система ограничений (1.4) примет вид

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + x_{n+1} = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \quad + x_{n+2} = b_2, \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \qquad + x_{n+m} = b_m. \end{cases} \tag{1.23}$$

Стандартная задача (1.4)–(1.6) в канонической форме выглядит так: *найти такое решение* $X = (x_1, x_2, \ldots, x_n, x_{n+1}, \ldots, x_{n+m})$, *удовлетворяющее системе (1.23) и условию (1.5), при котором функция (1.6) принимает максимальное значение.*

Глава 2

ЭЛЕМЕНТЫ ЛИНЕЙНОЙ АЛГЕБРЫ И ГЕОМЕТРИИ ВЫПУКЛЫХ МНОЖЕСТВ

2.1. Система m линейных уравнений с n переменными

Система m линейных уравнений с n переменными имеет вид

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1j}x_j + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2j}x_j + \ldots + a_{2n}x_n = b_2, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{ij}x_j + \ldots + a_{in}x_n = b_i, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mj}x_j + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{2.1}$$

или в краткой записи

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

В задачах линейного программирования представляют интерес системы, в которых ранг r матрицы системы $A = (a_{ij}),\ i = 1, 2, \ldots, m;$ $j = 1, 2, \ldots, n$ меньше числа переменных $(r < n)$. Будем полагать, что в системе (2.1) все m уравнений системы независимы $(r = m)$ и соответственно $m < n$.

Любые m переменных системы m линейных уравнений с n переменными $(m < n)$ называются ***основными*** *(или* ***базисными****), если определитель матрицы коэффициентов при них отличен от нуля Тогда остальные $n - m$ переменных называются* ***неосновными (свободными)***. Такой определитель называют *базисным минором* матрицы A.

Основными могут быть разные группы из n переменных. Максимально возможное число групп основных переменных равно числу способов выбора m переменных из общего числа n, т.е. числу сочетаний C_n^m.

Для решения системы (2.1) при условии $m < n$ существует теорема.

Теорема 2.1. *Если для системы m линейных уравнений с n переменными $(m < n)$ ранг матрицы коэффициентов при переменных равен m, то эта система является неопределенной, причем каждому произвольному набору значений неосновных переменных соответствует одно решение системы.*

Решение $X = (x_1, x_2, \ldots, x_n)$ системы (2.1) называется *допустимым*, если оно содержит лишь неотрицательные компоненты $(x_j \geq 0,\ j = 1, 2, \ldots, n)$. В противном случае решение называется *недопустимым*.

Среди бесконечного множества решений системы выделяют базисные решения.

Базисным *решением системы m линейных уравнений с n переменными называется решение, в котором все $n - m$ неосновных переменных равны нулю.*

В задачах линейного программирования особый интерес представляют *допустимые базисные решения,* или *опорные планы.* Число базисных решений является конечным и не превосходит C_n^m. Базисное решение, в котором хотя бы одна из основных переменных равна нулю, называется *вырожденным.*

Совместная система (2.1) имеет бесконечно много решений, из них базисных решений – конечное число, не превосходящее C_n^m.

2.2. Выпуклые множества точек

В математике *выпуклыми* называются многоугольники, целиком расположенные по одну сторону от прямых, на которых лежат их стороны (рис. 2.1).

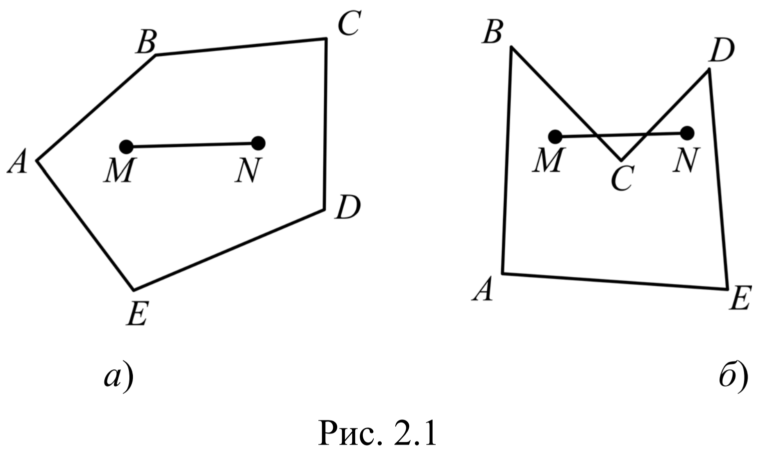


Рис. 2.1

Многоугольник на рис. 2.1,*а* – выпуклый, многоугольник на рис. 2.1,*б* не является выпуклым.

Множество точек называется ***выпуклым****, если оно вместе с любыми двумя своими точками содержит весь отрезок, соединяющий эти точки.*

Выпуклыми множествами могут быть не только многоугольники. На рис. 2.2 приведены примеры выпуклых множеств: круг, сектор, отрезок, многоугольная область, куб, пирамида. К выпуклым множествам относятся также многогранная область, прямая, полуплоскость, полупространство.

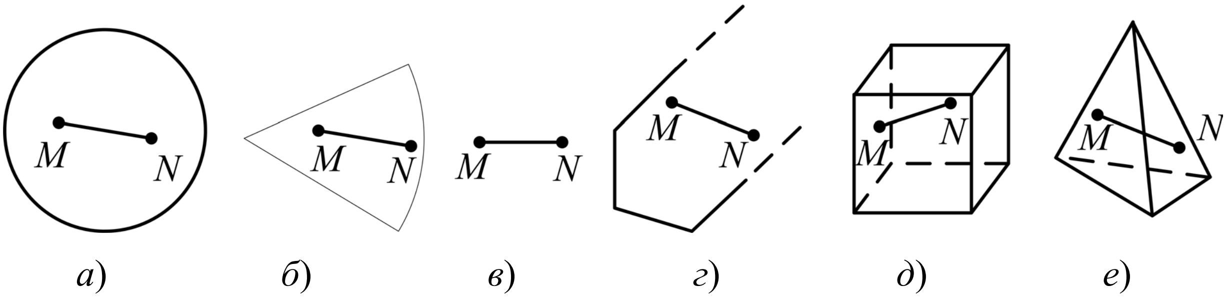


Рис. 2.2

Выпуклые множества обладают следующим свойством, которое можно сформулировать в виде теоремы.

Теорема 2.2. *Пересечение (общая часть) любого числа выпуклых множеств есть выпуклое множество.*

Введем понятие окрестности точки. Под *окрестностью* точки плоскости (пространства) подразумевается круг (шар) с центром в этой точке.

Среди точек выпуклого множества можно выделить внутренние, граничные и угловые точки.

Точка множества называется *внутренней*, если в некоторой ее окрестности содержатся точки только данного множества.

Точка множества называется *граничной,* если в любой ее окрестности содержатся как точки, принадлежащие данному множеству, так и точки, не принадлежащие ему.

Особый интерес в задачах линейного программирования представляют угловые точки. *Точка множества называется* ***угловой*** (***крайней***)*, если она не является внутренней ни для какого отрезка, целиком принадлежащего данному множеству.*

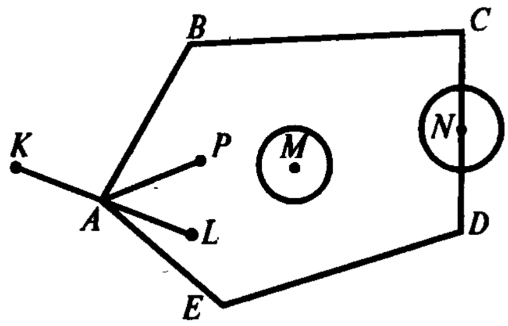


Рис. 2.3

На рис. 2.3 приведены примеры различных точек многоугольника: внутренней (точка M), граничной (точка N) и угловых точек (A, B, C, D, E). Для выпуклого множества угловые точки всегда совпадают с вершинами многоугольника, а для невыпуклого множества это не обязательно.

Множество точек называется *замкнутым*, если включает в себя все граничные точки. Множество точек называется *ограниченным*, если существует шар радиуса конечной длины с центром в любой точке множества, который полностью содержит в себе данное множество. В противном случае множество называется *неограниченным.*

Выпуклое замкнутое множество точек пространства, имеющее конечное число угловых точек, называется ***выпуклым многогранником,*** *если оно ограниченное, и* ***выпуклой многогранной областью****, если оно неограниченное.*

Множество всех точек $X = (x_1, x_2, \ldots, x_n)$ *образует* n*-мерное точечное (векторное) пространство.*

2.3. Геометрический смысл решений неравенств, уравнений и их систем

Рассмотрим решения неравенств. Справедлива теорема 2.3.

Теорема 2.3. *Множеством решений неравенства с двумя переменными*

$$a_{11}x_1 + a_{12}x_2 \le b_1 \tag{2.2}$$

является одна из двух полуплоскостей, на которые вся плоскость делится прямой $a_{11}x_1 + a_{12}x_2 = b_1$, *включая и эту прямую, а другая полуплоскость с той же прямой есть множество решений неравенства*

$$a_{11}x_1 + a_{12}x_2 \ge b_1. \tag{2.3}$$

Для определения искомой полуплоскости нужно задать произвольную контрольную точку, не лежащую на границе. Если неравенство выполняется в контрольной точке, то оно выполняется и во всех точках полуплоскости, содержащей контрольную точку.

Множество точек, удовлетворяющих уравнению

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \tag{2.4}$$

в n-мерном пространстве, называется *гиперплоскостью.* Теорему 2.3 можно обобщить.

Теорема 2.4. *Множеством всех решений линейного неравенства с n переменными*

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1$$

является одним из полупространств, на которые все пространство делится гиперплоскостью (2.4), включая и эту гиперплоскость.

Рассмотрим множество решений систем неравенств.

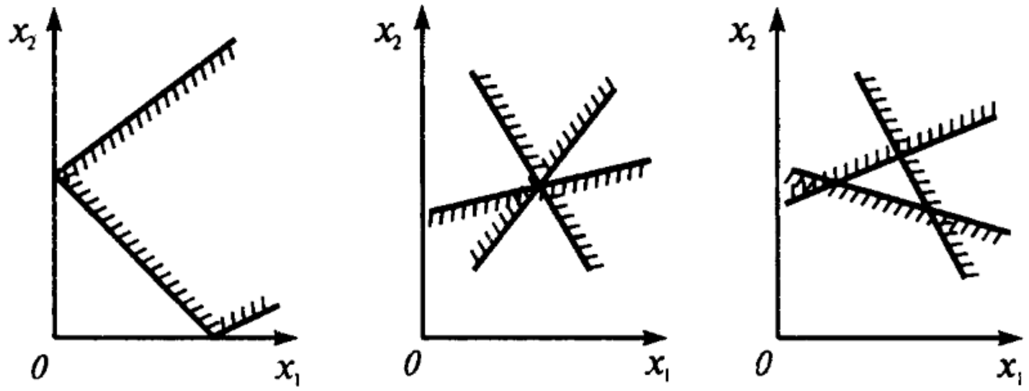
Теорема 2.5. *Множество решений совместной системы m линейных неравенств с двумя переменными*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \le b_1, \\ a_{21}x_1 + a_{22}x_2 \le b_2, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_{m1}x_1 + a_{m2}x_2 \le b_m \end{cases}$$

является выпуклым многоугольником.

Каждое из неравенств определяет одну из полуплоскостей, являющуюся выпуклым множеством точек. Множеством решений совместной системы линейных неравенств служат точки, которые принадлежат полуплоскостям решений всех неравенств, т.е. принадлежат их пересечению. Согласно теореме 2.2 о пересечении выпуклых множеств это множество является выпуклым и содержит конечное число угловых точек и является выпуклым многоугольником.

При построении областей решений систем неравенств могут встретиться другие случаи: множество решений – выпуклая многоугольная область (рис. 2.4,*а*); одна точка (рис. 2.4,*б*); пустое множество, когда система неравенств несовместна (рис.2.4,*в*).



а) *б*) *в*)

Рис. 2.4

Справедливы теорема 2.6 для множества решений совместных систем и теорема 2.7 – для множества допустимых решений.

Теорема 2.6. *Множество решений совместной системы m линейных неравенств с n переменными является выпуклым многогранником в n-мерном пространстве.*

Теорема 2.7. *Множество всех допустимых решений совместной системы m линейных уравнений с n переменными $(m < n)$ является выпуклым многогранником в n-мерном пространстве.*

Между допустимыми базисными решениями и угловыми точками множества допустимых решений системы линейных уравнений существует взаимнооднозначное соответствие.

Глава 3

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МЕТОДОВ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Рассмотрим теоретические основы методов линейного программирования.

3.1. Выпуклые множества в n-мерном пространстве

Необходимо определить выпуклое множество в n-мерном пространстве. Для этого нужно дать аналитическое определение отрезка в этом пространстве.

Рассмотрим двумерное пространство. Пусть $X_1 = (x_1^{(1)}, x_2^{(1)})$ и $X_2 = (x_1^{(2)}, x_2^{(2)})$ – точки плоскости Ox_1x_2, а $X = (x_1, x_2)$ – любая точка отрезка X_1X_2 (рис. 3.1). Отношение α длин отрезков XX_2 и X_1X_2 удовлетворяет условию $0 \le \alpha \le 1$. Запишем отношение α через координаты точек:

$$\alpha = \frac{x_1^{(2)} - x_1}{x_1^{(2)} - x_1^{(1)}} = \frac{x_2^{(2)} - x_2}{x_2^{(2)} - x_2^{(1)}},$$

откуда

$$\begin{cases} x_1 = \alpha x_1^{(1)} + (1-\alpha) x_1^{(2)}, \\ x_2 = \alpha x_2^{(1)} + (1-\alpha) x_2^{(2)}, \end{cases} \tag{3.1}$$

где

$$0 \le \alpha \le 1. \tag{3.2}$$

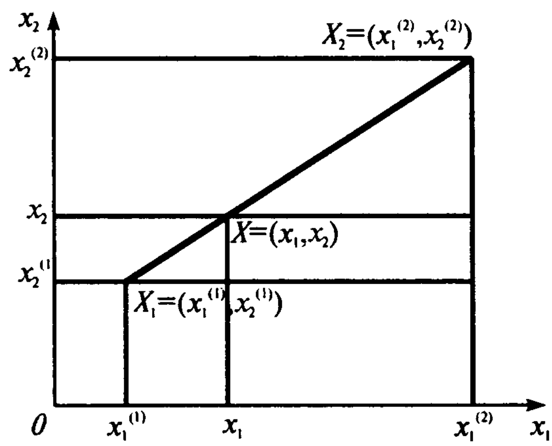


Рис. 3.1

Если обозначить $\alpha_1 = \alpha$ и $\alpha_2 = 1 - \alpha$, то условия (3.1), (3.2) примут вид

$$\begin{cases} x_1 = \alpha_1 x_1^{(1)} + \alpha_2 x_1^{(2)}, \\ x_2 = \alpha_1 x_2^{(1)} + \alpha_2 x_2^{(2)}, \end{cases} \tag{3.3}$$

$$\alpha_1 \geq 0, \alpha_2 \geq 0, \quad \alpha_1 + \alpha_2 = 1. \tag{3.4}$$

Равенство (3.3) можно записать в виде

$$X = \alpha_1 X_1 + \alpha_2 X_2. \tag{3.5}$$

Следовательно, ***отрезок*** $X_1 X_2$ *можно определить как множество точек, удовлетворяющих условиям (3.4) и (3.5).*

В случае n-мерного пространства определение отрезка будет таким же – множество точек, удовлетворяющих условиям (3.4) и (3.5), если под X_1 и X_2 подразумевать точки n-мерного пространства: $X_1 = (x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)})$ и $X_2 = (x_1^{(2)}, x_2^{(2)}, \ldots, x_n^{(2)})$.

Обобщением понятия отрезка для нескольких точек является их выпуклая линейная комбинация.

Точка X *называется* ***выпуклой линейной комбинацией*** *точек* $X_1, X_2, \ldots, X_n$, *если выполняются условия*

$$X = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_n X_n,$$

$$\alpha_j \geq 0 \quad (j = 1, 2, \ldots, n), \quad \sum_{j=1}^{n} \alpha_j = 1.$$

Множество точек является ***выпуклым****, если оно вместе с любыми своими двумя точками содержит их произвольную выпуклую линейную комбинацию.*

Обобщением является теорема о представлении выпуклого многогранника.

Теорема 3.1. *Выпуклый* n*-мерный многогранник является выпуклой линейной комбинацией своих угловых точек.*

Из теоремы 3.1 следует, что выпуклый многогранник порождается своими угловыми точками или вершинами: отрезок – двумя точками, треугольник – тремя, тетраэдр – четырьмя и т.д. Однако выпуклая многогранная область, являясь неограниченным множеством, не определяется однозначно своими угловыми точками: любую ее точку нельзя представить в виде выпуклой линейной комбинации угловых точек.

3.2. Свойства задачи линейного программирования

Любая задача линейного программирования может быть представлена в виде общей, канонической или стандартной задачи.

Рассмотрим каноническую задачу (1.20)–(1.22), в которой система ограничений есть система уравнений (2.1).

Для обоснования свойств задачи линейного программирования и методов ее решения целесообразно рассмотреть два вида записи канонической задачи.

Матричная форма записи:

$$F = CX \to \max(\min) \tag{3.6}$$

при ограничениях

$$AX = B, \tag{3.7}$$

$$X \geq 0, \tag{3.8}$$

где

$$C = (c_1, c_2, \ldots, c_n); \quad A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}; \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}; \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}.$$

Здесь C – матрица-строка, A – матрица системы; X – матрица-столбец переменных; B – матрица-столбец свободных членов.

Векторная форма записи:

$$F = CX \to \max(\min) \tag{3.9}$$

при ограничениях

$$P_1 x_1 + P_2 x_2 + \ldots + P_n x_n = P, \tag{3.10}$$

$$X \geq 0, \tag{3.11}$$

где CX – скалярное произведение векторов $C = (c_1, c_2, \ldots, c_n)$ и $X = (x_1, x_2, \ldots, x_n)$, векторы

$$P_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix}, \quad P_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix}, \ldots, \quad P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix}, \quad P = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$$

состоят из коэффициентов при переменных и свободных членах.

Была сформулирована теорема 2.7 о том, что множество всех допустимых решений задачи линейного программирования является выпуклым и представляет собой выпуклый многогранник или выпуклую многогранную область, которые в дальнейшем будем называть одним термином – *многогранником решений.*

Возникает вопрос: в какой точке многогранника решений возможно оптимальное решение задачи линейного программирования? Ответ на него дает следующая фундаментальная теорема.

Теорема 3.2. *Если задача линейного программирования имеет оптимальное решение, то линейная функция принимает максимальное (минимальное) значение в одной из угловых точек многогранника решений. Если линейная функция принимает максимальное значение более чем в одной угловой точке, то она принимает его в любой точке, являющейся выпуклой линейной комбинацией этих точек.*

Предположим, что многогранник решений является ограниченным. Обозначим его угловые точки через $X_1, X_2, \ldots, X_p$, а оптимальное решение – X^* (рис .3.2). Тогда $F(X^*) \geq F(X)$ для всех точек X многогранника решений. Если X^* – угловая точка, то первая часть теоремы доказана.

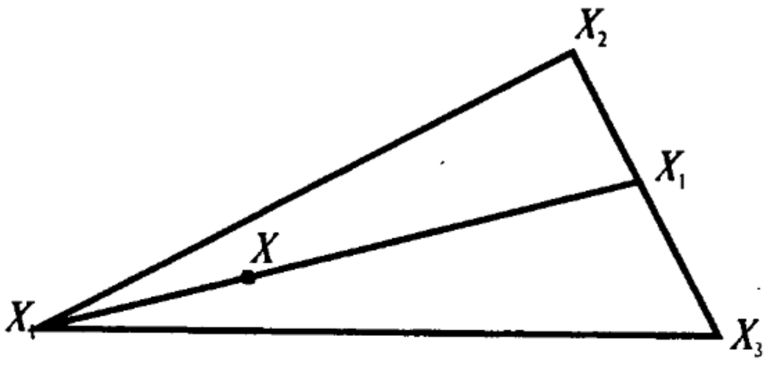


Рис. 3.2

Предположим, что X^* не является угловой точкой, тогда на основании теоремы 3.1 X^* можно представить как выпуклую линейную комбинацию угловых точек многогранника решений:

$$X^* = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_p X_P,$$

$$\alpha_j \geq 0, \quad j = 1, 2, \ldots, p; \quad \sum_{j=1}^{n} \alpha_j = 1.$$

Так как $F(X_j)$ – линейная функция, получаем

$$F(X^*) = F(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_p X_p) =$$
$$= \alpha_1 F(X_1) + \alpha_2 F(X_2) + \ldots + \alpha_p F(X_p). \qquad (3.12)$$

В этом разложении значений $F(X_j)\,(j = 1, 2, \ldots, p)$ выберем максимальное. Пусть оно соответствует угловой точке X_k $(1 \le k \le p)$; обозначим его M, т.е. $F(X_k) = M$. Заменим в выражении (3.12) каждое значение этим максимальным значением. Тогда

$$F(X^*) \le \alpha_1 M + \alpha_2 M + \ldots + \alpha_p M = M \sum_{j=1}^{p} \alpha_j = M.$$

По предположению X^* – оптимальное решение, поэтому, с одной стороны, $F(X^*) \ge F(X_k) = M$, но, с другой стороны, доказано, что $F(X^*) \le M$, следовательно, $F(X^*) = M = F(X_k)$, где X_k – угловая точка. Итак, существует угловая точка X_k, в которой линейная функция принимает максимальное значение.

Для доказательства второй части теоремы допустим, что $F(X)$ принимает максимальное значение более чем в одной угловой точке, например, в точках $X_1, X_2, \ldots, X_q$, где $1 \le q \le p$; тогда $F(X_1) = F(X_2) = \ldots = F(X_q) = M$.

