Лекция №3

ТЕМА. ОСНОВЫ СИМПЛЕКСНОГО МЕТОДА

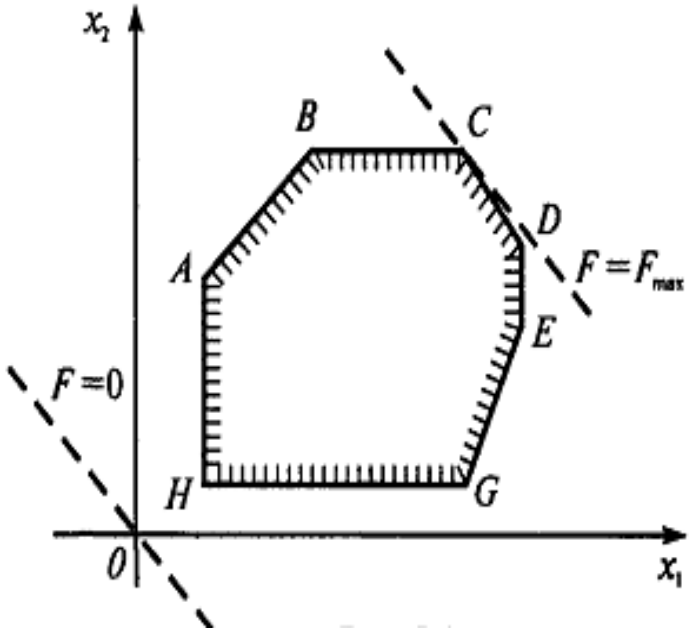
1. Геометрическая интерпретация симплексного метода

В гл. 3 рассмотрены основные теоремы линейного программирования, из которых следует, что если задача линейного программирования имеет оптимальное решение, то оно соответствует хотя бы одной угловой точке многогранника решений и совпадает, по крайней мере, с одним из допустимых базисных решений системы ограничений (см. теоремы 3.3, 3.4). Там же был указан путь решения любой задачи линейного программирования: перебрать конечное число допустимых базисных решений системы ограничений и выбрать среди них то, на котором функция цели принимает оптимальное решение. Геометрически это соответствует перебору всех угловых точек многогранника решений. Такой перебор в конце концов приведет к оптимальному решению (если оно существует), однако его практическое осуществление связано с огромными трудностями, так как для реальных задач число допустимых базисных решений хотя и конечно, но может быть чрезвычайно велико.

Число перебираемых допустимых базисных решений можно сократить, если производить перебор не беспорядочно, а с учетом изменений линейной функции, т.е. добиваясь того, чтобы каждое следующее решение было “лучше” (или, по крайней мере, “не хуже”), чем предыдущее, по значениям линейной функции (увеличение ее при отыскании максимума $F \to \max$, уменьшение — при отыскании минимума $F \to \min$).

Такой перебор позволяет сократить число шагов при отыскании оптимума. Поясним это на графическом примере.

Пусть область допустимых решений изображается многоугольником *ABCDEGH* (рис. 5.1). Предположим, что его угловая точка *A* соответствует исходному допустимому базисному решению. При беспорядочном переборе пришлось бы испытать семь допустимых базисных решений, соответствующих семи угловым точкам многоугольника. Однако из чертежа видно, что после вершины *A* выгодно перейти к соседней вершине *B*, а затем — к оптимальной точке *C*.



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

Идея последовательного улучшения решения легла в основу универсального метода решения задач линейного программирования — ***симплексного***[1] ***метода.***

Геометрический смысл симплексного метода состоит в последовательном переходе от одной вершины многогранника ограничений (называемой первоначальной) к соседней, в которой линейная функция принимает лучшее (по крайней мере, не худшее) значение

(по отношению к цели задачи) до тех пор, пока не будет найдено оптимальное решение — вершина, где достигается оптимальное значение функции цели (если задача имеет конечный оптимум).

Впервые симплексный метод был предложен американским ученым Дж. Данцигом в 1949 г., однако еще в 1939 г. идеи метода были разработаны российским ученым Л В.Канторовичем.

Симплексный метод, позволяющий решить любую задачу линейного программирования, универсален. В настоящее время он используется для компьютерных расчетов, однако несложные примеры с применением симплексного метода можно решать и вручную.

