1. Лекция 2. Линейное программирование

1.1.ПОСТАНОВКА ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

1.1.1. Формы задачи линейного программирования.

В общем виде задача линейного программирования* (в дальнейшем ЗЛП) может быть сформулирована как задача нахождения наибольшего значения линейной функции

$$f(x)=\sum_{j=1}^{n} c_j x_j = cx \qquad (1.1)$$

на некотором множестве $D \subset R^n$, где $x \in D$ удовлетворяют системе ограничений

$$\begin{aligned}
&a_{1,1}\ x_1 + a_{1,2}\ x_2 + \ldots + a_{1,n}\ x_n \le b_1, \\
&a_{2,1}\ x_1 + a_{2,2}\ x_2 + \ldots + a_{2,n}\ x_n \le b_2, \\
&\qquad\qquad \ldots \\
&a_{m,1}\ x_1 + a_{m,2}\ x_2 + \ldots + a_{m,n}\ x_n \le b_m, \qquad (1.2) \\
&a_{m+1,1}x_1 + a_{m+1,2}x_2 + \ldots + a_{m+1,n}x_n = b_{m+1}, \\
&\qquad\qquad \ldots \\
&a_{m+l,1}x_1 + a_{m+l,2}x_2 + \ldots + a_{m+l,n}x_n = b_{m+l},
\end{aligned}$$

и, возможно, ограничениям

$$x_1 \ge 0,\ x_2 \ge 0, \ldots,\ x_j \ge 0, \ldots,\ x_n \ge 0. \qquad (1.3)$$

* Напомним, что частные примеры, сводящиеся к задаче линейного программирования, были описаны во введении.

Не умаляя общности, можно считать, что в системе (1.2) первые m ограничений являются неравенствами, а последующие — l-уравнениями. Очевидно, этого всегда можно добиться за счет простого переупорядочения ограничений. Относительно направления знака неравенства будем предполагать, что левая часть меньше или равна правой. Добиться этого можно, умножив на (-1) обе части тех неравенств, которые имеют противоположный знак. Ограничения (1.3), вообще говоря, могут быть рассмотрены как частный случай ограничений в форме неравенств, но в силу особой структуры их обычно выделяют отдельно и называют *условиями неотрицательности* (или *тривиальными ограничениями).*

Дополнительно следует заметить, что выбор типа искомого экстремума (максимума или минимума) также носит относительный характер. Так, задача поиска максимума функции

$$f(x)=\sum_{j=1}^{n} c_j x_j \qquad (1.4)$$

эквивалентна задаче поиска минимума функции

$$-f(x)=\sum_{j=1}^{n} (-c_j) x_j\ . \qquad (1.5)$$

Часто условия задачи (1.1)-(1.3), содержащей ограничения только типа неравенств, бывает удобно записывать в сокращенной матричной форме

$$f(x) = cx \rightarrow \max,\ A_x \leq b,\ x \geq 0, \qquad (1.6)$$

где c и x — векторы из пространства R^n, b — вектор из пространства R^m, а A — матрица размерности m х n.

Задачу линейного программирования, записанную в форме (1.1)-(1.3), называют **общей задачей линейного программирования** (ОЗЛП).

Если все ограничения в задаче линейного программирования являются уравнениями и на все переменные x_j наложены условия неотрицательности, то она называется задачей линейного программирования в канонической форме, или **канонической задачей линейного программирования** (КЗЛП). В матричной форме КЗЛП можно записать в следующем виде:

$$f(x)=cx \rightarrow \max,\ D=\{x \in R^n \mid Ax=b,\ x \geq 0\}\ . \qquad (1.7)$$

Поскольку любая оптимизационная задача однозначно определяется целевой функцией f и областью D, на которой отыскивается оптимум (максимум), будем обозначать эту задачу парой (D, f).

Условимся относительно терминологии, которая используется в дальнейшем и является общепринятой в теории линейного программирования.

Планом ЗЛП называется всякий вектор x из пространства R^n.

Допустимым планом называется такой план ЗЛП, который удовлетворяет ограничениям (1.2)-(1.3), т. е. содержится в области D. Сама область D называется при этом *областью допустимых планов*. **Оптимальным планом** x^* называется такой допустимый план, при котором целевая функция достигает оптимального (в нашем случае — максимального) значения, т. е. план, удовлетворяющий условию

$$\max f(x) = f(x^*).$$

Величина $f^* = f(x^*)$ называется *оптимальным значением целевой функции*.

Решением задачи называется пара (x^*, f^*), состоящая из оптимального плана и оптимального значения целевой функции, а *процесс решения* заключается в отыскании множества всех решений ЗЛП.

1.1.2. Переход к канонической форме. Подавляющее большинство известных методов решения задач линейного программирования предназначены для канонических задач. Поэтому начальный этап решения всякой общей задачи линейного программирования обычно связан с приведением ее к некоторой эквивалентной канонической задаче.

