ПРЕДИСЛОВИЕ

Предлагаемое учебное пособие возникло на материале курса лекций, читавшегося автором в течение ряда лет на факультете вычислительной математики и кибернетики Московского государственного университета. Книга задумана как пособие для студентов, специализирующихся в области прикладной математики. Цель пособия — дать представление об основных принципах построения математических моделей экономических процессов и явлений и о специфических с математической точки зрения методах их исследования.

Из всего многообразия экономико-математических моделей выбрано несколько наиболее содержательных и в то же время не слишком громоздких. На этих примерах показаны самые существенные проблемы, возникающие в процессе моделирования экономической реальности, и результаты, которые можно получить на этом пути. Автор не стремился к наибольшей общности получаемых математических фактов и теорем, пытаясь добиться прозрачности изложения и обращая особое внимание на возможные экономические интерпретации.

Вместе с тем следует сказать, что исследование экономических проблем современного промышленного производства математическими методами требует прочного владения основами математического анализа, алгебры и в ряде случаев углубленных знаний функционального анализа, теории дифференциальных уравнений и других разделов математики.

Предполагается, что читатель знаком с основными разделами теории экстремальных задач: линейным и выпуклым программированием, принципом максимума Понтрягина для задач оптимального управления.

Это, впрочем, не означает, что материал книги недоступен тем, кто не получил специального математического образования. Все используемые абстрактные понятия определяются, а необходимые теоремы формулируются и снабжаются подробными пояснениями.

Книга состоит из трех частей. Это соответствует сложившейся экономической практике агрегирования экономических показателей. Так, раздел «Общее равновесие» относится к ситуации, когда никакого агрегирования исходных показателей не

$0 \leqslant y_1 \leqslant 1,\ y_2 = 0\}$. Функция полезности потребителя имеет вид

$$u(x_1, x_2) = x_1 + \sqrt{x_2}$$

и определена на множестве $X = \{(x_1, x_2) \mid x_1 \geqslant 0,\ x_2 \geqslant 0,\ x_2 - 1 \leqslant x_1 \leqslant x_2 + 1\}$. Весь доход $p_1 y_1 + p_2 y_2$ производственного сектора поступает в распоряжение потребителя.

13. Проверить, что в данной модели выполнены все условия теоремы Эрроу — Дебре, за одним исключением: у потребителя отсутствует начальная собственность.

14. Доказать, что в этой модели не существует конкурентного равновесия.

15. Показать, что если в качестве X взять $\mathbf{R}^n_+$, то равновесия также не существует.

16. Доказать, что множество Ω всех эффективных векторов (см. § 3 гл. 2) замкнуто.

ЧАСТЬ III

ПРОИЗВОДСТВЕННЫЕ ФУНКЦИИ

МАТЕМАТИЧЕСКОЕ ВВЕДЕНИЕ. ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Центральную роль при изучении нелинейных экономико-математических моделей играют математические методы оптимального управления и, более конкретно, различные формы принципа максимума как необходимого условия оптимальности. Нам потребуются две теоремы о необходимых условиях оптимальности в задачах управления — принцип максимума Понтрягина для двух задач. Данное математическое введение посвящено постановке и обсуждению задач оптимального управления в подходящей для нас формулировке и изложению без доказательства утверждений соответствующих теорем.

С понятием «управляемого объекта» мы уже встречались. Например, экономика, описываемая моделью Неймана, являлась таким управляемым объектом, причем параметром управления, с помощью которого регулировалось «движение» экономики во времени, являлся вектор интенсивностей. Состояние объекта описывалось вектором выпуска в данный момент времени. На управление — вектор интенсивностей — в каждый момент времени были наложены ограничения: оно должно быть таким, чтобы затраты в данный момент не превосходили объем выпуска в предыдущий.

Математическое описание управляемого объекта в общем случае таково. Состояние объекта задается в каждый момент времени n-мерным вектором $x \in \mathbf{R}^n$, координаты x_i, $i = 1, 2, \ldots, n$, которого называются *фазовыми координатами* объекта. Движение объекта состоит в том, что его координаты зависят от времени и от m величин $u_1, u_2, \ldots, u_m$, называемых *управлениями*. Управление u является m-мерным вектором $u = (u_1, u_2, \ldots, u_m)$, в свою очередь изменяющимся во времени. Однако изменение управляющего параметра u в определенной степени зависит только от управляющего субъекта.

Пусть задано некоторое замкнутое множество $U \subseteq \mathbf{R}^m$. Управление $u(t)$ называется *допустимым*, если в каждый момент времени t значение $u(t)$ принадлежит множеству U и вектор-функция $u(t)$ является кусочно непрерывной.

Пусть заданы начальное состояние объекта

$$x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$$

и закон движения в виде системы обыкновенных дифференциальных уравнений первого порядка:

$$\dot{x}_i = f_i(x, u, t), \quad i = 1, 2, \ldots, n, \tag{I}$$

где функции f_i предполагаются непрерывными по совокупности переменных и непрерывно дифференцируемыми по x_i, $i = 1, 2, \ldots, n$, и по t. Если теперь задать некоторое допустимое управление $u(t)$, то система дифференциальных уравнений (I) вместе с начальным условием $x(0) = x^0$ однозначно определит траекторию $x(t)$. Задача теории оптимального управления состоит в том, чтобы среди допустимых управлений $u(t)$ выбрать такое, которое определяло бы траекторию $x(t)$, максимизирующую некоторый заданный критерий. В зависимости от вида критерия, подлежащего максимизации, существует много различных типов задач оптимального управления. Среди них задачи экономической динамики выделяются определенной спецификой, связанной с наложением дополнительных ограничений на фазовые координаты: условие неотрицательности, требования на значения переменных в конце промежутка планирования и т. д.

Ниже мы формулируем несколько возможных постановок.

Задача 1. Для ее формулировки нам потребуется еще одно новое для нас понятие. *Гладкое многообразие* $M \subseteq \mathbb{R}^n$ размерности $n - k$ задается k уравнениями

$$g_i(x_1, x_2, \ldots, x_n) = 0, \quad i = 1, 2, \ldots, k,$$

где все функции g_i непрерывно дифференцируемы и все векторы

$$\operatorname{grad} g_i = \left(\frac{\partial g_i}{\partial x_1}, \frac{\partial g_i}{\partial x_2}, \ldots, \frac{\partial g_i}{\partial x_n}\right), \quad i = 1, 2, \ldots, k,$$

линейно независимы в каждой точке $x \in M$. Например, часто встречаются многообразия, задаваемые уравнениями вида

$$x_1 = \bar{x}_1, \quad x_2 = \bar{x}_2, \ldots, x_k = \bar{x}_k, \quad k \leqslant n,$$

где $\bar{x}_i$ — константы.

Будем говорить, что допустимое управление $u(t)$ переводит фазовую точку из положения x^0 на многообразие M, если $x(T) \in M$ для некоторого $T \geqslant 0$.

Задача формулируется следующим образом: среди всех допустимых управлений найти такое, которое переводит фазовую точку из положения x^0 на многообразие M за минимальное время. Для решения этой задачи вводятся новые (двойственные) переменные $\psi_1(t), \psi_2(t), \ldots, \psi_n(t)$ и строится *функция Гамильтона*

$$\mathscr{H}(\psi, x, u, t) = \sum_{i=1}^{n} \psi_i f_i(x, u, t).$$

Рассмотрим сопряженную систему дифференциальных уравнений:

$$\dot{\psi}_i = -\frac{\partial \mathscr{H}}{\partial x_i}, \quad i = 1, 2, \ldots, n. \qquad \text{(II)}$$

Взяв произвольное допустимое управление $u(t)$ и начальное условие x^0, мы из системы (I) можем найти соответствующую траекторию $x(t)$. После этого мы можем найти решение $\psi_1(t)$, $\psi_2(t), \ldots, \psi_n(t)$ сопряженной системы (II), задав какое-нибудь начальное условие. Естественно, что, задавая различные начальные условия, мы получим различные решения.

При фиксированных (постоянных) значениях ψ, x, t функция $\mathscr{H}$ является функцией параметра $u \in U$. Обозначим

$$\mathscr{M}(\psi, x, t) = \max_{u \in U} \mathscr{H}(\psi, x, u, t).$$

Теорема (принцип максимума Понтрягина [69]). *Если допустимое управление $u(t)$, переводящее фазовую точку из положения x^0 на многообразие M, оптимально, то существует такое ненулевое решение $\psi(t) = (\psi_1(t), \psi_2(t), \ldots, \psi_n(t))$ сопряженной системы, что выполнены условия:*

1) *при любом t функция $\mathscr{H}(\psi(t), x(t), u, t)$ переменной и достигает в точке $u = u(t)$ максимума:*

$$\mathscr{H}(\psi(t), x(t), u(t), t) = \mathscr{M}(\psi(t), x(t), t);$$

2) *вектор $\psi(T)$ ортогонален всем касательным векторам многообразия M в точке $x(T)$ (условие трансверсальности), т. е. вектор $\psi(T)$ является линейной комбинацией векторов* $\operatorname{grad} g_i(x(T))$, $i = 1, 2, \ldots, k$.

Задача 2. В этой задаче отрезок времени $[0, T]$ фиксирован заранее. Пусть задана функция $f_0(x, u, t)$, непрерывная вместе со своими частными производными по совокупности переменных. Среди всех допустимых управлений $u(t)$, $0 \leqslant t \leqslant T$, требуется определить такое, для которого соответствующая фазовая траектория удовлетворяет условию $x(T) \geqslant x^T$ (здесь x^T — заданный вектор, $x^T \in \mathbb{R}^n$) и максимизирует функционал $\int_0^T f_0(x, u, t)\,dt$.

Решение этой задачи проводится аналогично предыдущей, со следующими изменениями: функция Гамильтона имеет вид

$$\mathscr{H} = \psi_0 f_0(x, u, t) + \sum_{i=1}^{n} \psi_i f_i(x, u, t);$$

соответственно, сопряженная система (II) строится именно для такой функции Гамильтона.

Теорема [31]. *Оптимальное управление в задаче 2 удовлетворяет необходимому условию 1) принципа максимума Понтрягина. Кроме того, $\psi(T) \geqslant 0$, $\langle \psi(T), x(T) - x^T \rangle = 0$. При этом ψ_0 не зависит от t и можно считать, что ψ_0 равно либо 0, либо 1. Если $\psi_0 = 0$, то $\psi(t) \neq 0$ при всех t.*

Отметим, что принцип максимума Понтрягина дает лишь необходимые условия оптимальности, не отвечая на вопросы о существовании хотя бы одного допустимого управления, решающего задачу достижения заданного многообразия, о существовании оптимального управления и т. д. Как правило, существование допустимого управления, решающего задачу достижения заданного многообразия в задачах экономической динамики, очевидно из свойств рассматриваемых функций и содержательных соображений.

Вопрос о выделении с помощью необходимых условий оптимального управления решается обычно таким образом: доказывается, что существует лишь одно допустимое управление, удовлетворяющее этим условиям.

Наконец, проблема существования оптимального управления решается в работах А. Ф. Филиппова [79], и в тех моделях, которые мы будем рассматривать ниже, всякий раз можно с помощью его результатов показать, что оптимальное управление существует. Подробнее, однако, мы на этом останавливаться не будем.

ГЛАВА 1

ЭЛЕМЕНТЫ ОБЩЕЙ ТЕОРИИ ПРОИЗВОДСТВЕННЫХ ФУНКЦИЙ

§ 1. Основные понятия

1. Первая часть книги была посвящена изучению линейных замкнутых моделей производства — моделей Леонтьева, Неймана, Гейла. Оба из упомянутых свойств этих моделей — линейность и замкнутость — налагают определенные ограничения на круг изучаемых экономических явлений и на способность модели адекватно отражать действительность. Наиболее ограничительным предположением является требование линейности. Несомненно, что линейные модели, как наиболее простые, являются необходимым этапом в процессе моделирования экономической действительности. На их языке, как мы видели, удалось сформулировать некоторые основные экономико-математические понятия и связи между ними, получить интересные качественные выводы о динамике производства.

Переход к нелинейным моделям означает качественный скачок в теории, он не может совершаться гладко и безболезненно, скажем, путем построения нелинейных аналогов линейных моделей (что нетрудно) и обобщения соответствующих результатов (что если и возможно, то очень сложно). Поэтому естественный путь перехода к нелинейным моделям, как нам представляется, таков: отказываясь от линейности, упростить модель в каком-либо другом направлении — сделать технологическое

отображение однозначным, уменьшить число учитываемых производственных факторов и т. д.

Вместе с тем при изучении экономических процессов в современном крупномасштабном производстве бывает чрезвычайно трудно, если не невозможно, собрать необходимую статистику для практического построения модели, скажем типа модели Неймана, учитывающей внутреннюю структуру производства. С другой стороны, зачастую гораздо проще получить отчетные данные о поведении и взаимосвязи укрупненных экономических показателей, таких, как стоимость произведенного продукта, объем основных фондов, численность работников и т. п. Оказывается, что, оперируя даже такими укрупненными показателями и рассматривая производственный объект как «черный ящик» (т. е. изучая лишь связь между затраченными средствами и произведенным продуктом), можно получать определенные содержательные выводы.

Высказанные соображения лежат в основе теории производственных функций. Возникновение теории производственных функций принято относить к 1928 г., когда появилась статья [83] американских ученых экономиста П. Дугласа и математика Д. Кобба «Теория производства». В этой статье была предпринята попытка определить эмпирическим путем влияние величины затрачиваемого капитала и труда на объем выпускаемой продукции в обрабатывающей промышленности США. Были использованы статистические данные за 1899—1922 гг. и поставлены следующие задачи:

1. Определить параметрический класс функций, наиболее точно приближающий количественные соотношения между тремя выбранными характеристиками производственной деятельности.

2. Найти числовые параметры, задающие конкретную функцию этого класса.

3. Сравнить результаты, получаемые как значения функций, с фактическими данными.

Д. Коббом была предложена функция вида $Y = AK^{\alpha}L^{\beta}$, где Y — объем выпущенной продукции, K — объем основного капитала, L — затраты труда, A, α, b — числовые параметры. Заранее накладывались условия $A > 0$, α, $\beta \geqslant 0$, $\alpha + \beta = 1$.

Была составлена система уравнений

$$\ln Y_t = \ln A + \alpha \ln K_t + \beta \ln L_t, \quad 1899 \leqslant t \leqslant 1922,$$

где Y_t, K_t, L_t — фактическое значение соответствующих величин в год t. Методом наименьших квадратов отыскивались значения A, α, β, минимизирующие выражение

$$\sum_{t=1899}^{1922} (\ln Y_t - \ln A - \alpha \ln K_t - \beta \ln L_t)^2.$$

При этом оказалось, что $A = 1{,}01$, $\alpha = 0{,}25$, $\beta = 0{,}75$, и вся функция получила вид $Y = 1{,}01 \cdot K^{0{,}25}L^{0{,}75}$.

Сравнение величины $Y(K_t, L_t)$ с фактическими значениями Y_t показало, что полученная зависимость дает хорошее приближение к действительности. Отклонения величин Y_t и $Y(K_t, L_t)$ были связаны с периодами депрессии и оживления деловой активности.

Эти результаты привлекли внимание как экономистов, так и математиков, и к настоящему времени список литературы по производственным функциям насчитывает сотни наименований.

В нашей стране понятие производственной функции широко применяется для задач планирования и прогнозирования как в отдельных отраслях (см. [30, 51, 71, 72]), так и на уровне всего народного хозяйства (см. [1, 3, 77, 81]).

2. Вопрос об адекватном описании экономической реальности на языке производственных функций тесно связан с проблемой агрегирования экономических показателей. Для иллюстрации сказанного рассмотрим следующую формальную схему, предложенную в [51].

Пусть вектор $\xi = (\xi_1, \xi_2, \ldots, \xi_k) \in \mathbf{R}^k_+$ описывает набор ресурсов, необходимых для функционирования производственного процесса P, $\eta = (\eta_1, \eta_2, \ldots, \eta_s) \in \mathbf{R}^s_+$ представляет собой вектор товаров, выпускаемых данным процессом. Зависимость затраты — выпуск задается отображением $P\colon \mathbf{R}^k_+ \to 2^{\mathbf{R}^s_+}$, ставящим в соответствие всякому $\xi \in \mathbf{R}^k_+$ множество $P(\xi) \subseteq \mathbf{R}^s_+$. Предполагается, что отображение P достаточно подробно описывает реальную технологию, так что число k учитываемых видов ресурсов довольно велико.

Введем два однозначных отображения:

$$\mu\colon \mathbf{R}^k_+ \to \mathbf{R}^n_+, \quad \nu\colon \mathbf{R}^s_+ \to \mathbf{R}^1_+,$$

которые будем называть *отображениями агрегирования*. Функция μ ставит в соответствие вектору ресурсов $\xi \in \mathbf{R}^k_+$ вектор $x = \mu(\xi) \in \mathbf{R}^n_+$ с меньшим числом координат $(n < k)$, что интерпретируется как объединение, агрегирование нескольких различных видов ресурсов в один обобщенный ресурс, в результате чего число n видов таких обобщенных ресурсов оказывается небольшим. Например, пусть первые три координаты ξ_1, ξ_2, ξ_3 вектора ξ означают: $\xi_1 = 10$ тонн стального проката, $\xi_2 = 15$ тонн чугунного проката, $\xi_3 = 5$ тонн латунного проката.

При агрегировании с помощью отображения μ эти три координаты могут быть заменены одной координатой x_1 вектора $x = \mu(\xi)$, где $x_1 = \xi_1 + \xi_2 + \xi_3 = 30$ тонн проката черных и цветных металлов. Если же решено измерять данный обобщенный ресурс в стоимостных терминах, то координата x_1 вектора $x = \mu(\xi)$ может иметь вид $x_1 = p_1\xi_1 + p_2\xi_2 + p_3\xi_3$ — на такую сумму требуется проката черных и цветных металлов, где p_1, p_2, p_3 — цена одной тонны проката соответствующего вида.

Аналогично можно привести примеры агрегирования различных видов труда к одному виду с помощью либо простого суммирования числа работающих, либо подсчета суммарной заработной платы и т. д.

В нашем примере агрегирование производилось с помощью линейных отображений, но, вообще говоря, на отображение μ никаких априорных условий не накладывается.

$$\begin{tikzcd} x \arrow[r, "f"] & y \\ \xi \arrow[u, "\mu"] \arrow[r, "P"'] & \eta \arrow[u, "\nu"'] \end{tikzcd}$$

Рис. 1.1.

Отображение ν ставит в соответствие материальному результату $\eta \in P(\xi)$ производства число $y = \nu(\eta)$ — оценку этого результата (стоимость выпущенной продукции, прибыль и т. д.).

Предположим, что отображения P, μ, ν заданы. Назовем *производственной функцией* процесса P такую функцию $f\colon \mathbf{R}^n_+ \to \mathbf{R}^1_+$, что имеет место тождество

$$f(\mu(\xi)) = \nu(P(\xi)) \tag{1.1}$$

для любого $\xi \in \mathbf{R}^s_+$.

Соотношение между P, μ, ν, f можно выразить диаграммой (рис. 1.1). При заданных P, μ, ν отображение f является решением функционального уравнения

$$\mu * f = P * \nu, \tag{1.2}$$

где символ $*$ означает композицию отображений. В случае, когда выполнено (1.2), говорят, что диаграмма, изображенная на рис. 1.1, *коммутативна*.

Ясно, что функциональное уравнение (1.2) не всегда имеет решение. В нашей интерпретации это означает, что не для всякого производственного процесса P при заданных отображениях агрегирования μ, ν существует производственная функция. В связи с этим следует обратить внимание на ряд важных обстоятельств:

— Большое значение имеет выбор области определения функции f: может случиться, что тождество (1.1) выполняется не при произвольном $\xi \in \mathbf{R}^k_+$, а лишь в пределах некоторой области $U \subset \mathbf{R}^k_+$, так что функцию f следует рассматривать лишь на образе $D = \mu(U)$ этой области при отображении μ; в этом случае для адекватного описания процесса P может потребоваться несколько производственых функций с различными областями определения.

— Весьма важен правильный выбор отображений агрегирования μ, ν. Может случиться, что при некоторых μ, ν не существует производственной функции, моделирующей процесс P, в то время как подходящее задание отображений μ и ν позволяет такую функцию построить.

— Может оказаться, что существует функция f, для которой равенство (1.1) выполняется лишь приближенно, но с достаточной степенью точности, что вполне позволяет использовать его как модель производственного отображения P.

Для иллюстрации некоторых из высказанных соображений рассмотрим пример. Пусть производственное отображение P задается с помощью так называемой задачи составления плана производства, встречавшейся во второй части данной книги при изучении модели Вальда — Касселя.

Рассмотрим задачу линейного программирования

$$\begin{gathered} \max \langle c, \eta \rangle \\ A\eta \leqslant \xi, \quad \eta \geqslant 0. \end{gathered} \tag{1.3}$$

Здесь $\xi \in \mathbf{R}_+^k$ — вектор потребляемых ресурсов, A — технологическая $k \times s$-матрица, $\eta \in \mathbf{R}_+^s$ — вектор, описывающий выпуск товаров. В качестве ν естественно взять отображение, сопоставляющее всякому решению η задачи (1.3) число $\langle c, \eta \rangle$.

Предположим, что двойственная к (1.3) задача

$$\begin{gathered} \min \langle p, \xi \rangle \\ pA \geqslant c, \quad p \geqslant 0 \end{gathered} \tag{1.4}$$

имеет единственное решение p^*. В этом случае существует некоторая окрестность $U \subseteq \mathbf{R}_+^k$ вектора ξ такая, что p^* будет решением задачи вида (1.4), где в качестве ξ берется любой вектор $\xi' \in U$. Возьмем в качестве μ функцию

$$\mu(\xi) = p_1^*\xi_1 + p_2^*\xi_2 + \ldots + p_k^*\xi_k = \langle p^*, \xi \rangle.$$

Обозначим через $P(\xi)$ множество решений задачи (1.3). По теореме двойственности в линейном программировании (см. ч. I) имеем

$$\mu(\xi) = \langle c, \eta \rangle \tag{1.5}$$

для любых $\xi \in U, \eta \in P(\xi)$. Это означает, что функция $f(x) = x$, где x — число, является производственной функцией технологического процесса P, описываемого задачей (1.3). В самом деле, нетрудно убедиться, что равенство $f(\mu(\xi)) = \nu(P(\xi))$ эквивалентно (1.5). Областью определения производственной функции f служит числовой отрезок $D = \{x \mid x = \langle p^*, \xi \rangle, \ \ \xi \in U\}$.

На данном примере видно, что выбор отображений μ, ν и области определения D производственной функции сыграл решающую роль: для технологического процесса P удалось построить простейшую из возможных производственных функций — однофакторную функцию $f(x) = x$; как только становится известно значение фактора $x = \mu(\xi)$ (совокупная стоимость ресурсов), тут же определяется и оценка $\langle c, \eta \rangle$ результатов производства.

3. Определение производственной функции отличается от понятия технологического отображения лишь требованием однозначности. Пусть $\mathbf{R}_+^n$ — положительный ортант n-мерного пространства, каждый вектор $x = (x_1, x_2, \ldots, x_n)$ которого интерпретируется как вектор затрат производственных ресурсов (в стоимостном или натуральном выражении). В качестве факторов производства здесь могут выступать как первичные факторы в обычном понимании, так и продукты производства, внешнего по отношению к изучаемому, выступающие в данном случае как ресурсы или сырье.