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

- *способ определения какого-либо первоначального допустимого базисного решения задачи;*
- *правило перехода к лучшему (точнее, не худшему) решению;*
- *критерий проверки оптимальности найденного решения.*

Для использования симплексного метода задача линейного программирования должна быть приведена к каноническому виду, т.е. система ограничений должна быть представлена в виде уравнений. Алгоритм конкретной вычислительной реализации этих элементов рассмотрим на примерах.

2. Отыскание максимума линейной функции

В качестве первого примера рассмотрим задачу об использовании ресурсов, сформулированную в разд. 1.2 и уже решенную геометрически в задаче **4.1.**

5.1. Решить симплексным методом задачу:

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

при ограничениях:

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

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

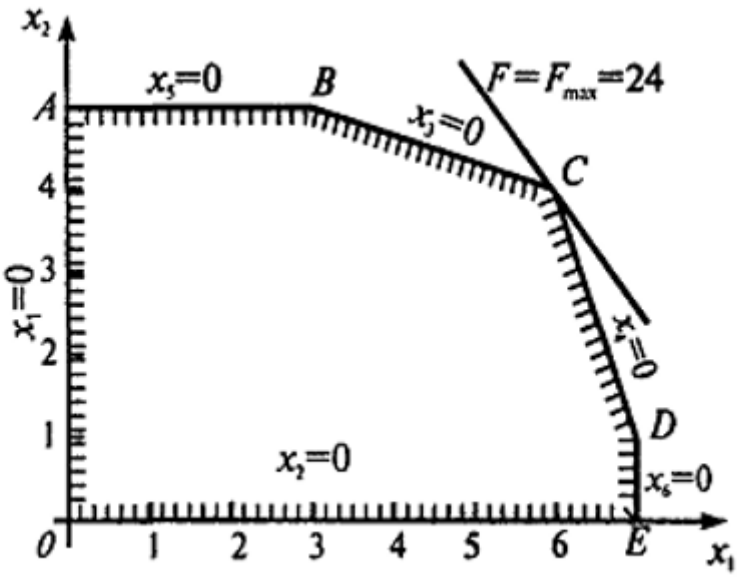


Рис. 5.2

Р е ш е н и е. С помощью дополнительных неотрицательных переменных перейдем к системе уравнений. В данном случае все дополнительные переменные вводятся со знаком "плюс", так как все неравенства имеют вид "$\le$" [*см.* систему ограничений (1.26)].

Получим систему ограничений в виде:

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

Для нахождения первоначального базисного решения разобьем переменные на две группы — основные и неосновные. Так как определитель, составленный из коэффициентов при дополнительных переменных x_3, x_4, x_5, x_6, отличен от нуля, то (*см.* разд. 2.1) эти переменные можно взять в качестве основных на первом шаге решения задачи. При выборе основных переменных на первом шаге не обязательно составлять определитель из их коэффициентов и проверять, равен ли он нулю. Достаточно воспользоваться следующим правилом:

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

Дополнительные переменные удовлетворяют этому правилу. Если выбранные по этому правилу переменные имеют те же знаки, что и соответствующие им свободные члены в правых частях уравнений, то полученное таким образом базисное решение будет допустимым. В данном случае так и получилось.

I ш а г . Основные переменные: x_3, x_4, x_5, x_6.
Неосновные переменные: x_1, x_2.

Выразим основные переменные через неосновные:

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

Положив неосновные переменные равными нулю, т.е. $x_1 = 0$, $x_2 = 0$, получим базисное решение $X_1 = (0; 0; 18; 16; 5; 21)$, которое является допустимым и соответствует вершине $O(0;0)$ многогранника $OABCDE$ на рис. 5.2. Поскольку это решение допустимое, нельзя отбросить возможность того, что оно оптимально. Выразим линейную функцию через неосновные переменные: $F = 2x_1 + 3x_2$. При решении X_1 значение функции равно $F(X_1)$. Легко понять, что функцию F можно увеличить за счет увеличения любой из неосновных переменных, входящих в выражение для F с положительным коэффициентом. Это можно осуществить, перейдя к такому новому допустимому базисному решению, в котором эта переменная будет неосновной, т.е. принимать не нулевое, а положительное значение (если новое решение будет вырождено, то функция цели сохранит свое значение). При таком переходе одна из основных переменных перейдет в неосновные, а геометрически произойдет переход к соседней вершине многоугольника, где значение линейной функции "лучше" (по крайней мере "не хуже"). В данном примере для увеличения F можно переводить в основные либо x_1, либо x_2, так как обе эти переменные входят в выражение для F со знаком "плюс". Для определенности в такой ситуации будем выбирать переменную, имеющую больший коэффициент, т.е. в данном случае x_2 (такое правило выбора не всегда дает наименее трудоемкое решение, иногда имеет смысл провести предварительные специальные оценки).