Общая идея перехода от ОЗЛП к КЗЛП достаточно проста:

- *ограничения в виде неравенств преобразуются в уравнения за счет добавления фиктивных неотрицательных переменных* x_i, $(i \in 1{:}m)$, *которые одновременно входят в целевую функцию с коэффициентом 0, т. е. не оказывают влияния на ее значение;*
- *переменные, на которые не наложено условие неотрицательности, представляются в виде разности двух новых неотрицательных переменных:*

$$x_j = \bar{x}_j - \bar{\bar{x}}_j, \ (\bar{x}_j \geq 0, \ \bar{\bar{x}}_j \geq 0).$$

Проиллюстрируем применение описанных выше рекомендаций на примере. Пусть задана задача линейного программирования (D, f) в общей форме с целевой функцией

$$f(x) = 5x_1 + 3x_2 + x_3 + 2x_4 - 2x_5 \rightarrow \max$$

и множеством допустимых планов D, определенным системой уравнений и неравенств,

$$\begin{array}{llllll}
2x_1 & + 4x_2 & + 5x_3 & & & = 7, \\
 & - 3x_2 & + 4x_3 & - 5x_4 & - 4x_5 & \leq 2, \\
3x_1 & & - 5x_3 & + 6x_4 & - 2x_5 & \leq 4, \\
x_1 \geq 0, & & x_3 \geq 0. & & &
\end{array}$$

Тогда в соответствии со сформулированными правилами эквивалентная каноническая задача будет иметь вид (D', f'), где:

$$f'(x') = 5x_1 + 3\bar{x}_2 - 3\bar{\bar{x}}_2 + x_3 + 2\bar{x}_4 - 2\bar{\bar{x}}_4 - 2\bar{x}_5 + 2\bar{\bar{x}}_5 + 0x'_2 + 0x'_3 \rightarrow \max,$$

а множество D' определено как:

$$
\begin{aligned}
&2x_1 + 4\bar{x}_2 - 4\bar{\bar{x}}_2 + 5x_3 = 7,\\
&\qquad -3\bar{x}_2 + 3\bar{\bar{x}}_2 + 4x_3 - 5\bar{x}_4 + 5\bar{\bar{x}}_4 - 4\bar{x}_5 + 4\bar{\bar{x}}_5 + x_2' = 2,\\
&3x_1 \qquad\qquad - 5x_3 + 6\bar{x}_4 - 6\bar{\bar{x}}_4 - 2\bar{x}_5 + 2\bar{\bar{x}}_5 + x_3' = 4,\\
&x_1 \ge 0,\ \bar{x}_2 \ge 0,\ \bar{\bar{x}}_2 \ge 0,\ x_3 \ge 0,\ \bar{x}_4 \ge 0,\ \bar{\bar{x}}_4 \ge 0,\ \bar{x}_5 \ge 0,\ \bar{\bar{x}}_5 \ge 0,\\
&x_2' \ge 0,\ x_3' \ge 0.
\end{aligned}
$$

Нетрудно заметить, что «платой» за переход от общей формы задачи линейного программирования к канонической является рост ее размерности, что, при прочих равных условиях, является фактором, усложняющим процесс решения.

1.2. ОСНОВНЫЕ СВОЙСТВА ЗЛП И ЕЕ ПЕРВАЯ ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ

1.2.1. Основные понятия линейной алгебры и выпуклого анализа, применяемые в теории математического программирования. Кратко напомним некоторые фундаментальные определения и теоремы линейной алгебры и выпуклого анализа, которые широко применяются при решении проблем как линейного, так и нелинейного программирования.

Фундаментальным понятием линейной алгебры является *линейное (вещественное) пространство*. Под ним подразумевается множество некоторых элементов (именуемых *векторами* или *точками*), для которых заданы операции сложения и умножения на вещественное число (скаляр), причем элементы, являющиеся результатом выполнения операций, также в соответствии с определением должны принадлежать исходному пространству.

Частными случаями линейных пространств являются вещественная прямая, плоскость, геометрическое трехмерное пространство.

Вектор $\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_m a^m$ называется *линейной комбинацией* векторов $a^1\ a^2,\ldots,\ a^m$ с коэффициентами $\lambda_1,\ \lambda_2,\ \lambda_m$,

Система векторов линейного пространства $a^1\ a^2,\ldots,\ a^m$ называется *линейно зависимой*, если существуют такие числа $\lambda_1,\ \lambda_2,\ \lambda_m$ не равные одновременно нулю, что их линейная комбинация $\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_m a^m$ равняется нулевому вектору (вектору, все компоненты которого равны нулю). В противном случае систему $a^1,\ a^2,\ldots,\ a^m$ называют *линейно независимой*, т. е. линейная комбинация данных векторов может быть равна нулевому вектору только при нулевых коэффициентах $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_m$

Максимально возможное количество векторов, которые могут образовывать линейно независимую систему в данном линейном пространстве, называют *размерностью пространства*, а любую систему линейно независимых векторов в количестве, равном размерности, — *базисом пространства*.

Линейное пространство обычно обозначают как R^n, где n — его размерность.

Любое подмножество данного линейного пространства, которое само обладает свойствами линейного пространства, называется *линейным подпространством.* Множество H, получаемое сдвигом некоторого линейного подпространства $L \subset R^n$ на вектор $a \in R^n$: $H=L+a$, называется *аффинным множеством* (пространством). Если фундаментальным свойством любого линейного пространства или подпространства является принадлежность ему нулевого вектора, то для аффинного множества это не всегда так. На плоскости примером подпространства является прямая, проходящая через начало координат, а аффинного множества — любая прямая на плоскости. Характеристическим свойством аффинного множества является принадлежность ему любой прямой, соединяющей две любые его точки. Размерность аффинного множества совпадает с размерностью того линейного подпространства, сдвигом которого оно получено.

Если рассматривается некоторое линейное пространство R^n, то принадлежащие ему аффинные множества размерности 1 называются *прямыми,* а размерности (n-1)—*гиперплоскостями*. Так, обычная плоскость является гиперплоскостью для трехмерного геометрического пространства R^3, а прямая — гиперплоскостью для плоскости R^2. Всякая гиперплоскость делит линейное пространство на два *полупространства.*

Множество V векторов (точек) линейного пространства R^n называется *выпуклым,* если оно содержит отрезок прямой, соединяющей две его любые точки, или, другими словами, из того, что $a \in V$ и $b \in V$, следует, что $x = (1-\lambda) \times a + \lambda \times b \in V$, где $0 \leq \lambda \leq 1$.