Пусть X – выпуклая линейная комбинация этих угловых точек, т.е. выполняются условия

$$X = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_q X_q, \; \alpha_j \ge 0, \quad j = 1, 2, \ldots, q; \quad \sum_{j=1}^{q} \alpha_j = 1.$$

Учитывая, что функция $F(X)$ – линейная, получим

$$F(X^*) = F(\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_q Xq) =$$
$$= \alpha_1 F(X_1) + \alpha_2 F(X_2) + \ldots + \alpha_q F(X_q) =$$
$$= \alpha_1 M + \alpha_2 M + \ldots \alpha_q M = M \sum_{j=1}^{q} \alpha_j = M,$$

линейная функция $F(X)$ принимает максимальное значение в произвольной точке X, являющейся выпуклой линейной комбинацией угловых точек $X_1, X_2, \ldots, X_q$.

Замечание. Требование ограниченности многогранника решений является существенным, так как в случае неограниченной многогранной области не каждую точку можно представить выпуклой линейной комбинацией ее угловых точек.

Теорема 3.2 является фундаментальной для решения задач линейного программирования. Согласно этой теореме вместо исследования бесконечного множества допустимых решений для нахождения оптимального решения необходимо исследовать конечное число угловых точек многогранника решений.

Рассмотрим теорему, посвященную аналитическому методу нахождения угловых точек.

Теорема 3.3. *Каждому допустимому базисному решению задачи линейного программирования соответствует угловая точка многогранника решений, и наоборот, каждой угловой точке многогранника решений соответствует допустимое базисное решение.*

Пусть $X = (x_1, x_2, \ldots, x_m; 0, 0, \ldots, 0)$ – допустимое базисное решение системы ограничений (3.10) задачи, в котором первые m компонент – основные переменные, а остальные $n-m$ компонент – неосновные переменные, равные нулю в базисном решении. Покажем, что X – угловая точка многогранника решений.

Предположим, что X не является угловой точкой. Тогда точку X можно представить внутренней точкой отрезка, соединяющего две различные точки

$$X_1 = (x_1^{(1)}, x_2^{(1)}, \ldots, x_m^{(1)}; 0, 0, \ldots, 0);$$
$$X_2 = (x_1^{(2)}, x_2^{(2)}, \ldots, x_m^{(2)}; 0, 0, \ldots, 0).$$

Другими словами, точку X можно представить линейной комбинацией точек X_1 и X_2 многогранника решений:

$$X = \alpha_1 X_1 + \alpha_2 X_2, \tag{3.13}$$

где $\alpha_1 > 0, \alpha_2 > 0, \alpha_1 + \alpha_2 = 1.$

Запишем векторное равенство (3.13) в координатном виде:

$$\begin{cases} x_1 = \alpha_1 x_1^{(1)} + \alpha_2 x_1^{(2)}, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ x_1 = \alpha_1 x_m^{(1)} + \alpha_2 x_m^{(2)}, \\ 0 = \alpha_1 x_{m+1}^{(1)} + \alpha_2 x_{m+1}^{(2)}, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ 0 = \alpha_1 x_n^{(1)} + \alpha_2 x_n^{(2)}. \end{cases}$$

Поскольку $x_j^{(1)} \ge 0, x_j^{(2)} \ge 0 \; (j = 1, 2, \ldots, n), \alpha_1 > 0, \alpha_2 > 0,$ то из последних $n - m$ равенств следует, что $x_{m+1}^{(1)} = 0, x_{m+1}^{(2)} = 0, \ldots, x_n^{(1)} = 0, x_n^{(2)} = 0$. Это означает, что в решениях X_1, X_2 и X системы уравнений (3.10) значения $n - m$ компонент равны в данном случае нулю. Эти компоненты можно считать значениями неосновных переменных. Однако значения неосновных переменных однозначно определяют значения основных, следовательно, $x_1^{(1)} = x_1^{(2)} = x_1, \ldots, x_m^{(1)} = x_m^{(2)} = x_2$. Таким образом, все n компонент в решениях X, X_1, X_2 совпадают и точки X_1, X_2 сливаются, что противоречит допущению. Следовательно, X – угловая точка многогранника решений.

Из теорем 3.2 и 3.3 вытекает важное следствие: *если задача линейного программирования имеет оптимальное решение, то оно совпадает, по крайней мере, с одним из ее допустимых базисных решений.*

Итак, *оптимум линейной функции задачи линейного программирования следует искать среди конечного числа ее допустимых базисных решений.*

Глава 4

ГЕОМЕТРИЧЕСКИЙ МЕТОД РЕШЕНИЯ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

В предыдущих главах было доказано, что множество допустимых решений задачи линейного программирования представляет собой выпуклый многогранник, а оптимальное решение задачи находится в одной из угловых точек многогранника решений.

Рассмотрим задачу в стандартной форме (1.4)–(1.6) с двумя переменными. Пусть геометрическим изображением системы ограничений является многоугольник $ABCDE$ (рис. 4.1). Необходимо среди точек этого многоугольника найти такую точку, в которой функция $F = c_1x_1 + c_2x_2$ принимает максимальное (или минимальное) значение.

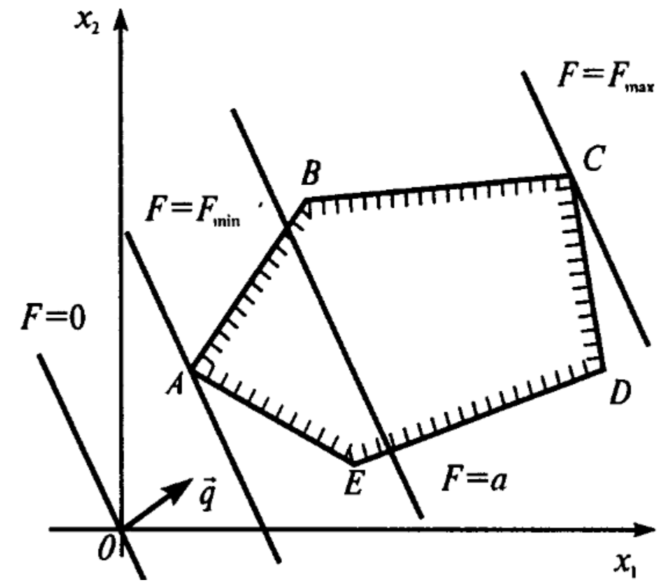


Рис. 4.1

Рассмотрим *линию уровня* линейной функции F, т.е. линию, вдоль которой эта функция принимает одно и то же фиксированное значение, $F = a$, или

$$c_1x_1 + c_2x_2 = a. \tag{4.1}$$

На многоугольнике решений следует найти точку, через которую проходит линия уровня функции F с наибольшим (если линейная функция максимизируется) или наименьшим (если функция минимизируется) уровнем. Уравнение линии уровня функции (4.1) есть уравнение прямой. При различных уровнях a линии уровня параллельны, поскольку их угловые коэффициенты определяются соотношением между коэффициентами c_1 и c_2, следовательно, равны.

Важное свойство линии уровня линейной функции состоит в том, что при параллельном смещении линии в одну сторону уровень только возрастает, а при смещении в другую сторону – только убывает.

Пусть имеются три линии уровня

$$F = c_1x_1 + c_2x_2 = a_1; \tag{4.1}$$

$$F = c_1x_1 + c_2x_2 = a_2; \tag{4.2}$$

$$F = c_1x_1 + c_2x_2 = a_3, \tag{4.3}$$

причем линия (4.2) заключена между линиями (4.1) и (4.3).Тогда $a_1 < a_2 < a_3$ или $a_1 > a_2 > a_3$.

Для определения направления возрастания рекомендуется изобразить две линии уровня и определить, на которой из них уровень больше. Например, если одну из линий можно взять проходящей через начало координат и линейная функция имеет вид $F = c_1x_1 + c_2x_2$ (без свободного члена), то это соответствует нулевому уровню. Другую линию можно провести произвольно, но так, чтобы она проходила через множество решений системы ограничений. Далее, определив направление возрастания линейной функции, найдем точку, в которой функция принимает максимальное или минимальное значение.

Задача 4.1. Решить геометрически следующую задачу.

Для изготовления двух видов продукции P_1 и P_2 используют четыре вида ресурсов S_1, S_2, S_3 и S_4. Запасы ресурсов, число единиц ресурсов, затрачиваемых на изготовление единицы продукции, приведены в табл. 4.1 (цифры условные).

Таблица 4.1

Вид ресурса	Запас ресурса	Число единиц ресурсов, затрачиваемых на изготовление единицы продукции	
		P_1	P_2
S_1	18	1	3
S_2	16	2	1
S_3	5	-	1
S_1	21	3	–

Прибыль, получаемая от единицы продукции P_1 и P_2, соответственно, 2 и 3 руб.

Линейная функция имеет вид (рис. 4.2)

$$F = 2x_1 + 3x_2 \to \max$$

при ограничениях:

$$\begin{cases} x_1 + 3x_2 \le 18\,(1), \\ 2x_1 + x_2 \le 16\,(2), \\ x_2 \le 5\,(3), \\ 3x_1 \le 21\,(4), \\ x_1 \ge 0,\ x_2 \ge 0\ \ (5,6). \end{cases}$$

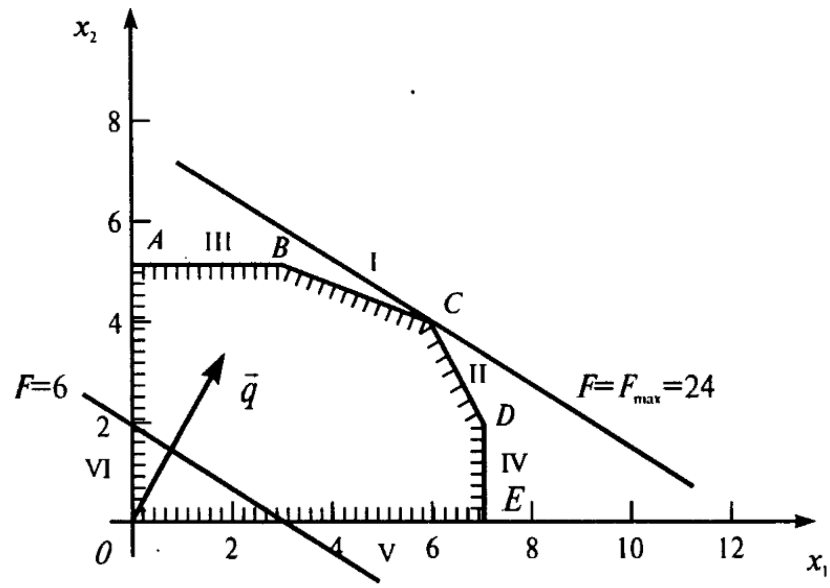


Рис. 4.2

Решение. Изобразим многоугольник решений на рис. 4.3. При $F = 0$ линия уровня $2x_1 + 3x_2 = 0$ проходит через начало координат. Зададим $F = 6$ и построим линию уровня $2x_1 + 3x_2 = 6$. Ее расположение указывает направление возрастания линейной функции (вектор $\vec{q}$). Рассматривается задача на отыскание максимума, и оптимальное решение находится в угловой точке C. Точка C стоит на пересечении прямых (1) и (2), координаты которой определяются решением системы уравнений

$$\begin{cases} x_1 + 3x_2 = 18, \\ 2x_1 + x_2 = 16, \end{cases}$$

откуда $x_1 = 6, \quad x_2 = 4,$ и точка $C - (6,4)$.

Максимальное значение линейной функции равно $F_{\max} = 2 \cdot 6 + 3 \cdot 4 = 24$. Это означает, что максимальная прибыль в 24 руб. может быть достигнута при производстве 6 единиц продукции P_1 и 4 единиц продукции P_2.

При геометрическом решении задач линейного программирования возможны случаи, когда условия задач противоречивы, т.е. область допустимых решений системы ограничений представляет пустое множество. В таких задачах нет оптимальных решений и нет смысла строить линию уровня.

Рассмотренный геометрический метод решения задач линейного программирования обладает рядом достоинств: простотой, наглядностью, позволяет быстро получить ответ.

Однако у этого метода есть недостатки. Возможны погрешности при приближенном построении графиков. Но самое главное – этот метод можно применять только в случае, когда число переменных в стандартной задаче равно двум. Поэтому необходимы аналитические методы, позволяющие решать задачи линейного программирования с любым числом переменных и выявлять экономический смысл входящих в них величин.

Глава 5

СИМПЛЕКСНЫЙ МЕТОД

5.1. Геометрическая интерпретация симплексного метода

В главе 3 рассмотрены основные теоремы линейного программирования, из которых следует, что если задача линейного программирования имеет оптимальное решение, то оно соответствует хотя бы одной угловой точке многогранника решений и совпадает с одним из допустимых базисных решений системы ограничений. Был указан путь решения любой задачи линейного программирования: перебрать конечное число допустимых базисных решений системы ограничений и выбрать из них то, на котором функция цели принимает оптимальное решение. Геометрически это соответствует перебору всех угловых точек многогранника решений. Для реальных задач число допустимых базисных решений может быть чрезвычайно велико, потребуется большое число вычислений.

Число перебираемых допустимых базисных решений можно сократить, если производить перебор не беспорядочно, а с учетом изменений линейной функции, добиваясь того, чтобы каждое следующее решение было лучше, чем предыдущее (увеличение при отыскании максимума $F \to \max$, уменьшение – при отыскании минимума $F \to \min$).

Поясним на примере. Пусть область допустимых решений изображается многоугольником $ABCDEGH$ (рис. 5.1). Предположим, что его угловая точка A соответствует исходному допустимому базисному решению. Из чертежа видно, что после вершины A выгодно перейти к соседней вершине B, а затем – к оптимальной точке C.

Вместо семи перебрали только три вершины, последовательно улучшая линейную функцию.

Идея последовательного улучшения решения лежит в основе универсального метода решения задач линейного программирования – ***симплексного метода.*** Симплекс (от лат. simplex – простой) – простейший выпуклый многогранник в n-мерном пространстве с $n+1$ вершиной; симплексом также является область допустимых решений неравенства $\sum\limits_{i=1}^{n} x_i \le 1$.

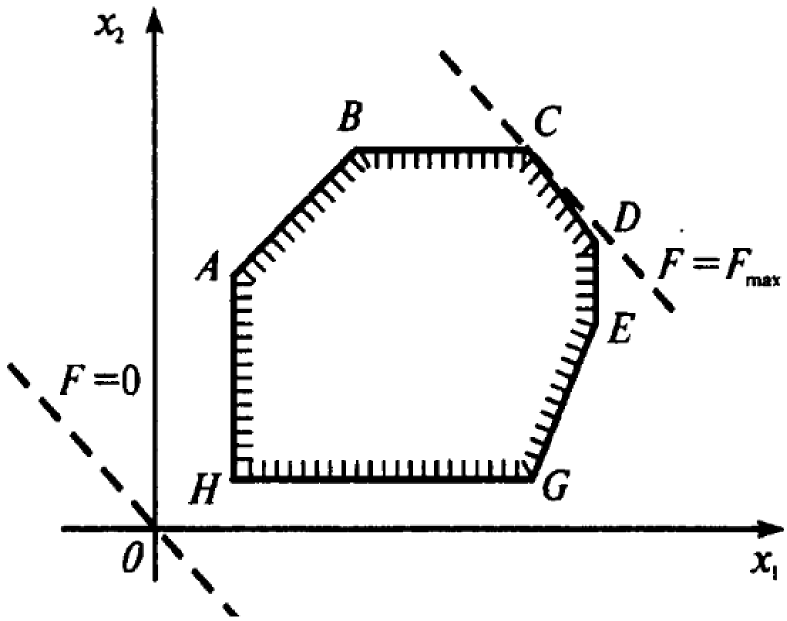


Рис. 5.1

Геометрический смысл симплексного метода состоит в последовательном переходе от одной вершины многогранника ограничений к соседней, в которой линейная функция принимает лучшее значение до тех пор, пока не будет найдено оптимальное решение – вершина, где достигается оптимальное значение функции цели.

Идеи симплексного метода были разработаны в 1939 году российским ученым Л. В. Канторовичем. В окончательном виде симплексный метод был предложен американским ученым Дж. Данцигом в 1949 году.

Для реализации симплексного метода – последовательного улучшения решения – необходимо освоить *три основных элемента*:

- *способ определения какого-либо первоначального допустимого базисного решения задачи;*
- *правило перехода к лучшему решению;*
- *критерий проверки оптимальности найденного решения.*

Для использования симплексного метода задача линейного программирования должна быть приведена к каноническому виду, система ограничений должна быть представлена в виде уравнений.

5.2. Отыскание максимума линейной функции

Рассмотрим алгоритм вычислительной реализации симплексного метода. В качестве примера рассмотрим задачу об использовании ресурсов, сформулированную в разд. 1.2 и решенную геометрически в задаче 4.1.

Задача. Решить симплексным методом

$$F = 2x_1 + 3x_2 \to \max$$

при ограничениях

$$\begin{cases} x_1 + 3x_2 \le 18, \\ 2x_1 + x_2 \le 16, \\ x_2 \le 5, \\ 3x_1 \le 21, \end{cases} \tag{5.1}$$

$$x_1 \ge 0, \quad x_2 \ge 0.$$

Решение. С помощью дополнительных неотрицательных переменных перейдем к системе уравнений. В данном случае все дополнительные переменные вводятся со знаком «плюс», так как все неравенства имеют вид «≤». Получим систему ограничений в виде

$$\begin{cases} x_1 + 3x_2 + x_3 = 18, \\ 2x_1 + x_2 + x_4 = 16, \\ x_2 + x_5 = 5, \\ 3x_1 + x_6 = 21. \end{cases} \tag{5.2}$$

Для нахождения первоначального базисного решения разобьем переменные на две группы – основные и неосновные. Поскольку определитель, составленный из коэффициентов при дополнительных переменных x_3, x_4, x_5, x_6, отличен от нуля, то эти переменные можно взять в качестве основных на первом шаге решения задачи. При выборе основных переменных на первом шаге можно воспользоваться следующим правилом:

в качестве основных переменных на первом шаге следует выбрать такие m переменных, каждая из которых входит только в одно из m уравнений системы ограничений, при этом нет таких уравнений системы, в которые не входит ни одна из этих переменных.

Если выбранные по этому правилу переменные имеют те же знаки, что и соответствующие им свободные члены в правых частях уравнений, то полученное базисное решение будет допустимым.

1-й шаг. Основные переменные: x_3, x_4, x_5, x_6.

Неосновные переменные: x_1, x_2.

Выразим основные переменные через неосновные:

$$\begin{cases} x_3 = 18 - x_1 - 3x_2, \\ x_4 = 16 - 2x_1 - x_2, \\ \boxed{x_5 = 5 \qquad - x_2,} \\ x_6 = 21 - 3x_1. \end{cases} \tag{5.3}$$

Положив неосновные переменные равными нулю $(x_1 = 0, x_2 = 0)$, получим базисное решение $X_1 = (0;0;18;16;5;21)$, которое является допустимым и соответствует вершине $O(0;0)$ многогранника $OABCDE$ на рис. 5.2.

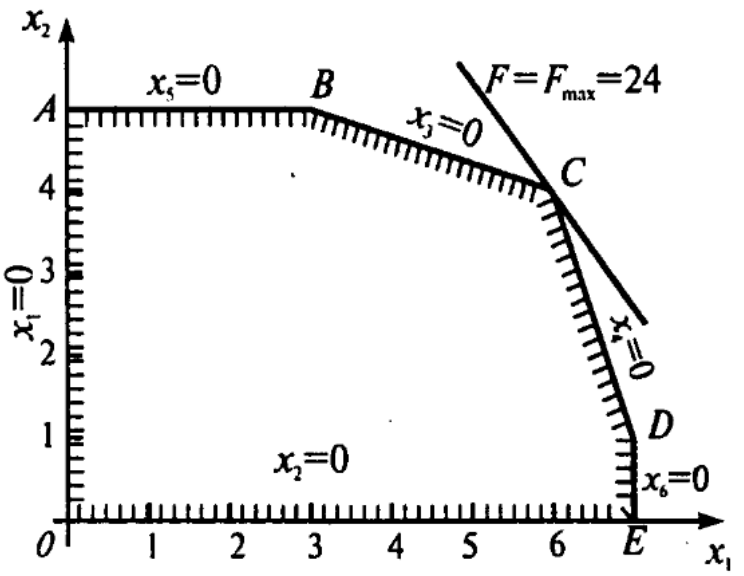


Рис 5.2

Поскольку это решение допустимое, то, возможно, оно оптимальное. Выразим линейную функцию через неосновные переменные: $F = 2x_1 + 3x_2$. При решении X_1 значение функции равно $F(X_1)$. Функцию F можно увеличить за счет увеличения любой из неосновных переменных, входящих в F с положительным коэффициентом. Это можно осуществить, перейдя к такому новому допустимому базисному решению, в котором эта переменная будет неосновной. При таком переходе одна из основных переменных перейдет в неосновные, а геометрически произойдет переход к соседней вершине многоугольника, где значение линейной функции «лучше». В данном примере для увеличения F можно переводить в основные либо x_1, либо x_2, так как обе эти переменные входят в F со знаком «плюс». Для определенности выберем переменную, имеющую больший коэффициент. В данном случае – это x_2.

Система (5.3) накладывает ограничения на рост переменной x_2. Необходимо сохранить допустимость решения, и все переменные должны оставаться неотрицательными. При этом $x_1 = 0$ – как неосновная переменная, и выполняются следующие неравенства:

$$\begin{cases} x_3 = 18 - 3x_2 \geq 0, \\ x_4 = 16 - x_2 \geq 0, \\ x_5 = 5 - x_2 \geq 0, \\ x_6 = 21, \end{cases} \quad \text{откуда} \begin{cases} x_2 \leq 18/3, \\ x_2 \leq 16/1, \\ x_2 \leq 5/1. \end{cases}$$

Каждое уравнение системы (5.3), кроме последнего , определяет оценочное отношение – границу роста переменной x_2. Эта граница определяется абсолютной величиной отношения свободного члена к коэффициенту при x_2.