Пусть $\mathbf{R}_+^m$ — положительный ортант m-мерного пространства, каждый вектор $y = (y_1, y_2, \ldots, y_m)$ которого интерпретируется как набор количественных оценок результатов производства при определенных затратах ресурсов. Такими количественными оценками могут служить, например, физический объем выпуска по каждому из наименований выпускаемой продукции, стоимостные показатели.

Если P — некоторый производственный процесс, то *производственной функцией* f этого процесса будем называть отображение $f\colon D \to U$, где $D \subseteq \mathbf{R}_+^n$, $U \subseteq \mathbf{R}_+^m$, моделирующее выпуск продукции в процессе P.

Введение множеств D, U обусловлено тем, что в реальных ситуациях построение функции f для процесса происходит всегда на основе ограниченного статистического материала. В этом случае отчетные данные, на основе которых строится функция, заполняют ограниченные участки соответствующих пространств. При этом может случиться, что удобная аналитическая форма функции f дает хорошие результаты (т. е. достаточно адекватно моделирует процесс P) только в пределах некоторых множеств D и U.

Отметим сразу, что, несмотря на широту приведенного выше определения производственной функции, до сих пор в научной и прикладной экономико-математической литературе в основном изучался случай $m = 1$, т. е. когда имеется единственная количественная оценка результатов производства. В этом случае производственную функцию естественно записывать как обычную функцию нескольких переменных: $y = f(x_1, x_2, \ldots, x_n)$. Всюду в дальнейшем мы будем иметь дело только с таким случаем.

Для введения основных математических характеристик производственной функции и выяснения их экономической интерпретации рассмотрим двухфакторную производственную функцию. Обозначим через K объем основных фондов либо в стоимостном выражении, либо в количественном (скажем, число станков). Пусть L — числовое выражение объема трудовых ресурсов, т. е. число рабочих, число человеко-дней, человеко-часов и т. д., Y — объем выпущенной продукции в стоимостном выражении либо в натуральном, если мы имеем дело с отраслью, выпускающей один продукт. Тогда производственная функция имеет вид

$$Y = F(K, L). \tag{1.6}$$

Ниже в качестве иллюстрации будем рассматривать одну из наиболее распространенных двухфакторных функций — функцию Кобба — Дугласа

$$Y = AK^{\alpha}L^{\beta}, \tag{1.7}$$

где $A > 0$ — константа, $0 < \alpha < 1$, $0 < \beta < 1$, $\alpha + \beta = 1$. Обычные требования на производственную функцию (1.6) заключаются в требовании гладкости и условиях

$$\frac{\partial F}{\partial K} > 0, \quad \frac{\partial F}{\partial L} > 0, \tag{1.8}$$

$$\frac{\partial^2 F}{\partial K^2} < 0, \quad \frac{\partial^2 F}{\partial L^2} < 0, \quad K,\ L > 0. \tag{1.9}$$

Смысл условий (1.8) ясен: при увеличении объема одного из факторов и неизменном объеме другого выпуск продукции возрастает. Условия (1.9) означают, что при фиксированном объеме одного из факторов последовательное увеличение другого приводит ко все меньшим приростам произведенного продукта.

Иногда отказываются от требования дифференцируемости производственной функции F. В таком случае вместо условий (1.8), (1.9) требуют монотонности и вогнутости F по каждому аргументу.

Перейдем к перечислению основных экономико-математических характеристик производственной функции. Важными характеристиками являются:

— *средняя производительность труда* $y = Y/L$ — отношение объема произведенного продукта к количеству затраченного труда;

— *средняя фондоотдача* $z = Y/K$ — отношение объема произведенного продукта к величине основных фондов.

Для функции Кобба — Дугласа, например, средняя производительность труда равна $AK^{\alpha}L^{\beta-1}$ и в силу условия $\beta < 1$ является убывающей функцией аргумента L. Другими словами, с увеличением затрат средняя производительность труда падает. Этот вывод допускает естественное объяснение: поскольку величина второго фактора K остается неизменной, то, значит, вновь привлекаемая рабочая сила не обеспечивается дополнительными средствами производства, что и приводит к снижению производительности труда. Таким образом, становится ясным и

значение такой характеристики, как фондовооруженность труда $k = K/L$, показывающая объем основных фондов, приходящийся на одного работника.

Наряду со средними показателями при анализе производственных функций играют роль и предельные характеристики функции. *Предельная производительность труда* $v = \partial F/\partial L$ характеризует величину дополнительного эффекта от каждой дополнительной единицы затраченного труда в данной точке (K, L) фазовой плоскости. Условия (1.9) показывают, что при неизменных основных фондах при увеличении численности работников предельная производительность труда, аналогично средней, падает. Для функции Кобба — Дугласа имеем $v = \beta AK^{\alpha}L^{\beta-1} = \beta Y/L$, т. е. предельная производительность труда пропорциональна средней производительности и всегда меньше ее ($\beta < 1$).

Предельная фондоотдача определяется аналогично: $r = \partial F/\partial K$.

Заметим, что такие характеристики, как предельная и средняя производительность труда и фондоотдача, являются размерными величинами, связанными с абсолютными приростами. Представляют также интерес величины, характеризующие процент прироста продукции при увеличении затрат ресурса на 1 %. Такие показатели именуются коэффициентами эластичности.

Коэффициент эластичности по фондам определяется равенством

$$\alpha = \frac{\partial F}{\partial K}\frac{K}{Y}. \tag{1.10}$$

Поясним формулу (1.10). Пусть приращению ΔK основных фондов при неизменном втором факторе — трудовых ресурсах — соответствует приращение ΔY объема выпуска. Тогда увеличению объема основных фондов на $(\Delta K/K)\cdot 100$ % соответствует увеличение выпуска на $(\Delta Y/Y)\cdot 100$ %. Следовательно, при увеличении объема основных фондов на 1 % объем выпуска увеличится на $\frac{\Delta Y}{\Delta K}\frac{K}{Y}$ %. Переходя к пределу при $\Delta K \to 0$, получаем выражение (1.10).

Коэффициент эластичности по труду $\beta = \frac{\partial F}{\partial L}\frac{L}{Y}$ имеет аналогичный смысл.

Видим, что параметры α и β в формуле (1.7), задающей функцию Кобба — Дугласа, являются как раз коэффициентами эластичности. Таким образом, коэффициенты эластичности по факторам для функции Кобба — Дугласа суть величины постоянные, не зависящие от значений факторов K, L.

§ 2. Неоклассические производственные функции. Условие однородности

На производственную функцию вида (1.6) часто накладывают дополнительные к (1.8), (1.9) ограничения. Основное из них состоит в том, что функцию $F(K, L)$ предполагают *однородной*, т. е. требуют выполнения соотношения

$$F(\lambda K, \lambda L) = \lambda^\gamma F(K, L)$$

для любого $\lambda > 0$. Число $\gamma > 0$ называется *степенью однородности* F. Показатель γ характеризует *эффект от расширения масштаба производства*: если $\gamma > 1$, то одновременное увеличение всех факторов в λ раз приводит к возрастанию объема выпуска больше чем в λ раз, т. е. эффект от расширения масштаба производства положителен. Наиболее употребительными являются линейно-однородные производственные функции, т. е. когда $\gamma = 1$ и

$$F(\lambda K, \lambda L) = \lambda F(K, L).$$

Применительно к капиталистической экономике понятие линейно-однородной производственной функции допускает следующую интерпретацию. Пусть Y обозначает суммарный реальный доход (национальный доход) всего общества. Воспользовавшись известной теоремой Эйлера об однородных функциях, можно написать

$$Y = \frac{\partial F}{\partial K} K + \frac{\partial F}{\partial L} L. \tag{1.11}$$

Будем считать, что общество состоит только из рабочих и капиталистов. Тогда доход Y распадается на две части: доход рабочих и прибыль капиталистов. Пусть средняя реальная заработная плата рабочего равна w. Следовательно, суммарный доход рабочих равен wL.

Теория предельной производительности утверждает, что в условиях совершенной конкуренции общая занятость рабочей силы и заработная плата w связаны соотношением $\partial F/\partial L = w$. В пользу этого утверждения можно привести следующее соображение. Нанимая рабочих, капиталист увеличивает их численность на своем предприятии до тех пор, пока дополнительный доход $\partial F/\partial L$, приносимый очередным дополнительным рабочим, не превосходит его заработной платы.

Таким образом, второе слагаемое в (1.11) имеет смысл суммарного дохода рабочих. Следовательно, первое слагаемое в (1.11) представляет собой прибыль капиталистов. При такой интерпретации величина $\partial F/\partial K$, характеризующая дополнительный доход от одной дополнительной единицы капитала, называется *нормой прибыли*.

При изучении линейно-однородных производственных функций естественно перейти к новым переменным: $k = K/L$, $y = Y/L$. Обозначив $f(k) = F(K/L, 1)$, получаем запись функции (1.6) в виде

$$y = f(k).$$

Выпишем введенные нами основные характеристики производственных функций для случая линейно-однородной функции $Y = F(K, L)$ или $y = f(k)$, где y — средняя производительность труда, k — фондовооруженность.

В новых переменных имеем:

— предельная производительность труда $v = f - kf'$;

— предельная фондоотдача $r = f'$;

— коэффициент эластичности по фондам $\alpha = kf'/f$ (эта формула верна для любой однородной функции);

— коэффициент эластичности по труду $\beta = 1 - kf/f$ (если степень однородности равна γ, то $\beta = \gamma - kf'/f$).

Величина v по своему смыслу и в силу (1.9) является возрастающей функцией аргумента k; величина r — убывающая функция. Относительно поведения величин α, β априори нельзя сказать ничего.

Представляет интерес вопрос о классах производственных функций, для которых какие-либо из этих величин являются постоянными, не зависящими от k. Оставляем читателю для доказательства следующий факт:

Если хотя бы один из коэффициентов α, β эластичности по ресурсам не зависит от k, то производственная функция является функцией Кобба — Дугласа.

Линейно-однородная производственная функция, удовлетворяющая условиям

$$f' > 0, \quad f'' < 0, \quad f(0) = 0,$$
$$\lim_{k\to\infty} f(k) = \infty, \quad \lim_{k\to 0} f'(k) = \infty, \quad \lim_{k\to\infty} f'(k) = 0, \tag{1.12}$$

называется *неоклассической*.

Перечисленные условия можно интерпретировать естественным образом: при отсутствии основных фондов выпуск равен нулю; при неограниченном росте фондовооруженности объем выпуска также растет неограниченно; при возрастании фондовооруженности от нуля происходит стремительный рост объема выпуска, который при дальнейшем росте фондовооруженности сходит на нет.

Отметим, что функция Кобба — Дугласа является неоклассической.

§ 3. Эластичность замены факторов. CES-функции

Существенной особенностью реальных процессов производства является возможность замещения одного фактора другим. Например, при отсутствии экскаватора (стоимость которого составляет часть основных фондов K) для рытья канавы его можно заменить определенным числом землекопов (увеличив значение фактора L, характеризующего объем трудовых ресурсов), и наоборот. Необходимость замены факторов вытекает из того, что тот или иной ресурс может быть дефицитным, откуда появляется стремление заменить его в процессе производства другим, более доступным. В рамках крупных производственных систем, как правило, имеются широкие возможности маневрирования ресурсами (факторами).

Введем для двухфакторной производственной функции $Y = F(K, L)$ показатели, характеризующие возможность замещения одного фактора другим. Предположим, что объем L трудовых ресурсов уменьшился на величину ΔL. Спрашивается, на какую величину ΔK следует увеличить объем K основных фондов, чтобы выпуск Y остался неизменным? Считая Y константой, возьмем полный дифференциал от обеих частей уравнения $Y = F(K, L)$:

$$0 = dY = \frac{\partial F}{\partial K} dK + \frac{\partial F}{\partial L} dL.$$

Предельной нормой замены S_K трудовых ресурсов L основными фондами K называется величина

$$S_K = -\frac{dK}{dL} = \frac{\partial F/\partial L}{\partial F/\partial K}.$$

Ясно, что можно ввести аналогичный показатель S_L, причем S_K и S_L связаны простым соотношением $S_K S_L = 1$.

Для однородной производственной функции можно получить более простое выражение для S_K. Поскольку в этом случае $Y = F(K, L) = L^\gamma f(K/L)$, то

$$\frac{\partial F}{\partial L} = L^{\gamma-1}(\gamma f(k) - kf'(k)), \quad \frac{\partial F}{\partial K} = L^{\gamma-1} f'(k),$$

откуда

$$S_K = \gamma \frac{f(k)}{f'(k)} - k. \tag{1.13}$$

Видим, что для однородной функции норма замены зависит только от величины k фондовооруженности.

Для функции Кобба — Дугласа $S_K = \frac{\beta}{\alpha} k$. Отсюда видно, что норма замены S_K прямо пропорциональна фондовооруженности, что представляется вполне естественным: чем выше фондовооруженность, тем больше требуется фондов для компенсации одной единицы трудовых ресурсов.

Для однородных производственных функций введем понятие эластичности замены. При изменении показателя k на 1 % значение нормы замены S_K меняется на $\frac{dS_K}{dk}\frac{k}{S_K}$ процентов. Следовательно, для того чтобы добиться изменения нормы замены на 1 %, необходимо изменить величину фондовооруженности на $\left(\frac{dS_K}{dk}\frac{k}{S_K}\right)^{-1}$ процентов. Данная величина называется *эластичностью замены* σ_K:

$$\sigma_K^{-1} = \frac{dS_K}{dk}\frac{k}{S_K}.$$

Нетрудно проверить, что для функции Кобба — Дугласа эластичность замены постоянна и равна 1.

Подставляя выражение для S_K, получаем

$$\sigma_K = -\frac{\gamma f'(f - kf')}{k[(1-\gamma)(f')^2 + ff'']}.$$

Аналогично можно ввести эластичность замены σ_L первого фактора K вторым фактором L по формуле

$$\sigma_L^{-1} = \frac{dS_L}{dk^{-1}}\frac{k^{-1}}{S_L}.$$

Нетрудно проверить, что при этом $\sigma_K = \sigma_L$, поэтому мы будем употреблять для обозначения эластичности символ σ, опуская значок, указывающий фактор.

Предоставляем читателю доказать, что если норма замены S некоторой линейно-однородной функции F является постоянной (не зависит от k), то $F(K, L) = AK + BL$, где A, B — константы.

Таким образом, особый интерес представляет лишь случай, когда постоянной является эластичность замены σ.

Соответственно характеру поведения показателя эластичности замены различают два класса производственных функций: VES (Variable Elastisity of Substitution — переменная эластичность замены) и CES (Constant Elastisity of Substitution — постоянная эластичность замены).

Класс функций с постоянной эластичностью замены допускает простое описание. Для его определения достаточно решить дифференциальное уравнение

$$\sigma = -\frac{f'(f - kf')}{k[(1-\gamma)(f')^2 + ff'']},$$

считая величину σ заданной константой, или эквивалентное ему уравнение

$$\frac{dk}{dS}\,\frac{S}{k}=\sigma. \qquad (1.14)$$

С целью избавить читателя от проведения утомительных выкладок приведем решение уравнения (1.14). Непосредственно из (1.14) имеем $S=Ck^{1/\sigma}$, где C — произвольная константа. Привлекая выражение для нормы замещения S, получаем уравнение

$$\frac{f'}{f}=\frac{\gamma}{k+Ck^{1/\sigma}},$$

или

$$\ln f=\gamma\int\frac{dk}{k+Ck^{1/\sigma}}. \qquad (1.15)$$

Интеграл в правой части формулы (1.15) вычисляется непосредственно заменой переменных $k=t^{\sigma/(\sigma-1)}$ и равен $\frac{\sigma}{\sigma-1}\ln C_1\times$ $\times\left(k^{(\sigma+1)/\sigma}+C\right)$, где C_1 — произвольная константа.

Таким образом, общее решение уравнения (1.14) имеет вид

$$f=C_1\left(k^{(\sigma-1)/\sigma}+C\right)^{\gamma\sigma/(\sigma-1)}.$$

Возвращаясь к переменным K и L, получаем

$$Y=C_1\left(K^{(\sigma-1)/\sigma}+CL^{(\sigma-1)/\sigma}\right)^{\gamma\sigma/(\sigma-1)}.$$

Обозначая $\rho=(1-\sigma)/\sigma$ и выбирая подходящим образом константы C_1 и C, общее решение уравнения (1.14) можно записать в виде

$$Y=F(K,L)=A[\delta K^{-\rho}+(1-\delta)L^{-\rho}]^{-\gamma/\rho}, \qquad (1.16)$$

$A>0$, $0<\gamma\leqslant 1$, $\rho>-1$. Условия $0<\gamma\leqslant 1$, $\rho>-1$ обеспечивают выполнение требований (1.8), (1.9). Форма записи (1.16) решения уравнения (1.14) является общепринятым видом производственной функции класса CES.

Приведенное нами решение годится лишь для случая $\sigma\neq 0$, $\sigma\neq 1$ (убедитесь в этом).

В случае $\sigma=1$ решение уравнения (1.14) дает уже знакомую нам функцию Кобба — Дугласа.

Случай $\sigma=0$ можно получить предельным переходом $\sigma\to 0$ ($\rho\to\infty$) из формулы (1.16):

$$\lim_{\rho\to\infty}\left[\delta K^{-\rho}+(1-\delta)L^{-\rho}\right]^{-\gamma/\rho}=\min\{K^\gamma,\ L^\gamma\}. \qquad (1.17)$$

Полученная функция называется производственной функцией *с фиксированными пропорциями*. Эту производственную функцию обычно записывают в виде $F(K,L)=\min\{aK^\gamma, bL^\gamma\}$.

При $\gamma=1$ получаем так называемую *функцию Леонтьева* $Y=\min\{aK, bL\}$; здесь коэффициент a имеет смысл фондоотдачи, коэффициент b соответствует производительности труда.

Отметим, что экономически равенство $\sigma=0$ означает отсутствие замещения факторов, что вполне согласуется с характером функции (1.17): в случае $K\neq L$ полностью используется фактор, имеющийся в минимальном количестве, а второй остается использованным не полностью.

§ 4. Конструирование производственных функций

Важнейшим этапом в построении производственной функции конкретного экономико-хозяйственного объекта является выбор конечнопараметрического класса функций от факторов производства. Ориентиром при этом служат наблюдаемые значения показателей деятельности производственного объекта.

Так, если с помощью имеющейся статистики (Y_t, K_t, L_t) за разные периоды времени составлен разностный аналог показателя эластичности по первому фактору $\frac{Y_t-Y_\tau}{K_t-K_\tau}\frac{K_t}{Y_t}$ и оказалось, что при малой разности K_t-K_τ данный показатель незначительно колеблется около постоянного числа α, то этот факт служит весомым аргументом за то, что следует выбрать класс функций Кобба — Дугласа.

Здесь следует сказать о том, что показатели α, β эластичности производственной функции по факторам имеют особенно наглядную экономическую интерпретацию, что позволяет, в частности, выбирать класс производственных функций, ориентируясь на поведение величин α, β.

Большая величина отношения α/β говорит о том, что эластичность по первому фактору существенно выше эластичности по второму, т. е. что однопроцентное увеличение показателя K приводит к существенно большему увеличению объема производства, нежели однопроцентное увеличение показателя L. На практике, как правило, подобное обстоятельство говорит о том, что данное производство работает в условиях дефицита первого фактора. Вместе с тем обычным для реальных производственных объектов является то, что время от времени дефицитным становится то один, то другой фактор и, следовательно, величина эластичностей α, β меняется в зависимости от соотношения K, L.

Пример функций, учитывающих два режима работы производства в указанном выше смысле, дает класс CES. В самом деле, как несложно вычислить, для CES-функции (1.16) показатель α зависит только от фондовооруженности $k=K/L$ и равен

$$\alpha(k)=\gamma\Big/\left(1+\frac{1-\delta}{\delta}k^\rho\right).$$

График функции $\alpha(k)$ при $\rho > 1$ см. на рис. 1.2. Величина $\bar{k}$ есть корень уравнения $\alpha''(k) = 0$, т. е. $\bar{k} = \left(\frac{\delta}{1-\delta}\,\frac{\rho-1}{\rho+1}\right)^{1/\rho}$. Таким образом, выделяются две характерные области для CES-функции (1.16) — область $(0, \bar{k})$ дефицита фактора K и область $(\bar{k}, \infty)$ дефицита фактора L.

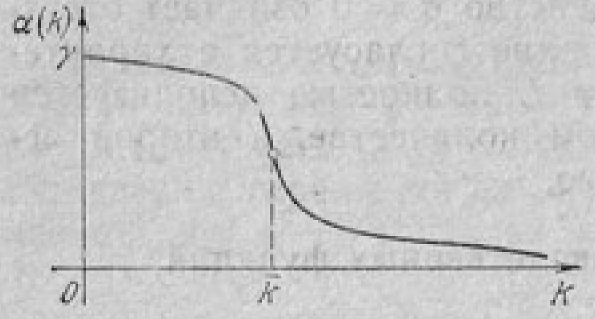

Рис. 1.2.

Предположим, что реальное поведение эластичности $\alpha(k)$ таково, что дефицитность фактора L носит не столь выраженный характер: $\lim\limits_{k\to\infty} \alpha(k) \neq 0$. Пусть, например, $\lim\limits_{k\to\infty} \alpha(k) = \alpha_0$, т. е. вид функции $\alpha(k)$ таков, как на рис. 1.3.

Подобную функцию можно аппроксимировать выражением вида

$$\alpha_1(k) = \gamma\Big/\left(1 + \frac{1-\delta}{\delta}k^{\rho}\right) + \alpha_0$$

и найти подходящую производственную функцию, решив уравнение

$$kf_1'/f_1 = \alpha_1(k). \qquad (1.18)$$

Разделяя переменные, получаем

$$\ln f_1 = \int \frac{\alpha_1(k)}{k}\,dk = \int \frac{\alpha(k)}{k}\,dk + \int \frac{\alpha_0}{k}\,dk.$$

Поскольку $\alpha(k)$ — эластичность некоторой CES-функции f, то $kf'/f = \alpha(k)$, или

$$\ln f = \int \frac{\alpha(k)}{k}\,dk.$$

Отсюда имеем $f_1 = Afk^{\alpha_0}$, где $f(k) = F(K/L, 1) = Y/L^{\gamma-1}$, $Y = F(K, L)$ — CES-функция степени однородности γ.

Следовательно, искомая функция Y_1 имеет вид

$$Y_1 = A(\delta K^{-\rho} + (1-\delta)L^{-\rho})^{-\gamma/\rho}K^{\alpha}L^{\beta},$$

т. е. является произведением CES-функции и функции Кобба — Дугласа. Степень ее однородности μ равна, очевидно, $\alpha + \beta + \gamma$.

Описанная процедура позволяет строить производственные функции, моделирующие производственную систему с двумя режимами работы. Однако, как отмечено в [49, 50], наиболее характерным для реальных экономико-хозяйственных объектов является наличие трех режимов работы: когда дефицитным является тот или иной ресурс и режим нормальной работы, когда

соотношение ресурсов сбалансировано согласно технологии изучаемого объекта.

Следовательно, наиболее реалистичным является вид зависимости эластичности по первому фактору, изображенный на рис. 1.4. Здесь области $(0, k_1)$, (k_2, ∞) отражают напряженный режим работы производственной системы при дефиците первого и второго факторов соответственно. Область (k_1, k_2) отвечает нормальному режиму работы.