Линейная комбинация

$$\sum_{i=1}^{m} \lambda_i a^i$$

векторов $a^1, a^2 \ldots a^m$ называется *выпуклой*, если $\lambda_i \geq 0$, $i \in 1{:}m$ и

$$\sum_{i=1}^{m} \lambda_i = 1 .$$

Множество, содержащее все возможные выпуклые комбинации точек некоторого множества M, называют *выпуклой оболочкой* данного множества. Можно показать, что выпуклая оболочка множества M является наименьшим выпуклым множеством, содержащим M.

Выпуклая оболочка конечного множества точек называется *выпуклым многогранником*, а непустое пересечение конечного числа замкнутых полупространств — *многогранным выпуклым множеством*. В отличие от выпуклого многогранника последнее может быть неограниченным.

Точка v выпуклого множества V называется его *угловой (крайней)* точкой, если она не является внутренней точкой ни для какого отрезка, концы которого принадлежат множеству V. Угловые точки выпуклого многогранника являются его *вершинами*, а сам он — *выпуклой оболочкой* своих вершин.

Множество K называется *конусом* с вершиной в точке x_0, если $x_0 \in K$, и из того, что некоторая точка x принадлежит K ($x \in K$), следует, что в K содержится и луч, начинающийся в x_0 и проходящий через x, т. е.

$$\{x \in R^n \mid x = (1-\lambda)x_0 + \lambda\bar{x},\ \lambda \geq 0\} \subset K$$

или

$$\{x \in R^n \mid x = x_0 + \lambda(\bar{x} - x_0),\ \lambda \geq 0\} \subset K .$$

Выпуклая оболочка конечного множества лучей, исходящих из одной точки, называется *многогранным выпуклым конусом* с вершиной в данной точке.

1.2.2. Первая геометрическая интерпретация ЗЛП и графический метод решения. Рассмотрим следующий пример. Пусть дана задача максимизации линейной целевой функции

$$f(x) = 3x_1 + x_2 \rightarrow \max$$

на множестве

$$D = \{x \in R^2 \mid \begin{array}{lll} x_1 & + x_2 & \leq 6, \\ x_1 & - x_2 & \leq 2, \\ x_1 & & \leq 3, \\ x_1 \geq 0, & x_2 \geq 0\}. & \end{array}$$

Так как количество переменных в неравенствах, задающих область допустимых планов задачи, равно двум, то ее можно изобразить на координатной плоскости (см. *рис. 1.1*).

На *рис. 1.1* показано, что каждое неравенство определяет некоторую полуплоскость. Соответствующие области для каждого ограничения отмечены штрихами. Пересечение D данных полуплоскостей (т. е. множество точек, которые одновременно принадлежат каждой их них) является областью допустимых планов задачи. Поведение целевой функции $f(x) = 3x_1 + x_2$ в рамках двумерной иллюстрации может быть охарактеризовано с помощью линий уровня.

Напомним, что *линией уровня функции* называется множество точек из ее области определения, в которых функция принимает одно и то же фиксированное значение. *Градиентом* функции $f(x)$ называется вектор

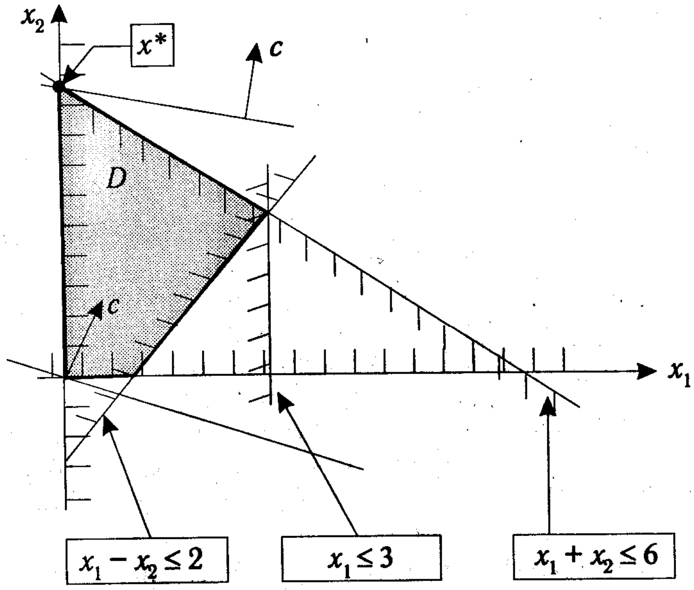

Рис. 1.1

$$\Delta f(x) = \frac{df}{dx_1}, \ldots, \frac{df}{dx_n}$$

указывающий направление наиболее быстрого возрастания функции, и, стало быть, ориентированный перпендикулярно линиям уровня.

Для линейной функции двух переменных линия уровня представляет собой прямую, перпендикулярную вектору *c*, который служит градиентом данной функции. Следовательно, если линия уровня определяется уравнением $f(x)=c_1x_1+ c_2x_2 =const$, то этот вектор имеет вид

$$c=\nabla(c_1x_1+c_2x_2)=(c_1,c_2) \qquad (1.8)$$

и указывает направление возрастания функции.

Таким образом, с геометрической точки зрения задача максимизации сводится к определению такой точки области *D*, через которую проходит линия уровня, соответствующая наибольшему из возможных значений. Последнее означает, что для нахождения точки экстремума в задаче линейного программирования мы должны сначала построить линию уровня для некоторого произвольного значения целевой функции. Затем необходимо осуществлять ее параллельное передвижение (так, чтобы она оставалась перпендикулярной вектору *c*) до тех пор, пока не достигнем такой точки области допустимых планов *D*, из которой смещение в направлении вектора *c* было бы невозможно. Такой метод решения получил название *графического*. Заметим, что решение задачи поиска минимума линейной функции осуществляется аналогично, с той лишь разницей, что движение по линиям уровня должно производиться в направлении, обратном градиенту целевой функции, т. е. по вектору (*-c*).