Сохранение неотрицательности всех переменных (допустимость решения) возможно, если не нарушается ни одна из полученных во всех уравнениях границ. В данном случае наибольшее возможное значение для переменной x_2 определяется как $x_2 = \min\{18/3; 16/1; 5/1; \infty\} = 5$. При $x_2 = 5$ переменная x_5 обращается в нуль и переходит в неосновные.

Уравнение, в котором достигается наибольшее возможное значение переменной, переводимой в основные (т.е. где оценка минимальна), называется *разрешающим.* Разрешающее уравнение будем выделять рамкой в системе ограничений.

2-й шаг. Основные переменные: x_2, x_3, x_4, x_6.

Неосновные переменные: x_1, x_5.

Выразим новые основные переменные через новые неосновные, начиная с разрешающего уравнения:

$$\begin{cases} x_2 = 5 - x_5, \\ x_3 = 18 - x_1 - 3(5 - x_5), \\ x_4 = 16 - 2x_1 - (5 - x_5), \\ x_5 = 21 - 3x_1, \end{cases} \tag{5.4}$$

или

$$\begin{cases} x_2 = 5 - x_5, \\ \boxed{x_3 = 3 - x_1 + 3x_5,} \\ x_4 = 11 - 2x_1 + x_5, \\ x_5 = 21 - 3x_1. \end{cases}$$

Второе базисное решение $X_2 = (0;5;3;11;0;21)$ является допустимым и соответствует вершине $A(0;5)$ на рис. 5.2. Геометрическая интерпретация перехода от X_1 к X_2 – переход от вершины O к соседней вершине A на многоугольнике решений $OABCDE$.

Выразив линейную функцию через неосновные переменные на этом шаге, получаем:

$$F = 2x_1 + 3x_2 = 2x_1 + 3(5 - x_5) = 15 + 2x_1 - 3x_5.$$

Значение линейной функции $F_2 = F(X_2) = 15$. Изменение значения линейной функции можно определить заранее как произведение наибольшего возможного значения переменной, переводимой в основные, на ее коэффициент в выражении для линейной функции; в данном случае $\Delta F_1 = 5 \cdot 3 = 15, \quad F_2 = F_1 + \Delta F_1 = 0 + 15 = 15.$

Однако значение F_2 не является максимальным, так как, повторяя рассуждения первого шага, обнаруживаем возможность дальнейшего увеличения линейной функции за счет переменной x_1, которая входит в F с положительным знаком. Система уравнений (5.4) определяет наибольшее возможное значение для x_1: $x_1 = \min\{\infty; 3/1; 11/2; \infty\} = 3.$ Второе уравнение является разрешающим, переменная x_3 переходит в неосновные, при этом $\Delta F_2 = 3 \cdot 2 = 6$.

3-й шаг. Основные переменные: x_1, x_2, x_4, x_6.

Неосновные переменные: x_3, x_5.

Выражаем новые основные переменные через неосновные. После преобразований получаем:

$$\begin{cases} x_1 = 3 - x_3 + 3x_5, \\ x_2 = 5 - x_5, \\ \boxed{x_4 = 5 + 2x_3 - 5x_5,} \\ x_6 = 12 + 3x_3 - 9x_5. \end{cases} \tag{5.5}$$

Базисное решение $X_3 = (3;5;0;5;12)$ соответствует вершине $B(3;5)$. Выражаем линейную функцию через неосновные переменные: $F = 2x_1 + 3x_2 = 2(3 - x_3 + 3x_5) + 3(5 - x_5) = 21 - 2x_3 + 3x_5$, $F_3 = F(X_3) = 21$. Проверяем: $F_3 - F_2 = 21 - 15 = 6 = \Delta F_2$. Третье допустимое базисное решение тоже не является оптимальным, поскольку при неосновной переменной x_5 в выражении линейной функции через неосновные переменные содержится положительный коэффициент. Переводим x_5 в основную переменную. При определении наибольшего возможного значения для x_5 следует обратить внимание на первое уравнение в системе (5.5). Оно не дает ограничений на рост x_5. Поэтому $x_5 = \min\{\infty; 5; 1; 12/9\} = 1$. Третье уравнение является разрешающим, переменная x_4 переходит в неосновные; $\Delta F_3 = 1 \cdot 3 = 3$.

4-й шаг. Основные переменные: x_1, x_2, x_5, x_6.

Неосновные переменные: x_3, x_4.

После преобразований получим:

$$\begin{cases} x_1 = 6 + \frac{1}{5}x_3 - \frac{3}{5}x_4, \\ x_2 = 4 - \frac{2}{5}x_3 + \frac{1}{5}x_4, \\ x_5 = 1 + \frac{2}{5}x_3 - \frac{1}{5}x_4, \\ x_6 = 3 - \frac{3}{5}x_3 + \frac{9}{5}x_4. \end{cases}$$

Базисное решение $X_4 = (6;4;0;0;1;3)$ соответствует вершине $C(6;4)$ на рис. 5.2.

Линейная функция, выраженная через неосновные переменные, имеет вид $F = 24 - \frac{4}{5}x_3 - \frac{3}{5}x_4$. Это выражение не содержит положительных коэффициентов при неосновных переменных, поэтому значение $F_4 = F(X_4) = 24$ максимальное. Функцию F невозможно увеличить, и решение X_4 оптимальное. Вспоминая экономический смысл всех переменных, можно сделать выводы.

Прибыль предприятия принимает максимальное значение 24 руб. при реализации 6 единиц продукции $P_1 (x_1 = 6)$ и 4 единиц продукции $P_2(x_2=4)$. Дополнительные переменные x_3, x_4, x_5, x_6 показывают разницу между запасами ресурсов каждого вида и их потреблением, т.е. остатки ресурсов. При оптимальном плане производства $x_3 = x_4 = 0$ остатки ресурсов S_1 и S_2 равны нулю, а остатки ресурсов S_3 и S_4 равны соответственно 1 и 3.

Сформулируем критерий оптимальности решения при отыскании максимума линейной функции: *если в выражении линейной функции через неосновные переменные отсутствуют положительные коэффициенты при неосновных переменных, то решение оптимально.*

5.3. Отыскание минимума линейной функции

При определении минимума линейной функции Z возможны два пути:

1) отыскать максимум функции F, полагая $F=-Z$ и учитывая, что $Z_{\min} = -F_{\max}$;

2) модифицировать симплексный метод: на каждом шаге уменьшать линейную функцию за счет той неосновной переменной, которая входит в выражение линейной функции с отрицательным коэффициентом.

Рассмотрим следующий пример.

Задача. Решить симплексным методом

$$Z = 18y_1 + 16y_2 + 5y_3 + 21y_4 \to \min$$

при ограничениях:

$$\begin{cases} y_1 + 2y_2 + 3y_4 \geq 2, \\ 3y_1 + y_2 + y_3 \geq 3, \end{cases}$$

$$y_i \geq 0, \quad i = 1,2,3,4.$$

Решение. Введем дополнительные неотрицательные переменные X_n и y_6 со знаком «минус». Получим систему уравнений:

$$\begin{cases} y_1 + 2y_2 + 3y_4 - y_5 = 2, \\ 3y_1 + y_2 + y_3 - y_6 = 3. \end{cases}$$

Если на первом шаге в качестве основных взять дополнительные переменные, то получим недопустимое базисное решение:

(0;0;0;0;–2;–3). В данном случае в качестве основных удобно взять переменные y_3 и y_4. Коэффициенты при y_3 и y_4 положительные, поэтому в качестве первоначального получим допустимое базисное решение.

1-й шаг. Основные переменные: y_3, y_4.

Неосновные переменные: y_1, y_2, y_5, y_6.

Выразим основные переменные через неосновные:

$$\begin{cases} y_3 = 3 - 3y_1 - y_2 + y_6, \\ y_4 = \left(2/3\right) - \left(1/3\right)y_1 - \left(2/3\right)y_2 + \left(1/3\right)y_5. \end{cases}$$

$Y_1 = \left(0;0;3;\dfrac{2}{3};0;0\right)$ – первое базисное решение. Оно допустимое. Выразим линейную функцию через неосновные переменные:

$$Z = 18y_1 + 16y_2 + 5y_3 + 21y_4 =$$

$$= 18y_1 + 16y_2 + 5\left(3 - 3y_1 - y_2 + y_6\right) + 21\left(2/3 - 1/3\, y_1 - 2/3\, y_2 + 1/3\, y_5\right) =$$

$$= 29 - 4y_1 - 3y_2 + 7y_5 + 5y_6.$$

$Z_1 = Z(Y_1) = 29$ – это значение не является минимальным, так как функцию Z можно уменьшить за счет перевода в основные любой из переменных y_1 или y_2, имеющих в выражении для Z отрицательные коэффициенты. Поскольку y_1 имеет больший по абсолютному значению коэффициент, то начнем с этой переменной. Для нее наибольшее возможное значение: $y_1 = \min\left\{\dfrac{3}{3};\dfrac{2}{3};\dfrac{1}{3}\right\} = 1$, т.е. первое уравнение является разрешающим; y_3 становится неосновной переменной, $\Delta Z_1 = -4 \cdot 1 = -4$.

2-й шаг. Основные переменные: y_1, y_4.

Неосновные переменные: y_2, y_3, y_5, y_6.

Получим после преобразований:

$$\begin{cases} y_1 = 1 - \dfrac{1}{3}y_2 - \dfrac{1}{3}y_3 + \dfrac{1}{3}y_6, \\ y_4 = \dfrac{1}{3} - \dfrac{5}{9}y_2 + \dfrac{1}{9}y_3 + \dfrac{1}{3}y_5 - \dfrac{1}{9}y_6. \end{cases}$$

$Z=25-\frac{5}{3}y_2+\frac{4}{3}y_3+7y_5+\frac{11}{3}y_6-$ линейная функция. При базисном решении $Y_2=\left(1;0;0;\frac{1}{3};0;0\right)$ получаем $Z_2=Z(Y_2)=25.$ $Z_2-Z_1=25-29=-4=\Delta Z_1$. Переменную y_2 переводим в основные, так как в выражение для Z она входит с отрицательным коэффициентом. Наименьшее возможное значение $y_2=\min\left\{3;\frac{3}{5}\right\}=\frac{3}{5}$, второе уравнение разрешающее и y_4 переходит в неосновные переменные; $\Delta Z_2=\frac{3}{5}-\frac{5}{3}=-1.$

3-й шаг. Основные переменные: y_1, y_2.

Неосновные переменные: y_3, y_4, y_5, y_6.

После преобразований получим:

$$\begin{cases} y_1=\frac{4}{5}-\frac{2}{5}y_3+\frac{3}{5}y_4-\frac{1}{5}y_5+\frac{2}{5}y_6, \\ y_2=\frac{3}{5}+\frac{1}{5}y_3-\frac{9}{5}y_4+\frac{3}{5}y_5-\frac{1}{5}y_6, \end{cases}$$

$Z=24+y_3+3y_4+6y_5+4y_6$. Базисное решение $Y_3=\left(\frac{4}{5};\frac{3}{5};0;0;0;0\right)$ оптимальное, так как в выражении для Z нет неосновных переменных с отрицательными коэффициентами. Поэтому $Z_{\min}=Z_3=Z(Y_3)=24.$ $Z_3-Z_2=24-25=-1=\Delta Z_2$.

Критерий оптимальности при отыскании минимума линейной функции: *если в выражении линейной функции через неосновные переменные отсутствуют отрицательные коэффициенты при неосновных переменных, то решение оптимально.*

Замечание. На каждом шаге симплексного метода какая-либо неосновная переменная переводится в основные, при этом каждое уравнение системы ограничений определяет конечное или бесконечное наибольшее возможное значение этой переменной – оценочное отношение. В разобранных задачах встречались различные случаи оценки роста неосновной переменной. Сформулируем все возможные случаи. Обозначим: x_i – переводимая неосновная переменная, b_j – свободный член, a_{ij} – коэффициент при x_i. В общем виде уравнение

$x_j = b_j + \ldots + a_{ij} x_i + \ldots$ определяет наибольшее возможное значение x_i по следующим правилам:

1) $x_i = \left|\dfrac{b_j}{a_{ij}}\right|$, если b_j и a_{ij} разного знака и $a_{ij} \neq 0, b_j \neq 0$. Например: $x_3 = 8 - 2x_2 + \ldots; \; x_2 = \dfrac{8}{2} = 4$ или $x_3 = -8 + 2x_2 + \ldots; \; x_2 = \dfrac{8}{2} = 4.$

2) $x_i = \infty$, если b_j и a_{ij} одного знака и $a_{ij} \neq 0, b_j \neq 0$; например, $x_3 = 8 + 2x_2 + \ldots; x_2 = \infty.$

3) $x_i = 0$, если $b_j = 0$ и $a_{ij} < 0$; например, $x_3 = 0 - 2x_2 + \ldots;\; x_2 = 0.$

4) $x_i = \infty$, если $b_j = 0$ и $a_{ij} > 0$; например, $x_3 = 0 + 2x_2 + \ldots;\; x_2 = \infty.$

5) $x_i = \infty$, если $a_{ij} = 0$; например, $x_3 = 5 + 0 \cdot x_2 + \ldots;\; x_2 = \infty.$

5.4. Определение первоначального допустимого базисного решения

В рассмотренных примерах оптимальное решение получено путем последовательного перехода от первоначального допустимого базисного решения к следующему, лучшему, и так до достижения критерия оптимальности. Однако не всегда на первом шаге получается допустимое базисное решение. Рассмотрим на примере один из возможных алгоритмов получения допустимого базисного решения.

Задача. Решить симплексным методом

$$F = x_1 + 2x_2 \to \max$$

при ограничениях

$$\begin{cases} x_1 - x_2 \le -1, \\ x_1 - x_2 \ge -3, \\ x_1 \le 3, \end{cases}$$

$$x_1 \ge 0, \quad x_2 \ge 0.$$

Решение. Вводим дополнительные неотрицательные переменные x_3, x_4, x_5 с соответствующими знаками:

$$\begin{cases} x_1 - x_2 + x_3 = -1, \\ x_1 - x_2 - x_4 = -3, \\ x_1 + x_5 = 3. \end{cases}$$

В соответствии с правилом, сформулированным в разд. 5.2, на первом шаге в качестве основных возьмем дополнительные переменные.

1-й шаг. Основные переменные: x_3, x_4, x_5.

Неосновные переменные: x_1, x_2.

Выразим основные переменные через неосновные:

$$\begin{cases} x_3 = -1 - x_1 + x_2, \\ x_4 = 3 + x_1 - x_2, \\ x_5 = 3 - x_1. \end{cases}$$

$X_1 = (0;0;-1;3;3)$ – первое базисное решение недопустимое (имеется отрицательная компонента), оно не может быть оптимальным. Линейную функцию на недопустимом решении не рассматриваем. В исходной системе выберем то уравнение, которое содержит отрицательный свободный член, т.е. первое уравнение (если таких уравнений несколько, выбираем любое из них).

Поскольку переменная $x_1 \geq 0, x_2 \geq 0, \ldots x_n \geq 0$ принимает отрицательное значение при первом базисном решении, то ее необходимо увеличить. Это можно сделать за счет увеличения любой из неосновных переменных, входящих в первое уравнение с положительным коэффициентом, в данном случае – переменной x_2. Если перевести эту переменную в основные, то она увеличит переменную x_3. Как только x_2 достигнет значения 1, то x_3 обратится в 0 и исчезнет отрицательная компонента в решении. Можно считать, что переменная x_3 станет неосновной. Рост переменной x_2 ограничен условиями неотрицательности остальных переменных, которые определяют $Y = (y_1, y_2, \ldots, y_m)$. Первое уравнение разрешающее. При $x_2 = 1$ переменная $x_3 = 0$ и переходит в неосновные переменные.

2-й шаг. Основные переменные: x_2, x_4, x_5.

Неосновные переменные: x_1, x_3.

Выражая новые основные переменные через новые неосновные, получаем

$$\begin{cases} x_2 = 1 + x_1 + x_3, \\ x_4 = 2 - x_3, \\ x_5 = 3 - x_1. \end{cases}$$

и базисное решение $X_2 = (0;1;0;2;3)$ является допустимым. Поэтому выражаем через неосновные переменные линейную функцию $F = x_1 + x_2 = 1 + 2x_1 + x_3$ и продолжаем решение в соответствии с алгоритмом, изложенным в разд. 5.2.

5.5. Особые случаи симплексного метода

Рассмотрим особые случаи, которые могут возникнуть при решении задачи линейного программирования симплексным методом.

Неединственность оптимального решения (альтернативный оптимум)

Задача. Решить симплексным методом

$$F = 3x_1 + 3x_2 \to \max$$

при ограничениях

$$\begin{cases} x_1 + x_2 \le 8, \\ 2x_2 - x_2 \ge 1, \\ x_1 - 2x_2 \le 2, \end{cases}$$

$$x_1 \ge 0, \quad x_2 \ge 0.$$

Решение. Геометрическое решение задачи показывает, что оптимум достигается в любой точке отрезка AB (см. рис 4.3), так как линия уровня параллельна этому отрезку. Посмотрим, как проявится альтернативный оптимум при решении задачи симплексным методом.

На очередном шаге получаем: основные переменные: x_1, x_2, x_5, неосновные переменные: x_3, x_4.

Выражаем основные переменные через неосновные:

$$\begin{cases} x_1 = 5 - \dfrac{2}{3}x_3 - \dfrac{1}{3}x_4, \\ x_2 = 3 - \dfrac{1}{3}x_3 + \dfrac{1}{3}x_4, \\ x_5 = 9 - x_3 - x_4. \end{cases}$$

$X_1 = (3;5;0;0;9)$ – допустимое базисное решение, соответствующее угловой точке $A(3;5)$. Линейная функция $F = 24 - x_3$. В этом выражении отсутствуют положительные коэффициенты при неосновных переменных, значит, критерий оптимальности выполнен, X_1 – оптимальное базисное решение, $F_{\max} = F(X_1) = 24$. Однако в последнем выражении для F отсутствует неосновная переменная x_4 (формально входит с нулевым коэффициентом), поэтому изменение этой переменной не повлечет за собой изменение линейной функции. Например, можно перевести в основную переменную x_4; $x_4 = \min\{15;\infty;9\} = 9$. Переменная x_5 перейдет в неосновные, однако изменения линейной функции не произойдет: $\Delta F = 9 \cdot 0 = 0$. Из системы уравнений можно получить все множество оптимальных решений задачи. Положим, $\Delta F < 0$, где $t \in [0;9]$. Тогда множество оптимальных решений: $x_1 = 3 + \frac{1}{3}t$, $x_2 = 5 - \frac{1}{3}t$, $x_3 = 0$, A_1 $x_5 = 9 + t (t \in [0;9])$.

Замечание. В соответствии с теоремами 3.3 и 3.4 *множество оптимальных решений можно представить как выпуклую линейную комбинацию* X *базисных решений* $X_1 = (3;5;0;0;9)$ и $X_2 = (6;2;0;9;0)$, т.е. имеем: $X = \alpha X_1 + (1-\alpha) X_2$, где $0 \le \alpha \le 1$.

Глава 6

ДВОЙСТВЕННЫЕ ЗАДАЧИ

Каждой задаче линейного программирования соответствует другая задача, называемая *двойственной* или *сопряженной* по отношению к исходной. Теория двойственности оказалась полезной для проведения качественных исследований задач линейного программирования.

6.1. Экономическая интерпретация задачи двойственной задаче об использовании ресурсов

В главе 1 рассмотрена задача об использовании ресурсов. Экономико-математическая модель и содержательная интерпретация задачи 1 представлена в левой части табл. 6.1.

Таблица 6.1

Задача 1 (исходная)	Задача 2 (двойственная)
$F = c_1x_1 + c_2x_2 + \ldots + c_nx_n \to \max \qquad (6.1)$ при ограничениях $\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m, \end{cases} \qquad (6.2)$ и условии неотрицательности $x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0. \qquad (6.3)$ *Составить такой план выпуска продукции* $X = (x_1, x_2, \ldots, x_n)$, *при котором прибыль от реализации продукции будет максимальной при условии, что потребление ресурсов по каждому виду продукции не превзойдет имеющихся запасов*	$Z = b_1y_1 + b_2y_2 + \ldots + b_my_m \to \min \qquad (6.4)$ при ограничениях $\begin{cases} a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_1 \ge c_1, \\ a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_2 \ge c_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m \ge c_n, \end{cases} \qquad (6.5)$ и условии неотрицательности $y_1 \ge 0, y_2 \ge 0, \ldots, y_m \ge 0. \qquad (6.6)$ *Найти такой набор цен ресурсов* $Y = (y_1, y_2, \ldots, y_m)$, *при котором общие затраты на ресурсы будут минимальными при условии, что затраты на ресурсы при производстве каждого вида продукции будут не менее прибыли от реализации этой продукции*

В приведенной модели b_i $(i = 1, 2, \ldots, m)$ означает запас ресурса S_i; a_{ij} – число единиц ресурса S_i, потребляемого при производстве единицы продукции P_j $(j=1,2,\ldots,n)$; c_j – прибыль от реализации единицы продукции P_j (или цена продукции P_j).

Предположим, что некоторая организация решила закупить ресурсы $S_1, S_2, \ldots, S_m$ предприятия и необходимо установить оптимальные цены на эти ресурсы $y_1, y_2, \ldots, y_m$.

Очевидно, что покупающая организация заинтересована в том, чтобы затраты на все ресурсы Z в количествах $b_1, b_2, \ldots, b_m$ по ценам соответственно $y_1, y_2, \ldots, y_m$ были минимальными, т.е.

$$Z = b_1 y_1 + b_2 y_2 + \ldots + b_m y_m \to \min.$$

С другой стороны, предприятие, продающее ресурсы, заинтересовано в том, чтобы полученная выручка была не менее той суммы, которую предприятие может получить при переработке ресурсов в готовую продукцию. На изготовление единицы продукции P_1 расходуется a_{11} единиц ресурса S_1, a_{21} – единиц ресурса $S_2, \ldots$, a_{i1} единиц ресурса – $S_i, \ldots$, a_{m1} единиц ресурса – S_m по цене соответственно $y_1, y_2, \ldots, y_m$. Поэтому для удовлетворения требований продавца затраты на ресурсы, потребляемые при изготовлении единицы продукции P_1, должны быть не менее ее цены c_1, т.е.

$$a_{11} y_1 + a_{12} y_2 + \ldots + a_{m1} y_m \ge c_1.$$

Аналогично можно составить ограничения в виде неравенств по каждому виду продукции $P_1, P_2, \ldots, P_n$. Экономико-математическая модель и содержательная интерпретация полученной двойственной задачи приведены в табл. 6.1а.