Зависимость типа той, что изображена на рис. 1.4, можно получить, положив $\alpha(k) = \alpha_1(k) + \alpha_2(k) + \alpha_3(k)$, где $\alpha_1(k)$,

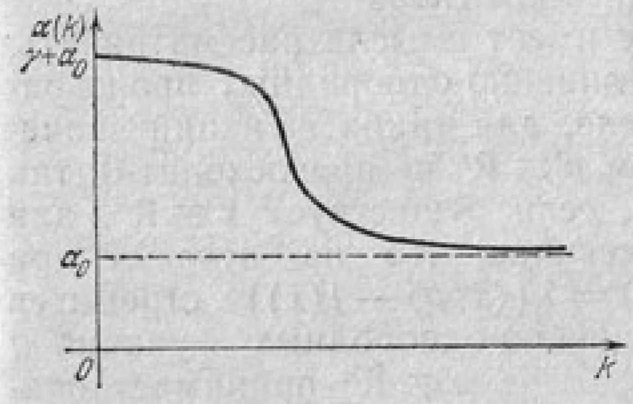

Рис. 1.3.

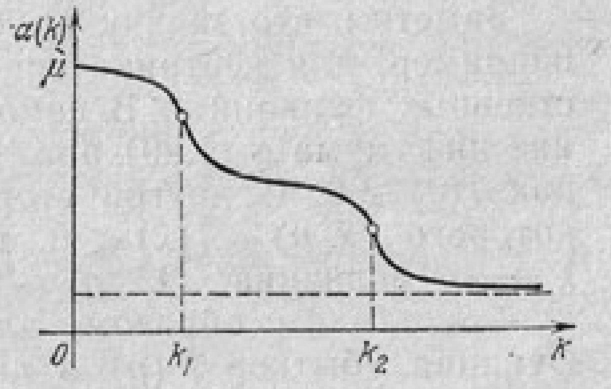

Рис. 1.4.

$\alpha_2(k)$ — эластичности подходящих CES-функций, $\alpha_3(k)$ — константа (эластичность Кобба — Дугласа). Рассмотрения, полностью аналогичные вышеизложенным, дают зависимость

$$F(K, L) = A\left(\delta_1 K^{-\rho_1} + (1-\delta_1)L^{-\rho_1}\right)^{-\gamma_1/\rho_1} \times$$
$$\times \left(\delta_2 K^{-\rho_2} + (1-\delta_2)L^{-\rho_2}\right)^{-\gamma_2/\rho_2}K^{\alpha}L^{\beta}, \qquad (1.19)$$

степень однородности которой равна $\mu = \gamma_1 + \gamma_2 + \alpha + \beta$.

Убедимся, что функция (1.19) удовлетворяет условиям (1.8), (1.9). Поскольку F однородна степени μ, то $F(K, L) = L^{\mu}f(k)$, где $f(k) = F(K/L, 1)$, $k = K/L$. Отсюда $f'(k) = \frac{f}{k}\alpha(k)$. По построению $\alpha(k) > 0$, следовательно, $f'(k) > 0$. Далее, $f''(k) = \frac{f}{k^2} \times$ $\times (k\alpha'(k) + \alpha(k)(\alpha(k) - 1))$. Поскольку $\alpha(k)$ — убывающая функция, то $\alpha'(k) < 0$, и вторая производная $f''(k)$ будет отрицательной, если степень однородности μ будет невысокой (например, при $\mu \leqslant 1$, ибо в этом случае $\alpha(k) \leqslant 1$ при всех $k \geqslant 0$).

§ 5. Двойственность для производственных функций

1. Пусть $y = f(x_1, x_2, \ldots, x_n)$ — вогнутая производственная функция, определенная для всех $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n_+$ и принимающая значения $y \in \mathbf{R}^1$. Предположим, что цена выпускаемой продукции принята за 1, а цена на j-й фактор равна

p_j, $j = 1, 2, \ldots, n$. Рассмотрим задачу

$$\inf_x (\langle x, p\rangle - f(x)), \tag{1.20}$$

которую можно интерпретировать как задачу планирования производства с целью максимизации прибыли (точнее, с целью минимизации убытков) — разности от стоимости $f(x)$ выпущенной продукции и затрат $\langle x, p\rangle$ на приобретение набора x факторов производства.

Значение $f^*(p)$ инфимума в (1.20) будем называть функцией, *сопряженной* к f, или *функцией убытков*.

Заметим, что задачу (1.20) не имеет смысла рассматривать, например, для неотрицательных линейно-однородных производственных функций f. В самом деле, для таких функций значение инфимума в (1.20) при любом $p \in \mathbf{R}^n$ не превосходит 0, так как $f(0) \geqslant 0$. С другой стороны, если существует $\bar{x} \in \mathbf{R}^n$, для которого $\langle \bar{x}, p\rangle - f(\bar{x}) < 0$, то очевидно, что на луче $\lambda\bar{x}$ при $\lambda \to \infty$ величина $\langle \lambda\bar{x}, p\rangle - f(\lambda\bar{x}) = \lambda(\langle \bar{x}, p\rangle - f(\bar{x}))$ стремится к $-\infty$. Таким образом, для линейно-однородных функций f функция убытков $f^*(p)$ в любой точке $p \in \mathbf{R}^n$ принимает значение либо 0, либо $-\infty$.

Наложив некоторые оправданные с содержательной точки зрения ограничения на класс функций f и рассматривая только положительные цены на факторы, можно избежать нежелательной ситуации, описанной выше. Проще, однако, поступить наоборот — допустить к рассмотрению функции f, принимающие значения $-\infty$.

Всюду ниже под расширенной действительной прямой $\bar{\mathbf{R}}^1$ понимается множество $\mathbf{R}^1$ действительных чисел, пополненное символом $-\infty$, так что для всякого $t \in \mathbf{R}^1$ выполняются соотношения

$$t > -\infty,\ t + (-\infty) = -\infty;\quad 0\cdot(-\infty) = 0,\ t\cdot(-\infty) = -\infty$$

при $t > 0$. Под словом «функция» будем понимать отображение в расширенную прямую $\bar{\mathbf{R}}^1$.

Введем обозначение

$$\operatorname{dom} f = \{x \in \mathbf{R}^n \mid f(x) > -\infty\}. \tag{1.21}$$

Определение. Функцию f назовем *собственной*, если она не равна тождественно $-\infty$, т. е. если $\operatorname{dom} f \neq \varnothing$.

Ослабим также обычное требование непрерывности функции f, для чего введем ряд новых понятий.

Определение. *Подграфиком* Γ_f функции f назовем следующее множество в $\mathbf{R}^n \times \bar{\mathbf{R}}^1$:

$$\Gamma_f = \{(x, t) \mid x \in \mathbf{R}^n,\quad t \in \bar{\mathbf{R}}^1,\quad t \leqslant f(x)\}.$$

На пространство $\mathbf{R}^n \times \bar{\mathbf{R}}^1$ без труда переносятся все понятия и основные факты выпуклого анализа: определение выпуклого множества, теоремы об отделимости и т. д.

Предоставим читателю доказать, что если функция $f(x)$ вогнута, то ее подграфик является выпуклым множеством в $\mathbf{R}^n \times \bar{\mathbf{R}}^1$.

Определение. Функцию f назовем *замкнутой*, если ее подграфик Γ_f замкнут в $\mathbf{R}^n \times \bar{\mathbf{R}}^1$.

Очевидно, что если f непрерывна, то она замкнута.

Основным объектом изучения в данном параграфе являются собственные вогнутые замкнутые функции.

Будем понимать задачу (1.20) так, что инфимум в ней берется по всем $x \in \operatorname{dom} f$.

Отметим простые свойства сопряженных функций.

1. $f(x) + f^*(p) \leqslant \langle x, p\rangle$ (*неравенство Юнга — Фенхеля*) — непосредственно следует из (1.20).

2. $f^*(p)$ — вогнутая функция. В самом деле, пусть $p_1, p_2 \in \mathbf{R}^n$, $\alpha, \beta \geqslant 0$, $\alpha + \beta = 1$. Тогда

$$f^*(\alpha p_1 + \beta p_2) = \inf_x (\langle \alpha p_1 + \beta p_2, x\rangle - f(x)) =$$
$$= \inf_x (\alpha(\langle p_1, x\rangle - f(x)) + \beta(\langle p_2, x\rangle - f(x))).$$

Поскольку для любых функций φ, ψ

$$\inf_x (\varphi(x) + \psi(x)) \geqslant \inf_x \varphi(x) + \inf_x \psi(x),$$

то

$$f^*(\alpha p_1 + \beta p_2) \geqslant \alpha f^*(p_1) + \beta f^*(p_2),$$

что и означает вогнутость функции f^*.

3. Если f — собственная вогнутая замкнутая функция, то f^* — собственная функция. Пусть $\bar{x} \in \operatorname{dom} f$. Тогда $(\bar{x}, f(\bar{x}) + 1) \notin \Gamma_f$. Поскольку Γ_f — непустое замкнутое множество, то его можно отделить от точки $(\bar{x}, f(\bar{x}) + 1)$, т. е. существует пара (p, β), для которой $\langle p, \bar{x}\rangle + \beta(f(\bar{x}) + 1) < \langle p, x\rangle + \beta t$ для любого $(x, t) \in \Gamma_f$. Допустим, что $\beta = 0$. Тогда неравенство $\langle p, \bar{x}\rangle < \langle p, x\rangle$ должно выполняться для всех $x \in \mathbf{R}^n$, что невозможно, так как оно нарушается для $x = \bar{x}$. С другой стороны, β не может быть положительным числом, так как в таком случае выражение $\langle p, x\rangle + \beta t$ становится сколь угодно малым при $t \to -\infty$. Следовательно, $\beta < 0$, откуда

$$\langle p|\beta|^{-1}, \bar{x}\rangle - (f(\bar{x}) + 1) < \langle p|\beta|^{-1}, x\rangle - t$$

для всех $(x, t) \in \Gamma_f$, или, иначе,

$$\inf_{x \in \operatorname{dom} f} (\langle \bar{p}, x\rangle - f(x)) > -\infty,$$

где $\bar{p} = p|\beta|^{-1}$. Это означает, что $f^*(\bar{p}) > -\infty$, т. е. f^* — собственная функция.

4. Сопряженная функция f^* замкнута. Действительно, пусть $(p_k, t_k) \in \Gamma_{f^*}$, причем $\lim\limits_{k\to\infty}(p_k, t_k) = (p_0, t_0)$. По определению Γ_{f^*}

$$f^*(p_k) \geqslant t_k, \quad k = 1, 2, \ldots,$$

откуда $\langle x, p_k\rangle - f(x) \geqslant t_k$, $k = 1, 2, \ldots$, для всех $x \in \operatorname{dom} f$. Переходя в этом неравенстве к пределу при каждом x, имеем $\langle x, p_0\rangle - f(x) \geqslant t_0$, откуда

$$\inf_x (\langle x, p_0\rangle - f(x)) \geqslant t_0,$$

или $f^*(p_0) \geqslant t_0$, что означает включение $(p_0, t_0) \in \Gamma_{f^*}$ и замкнутость Γ_{f^*}.

Таким образом, свойства 2, 3, 4 показывают, что класс собственных замкнутых вогнутых функций инвариантен по отношению к операции взятия сопряженной функции.

Для функции $f^*(p)$ можно также построить сопряженную:

$$f^{**}(x) = \inf_{p \in \operatorname{dom} f^*} (\langle x, p\rangle - f^*(p)). \tag{1.22}$$

Отметим очевидное соотношение между f и f^{**}:

$$f^{**}(x) \geqslant f(x). \tag{1.23}$$

В самом деле, согласно неравенству Юнга — Фенхеля

$$\langle x, p\rangle - f^*(p) \geqslant f(x);$$

следовательно,

$$f^{**}(x) = \inf_p (\langle x, p\rangle - f^*(p)) \geqslant f(x).$$

Предложим следующую интерпретацию задачи (1.22). Рассмотрим класс производственных функций $g(x)$, характеризующихся такой же, как для f, сопряженной функцией $f^*(p)$. Например, известно, что в условиях современной конкуренции все однотипные предприятия, фирмы имеют одинаковую прибыль, которая в нашем случае выражается величиной $-f^*(p)$. Зададимся вопросом: какова должна быть функция $g(x)$ из нашего класса, чтобы она при любых ценах гарантировала выпуск максимального количества продукта? Ответ на данный вопрос дается формулой (1.22): при ценах p максимально возможный выпуск продукта с использованием набора факторов x равен $\langle x, p\rangle + (-f^*(p))$ — сумме затрат и дохода. В таком случае гарантированная при любых ценах p величина выпуска выражается функцией $f^{**}(x)$. Таким образом, $f^{**}(x)$ — в некотором смысле наилучшая производственная функция, отвечающая заданной функции прибыли $-f^*(p)$. С этой точки зрения особенно интересен следующий замечательный факт.

Теорема 1.1. *Если f — собственная вогнутая замкнутая функция, то $f^{**} = f$.*

Доказательство. Как следует из (1.23), достаточно доказать неравенство $f^{**} \leqslant f$. Допустим, что $f^{**}(\bar x) > f(\bar x)$ в некоторой точке $\bar x \in \operatorname{dom} f^{**}$. Тогда подграфик Γ_f функции f, как непустое выпуклое и замкнутое множество, можно сильно отделить от точки $(\bar x, f^{**}(\bar x))$, т. е. существует пара $(p, \beta) \in \mathbf{R}^n \times \overline{\mathbf{R}}^1$, для которой

$$\langle p, \bar x\rangle + \beta f^{**}(\bar x) < \inf(\langle p, x\rangle + \beta t)$$

при всех $(x, t) \in \Gamma_f$. Очевидно, что $\beta < 0$. Разделив неравенство на $|\beta|$ и обозначая $\bar p = p|\beta|^{-1}$, имеем $\langle \bar p, \bar x\rangle - f^{**}(\bar x) < \inf(\langle \bar p, x\rangle - t)$ при всех $(x, t) \in \Gamma_f$. В частности,

$$\langle \bar p, \bar x\rangle - f^{**}(\bar x) < \inf_x (\langle \bar p, x\rangle - f(x)). \tag{1.24}$$

Очевидно, что (1.24) эквивалентно условию

$$\langle \bar p, \bar x\rangle - f^{**}(\bar x) < f^*(\bar p),$$

которое противоречит неравенству Юнга — Фенхеля. Теорема доказана.

Следствие. *Если $f^* = g^*$, то $f = g$.*

Данное следствие говорит о том, что производственная функция f полностью характеризуется своей функцией прибыли $-f^*$, что подтверждает целесообразность использования в экономической практике не только показателей выпуска, но и стоимостных оценок.

Изложенная конструкция сопряженной функции и теорема 1.1 носят название *теории Фенхеля — Моро*, по имени авторов, получивших основные результаты (в более общей ситуации, нежели изложенная нами).

2. Поскольку понятие неотрицательной линейно-однородной производственной функции широко используется в экономико-математическом моделировании, то неприменимость конструкции (1.20) к таким функциям заставляет искать иные формы изучения соотношений затраты — выпуск. Г. Б. Клейнером в [51] предложен следующий способ.

Рассматривается класс $\mathcal{F}$ неотрицательных непрерывных вогнутых функций $f(x) \not\equiv 0$, определенных для всех $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n_+$. Цены $p = (p_1, p_2, \ldots, p_n)$ на факторы также предполагаются неотрицательными. Обозначим через X_f множество

$$X_f = \{x \in \mathbf{R}^n_+ \mid f(x) > 0\}.$$

По предположению $X_f \neq \varnothing$.

Рассмотрим задачу

$$\inf_{x \in X_f} \frac{\langle p, x\rangle}{f(x)}, \tag{1.25}$$

которую можно интерпретировать как задачу планирования производства с целью обеспечить минимум затрат на производство количества продукции стоимостью в 1 руб. (считая, что цена выпускаемого продукта принята за единицу). Значение задачи (1.25) будем обозначать символом $f^\times(p)$. Отметим некоторые свойства функции $f^\times(p)$.

1. Областью определения $f^\times(p)$ служит весь неотрицательный ортант R^n_+. В самом деле, поскольку $f(x)$ не равна тождественно нулю, то, взяв $x \in R^n_+$, для которого $f(x) > 0$, имеем

$$f^\times(p) \leqslant \langle p, x\rangle / f(x) < \infty.$$

2. $f^\times(p) \geqslant 0$.

3. $f^\times(p)$ линейно-однородна. Действительно, при любом $\lambda > 0$

$$f^\times(\lambda p) = \inf_x \frac{\langle \lambda p,\ x\rangle}{f(x)} = \lambda \inf_x \frac{\langle p,\ x\rangle}{f(x)} = \lambda f^\times(p).$$

4. $f^\times(p)$ вогнута. Пусть $p_1, p_2 \in R^n_+$, α, $\beta \geqslant 0$, $\alpha + \beta = 1$. Тогда

$$f^\times(\alpha p_1 + \beta p_2) = \inf_x \frac{\langle \alpha p_1 + \beta p_2,\ x\rangle}{f(x)} = \inf_x \left(\alpha \frac{\langle p_1,\ x\rangle}{f(x)} + \beta \frac{\langle p_2,\ x\rangle}{f(x)}\right) \geqslant$$

$$\geqslant \alpha \inf_x \frac{\langle p_1,\ x\rangle}{f(x)} + \beta \inf_x \frac{\langle p_2,\ x\rangle}{f(x)} = \alpha f^\times(p_1) + \beta f^\times(p_2).$$

5. $f^\times(p) f(x) \leqslant \langle p,\ x\rangle$ для всех x, $p \in R^n_+$.

6. $f^\times(p)$ не равна тождественно нулю.

Действительно, поскольку функция $f(x)$ непрерывна, то ее подграфик $\Gamma_f = \{(x, t) \mid t \leqslant f(x)\}$ замкнут. Кроме того, в силу вогнутости f множество Γ_f выпукло. Это означает, что для произвольной граничной точки $(x_0, f(x_0))$ множества Γ_f существует опорная гиперплоскость. Замечая, что в силу линейной однородности функции f множество Γ_f является конусом в пространстве $R^n \times R^1$, получаем, что опорная гиперплоскость должна проходить через 0. Таким образом, существует пара $(p, \beta) \in R^n \times R^1$, для которой

$$\langle p, x_0\rangle + \beta f(x_0) = 0, \quad \langle p, x\rangle + \beta t \geqslant 0$$

для всех $(x, t) \in \Gamma_f$. В таком случае очевидно, что $p \geqslant 0$. Отсюда вытекает, что $\beta < 0$. Обозначим $\bar p = p|\beta|^{-1}$. Тогда

$$\langle \bar p, x_0\rangle = f(x_0), \quad \langle \bar p, x\rangle \geqslant f(x)$$

для всех $x \in R^n_+$. Следовательно, если $f(x) > 0$, то $\langle \bar p, x\rangle / f(x) \geqslant 1$, и

$$\inf_{x \in X_f} \frac{\langle \bar p,\ x\rangle}{f(x)} \geqslant 1,$$

т. е. $f^\times(\bar p) \geqslant 1$.

Следующее важное свойство функции $f^\times(p)$ сформулируем в виде леммы.

Лемма 1.1. *Функция $f^\times(p)$ непрерывна на R^n_+.*

Доказательство. Напомним, что произвольная функция $g(p)$ называется *полунепрерывной сверху* в точке p_0, если для любой последовательности p_k, $k = 1, 2, \ldots$, для которой $\lim\limits_{k\to\infty} p_k = p_0$ и предел $\lim\limits_{k\to\infty} g(p_k)$ существует, имеет место неравенство $\lim\limits_{k\to\infty} g(p_k) \leqslant g(p_0)$. Аналогично, если $\lim\limits_{k\to\infty} g(p_k) \geqslant g(p_0)$, то функция называется *полунепрерывной снизу* в точке p_0. Если функция полунепрерывна сверху и снизу, то она непрерывна.

Докажем вначале полунепрерывность сверху функции $f^\times(p)$. По определению

$$f^\times(p_k) \leqslant \langle p_k, x\rangle / f(x)$$

для всех $x \in X_f$. Следовательно, при каждом x

$$\lim_{k\to\infty} p^\times(p_k) \leqslant \langle p_0, x\rangle / f(x),$$

откуда

$$\lim_{k\to\infty} f^\times(p_k) \leqslant \inf_{x \in X_f} \frac{\langle p_0, x\rangle}{f(x)},$$

т. е. $\lim\limits_{k\to\infty} f^\times(p_k) \leqslant f^\times(p_0)$.

Перейдем к доказательству полунепрерывности снизу функции $f^\times(p)$. Пусть $p_0 \neq 0$ — произвольный вектор. Обозначим через $e_1, e_2, \ldots, e_n$ орты пространства R^n. Рассмотрим произвольные наборы вида $\sigma = \{p_0, e_{s_1}, e_{s_2}, \ldots, e_{s_{n-1}}\}$ и выберем среди них линейно независимые. Рассмотрим для каждого линейно независимого набора σ конус K_σ, натянутый на образующие этот набор векторы:

$$K_\sigma = \left\{p \mid p = \alpha_0 p_0 + \sum_{i=1}^{n-1} \alpha_i e_{s_i}, \alpha_i \geqslant 0, i = 0, 1, 2, \ldots, n-1\right\}.$$

Рисунки 1.5 и 1.6 иллюстрируют данную конструкцию для случая $n = 3$. На рис. 1.5 вектору p_0 соответствуют три конуса: $K(p_0, e_1, e_2)$, $K(p_0, e_2, e_3)$, $K(p_0, e_1, e_3)$. В то же время ситуации, изображенной на рис. 1.6, соответствуют только два конуса: $K(p_0, e_1, e_2)$, $K(p_0, e_2, e_3)$.

Поскольку всего конусов K_σ — конечное число, то по крайней мере в одном из них содержится бесконечное число членов последовательности $\{p_k\}$. Для простоты будем считать, что последовательность $\{p_k\}$ целиком лежит в конусе K_σ, где $\sigma = \{p_0, e_1, e_2, \ldots, e_{n-1}\}$. Разложим каждый вектор p_k по базису σ:

$$p_k = \alpha_0(k) p_0 + \sum_{i=1}^{n-1} \alpha_i(k) e_i.$$

Поскольку $f^{\times}$ вогнута и линейно-однородна, то

$$f^{\times}(p_k) \geqslant \alpha_0(k) f^{\times}(p_0) + \sum_{i=1}^{n-1} \alpha_i(k) f^{\times}(e_i) \tag{1.26}$$

при всех $k = 1, 2, \ldots$

Так как σ — базис, то из условия $\lim\limits_{k\to\infty} p_k = p_0$ вытекает, что $\lim\limits_{k\to\infty} \alpha_0(k) = 1$, $\lim\limits_{k\to\infty} \alpha_i(k) = 0$. В таком случае, переходя в (1.26) к пределу, получаем $\lim\limits_{k\to\infty} f^{\times}(p_k) \geqslant f^{\times}(p_0)$, что и требовалось доказать. Непрерывность $f^{\times}$ в точке $p = 0$ очевидна. Лемма доказана.

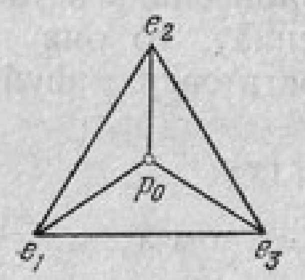

Рис. 1.5.

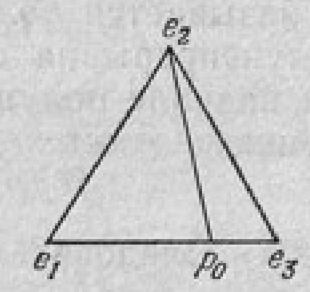

Рис. 1.6.