На *рис. 1.1* изображен некоторый частный случай, для которого решение

ЗЛП достигается в угловой точке $x^* = (0, 6)$ области D. Нетрудно представить, что возможны и другие варианты. Они изображены на *рис. 1.2*.

Рисунок (*a*) иллюстрирует ситуацию неограниченности целевой функции $f(x)=cx$ на множестве D, т.е. сколько бы мы ни перемещались по линиям уровня в направлении вектора c, ее значение будет возрастать.

В случае, изображенном на рисунке (*b*), линия уровня, соответствующая максимальному значению $f(x)$, касается грани множества D, и, соответственно, все точки, лежащие на этой грани, являются оптимальными планами.

Во всех рассмотренных иллюстрациях допустимые планы ЗЛП представлялись в виде некоторого многогранного выпуклого множества на плоскости. Такое их представление в литературе получило название *первой геометрической интерпретации задачи линейного программирования*.

Заметим также, что аналогичным образом могут быть построены интерпретации ЗЛП для случая трехмерного пространства R^3, где множеству D будет соответствовать некоторый ограниченный или неограниченный многогранник, а поведение целевой функции будет характеризоваться поверхностями (плоскостями) уровня.

Несмотря на свою очевидную ограниченность, графический метод решения ЗЛП часто оказывается полезным. В частности, он может быть применен не только к задачам с двумя переменными и ограничениями в виде неравенств, но и к каноническим задачам вида (1.7), у которых $n - m = 2$, где n — количество переменных, а m — ранг матрицы A.

Действительно, можно выбрать две произвольные переменные x_{j1}, x_{j2} и, используя систему уравнений, выразить через них остальные переменные

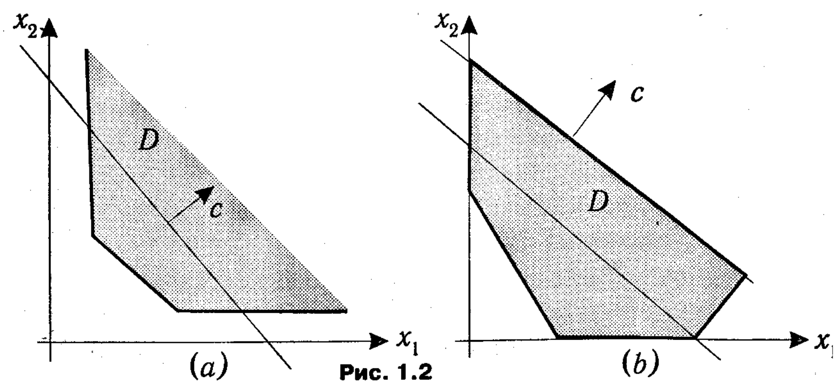

Рис. 1.2

$$x_j = \varphi_j(x_{j_1}, x_{j_2}),\ j \in 1:n \setminus \{j_1, j_2\}, \qquad (1.9)$$

где $\varphi_j(x_{j1}, x_{j2})$ —линейные функции.

Подставив выражения (1.9) в целевую функцию, мы получим эквивалентную задачу

$$f'(x_{j_1}, x_{j_2}) = f(\varphi_j(x_{j_1}, x_{j_2})|_{j \in 1:n \setminus \{j_1, j_2\}}, x_{j_1}, x_{j_2}) \to \max,$$

при ограничениях

$$\varphi_j(x_{j_1}, x_{j_2}) \geq 0,\ j \in 1:n \setminus \{j_1, j_2\},\ x_{j_1} \geq 0,\ x_{j_2} \geq 0.$$

Последняя ЗЛП может быть решена графически.

1.2.3. Основные теоремы линейного программирования. Рассмотрим некоторые теоремы. Отражающие фундаментальные свойства задач линейного программирования и лежащие в основе методов их решения. Они по существу обобщают на случай задач с произвольным количеством переменных те свойства, которые мы наблюдали в двумерном случае.

***Теорема 1.1**. Если целевая функция f принимает максимальное значение в некоторой точке множества допустимых планов D, то она принимает это значение и в некоторой угловой точке данного множества.*

Доказательство.

Чтобы не усложнять изложение, ограничимся тем случаем, когда множество D ограничено, и, следовательно, является выпуклым многогранником.

Для доказательства воспользуемся следующим известным свойством ограниченных выпуклых множеств:

Если D — замкнутое ограниченное выпуклое множество, имеющее конечное число угловых точек, то любая точка $x \in D$ может быть представлена в виде выпуклой комбинации угловых точек D.

Пусть $x^1, x^2, \ldots, x^m$ — угловые точки множества D, а x^* — точка, в которой целевая функция f достигает максимума.