*Таблица 6.1*а

<table>
<tr><th>Задача 1 (исходная)</th><th>Задача 2 (двойственная)</th></tr>
<tr><td>$F = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \to \max$ (6.1)
при ограничениях
$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m \end{cases}$ (6.2)
и условии неотрицательности
$x_1 \ge 0, x_2 \ge 0, \ldots x_n \ge 0.$ (6.3)</td><td>$Z = b_1 y_1 + b_2 y_2 + b_m y_m \to \min$ (6.4)
при ограничениях
$\begin{cases} a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_1 \ge c_1; \\ a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_2 \ge c_2; \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{nm}y_m \ge c_n, \end{cases}$ (6.5)
и условии неотрицательности
$y_1 \ge 0, y_2 \ge 0, \ldots, y_m \ge 0.$ (6.6)</td></tr>
</table>

*Окончание табл. 6.1*а

Задача 1 (исходная)	Задача 2 (двойственная)
Составить такой план выпуска продукции $X=(x_1,x_2,\dots,x_n)$, *при котором прибыль (выручка) от реализации продукции будет максимальной при условии, что потребление ресурсов по каждому виду продукции не превзойдет имеющихся запасов*	*Найти такой набор цен (оценок) ресурсов* $Y=(y_1,y_2,\dots,y_m)$, *при котором общие затраты на ресурсы будут минимальными при условии, что затраты на ресурсы при производстве каждого вида продукции будут не менее прибыли (выручки) от реализации этой продукции*

Цены ресурсов $y_1,y_2,\dots,y_m$ в экономической литературе получили различные названия: *учетные, неявные, теневые.* Смысл этих названий состоит в том, что это условные, ненастоящие цены. Цены $c_1,c_2,\dots,c_n$ называются внешними и известны до начала производства; цены ресурсов $y_1,y_2,\dots,y_m$ называются внутренними, определяются в результате решения задачи и часто их называют оценками ресурсов.

6.2. Взаимно двойственные задачи линейного программирования и их свойства

Рассмотрим формально две задачи 1 и 2 линейного программирования, представленные в табл. 6.1. Обе задачи обладают следующими свойствами:

1. В одной задаче ищут максимум линейной функции, в другой – минимум.

2. Коэффициенты при переменных в линейной функции одной задачи являются свободными членами ограничений в другой.

3. Каждая из задач задана в стандартной форме, причем в задаче максимизации все неравенства вида «$\le$», а в задаче минимизации – все неравенства вида «$\ge$».

4. Матрицы коэффициентов при переменных в системах ограничений обеих задач являются транспонированными друг к другу:

$$\text{для задачи 1 } A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

$$\textit{для задачи 2 } A' = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

5. Число неравенств в системе ограничений одной задачи совпадает с числом переменных в другой задаче.

6. Условия неотрицательности переменных имеются в обеих задачах.

Две задачи 1 и 2 линейного программирования, обладающие указанными свойствами, называются ***симметричными взаимно двойственными задачами*** и образуют последовательность решений:

1. Привести все неравенства системы ограничений исходной задачи к одному смыслу: если в исходной задаче ищут максимум линейной функции, то все неравенства системы ограничений привести к виду «$\le$», а если минимум – к виду «$\ge$». Для этого неравенства, в которых данное требование не выполняется, умножить на –1.

2. Составить расширенную матрицу системы A_1, в которую включить матрицу коэффициентов при переменных A, столбец свободных членов системы ограничений и строку коэффициентов при переменных в линейной функции.

3. Найти матрицу A_1', транспонированную к матрице A_1.

4. Сформулировать двойственную задачу на основании полученной матрицы A_1' и условия неотрицательности переменных.

Составить задачу, двойственную следующей задаче:

$$F = -x_1 + 2x_2 \to \max$$

при ограничениях

$$\begin{cases} 2x_1 - x_2 \ge 1, \\ -x_1 + 4x_2 \le 24, \\ x_1 - x_2 \le 3, \\ x_1 + x_2 \ge 5, \\ x_1 \ge 0, \quad x_2 \ge 0. \end{cases}$$

Решение. 1. Исходная задача на максимизацию, поэтому приведем все неравенства системы ограничений к виду «$\le$», для этого умножим обе части первого и четвертого неравенств на –1. Получим

$$\begin{cases} -2x_1 + x_2 \le -1, \\ -x_1 + 4x_2 \le 24, \\ x_1 - x_2 \le 3, \\ -x_1 - x_2 \le -5. \end{cases}$$

2. Составим расширенную матрицу системы:

$$A_1 = \begin{pmatrix} -2 & 1 & -1 \\ -1 & 4 & 24 \\ 1 & -1 & 3 \\ -1 & -1 & -5 \\ -1 & 2 & F \end{pmatrix}$$

3. Найдем матрицу A_1', транспонированную к матрице A_1:

$$A_1' = \begin{pmatrix} -2 & -1 & 1 & -1 & -1 \\ 1 & 4 & -1 & -1 & 2 \\ -1 & 24 & 3 & -5 & Z \end{pmatrix}.$$

4. Сформулируем двойственную задачу:

$$Z = -y_1 + 24y_2 + 3y_3 - 5y_4 \to \min$$

при ограничениях

$$\begin{cases} 2y_1 - y_2 + y_3 - y_4 \ge -1, \\ y_1 + 4y_2 - y_3 - y_4 \ge 2, \\ y_1 \ge 0, y_2 \ge 0, y_3 \ge 0, y_4 \ge 0. \end{cases}$$

6.3. Первая теорема двойственности

Связь между оптимальными решениями двойственных задач устанавливается с помощью теорем двойственности. Рассмотрим вспомогательное утверждение.

Основное неравенство теории двойственности. Пусть имеется пара двойственных задач 1 и 2. Покажем, что для любых допустимых решений $X = (x_1, x_2, \dots, x_n)$ и $Y = (y_1, y_2, \dots, y_m)$ исходной и двойственной задач справедливо неравенство

$$F(X) \le Z(Y) \text{ или } \sum_{j=1}^{n} c_j x_j \le \sum_{i=1}^{m} b_i y_i. \qquad (6.7)$$

Доказательство. Умножим неравенства системы ограничений (6.2) исходной задачи соответственно на переменные $y_1, y_2, \ldots, y_i, \ldots, y_m$ и сложим правые и левые части полученных неравенств, получим

$$\sum_{i=1}^{m} y_j \sum_{j=1}^{n} a_{ij} y_i \le \sum_{i=2}^{n} b_j y_j, \qquad (6.8)$$

Аналогично преобразуем систему ограничений (6.5) двойственной задачи, умножив ее на $x_1, x_2, ..., x_j, ..., x_n$, и сложив неравенства, получим

$$\sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} y_i \ge \sum_{j=1}^{n} c_j x_j. \qquad (6.9)$$

Так как левые части неравенств (6.8) и (6.9) равны, получим доказываемое неравенство (6.7).

Рассмотрим признаки оптимальности решений.

Достаточный признак оптимальности.

Теорема. *Если* $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ *и* $Y^* = (y_1^*, y_2^*, \ldots, y_m^*)$ *–- допустимые решения взаимно двойственных задач, для которых выполняется равенство*

$$F(X^*) = Z(Y^*), \qquad (6.10)$$

то X^* *– оптимальное решение исходной задачи 1, а* Y^* *– оптимальное решение двойственной задачи 2.*

Доказательство. Пусть X_1 – любое допустимое решение исходной задачи 1. Тогда на основании основного неравенства (6.7) получим $F(X_1) \le Z(Y^*)$. Однако X_1 – произвольное решение задачи 1, поэтому в силу равенства (6.10) выполняется неравенство $F(X_1) \le F(X^*)$, т.е. X^* – оптимальное решение задачи 1. Аналогично доказывается, что решение оптимально для задачи 2.

Кроме достаточного признака оптимальности взаимно двойственных задач существуют и другие важные соотношения между их решениями. Возникают следующие вопросы: всегда ли для каждой пары двойственных задач есть одновременно оптимальные решения; возможно ли, что одна из двойственных задач имеет решение, а другая нет. На эти вопросы отвечает следующая теорема.

Первая (основная) теорема двойственности. *Если одна из взаимно двойственных задач имеет оптимальное решение, то его* *имеет и другая, причем оптимальные значения их линейных функций равны:*

$$F_{\max} = Z_{\min} \text{ или } F(X^*) = Z(Y^*). \qquad (6.11)$$

Если линейная функция одной из задач не ограничена, то условия другой задачи противоречивы.

Экономический смысл первой (основной) теоремы двойственности. *План производства* $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ *и набор цен (оценок) ресурсов* $Y^* = (y_1^*, y_2^*, \ldots, y_m^*)$ *оказываются оптимальными тогда и только тогда, когда прибыль (выручка) от продукции, найденная при «внешних» (известных заранее) ценах* $c_1, c_2, \ldots, c_n$, *равна затратам на ресурсы по «внутренним» (определяемым только из решения задачи) ценам* $y_1, y_2, \ldots, y_m$.

Для всех других планов X и Y обеих задач в соответствии с основным неравенством (6.7) теории двойственности прибыль (выручка) от продукции всегда меньше (или равна) затрат на ресурсы.

Экономический смысл первой теоремы двойственности можно интерпретировать так: предприятию безразлично, производить продукцию по оптимальному плану $X^* = (x_1^*, x_2^*, \ldots, x_n^*)$ и получить максимальную прибыль $F_{\max}$ или продавать ресурсы по оптимальным ценам $Y^* = (y_1^*, y_2^*, \ldots, y_m^*)$ и возместить от продажи равные ей минимальные затраты на ресурсы $Z_{\min}$.

6.4. Вторая теорема двойственности

Пусть даны две взаимно двойственные задачи: 1 – (6.1)–(6.3) и задача 2 – (6.4)–(6.6). Если каждую из этих задач решать симплексным методом, то необходимо их привести к каноническому виду. Для этого в систему ограничений (6.2) задачи 1 следует ввести m неотрицательных переменных $x_{n+1}, x_{n+2}, \ldots, x_{n+i}, \ldots, x_{n+m}$, а в систему ограничений (6.5) задачи 2 – n неотрицательных переменных $y_{m+1}, y_{m+2}, \ldots, y_{m+j}, \ldots, y_{n+m}$.

Системы ограничений каждой из взаимно двойственных задач примут вид

$$\sum_{j=1}^{n} a_{ij} x_j + x_{n+1} = b_i, \quad i = 1, 2, \ldots, m; \qquad (6.12)$$

$$\sum_{j=1}^{m} a_{ij} y_i - y_{m+j} = c_j, \quad j = 1, 2, \ldots, n. \tag{6.13}$$

Установим соответствие между первоначальными переменными одной из двойственных задач и дополнительными переменными другой задачи (табл. 6.2).

Таблица 6.2

<table>
<tr><td colspan="2">Переменные исходной задачи 1</td></tr>
<tr><td>Первоначальные</td><td>Дополнительные</td></tr>
<tr><td>$\begin{array}{ccccccc} x_1 & x_2 & \ldots & x_j & \ldots & x_n \\ \updownarrow & \updownarrow & & \updownarrow & & \updownarrow \\ y_{m+1} & y_{m+2} & \ldots & y_{m+j} & \ldots & y_{m+n} \end{array}$</td><td>$\begin{array}{ccccccc} x_{n+1} & x_{n+2} & \ldots & x_{n+i} & \ldots & x_{n+m} \\ \updownarrow & \updownarrow & & \updownarrow & & \updownarrow \\ y_1 & y_2 & \ldots & y_j & \ldots & y_m \end{array}$ (6.14)</td></tr>
<tr><td>Дополнительные</td><td>Первоначальные</td></tr>
<tr><td colspan="2">Переменные двойственной задачи 2</td></tr>
</table>

Теорема. *Положительным (ненулевым) компонентам оптимального решения одной из взаимно двойственных задач соответствуют нулевые компоненты оптимального решения другой задачи, т.е. для любых* $i = 1, 2, \ldots, m$ *и* $j = 1, 2, \ldots, n$*: если* $x_j^* > 0$, *то* $y_{m+j}^* = 0$; *если* $x_{n+i}^* > 0$, *то* $y_i^* = 0$, *и аналогично, если* $y_i^* > 0$, *то* $x_{n+i}^* = 0$; *если* $y_{m+j}^* > 0$, *то* $x_j^* = 0$.

Из этой теоремы следует важный вывод о том, что введенное *соответствие* (6.14) *между переменными взаимно двойственных задач при достижении оптимума представляет соответствие между* ***основными*** *переменными одной из двойственных задач и* ***неосновными*** *переменными другой задачи, когда они допускают допустимые базисные решения.*

Рассмотренная теорема является следствием следующей теоремы.

Вторая теорема двойственности. *Компоненты оптимального решения двойственной задачи равны абсолютным значениям коэффициентов при соответствующих переменных линейной функции исходной задачи, выраженной через неосновные переменные ее оптимального решения.*

Замечание. *Если в одной из взаимно двойственных задач нарушается единственность оптимального решения, то решение двойственной задачи вырожденное.* Это связано с тем, что при нарушении единственности оптимального решения исходной задачи в выражении линейной функции ее оптимального решения через неосновные переменные отсутствует хотя бы одна из основных переменных.

С помощью теорем двойственности можно, решив симплексным методом исходную задачу, найти оптимум и оптимальное решение двойственной задачи.

Задача. Решить симплексным методом

$$F = -x_1 + 2x_2 \to \max$$

при ограничениях

$$\begin{cases} -2x_1 + x_2 \le -1, \\ x_1 - 4x_2 \ge -24, \\ x_1 - x_2 \le 3, \\ x_1 + x_2 \ge 5, \\ x_1 \ge 0, \quad x_2 \ge 0 \end{cases}$$

и найти оптимум и оптимальное решение двойственной задачи, используя теоремы двойственности.

Решение. Решая задачу симплексным методом, получим на последнем шаге решения: $F = 10 - \frac{2}{7}x_3 - \frac{3}{7}x_4$, т.е. $F_{\max} = 10$ при оптимальном базисном решении $X^* = (4; 7; 0; 0; 6; 6)$.

В данной задаче соответствие (6.14) между переменными примет вид

$$\begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ y_5 & y_6 & y_1 & y_2 & y_3 & y_4. \end{array}$$

На основании первой теоремы двойственности $Z_{\min} = F_{\max} = 10$. На основании второй теоремы двойственности оптимальное решение двойственной задачи $Y^* = \left(\frac{2}{7}; \frac{3}{7}; 0; 0; 0; 0\right)$, так как $y_1^* = \frac{2}{7}$, т.е. коэффициенту при соответствующей переменной x_3 в выражении линейной функции $F(x)$, $y_2^* = \frac{3}{7}$, т.е. коэффициенту при переменной x_4, $y_3^* = y_4^* = y_5^* = y_6^* = 0$ (из-за отсутствия соответствующих переменных

x_5, x_6, x_1, x_2 в выражении $F(x)$ их коэффициенты равны нулю). Итак, в двойственной задаче $Z_{\min} = 10$ при $Y^* = \left(\frac{2}{7}; \frac{3}{7}; 0; 0; 0; 0\right)$.

Метод, при котором вначале симплексным методом решается двойственная задача, а затем оптимум и оптимальное решение исходной задачи находятся с помощью теорем двойственности, называется ***двойственным симплексным методом***. Этот метод выгодно применять, когда первое базисное решение исходной задачи недопустимое или когда число ее ограничений m больше числа переменных n.

Глава 7

ТРАНСПОРТНАЯ ЗАДАЧА

7.1. Экономико-математическая модель транспортной задачи

Частным случаем линейного программирования является транспортная задача.

Построим экономико-математическую модель следующей задачи. Имеются три поставщика и четыре потребителя. Мощность поставщиков и спросы потребителей, а также затраты на перевозку единицы груза для каждой пары «поставщик-потребитель» сведены в таблицу поставок (табл. 7.1).

Таблица 7.1

Поставщики	Мощность поставщиков	Потребители и их спрос			
		1	2	3	4
		20	110	40	110
1	60	1 x_{11}	2 x_{12}	5 x_{13}	3 x_{14}
2	120	1 x_{21}	6 x_{22}	5 x_{23}	2 x_{24}
3	100	6 x_{31}	3 x_{32}	7 x_{33}	4 x_{34}

В левом верхнем углу произвольной (i, j)-клетки (i – номер строки, j – номер столбца) стоит *коэффициент затрат* – затраты на перевозку единицы груза от i-го поставщика к j-му потребителю.

Задача. *Найти объемы перевозок для каждой пары «поставщик-потребитель» так, чтобы:*

1) *мощности всех поставщиков были реализованы;*
2) *спросы всех потребителей были удовлетворены;*
3) *суммарные затраты на перевозку были минимальными.*

Решение. Искомый объем перевозки от i-го поставщика к j-му потребителю обозначим через x_{ij} и назовем *поставкой клетки с номерами* (i, j). Заданные мощности поставщиков и спросы потребителей накладывают ограничения на значения неизвестных x_{ij}. Так, например, объем груза, забираемого от 1-го поставщика, должен быть равен мощности этого поставщика – 60 единицам, т.е. $x_{11}+x_{12}+x_{13}+x_{14}=60$ (уравнение баланса по первой строке).Чтобы мощность каждого из поставщиков была реализована, необходимо составить уравнения баланса для каждой строки таблицы поставок, т.е.

$$\begin{cases} x_{11}+x_{12}+x_{13}+x_{14}=60, \\ x_{21}+x_{22}+x_{23}+x_{24}=120, \\ x_{31}+x_{32}+x_{33}+x_{34}=100. \end{cases} \tag{7.1}$$

Аналогично, чтобы спрос каждого из потребителей был удовлетворен, подобные уравнения баланса составляем для каждого столбца таблицы поставок:

$$\begin{cases} x_{11}+x_{21}+x_{31}=20, \\ x_{12}+x_{22}+x_{32}=110, \\ x_{13}+x_{23}+x_{33}=40, \\ x_{14}+x_{24}+x_{34}=110. \end{cases} \tag{7.2}$$

Очевидно, что объем перевозимого груза не может быть отрицательным, поэтому следует дополнительно предположить, что

$$x_{ij}\ge 0 \quad (i=1,2,3;\ j=1,2,3,4).$$

Суммарные затраты F на перевозку выражаются через коэффициенты затрат и поставки следующим образом:

$$F=1\cdot x_{11}+2x_{12}+5x_{13}+3x_{14}+1\cdot x_{21}+6x_{22}+5x_{23}+$$
$$+2x_{24}+6x_{31}+3x_{32}+7x_{33}+4x_{34}. \tag{7.3}$$

Дадим математическую формулировку задачи. *На множестве неотрицательных решений системы ограничений (7.1) и (7.2) найти такое решение* $X=(x_{11},x_{12},\ldots,x_{33},x_{44})$, *при котором линейная функция (7.3) принимает минимальное значение.*

Особенности экономико-математической модели транспортной задачи:

- *система ограничений есть система уравнений (транспортная задача задана в канонической форме);*
- *коэффициенты при переменных системы ограничений равны единице или нулю;*

• *каждая переменная входит в систему ограничений два раза: один раз – в систему (7.1) и один раз – в систему (7.2).*

Для математической формулировки транспортной задачи обозначим через c_{ij} коэффициенты затрат, через M_i – мощности поставщиков, через N_j – мощности потребителей, где $i=1,2,\dots,m$; $j=1,2,\dots,n$; m – число поставщиков, n – число потребителей. Тогда система ограничений примет вид

$$\sum_{j=1}^{n} x_{ij} = M_i,\ i=1,2,\dots,m, \tag{7.4}$$

$$\sum_{i=1}^{m} x_{ij} = N_j,\ j=1,2,\dots,n. \tag{7.5}$$

Система (7.4) включает в себя уравнение баланса по строкам, а система (7.5) – по столбцам таблицы поставок. Линейная функция в данном случае

$$F=\sum_{j=1}^{n}\sum_{i=1}^{m} c_{ij}x_{ij}. \tag{7.6}$$

Математическая формулировка транспортной задачи в общей постановке будет следующей: *на множестве неотрицательных решений системы ограничений (7.4),(7.5) найти такое решение* $X=(x_{11},x_{12},\dots,x_{ij},\dots,x_{mn})$*, при котором значение линейной функции (7.6) минимально.*

Произвольное допустимое решение $X=(x_{11},x_{12},\dots,x_{ij},\dots,x_{mn})$ системы ограничений (7.4), (7.5) называется *распределением поставок.*

Транспортная задача, приведенная в примере, обладает важной особенностью: суммарная мощность поставщиков равна суммарной мощности потребителей, т.е.

$$\sum_{i=1}^{m} M_i = \sum_{j=1}^{n} N_j. \tag{7.7}$$

Такие транспортные задачи называют *закрытыми.* В противном случае транспортная задача называется *открытой*.

Рассмотрим закрытую транспортную задачу. Являясь задачей линейного программирования, транспортная задача может быть решена симплексным методом. Однако специфическая форма системы ограничений данной задачи позволяет упростить обычный симплексный метод. Модификация симплексного метода применительно к транспортной задаче называется *распределительным методом.* По аналогии с общим случаем решение в нем осуществляется по шагам, и каждому шагу соответствует разбиение переменных на основные (базисные) и неосновные (свободные).

Число r основных переменных транспортной задачи равно рангу системы линейных уравнений (максимальному числу линейно независимых уравнений в системе ограничений).