Из свойств 1—4, 6 и леммы 1.1 следует, что класс $\mathscr{F}$ неотрицательных линейно-однородных вогнутых функций на $\mathbf{R}^n_+$, не равных тождественно нулю, замкнут относительно операции $\times$.

Задаче (1.25) сопоставим двойственную:

$$f^{\times\times}(x) = \inf_{p \in X_f^{\times}} \frac{\langle p, x\rangle}{f^{\times}(p)}. \tag{1.27}$$

По доказанному $f^{\times\times}(x) \in \mathscr{F}$.

Интерпретация задачи (1.27) близка к нашему толкованию задачи (1.22): среди всех производственных функций, характеризующихся одинаковой зависимостью $f^{\times}(p)$ затрат на 1 руб. продукции, найти ту, которая обеспечивает при наихудшем соотношении цен на факторы наибольший выпуск продукции.

Отметим очевидную связь между $f^{\times\times}$ и f:

$$f^{\times\times} \geqslant f, \tag{1.28}$$

вытекающую из свойства 5.

Теорема 1.2. *Если f — вогнутая непрерывная неотрицательная линейно-однородная функция, не равная тождественно 0, то $f^{\times\times} = f$.*

Доказательство. Предположим, что существует $\bar{x}$, для которого $f^{\times\times}(\bar{x}) > f(\bar{x})$. Тогда подграфик Γ_f функции f, как выпуклое замкнутое множество в $\mathbf{R}^n \times \mathbf{R}^1$, можно строго отделить от точки $(\bar{x}, f^{\times\times}(\bar{x}))$ с помощью опорной к Γ_f гиперплоскости. Следовательно, существует пара $(p, \beta) \in \mathbf{R}^n \times \mathbf{R}^1$, для которой

$$\langle p, \bar{x}\rangle + \beta f^{\times\times}(\bar{x}) < 0, \quad \langle p, x\rangle + \beta t \geqslant 0$$

для всех $(x, t) \in \Gamma_f$. Отсюда следует, что $\beta < 0$, $p \geqslant 0$. Обозначив $\bar{p} = p|\beta|^{-1}$, получаем

$$\langle \bar{p}, \bar{x}\rangle < f^{\times\times}(\bar{x}), \tag{1.29}$$

$$\langle \bar{p}, x\rangle \geqslant f(x) \tag{1.30}$$

для всех $x \in \mathbf{R}^n_+$.

Из (1.30) имеем

$$f^{\times}(\bar{p}) = \inf_{x \in X_f} \frac{\langle \bar{p}, x\rangle}{f(x)} \geqslant 1.$$

Тогда с помощью (1.29) получаем $\langle \bar{p}, \bar{x}\rangle < f^{\times\times}(\bar{x}) f^{\times}(\bar{p})$, что противоречит свойству 5.

Следствие. *Если $f^{\times} = g^{\times}$, то $f = g$.*

Данный факт означает, что производственная функция полностью характеризуется своими стоимостными показателями — функцией затрат на 1 руб. продукции, что оправдывает существующую практику оценки работы экономико-хозяйственных объектов стоимостными характеристиками подобного типа.

Теоремы, доказанные в этом параграфе, аналогичны теоремам теории двойственности в линейном программировании. В первой части настоящей книги мы видели немало примеров, демонстрирующих мощь теории двойственности при исследовании качественных свойств линейных задач. Можно ожидать, что результаты этого параграфа сыграют такую же роль при изучении производственных функций.

Как следует из теорем 1.1 и 1.2 данного параграфа, вогнутая производственная функция $f(x)$ является в некотором смысле наилучшей производственной функцией среди класса всех функций, характеризующихся одинаковыми экономическими показателями: величинами максимальной прибыли и минимума затрат на 1 руб. продукции при ценах p. Из этого следует вывод, что гипотеза о вогнутости производственной функции $y = f(x)$ включает в себя предположение об оптимальном использовании имеющихся ресурсов, основных фондов и труда, описываемых вектором x, при выпуске объема y продукции.

В качестве подтверждения этого положения рассмотрим следующую линейную модель планирования производства. Имеется производственная система, выпускающая n наименований товаров, используя m видов ресурсов (основные фонды, трудовые ресурсы и т. д.). Коэффициент a_{ij} показывает необходимые затраты ресурса i-го вида на выпуск единицы j-го продукта. Пусть $b = (b_1, b_2, \ldots, b_m)$ — вектор ресурсов, имеющихся в наличии, $c = (c_1, c_2, \ldots, c_n)$ — вектор цен на товары, $x = (x_1, x_2, \ldots, x_n)$ — план выпуска.

Рассмотрим задачу $\Pi(b)$:

$$\max \langle c, x\rangle$$

$$Ax \leqslant b, \quad x \geqslant 0.$$

Пусть x^0 — решение этой задачи. Обозначим $f(b)=\langle x^0, c\rangle$. В таком случае $f(b)$ — вогнутая функция. В самом деле, пусть b', $b''\in\mathbf{R}^m$, $0\leqslant\alpha\leqslant 1$. Надо показать, что

$$f(\alpha b'+(1-\alpha)b'')\geqslant \alpha f(b')+(1-\alpha)f(b'').$$

Действительно, пусть x', x'' таковы, что $f(b')=\langle x', c\rangle$, $f(b'')=\langle x'', c\rangle$. Тогда

$$\alpha f(b')+(1-\alpha)f(b'')=\langle \alpha x'+(1-\alpha)x'', c\rangle.$$

Поскольку векторы x' и x'' допустимы для задач $\Pi(b')$ и $\Pi(b'')$ соответственно, то $Ax'\leqslant b'$, $Ax''\leqslant b''$, откуда $A(\alpha x'+(1-\alpha)x'')\leqslant \alpha b'+(1-\alpha)b''$, т. е. вектор $\alpha x'+(1-\alpha)x''$ допустим для задачи $\Pi(\alpha b'+(1-\alpha)b'')$, откуда следует, что значение целевой функции на этом векторе не превосходит ее максимальной величины:

$$\alpha f(b')+(1-\alpha)f(b'')\leqslant f(\alpha b'+(1-\alpha)b'').$$

Заметим, что если в качестве плана выбирать произвольный вектор $x(b)$, удовлетворяющий ограничениям задачи $\Pi(b)$, то функция $\varphi(b)=\langle c, x(b)\rangle$ не обязана быть вогнутой.

§ 6. Производственные функции и экзогенный научно-технический прогресс

1. До этого параграфа при обсуждении производственных функций мы не учитывали тот факт, что производственная структура любого участка народного хозяйства подвержена со временем изменениям, обусловленным разнообразными причинами, которые мы условно назовем научно-техническим прогрессом. Это явление было обнаружено при первых же попытках построения производственных функций и использования их как инструмента для получения прогнозов. Поэтому в настоящее время основным является понятие динамической производственной функции. Например, однопродуктовая двухфакторная производственная функция имеет вид $Y(t)=F(K(t), L(t), t)$. Мы по-прежнему будем предполагать функцию F линейно-однородной по факторам производства K и L:

$$F(\lambda K, \lambda L, t)=\lambda F(K, L, t).$$

На протяжении этого параграфа остаются действительными наши обозначения из § 1 для основных характеристик функции $F(K, L, t)$. Под *техническим прогрессом* будем понимать монотонное возрастание функции $F(K, L, t)$ по аргументу t.

Примерами производственных функций такого типа являются:

а) $Y=A(t)F_0(K, L)$, (1.31)

б) $Y=F_0(K, A(t)L)$, (1.32)

в) $Y=F_0(A(t)K, L)$. (1.33)

По мере практического изучения производственных функций было выдвинуто несколько гипотез относительно характера воздействия научно-технического прогресса на общественное производство. Как правило, эти гипотезы формулируются в терминах величин y, k, v, r, α, β, S, σ. Предполагается, что технический прогресс приводит к изменению этих величин. Например, качественное изменение состава основных фондов приводит к росту производительности труда, т. е. к увеличению показателей y и v. Естественно, однако, что технический прогресс воздействует, как правило, сразу на несколько, если не на все, из перечисленных показателей. Поэтому основным для классификации различных типов прогресса является сохранение во времени определенных зависимостей между этими величинами.

Назовем технический прогресс *Φ-нейтральным*, если $\Phi(y, k, v, r, \alpha, \beta, S, \sigma)=0$, где Φ — некоторая функция. Частными случаями понятия Φ-нейтральности являются следующие три случая.

Технический прогресс называется *нейтральным по Хиксу*, если предельная норма замещения S является постоянной во времени функцией фондовооруженности, т. е.

$$S=\varphi(k), \tag{1.34}$$

где φ — некоторая функция. Соотношение (1.34) представляет собой дифференциальное уравнение первого порядка вида

$$\frac{f}{f'}-k=\varphi(k).$$

Пусть $f_0(k)$ — некоторое частное решение этого уравнения. Общее решение должно содержать одну произвольную константу (в нашем случае — произвольную функцию времени, не зависящую от k). Нетрудно видеть, что функция вида $A(t)f_0(k)$ является решением данного уравнения при любом $A(t)$. Таким образом, функция $y=A(t)f_0(k)$ является общим решением уравнения (1.34). Возвращаясь к переменным Y, K, L, получаем функцию вида (1.31).

Множитель $A(t)$ иногда интерпретируется как мера современного состояния технических знаний. Обычно считают $A(t)>0$, $A'(t)>0$.

Технический прогресс называется *нейтральным по Харроду*, если предельная фондоотдача является постоянной во времени функцией средней фондоотдачи: $r=\psi(z)$, или $f'=\psi(f/k)$. Пусть $f_0(k)$ — частное решение данного уравнения. Нетрудно видеть, что функция $f=A(t)f_0(k/A(t))$ удовлетворяет данному дифференциальному уравнению. В самом деле,

$$f'=f_0'\left(\frac{k}{A(t)}\right), \quad \psi\left(\frac{f}{k}\right)=\psi\left(\frac{f_0(k/A(t))}{k/A(t)}\right).$$

По определению функции $f_0(k)$ имеем

$$f_0'\left(\frac{k}{A(t)}\right) = \psi\left(\frac{f_0(k/A(t))}{k/A(t)}\right),$$

откуда $f' = \psi(f/k)$. Следовательно, функция $f = A(t) f_0(k/A(t))$ — общее решение исходного дифференциального уравнения. Возвращаясь к переменным Y, K, L, получаем, что в случае, когда технический прогресс нейтрален по Харроду, производственная функция имеет вид (1.32). Нейтральный по Харроду технический прогресс эквивалентен росту во времени численности рабочей силы.

Технический прогресс называется *нейтральным по Солоу*, если предельная производительность труда является постоянной во времени функцией средней производительности.

Рассмотрения, полностью аналогичные предыдущим, дают общий вид производственной функции, отражающей технический прогресс, нейтральный по Солоу, совпадающий с (1.33). Прогресс такого типа эквивалентен росту во времени основных производственных фондов.

Отметим, что перечисленные типы научно-технического прогресса отражают так называемый *экзогенный прогресс*, т. е. технологические изменения, которые не обуславливаются самой моделью (производственной функцией), а лишь учитываются ею. Так, технический прогресс, нейтральный по Хиксу, можно трактовать как учет растущего опыта работников, отражение современного состояния научно-технических знаний. При этом предполагается, что возрастание коэффициента $A(t)$ (*мультипликатора прогресса*) происходит независимо от остальных параметров модели, от их динамики.

Существуют, однако, модели, в которых происхождение научно-технического прогресса определяется внутри данной модели и зависит от действий участников экономической системы, моделируемой производственной функцией. Технический прогресс такого типа называется *эндогенным*. Рассмотрение эндогенного научно-технического прогресса приводит к весьма интересным и содержательным постановкам различных оптимизационных динамических задач. Примеры таких задач будут разобраны в гл. 2.

2. Введение понятия динамической производственной функции сразу же ставит перед исследователем новые по сравнению со статическим случаем проблемы. Здесь мы вкратце остановимся на некоторых из них. При описании однопродуктовой двухфакторной производственной функции мы интерпретировали величину основных производственных фондов K как число однотипных машин, а L — как число одинаковых в производственном смысле работников. Это предположение, представляющее собой существенную идеализацию и в случае статической производственной функции, конечно, становится ограничительным в динамическом варианте.

Действительно, при изучении динамики производства, выражающейся, в частности, в росте и появлении новых производственных фондов, нет основания считать, скажем, машины, созданные в разные годы, однотипными. Кроме того, появляется необходимость учитывать факт выбывания (амортизацию) старых машин. Вместе с тем нельзя уже считать однородным и труд людей, работающих на разнотипных машинах.

Пусть $K(t,\tau)$ — часть основных фондов, введенных в эксплуатацию в году τ и продолжающих функционировать к году t, $t \geqslant \tau$, $K(t,\tau) \leqslant K(\tau,\tau)$. Пусть $L(t,\tau)$ — часть общей рабочей силы $L(t)$, обслуживающая указанные машины. Тогда часть $Y(t,\tau)$ национального дохода $Y(t)$, произведенного в момент t на машинах выпуска года τ, равна

$$Y(t,\tau) = F(K(t,\tau),\ L(t,\tau),\tau). \tag{1.35}$$

Обратим внимание на то, что в (1.35) стоит функция $F(K, L, \tau)$, относящаяся к моменту τ. Тем самым неявно предполагается, что машины, произведенные в году τ, эксплуатируются в году t так же, как и в момент τ их создания. Тем самым здесь игнорируется возможность усовершенствования навыков рабочих, изменения организации работы и т. д. Отсюда валовый национальный доход $Y(t)$ в год t равен

$$Y(t) = \int_{-\infty}^{t} Y(t,\tau)\,d\tau. \tag{1.36}$$

В формулах (1.35), (1.36) остается открытым вопрос о выборе распределения $L(t,\tau)$ рабочей силы. Считая, что распределение рабочей силы должно быть подчинено задаче максимизации валового национального дохода $Y(t)$, и предполагая, что производственная функция (1.35) допускает перераспределение рабочей силы (т. е. эластичность замещения $\sigma > 0$), рассмотрим следующую задачу:

$$\max_{L(t,\tau)} \int_{-\infty}^{t} F(K(t,\tau), L(t,\tau),\tau)\,d\tau, \tag{1.37}$$

$$\int_{-\infty}^{t} L(t,\tau)\,d\tau = L(t), \tag{1.38}$$

где функции $K(t,\tau)$, $L(t,\tau)$ считаем непрерывными.

Задача (1.37), (1.38) представляет собой обычную изопериметрическую задачу вариационного исчисления, для которой уравнение Эйлера имеет вид

$$\frac{\partial F}{\partial L} = \mu(t), \tag{1.39}$$

где $\mu(t)$ — не зависящий от τ множитель Лагранжа. Как мы уже видели в § 2, для линейно-однородной производственной функции предельная производительность труда $\partial F/\partial L$ является функцией фондовооруженности k:

$$\frac{\partial F}{\partial L} = f(k, \tau) - k\frac{\partial f}{\partial k} = v(k, \tau).$$

Из основного требования (1.9) для производственных функций следует, что функция $v(k, \tau)$ при каждом фиксированном τ представляет собой монотонно возрастающую функцию k.

Из этого вытекает, что уравнение (1.39), которое мы запишем в виде $v(k, \tau) = \mu(t)$, имеет единственное решение $k^* = k^*(\mu(t), \tau)$; искомое распределение $L^*(t, \tau)$ рабочей силы определяется равенством

$$L^*(t, \tau) = \frac{K(t, \tau)}{k^*(\mu(t), \tau)}, \tag{1.40}$$

где $\mu(t)$ находится из уравнения (1.38):

$$\int_{-\infty}^{t} \frac{K(t, \tau)}{k^*(\mu(t), \tau)}\, d\tau = L(t). \tag{1.41}$$

Проиллюстрируем решение задачи (1.37), (1.38) на примере функции Кобба — Дугласа. Пусть

$$Y(t, \tau) = F(K(t, \tau), L(t, \tau), \tau) = A(\tau) K^{\alpha}(t, \tau) L^{\beta}(t, \tau),$$

где параметры α, β постоянны и, как обычно, $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$. Тогда $\partial F/\partial L = \beta A(\tau) k^{\alpha}(t, \tau)$, и решение соответствующего уравнения (1.39) имеет вид

$$k^*(t, \tau) = \left[\frac{\mu(t)}{\beta A(\tau)}\right]^{1/\alpha}.$$

Подставляя данное выражение в (1.41), получаем выражение для $\mu(t)$:

$$\mu(t) = \beta\left\{\frac{1}{L(t)}\int_{-\infty}^{t} K(t, \tau) A^{1/\alpha}(\tau)\, d\tau\right\}^{\alpha}.$$

Тогда

$$k^*(t, \tau) = \frac{1}{L(t) A^{1/\alpha}(\tau)}\int_{-\infty}^{t} K(t, \tau) A^{1/\alpha}(\tau)\, d\tau.$$

Окончательно по формуле (1.40) находим оптимальный вариант распределения рабочей силы:

$$L^*(t, \tau) = L(t)\frac{K(t, \tau) A^{1/\alpha}(\tau)}{\int_{-\infty}^{t} K(t, \tau) A^{1/\alpha}(\tau)\, d\tau}.$$

Подставляя выражение для $L^*(t, \tau)$ в (1.35) и (1.36), получаем значение величины $Y(t)$ валового национального дохода в момент времени t при оптимальном использовании трудовых ресурсов:

$$Y(t) = \left\{\int_{-\infty}^{t} A^{1/\alpha}(\tau) K(t, \tau)\, d\tau\right\}^{\alpha} L^{\beta}(t). \tag{1.42}$$

Если здесь обозначить

$$K(t) = \int_{-\infty}^{t} A^{1/\alpha}(\tau) K(t, \tau)\, d\tau, \tag{1.43}$$

то формула (1.42) принимает вид

$$Y(t) = K^{\alpha}(t) L^{\beta}(t). \tag{1.44}$$

Естественно считать производственной функцией именно выражение для значения валового национального дохода при оптимальном распределении ресурсов.

Производственная функция (1.44) имеет вид функции Кобба — Дугласа, если считать, что величина $K(t)$, определяемая равенством (1.43), представляет собой количество основных фондов в момент времени t. Отсюда возникает идея рассматривать функцию $A_0(\tau) = A^{1/\alpha}(\tau)$ как некоторый коэффициент переоценки основных фондов с целью иметь возможность сравнения качественно разнородных фондов.

Используя последнее обозначение, запишем нашу производственную функцию в виде

$$Y(t, \tau) = [A_0(\tau) K(t, \tau)]^{\alpha} L^{\beta}(t, \tau).$$

Видим, что в данном случае технологические изменения носят характер технического прогресса, нейтрального по Солоу. Оказывается, данный факт не случаен. Именно, если рассматривать задачу (1.37), (1.38) для линейно-однородной функции от двух факторов K и L и ввести класс функций

$$K(t) = \int_{-\infty}^{t} A(\tau) K(t, \tau)\, d\tau,$$

где $A(\tau)$ — произвольная непрерывная весовая функция, то можно доказать следующее утверждение: равенство

$$\max_{L(t,\tau)} \int_{-\infty}^{t} F(K(t,\tau), L(t,\tau), \tau)\, d\tau = H(K(t), L(t)),$$

где H — некоторая линейно-однородная функция, возможно тогда и только тогда, когда $F(K, L, t) = H(A(t)K, L)$, т. е. если технический прогресс нейтрален по Солоу.

3. Обратим внимание на одно существенное предположение в рассмотренной модели, неявно выраженное формулой (1.38). Именно, считается возможным свободное перераспределение трудовых ресурсов в любой момент времени t по всем имеющимся фондам. Данная гипотеза, конечно, идеализирует реальное положение дел. Мы имеем в виду тот факт, что в момент τ создания основных фондов, как правило, фиксируется раз и навсегда необходимое количество труда $L(\tau)$ для обслуживания конкретных машин и станков, составляющих эти основные фонды. Поэтому в задачу (1.37), (1.38) следовало бы добавить ограничение вида $L(t,\tau) \leqslant L(\tau)$, которое разрешает загружать основные фонды в меньшей степени, чем это возможно, но запрещает увеличивать объем занятых на них трудовых ресурсов больше нормы, поскольку это не приведет к росту выпуска.

Более реалистичным в этом смысле является подход, развиваемый Л. В. Канторовичем и его учениками (см. [42, 43, 38]) при исследовании чрезвычайно важной с практической точки зрения проблемы определения оптимального срока службы фондов.

Пусть $[m(t), t]$, где $m(t) \leqslant t$, — отрезок времени, характеризующийся тем, что в момент t функционируют лишь те станки и машины, которые были созданы в момент $\tau \in [m(t), t]$. Остальные фонды, созданные в более ранние периоды, считаются отслужившими свой срок и не используются в момент t.

Пусть $K(\tau)d\tau$ — объем основных фондов, созданных в промежуток времени $[\tau, \tau + d\tau]$, $L(\tau)d\tau$ — объем трудовых ресурсов, обслуживающих данные основные фонды, $F(K, L, \tau)$ — однородная по K и L производственная функция, значение которой равно величине выпуска, т. е.

$$F(K(\tau)d\tau,\ L(\tau)d\tau, \tau) = F(K(\tau),\ L(\tau),\ \tau)d\tau.$$

Тогда объем выпуска в момент t подсчитывается по формуле

$$Y(t) = \int_{m(t)}^{t} F(K(\tau), L(\tau), \tau)\, d\tau. \tag{1.45}$$

Если $l(t)$ — объем трудовых ресурсов в момент t, то в условиях полной занятости должно выполняться соотношение

$$\int_{m(t)}^{t} L(\tau)\, d\tau = l(t); \tag{1.46}$$

функция $l(t)$ предполагается заданной экзогенно, обычно считают, что

$$l(t) = l(t_0)\, e^{\lambda(t-t_0)},$$

где λ — темп роста населения.

В промежуток $[t, t+dt]$ создаются новые фонды $K(t)dt$, на изготовление которых тратится некоторая доля $\gamma(t)$, $0 \leqslant \gamma(t) \leqslant 1$, национального продукта:

$$K(t)\,dt = \gamma(t)\,Y(t)\,dt,$$

т. е.

$$K(t) = \gamma(t) \int_{m(t)}^{t} F(K(\tau), L(\tau), \tau)\, d\tau. \tag{1.47}$$

Если предположить, что функция $\gamma(t)$ задана заранее (например, $\gamma(t) \equiv \gamma$, где γ — определенная константа, $0 < \gamma < 1$), а функцию $m(t)$ считать непрерывной, то видно, что значение $Y(t)$ полностью определяется предысторией процесса, на которую в момент t повлиять уже невозможно. Тем не менее задача (1.45)—(1.47) оставляет возможность для оптимизации: задавать в момент t не значение функции (управления) $m(t)$, но значение ее производной с целью максимизировать темп роста (производной) величины $Y(t)$. В [42] выведено дифференциальное уравнение для $m'(t)$, дающее решение следующей задачи: как следует выбирать в момент t значение $m'(t)$, чтобы максимизировать $Y'(t)$? Это уравнение выглядит следующим образом:

$$\frac{\partial F(K(t), L(t), t)}{\partial L} = \frac{F(K(m(t)),\ L(m(t)), m(t))}{L(m, t)}. \tag{1.48}$$

В левой части данного уравнения стоит предельная производительность труда на фондах, созданных в момент t, т. е. норма эффективности труда в настоящий момент. Правая часть формулы (1.48) — средняя производительность труда на фондах, созданных в момент $m(t)$.