На основе сформулированного выше утверждения точку x^* можно представить в виде выпуклой комбинации угловых точек $x^1, x^2, \ldots, x^m$

$$x^* = \sum_{i=1}^{m} \lambda_i x^i, \text{ где } \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0,\ \ i \in 1:m.$$

Так как x^* — точка максимума, то для любого $x \in D$ $cx^* \geq cx$. Функция $f(x)$ — линейная, поэтому

$$f\left(\sum_{i=1}^{m} \lambda_i x^i\right) = \sum_{i=1}^{m} \lambda_i f(x^i),$$

следовательно,

$$cx^* = c\sum_{i=1}^{m} \lambda_i x^i = \sum_{i=1}^{m} \lambda_i (cx^i) \le$$
$$\le \sum_{i=1}^{m} \lambda_i \ (cx^r) = cx^r \sum_{i=1}^{m} \lambda_i = cx^r, \qquad (1.10)$$

где x^r — угловая точка, удовлетворяющая условию

$$cx^r = \max_{1 \le i \le m} \{cx^i\}.$$

Из (1.10) видно, что $cx^* \le cx^r$. В то же время справедливо обратное неравенство: $cx^* \ge cx^r$. Откуда следует, что $cx^* = cx^r$, т. е. существует по крайней мере одна угловая точка x^r, в которой целевая функция принимает максимальное значение. ✌

Теорема 1.2. *Если целевая функция f принимает максимальное значение в нескольких точках множества D, то она принимает это же значение в любой точке, являющейся их выпуклой комбинацией.*

Доказательство.

Пусть максимальное значение функции f достигается в точках $\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^s$, т. е. $c\tilde{x}^i = f^*$, $i \in 1{:}s$. Рассмотрим произвольную выпуклую комбинацию этих точек

$$x^* = \sum_{i=1}^{s} \lambda_i \tilde{x}^i, \text{ где } \sum_{i=1}^{s} \lambda_i = 1, \ \ \lambda_i \ge 0, \ \ i \in 1{:}s. \qquad (1.11)$$

Найдем значение целевой функции в точке x^*

$$f(x^*) = cx^* = \sum_{i=1}^{s} c(\lambda_i \tilde{x}^i) = \sum_{i=1}^{s} \lambda_i (c\tilde{x}^i) = \sum_{i=1}^{s} \lambda_i f^* = f^* \sum_{i=1}^{s} \lambda_i = f^*. \qquad (1.12)$$

Итак, для произвольной выпуклой комбинации x^* точек $\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^s$ справедливо равенство

$$f(x^*) = cx^* = f^*. \ ✌ \qquad (1.13)$$

1.3. БАЗИСНЫЕ РЕШЕНИЯ И ВТОРАЯ ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ ЗЛП

1.3.1. Векторная форма записи КЗЛП и ее применение. Рассмотрим каноническую задачу линейного программирования

$$f(x) = cx \to \max, \ D = \{x \in R^n \mid Ax = b, \ x \geq 0\}.$$

Обозначим через a^j столбцы матрицы A и будем рассматривать их как векторы пространства R^m. Тогда каждому допустимому плану КЗЛП — n-мерному вектору x — соответствует неотрицательная линейная комбинация столбцов a^j, равная столбцу $b \in R^m$:

$$x_1a^1 + x_2a^2 + \ldots + x_na^n = b, \quad x_i \geq 0, \quad i \in 1:n. \qquad (1.14)$$

Такое представление ограничений КЗЛП обычно называют *векторной формой записи.*

Векторы a^j, $j \in 1{:}n$ будем называть *векторами требований* задачи (D, f), а вектор b — *вектором ограничений*. Множество всех неотрицательных линейных комбинаций столбцов a^j с геометрической точки зрения может быть представлено как многогранный выпуклый конус, натянутый на систему векторов a^j в пространстве R^m *(рис. 1.3).*

Соответственно, вопрос о существовании допустимого плана задачи (D, f) равнозначен вопросу о принадлежности вектора b данному конусу, а компоненты x_j некоторого допустимого плана $x \in D$ являются не чем иным, как *коэффициентами разложения вектора ограничений задачи* b *по векторам, требований* a^j.

Такое представление КЗЛП получило название ***второй геометрической интерпретации.***

В дальнейшем без ограничения общности можем предполагать, что число уравнений, задающих множество D, меньше или равно числу переменных задачи ($m \leq n$). Действительно, если это не так, то либо система уравнений $Ax = b$ несовместна (и, значит, множество D пустое), либо содержит избыточные (линейно зависимые) уравнения.

Если некоторые m столбцов $a^{j1}, a^{j2}, \ldots, a^{jm}$ матрицы A являются линейно независимыми, то они образуют базис в пространстве R^m, и их, вообще говоря, будет достаточно для представления вектора b в виде линейной комбинации указанных столбцов. Это означает, что остальные столбцы войдут в данное разложение с нулевыми коэффициентами. Если к тому же коэффициенты линейной комбинации окажутся неотрицательными, то мы получаем так называемый базисный допустимый план x, у которого не более

m компонентов отличны от нуля. Сформулируем определение базисного плана более строго, так как это одно из фундаментальных понятий теории линейного программирования.

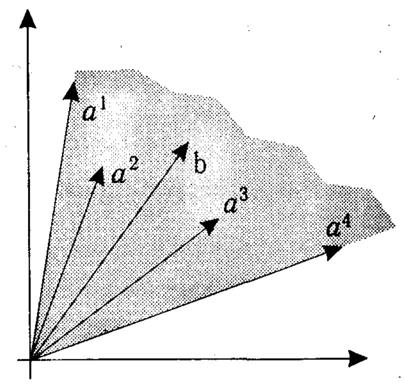

Рис. 1.3

☞ *Пусть задана некоторая каноническая ЗЛП (D,f), A —матрица системы ограничений задачи, и* $\beta = \{a^{j1}, aj2,..., a^{jm}\}$ *—линейно независимая система столбцов матрицы A, образующая базис* R^m*. Обозначим множество номеров столбцов, входящих в систему b, через* $N(\beta) = \{j_1, j_2,..., j_m\}$*. План x называется* **базисным планом** *задачи (D,f), если его компоненты, соответствующие базисным столбцам и называемые базисными компонентами, больше или равны нулю* ($x_j \geq 0,\ j \in N(\beta)$)*, а все остальные компоненты* **(небазисные)** *— равны нулю* ($x_j = 0,\ j \notin N(\beta)$)*.*

☞ *Базисный план x называется* **невырожденным**, *если все его базисные компоненты строго положительны, и* **вырожденным** *в противном случае.*

1.3.2. Свойства базисных планов. Следующая теорема трактует понятие базисного плана в терминах первой геометрической интерпретации ЗЛП.