Теорема 7.1. *Ранг r системы уравнений (7.4),(7.5) при условии (7.7) равен $m+n-1$.*

Основное следствие теоремы 7.1. *Число r основных (базисных) переменных закрытой транспортной задачи равно $m+n-1$, где m – число поставщиков, n – число потребителей.* Каждому разбиению переменных x_{ij} задачи на основные (базисные) и неосновные (свободные) соответствует базисное решение и заполнение таблицы поставок. Другими словами, *распределение поставок называется базисным, если переменные, соответствующие заполненным клеткам, можно принять за основные переменные.* Поскольку в дальнейшем мы используем исключительно базисные распределения поставок, то термины «базисная клетка» и «заполненная клетка» будут считаться равнозначными.

Подобно тому, как это было в симплексном методе, в распределительном методе решения транспортной задачи будем переходить от одного базисного распределения поставок к другому в сторону невозрастания целевой функции вплоть до оптимального решения. Для начала такого движения потребуется исходное базисное распределение поставок – *опорный план.*

7.2. Нахождение первоначального базисного распределения поставок

Одним из возможных методов нахождения первоначального базисного распределения поставок является *метод «северо-западного угла».* Рассмотрим его на примере. Используем условия задачи из разд. 7.1 и найдем первоначальное базисное распределение поставок. Для этого дадим переменной x_{11} максимально возможную поставку в клетку (1,1) – «северо-западный угол» таблицы поставок: $x_{11}=\min\{60,20\}=20$. После этого спрос 1-го потребителя будет полностью удовлетворен, в результате чего первый столбец таблицы поставок выпадет из дальнейшего рассмотрения. В таблице поставок найдем новый «северо-западный угол» – клетку (1,2) и дадим в нее максимально возможное значение: $x_{12} = \min\{40,110\} = 40$. После этого мощность 1-го поставщика полностью реализована и из рассмотрения выпадет первая строка таблицы поставок. В полученной таблице снова находим «северо-западный угол» и т.д. В результате получаем исходное распределение поставок, представленное в табл. 7.2.

Таблица 7.2

	20	110	40	110
60	1 20	2 40	5	3
120	1	6 70	5 40	2 10
100	6	3	7	4 100

Число заполненных клеток в полученном распределении оказалось равным $m+n-1=3+4-1=6$, т.е. числу основных (базисных) переменных. Действительно, на каждом шаге из рассмотрения выпадали либо строка, либо столбец, а на последнем шаге и столбец, и строка. Поэтому число заполненных клеток на единицу меньше, чем сумма числа строк и столбцов таблицы поставок, т.е. равно $m+n-1$. Эта особенность шагов метода «северо-западного угла» служит причиной того, что полученное распределение является базисным.

Недостатком метода «северо-западного угла» является то, что он построен без учета значений коэффициентов затрат задачи. Однако данный метод допускает модификацию, лишенную этого недостатка. Она состоит в том, что на каждом шаге максимально возможную поставку следует давать не в «северо-западную» клетку оставшейся таблицы, а в клетку с наименьшим коэффициентом затрат. Такой метод получения опорного плана называется *методом наименьших затрат.*

Теорема 7.2. *Пусть на каждом шаге заполнения таблицы поставок возникает одна заполненная клетка, причем из рассмотрения на каждом (кроме последнего) шаге выпадает либо одна строка, либо один столбец. Тогда переменные, соответствующие заполненным клеткам, можно принять за базисные.*

Доказательство. Предположим, что переменные заполненных клеток, возникающие на первых t шагах метода, где $t-1,2,\ldots,m+n-2$, можно выразить через переменные, соответствующие свободным клеткам тех строк и столбцов, которые были вычеркнуты на первых t шагах. Пусть на $(t+1)$-м шаге метода заполнена (p,q)-я клетка и из рассмотрения выпала p-я строка. Выразим переменную x_{pq} из уравнения баланса по p-й строке:

$$x_{pq} = M_p - \sum_{\substack{j=1 \\ j \neq q}}^{n} x_{pj}.$$

Пусть среди переменных правой части последнего равенства есть переменные клеток, заполненные на одном из первых t шагов. Тогда их можно выразить через переменные свободных клеток тех строк и столбцов, которые были вычеркнуты на первых t шагах. В случае, если на $(t+1)$-м шаге из рассмотрения выпал q-й столбец, x_{pq} следует выразить из уравнения баланса по q-му столбцу. Подобные рассуждения следует последовательно провести для каждого шага заполнения таблицы поставок.

Из теоремы 7.2 следует, что методы «северо-западного угла» и наименьших затрат приводят к базисным распределениям поставок, если на каждом шаге из рассмотрения выпадает либо одна строка, либо один столбец.

7.3. Критерий оптимальности базисного распределения поставок

Оптимальность базисного решения была рассмотрена в главе, посвященной симплексному методу. Вначале следует выразить линейную функцию задачи через свободные переменные. Поскольку транспортная задача – задача на минимум, то оптимум достигнут только тогда, когда все коэффициенты при свободных переменных в выражении линейной функции неотрицательны. В транспортной задаче произвольная переменная x_{ij} отождествляется с содержимым соответствующей клетки (i,j) таблицы поставок. Коэффициент β_{ij} при свободной переменной x_{ij} в выражении линейной функции F через свободные переменные называется *оценкой свободной клетки* (i,j). Тогда **критерий оптимальности** формулируется так: *базисное* *распределение поставок оптимально тогда и только тогда, когда оценки всех свободных клеток неотрицательны.*

Таким образом, на первый план выходит задача о нахождении оценок свободных клеток для фиксированного базисного распределения поставок.

Пусть фиксировано некоторое базисное распределение поставок, при этом клетка (i, j) – свободная, β_{ij} – оценка клетки (i, j) или коэффициент при x_{ij} в выражении линейной функции F через свободные переменные, т.е.

$$F = F_0 + \beta_{ij} x_{ij} + \dots, \tag{7.8}$$

где многоточием обозначены слагаемые, отвечающие свободным переменным, отличным от x_{ij}; F_0 – суммарные затраты на перевозку данного распределения поставок. Из выражения (7.8) следует: *оценка* β_{ij} свободной клетки (i, j) равна приращению ΔF суммарных затрат на перевозку при переводе в клетку (i, j) единичной поставки. Очевидно, что $\Delta F > 0$, если $\beta_{ij} > 0$; $\Delta F < 0$, если $\beta_{ij} < 0$. Это определение оценки свободной клетки называют *экономическим смыслом оценки свободной клетки.*

Последовательность действий при решении произвольной транспортной задачи можно изложить в виде следующего алгоритма.

1. Для данного базисного распределения поставок подбираем потенциалы строк и столбцов таблицы поставок так, чтобы коэффициенты затрат заполненных клеток стали равны нулю. Составляем матрицу оценок.

2. Если оценки всех свободных клеток неотрицательны, то найденное распределение оптимально – решение закончено. Если среди оценок свободных клеток есть отрицательные, то выбираем одну из них для передачи в нее поставки (для определенности можно брать , например, одну из клеток с наименьшей оценкой).

3. Для избранной свободной клетки строится цикл пересчета. Поставка z, передаваемая по циклу, определяется как минимум среди поставок в клетках со знаком «–». Найденная поставка передвигается по циклу. При этом поставка в клетках со знаком «+» увеличивается на z, а в клетках со знаком «–» уменьшается на z. Клетка, поставка в которой при этом станет равной нулю, будет считаться свободной, остальные клетки цикла – заполненными. Таким образом, получено новое базисное распределение поставок.

РАЗДЕЛ 2

МОДЕЛИ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Глава 8

КЛАССИЧЕСКИЕ МЕТОДЫ ОПТИМИЗАЦИИ

8.1. Методы определения экстремумов

Многие экономические модели можно считать линейными только в первом приближении, при более детальном рассмотрении обнаруживается их нелинейность. Такие показатели, как прибыль, себестоимость, капитальные затраты на производство зависят от объема производства и расхода производства нелинейно. Возникает задача нелинейного программирования.

Рассмотрим класс нелинейных задач, которые относятся к *классическим методам оптимизации.* Допустим, что переменные не являются дискретными, $m < n$. Функции $\varphi_i(X)$ и $f(X)$ непрерывны и имеют частные производные, по крайней мере, второго порядка. В этом случае задачу оптимизации можно сформулировать так: *найти переменные* $x_1, x_2, \ldots, x_n$, *удовлетворяющие системе уравнений*

$$\varphi_i(x_1, x_2, \ldots, x_n) = b_i, \quad i = 1, 2, \ldots, m \tag{8.1}$$

и обращающие в максимум (минимум) целевую функцию

$$z = f(x_1, x_2, \ldots, x_n). \tag{8.2}$$

Такие задачи можно решать классическими методами дифференциального исчисления. Однако на этом пути встречаются такие вычислительные трудности, которые делают необходимым поиск других методов решения. Поэтому классические методы используются не в качестве вычислительного средства, а как основа для теоретического анализа.

Объем производства (выраженный в натуральных или стоимостных единицах) является функцией затрат производства $z = f(x_1, x_2)$. Эта зависимость называется ***производственной функцией.*** Издержки зависят от расхода обоих факторов (x_1 и x_2) и от цен этих факторов (c_1 и c_2). Требуется при данных совокупных издержках определить такое количество факторов производства, которое максимизирует объем продукции z.

Математическая модель этой задачи имеет вид: определить такие переменные x_1 и x_2, удовлетворяющие условиям

$$\begin{aligned} &c_1x_1 + c_2x_2 = b, \\ &x_1 \ge 0, \quad x_2 \ge 0, \end{aligned} \tag{8.3}$$

при которых функция

$$z = f(x_1, x_2) \tag{8.4}$$

достигает максимума.

Функция (8.4) может иметь произвольный нелинейный вид.

Используя классические методы оптимизации, следует представлять различие между *локальным* экстремумом функции, *глобальным* экстремумом и *условным* экстремумом. Понятие условного экстремума вводится для случая, когда число переменных n не меньше 2, $(n \ge 2)$.

Будем полагать, что функция $z = f(x_1, x_2, \ldots, x_n) = f(X)$ дважды дифференцируема в точке $X^* = (x_1^*, x_2^*, \ldots, x_n^*), (X^* \in D(f)$ и в некоторой ее окрестности. Если для всех точек X этой окрестности $f(X^*) \ge f(X)$ или $f(X^*) \le f(X)$, то говорят, что функция $f(X)$ имеет экстремум в точке X^*.

Точка X^*, в которой все частные производные функции $z = f(X)$ равны нулю, называется *стационарной точкой.*

Необходимое условие экстремума. *Если в точке* X^* функция $z = f(X)$ *имеет экстремум, то частные производные функции в этой точке равны нулю:*

$$f'_{x_i}(X^*) = 0, \quad i = 1, 2, \ldots, n.$$

Следовательно, точки экстремума функции $z = f(X)$ удовлетворяют системе уравнений:

$$\begin{cases} f'_{x_1}(x_1, x_2, \ldots, x_n) = 0, \\ f'_{x_2}(x_1, x_2, \ldots, x_n) = 0, \\ \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ f'_{x_n}(x_1, x_2, \ldots, x_n) = 0. \end{cases} \tag{8.5}$$

Необходимое условие не является достаточным для того, чтобы стационарная точка была точкой экстремума. Для получения достаточных условий следует определить в стационарной точке знак дифференциала второго порядка

$$d^2 f(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} f''_{x_i x_j}(X)\Delta x_i \Delta x_j.$$

Достаточные условия экстремума:

а) в стационарной точке X^0 функция $z = f(X)$ имеет максимум, если $d^2 f(X^0) < 0$, и минимум, если $d^2 f(X^0) > 0$, при любых Δx_i и Δx_j, не обращающихся в нуль одновременно;

б) если $d^2 f(X^0)$ может принимать в зависимости от Δx_i и Δx_j и положительные, и отрицательные значения, то в точке X^0 экстремума нет;

в) если $d^2 f(X^0)$ может обращаться в нуль не только при нулевых приращениях Δx_i и Δx_j, то вопрос об экстремуме остается открытым.

Для функции двух переменных $z = f(x_1, x_2)$ достаточные условия выглядят следующим образом. Находим значения частных производных второго порядка в стационарной точке $X^0(x_1^0, x_2^0)$:

$$a_{11} = f''_{x_1^2}(X^0);\ a_{12} = f''_{x_1 x_2}(X^0);\ a_{21} = f''_{x_2 x_1}(X^0);\ a_{22} = f''_{x_2^2}(X^0).$$

Легко убедиться, что для непрерывных функций $a_{12} = a_{21}$. Обозначим через Δ определитель, составленный из $a_{ij}\,(i, j = 1, 2)$:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Тогда достаточные условия экстремума функции двух переменных имеют вид:

а) если $\Delta > 0$ и $a_{11} < 0\ (a_{22} < 0)$, то в точке X^0 функция имеет максимум; если $\Delta > 0$, то в точке X^0 – минимум;

б) если $\Delta < 0$, то экстремума нет;

в) если $\Delta = 0$, то вопрос об экстремуме остается открытым.

Схема определения экстремума функции n переменных совпадает с правилами определения локального экстремума функции двух переменных.

В практических задачах необходимо определить наибольшее и наименьшее значения функции (глобальный экстремум) в некоторой области.

Функция $z = f(X)$ имеет в точке X^0 заданной области D *глобальный максимум (наибольшее значение)* или *глобальный минимум (наименьшее значение)*, если выполняется одно из неравенств $f(X) \le f(X^0)$ или $f(X) \ge f(X^0)$ соответственно для любой точки $X \in D$.

Справедлива теорема К. Т. Вейерштрасса.

Теорема. *Если область D замкнута и ограничена, то дифференцируемая функция $z = f(X)$ достигает в этой области своих наибольшего и наименьшего значений в стационарной точке или в граничной точке области.*

Следовательно, чтобы найти наибольшее (наименьшее) значение функции $z = f(X)$ в области D, нужно:

1) найти все стационарные точки внутри области D и вычислить значения функции в них;

2) исследовать функцию на экстремум на границе области D;

3) из полученных значений функции выбрать наибольшее (наименьшее).

Исследуя экстремальные свойства функции на границе, необходимо решить задачу определения условного экстремума.

Условный экстремум.

Необходимо найти экстремум функции $z = f(x_1, x_2, \dots, x_n)$ при условии, что переменные $x_1, x_2, \dots, x_n$ удовлетворяют уравнениям

$$\varphi_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, 2, \dots, m, \quad m < n. \tag{8.6}$$

Предполагается, что функции f и φ_i имеют непрерывные частные производные по всем переменным. Уравнения (8.6) называются *уравнениями связи.*

Функция $z = f(x_1, x_2, \dots, x_n)$ имеет в точке $X^0 = (x_1^0, x_2^0, \dots, x_n^0)$, удовлетворяющей уравнениям связи (8.6), *условный максимум (минимум)*, если выполняется неравенство $f(X) \ge f(X^0)$ ($f(X) \le f(X^0)$).

Легко заметить, что задача определения условного экстремума совпадает с задачей нелинейного программирования (8.1), (8.2).

Один из способов определения условного экстремума применяется в том случае, если из уравнений связи (8.6) m переменных можно выразить через оставшиеся $n-m$ переменных:

$$x_i = \psi_i(x_{m+1}, \ldots, x_n), \quad i = 1, 2, \ldots, m. \tag{8.7}$$

Подставив полученные выражения для x_i в функцию z, получим

$$z = f(\psi_1(x_{m+1}, ..., x_n), \ldots, \psi_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n),$$

или

$$z = F(x_{m+1}, \ldots, x_n). \tag{8.8}$$

Задача сведена к нахождению локального (глобального) экстремума для функции (8.8) от $n-m$ переменных.

8.2. Метод множителей Лагранжа

Второй способ определения условного экстремума начинается с построения вспомогательной функции Ж. Л. Лагранжа. В области допустимых решений функция Лагранжа достигает экстремума в тех же точках, что и целевая функция z.

Решим **задачу** определения условного экстремума функции $z = f(X)$ при ограничениях (8.6).

Составим функцию

$$L(X) = f(X) + \sum_{i=1}^{m} \lambda_i \varphi_i(X), \tag{8.9}$$

которая называется *функцией Лагранжа*; где λ_i – постоянные множители. Множителям Лагранжа можно придать экономический смысл. Если $f(x_1, x_2, \ldots, x_n)$ – доход, соответствующий плану $X = (x_1, x_2, \ldots, x_n)$, а функция $\varphi_i(x_1, x_2, \ldots, x_n)$ – издержки i-го ресурса, соответствующие этому плану, то λ_i – цена (оценка) i-го ресурса, характеризующая изменение экстремального значения целевой функции в зависимости от изменения размера i-го ресурса. $L(X)$ – функция $n+m$ переменных $(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$. Определение стационарных точек этой функции приводит к решению системы уравнений

$$\begin{cases} \dfrac{\partial L(X)}{\partial x_j} = 0, \quad j = 1, 2, \ldots, n, \\ \dfrac{\partial L(X)}{\partial \lambda_i} = 0, \quad i = 1, 2, \ldots, m. \end{cases} \tag{8.10}$$

Отметим, что $L'_{\lambda_i}(X) = \varphi_i(X)$, т.е. в систему (8.10) входят уравнения связи. Таким образом, задача нахождения условного экстремума функции $z = f(X)$ сводится к нахождению локального экстремума функции $L(X)$. Если стационарная точка найдена, то вопрос о существовании экстремума решается на основании достаточных условий экстремума – исследовании знака второго дифференциала $d^2L(X)$ в стационарной точке при условии, что переменные приращения Δx_j связаны соотношениями

$$\sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial x_j} \Delta x_j = 0, \quad i = 1, 2, \ldots, m, \tag{8.11}$$

полученными путем дифференцирования уравнений связи.

Глава 9

МОДЕЛИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

В этой главе рассматривается задача нелинейного программирования (8.1) и (8.2) при условии, что функции f и φ_i $(i=1,2,\ldots,m)$ являются выпуклыми.

9.1. Производная по направлению и градиент. Выпуклые функции

Производной $\dfrac{\partial F}{\partial l}$ *функции* $F(X)=F(x_1,\ldots,x_n)$ *по направлению* l *в точке* X называется предел

$$\frac{\partial F}{\partial l}=\lim_{\lambda\to+0}\frac{F(X+\lambda l)-F(X)}{\lambda}.$$

Направление l задается вектором $l=(l_1,\ldots,l_n)$.

Если функция F дифференцируема в точке X, то она имеет в этой точке производную по любому направлению l, которая выражается через частные производные по формуле

$$\frac{\partial F}{\partial l}=\frac{1}{|l|}\sum_{i=1}^{n}l_i\frac{\partial F}{\partial x_i}, \qquad (9.1)$$

где длина вектора $|l|=\sqrt{l_1^2+\ldots+l_n^2}$.

Абсолютная величина производной по направлению дает скорость изменения функции в этом направлении, а знак показывает характер изменения функции (возрастание или убывание).

Градиентом ∇F *функции* $F(X)$ *называется вектор, проекциями которого на координатные оси служат соответствующие частные производные:*

$$\nabla F=\left(\frac{\partial F}{\partial x_1},\frac{\partial F}{\partial x_2},\ldots,\frac{\partial F}{\partial x_n}\right).$$

Известно, что $\max\left(\dfrac{\partial F}{\partial l}\right)$ достигается тогда, когда направление l совпадает с направлением ∇F. По формуле (9.1) производная функции F по направлению градиента ∇F равна

$$\frac{\partial F}{\partial(\nabla F)}=\frac{1}{|\nabla F|}\sum_{i=1}^{n}\frac{\partial F}{\partial x_i}\cdot\frac{\partial F}{\partial x_i}=|\nabla F|.$$

Итак, в каждой точке X *направление градиента является направлением наибольшего возрастания функции*, а длина градиента равна наибольшей скорости возрастания функции в этой точке.

Множество точек называется *выпуклым*, если оно вместе с любыми своими двумя точками содержит и весь отрезок, соединяющий эти точки. Если $[a,b]$ – отрезок на числовой прямой и $x\in[a,b]$, то $x=\alpha a+(1-\alpha)b,\ 0\le\alpha\le 1$, или

$$z=\alpha_1 a+\alpha_2 b,\quad \alpha_1+\alpha_2=1,\quad \alpha_1\ge 0,\quad \alpha_2\ge 0. \qquad (9.2)$$

Справедливо и обратное: если выполняется (9.2), то $x\in[a,b]$. Выпуклое множество – это множество, которое вместе с любой парой своих точек a,b содержит и все точки x, для которых выполняется условие (9.2). Эти определения отрезка и выпуклого множества сохраняются для случая, когда a,b,x – точки n-мерного пространства.

Используя равенство (9.2), можно показать, что если M – выпуклое пространство, то $\sum_{i=1}^{r} t_i X_i\in M$ для любых точек $X_1,\dots,X_r\in M$ и любых действительных точек $t_i\ge 0$, удовлетворяющих условию $\sum_{i=1}^{r} t_i=1$.

Функция $F(X)=F(x_1,\dots,x_n)$*, определенная на выпуклом множестве* M n*-мерного пространства, называется* ***выпуклой*** *на этом множестве, если*

$$F(\alpha X_1+(1-\alpha)X_2)\le\alpha F(X_1)+(1-\alpha)F(X_2) \qquad (9.3)$$

для любых точек $X_1,X_2\in M$ *и любого числа* $\alpha\in[0,1]$.

Если в условии (9.2) изменить знак неравенства с ≤ на ≥, то получим определение ***вогнутой*** функции. Если в условии (9.3) неравенство выполняется как строгое, то функция называется ***строго выпуклой*** (или ***строго вогнутой***).

Для любой пары X_1,X_2 значений аргумента произвольную точку $X\in[X_1,X_2]$ можно задать в виде $X=\alpha X_1+(1-\alpha)X_2$. Согласно рис. 9.1 неравенство (9.3) описывает отрезок, соединяющий точки $(X_1;F(X_1))$ и $(X_2;F(X_2))$, расположенный не ниже графика функции на этом участке (для строго выпуклой функции этот отрезок лежит выше графика функции).