ГЛАВА 2

МОДЕЛИРОВАНИЕ ПРОЦЕССОВ РАСПРЕДЕЛЕНИЯ КАПИТАЛОВЛОЖЕНИЙ

§ 1. Потребление и накопление: оптимальные пропорции

1. Ряд вопросов в рамках теории производственных функций при исследовании их динамики связан с изучением потребления одновременно с процессом производства национального дохода. В настоящий момент существует большое количество моделей, в которых с разной степенью подробности описывается все многообразие взаимоотношений процессов производства и потребления. Чтобы дать представление о характере подобных моделей и проблемах, возникающих при их исследовании, изложим один из вариантов модели, предложенной Ф. Рамсеем (см. [93]).

Предполагается, что в каждый момент времени выпуск Y делится на потребление C и инвестиции (капиталовложения) I:

$$Y(t) = C(t) + I(t) = (1 - s(t))\,Y(t) + s(t)\,Y(t), \qquad (2.1)$$

$$0 \leqslant s(t) \leqslant 1.$$

Производственные фонды амортизируют с темпом $\mu > 0$. Это означает, что за единицу времени из строя выбывает μ-я часть имеющихся основных фондов. Таким образом, выполняется уравнение

$$\dot{K}(t) = s(t)\,F(K(t),\, L) - \mu K(t). \qquad (2.2)$$

Здесь неявно предполагается, что основные фонды однородны в течение всего рассматриваемого промежутка времени, технологических изменений производственной функции не происходит и трудовые ресурсы находятся на постоянном уровне.

В качестве критерия, подлежащего максимизации в плановом периоде $[0, T]$, примем общее удельное потребление (на одного работника) с дисконтированием:

$$\int_0^T \frac{C(t)}{L}\, e^{-\delta t}\, dt, \qquad (2.3)$$

где $\delta > 0$ — константа. Наложим также условие

$$\frac{K(T)}{L} \geqslant k_T > 0, \qquad (2.4)$$

которое интерпретируется как условие «экономического горизонта»: планирование накопления и потребления должно обеспечить определенный экономический потенциал за пределами рассматриваемого периода. Задача планирующего органа состоит в выборе функции $s(t)$, $0 \leqslant s(t) \leqslant 1$, $0 \leqslant t \leqslant T$, максимизирующей функционал (2.3) при ограничениях (2.1), (2.2), (2.4).

Перейдем в описанной задаче к новым переменным — фондовооруженности $k = K/L$ и удельному потреблению $c = C/L$, обозначив при этом, как обычно, $f(k) = F(K/L, 1)$.

Нужно максимизировать

$$\int_0^T (1 - s(t))\, f(k(t))\, e^{-\delta t}\, dt \qquad (2.5)$$

при ограничениях

$$\dot{k}(t) = s(t)\,f(k(t)) - \mu k(t), \qquad (2.6)$$

$$0 \leqslant s(t) \leqslant 1, \quad k(0) = k_0 > 0, \quad k(T) \geqslant k_T > 0. \qquad (2.7)$$

Всюду ниже будем предполагать, что функция $f(k)$ удовлетворяет неоклассическим условиям.

Прежде чем исследовать модель (2.5)—(2.7) с применением принципа максимума Понтрягина, поставим перед собой более скромную цель — изучить траектории, определяемые дифференциальным уравнением (2.6) и начальным условием $k(0) = k_0$ (классическую задачу Коши), считая управление $s(t)$ константой, $s(t) = s$, $0 < s < 1$:

$$\dot{k} = sf(k) - \mu k, \quad k(0) = k_0. \qquad (2.8)$$

Несомненно, что, используя аппарат теории дифференциальных уравнений, можно получить исчерпывающую информацию о траекториях данной задачи Коши. Однако, желая облегчить понимание материала и пользуясь тем, что встретившаяся нам задача чрезвычайно проста, мы не станем отсылать читателя к специальным руководствам, а проведем необходимое исследование полностью.

Вначале мы отвлечемся от начального условия $k(0) = k_0$ и рассмотрим простейшие из возможных траекторий дифференциального уравнения (2.8) — стационарные траектории вида $k(t) \equiv k_s$, где $k_s > 0$ — некоторое число. Индекс s нам напоминает, что изучаются траектории при фиксированном постоянном управлении s. Понятно, что число k_s определяет стационарную траекторию в том и только том случае, когда $0 = \dot{k}(t) = sf(k_s) - \mu k_s$, т. е. если k_s является корнем уравнения $f(k)/k = \mu/s$. Покажем, что решение k_s данного уравнения существует и единственно для любого $s > 0$. Из неоклассических условий получаем, что функция $\alpha(k) = f(k)/k$ монотонно убывает от ∞ до 0. В самом деле, с помощью правила Лопиталя для вычисления пределов получаем

$$\lim_{k \to 0} \alpha(k) = \lim_{k \to 0} f'(k) = \infty,$$

$$\lim_{k \to \infty} \alpha(k) = \lim_{k \to \infty} f'(k) = 0.$$

Осталось убедиться в монотонном убывании функции $\alpha(k)$. Однако $\alpha'(k) = \frac{f'(k)\,k - f(k)}{k^2}$, и нужно лишь доказать, что $f'(k)\,k < f(k)$ при любом $k > 0$. Воспользуемся формулой Ньютона — Лейбница

$$f(k) = \int_0^k f'(u)\, du.$$

Поскольку f' — убывающая функция (требование $f'' < 0$ в неоклассических условиях), то $f'(u) > f'(k)$ для всех $u \in [0, k]$. В таком случае

$$f(k) = \int_0^k f'(u)\,du > \int_0^k f'(k)\,du = f'(k)\,k.$$

Из проведенных рассуждений вытекает существование и единственность стационарного решения $k(t) \equiv k_s$ дифференциального уравнения (2.8) при любом $s > 0$.

Обозначим $\varphi(k) = sf(k) - \mu k$.

Нам понадобится утверждение общего характера. Пусть $\varphi(k)$ — дифференцируемая функция, обладающая следующими свойствами:

$\varphi(k) > 0$, если $0 < k < k_s$,

$\varphi(k) = 0$, если $k = k_s$,

$\varphi(k) < 0$, если $k > k_s$.

Тогда решение $k(t)$ задачи Коши $\dot{k} = \varphi(k)$, $k(0) = k_0$ при любом $k_0 > 0$ существует и единственно на всей полуоси $t \in [0, \infty)$, и при этом $\lim\limits_{t\to\infty} k(t) = k_s$. Действительно, записав дифференциальное уравнение в виде $dk/dt = \varphi(k)$, убеждаемся, что задача Коши эквивалентна уравнению

$$\int_{k_0}^{k(t)} \frac{du}{\varphi(u)} = t. \tag{2.8'}$$

Напомним, что интеграл

$$\int_{k_0}^{k_s} \frac{du}{(u - k_s)^{\lambda}}$$

расходится при $\lambda \geqslant 1$. Поскольку функция $\varphi(u)$ дифференцируема, то $\varphi(u) = \varphi(k_s) + \varphi'(k_s)(u - k_s) + o(u - k_s) = \varphi'(k_s)(u - k_s) + o(u - k_s)$. Отсюда очевидно, что найдется константа $c < 0$, для которой $\varphi(u) \leqslant c(u - k_s)$ при всех u, $k_0 \leqslant u \leqslant k_s$. Тогда $\frac{1}{\varphi(u)} \geqslant \frac{1}{c(u-k_s)}$, откуда заключаем, что интеграл $\int_{k_0}^{k_s} \frac{du}{\varphi(u)}$ расходится. Используя знакоопределенность функции $\varphi(u)$ в любом промежутке, не содержащем точки k_s, можно утверждать, что функция

$$I(k) = \int_{k_0}^{k} \frac{du}{\varphi(u)}$$

обладает следующими свойствами:

1) $I(k)$ определена при любом k, если $k_s \notin [k_0, k]$;

2) $I(k)$ — монотонная функция, неограниченно возрастающая при $k \to k_s$.

Следовательно, если $k_0 > 0$, $k_0 \neq k_s$, то для любого $t > 0$ существует единственное число $k(t)$, удовлетворяющее уравнению (2.8′). Очевидно, $k(t) \to k_s$ при $t \to \infty$. Если теперь $k_0 = k_s$, то из определения k_s видно, что в этом случае единственным решением задачи Коши служит траектория $k(t) \equiv k_s$.

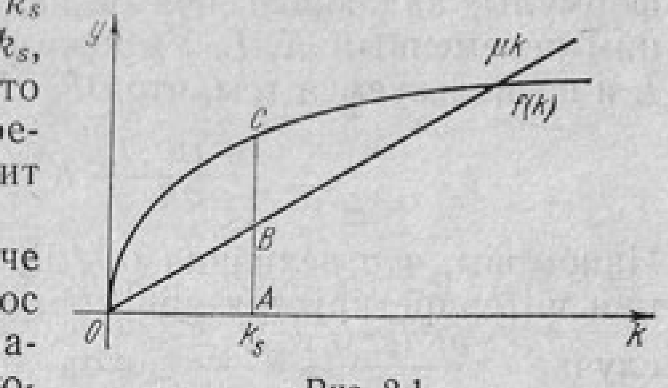

Рис. 2.1.

Теперь вернемся к задаче (2.5) — (2.7) и поставим вопрос об отыскании стационарной траектории $k(t) \equiv k_s$, максимизирующей потребление. Как следует из (2.5), при $s(t) \equiv s$, $k(t) \equiv k_s$ для максимизации постоянного потребления достаточно найти величину s, доставляющую наибольшее значение выражению $(1 - s) f(k_s)$.

Наиболее наглядным является графическое решение этой задачи. На рис. 2.1 построены графики функций $y = \mu k$ и $y = f(k)$. Пусть k_s — стационарная точка. Тогда, по определению k_s, $\mu k_s = sf(k_s)$; следовательно, вертикальный отрезок AB под прямой $y = \mu k$ равен по длине как раз $sf(k_s)$. Поэтому длина отрезка BC равна $f(k_s) - sf(k_s) = (1 - s) f(k_s)$ — максимизируемой функции. Поэтому проблема максимизации интегрального потребления (2.5) вдоль стационарных траекторий свелась к поиску точки k^*, для которой длина отрезка BC наибольшая.

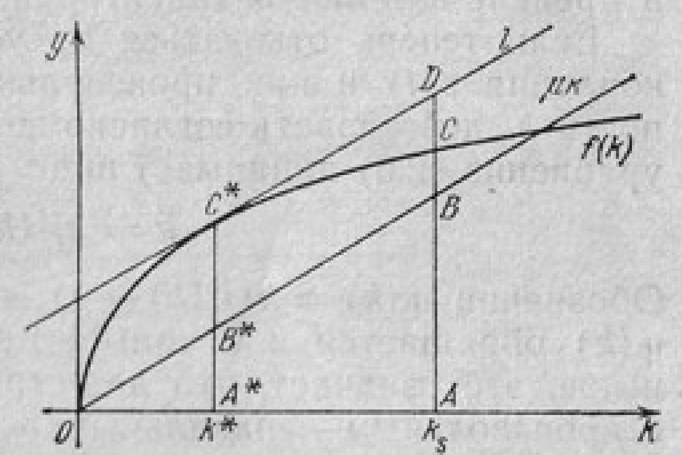

Рис. 2.2.

На рис. 2.2 точка k^* выбрана таким образом, что касательная l к графику функции $y = f(k)$ в точке C^* параллельна прямой $y = \mu k$. Поскольку область, находящаяся под графиком вогнутой функции $f(k)$, представляет собой выпуклое множество, для которого касательная l является гиперплоскостью, то точка C с координатами $(k, f(k))$ при $k \neq k^*$ лежит под прямой ниже точки D. Отрезки B^*C^* и BD, очевидно, равны, откуда получаем $|B^*C^*| > |BC|$.

Таким образом, стационарная точка k^* находится из условия равенства производных функций $y = f(x)$ и $y = \mu k$: $f'(k) = \mu$. Ясно, что решение k^* этого уравнения существует и единственно. Соответствующее управление s^*, оптимальное для стационарных траекторий, определяется тогда по формуле

$$s^* = \frac{\mu k^*}{f(k^*)} = \frac{f'(k^*)\,k^*}{f(k^*)}.$$

Для выяснения экономической интерпретации полученной формулы, задающей норму накопления s^*, вернемся к прежним переменным K, L. Умножим равенство $k^* f'(k^*) = s^* f(k^*)$ на L и воспользуемся тем, что $\partial F/\partial K = f'(k)$. Тогда

$$\frac{\partial F(K^*, L)}{\partial K} K^* = s^* F(K^*, L).$$

Напомним, что величина $\partial F/\partial K$ для линейно-однородной функции интерпретируется как норма прибыли с капитала. В таком случае $\frac{\partial F(K^*, L)}{\partial K} K^*$ — доход от капитала в фазовой точке (K^*, L). Отсюда получаем *золотое правило накопления*, сформулированное Е. Фелпсом в [92]:

Инвестиции в основные фонды должны равняться доходу, получаемому от капитала.

Как было отмечено выше, при управлении s^* для любой траектории с началом k_0 выполняется условие $\lim\limits_{t\to\infty} k(t) = k^*$. Это означает, что золотое правило автоматически выполняется в пределе для любой траектории.

Если теперь отказаться от условия постоянства нормы накопления $s(t)$ и при произвольной начальной фондовооруженности k_0 действовать согласно золотому правилу накопления, то уравнение (2.6) принимает вид

$$\dot k = k f'(k) - \mu k.$$

Обозначив $\varphi(k) = k(f'(k) - \mu)$, видим, что при $k > 0$ функция $\varphi(k)$ обращается в 0 только при $k = k^*$. Как было показано выше, это означает, что для траектории $k(t)$ этого уравнения с произвольным начальным $k_0 > 0$ выполняется условие $\lim\limits_{t\to\infty} k(t) = k^*$.

2. В работе [53] приведены аналогичные рассмотрения для двухсекторной модели экономики, берущей свое начало от схем простого и расширенного воспроизводства К. Маркса.

Рассматривается экономика, состоящая из двух секторов, характеризующихся производственными функциями $Y_1 = F_1(K_1, L_1)$, $Y_2 = F_2(K_2, L_2)$. Первый сектор производит средства производства, второй — предметы потребления. В соответствии со сказанным инвестиции в основные фонды как первого, так и второго сектора осуществляются только за счет выпуска Y_1 первого сектора, а потребление C совпадает с выпуском Y_2 второго сектора. Для простоты считаем общий объем $L = L_1 + L_2$ трудовых ресурсов постоянным и подлежащим распределению между секторами в каждый момент времени. Исходя из высказанных предпосылок, получаем систему уравнений

$$\begin{aligned}
\dot K_1 &= sF_1(K_1, L_1) - \mu K_1,\\
\dot K_2 &= (1-s) F_1(K_1, L_1) - \mu K_2,\\
C &= F_2(K_2, L_2),\\
L_1 &= qL,\\
L_2 &= (1-q) L,
\end{aligned}$$

где $0 \leqslant s \leqslant 1$, $0 \leqslant q \leqslant 1$.

Обозначив, как обычно, $k_1 = K_1/L_1$, $k_2 = K_2/L_2$, $f_1(k_1) = F_1(k_1, 1)$, $f_2(k_2) = F_2(k_2, 1)$, из первых двух уравнений системы получаем

$$\dot k_1 = s f_1(k_1) - \mu k_1, \quad \dot k_2 = \frac{q(1-s)}{1-q} f_1(k_1) - \mu k_2.$$

Предполагаем, что f_1, f_2 — неоклассические функции.

Вновь рассмотрим случай, когда управления постоянны: $s(t) \equiv s$, $q(t) \equiv q$, $0 \leqslant s, q \leqslant 1$. Так же, как и выше, можно показать, что при фиксированных s, q наша система имеет единственное стационарное решение $\bar k_1 = k_1(s, q)$, $\bar k_2 = k_2(s, q)$, определяемое из уравнений $\frac{f_1(k_1)}{k_1} = \frac{\mu}{s}$, $\frac{f_1(k_1)}{k_2} = \frac{\mu(1-q)}{q(1-s)}$.

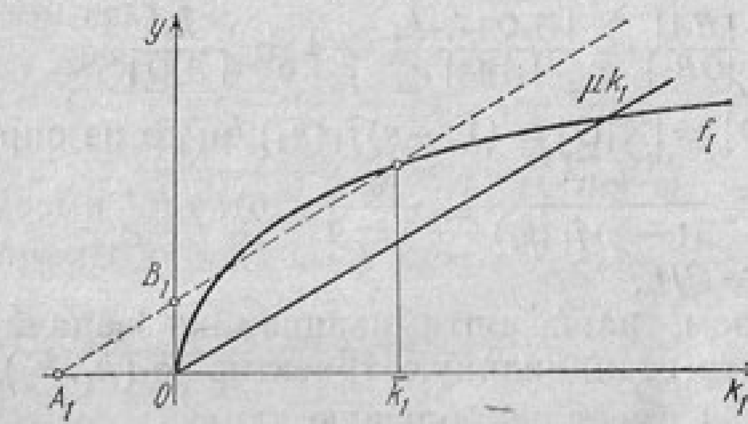

Рис. 2.3.

Поставим задачу максимизации потребления на душу населения на стационарной траектории:

$$\max \frac{C}{L} = \max \frac{F_2(K_2, L_2)}{\frac{1}{1-q} L_2} = \max (1-q) f_2(k_2).$$

Обратимся вновь к графическим рассмотрениям. На рис. 2.3 построены график функции $y = f_1(k_1)$ и прямая $y = \mu k_1$. Штриховая прямая, параллельная прямой $y = \mu k_1$, задается уравнением $y = \mu(k_1 - \bar k_1) + f_1(\bar k_1)$ и отсекает на оси y отрезок OB_1, длина которого равна $|OB_1| = -\mu\bar{k}_1 + f_1(\bar{k}_1) = (1-s) f_1(\bar{k}_1)$. Длина отрезка OA_1, отсекаемого этой прямой на оси k_1, равна $|OA_1| = (1-s) f_1(\bar{k}_1)/\mu$.

На рис. 2.4 построен график функции $y = f_2(k_2)$ и на оси k_2 отложена точка A_2 так, что $|A_2O| = |A_1O|$. Затем точка A_2 соединена штриховой прямой с точкой $E(\bar{k}_2, f_2(\bar{k}_2))$ графика $y = f_2(k_2)$. Вычислим координату точки пересечения B_2 этой

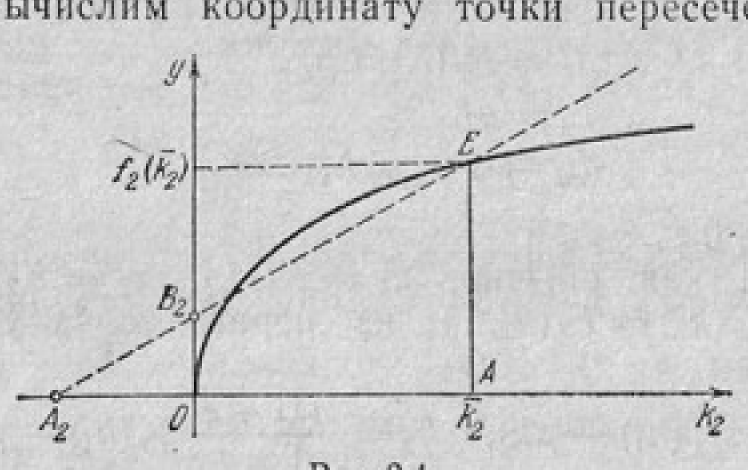

Рис. 2.4.

прямой с осью y. Из подобия треугольников A_2OB_2 и A_2AE находим

$$\frac{|EA|}{|OB_2|} = \frac{|A_2O| + \bar{k}_2}{|A_2O|} = 1 + \frac{\bar{k}_2}{|A_2O|}.$$

Поскольку $|A_2O| = |A_1O| = (1-s) f_1(\bar{k}_1)/\mu$, то из определения $\bar{k}_2$ получаем $\frac{\bar{k}_2}{|A_2O|} = \frac{\bar{k}_2\mu}{(1-s) f_1(\bar{k}_1)} = \frac{q}{1-q}$, откуда имеем $|OB_2| = (1-q) f_2(\bar{k}_2) = C/L$.

Таким образом, наша оптимизационная задача свелась к следующей: найти стационарную траекторию (k_1^*, k_2^*), для которой отрезок $|OB_2|$ имеет наибольшую длину.

Заметим, что длина отрезка A_1O (а значит, и A_2O) зависит только от точки $\bar{k}_1$, т. е. от выбора управлений s и q. При фиксированном A_2O отрезок OB_2 максимален, когда штриховая прямая на рис. 2.4 касается графика функции f_2. При увеличении отрезка OA_2 (т. е. OA_1) отрезок OB_2 увеличивается. Отрезок OA_1 максимален, когда штриховая прямая на рис. 2.3 параллельна прямой $y = \mu k_1$.

Получаем следующий алгоритм нахождения оптимальной стационарной траектории (k_1^*, k_2^*). Сначала надо провести касательную к графику $y = f_1(k_1)$ параллельно прямой $y = \mu k_1$, т. е. найти решение k_1^* уравнения $f_1'(k_1) = \mu$. Затем отложить на оси k_2 отрезок OA_2^*, равный OA_1^*. Из точки A_2^* следует провести касательную к графику $y = f_2(k_2)$. Абсцисса точки касания и есть k_2^*.

Оптимальные для стационарных траекторий постоянные управления s^* и q^* после этого находятся по формулам

$$s^* = \frac{\mu k_1^*}{f_1(k_1^*)} = \frac{k_1^* f_1'(k_1^*)}{f_1(k_1^*)}, \quad q^* = \frac{\mu k_2^*}{f_2(k_2^*)} = \frac{k_2^* f_2'(k_2^*)}{f_2(k_2^*)}.$$

Обсудим экономический смысл полученного оптимального решения. Возвращаясь к переменным K, L, получаем

$$\frac{\partial F_1(K_1^*, L_1^*)}{\partial K_1} K_1^* = s^* F_1(K_1^*, L_1^*).$$

Таким образом, инвестиции в первый сектор равны доходу с капитала первого сектора. Воспользовавшись теоремой Эйлера об однородных функциях:

$$\frac{\partial F_1}{\partial K_1} K_1 + \frac{\partial F_1}{\partial L_1} L_1 = F_1,$$

получаем

$$(1-s^*) F_1(K_1^*, L_1^*) = \frac{\partial F(K_1^*, L_1^*)}{\partial L_1} L_1^*.$$

Следовательно, инвестиции во второй сектор равны доходу от труда в первом секторе.

Поскольку $f_2'(k_2^*) k_2^* = q^* f_2(k_2^*)$, то

$$\frac{\partial F_2(K_2^*, L_2^*)}{\partial K_2} K_2^* = q^* F_2(K_2^*, L_2^*).$$

Вновь из теоремы Эйлера получаем

$$\frac{\partial F_2(K_2^*, L_2^*)}{\partial L_2} L_2^* = (1-q^*) F_2(K_2^*, L_2^*).$$

Отсюда следует, что отношение $q^*/(1-q^*)$ трудовых ресурсов, занятых в первом и втором секторах, равно отношению доходов с капитала второго сектора к доходам от труда во втором секторе.

Сформулируем теперь полностью *золотое правило накопления* в применении к рассматриваемой двухсекторной модели экономики.