Теорема 1.3. *Каждый допустимый базисный план является угловой точкой множества допустимых планов D.*

Доказательство.

Ради простоты положим, что базисными векторами являются первые m столбцов матрицы A, т. е. $\beta=\{a^1, a^2,...,a^m\}$. Тогда утверждение теоремы 1.3 может быть переформулировано следующим образом:

Если существует такой n-мерный вектор

$$x = (x_1, x_2, ... x_k, \underbrace{0, ..., 0}_{n-k}), \quad x_j > 0, \quad j \in 1:k, \quad (k \leq m),$$

что $x_1 a^1 + x_2 a^2 + ... + x_k a^k = b$, то x есть угловая точка множества D.

Проведем доказательство от противного, т. е. предположим, что рассматриваемый базисный план x не является угловой точкой множества D. Тогда ее можно представить в виде выпуклой комбинации некоторых двух различных допустимых планов x^1 и x^2:

$$x = \lambda x^1 + (1-\lambda)x^2,\ 0<\lambda<1.$$

В координатной форме последнее утверждение означает

$$x_j = \lambda x_j^1 + (1-\lambda)x_j^2,\ 0<\lambda<1,\ \ j \in 1:n.$$

Поскольку последние $(n - k)$ компоненты вектора x по предположению равны 0, а числа $x_j^1, x_j^2 \geq 0$ и λ, $(1-\lambda)>0$, то эти же компоненты в векторах x_1 и x_2 также равны 0. Поэтому, с учетом допустимости планов x^1 и x^2, можно утверждать, что

$$a^1 x_1^1 + a^2 x_2^1 + ... + a^k x_k^1 = b; \qquad (1.15)$$

$$a^1 x_1^2 + a^2 x_2^2 + ... + a^k x_k^2 = b. \qquad (1.16)$$

Вычитая из (1.15) (1.16), получим

$$a^1(x_1^1 - x_1^2) + a^2(x_2^1 - x_2^2) + ... + a^k(x_k^1 - x_k^2) = 0. \qquad (1.17)$$

Так как векторы $a^1, a^2, ..., a^k$ — линейно независимы, то коэффициенты $x^1{}_1 - x^2{}_1 = 0, ..., x^1{}_k - x^2{}_k = 0$, из чего следует, что $x^1 = x^2$. Это противоречит предположению, что x^1 и x^2 являются различными угловыми точками множества D. Следовательно, x не может быть представлен в виде выпуклой комбинации двух точек D и по определению является угловой точкой данного множества. ✌

Интересно отметить, что справедливо и обратное утверждение, которое приведем без доказательства:

Если x — угловая точка множества D, то она является допустимым базисным планом задачи (D, f).

В завершение параграфа необходимо отметить практическое значение установленной связи между угловыми точками и допустимыми базисными решениями: она позволяет формализовать (и тем самым существенно упростить) процесс перехода от одной угловой точки к другой.

1.4. СИМПЛЕКС-МЕТОД

1.4.1. Основные этапы симплекс-метода. Исходя из свойств линейных экстремальных задач можно заключить, что на принципиальном уровне поиск их решений сводится к последовательному перебору угловых точек множества допустимых планов или, что то же самое, перебору соответствующих допустимых базисных планов. Следует подчеркнуть, что такой перебор для реальных многомерных задач возможен только в теоретическом плане и на практике неосуществим (или крайне неэффективен) даже при условии использования мощной вычислительной техники. Средством решения данной проблемы явились прикладные оптимизационные методы, основанные на последовательном, целенаправленном переборе базисных планов ЗЛП.

Классическим методом решения ЗЛП стал *симплекс-метод*, получивший также в литературе название *метода последовательного улучшения плана*, разработанный в 1947г. американским математиком Джорджем Данцигом.

Идея такого перехода от одного базисного плана к другому, при котором происходит «улучшение» значения целевой фунции, может быть продемонстрирована для случая $m = 2$ с помощью *рис. 1.4.* Если вектор ограничений b принадлежит конусу, натянутому на некоторые два базисных вектора условий $\{a^{j1}, a^{j2}\}$, то существует такой базисный план x с базисными компонентами x_{j1}, x_{j2}, что $b = x_{j1}a^{j1} + x_{j2}a^{j2}$. Разумеется, таких планов может быть несколько в зависимости от выбора системы базисных векторов. Чтобы различать их по соответствующей величине целевой функции $f(x)=c_{j1}x_{j1} + c_{j2}x_{j2}$, вводятся так называемые *расширенные векторы условий* и *ограничений*. В общем случае расширенный столбец условий $\bar{a}^j$ получается соединением коэффициента целевой функции c^j и столбца a^j:

$$\bar{a}^j = (c_j,\ a^j) = (c_j,\ a_{1,j},\ a_{2,j}, \ldots,\ a_{m,j}), \qquad (1.18)$$

расширенный вектор ограничений определим как

$$\bar{b} = (0,\ b_1,\ b_2,\ \ldots,\ b_m). \qquad (1.19)$$

В дальнейшем для удобства нумерации элементов будем считать, что добавляемый коэффициент целевой функции c_j является нулевым элементом j-го расширенного столбца условий, т. е. $\bar{a}_{0,j} = c_j$. При изображении расширенных векторов нулевая координата откладывается вдоль вертикальной оси — *оси аппликат.*