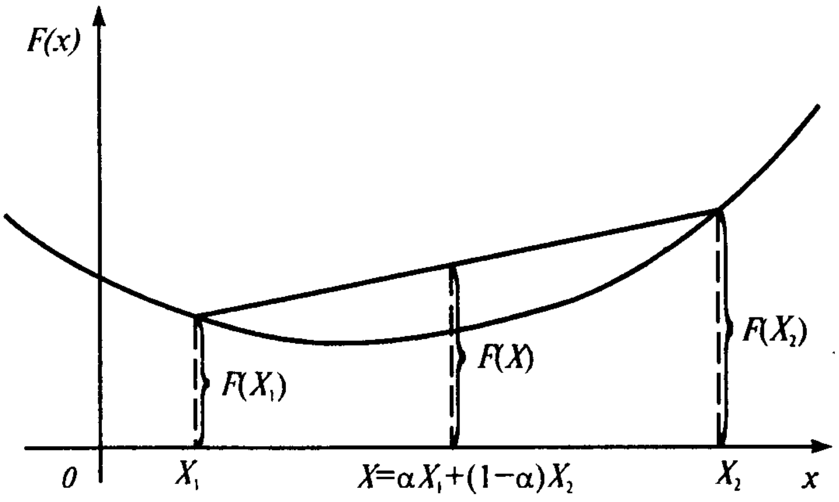


Рис. 9.1

Алгебраические и аналитические свойства выпуклых функций:

1. Если функция $F(X)$ выпукла, то функция $-F(X)$ вогнута.

2. Функция $F(x)=C$ и линейная функция $F(X)=aX+b$ являются всюду выпуклыми.

3. Если функции $F_i(X), i=1,\ldots,m$, выпуклы, то при любых действительных числах $\alpha_i \geq 0$ функция $\sum_{i=1}^{m}\alpha_i F_i(X)$ также является выпуклой.

4. Если функция $F(X)$ выпукла, то для любого числа α область решений неравенства $F(X)<\alpha$ является либо выпуклым множеством, либо пустым.

5. Если функции $\varphi_i(X)$ выпуклые при всех неотрицательных значениях переменных, то область решений системы неравенств $\varphi_i(X) \leq b_i, i=1,\ldots,m$, является выпуклым множеством (если она не пуста).

6. Выпуклая (вогнутая) функция, определенная на выпуклом множестве M, непрерывна в каждой внутренней точке этого множества.

7. Всякая дифференцируемая строго выпуклая (вогнутая) функция имеет не более одной стационарной точки (т.е. точки, в которой равны нулю все частные производные). При этом для выпуклой

(вогнутой) функции стационарная точка всегда является точкой локального и глобального минимума (максимума).

8. Дважды дифференцируемая функция $F(X)=F(x_1,\ldots,x)$ является выпуклой в том и только в том случае, когда

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 F(X)}{\partial x_i \partial x_j}\Delta x_i \Delta x_j \ge 0 \tag{9.4}$$

для любых $X\in M$ и $\Delta x_i, \Delta x_j$, не обращающихся в нуль одновременно.

Чтобы использовать это условие для определения выпуклости функции, часто применяют критерий Дж. Сильвестра: условие (9.4) выполняется тогда и только тогда, когда неотрицательны все главные миноры Δ_k матрицы вторых частных производных, т.е. определители

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}, \quad a_{ij}=\frac{\partial^2 F(X)}{\partial x_i \partial x_j}, \quad k=1,2,\ldots,n. \tag{9.5}$$

Если все $\Delta_k > 0$, то неравенство (9.4) выполняется как строгое, и тогда функция F является строго выпуклой.

9.2. Задача выпуклого программирования

Пусть дана система неравенств вида

$$\varphi_i(x_1,x_2,\ldots,x_n)\le b_i, \quad i=1,\ldots,m, \tag{9.6}$$

и функция

$$Z=f(x_1,x_2,\ldots,x_n), \tag{9.7}$$

причем все функции $\varphi_i(X)$ являются выпуклыми на некотором выпуклом множестве M, а функция Z либо выпукла на множестве M, либо вогнута. *Задача выпуклого программирования (ВП) состоит в отыскании такого решения системы ограничений (9.6), при котором целевая функция Z достигает минимального значения или вогнутая функция Z достигает максимального значения.*

Ввиду свойства 3 (см. разд. 9.1) всякая задача линейного программирования является частным случаем задачи ВП. В общем случае задачи ВП являются задачами нелинейного программирования.

Выделение задач ВП в специальный класс объясняется экстремальными свойствами выпуклых функций: всякий локальный минимум выпуклой функции (локальный максимум вогнутой функции) является одновременно и глобальным; кроме того, ввиду свойства 2 выпуклая (вогнутая) функция, заданная на замкнутом ограниченном множестве, достигает на этом множестве глобального максимума и глобального минимума. Отсюда вытекает, что *если целевая функция Z является строго выпуклой (строго вогнутой) и если область решений системы ограничений не пуста и ограничена, то задача ВП всегда имеет единственное решение.* В этом случае минимум выпуклой (максимум вогнутой) функции достигается внутри области решений, если там имеется стационарная точка, или на границе этой области, если внутри нее нет стационарной точки.

9.3. Приближенное решение задач выпуклого программирования методом кусочно-линейной аппроксимации

Функция $F(X) = F(x_1, \ldots, x)$ называется ***сепарабельной**, если ее можно представить в виде суммы функций, каждая из которых зависит только от одной переменной*, т.е. если

$$F(X) = \sum_{i=1}^{n} F_i(x_i). \tag{9.8}$$

Предположим, что в задаче ВП (9.6), (9.7) функция цели Z и все ограничения φ_i являются сепарабельными. Тогда задача имеет вид: *найти минимум выпуклой (максимум вогнутой) функции* $Z = \sum_{j=1}^{n} f_j(x_j)$ *при ограничениях*

$$\sum_{j=1}^{n} \varphi_{ij}(x_j) \le b_i, \quad i = 1, \ldots, m. \tag{9.9}$$

Идея метода кусочно-линейной аппроксимации состоит в том, что все f_i и все φ_{ij} заменяются ломаными линиями, состоящими из прямолинейных отрезков. При этом исходная задача ВП заменяется новой, приближенной задачей, которая является задачей линейного программирования. Эта задача решается симплексным методом и ее решение является приближенным решением исходной задачи ВП.

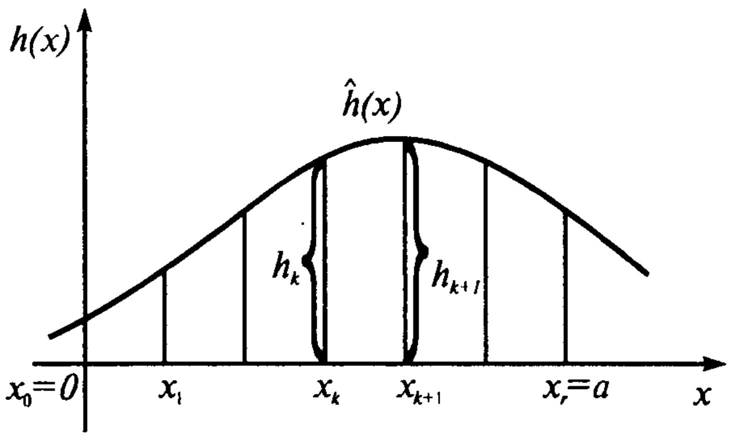


Рис. 9.2

Для построения приближенной задачи рассмотрим кусочно-линейную аппроксимацию функции одной переменной $h(x)$, заданной на отрезке $[0,a]$. Разобьем этот отрезок на r частей точками $x_0 < x_1 < \ldots < x_r$ так, чтобы $x_0 = 0, x_r = a$ (рис. 9.2). Вычислим значения функции $h_k(x)\,(k = 0,\ldots,r)$ в этих точках. Соединим попарно точки $(x_k; h_k)$ и $(x_{k+1}; h_{k+1})$ отрезками прямых. Состоящая из этих отрезков ломаная $\hat{h}(x)$ аппроксимирует функцию $h(x)$ на отрезке $[0,a]$. (Точность приближения можно увеличить за счет более мелкого разбиения отрезка).

Уравнение участка ломаной $\hat{h}(x)$ между точками $(x_k; h_k)$ и $(x_{k+1}; h_{k+1})$ имеет вид $\dfrac{\hat{h}(x) - h_k}{h_{k+1} - h_k} = \dfrac{x - x_k}{x_{k+1} - x_k}$. Если каждое из отношений в этом равенстве обозначить через λ, то получим:

$$\begin{cases} x = \lambda x_{k+1} + (1-\lambda)x_k, \\ \hat{h}(x) = \lambda h_{k+1} + (1-\lambda)h_k, \end{cases} \tag{9.10}$$

причем $0 \le \lambda \le 1$.

Для каждого $x \in [x_k, x_{k+1}]$ существует единственное значение λ, удовлетворяющее условиям (9.10). Обозначим $1-\lambda = \lambda_k, \lambda = \lambda_{k+1}$, перепишем систему (9.10) в виде

$$\begin{cases} x = \lambda_k x_k + \lambda_{k+1} x_{k+1}, \\ \hat{h}(x) = \lambda_k h_k + \lambda_{k+1} h_{k+1}, \end{cases} \tag{9.11}$$

где $\lambda_k + \lambda_{k+1} = 1, \quad \lambda_k \ge 0, \quad \lambda_{k+1} \ge 0.$

Уравнения (9.11) называются *параметрическими уравнениями отрезка.*

Итак, для любого $x \in [0, a]$ уравнение ломаной можно записать в виде

$$\begin{cases} x = \sum_{r=0}^{r} \lambda_k x_k, \\ \hat{h}(x) = \sum_{k=0}^{r} \lambda_k h_k, \quad \sum_{k=0}^{r} \lambda_k = 1, \quad \lambda_k \geq 0 \, (k = 0, \ldots, r), \end{cases} \tag{9.12}$$

причем всегда отличны от нуля два значения λ_k (если x является внутренней точкой k-го отрезка разбиения) или одно (если x совпадает с концом отрезка).

Решая задачу ВП с сепарабельными функциями, нужно прежде всего определить интервал изменения каждой переменной x_j. Затем каждый интервал разбить на части точками x_{jk} и с помощью формул (9.12) построить кусочно-линейную аппроксимацию функций f_j и φ_{ij}. После этого можно для исходной задачи (9.9) записать приближенную задачу:

найти максимум функции $\hat{Z} = \sum_{j=1}^{n} \hat{f}_j(x_j)$

при ограничениях $\sum_{j=1}^{n} \hat{\varphi}_{ij}(x_j) \leq b_i, \, i = 1, \ldots, m.$ (9.13)

$x_j \geq 0, \; j = 1, \ldots, n.$

Приближенная задача (9.13) является задачей линейного программирования и обычно решается симплексным методом. Отдельно записаны условия неотрицательности переменных.

Улучшить точность приближения можно, разбивая на более мелкие части не исходные отрезки изменения переменных, а меньшие, взятые в окрестности полученного первого приближения. Недостатком метода является большое увеличение размерности задачи (т.е. числа переменных) при переходе к приближенной линейной модели.

9.4. Методы спуска. Приближенное решение задач выпуклого программирования градиентным методом

Схема решения задач математического программирования методами спуска состоит в построении последовательности

$$X_0, X_1, \ldots, X_k, \ldots \qquad (9.14)$$

решений системы ограничений данной задачи по следующему принципу: в качестве X_0 выбирается любая точка области решений и затем каждая последующая точка получается из предыдущей по формуле

$$X_{k+1} = X_k + \lambda l, \qquad (9.15)$$

где $l = (l_1, \ldots, l_n)$ – некоторое направление, а λ – число. При этом направление l и длина шага λ выбираются так, чтобы обеспечить сходимость последовательности (9.13) к оптимальному решению X^*. В общем случае процесс получения последовательных приближений X_k бесконечен (тогда некоторое X_{k_0} берется за приближенное значение оптимального решения X^*), однако иногда процесс может завершиться за конечное число шагов, приводя к локальному, а в задачах ВП и глобальному оптимуму.

Находя производную по направлению $\frac{\partial Z}{\partial l}$, можно определить, является ли направление l «выгодным» или «невыгодным» в смысле приближения к оптимуму.

Так как направление градиента ∇Z целевой функции является направлением ее наискорейшего роста, то при отыскании максимума вогнутой функции (максимума выпуклой функции) в качестве l берется ∇Z $(-\nabla Z)$, и тогда формула (9.15) принимает вид

$$X_{k+1} = X_k + \lambda \cdot \nabla Z(X_k), \quad \lambda > 0 \quad (\text{если ищется } Z_{\max}) \qquad (9.16)$$

или

$$X_{k+1} = X_k - \lambda \cdot \nabla Z(X_k), \quad \lambda > 0 \quad (\text{если ищется } Z_{\min}). \qquad (9.16^*)$$

Методы спуска, в которых итерационная последовательность (9.14) находится по формуле (9.16) или (9.16*), называются *градиентными*. Друг от друга они отличаются способами выбора длины шага λ и алгоритмами нахождения точки X_{k+1}, если X_k находится на границе области решений и формула (9.16) выводит X_{k+1} за пределы этой области. Если величина λ выбирается так, чтобы приращение ΔZ при перемещении из точки X_k в точку X_{k+1} было наибольшим (при отыскании $Z_{\max}$) или наименьшим (при отыскании $Z_{\min}$), то градиентный метод называется *методом скорейшего спуска.* Таким образом, по методу скорейшего спуска длина шага λ в формуле (9.16) или (9.16*) выбирается так, чтобы при этом λ достигался экстремум функции $\Delta Z = Z(X_{r+1}) - Z(X_k)$. Продифференцировав функцию ΔZ с учетом выражения X_{k+1} по формуле (9.16) и выражения градиента в точке X_k, $\nabla Z(X_k) = \left(\frac{\partial Z}{\partial x_1}(X_k), \frac{\partial Z}{\partial x_2}(X_k), \ldots, \frac{\partial Z}{\partial x_n}(X_k) \right)$, получим, что необходимое условие экстремума примет вид

$$\frac{\partial Z}{\partial x_1}(X_{k+1})\frac{\partial Z}{\partial x_1}(X_k) + \frac{\partial Z}{\partial x_2}(X_{k+1})\frac{\partial Z}{\partial x_2}(X_k) + \ldots +$$

$$+\frac{\partial Z}{\partial x_n}(X_{r+1})\frac{\partial Z}{\partial x_n}(X_k) = 0. \qquad (9.17)$$

Этому равенству можно придать компактную форму, если использовать скалярное произведение векторов:

$$\nabla Z(X_{k+1}) \nabla Z(X_k) = 0. \qquad (9.17^*)$$

Известно, что скалярное произведение векторов равно 0 тогда и только тогда, когда они ортогональны.

Если оптимум достигается внутри области решений системы ограничений данной задачи ВП, то нет опасности, что точка X_{k+1}, найденная по формуле (9.16) или (9.16*), выйдет за пределы этой области, и длину шага λ определяем по формуле (9.17).

Для случая функции двух переменных метод скорейшего спуска имеет простую геометрическую интерпретацию: для любого k луч, идущий из точки X_k к точке X_{k+1}, перпендикулярен к линии уровня функции Z, проходящей через точку X_k (так как направлен по градиенту), и касается линии уровня, проходящей через точку X_{k+1} (так как ввиду условия (9.16*) он перпендикулярен к следующему лучу, который, в свою очередь, перпендикулярен к этой линии уровня). Скорейший спуск на плоскости происходит по двум взаимно перпендикулярным направлениям, как показано на рис. 9.3.

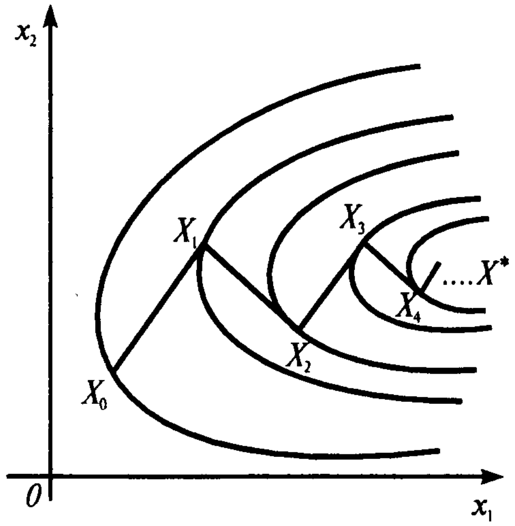


Рис. 9.3

Рассмотрим задачу ВП, когда *оптимум целевой функции достигается на границе области решений системы ограничений.* В качестве исходной точки X_0 можно взять любую точку из области решений, последующие точки находим по формуле (9.16) или (9.16*). На некотором шаге получим, что X_k уже не лежит в области решений (рис. 9.4,*а*). Тогда вместо X_k берем X'_k, которая лежит на пересечении направления спуска с границей области решений, а все последующие точки находятся путем проектирования на эту границу точек, получаемых обычным методом скорейшего спуска.

Рассмотрим случай, когда система ограничений линейная, т.е. область решений задачи для двух переменных ограничена отрезками прямых (рис. 9.4,*б*).

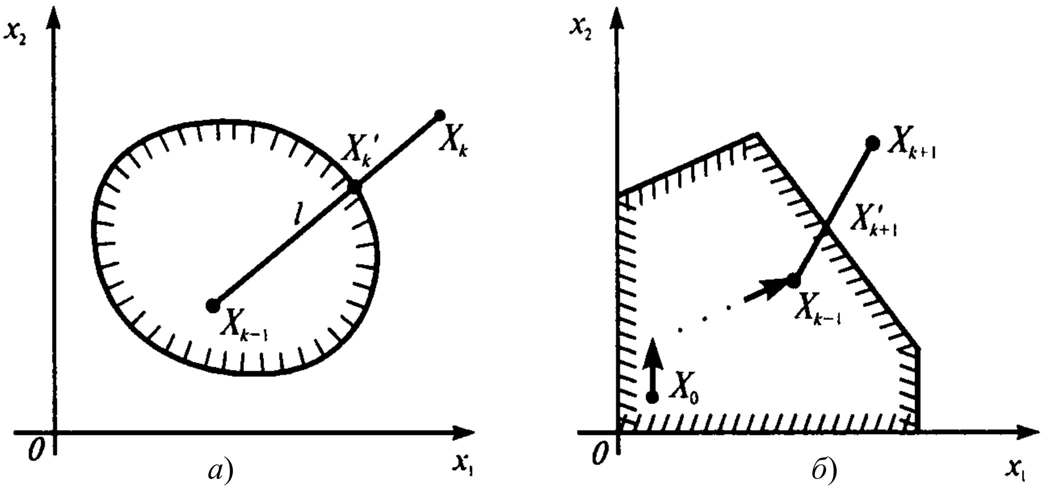


Рис. 9.4

В этом случае система ограничений данной задачи примет вид

$$\sum_{o=1}^{m} a_{ij}x_j \le b_i, \quad i=1,\dots,m. \tag{9.18}$$

Пусть по методу скорейшего спуска мы построили точки $X_0,\dots,X_k,X_{k+1}$ и убедились, что $X_0,\dots,X_k$ лежат в области решений, а точка X_{k+1} уже не лежит в ней. Значит, координаты точки X_{k+1} не удовлетворяют хотя бы одному неравенству системы (9.18). Пусть $X_k=(x_1^{(k)},\dots,x_n^{(k)}),\quad X_{k+1}=(x_1^{(k+1)},\dots,x_n^{(k+1)})$ и $X_{k+1}=X_k+\lambda l$, т.е.

$$x_j^{(k+1)}=x_j^{(k)}+\lambda l_j, \quad j=1,\dots,n. \tag{9.19}$$

Здесь $l=(l_1,\dots,l_n)$ – направление спуска, которое совпадает с направлением $\nabla Z(X_n)$ при отыскании максимума или с направлением $-\nabla Z(X_n)$ при отыскании минимума. Чтобы остаться в пределах области решений, следует вместо точки X_{k+1} взять точку, лежащую на том же направлении спуска, т.е. с координатами, удовлетворяющими условию (9.19), но с меньшей длиной шага λ. Значение λ нужно выбирать так, чтобы точка оказалась на границе области решений, чтобы выполнялось равенство $\sum_{j=1}^{n} a_{ij}x_j^{(k+1)}=b_j$. Учитывая (9.19), получаем $\sum_{j=1}^{n} a_{ij}(x_j^{(k)}+\lambda l_j)=b_j$ или $\sum_{j=1}^{n} a_{ij}x_j^{(k)}-b_i+\lambda\sum_{j=1}^{n} a_{ij}l_j=0$. Тогда

$$\lambda=-\frac{\sum_{j=1}^{n} a_{ij}x_j^{(k)}-b_i}{\sum_{j=1}^{n} a_{ij}l_j}. \tag{9.20}$$

Формула (9.20) дает значение λ, при котором направление спуска пересекает границу области решений. Если координаты точки X_{k+1} не удовлетворяют нескольким неравенствам системы (9.18), то нужно найти λ по формуле (9.20) для каждого из этих неравенств и взять наименьшее из найденных значений. Подставляя его в (9.19), найдем координаты точки X'_{k+1}, которая будет исходной для следующего шага решения. Так как проекция градиента $\nabla Z(X'_{k+1})$ задачи с линейными ограничениями лежит на границе области решений, а для случая двух переменных – просто на прямой $a_{i1}x_1 + a_{i2}x_2 = b_i$, то направление спуска l на следующем шаге берется на этой границе. Оптимум достигается в той точке, в которой градиент перпендикулярен границе области решений.

Замечание. Если прямая на плоскости имеет уравнение $Ax + By + C = 0$, то ее направление задается вектором $0 \le x_k \le s_{k-1}$.

9.5. Понятие о параметрическом и стохастическом программировании

Параметрическое программирование. В экономике нередко возникают задачи, в математических моделях которых коэффициенты линейной формы или системы ограничений не являются постоянными числами. Они меняются в зависимости от некоторых параметров.