При оптимальном стационарном развитии экономики:

а) инвестиции в первый сектор равны доходу с капитала первого сектора, а инвестиции во второй сектор равны доходу от труда первого сектора;

б) трудовые ресурсы между первым и вторым секторами делятся пропорционально доходу с капитала второго сектора и доходу от труда во втором секторе.

Отметим, что если рассматривать не обязательно стационарные траектории (но с постоянными управлениями s и q), то для всякой траектории с произвольными начальными данными золотое правило выполняется в пределе при $t \to \infty$.

Рассмотрим один пример. Пусть производственные функции F_1 и F_2 суть функции Кобба — Дугласа, т. е. $F_1 = A_1 K_1^{\alpha_1} L_1^{\beta_1}$, $F = A_2 K_2^{\alpha_2} L_2^{\beta_2}$, где $\alpha_i + \beta_i = 1$, $\alpha_i, \beta_i \geqslant 0$. В этом случае золотое правило принимает следующий вид: а) инвестиции в первый сектор составляют α_1-ю часть выпуска первого сектора, а инвестиции во второй сектор — β_1-ю часть выпуска первого сектора; б) трудовые ресурсы первого сектора должны составлять α_2-ю часть, а трудовые ресурсы второго сектора — β_2-ю часть всех трудовых ресурсов.

Что произойдет, если отказаться от постоянства управлений s и q и воспользоваться соотношениями

$$s(t) = \frac{\partial F_1}{\partial K_1}\frac{K_1}{F_1}, \quad q(t) = \frac{\partial F_2}{\partial K_2}\frac{K_2}{F_2},$$

подсказанными золотым правилом? В общем случае ответ на этот вопрос не столь прост, однако для случая с функциями Кобба — Дугласа проблем не возникает, так как $s(t) = \alpha_1$, $q(t) = \beta_1$, т. е. при действиях по золотому правилу управления оказываются постоянными.

3. Перейдем к исследованию модели (2.5) — (2.7), не накладывая априори каких-либо ограничений на функцию $s(t)$, $0 \leqslant s(t) \leqslant 1$. Решение этой задачи проводится на основе принципа максимума Понтрягина. Можно считать, что двойственная переменная ψ_0 равна 1. В самом деле, если $\psi_0 = 0$, то, как нетрудно показать (это мы оставляем читателю), оптимальным будет управление $s(t) \equiv 1$. Это означает, что плановое задание k_T слишком велико — все ресурсы уходят на его выполнение. Поскольку этот случай нереалистичен, мы его отбрасываем.

Пусть $\psi(t)$ — двойственная переменная, соответствующая уравнению (2.6). Тогда гамильтониан нашей задачи имеет вид

$$\mathscr{H}(\psi, k, s, t) = (1 - s) e^{-\delta t} f(k) + \psi[s f(k) - \mu k].$$

Если план $s(t)$ оптимален, то в силу принципа максимума существует такая непрерывная функция $\psi(t)$, что выполнены условия (2.6), (2.7) и

$$\dot\psi(t) = -\frac{\partial \mathscr{H}}{\partial k} = -(1 - s) e^{-\delta t} f'(k) - \psi[s f'(k) - \mu]; \qquad (2.9)$$

функция $s(t)$ максимизирует выражение

$$(1 - s) e^{-\delta t} + \psi s \qquad (2.10)$$

при ограничениях $0 \leqslant s \leqslant 1$ и

$$\psi(T)[k(T) - k_T] = 0. \qquad (2.11)$$

Выражение (2.10) получается из $\mathscr{H}$ отбрасыванием членов, не зависящих от s. Введя обозначение $q = \psi e^{\delta t}$ (величина q имеет смысл «объективно обусловленной оценки» производственных фондов в момент времени t), из (2.10) легко получаем

$$s(t) = \begin{cases} 1, & \text{если } q > 1, \\ 0, & \text{если } q < 1, \\ 0 \leqslant s(t) \leqslant 1, & \text{есил } q = 1. \end{cases} \qquad (2.12)$$

Уравнение (2.9) при переходе к новой переменной принимает вид

$$\dot q(t) = (\delta + \mu) q - [(1 - s) + sq] f'(k). \qquad (2.13)$$

Таким образом, из принципа максимума вытекает, что если $s(t)$ — оптимальное управление, то оно удовлетворяет (2.12), и система дифференциальных уравнений

$$\begin{aligned} \dot k(t) &= s f(k) - \mu k, \\ \dot q(t) &= (\delta + \mu) q - (1 - s + sq) f'(k), \\ k(0) &= k_0, \quad k(T) \geqslant k_T, \\ q(t)&[k(T) - k_T] = 0 \end{aligned} \qquad (2.14)$$

имеет решение.

Отметим, что первое уравнение системы (2.14) связано со вторым именно через функцию $s(t)$, удовлетворяющую условию (2.12). Анализ системы дифференциальных уравнений (2.14) облегчается тем, что она автономна — в правые части уравнений не входит в явном виде время t. С целью исследовать качественное поведение решений системы (2.14) найдем участки знакопостоянства производных $\dot k(t)$ и $\dot q(t)$. Рассмотрим отдельно два случая.

А. $q > 1$. Тогда $\dot q(t) = (\delta + \mu) q - q f'(k)$, и уравнение $\dot q(t) = 0$ приводит к $f'(k) = \delta + \mu$. Учитывая условия, наложенные на неоклассическую производственную функцию $f(k)$, можно заключить, что данное уравнение имеет единственное решение $k^* > 0$. Далее, уравнение $\dot k(t) = f(k) - \mu k = 0$ с учетом неоклассических условий также дает единственный корень $\bar k > 0$. Нетрудно показать, что $k^* < \bar k$. В самом деле, из того, что $f'(k)$ является убывающей функцией, имеем

$$f(k^*) = \int_0^{k^*} f'(k)\, dk > \int_0^{k^*} (\delta + \mu)\, dk = (\delta + \mu) k^* > \mu k^*.$$

Таким образом, точка k^* удовлетворяет неравенству $f(k^*) > \mu k^*$. Поскольку функция $f(k)$ вогнута ($f''(k) < 0$), это условие вместе с равенством $f(\bar k) = \mu \bar k$ и дает $k^* < \bar k$.

Следовательно, в случае А имеем:

$$\begin{aligned} &\dot q(t) < 0 \text{ при } k < k^*, && \dot k(t) > 0 \text{ при } k < \tilde k,\\ &\dot q(t) = 0 \text{ при } k = k^*, && \dot k(t) = 0 \text{ при } k = \tilde k, \\ &\dot q(t) > 0 \text{ при } k > k^*; && \dot k(t) < 0 \text{ при } k > \tilde k. \end{aligned} \tag{2.15}$$

Б. $q < 1$. Тогда $\dot q(t) = (\delta + \mu) q - f'(k)$, и уравнение $\dot q(t) = 0$ дает

$$q = \frac{f'(k)}{\delta + \mu} > 0. \tag{2.16}$$

Кроме того, $\dot k(t) = -\mu k < 0$. Полученные сведения позволяют нам нарисовать фазовую диаграмму для системы дифференциальных уравнений (2.14), т. е. изобразить на плоскости (k, q)

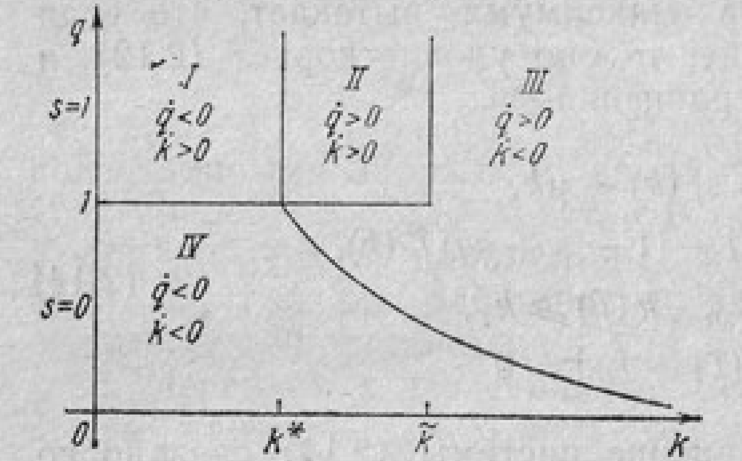

Рис. 2.5.

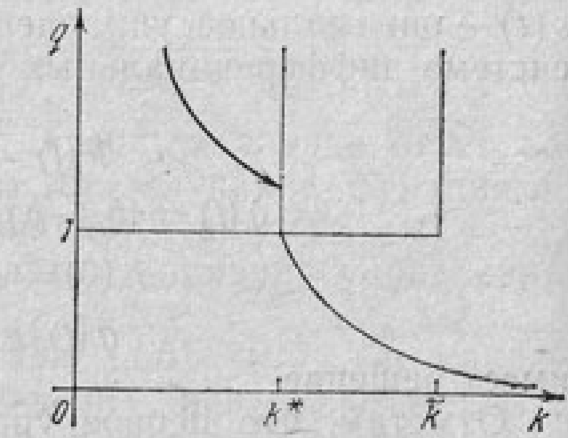

Рис. 2.6.

проекцию траекторий $(k(t),\ q(t))$, являющихся решениями системы.

С этой целью вначале сведем полученную информацию воедино. На рис. 2.5 изображено положение областей знакопостоянства производных $\dot q$ и $\dot k$ на плоскости (k, q). Здесь кривая, отделяющая области *IV* и *III*, задается уравнением (2.16).

Теперь нетрудно нарисовать фазовые траектории решений нашей системы в зависимости от начальных условий $k(0) = k_0$ и $q(0) = q_0$. Так, если точка (k_0, q_0) принадлежит области *I*, то система уравнений (2.14) эквивалентна системе

$$\dot k(t) = f(k) - \mu k, \quad \dot q(t) = (\delta + \mu) q - q f'(k).$$

Из общих теорем о существовании и единственности решения системы дифференциальных уравнений вытекает, что через заданную точку (k_0, q_0) области *I* проходит единственная кривая $(k(t),\ q(t))$, являющаяся решением системы. При этом фазовая траектория будет иметь один из видов, изображенных на рис. 2.6, 2.7 и 2.8. Подобные же рассуждения (и иллюстрации) можно было бы провести для каждого из случаев принадлежности точки (k_0, q_0) той или иной области.

Приведем сводную иллюстрацию типичных траекторий решения системы (2.14). На рис. 2.9 изображены траектории, не проходящие через точку $(k^*, 1)$. Из предыдущих рассуждений

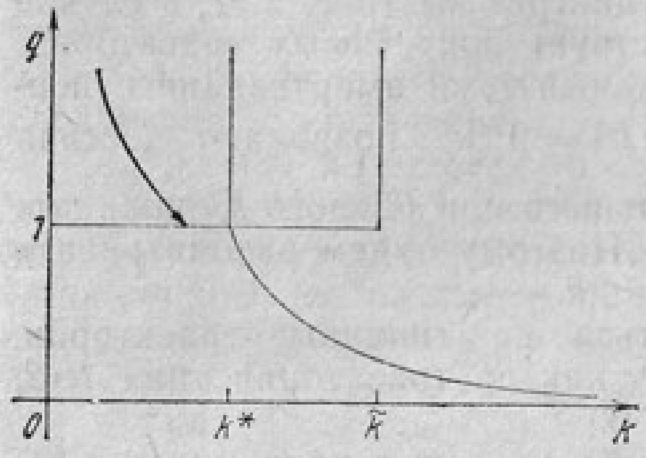

Рис. 2.7.

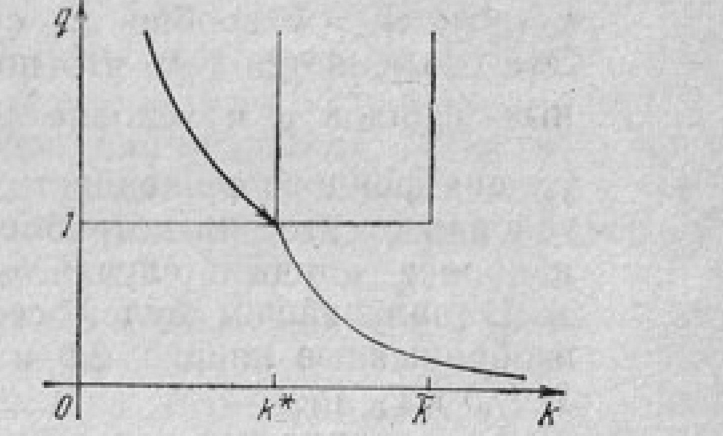

Рис. 2.8.

ясно, что каждая такая траектория полностью определяется начальными условиями (k_0, q_0). Здесь под словами «полностью определяется» мы имеем в виду не только график зависимости $q(k)$, но и «скорость» его прохождения траекторией $(k(t),\ q(t))$. Имеются всего две траектории, «входящие» в точку $(k^*, 1)$, и две траектории, «выходящие» из нее. Они условно изображены

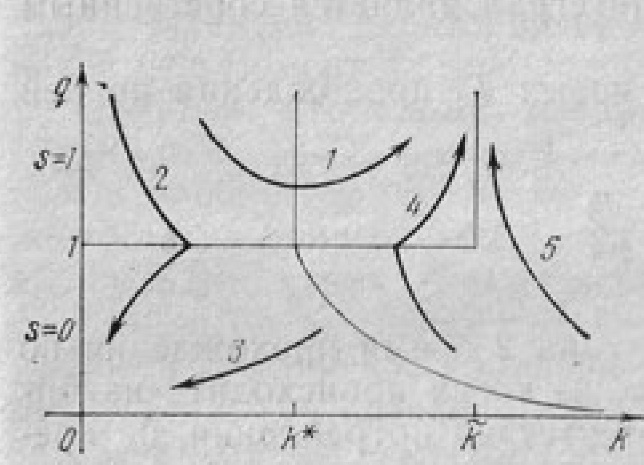

Рис. 2.9.

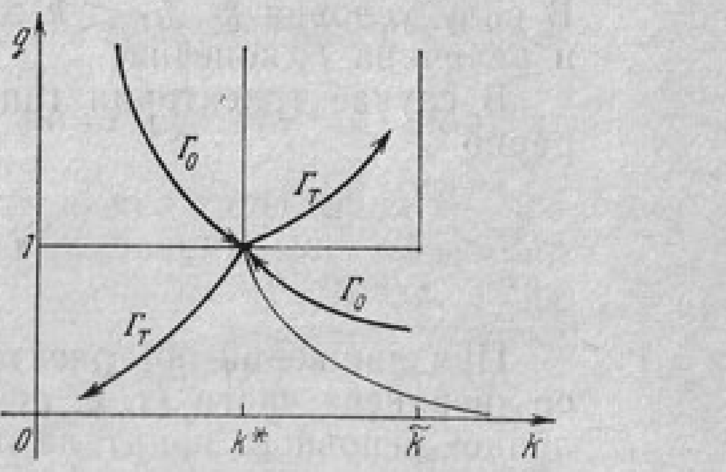

Рис. 2.10.

на рис. 2.10. Точка $(k^*, 1)$, которую мы назовем *особой* или *узловой*, обладает тем свойством, что, придя в нее, мы можем оставаться в ней сколь угодно долго, ибо она является общим корнем уравнений $\dot k(t) = 0,\ \dot q(t) = 0$. Для этого достаточно применять управление s^*, которое легко определить из этой системы уравнений: $s f(k) - \mu k = 0$, т. е. $s = \mu \frac{k}{f(k)}$, и $f'(k) = \delta + \mu$, т. е. $k = k^*$. Окончательно:

$$s^* = \mu \frac{k^*}{f(k^*)}.$$

Из предыдущих рассуждений легко вытекает, что $0 < s^* < 1$.

Отсюда ясно, что траектории, входящие в узловую точку $(k^*, 1)$, не определяются полностью начальными условиями, ибо время пребывания в этой точке зависит от управляющего субъекта. Как видно из фазовой диаграммы (рис. 2.9), в случае $k_0 < \bar{k}$, $k_T > \bar{k}$ вообще не существует допустимых траекторий. Это объясняется тем, что постоянный темп амортизации основных фондов μ и условие $\lim\limits_{t\to\infty} \dot{k}(t) = 0$ не позволяют достичь уровня фондовооруженности, большего или равного $\bar{k}$, даже при условии отсутствия потребления. Поэтому будем рассматривать наиболее типичный случай k_0, $k_T < \bar{k}$.

В дальнейшем будем ссылаться на типичные траектории, изображенные на рис. 2.9 и 2.10, как на траектории типа *1*, *2*, *3*, *4*, *5* и Γ_0, Γ_T.

Если начальное условие q_0 выбрано таким образом, что фазовая траектория является траекторией типа *1*, то время движения по этой траектории полностью определяется начальным условием k_0 и уравнением $\dot{k}(t) = f(k(t)) - \mu k(t)$ и равно

$$T_1 = \int\limits_0^{k_T} \frac{dk}{f(k) - \mu k}.$$

В силу условия k_0, $k_T < \bar{k}$ этот интеграл является собственным и величина T_1 конечна.

В случае траектории типа *3* время T_3 прохождения по ней равно

$$T_3 = \int\limits_{k_T}^{k_0} \frac{dk}{\mu k}.$$

При движении по траектории типа *2* время прохождения по ее «верхней» части (т. е. при $s = 1$), когда происходит «накопление» основного капитала при отсутствии потребления, не превосходит величины

$$T_2' = \int\limits_{k_0}^{k^*} \frac{dk}{f(k) - \mu k}.$$

Время же движения по «нижней» части траектории типа *2*, как нетрудно видеть, не превосходит величины

$$T_2'' = \int\limits_{k^*}^{k_T} \frac{dk}{\mu k}.$$

Таким образом, время прохождения всей траектории типа *2* ограничено числом $T_2 = T_2' + T_2''$.

Аналогичные рассуждения относительно траекторий типа *4*, *5* дают верхнюю оценку времени их прохождения, которые мы обозначим соответственно T_4 и T_5.

Пусть $T_0 = \max\{T_1, T_2, T_3, T_4, T_5\}$. Очевидно, что если $T > T_0$, то ни одна из траекторий типа *1*, *2*, *3*, *4*, *5* не может быть оптимальной. Проведенное исследование позволяет сформулировать следующий важный факт.

Теорема 2.1. *Пусть промежуток планирования T достаточно велик ($T > T_0$). Тогда, если в задаче оптимального планирования (2.5)—(2.8) существуют допустимые траектории, то существует и оптимальное управление $s(t)$, которое устроено следующим образом: в начале периода (при $0 \leqslant t \leqslant T^*$) и в конце (при $T^{**} \leqslant t \leqslant T$) $s(t)$ равно либо 0, либо 1. Все остальное время ($T^* \leqslant t \leqslant T^{**}$) $s(t) = s^*$. При этом величины T^* и T^{**} не зависят от T.*

Для доказательства теоремы к предыдущим рассуждениям достаточно добавить явное указание оптимальной траектории.

Следует выбрать начальное условие q_0 таким образом, чтобы точка (k_0, q_0) оказалась на траектории Γ_0, идущей в узловую точку $(k^*, 1)$. При этом управление равно либо 0, либо 1, в зависимости от соотношений величин k_0 и k_T. Отметим, что время T^* движения по траектории Γ_0 зависит только от k_0 и k^*.

Затем следует оставаться в точке $(k^*, 1)$ ровно столько, чтобы затем по траектории Γ_T попасть в точку с первой координатой k_T в момент T.

Отметим, что время движения по траектории Γ_T зависит только от k_T (и k^*, т. е. от функции f).

Сделаем оговорку. Может случиться, что траектория Γ_T, выходящая из равновесной точки $(k^*, 1)$ в направлении убывания k, проходит через точку вида $(\hat{k}, 0)$, где $\hat{k} > 0$. Тогда, если $k_T < \hat{k}$, то оптимальная траектория на своем конечном участке устроена следующим образом: надо находиться в состоянии равновесия $(k^*, 1)$ ровно столько, чтобы, выйдя из него, оказаться в момент T как раз в точке $(\hat{k}, 0)$. При этом терминальное условие $k(T) \geqslant k_T$ будет выполняться в виде строгого неравенства.

Содержательно это значает следующее: оказавшись в точке равновесия, нам выгодно поддерживать уровень фондовооруженности k^*, даже если потом (при $s = 0$) он не успеет амортизировать до уровня k_T.

Дело в том, что доля $(1 - s)k^*$, отчисляемая на потребление, выше, чем величина $k(t)$ при движении по рассматриваемой траектории Γ_T (точнее, надо говорить об интегралах от этих величин).

Чтобы показать, что такая возможность действительно имеет место, рассмотрим в качестве производственной функции f функцию Кобба — Дугласа вида k^α, где $0 < \alpha < 1$.

Система дифференциальных уравнений, описывающих рассматриваемую траекторию Γ_T, такова:

$$\dot{k} = -\mu k, \quad \dot{q} = (\mu + \delta) - f'(k),$$
$$k(0) = k^*, \ q(0) = 1.$$

Решение первого уравнения находится без труда: $k(t) = k^* e^{-\mu t}$. Решение однородного уравнения, соответствующего второму уравнению равно $q(t) = Ce^{(\mu+\delta)t}$. Отыскивая решение неоднородного уравнения методом вариации постоянной, получаем

$$q(t) = C(t)\, e^{(\mu+\delta)t},$$

где $C(t)$ удовлетворяет уравнению

$$\dot{C}(t) = -f'(k)\, e^{-(\mu+\delta)t}, \quad C(0) = 1.$$

Таким образом,

$$C(t) = 1 - \int_0^t f'(k(\tau))\, e^{-(\mu+\delta)\tau}\, d\tau.$$

Функция $q(t)$ обращается в нуль одновременно с функцией $C(t)$. Может ли функция $C(t)$ стать равной нулю, полностью зависит от поведения интеграла $\int_0^t f'(k(\tau))\, e^{-(\mu+\delta)\tau}\, d\tau$ как функции t. Решим уравнение $C(t) = 0$, или

$$\int_0^t f'(k(\tau))\, e^{-(\mu+\delta)\tau}\, d\tau = 1.$$

Вспоминая вид функций $f(k)$ и $k(\tau)$, получаем

$$\alpha (k^*)^{\alpha-1} \int_0^t e^{-\mu(\alpha-1)\tau} e^{-(\mu+\delta)\tau}\, d\tau = 1.$$

Напомним, что k^* является решением уравнения $f'(k) = \delta + \mu$. Окончательно находим решение $\hat{t}$ уравнения $C(t) = 0$:

$$\hat{t} = -\frac{1}{\alpha\mu + \delta} \ln\left(1 - \frac{\alpha\mu + \delta}{\mu + \delta}\right).$$

В силу условия $0 < \alpha < 1$ выражение под знаком логарифма лежит в интервале $(0, 1)$; следовательно, величина $\hat{t}$ действительна и положительна. Таким образом, функция Кобба — Дугласа дает нам требуемый пример.

Возвращаясь к переменным K, L и используя тот факт, что $f'(k^*) = \delta + \mu$, получаем

$$s^* F(K^*, L) = \frac{\partial F(K^*, L)}{\partial K} K^* - \delta K^*.$$

Видим, что в стационарной точке K^* для оптимальных траекторий модели (2.1)—(2.4) золотое правило накопления отличается от золотого правила Фелпса для стационарных траекторий: инвестиции уже не равны доходу от капитала, а меньше этого дохода на величину, зависящую от коэффициента дисконтирования δ. В том случае, когда $\delta = 0$, остается в силе правило Фелпса.