Рассмотрим также вектор

$$\bar{\bar{b}} = x_{j_1}\bar{a}^{j_1} + x_{j_2}\bar{a}^{j_2} + \ldots + x_{j_m}\bar{a}^{j_m}. \qquad (1.20)$$

=

Геометрически определение вектора b означает, что он принадлежит конусу, натянутому на расширенные векторы, а вектор служит его проекцией. Нулевая координата вектора b имеет вид:

$$\bar{\bar{b}}_0 = x_{j_1}\bar{a}_{0,j_1} + x_{j_2}\bar{a}_{0,j_2} + \ldots + x_{j_m}\bar{a}_{0,j_m} =$$
$$= x_{j_1}c_{j_1} + x_{j_2}c_{j_2} + \ldots + x_{j_m}c_{j_m} = f(x), \qquad (1.21)$$

т. е. равна значению целевой функции для выбранного базисного плана.

Из геометрической иллюстрации следует, что *для решения задачи мы должны среди векторов a^j выбрать такой набор $\{a^{j1}, a^{j2}\}$, чтобы, прямая, проведенная через конец вектора параллельно оси аппликат, пересекала конус, натянутый на систему соответствующих расширенных столбцов $\{\bar{a}^{j1}, \bar{a}^{j2}\}$, в «наивысшей» точке.*

На *рис. 1.4* выделен конус, натянутый на систему расширенных столбцов $\bar{a}^2$ и $\bar{a}^3$, отвечающих текущему допустимому базису. Нетрудно заметить, что данный базис не является оптимальным, например, для базиса $\{a^3, a^4\}$ точка пересечения соответствующего конуса и прямой будет находиться выше. Можно, наоборот, указать базис с «худшим» значением целевой функции: $\{a^1\ a^2\}$. Наконец, рассматриваемая геометрическая интерпретация КЗЛП иллюстрирует и общую идею критерия, который используется в симплекс-методе для определения оптимальности (или неоптимальности) текущего плана: *если существуют векторы-столбцы, лежащие выше плоскости, проходящей через векторы текущего базиса, то он не является оптимальным и может быть «улучшен».*

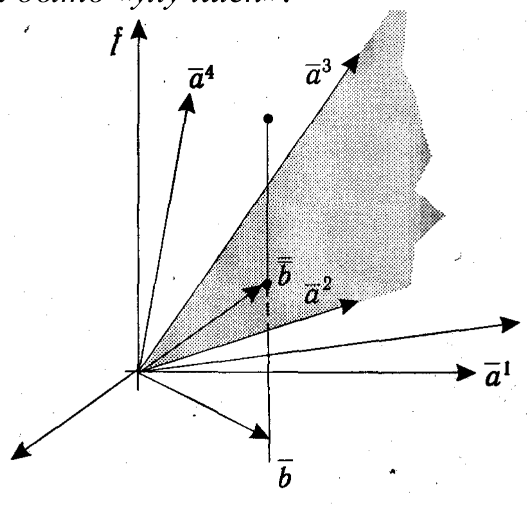

Рис. 1.4

С вычислительной точки зрения критерий оптимальности в симплекс-методе реализуется через нахождение специальных оценок для небазисных векторов-столбцов, относительно текущего базиса.

Для удобства дальнейшего изложения введем некоторые обозначения. Учитывая то, что симплекс-метод представляет собой некоторый итеративный вычислительный процесс, то через q будем обозначать номер очередной (текущей) итерации. Соответственно, набор базисных столбцов, получаемых на q-й итерации, будет обозначаться как $\beta^{(q)}$:

$$\beta^{(q)} = \{a^{j_1}, a^{j_2}, \ldots, a^{j_m}\}. \qquad (1.22)$$

Для того чтобы было легче отличить номер итерации от номеров компонентов матриц и векторов, он заключен в круглые скобки. Номера столбцов, входящих в базис, обозначим через $N(\beta^{(q)})$, а именно

$$N(\beta^{(q)}) = \{j_1, j_2, \ldots, j_m\}. \qquad (1.23)$$

При этом $N_r(\beta^{(q)}) = j_r$ — номер столбца, занимающего r-ю позицию в базисе. Тогда текущий базисный план x имеет вид:

$$x = (x_1, x_2, \ldots, \ x_n), \quad \text{где } \ x_j = 0 \text{ для } j \notin N(\beta^{(q)}).$$

Обозначим через $\Delta(\beta^{(q)})$ матрицу, составленную из столбцов $\{a^{j1}, a^{j2}, \ldots, a^{jm}\}$, образующих базис, через $\Delta(\beta^{(q)})$ матрицу, образованную из соответствующих расширенных столбцов и дополненную слева столбцом специального вида:

$$\bar{\Delta}(\beta^{(q)}) = \left[\begin{array}{c|cccc} -1 & c_{j_1} & c_{j_2} & \cdots & c_{j_m} \\ \hline 0 & a_{1,j_1} & a_{1,j_2} & \cdots & a_{1,j_m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{m,j_1} & a_{m,j_2} & \cdots & a_{m,j_m} \end{array}\right], \qquad (1.24)$$

а через $\Delta^{-1}(\beta^{(q)})$ и $\bar{\Delta}^{-1}(\beta^{(q)})$ — матрицы, обратные по отношению к ним. Также представляется удобным ввести отдельные обозначения для элементов матрицы $\Delta^{-1}(\beta^{(q)})$:

$\delta_i(\beta^{(q)})$ — i-я строка матрицы $\bar{\Delta}^{-1}(\beta^{(q)})$, ($i \in 0: m$);

$\delta_{ij}(\beta^{(q)})$ — элемент матрицы $\Delta^{-1}(\beta^{(q)})$, находящийся в i-й строке j-го столбца ($i \in 0: m, j \in 0 : m$).