Параметрическое программирование рассматривает экстремальные задачи с целевыми функциями и ограничениями, зависящими от параметров, разрабатывает методы нахождения оптимальных решений для совокупности значений параметров и изучает поведение оптимальных планов этих задач при изменении параметров.

Наиболее простой и хорошо изученной является *задача линейного параметрического программирования* с одним параметром, от которого зависят только коэффициенты. Эта задача состоит в максимизации линейной функции $F(t) = \sum_{j=1}^{n} (c_j + d_j t) x_j$ при ограничениях $\sum_{j=1}^{n} a_{ij} x_j = b_i \quad (i = 1, \ldots, m)$ и условии неотрицательности переменных $x_j \ge 0 \quad (j = 1, \ldots, n)$, где c_j, d_j, a_{ij}, b_i – заданные постоянные, а t – параметр $t \in [\alpha, \beta]$.

В результате решения задачи отрезок $[\alpha, \beta]$ разбивается на конечное число отрезков значений параметра t таким образом, чтобы для каждого из них максимальное значение линейной функции $F(t)$ достигалось в одной и той же вершине многогранника решений. Для каждого промежутка значений параметра находится оптимум и оптимальное решение.

Стохастическое программирование. Представляет собой совокупность методов решения оптимизационных задач вероятностного характера. Задача становится стохастической, если параметры целевой функции или системы ограничений рассматривать как случайные величины.

При решении стохастических задач проще всего найти средние значения всех случайных параметров и свести такие задачи к обычным, детерминированным задачам математического программирования. Однако такой подход не всегда эффективен, так как при некоторых реализациях случайных величин можно прийти к решению, далекому от оптимального.

Другой подход состоит в том, что на первом этапе устанавливается предварительный оптимальный план на основе решения детерминированной задачи, который и реализуется на этом этапе. На втором этапе этот план корректируется в соответствии с реальными статистическими характеристиками параметров.

Глава 10

МОДЕЛИ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

10.1. Общая постановка задачи динамического программирования

Динамическое программирование (ДП) – метод оптимизации, приспособленный к операциям, в которых процесс принятия решения может быть разбит на этапы (шаги). Начало развития ДП относится к 50-м годам XX века и связано с именем американского математика Р.Э. Беллмана.

Модели линейного программирования можно использовать в экономике для принятия крупномасштабных плановых решений в сложных ситуациях. Модели ДП применяются при решении задач значительно меньшего масштаба.

В реально функционирующих больших экономических системах еженедельно требуется принимать микроэкономические решения. Модели ДП ценны тем, что позволяют на основе стандартного подхода с использованием минимального вмешательства человека принимать решения.

Приведем *общую постановку задачи ДП*. Рассматривается управляемый процесс. В результате управления система S переводится из начального состояния s_0 в состояние $\hat{s}$. Предположим, что управление можно разбить на n шагов, т.е. решение принимается на каждом шаге, а управление, переводящее систему S из начального состояния в конечное, представляет собой совокупность n пошаговых управлений.

Обозначим через X_k управление на k-м шаге $(k = 1, 2, \ldots, n)$. Переменные X_k удовлетворяют некоторым ограничениям и называются *допустимыми*.

Пусть $X(X_1, X_2, \ldots, X_n)$ – управление, переводящее систему S из состояния s_0 в состояние $\hat{s}$. Обозначим через s_k состояние системы после k-го шага управления. Получаем последовательность состояний $s_0, s_1, \ldots, s_{k-1}, s_k, \ldots, s_{n-1}, s_n = \hat{s}$, которую изобразим кружками (рис. 10.1).

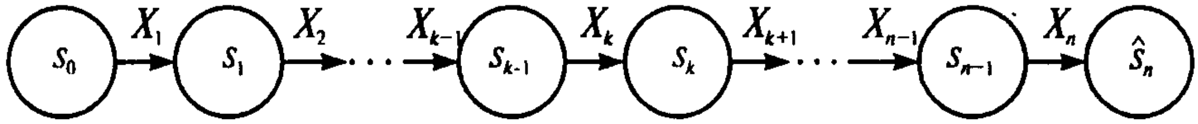


Рис. 10.1

Показатель эффективности рассматриваемой управляемой операции – целевая функция – зависит от начального состояния и управления:

$$Z = F(s_0, X). \tag{10.1}$$

Сделаем несколько предположений.

1. Состояние s_k системы в конце k-го шага зависит только от предшествующего состояния s_{k-1} и управления на k-м шаге X_k и не зависит от предшествующих состояний и управлений. Это требование называется «отсутствием последействия» и записывается в виде уравнений

$$s_k = \varphi_k(s_{k-1}, X_k), \quad k = 1, 2, \ldots, n, \tag{10.2}$$

которые называются *уравнениями состояния.*

2. Целевая функция (10.1) является аддитивной от показателя эффективности каждого шага. Обозначим показатель эффективности k- го шага через

$$Z_k = f_k(s_{k-1}, X_k), \quad k = 1, 2, \ldots, n, \tag{10.3}$$

тогда

$$Z = \sum_{k=1}^{n} f_k(s_{k-1}, X_k). \tag{10.4}$$

Задача пошаговой оптимизации (задача ДП) формируется так: *определить такое допустимое управление X, переводящее систему S из состояния s_0 в состояние $\hat{s}$, при котором целевая функция (10.4) принимает наибольшее (наименьшее) значение.*

Выделим особенности модели ДП:

1. *Задача оптимизации интерпретируется как n-шаговый процесс управления.*

2. *Целевая функция равна сумме целевых функций каждого шага.*

3. *Выбор управления на k-м шаге зависит только от состояния системы к этому шагу, не влияет на предшествующие шаги (нет обратной связи).*

4. *Состояние* s_k *после* k*-го шага управления зависит только от предшествующего состояния* s_{k-1} *и управления* X_k *(отсутствие последствия).*

5. *На каждом шаге управление* X_k *зависит от конечного числа управляющих переменных, а состояние* s_k *– от конечного числа параметров.*

Существуют различные способы решения подобных задач, применяемые в зависимости от вида функций, ограничений, размерности и т.д. Рассмотрим вычислительную схему ДП, которая окажется безразличной к способам задания функций и ограничений. Вычислительная схема связана с принципом оптимальности и использует рекуррентные соотношения.

10.2. Принцип оптимальности и уравнения Беллмана

Принцип оптимальности. Он был сформулирован Р. Беллманом в 1953 г. : к*аково бы ни было состояние системы* s *в результате какого-либо числа шагов, на ближайшем шаге нужно выбрать управление так, чтобы оно в совокупности с оптимальным управлением на всех последующих шагах приводило к оптимальному выигрышу на всех оставшихся шагах, включая данный.* Беллманом четко были сформулированы и условия, при которых принцип верен. Основное требование – процесс управления должен быть без обратной связи, т.е. управление на данном шаге не должно оказывать влияния на предшествующие шаги.

Принцип оптимальности утверждает, что для любого процесса без обратной связи оптимальное управление таково, что оно является оптимальным для любого подпроцесса по отношению к исходному состоянию этого подпроцесса. Поэтому решение на каждом шаге оказывается наилучшим с точки зрения управления в целом.

Уравнения Беллмана. Вместо исходной задачи ДП с фиксированным числом шагов n и начальным состоянием s_0 рассмотрим последовательность задач, полагая последовательно $n = 1, 2, \ldots$ при различных s и используя принцип оптимальности.

Введем ряд новых обозначений. На каждом шаге любого состояния системы s_{k-1} решение X_k нужно выбирать «с оглядкой», потому что этот выбор влияет на последующее состояние s_k и дальнейший процесс управления.

Однако есть один шаг, который можно для любого состояния s_{n-1} планировать локально-оптимально, исходя только из соображений этого шага.

Рассмотрим n-й шаг: s_{n-1} – состояние системы к началу n-го шага, $s_n = \hat{s}$ – конечное состояние, X_n – управление на n-м шаге. $f_n(s_{n-1}, X_n)$ – целевая функция (выигрыш) n-го шага.

Согласно принципу оптимальности, X_n нужно выбирать так, чтобы для любых состояний s_{n-1} получить максимум целевой функции на этом шаге.

Обозначим через $Z_n^*(s_{n-1})$ максимум целевой функции – показателя эффективности n-го шага при условии, что к началу последнего шага система S была в произвольном состоянии s_{n-1}, а на последнем шаге управление было оптимальным.

$Z_n^*(s_{n-1})$ называется *условным максимумом целевой функции на n-м шаге.* Очевидно, что

$$Z_n^*(s_{n-1}) = \max_{\{X_n\}} f_n(s_{n-1}, X_n). \qquad (10.5)$$

Максимизация ведется по всем допустимым управлениям X_n.

Решение X_n, при котором достигается $Z_n^*(s_{n-1})$, также зависит от s_{n-1} и называется *условным оптимальным управлением на n-м шаге*. Оно обозначается $X_n^*(s_{n-1})$.

Решив одномерную задачу локальной оптимизации по уравнению (10.5), найдем для всех возможных состояний s_{n-1} две функции: $Z_n^*(s_{n-1})$ и $X_n^*(s_{n-1})$.

Рассмотрим двухшаговую задачу: присоединим к n-му шагу $(n-1)$-й шаг (рис. 10.2).

Для любых состояний s_{n-2} произвольных управлений X_{n-1} и оптимальном управлении на n-м шаге значение целевой функции на двух последних шагах равно:

$$f_{n-1}(s_{n-2}, X_{n-1}) + Z_n^*(s_{n-1}). \qquad (10.6)$$

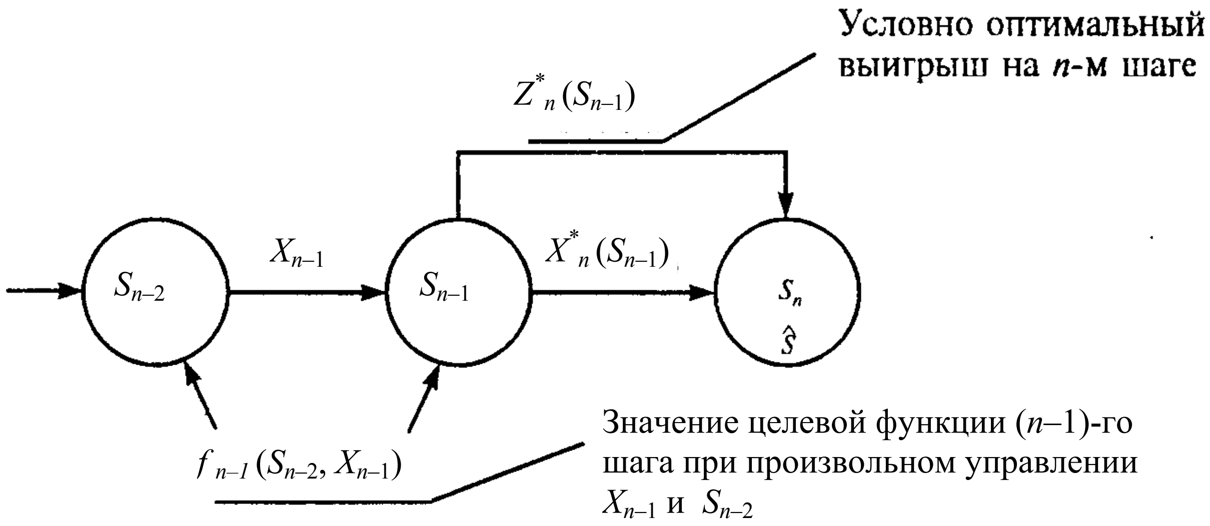


Рис. 10.2

Согласно принципу оптимальности для любых s_{n-2} решение нужно выбирать так, чтобы они вместе с оптимальным управлением на последнем $(n-м)$-м шаге приводилЬ бы к максимуму целевой функции на двух последних шагах. Следовательно, нужно найти максимум выражения (10.6) по всем допустимым управлениям X_{n-1}. Максимум этой суммы зависит от s_{n-2}, обозначается через $Z^*_{n-1}(s_{n-2})$ и называется *условным максимумом целевой функции при оптимальном управлении на двух последних шагах*. Соответствующее управление X_{n-1} на $(n-1)$-м шаге обозначается через $X^*_{n-1}(s_{n-2})$ и называется *условным оптимальным управлением на* $(n-1)$-м шаге:

$$Z^*_{n-1}(s_{n-2}) = \max_{\{X_{n-1}\}} \{f_{n-1}(s_{n-2}, X_{n-1}) + Z^*_n(s_{n-1})\}. \qquad (10.7)$$

Следует обратить внимание на то, что выражение, стоящее в фигурных скобках (10.7), зависит только от s_{n-2} и X_{n-1}, так как s_{n-1} можно найти из уравнения состояний (10.2) при $k = n-1$:

$$s_{n-1} = \varphi_{n-1}(s_{n-2}, X_{n-1})$$

и подставить вместо s_{n-1} в функцию $Z^*_n(s_{n-1})$.

В результате максимизации только по одной переменной X_{n-1} согласно уравнению (10.7) вновь получаются две функции:

$$Z^*_{n-1}(s_{n-2}) \text{ и } X^*_{n-1}(s_{n-2}).$$

Далее рассматривается трехшаговая задача: к двум последним шагам присоединяется $(n-2)$-й и т.д.

Обозначим через $Z_k^*(s_{k-1})$ *условный максимум целевой функции, полученный при оптимальном управлении на* $(n-k+1)$*-х шагах, начиная с* k*-го до конца, при условии, что к началу* k*-го шага система находилась в состоянии* s_{k-1}. Фактически эта функция равна

$$Z_k^*(s_{k-1}) = \max_{\{(x_k,\dots,x_n)\}} \sum_{i=k}^{n} f_i(s_{i-1}, X_i).$$

Тогда

$$Z_{k+1}^*(s_k) = \max_{\{(x_{k+1},\dots,x_n)\}} \sum_{i=k+1}^{n} f_i(s_{i-1}, X_i).$$

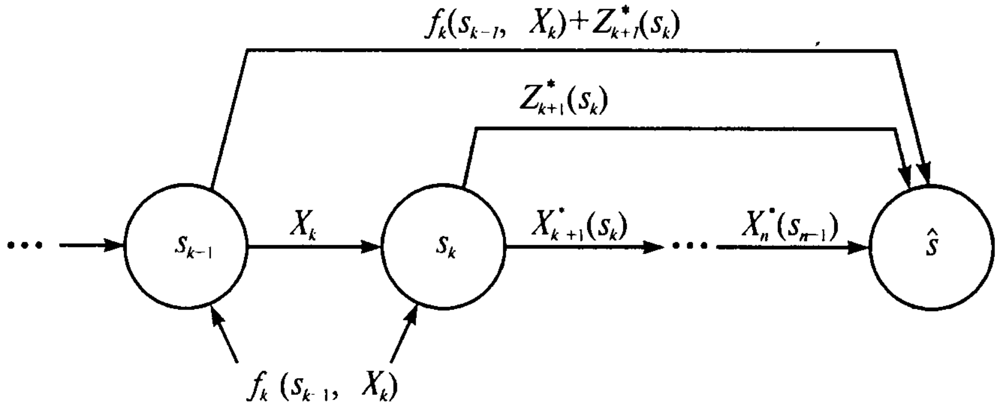


Рис. 10.3

Целевая функция на $(n-k)$-х последних шагах (рис. 10.3) при произвольном управлении X_k на k-м шаге и оптимальном управлении на последующих $(n-k)$-х шагах равна

$$f_k(s_{k-1}, X_k) + Z_{k+1}^*(s_k).$$

Согласно принципу оптимальности, X_k выбирается из условия максимума этой суммы, т.е.

$$Z_k^*(s_{k-1}) = \max_{\{X_k\}} \left\{ f_k(s_{k-1}, X_k) + Z_{k+1}^*(s_k) \right\}, \quad k = n-1, n-2, \dots, 2, 1. \quad (10.8)$$

Управление X_k на k-м шаге, при котором достигается максимум в (10.8), обозначается через $X_k^*(s_{k-1})$ и называется *условным* *оптимальным управлением на* k*-м шаге* (в правую часть уравнения (10.8) следует вместо s_k подставить выражение $s_k = \varphi_k(s_{k-1}, X_k)$, найденное из уравнения состояния).

Уравнения (10.8) называют *уравнениями Беллмана*. Это рекуррентные соотношения, позволяющие найти предыдущее значение функции, зная последующие. Если из (10.5) найти $Z_n^*(s_{n-1})$, то при $k = n-1$ из (10.8) можно определить, решив задачу максимизации для всех возможных значений s_{n-2}, выражения для $Z_{n-1}^*(s_{n-2})$ и соответствующее $X_{n-1}^*(s_{n-2})$. Далее, зная $Z_{n-1}^*(s_{n-2})$, находим, используя (10.8) и (10.2), уравнения состояний.

Процесс решения уравнений (10.5) и (10.8) называется *условной оптимизацией.* Описанный способ решения задачи ДП, начинающийся с последнего шага, называется «обратная схема». Если n-й и 1-й шаги поменять местами, то получится «прямая схема».

В результате условной оптимизации получаются две последовательности:

$$Z_n^*(s_{n-1}), Z_{n-1}^*(s_{n-2}), \ldots, Z_2^*(s_1), Z_1^*(s_0) \ –$$

условные максимумы целевой функции на последнем, на двух последних, на n шагах, и

$$X_n^*(s_{n-1}), X_{n-1}^*(s_{n-2}), \ldots, X_2^*(s_1), X_1^*(s_0) \ –$$

условные оптимальные управления на n-м, $(n-1)$-м, …, 1-м шагах.

Используя эти последовательности, можно найти решение задачи ДП при данных n и s_0. По определению $Z_1^*(s_0)$ – условный максимум целевой функции за n шагов при условии, что к началу первого шага система была в состоянии s_0, т.е.

$$Z_{\max} = Z_1^*(s_0). \qquad (10.9)$$

Далее следует использовать последовательность условных оптимальных управлений и уравнения состояний (10.2).

При фиксированном s_0 получаем $X_1^* = X_1^*(s_0)$. Далее из уравнений (10.2) находим $s_1^* = \varphi_1(s_0, X_1^*)$ и подставляем это выражение в последовательность условных оптимальных управлений $X_2^* = X_2^*(s_1^*)$ и так далее по цепочке:

$$X_1^* = X_1^*(s_0) \to s_1^* = \varphi_1(s_0, X_1^*) \to X_2^* = X_2^*(s_1^*) \to$$
$$\to s_2^* = \varphi_2(s_1^*, X_2^*) \Rightarrow X_3^* = X_3^*(s_2^*) \to \ldots \to$$
$$\to s_{n-1}^* = \varphi_{n-1}(s_{n-2}^*, X_{n-1}^*) \Rightarrow X_n^* = X_n^*(s_{n-1}^*).$$

Через s_k^* обозначено состояние системы после k-го шага при условии, что на k-м шаге выбрано оптимальное управление.

Получаем оптимальное решение задачи ДП:

$$X^* = (X_1^*, X_2^*, \ldots, X_n^*).$$

(Стрелка $\to$ означает использование уравнений состояния, а стрелка $\Rightarrow$ – последовательности условных оптимальных управлений).

10.3. Задача о распределении средств между предприятиями

Рассмотрим предложенную схему на конкретной задаче о распределении средств между предприятиями.

Планируется деятельность четырех промышленных предприятий (системы) на очередной год. Начальные средства: $s_0 = 5$ усл. ед. Размеры вложения в каждое предприятие кратны 1 усл. ед. Средства x, выделенные k-му предприятию $(k = 1, 2, 3, 4)$, приносят в конце года прибыль $f_k(x)$. Функции $f_k(x)$ заданы таблично (табл. 10.1). Принято считать, что:

1) прибыль $f_k(x)$ не зависит от вложения средств в другие предприятия;

2) прибыль от каждого предприятия выражается в одних условных единицах;

3) суммарная прибыль равна сумме прибылей, полученных от каждого предприятия.

Определить, какое количество средств нужно выделить каждому предприятию, чтобы суммарная прибыль была наибольшей.

Таблица 10.1

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
1	2	3	4	5
1	8	6	3	4
2	10	9	4	6

1	2	3	4	5
3	11	11	7	8
4	12	13	11	13
5	18	15	18	16

Решение. Обозначим через x_k количество средств, выделенных k-му предприятию. Суммарная прибыль равна

$$Z = \sum_{k=1}^{4} f_k(x_k). \tag{10.10}$$

Переменные x удовлетворяют ограничениям:

$$\sum_{k=1}^{n} x_k = 5, \quad x_k \geq 0, \quad k = 1,2,3,4. \tag{10.11}$$

Требуется найти переменные x_1, x_2, x_3, x_4, удовлетворяющие системе ограничений (10.11) и обращающие в максимум функцию (10.10).

Особенности модели. Ограничения линейные, но переменные целочисленные, а функции $f_k(x_k)$ заданы таблично, поэтому нельзя применить методы целочисленного линейного программирования.

Схема решения задачи методом ДП имеет следующий вид: процесс решения распределения средств $s_0 = 5$ можно рассматривать как 4-шаговый, номер шага совпадает с номером предприятия; выбор переменных x_1, x_2, x_3, x_4 – управление соответственно на 1,2,3, 4-м шагах; $\hat{s}$ – конечное состояние процесса распределения, равное нулю, так как все средства должны быть вложены в производство, $\hat{s} = 0$. Схема распределения показана на рис. 10.4.

Уравнения состояний (10.2) в данной задаче имеют вид

$$s_k = s_{k-1} - x_k, \quad k = 1,2,3,4. \tag{10.12}$$

где s_k – параметр состояния – количество средств, оставшихся после k-го шага, т.е. средства, которые остается распределить между оставшимися $4 - k$ предприятиями.