Если с самого начала решения задачи (2.5)—(2.7) отказаться от поиска оптимальной траектории, а действовать согласно выведенному модифицированному золотому правилу накопления, то уравнение (2.6) примет вид

$$\dot{k} = (f'(k) - \delta - \mu)\, k$$

и, как было показано выше, любое его решение $k(t)$ при произвольном $k_0 > 0$ будет удовлетворять условию $\lim_{t\to\infty} k(t) = k^*$. При этом, однако, может оказаться невыполненным условие на конец периода T (если $k_T > k^*$).

Теорема 2.1 весьма напоминает магистральные теоремы, знакомые нам по ч. I и относящиеся к многопродуктовым моделям производства Неймана и Гейла. Этой теореме можно придать и такую формулировку: при увеличении продолжительности T планового периода доля времени, когда на оптимальной траектории не выполняется золотое правило накопления, стремится к нулю.

Можно было бы рассмотреть более сложный вариант описанной модели. Если считать, что объем L трудовых ресурсов меняется во времени экзогенным образом, скажем, по экспоненциальному закону $L(t) = Le^{\lambda t}$, то анализ такой модели ничем не отличался бы от проведенного нами. Существенно более сложным является случай, когда учитывается технический прогресс, меняющий производственную функцию в зависимости от времени, скажем, если имеет место технический прогресс, нейтральный по Хиксу. Объясняется это в основном тем, что соответствующая система дифференциальных уравнений перестает быть автономной.

§ 2. Производственные функции в экологической модели

Естественно, что в начальный период своего существования теория производственных функций имела дело в основном только с вопросами собственно производства. Затем в моделях, описываемых производственными функциями, стали явно учитывать процессы потребления. Исследование наиболее рационального соотношения между производством и потреблением дало интересные качественные выводы, с которыми мы познакомились в § 1. В самое последнее время в связи с проблемой охраны окружающей среды появилось определенное число моделей, посвященных данному вопросу. Поскольку основным источником загрязнения среды является современное производство, то изучение возможностей установления контроля над загрязнением окружающей среды происходит в рамках теории производственных функций. Здесь мы изложим одну из таких моделей, построенную и исследованную в [48].

Мы не будем давать сколько-нибудь общего определения понятия загрязнения окружающей среды. Отметим только, что оно включает в себя чрезвычайно широкий круг явлений. Загрязнение воздуха и воды, увеличение количества ДДТ в почве, радиация — все это загрязнение в широком смысле.

Предполагается, что интересы общества выражаются функцией полезности $u(c, P)$, зависящей от двух аргументов: c — объем потребления, P — переменная, характеризующая объем загрязнения. При этом

$$\frac{\partial u}{\partial c} > 0, \quad \frac{\partial u}{\partial P} < 0, \quad \frac{\partial^2 u}{\partial c^2} < 0, \quad \frac{\partial^2 u}{\partial P^2} < 0, \qquad (2.17)$$
$$\lim_{c \to 0} \frac{\partial u}{\partial c} = \infty.$$

Интерпретация этих соотношений затруднений не вызывает, и мы на этом останавливаться не будем.

Сделаем дополнительно еще одно предположение о характере функции $u(c, P)$. Оно касается явления «замещения». Именно, если потребление c уменьшится на некоторую величину Δc, то для того, чтобы значение функции полезности не изменилось, требуется объем загрязнения P также уменьшить на некоторое ΔP. Как и в случае производственных функций, данный эффект определяется предельной нормой замещения

$$S = \frac{dP}{dc} = -\frac{\partial u/\partial c}{\partial u/\partial P}.$$

Наше предположение, накладываемое на функцию $u(c, P)$, будет состоять в том, что при малом уровне потребления для возмещения уменьшения c на одну единицу требуется уменьшить объем загрязнения на очень большую величину и, наоборот, при неограниченном возрастании c величина ΔP, необходимая для возмещения одной единицы потребления, стремится к нулю. Это означает, что линии уровня функции

$$S = -\frac{\partial u/\partial c}{\partial u/\partial P},$$

определяемые уравнениями вида $S(c, P) = S_0$, где $S_0 > 0$, имеют вид гипербол. Именно, кривая $P = P(c, S_0)$, определяемая данным уравнением, должна обладать следующими свойствами:

$$\lim_{c \to 0} P(c, S_0) = \infty, \quad \lim_{c \to \infty} P(c, S_0) = 0,$$

и $P(c, S_0)$ монотонно убывает при возрастании c.

Выведем формальные условия, обеспечивающие монотонное убывание функции $P(c, S_0)$. Рассмотрим вновь уравнение $S(c, P) = S_0$, задающее кривую $P = P(c, S_0)$, и вычислим производную

$$\frac{dP}{dc} = -\frac{\partial S/\partial c}{\partial S/\partial P}.$$

Тогда условие убывания функции $P = P(c, S_0)$ принимает вид

$$-\frac{\partial S/\partial c}{\partial S/\partial P} < 0.$$

Вычисляя частные производные функции $S = S(c, P)$, получаем, что требуемое условие выполняется при

$$\left(\frac{\partial^2 u}{\partial c^2} - \frac{\partial u/\partial c}{\partial u/\partial P}\,\frac{\partial^2 u}{\partial c\,\partial P}\right)\left(\frac{\partial^2 u}{\partial P^2}\,\frac{\partial u/\partial c}{\partial u/\partial P} - \frac{\partial^2 u}{\partial c\,\partial P}\right) < 0.$$

Данное неравенство справедливо, в частности, если $\frac{\partial^2 u}{\partial c\,\partial P} \leqslant 0$. Примером функции полезности, удовлетворяющей указанным условиям, служит функция $u(c, P) = Ac^{\alpha} - BP^{\beta}$, где $A, B > 0$; $0 < \alpha < 1$, $\beta > 1$.

В качестве критерия, подлежащего максимизации, принимается интеграл от функции полезности вдоль конкретной траектории $c(t)$ и $P(t)$ с учетом дисконтирования:

$$w = \int_0^T u(c, P)\, e^{-rt}\, dt. \qquad (2.18)$$

В качестве производственной функции рассматривается уже знакомая нам неоклассическая однопродуктовая двухфакторная функция, аргументами которой служат, как обычно, объем K основного капитала и объем L трудовых ресурсов. Вновь, с целью не усложнять изложение, будем считать, что предложение трудовых ресурсов не изменяется во времени. Предполагаем, что основной капитал амортизирует с постоянным темпом $\mu > 0$.

Загрязнитель не используется в производстве как полезный продукт (что, например, имеет место в случае с ДДТ), а является его побочным продуктом. Считаем, что объем загрязнителя прямо пропорционален объему продукта производства и составляет от него долю ε, $0 < \varepsilon < 1$. Примером подобного производства может служить, например, металлургическая отрасль, производство бумаги и т. д. Таким образом, загрязнение измеряется в тех же единицах, что и основная продукция.

Как известно, окружающая среда обладает определенной способностью ассимилировать отходы производства. Будем считать, что естественная убыль отходов в каждый момент времени составляет долю γ от их общего количества. Общество, в свою очередь, может выделять часть произведенного общественного продукта на борьбу с загрязнением. Предполагается, что эффективность (производительность) затрат на уменьшение загрязнения постоянна. При этом затрата одной единицы продукции уменьшает загрязнение на δ единиц (будем считать $\delta > 1$).

Задача управления состоит в определении долей α и β выпуска, предназначенных на потребление и борьбу с загрязнением соответственно. Имеем следующую систему уравнений:

$$c = \alpha F(K, L), \tag{2.19}$$

$$\dot{K} = (1 - \alpha - \beta) F(K, L) - \mu K, \tag{2.20}$$

$$\dot{P} = (\varepsilon - \delta\beta) F(K, L) - \gamma P; \tag{2.21}$$

здесь $0 \leqslant \alpha(t) \leqslant 1,\ 0 \leqslant \beta(t) \leqslant 1,\ \alpha(t) + \beta(t) \leqslant 1$.

Будем решать задачу нахождения оптимальных управлений $\alpha(t)$ и $\beta(t)$ с помощью принципа максимума Понтрягина. Обозначим через ψ_1 двойственную переменную, соответствующую уравнению (2.20) ($\psi_1(t)$ — объективно обусловленная оценка капитала $K(t)$ в момент t), через $\psi_2(t)$ — объективно обусловленную оценку загрязнения $P(t)$. Тогда гамильтониан $\mathscr{H}$ имеет вид

$$\mathscr{H} = u(c, P) e^{-rt} + \psi_1[(1 - \alpha - \beta) F(K, L) - \mu K] + \psi_2[(\varepsilon - \delta\beta) F(K, L) - \gamma P].$$

Двойственная система уравнений такова:

$$\dot{\psi}_1 = -\left\{\frac{\partial u}{\partial c}\alpha\frac{\partial F}{\partial K}e^{-rt} + \psi_1\left[(1 - \alpha - \beta)\frac{\partial F}{\partial K} - \mu\right] + \psi_2(\varepsilon - \delta\beta)\frac{\partial F}{\partial K}\right\},$$

$$\dot{\psi}_2 = -\frac{\partial u}{\partial P}e^{-rt} + \psi_2\gamma.$$

Перенормируем двойственные оценки:

$$q_1 = \psi_1 e^{rt}, \quad q_2 = \psi_2 e^{rt}.$$

Тогда двойственная система примет вид

$$\dot{q}_1 = -\frac{\partial u}{\partial c}\alpha\frac{\partial F}{\partial K} + q_1\left[r + \mu - (1 - \alpha - \beta)\frac{\partial F}{\partial K}\right] + q_2(\delta\beta - \varepsilon)\frac{\partial F}{\partial K}, \tag{2.22}$$

$$\dot{q}_2 = -\frac{\partial u}{\partial P} + q_2(\gamma + r), \tag{2.23}$$

а гамильтониан $\mathscr{H}$ будет равен

$$\mathscr{H} = e^{-rt}\{u(\alpha F(K, L), P) + q_1[(1 - \alpha - \beta) F(K, L) - \mu K] + q_2[(\varepsilon - \delta\beta) F(K, L) - \gamma P]\}.$$

В силу принципа максимума, если управление $\alpha(t)$ и $\beta(t)$ оптимально, то существуют непрерывные функции $q_1(t)$ и $q_2(t)$, удовлетворяющие (2.22), (2.23). При этом функции $\alpha(t)$ и $\beta(t)$ максимизируют значение формы $\mathscr{H}(q_1(t), q_2(t), K(t), P(t), \alpha, \beta)$ в момент времени t.

Выпишем функцию $\mathscr{H}$, перегруппировав ее члены:

$$\mathscr{H} = e^{-rt}\{u(\alpha F(K, L), P) - q_1\alpha F(K, L) - (q_1 + q_2\delta) F(K, L)\beta + q_1(F(K, L) - \mu K) + q_2(\varepsilon F(K, L) - \gamma P)\}.$$

Видно, что для получения максимума функции $\mathscr{H}$ по переменным α, β достаточно максимизировать выражение

$$\varphi(\alpha, \beta) = v(\alpha) + \theta\beta \tag{2.24}$$

в области $\alpha \geqslant 0,\ \beta \geqslant 0,\ \alpha + \beta \leqslant 1$, где

$$v(\alpha) = u(\alpha F(K, L), P) - q_1\alpha F(K, L),$$

$$\theta = -(q_1 + q_2\delta) F(K, L).$$

В случае $\theta > 0$ очевидно, что максимум функции $\varphi(\alpha, \beta)$ достигается при $\alpha + \beta = 1$.

Если $\theta < 0$, максимум $\varphi(\alpha, \beta)$ достигается при $\beta = 0$.

Когда $\theta = 0$, то β произвольно, а α либо равно 1, либо является решением уравнения $\partial v/\partial\alpha = 0$. (Максимум функции $v(\alpha)$ не может достигаться в точке $\alpha = 0$ в силу условий на функцию u: $\left.\frac{\partial u}{\partial c}\right|_{c=0} = \infty$.)

Полный анализ поведения оптимальных траекторий в данной модели довольно сложен, поскольку здесь имеются два управляющих параметра. Однако зная, какую важную роль играют при описании оптимальных траекторий точки равновесия (см. предыдущий параграф), попытаемся ответить на следующий вопрос: существуют ли в данной модели траектории сбалансированного роста (точки равновесия), удовлетворяющие необходимым условиям принципа максимума, и сколько их?

Мы покажем, что существуют ровно две точки равновесия. В одном из таких положений никаких средств на борьбу с загрязнением не тратится. Такое состояние авторы модели назвали «равновесием темного века». Оно характеризуется высоким уровнем производства (большим объемом основного капитала), высоким уровнем потребления и крайне высоким уровнем загрязнения, который регулируется лишь естественными процессами очистки.

Состояние равновесия, в котором производятся расходы как на потребление, так и на борьбу с загрязнением, называют «равновесием золотого века». Оно отличается более низкими уровнями капитала, потребления и загрязнения, чем равновесие темного века.

Состояние равновесия описывается условиями $\dot{K}=0$, $\dot{P}=0$, т. е.

$$(1-\alpha-\beta)F(K, L)=\mu K, \tag{2.25}$$

$$(\varepsilon-\delta\beta)F(K, L)=\gamma P. \tag{2.26}$$

Поскольку величина K в состоянии равновесия постоянна и, естественно, положительна, то из (2.25) вытекает, что сумма $\alpha+\beta$ соответствующих оптимальных управлений постоянна и строго меньше единицы. Следовательно, $\alpha<1$, $\beta<1$ и имеет место случай $\theta\leqslant 0$. Из (2.26) получаем, что в точке равновесия управление β (а следовательно, и α) постоянно. Поскольку $\alpha<1$, то, значит, $\partial v/\partial\alpha=0$, т. е.

$$\frac{\partial u\,(\alpha F\,(K,\ L),\ P)}{\partial c}=q_1, \tag{2.27}$$

откуда вытекает, что функция $q_1(t)$ также является константой. Формула (2.22) показывает, что в таком случае $q_2(t)$ также постоянно. Из формулы (2.23) получаем выражение q_2 через функцию u:

$$q_2=\frac{1}{\gamma+r}\,\frac{\partial u}{\partial P}.$$

Уравнение (2.22) принимает вид

$$q_1\left(r+\mu-\frac{\partial F}{\partial K}+\beta\,\frac{\partial F}{\partial K}\right)-q_2(\varepsilon-\delta\beta)\,\frac{\partial F}{\partial K}=0. \tag{2.28}$$

Рассмотрим теперь два случая равновесия. *Равновесие золотого века* — это равновесие при $\theta=0$. *Равновесие темного века* — когда $\theta<0$.

Случай 1 (золотой век). Поскольку $\theta=0$, то $q_2=-q_1/\delta$. Подставляя полученное выражение для q_2 в (2.28), получаем $q_1\left[r+\mu-\left(1-\frac{\varepsilon}{\delta}\right)\frac{\partial F}{\partial K}\right]=0$. Если $q_1=0$, то и $q_2=0$, что противоречит формулировке принципа максимума. Следовательно, равно нулю выражение в квадратных скобках, откуда получаем

$$\frac{\partial F}{\partial K}=\frac{r+\mu}{1-\varepsilon/\delta}. \tag{2.29}$$

Условие $\delta>1>\varepsilon$ обеспечивает неравенство $\dfrac{r+\mu}{1-\varepsilon/\delta}>0$. Вместе с обычными неоклассическими условиями

$$\lim_{K\to 0}\frac{\partial F}{\partial K}=\infty,\quad \lim_{K\to\infty}\frac{\partial F}{\partial K}=0,\quad \frac{\partial^2 F}{\partial K^2}<0$$

это обеспечивает существование единственного решения K^* уравнения (2.29).

Условие же $\dot{q}_2=0$ вместе с (2.23) дает

$$\frac{\partial u}{\partial P}=(r+\gamma)\,q_2=-\frac{r+\gamma}{\delta}\,\frac{\partial u}{\partial c},$$

или

$$-\frac{\partial u/\partial P}{\partial u/\partial c}=\frac{r+\gamma}{\delta}. \tag{2.30}$$

Поскольку $(r+\gamma)/\delta>0$, то наше предположение о характере замещения функции $u(c, P)$, сделанное в начале данного параграфа, позволяет утверждать, что кривая, определяемая уравнением (2.30) в плоскости (c, P), монотонно убывает от ∞ до 0 при возрастании c от 0 до ∞.

Исключая β из системы уравнений (2.25), (2.26), находим зависимость между K^* и P:

$$\alpha\delta F(K^*, L)=-\mu\delta K^*+\gamma P+(\delta-\varepsilon)F(K^*, L).$$

Вспоминая, что $c=\alpha F(K^*, L)$, получаем зависимость между c и P:

$$\delta c=-\mu\delta K^*+\gamma P+(\delta-\varepsilon)F(K^*, L). \tag{2.31}$$

Поскольку, как мы уже знаем, значение K^* определено и единственно, то уравнение (2.31) определяет в плоскости (c, P) прямую с углом наклона $\delta/\gamma>0$. Следовательно, система уравнений (2.30), (2.31) имеет единственное решение (c^*, P^*), где $c^*>0$, $P^*>0$.

Случай 2 (темный век). Здесь $\theta<0$ и потому $\beta=0$. Из (2.28) тогда получаем

$$\frac{\partial F}{\partial K}=\frac{r+\mu}{1+q_2/q_1},$$

или

$$\frac{\partial F}{\partial K}=(r+\mu)\left(1+\frac{1}{\gamma+r}\,\frac{\partial u/\partial P}{\partial u/\partial c}\right)^{-1}. \tag{2.32}$$

При $\beta=0$ имеем $P=\frac{\varepsilon}{\gamma}F(K, L)$. Уравнение (2.32) определяет P как неявную функцию от K, и можно вычислить производную dP/dK. Мы не будем приводить здесь эти несложные, но громоздкие преобразования. Скажем только, что из явного вида dP/dK и предположения о характере замещения функции u вытекает, что $dP/dK<0$. Вместе с тем для функции

$$P=\frac{\varepsilon}{\gamma}F(K,\ L) \tag{2.33}$$

справедливо $dP/dK>0$. Следовательно, существует единственное решение K^{**}, P^{**} системы уравнений (2.32), (2.33). В таком случае из уравнения (2.25) при $\beta=0$ определяется единственное значение c^{**}: $c^{**}=\alpha^{**}F(K^{**},L)=F(K^{**},L)-\mu K^{**}$. Итак, в случае 2 также существует единственное положение равновесия (c^{**},P^{**}), где $P^{**}=\frac{1}{\gamma}F(K^{**},L)$.

§ 3. Эндогенный научно-технический прогресс

В этом параграфе мы рассмотрим модель, посвященную одному из самых интересных, на наш взгляд, вопросов — изучению управляемого эндогенного научно-технического прогресса в агрегированных макроэкономических моделях, базирующихся на понятии производственной функции.

Мы уже встречались в гл. 1 с различными типами научно-технического прогресса, выражающегося в изменении со временем производственной функции. В моделях § 1, 2 можно было бы ввести учет научно-технического прогресса, скажем, нейтрального по Хиксу (например, рассматривая производственную функцию вида $e^{\lambda t}F(K,L)$), что не слишком усложнило бы исследование этих моделей. Однако технические изменения, учитываемые таким образом, являются экзогенными (внешними) по отношению к модели. Они обусловлены факторами, не входящими в модель. Вместе с тем, даже если речь идет не обо всем народном хозяйстве, а лишь об отдельной его отрасли, то ее технический прогресс во многом определяется внутри самой отрасли — политикой капиталовложений в отраслевые НИИ, КБ, опытное производство и т. д.

Изучение влияния научных исследований на темпы роста промышленного производства и связанные с этим вопросы эффективности капиталовложений «в науку», определения рациональной доли общественного продукта, выделяемой на научные исследования, — все это приобретает в последнее время все большее значение.

Приведем одну из моделей, описывающую влияние эндогенного научно-технического прогресса на общественное производство (см. [39]). Рассмотрим односекторную замкнутую модель экономики, динамика которой определяется тем, в какой пропорции делится национальный доход между капиталовложениями на расширение основных фондов и на улучшение производства («на науку»). Для простоты считаем, что потребление отсутствует. (Можно было бы, например, считать, что норма потребления фиксирована, т. е. составляет заданную долю национального дохода, но это то же самое, что рассматривать другую производственную функцию.) Будем считать, что вложения «в науку» действуют таким образом, что мы имеем дело с прогрессом, нейтральным по Хиксу, т. е. величина Y выпуска определяется формулой

$$Y=A(Q)F(K,L), \tag{2.34}$$

где K и L, как обычно, — объем основных фондов и трудовых ресурсов; $F(K,L)$ — производственная функция; Q — суммарный объем капиталовложений в научно-технический прогресс; $A(Q)$ — мультипликатор прогресса, показывающий эффективность затрачиваемых «на науку» средств.

Относительно производственной функции $F(K,L)$ будем по-прежнему считать, что она удовлетворяет неоклассическим условиям (см. § 2) и, кроме того, имеет вид $F(K,L)=g(K)h(L)$. Например, функция Кобба — Дугласа удовлетворяет этим условиям.

На мультипликатор прогресса $A(Q)$ также наложим условия, аналогичные неоклассическим:

$$A'(Q)>0,\ A''(Q)<0,\ A(0)=1,$$
$$\lim_{Q\to\infty}A(Q)=\infty,\quad \lim_{Q\to\infty}A'(Q)=0,\quad \lim_{Q\to 0}A'(Q)=\infty.$$

Пусть u, $0\leqslant u\leqslant 1$, — норма накопления, т. е. доля национального дохода Y, направляемая на увеличение основных фондов. Тогда

$$\dot K=uY=uA(Q)g(K)h(L). \tag{2.35}$$

Величина $1-u$ представляет собой долю национального дохода Y, направляемую «на науку», в научно-технический прогресс. Следовательно,

$$\dot Q=(1-u)Y=(1-u)A(Q)g(K)h(L). \tag{2.36}$$

Наконец, относительно трудовых ресурсов L будем считать, что их рост подчиняется закону

$$\dot L=p(L), \tag{2.37}$$

где $p(L)$ — такая функция, что величина L не может стать бесконечной за конечное время.

Пусть (K_0,Q_0,L_0) — произвольное начальное состояние. Задание управления $u(t)$, $t\geqslant 0$, $0\leqslant u(t)\leqslant 1$, позволяет из уравнений (2.35) — (2.37) определить траекторию $(K(t),Q(t),L(t))$.

Рассмотрим следующую экстремальную задачу: за кратчайшее время достигнуть заданного уровня $\bar K$ объема основных фондов.

Решение этой задачи проведем с помощью принципа максимума.

Предварительно отметим следующее. Описанная в математическом введении задача 1 на быстродействие обладает тем очевидным свойством, что если оптимальная траектория с началом x^0 проходит через точку x^1, то отрезок этой траектории с началом x^1 до многообразия M будет также оптимальным. Отсюда, в свою очередь, вытекает, что оптимальное управление в каждой точке фазовой плоскости зависит лишь от самой этой точки (ее координат). Поэтому часто оказывается естественным искать оптимальное управление в задаче 1 не в виде $u(t)$, а в виде $u(x)$, где x — точка фазовой плоскости. Оптимальное управление вида $u(x)$ называется *синтезирующим*.

Все сказанное относится и к рассматриваемой задаче. Будем искать оптимальное управление в виде $u(K, Q, L)$. На самом деле окажется, что оптимальное управление в данном случае зависит только от K, Q.

В дальнейшем будем любую траекторию, удовлетворяющую необходимым условиям оптимальности — принципу максимума, — называть *экстремалью*.