Система (5.3) накладывает ограничения на рост переменной x_2. Поскольку необходимо сохранять допустимость решений, т.е. все переменные должны оставаться неотрицательными, то должны выполняться следующие неравенства (при этом $x_1 = 0$ как неосновная переменная):

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

Каждое уравнение системы (5.3), кроме последнего, определяет оценочное отношение — границу роста переменной x_2, сохраняющую неотрицательность соответствующей переменной. Эта граница определяется абсолютной величиной отношения свободного члена к коэффициенту при x_2 при условии, что эти числа имеют разные знаки. Последнее уравнение системы не ограничивает рост переменной x_2, так как данная переменная в него не входит (или формально входит с нулевым коэффициентом). В этом случае условимся обозначать границу символом ∞. Такой же символ будем использовать, когда свободный член и коэффициент при переменной в уравнении имеют одинаковые знаки, так как и в этом случае нет ограничений на рост переменной.

Не накладывает ограничений на рост переменной, переводимой в основные, и такое уравнение, где свободный член отсутствует (т.е. равен 0), а переводимая переменная имеет положительный коэффициент. И в этом случае граница обозначается символом ∞. Обратите внимание, что при нулевом свободном члене и отрицательном коэффициенте при переводимой переменной уравнение ограничивает рост этой переменной нулем! (любое положительное ее значение вносит отрицательную компоненту в следующее базисное решение). Все возможные случаи, которые возникают при оценке переменной, переводимой в основные, перечислены в конце разд. 5.3.

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

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

II шаг. Основные переменные: x_2, x_3, x_4, x_6.
Неосновные переменные: x_1, x_5.

Выразим новые основные переменные через новые неосновные, начиная с разрешающего уравнения (его используем при записи выражения для x_2):

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

или

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

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

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

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

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

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

III шаг. Основные переменные: x_1, x_2, x_4, x_6.
Неосновные переменные: x_3, x_5.

Как и на II шаге, выражаем новые основные переменные через новые неосновные, начиная с разрешающего уравнения (его используем при записи выражения для x_1). После преобразований получаем:

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

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

IV ш а г . Основные переменные: x_1, x_2, x_5, x_6.
Неосновные переменные: x_3, x_4.

После преобразований получим:

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

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

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

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

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

3. Отыскание минимума линейной функции

При определении минимума линейной функции Z возможны два пути:

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

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

$Y_1 = (0;\ 0;\ 3;\ \frac{2}{3};\ 0;\ 0)$ — первое базисное решение, оно допустимое. Выражаем линейную функцию через неосновные переменные: $Z = 18y_1 + 16y_2 + 5y_3 + 21y_4 = 18y_1 + 16y_2 + 5(3 - 3y_1 - y_2 + y_6) + 21(2/3 - (1/3)y_1 - (2/3)y_2 + (1/3)y_5) = 29 - 4y_1 - 3y_2 + 7y_5 + 5y_6$.

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

II шаг. Основные переменные: y_1, y_4.
Неосновные переменные: y_2, y_3, y_5, y_6.

Получим после преобразований:

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

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

III шаг. Основные переменные: y_1, y_2.
Неосновные переменные: y_3, y_4, y_5, y_6.

Получим после преобразований:

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

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

Замечание. На каждом шаге симплексного метода какая-либо неосновная переменная переводится в основные, при этом каждое уравнение системы ограничений определяет конечное или бесконечное наибольшее возможное значение этой переменной — оценочное отношение. В задачах **5.1** и **5.2** встречались различные случаи оценки роста неосновной переменной, которые зависели от знаков и значений свободного члена и коэффициента при переводимой переменной. Сформулируем все возможные случаи. Обозначим: x_i — переводимая неосновная переменная, b_j — свободный член, a_{ij} — коэффициент при x_i. В общем виде уравнение $x_j = b_j + \ldots + a_{ij}x_i + \ldots$ определяет наибольшее возможное значение x_i по следующим правилам:

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

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

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

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

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