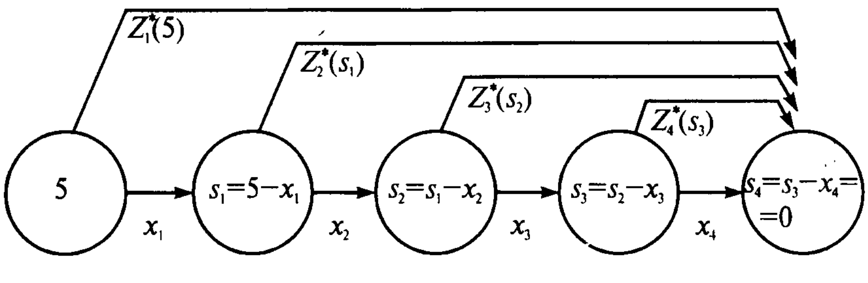


Рис. 10.4

Введем в рассмотрение функцию $Z_k^*(s_{k-1})$ – условную оптимальную прибыль, полученную от k-го, $(k+1)$-го, …, 4-го предприятия, если между ними распределялись оптимальным образом средства $s_{k-1}(0 \le s_{k-1} \le 5)$. Допустимые управления на k-м шаге удовлетворяют условию $0 \le x_k \le s_{k-1}$ (либо k-му предприятию ничего не выделяем, $x_k = 0$, либо не больше того, что имеем к k-му шагу, $x_k \le s_{k-1}$).

Уравнения (10.5)–(10.8) имеют вид:

$$k = 4, \quad s_4 = 0 \Rightarrow Z_4^*(s_3) = \max_{0 \le x_4 \le s_3} f_4(x_4), \qquad (10.5а)$$

$$Z_3^*(s_2) = \max_{0 \le x_3 \le s_2} \left\{ f_3(x_3) + Z_4^*(s_3) \right\}, \qquad (10.6а)$$

$$Z_2^*(s_1) = \max_{0 \le x_2 \le s_1} \left\{ f_2(x_2) + Z_3^*(s_2) \right\}, \qquad (10.7а)$$

$$Z_1^*(5) = \max_{0 \le x_1 \le 5} \left\{ f_1(x_1) + Z_2^*(s_1) \right\}. \qquad (10.8а)$$

Последовательно решаем записанные уравнения, проводя условную оптимизацию каждого шага.

4-й шаг. В табл. 10.1 $f_4(x)$ прибыли монотонно возрастают, поэтому все средства, оставшиеся к 4-му шагу, следует вложить в 4-е предприятие. При этом для возможных значений $s_3 = 0, 1, \ldots, 5$ получим:

$$Z_4^*(s_3) = f_4(s_3) \text{ и } x_4^*(s_3) = s_3.$$

3-й шаг. Делаем все предположения относительно остатка средств s_2 к 3-му шагу (т.е. после выбора x_1 и x_2). s_2 может принимать значения $0, 1, 2, 3, 4, 5$ (например, $s_2 = 0$, если все средства отданы 1-му и 2-му предприятиям, $s_2 = 5$, если 1-е и 2-е предприятия ничего не получили, и т.д.). В зависимости от этого выбираем $0 \le x_3 \le s_2$, находим $s_3 = s_2 - x_3$ и сравниваем для разных x_3 при фиксированном s_2 значения суммы $f_3(x_3) + Z_4^*(s_3)$. Для каждого s_2 наибольшее из этих значений есть $Z_3^*(s_2)$ – условная оптимальная прибыль, полученная при оптимальном распределении средств s_2 между 3-м и 4-м предприятиями. Оптимизация дана в табл. 10.2 при $k = 3$. Для каждого значения s_2 $Z_3^*(s_2)$ и $X_3^*(s_2)$ помещены в графах 5 и 6 соответственно.

2-й шаг. Условная оптимизация, согласно уравнению (10.7а), проведена в табл. 10.2 при $k = 2$. Для всех возможных значений s_1 значения $Z_2^*(s_1)$ и $X_2^*(s_1)$ находятся в столбцах 8 и 9 соответственно; первые слагаемые в столбце 7 – значения $f_2(x_2)$ взяты из табл. 10.1, а вторые слагаемые взяты из столбца 5 табл. 10.2 при $s_2 = s_1 - x_2$.

1-й шаг. Условная оптимизация (уравнение (10.8а)) проведена в табл. 10.2 при $k = 1$ для $s_0 = 5$. На 1-м шаге условной оптимизации достаточно заполнить раздел таблицы, соответствующий $s_0 = 5$. Поясним решение подробно: если $x_1 = 0$, то $s_1 = 5$, прибыль, полученная от четырех предприятий при условии, что $s_1 = 5$ единиц средств между оставшимися тремя предприятиями будут распределены оптимально, равна $f_1(0) + Z_2^*(5) = 0 + 19 = 19$ ($Z_2^*(5)$ взято из столбца 9 табл. 10.2 при $s_1 = 5$). Если $x_1 = 1$, то $s_2 = 4$. Суммарная прибыль при условии, что $s_2 = 4$ единицам средств между оставшимися тремя предприятиями, будет распределена оптимально и равна $f_1(1) + Z_2^*(4) = 8 + 16 = 24$ ($f_1(1)$ взято из табл. 10.1, а $Z_2^*(4)$ – из столбца 9 табл. 10.2). Аналогично

$$\text{при } x_1 = 2, s_2 = 3 \text{ и } f_1(2) + Z_2^*(3) = 10 + 13 = 23;$$

$$\text{при } x_1 = 3, s_2 = 2 \text{ и } f_1(3) + Z_2^*(3) = 11 + 10 = 21;$$

$$\text{при } x_1 = 4, s_2 = 1 \text{ и } f_1(4) + Z_2^*(1) = 12 + 16 = 28;$$

$$\text{при } x_1 = 5, s_2 = 0 \text{ и } f_1(5) + Z_2^*(0) = 18 + 0 = 18.$$

Сравнивая подчеркнутые числа, получим $Z_1^*(5) = 24 \text{ усл. ед.} = Z_{\max}$ при $x_1^* = x_1^*(5) = 1$.

Используя уравнения (10.12), получим $s_1^* = 5 - 1 = 4$, по табл. 10.2 в столбце 9 находим $x_2^* = x_2^*(4) = 2$. Далее находим $s_2^* = 4 - 2 = 2$, а по табл. 10.2 в столбце 6 находим $x_3^* = x_3^*(2) = 1$. Наконец, $s_3^* = 2 - 1 = 1$ и $x_4^* = x_4^*(1) = 1$, т.е. $X^*(1; 2; 1; 1)$.

Максимум суммарной прибыли равен 24 усл. ед. средств при условии, что первому предприятию выделено 1 усл. ед.; второму предприятию – 2 усл. ед.; третьему предприятию – 1 усл. ед.; четвертому предприятию – 1 усл. ед.

Замечание 1. Решение четырехмерной задачи на определение условного экстремума сведено к решению четырех одномерных задач: на каждом шаге определялась одна переменная x.

Замечание 2. В задаче видно, что метод ДП безразличен к виду и способу задания функции: $f_k(x)$ были заданы таблично, поэтому и $Z_k^*(s)$ и $X_k^*(s)$ принимали дискретные значения, представленные в табл. 10.2.

Замечание 3. Альтернативой ДП для подобной дискретной задачи является метод перебора. Метод ДП предпочтительнее, так как на этапе условной оптимизации отбрасываются заведомо негодные варианты.

Замечание 4. Достоинством метода является возможность анализа решения на чувствительность к изменению s_0 и числа шагов n.

Замечание 5. К недостаткам метода следует отнести возникновение технических трудностей при увеличении размерности. Если каждое управление X_k^* будет зависеть от r переменных, а состояние s_k^* – от p параметров, то на каждом шаге возникает rp-мерная задача оптимизации. При реализации метода ДП на ЭВМ практически можно решать задачи для небольших r, p, n.

Таблица 10.2

s_{k-1}	x_k	s_k	$k=3$: $f_3(x_3)+$ $+Z_4^*(s_3)$	Z_3^* (s_2)	x_3^* (s_2)	$k=2$: $f_2(x_2)+$ $+Z_3^*(s_2)$	Z_2^* (s_1)	x_2^* (s_1)	$k=1$: $f_1(x_1)+$ $+Z_2^*(s_1)$	Z_1^* (s_0)	x_1^* (s_0)
1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0+1=1			0+4=4			0+6=6	8	1
	1	0	3+0=3	4	0	6+0=6	6	1	8+0=8		
2	0	2	0+6=6			0+7=7			0+10=10		
	1	1	3+4=7	7	1	6+4=10	10	1	8+6=14	14	1
	2	0	4+0=4			9+0=9			10+0=10		
	0	3	0+8=8			0+9=9			0+13=13		
3	1	2	3+6=9	9	1	6+7=13	13	1	8+10=18	18	1
	2	1	4+4=8			9+4=13		2	10+6=16		
	3	0	7+0=7			11+0=11			11+0=11		
	0	4	0+13=13			0+13=13			0+16=16		
	1	3	3+8=11			6+9=15			8+13=21		
4	2	2	4+6=10	13	0	9+7=16	16	2	10+10=20	21	1
	3	1	7+4=11			11+4=15			11+6=17		
	4	0	11+0=11			13+0=13			12+0=12		

1	2	3	4	5	6	7	8	9	10	11	12
5	0 1 2 3 4 5	5 4 3 2 1 0	0+16=16 3+13=16 4+8=12 7+6=13 11+4=15 18+0=18	18	5	0+18=18 6+13=19 9+9=18 11+7=18 13+4=17 15+0=15	19	1	0+19=19 8+16=24 10+13=23 11+10=21 12+6=18 18+0=18	24	1

10.4. Общая схема применения метода ДП. Задача об оптимальном распределении ресурсов между отраслями на *n* лет

Рассмотрим общую схему применения метода ДП.

Предположим, что все требования, предъявляемые к задаче методом ДП, выполнены. Эти требования сформулированы в разд. 10.1 Построение модели ДП и применение метода ДП для решения сводятся к следующим пунктам:

1. Выбирают способ деления процесса управления на шаги.
2. Определяют параметры состояния s_k и переменные X_k на каждом шаге.
3. Записывают уравнения состояний.
4. Вводят целевые функции k-го шага и суммарную целевую функцию.
5. Вводят в рассмотрение условные максимумы (минимумы) $Z_k^*(s_{k-1})$ и условное оптимальное управление на k-м шаге: $X_k^*(s_{k-1}), k = n, n-1, \ldots, 2, 1$.
6. Записывают основные для вычислительной схемы ДП уравнения Беллмана для $Z_n^*(s_{n-1})$ и $Z_k^*(s_{k-1}), k = n-1, \ldots, 1$.
7. Решают последовательно уравнения Беллмана (условная оптимизация) и получают две последовательности функций:

$$\{Z_k^*(s_{k-1})\} \text{ и } \{X_k^*(s_{k-1})\}.$$

8. После выполнения условной оптимизации получают оптимальное решение для конкретного состояния s_0:

а) $Z_{\max} = Z_1^*(s_0)$;

б) по цепочке $s_0 \Rightarrow X_1^* \to s_1^* \Rightarrow X_2^* \to s_2^* \Rightarrow \ldots \to X_{n-1}^* \to s_{n-1}^* \Rightarrow$ $\Rightarrow X_n^* \to s_n^*$ оптимальное управление: $X^*(X_1^*, X_2^*, \ldots, X_n^*)$.

Рассмотрим, как работает схема на примере задачи об оптимальном распределении ресурсов между двумя отраслями на n лет вперед.

Планируется деятельность двух отраслей производства на n лет. Начальные ресурсы s_0. Средства x, вложенные в первую отрасль в начале года, дают в конце года прибыль $f_1(x)$ и возвращаются в конце года в размере $q_1(x) < x$; аналогично для второй отрасли функция прибыли равна $f_2(x)$, а возврата – $q_2(x)$ $(q_2(x) < x)$. В конце года все возвращенные средства заново перераспределяются между первой и второй отраслями, новые средства не поступают, прибыль в производство не вкладывается. Последние условия определяют вид уравнений состояний; если поступают новые средства или часть прибыли вкладывается в производство, это можно учесть, так как алгоритм метода ДП не изменится.

Требуется распределить имеющиеся средства s_0 между двумя отраслями производства на n лет так, чтобы суммарная прибыль от обеих отраслей за n лет оказалась максимальной.

Необходимо построить модель ДП для задачи и вычислительную схему.

Процесс распределения средств между двумя отраслями производства разворачивается во времени, решения принимаются в начале каждого года, следовательно, осуществляется деление на шаги: номер шага – номер года. Управляемая система – две отрасли производства, а управление состоит в выделении средств каждой отрасли в очередном году. Параметры состояния к началу k-го года – $s_{k-1}(k = 1, 2, \ldots, n)$ – количество средств, подлежащих распределению. Переменных управления на каждом шаге две: x_k – количество средств, выделенных первой отрасли и y_k – второй отрасли. Поскольку все средства s_{k-1} распределяются, то $y_k = s_{k-1} - x_k$, и поэтому управление на k-м шаге зависит от одной переменной x_k, т.е. $X_k(x_k, s_{k-1} - x_k)$.

Уравнения состояний

$$s_k = q_1(x_k) + q_2(s_{k-1} - x_k) \tag{10.9}$$

выражают остаток средств, возвращенных в конце k-го года.

Показатель эффективности k-го шага – прибыль, полученная в конце k-го года от обеих отраслей:

$$f_1(x_k) + f_2(s_{k-1} - x_k). \qquad (10.10)$$

Суммарный показатель эффективности – целевая функция задачи – прибыль за n лет:

$$Z = \sum_{k=1}^{n} (f_1(x_k) + f_2(s_{k-1} - x_k)). \qquad (10.11)$$

Пусть $Z_k^*(s_{k-1})$ – условная оптимальная прибыль за $n - k + 1$ год, начиная с k-го включительно, при условии, что имеющиеся на начало k-го года средства s_{k-1} в дальнейшем распределялись оптимально. Тогда оптимальная прибыль за n лет $Z_{\max} = Z_1^*(s_0)$.

Уравнения Беллмана имеют вид

$$Z_n^*(s_{n-1}) = \max_{0 \le x_k \le s_{k-1}} \left\{ f_1(x_n) + f_2(s_{n-1} - x_n) \right\}, \qquad (10.12)$$

$$Z_k^*(s_{k-1}) = \max_{0 \le x_k \le s_{k-1}} \{ f_1(x_k) + f_2(s_{kn-1} - x_k) + Z_{k+1}^*(s_k) \}, \qquad (10.13)$$

$(k = n-1, n-2, \ldots, 2)$.

Рассмотрим пример на конкретных данных.

Уравнение состояний (10.9) примет вид

$$s_k = 0,7x_k + 0,8(s_{k-1} - x_k), \text{ или } s_k = 0,8s_{k-1} - 0,1x_k. \qquad (10.14)$$

Целевая функция k-го шага (10.10)

$$Z = \sum_{k=1}^{4} 0,5s_{k-1} + 0,1x_k. \qquad (10.15)$$

Функциональные уравнения

$$Z_4^*(s_3) = \max_{0 \le x_4 \le s_3} \{0,5s_3 + 0,1x_4\}, \qquad (10.16)$$

и

$$Z_k^*(s_{k-1}) = \max_{0 \le x_k \le s_{k-1}} \{0,1x_k + 0,5s_{k-1} + Z_{k+1}^*(s_k)\}. \qquad (10.17)$$

Проведем условную оптимизацию.

4-й шаг. Используем уравнение (10.16). Обозначим через Z_4 функцию, стоящую в скобках, $Z_4 = 0,1x_4 + 0,5s_3$; Функция Z_4 –

линейная, возрастающая, с угловым коэффициентом 0,1. Поэтому максимум достигается на конце интервала $[0; s_3]$ (рис. 10.5). Следовательно, $Z_4^*(s_3) = 0,6s_3$ при $x_4^*(s_3) = s_3$.

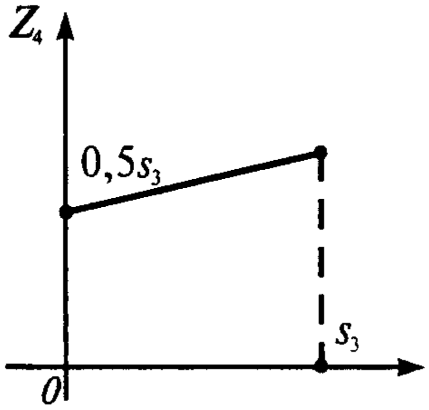


Рис.10.5

3-й шаг. Из уравнения оптимального решения

$$Z_3^*(s_3) = \max_{0 \le x_3 \le s_2} \{0,1x_3 + 0,5s_2 + 0,6s_3\}$$

найдем s_3 из уравнений состояний (10.14): $s_3 = 0,8s_2 - 0,1x_3$. Подставим это выражение в правую часть уравнения, получим

$$Z_3^*(s_2) = \max_{0 \le x_3 \le s_2} \{0,04x_3 + 0,98s_2\}.$$

Максимум достигается при $x_3 = s_2$; т.е. $Z_3^*(s_2) = 1,02s_2$ при $x_3^*(s_2) = s_2$.

2-й шаг. Из уравнения состояния имеем $s_2 = 0,8s_1 - 0,1x_2$. Поэтому уравнение (10.16) при $k = 2$ примет вид

$$Z_2^*(s_1) = \max_{0 \le x_2 \le s_1} \{1,316s_1 - 0,002x_2\}.$$

Это линейная относительно x_2 функция и она убывает на отрезке $[0; s_1]$. Поэтому ее максимум достигается при $x_2 = 0$ (рис. 10.6):

$$Z_2^*(s_1) = 1,316s_1 \text{ при } x_2^*(s_1) = 0.$$

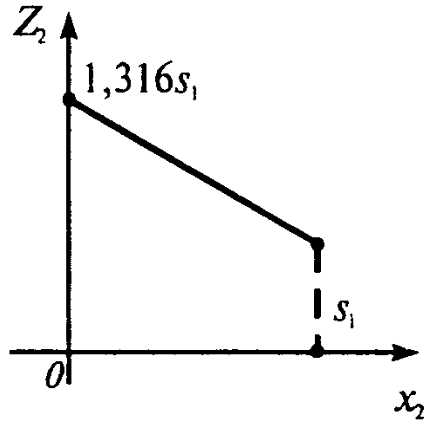


Рис. 10.6

1-й шаг. $s_1 = 0{,}8s_0 - 0{,}1x_1$. Уравнение (10.16) при $k = 1$ имеет вид

$$Z_1^*(s_0) = \max_{0 \le x_1 \le s_0} \{1{,}5528s_0 - 0{,}0316x_1\}.$$

Как и в предыдущем случае, максимум достигается в начале отрезка, т.е.

$$Z_1^*(s_0) = 1{,}5528s_0 \text{ при } x_1^*(s_0) = 0.$$

Условная оптимизация закончена. Используя ее результат и исходные данные, получим $Z_{\max} = Z_1^*(10\,000)$, $Z_{\max} = 15\,528$,

$$x_1^* = 0, \quad y_1^* = s_0 = 10\,000$$

(все средства выделяются второй отрасли)$\rightarrow$

$$s_1^* = 0{,}8 \cdot 10\,000 - 0{,}1 \cdot 0 = 8000 \Rightarrow x_2^* = 0,\ y_2^* = s_1 = 8000$$

(все средства выделяются второй отрасли)$\rightarrow$

$$\rightarrow s_2^* = 0{,}8 \cdot 8000 - 0{,}1 \cdot 0 = 6400 \Rightarrow x_3^* = 6400,\ y_3^* = 0 \rightarrow$$

(все средства выделяются первой отрасли)$\rightarrow$

$$\rightarrow s_3^* = 0{,}8 \cdot 6400 - 0{,}1 \cdot 6400 = 4480 \Rightarrow x_4^* = 4480,\ y_4^* = 0$$

(все средства выделяются первой отрасли).

Оптимальная прибыль за 4 года, полученная от двух отраслей производства при начальных средствах 10 000 ед., равна 15 528 ед. при условии, что первая отрасль получает по годам (0;0; 6400;4480), а вторая отрасль – соответственно (10 000; 8000; 0; 0).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Акулич, И. Л. Математическое программирование в примерах и задачах / И. Л. Акулич. – М. : Высш. шк., 1986.

2. Банди, Б. Основы линейного программирования / Б. Банди. – М. : Радио и связь, 1989.

3. Вентцель, Е. С. Исследование операций / Е. С. Вентцель. – М. : Сов. радио, 1972.

4. Вентцель, Е. С. Исследование операций. Задачи, принципы, методология / Е. С. Вентцель. – М. : Наука, 1980.

5. Горелик, В. А. Исследование операций / В. А. Горелик, И. А. Ушаков. – М. : Машиностроение, 1986.

6. Горчаков, А. А. Компьютерные экономико-математические модели / А. А. Горчаков, И. В. Орлова. – М. : Компьютер : ЮНИТИ, 1995.

7. Исследование операций : в 2 т. : пер. с англ. / под ред. Дж. Моудера и С. Элмаграби.– М. : Мир, 1981.

8. Исследование операций / под ред. М. А. Войтенко и Н. Ш. Кремера. – М. : Экономическое образование, 1992.

9. Калихман, И. Л. Динамическое программирование / И. Л. Калихман, М. А. Войтенко. – М. : Высш. шк., 1979.

10. Карасев, А. И. Математические методы и модели в планировании / А. И. Карасев, Н. Ш. Кремер, Т. И. Савельева. – М. : Экономика, 1987.

11. Карпелевич, Ф. И. Элементы линейной алгебры и линейного программирования / Ф. И. Карпелевич, Л. Е. Садовский. – М. : Наука, 1967.

12. Кофман, А. Методы и модели исследования операций : пер. с франц. / А. Кофман. – М. : Мир, 1966.

Учебное издание

Добрынина Наталия Филипповна,
Тарасов Дмитрий Викторович

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

Редактор *Т. В. Веденеева*
Технический редактор *Ю. В. Анурова*
Компьютерная верстка *Ю. В. Ануровой*
Дизайн обложки *А. А. Стаценко*

Подписано в печать 22.08.2017.

Формат $60\times84^{1}/_{16}$. Усл. печ. л. 6,04.

Заказ № 484. Тираж 28.

440026, Пенза, Красная, 40. Издательство ПГУ

Тел./факс: (8412) 56-47-33; e-mail: iic@pnzgu.ru