В нашей задаче терминальное многообразие задается уравнением

$$K - \bar{K} = 0. \tag{2.38}$$

Пусть ψ_1, ψ_2, ψ_3 — двойственные переменные, соответствующие уравнениям (2.35)—(2.37). Тогда функция Гамильтона для нашей задачи имеет вид

$$\mathscr{H} = \psi_1 uY + \psi_2(1-u)Y + \psi_3 p(L), \tag{2.39}$$

где $Y = A(Q)g(K)h(L)$; сопряженная система такова:

$$\dot{\psi}_1 = -\psi_1 u \frac{\partial Y}{\partial K} - \psi_2(1-u)\frac{\partial Y}{\partial K}, \tag{2.40}$$

$$\dot{\psi}_2 = -\psi_1 u \frac{\partial Y}{\partial Q} - \psi_2(1-u)\frac{\partial Y}{\partial Q}. \tag{2.41}$$

$$\dot{\psi}_3 = -\psi_1 u \frac{\partial Y}{\partial L} - \psi_2(1-u)\frac{\partial Y}{\partial L} - \psi_3 p'(L). \tag{2.42}$$

Условия трансверсальности для многообразия (2.39) имеют вид

$$\psi_1(T) = 1, \quad \psi_2(T) = \psi_3(T) = 0. \tag{2.43}$$

Если $u(t)$ — оптимальное управление, то оно в каждый момент времени максимизирует функцию $\mathscr{H}$.

Введем в рассмотрение функцию $\varphi(t) = \psi_1(t) - \psi_2(t)$. Из (2.39) следует, что

$$u = \begin{cases} 0 & \text{при } \varphi < 0, \\ 1 & \text{при } \varphi > 0. \end{cases}$$

В случае $\psi_1 = \psi_2$ значение управления u непосредственно из условия максимума функции $\mathscr{H}$ определить нельзя. Имеются две возможности: либо вдоль оптимальной траектории равенство $\psi_1 = \psi_2$ сохраняется тождественно в течение некоторого интервала времени, либо оно имеет место лишь в одной точке. Рассмотрим первую возможность: $\psi_1 \equiv \psi_2$. В таком случае $\dot{\psi}_1 = \dot{\psi}_2$, и из (2.40) и (2.41) получаем

$$\frac{\partial Y}{\partial K} = \frac{\partial Y}{\partial Q}. \tag{2.44}$$

Поскольку $Y = A(Q)g(K)h(L)$, то

$$\frac{\partial Y}{\partial K} - \frac{\partial Y}{\partial Q} = A'(Q)g'(K)h(L)\left(\frac{A(Q)}{A'(Q)} - \frac{g(K)}{g'(K)}\right)$$

и (2.44) эквивалентно уравнению

$$\frac{A(Q)}{A'(Q)} = \frac{g(K)}{g'(K)} \tag{2.45}$$

(так как $A'(Q)$, $g'(K)$, $h(L) > 0$).

Поверхность в фазовом пространстве, определяемую уравнением (2.45), назовем *особой поверхностью*. Поскольку переменная L не участвует в уравнении (2.45), то особая поверхность является цилиндрической и ее ось параллельна оси координат OL. Направляющую кривую для особой поверхности в плоскости (K, Q), уравнение которой совпадает с (2.45), назовем *особой кривой*.

Выясним вид этой кривой. Пусть $\tilde{A}(Q) = A(Q)/A'(Q)$, $\tilde{g}(K) = g(K)/g'(K)$. Из условий на $A(Q)$ и $g(K)$ легко следует, что $\tilde{A}(Q)$ и $\tilde{g}(K)$ — монотонные функции, строго возрастающие от 0 до ∞ при стремлении их аргументов к бесконечности. Обозначив через $\tilde{A}^{-1}$ функцию, обратную к $\tilde{A}$, из (2.45) получаем $Q = \tilde{A}^{-1}\tilde{g}(K)$. Поскольку функция, обратная к возрастающей, сама возрастает, видим, что Q есть возрастающая функция аргумента K, определенная при всех $K > 0$. Из (2.45) нетрудно также видеть, что $Q \to 0$ при $K \to 0$.

Таким образом, если оптимальная траектория $(Q(t), K(t), L(t))$ такова, что в течение некоторого времени $\psi_1(t) \equiv \psi_2(t)$, то эта траектория проходит по особой поверхности (2.45).

Отметим, что, как можно видеть из (2.45), выше особой кривой $\frac{A(Q)}{A'(Q)} - \frac{g(K)}{g'(K)} > 0$, ниже ее $\frac{A(Q)}{A'(Q)} - \frac{g(K)}{g'(K)} < 0$. Это же верно для выражения $\frac{\partial Y}{\partial K} - \frac{\partial Y}{\partial Q}$.

Займемся подробнее изучением поведения функций $\psi_1(t)$ и $\psi_2(t)$. Учитывая, что все функции $u(t)$, $1-u(t)$, $\partial Y/\partial K$, $\partial Y/\partial Q$ неотрицательны, $\psi_1(T) = 1$, $\psi_2(T) = 0$, из системы уравнений (2.40), (2.41) можно получить, что $\psi_1(t) > 0$, $\psi_2(t) > 0$ при $t < T$. Из (2.40), (2.41) имеем

$$\dot{\varphi} = \dot{\psi}_1 - \dot{\psi}_2 = (\psi_1 u + \psi_2(1-u))\left(\frac{\partial Y}{\partial Q} - \frac{\partial Y}{\partial K}\right).$$

Отсюда видно, что знак $\dot{\varphi}$ определяется знаком $\frac{\partial Y}{\partial Q} - \frac{\partial Y}{\partial K}$.

Следовательно, выше особой кривой $\dot{\varphi} < 0$, ниже особой кривой $\dot{\varphi} > 0$.

В дальнейшем мы покажем, что если экстремальная траектория проходит через точку фазовой плоскости, лежащую выше особой кривой, то она после этого обязательно пересечет особую кривую. Отсюда вместе с условием $\dot{\varphi} < 0$ будет вытекать, что всюду выше особой кривой на экстремальных траекториях $u \equiv 1$.

Для выяснения поведения экстремальных траекторий ниже особой кривой попытаемся найти множество точек, в которых функция φ обращается в нуль.

На терминальном многообразии, как мы видели, $\varphi = 1$. В силу непрерывности функций $\psi_1(t)$ и $\psi_2(t)$, в некоторой окрестности терминального многообразия имеет место неравенство $\varphi > 0$, т. е. оптимальное управление $u(t)$ равно 1.

Изучим поведение экстремальных траекторий при $u \equiv 1$. Прямая и сопряженная системы при $u \equiv 1$ имеют вид

$$\dot{K} = Y = A(Q)g(K)h(L), \quad \dot{Q} = 0, \quad \dot{L} = p(L), \qquad (2.46)$$

$$\dot{\psi}_1 = -\psi_1 \frac{\partial Y}{\partial K} = -\psi_1 A(Q)\,g'(K)\,h(L), \qquad (2.47)$$

$$\dot{\psi}_2 = -\psi_1 \frac{\partial Y}{\partial Q} = -\psi_1 A'(Q)\,g(K)\,h(L), \qquad (2.48)$$

$$\dot{\psi}_3 = -\psi_1 \frac{\partial Y}{\partial L} - \psi_3 p'(L).$$

Вычислим функцию $\varphi = \psi_1 - \psi_2$ в этом случае. Из уравнений (2.46) и (2.47) получаем

$$\frac{dK}{g(K)} = A(Q)\,h(L)\,dt = -\frac{d\psi_1}{g'(K)\,\psi_1},$$

или $dg(K)/g(K) = -d\psi_1/\psi_1$. Интегрируя по времени от t до T и учитывая, что $K(T) = \bar{K}$, $\psi_1(T) = 1$, получаем $\psi_1(t) = g(\bar{K})/g(K)$. Подставляя это выражение в (2.48) и вновь интегрируя с учетом того, что $\psi_2(T) = 0$, $Q(t) \equiv Q(T)$, имеем

$$\psi_2(t) = g(\bar{K})\,A'(Q(T)) \int_t^T h(L(\tau))\,d\tau.$$

Отсюда

$$\varphi = g(\bar{K}) \left[\frac{1}{g(K)} - A'(Q(T)) \int_t^T h(L(\tau))\,d\tau \right].$$

Как и должно быть, при $t = T$ получаем $\varphi = 1$. При уменьшении t оптимальное управление остается равным 1 до тех пор, пока $\varphi > 0$, т. е. до момента $\hat{t}$, являющегося наибольшим корнем уравнения

$$\int_0^T h(L(\tau))\,d\tau = \frac{1}{g(K)\,A'(Q(T))}. \qquad (2.49)$$

Найдем уравнение множества точек в фазовом пространстве, соответствующих точкам экстремальных траекторий в момент $\hat{t}$.

Для этого выразим интеграл от $h(L)$, проинтегрировав уравнение (2.46):

$$\frac{1}{A(Q)} \int_K^{\bar{K}} \frac{dK}{g(K)} = \int_t^T h(L(\tau))\,d\tau.$$

Здесь мы вновь учли, что при $u \equiv 1$ величина Q постоянна. Отсюда и из (2.49) получаем

$$g(K) \int_K^{\bar{K}} \frac{dK}{g(K)} = \frac{A(Q)}{A'(Q)}. \qquad (2.50)$$

Перепишем (2.50) в виде

$$\tilde{A}(Q) = g(K) \int_K^{\bar{K}} \frac{dK}{g(K)}$$

или

$$Q = \tilde{A}^{-1}\left(g(K) \int_K^{\bar{K}} \frac{dK}{g(K)} \right). \qquad (2.51)$$

Уравнение (2.51) задает в плоскости (K, Q) некоторую кривую, вид которой нам необходимо выяснить. Из (2.50) непосредственно видно, что если $K = 0$ или $K = \bar{K}$, то $Q = 0$.

Особая кривая (2.45) и кривая (2.51) имеют единственную общую точку пересечения (если не считать точку $K = 0$), определяемую из уравнения

$$\int_K^{\bar{K}} \frac{dK}{g(K)} = \frac{1}{g'(K)}. \qquad (2.52)$$

Поскольку $g(K)$ растет с ростом K, а $g'(K)$ убывает, то левая часть уравнения (2.52) является функцией, убывающей от ∞ до 0, правая — возрастающая от 0 до ∞. Следовательно, уравнение (2.52) имеет единственный корень, который мы обозначим K_p. Соответствующее значение Q, определяемое из (2.51), обозначим Q_p. Вычислим производную функции (2.51):

$$Q'_K = \left(\frac{A(Q)}{A'(Q)} \frac{g'(K)}{g(K)} - 1 \right) \Big/ \left(\frac{A(Q)}{A'(Q)} \right)'.$$

Данное выражение несложно получается, если дифференцировать неявно заданную функцию Q (2.50).

Видно, что $Q'_K = 0$, как раз когда $A(Q)/A'(Q) = g(K)/g'(K)$, т. е. в точке K_p. При $K < K_p$ имеем $Q'_K > 0$, при $K > K_p$ имеем $Q'_K < 0$. Таким образом, K_p — вершина кривой (2.51). Если

(P_0, Q_0) лежит выше особой кривой, то возможны два случая: $Q_0 > Q_p$ и $Q_0 \leqslant Q_p$. В первом случае уравнение (2.50) решений не имеет и поэтому для экстремальной траектории $u \equiv 1$. Во втором случае экстремаль обязательно пересекает особую кривую, что означает, как отмечено выше, что всюду над особой кривой $u \equiv 1$. Правую ветвь кривой (2.51), т. е. при $K \geqslant K_p$, назовем *кривой переключения*. На рис. 2.11 сплошной линией изображена особая кривая, штриховой — кривая переключения. Непрерывную кривую, составленную из особой кривой при $0 \leqslant K \leqslant K_p$ и кривой переключения, назовем *обобщенной кривой переключения* и обозначим Λ.

Как мы видели, выше кривой переключения на экстремальных траекториях $\varphi > 0$ (и, значит, $u \equiv 1$). На самой кривой переключения $\varphi = 0$. Вспоминая, что ниже особой кривой $\varphi > 0$, получаем, что в области под кривой Λ $\varphi < 0$, т. е. на экстремальных траекториях $u \equiv 0$.

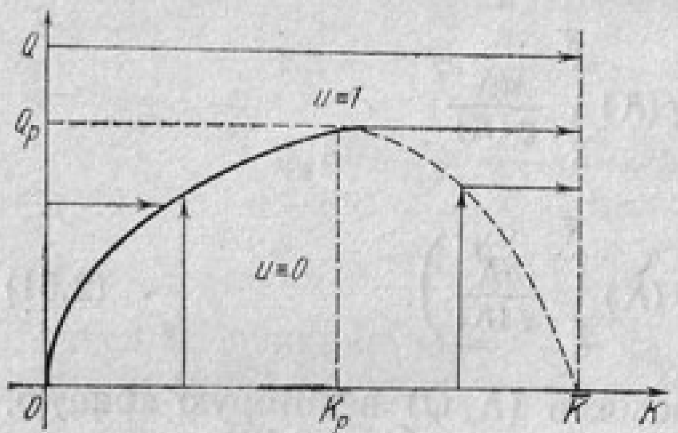

Рис. 2.11.

Таким образом, мы почти полностью выяснили поведение оптимальных управлений во всех точках фазовой плоскости (K, Q): выше кривой Λ $u \equiv 1$, ниже кривой Λ $u \equiv 0$.

Значение управлений в точках кривой переключения не играет роли — можно считать, например, $u = 1$. Осталось найти значение оптимального управления на особой кривой. Это нетрудно. Из уравнения (2.45) можно найти dQ/dK, т. е. выяснить, каким образом должны быть связаны скорости изменения величин Q и K, чтобы точка (K, Q) двигалась вдоль особой кривой:

$$\frac{dQ}{dK} = \frac{\tilde{g}'(K)}{\tilde{A}'(Q)}.$$

С другой стороны, из (2.35), (2.36) получаем $dQ/dK = (1-u)/u$. Отсюда $u^* = \tilde{A}'(Q)/(\tilde{A}'(Q) + \tilde{g}'(K))$. На рис. 2.11 стрелками показаны типичные экстремальные траектории нашей задачи. Окончательно имеем

$$u(K, Q, L) = \begin{cases} 1, & \text{если } (K, Q) \text{ выше } \Lambda, \\ 0, & \text{если } (K, Q) \text{ ниже } \Lambda, \\ u^*, & \text{если } (K, Q) \in \Lambda. \end{cases} \qquad (2.53)$$

Видим, что из каждой точки (K_0, Q_0, L_0) фазового пространства идет единственная траектория, удовлетворяющая принципу мак-

симума Понтрягина. Используя теорему Филиппова, можно показать, что в нашей задаче существует оптимальная траектория для любого начального состояния (K_0, Q_0, L_0) и любого $\bar{K} > K_0$. Отсюда вытекает, что единственная экстремальная траектория является оптимальной и формула (2.53) описывает оптимальный синтез. Тем самым поставленная нами задача решена.

Обсудим это решение с экономической точки зрения. Из полученного описания оптимальных траекторий видно, что типичным является следующий случай: если K_0 много меньше $\bar{K}$, то вначале траектория выходит на особую поверхность и движется по ней до пересечения с поверхностью переключения, после чего сходит с особой поверхности.

Из (2.52) можно показать, что величина K_p неограниченно растет с ростом $\bar{K}$. В самом деле, используя условие $\lim\limits_{k\to\infty} g'(K) = 0$, получаем, что при больших K $g(K) < K$, т. е. $1/g(K) > 1/K$. Отсюда уж непосредственно видно, что интеграл $\int\limits_{\bar{K}}^{\bar{K}} \frac{dg}{g(K)}$ неограниченно растет с ростом $\bar{K}$ при фиксированном $\bar{K}$. Поэтому, если предположить, что $K_p(\bar{K})$ есть величина ограниченная, скажем, числом $\bar{K}_p$, то получим противоречивое неравенство

$$\int\limits_{\tilde{K}_p}^{\infty} \frac{dK}{g(K)} \leqslant \frac{1}{g'(\tilde{K}_p)}.$$

Неограниченный рост величины $K_p(\bar{K})$ означает, в частности, что при увеличении величины $\bar{K}$ отрезок времени, в течение которого оптимальная траектория находится на особой поверхности, неограниченно увеличивается. По аналогии с тем же феноменом при исследовании оптимальных траекторий в линейных моделях назовем особую поверхность *магистралью*.

Поскольку уравнение магистрали имеет вид $\partial Y/\partial K = \partial Y/\partial Q$, то ее экономическая характеристика такова: на этой поверхности, и только на ней, предельная фондоотдача (норма эффективности накопления) и норма эффективности капиталовложений в научно-технический прогресс равны. Следовательно, оптимальное управление подчиняется такому правилу: если цель достаточно удалена, то все капиталовложения следует направлять в ту область, где норма эффективности больше. На магистрали капиталовложения должны распределяться в такой пропорции, чтобы равенство норм эффективности сохранялось тождественно.

Если же руководствоваться двойственными переменными, интерпретируя их как цены оптимального плана (с точностью до положительного множителя), то здесь алгоритм оптимальности еще проще: независимо от соотношения K_0 и $\bar{K}$ определяется оптимальное управление u по соотношению ψ_1 и ψ_2.

Интересен тот факт, что уравнение особой поверхности и оптимальное управление на ней сохраняются для обширного класса целевых функций. В качестве иллюстрации к этому утверждению рассмотрим еще одну постановку нашей задачи: за кратчайшее время достигнуть заданную величину $\bar{Y}$ выпуска. В этом случае гамильтонова система уравнений остается прежней: (2.35)—(2.37), (2.40)—(2.42), меняется лишь уравнение терминального многообразия, которое имеет вид $Y(Q, K, L) = \bar{Y}$.

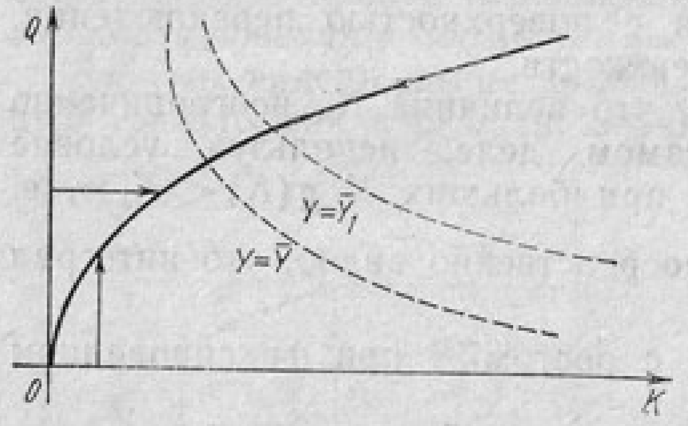

Рис. 2.12.

В соответствии с этим уравнение особой поверхности имеет такой же вид. Условие трансверсальности в данном случае дает $\psi_1(T) = \partial Y/\partial K$, $\psi_2(T) = \partial Y/\partial Q$, $\psi_3(T) = \partial Y/\partial L$. Поскольку все эти величины положительны, то остается справедливым вывод о том, что выше особой поверхности функция φ убывает, ниже ее — растет.

Если экстремальная траектория выходит на терминальное многообразие таким образом, что проекция точки $(K(T), Q(T), L(T))$ на плоскость (K, Q) лежит выше магистральной (особой) кривой, то, как мы видели, в этой точке $\partial Y/\partial K > \partial Y/\partial Q$, т. е. $\varphi = \psi_1 - \psi_2 > 0$. Следовательно, в любой точке выше магистрали $\varphi > 0$ и оптимальное управление $u \equiv 1$.

Аналогично показывается, что ниже магистральной кривой $u \equiv 0$. Следовательно, экстремальные траектории в данной задаче устроены проще, чем в предыдущей. Поскольку из каждой начальной точки (K_0, Q_0, L_0) идет только одна экстремальная кривая, то она является оптимальной. Проекция оптимального синтеза изображена на рис. 2.12.

Следует обратить внимание на то, что оптимальное управление не зависит от того, чему равен желаемый уровень выпуска $\bar{Y}$: для любого начального состояния (K_0, Q_0, L_0) выгодно выйти на магистраль и идти вдоль нее, даже если в процессе развития решено изменить цель и достичь за кратчайшее время другого уровня выпуска $\bar{Y}_1$. Этот вывод является подтверждением целесообразности практики скользящего планирования.

ЗАДАЧИ

1. Показать, что если хотя бы один из коэффициентов α, β эластичности по ресурсам линейно-однородной двухфакторной производственной функции $F(K, L)$ не зависит от K, L, то $F(K, L)$ является функцией Кобба — Дугласа.

2. Показать, что если предельная норма замещения S для линейно-однородной двухфакторной производственной функции $F(K, L)$ не зависит от K, L, то $F(K, L) = AK + BL$, где A, B — константы.

3. Показать, что функция

$$F = A[\delta K^{-\rho} + (1-\delta)L^{-\rho}]^{-1/\rho},$$

где $A > 0$, $0 < \delta < 1$, при $\rho < -1$ не является вогнутой.

4. Для однородной функции $F(K, L)$ выяснить связь решений задач

$$\max F(K, L) \qquad \min(p_1K + p_2L)$$

$$p_1K + p_2L = c; \qquad F(K, L) = Y.$$

5. Доказать, что в случае технического прогресса, нейтрального по Солоу, производственная функция имеет вид (1.33).

6. Доказать, что однородная степени α функция одной переменной имеет вид Ax^{α}.

7. Доказать теорему Эйлера: если $f(x)$ — дифференцируемая однородная степени α функция, то

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} x_i = \alpha f(x).$$

Указание. Вычислить производные функций $\varphi(t) = f(tx)$ и $\psi(t) = t^{\alpha} f(x)$ в точке $t = 1$.

8. Доказать, что если $f(x)$ — дифференцируемая однородная степени α функция, то функции $\partial f/\partial x_i$, $i = 1, 2, \ldots, n$, однородны степени $\alpha - 1$.

9. Пусть $F(K, L)$ — линейно-однородная производственная функция. Показать, что $\partial^2 F/\partial K\,\partial L > 0$.

10. Показать, что для производственной функции Леонтьева $F(K, L) = \min\{aK, bL\}$ предельная норма замещения S равна ∞, т. е. факторы незаменяемы. При этом, естественно, эластичность замещения σ равна 0.

11. Доказать

$$\lim_{\rho\to\infty} [\delta K^{-\rho} + (1-\delta) L^{-\rho}]^{-1/\rho} = \min\{K, L\}.$$

Напомним, что $\rho \to \infty$ при $\sigma \to 0$ $(\rho = (1-\sigma)/\sigma)$.

12. Пусть $x \in \mathbf{R}^n_+$, $\lambda > 0$. Обозначим через $\lambda(i)$ отображение $\lambda(i)$: $\mathbf{R}^n_+ \to \mathbf{R}^n_+$, определяемое следующим образом. Если $x = (x_1, x_2, \ldots, x_n)$, то

$$\lambda(i)x = (\lambda x_1, \ldots, \lambda x_{i-1}, x_i, \lambda x_{i+1}, \ldots, \lambda x_n).$$

Доказать, что если функция $f(x)$ такова, что $f(\lambda(i)x) = \lambda^{\alpha_i} f(x)$ для всех $i = 1, 2, \ldots, n$, то

$$f(x) = Ax_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}.$$

13. Вычислить сопряженную функцию $f^{\times}$ (см. гл. 1, § 4) для каждой из функций $f = AK^{\alpha}L^{\beta}$, где $\alpha + \beta = 1$, $\alpha, \beta \geqslant 0$,

$$f = aK + bL, \quad f = \min\{aK, bL\}.$$