Расширенный вектор ограничений b представляется в виде линейной комбинации расширенных векторов условий с коэффициентами, равными базисным компонентам текущего базисного плана:

$$\bar{b} = x_{j_1}\bar{a}^{j_1} + x_{j_2}\bar{a}^{j_2} + \ldots \ + x_{j_m}\bar{a}^{j_m}.$$

Если интерпретировать компоненты векторов $\bar{a}^j$ и $\bar{b}$ как координаты в ортогональном базисе, то их столбцы координат относительно произвольного базиса ($\beta^{(q)}$), дополненного единичным вектором (-1,0,..., 0), направленным противоположно оси аппликат примут вид:

$$\bar{a}^j(\beta^{(q)}) = \bar{\Delta}^{-1}(\beta^{(q)})\bar{a}^j,\ \bar{b}(\beta^{(q)}) = \bar{\Delta}^{-1}(\beta^{(q)})\bar{b}, \qquad (1.25)$$

а для расширенной матрицы задачи в целом можно записать

$$\bar{A}(\beta^{(q)}) = \bar{\Delta}^{-1}(\beta^{(q)})\bar{A}. \qquad (1.26)$$

Нулевая строка данной матрицы $a_0(\beta^{(q)})$ содержит координаты расширенных векторов условий по оси аппликат. Согласно построению, элементы данной строки имеют следующие знаки:

- $a_{0,j}$ ($\beta^{(q)}$) < 0 — *для расширенных векторов условий, расположенных выше плоскости, натянутой на систему расширенных базисных векторов;*
- $a_{0,j}$ ($\beta^{(q)}$) > 0 — *для расширенных векторов условий, расположенных ниже плоскости, натянутой на систему расширенных базисных векторов;*
- $a_{0,j}$ ($\beta^{(q)}$) = 0 — *для расширенных базисных векторов.*

Подводя итог сказанному, сформулируем ***критерий оптимальности допустимого базисного плана*** в симплекс-методе:

☞ *план является оптимальным, если для всех $j \in 1: n$ $a_{0,j}$ ($\beta^{(q)}$) ≥ 0, и неоптимальным в противном случае, т. е. если существует такое $l \in 1 : n$,* что a_{0l} ($\beta^{(q)}$) < 0.

Значения $a_{0,j}$ ($\beta^{(q)}$) также называют *оценками столбцов матрицы A* относительно текущего базиса, или *симплекс-разностями.*

В случае неоптимальности текущего базиса в алгоритме симплекс-метода осуществляется переход к следующему базису. Это делается за счет вывода одного столбца из базиса и ввода другого. Для обеспечения улучшения значения целевой функции в базис должен быть введен вектор-столбец, имеющий отрицательную оценку. Если таких столбцов несколько, то ***для ввода рекомендуется выбирать столбец***, имеющий *максимальную по модулю оценку*. Отметим, что данное правило носит относительный характер и не гарантирует наилучшего выбора вводимого столбца. Одновременно на этой стадии требуется принять решение о том, какой столбец следует вывести из базиса. Сделать это нужно таким образом, чтобы вновь формируемый базис оказался допустимым. Данное требование может быть легко проиллюстрировано для случая m = 2. Например, на *рис. 1.3* векторы {a^2,a^3} образуют допустимый базис, а векторы {a^3,a^4} —недопустимый, т. к. разложение b по a^3 и a^4 содержит один отрицательный компонент плана, что противоречит условиям КЗЛП.

Можно доказать, что допустимость базиса, к которому осуществляется переход, обеспечивается следующим ***правилом вывода столбца из***

текущего базиса:

☞ *для столбца l, претендующего на ввод в базис, и вектора ограничений b рассматриваются отношения*

$$\lambda_i = \frac{b_i(\beta^{(q)})}{a_{i,l}(\beta^{(q)})}, \text{ где } i \in \{1:m \mid a_{i,l}(\beta^{(q)}) > 0\}$$

и определяется такая строка r, что

$$\lambda_r = \min_i \{\lambda_i\}. \qquad (1.27)$$

Полученный индекс r определяет номер столбца в $N(\beta^{(q)})$, *выводимого из базиса, а именно,* $N(\beta^{(q)})$.

Таким образом, если базис на q-й итерации включал столбцы с номерами

$$N(\beta^{(q)}) = \{j_1, j_2, \ldots, j_{r-1}, j_r, j_{r+1}, \ldots, j_m\},$$

то базис на итерации $q + 1$ будет состоять из столбцов с номерами:

$$N(\beta^{(q+1)}) = \{j_1, j_2, \ldots, j_{r-1}, j_l, j_{r+1}, \ldots, j_m\}.$$

Отдельно следует обсудить тот случай, когда столбец a_l $(\beta^{(q)})$, претендующий на ввод в базис, не содержит положительных компонентов a_l $(\beta^{(q)}) \leq 0)$. Это означает, что целевая функция в задаче не ограничена на множестве допустимых значений, т. е. может достигать сколь угодно большого значения. Последнее, очевидно, означает завершение процесса вычислений ввиду *отсутствия оптимального плана.* Геометрически ситуация, когда a^l $(\beta^{(q)}) \leq 0$, соответствует тому, что ось ординат оказывается внутри конуса, натянутого на систему расширенных столбцов a^j, а значит, прямая, проведенная через конец вектора b параллельно оси аппликат, однажды «войдя» в этот конус, более никогда из него «не выходит».

– Вообще говоря, после перехода от базиса $\beta^{(q)}$ к базису $\beta^{(q+1)}$ мы можем заново сформировать матрицы Δ $(\beta^{(q+1)})$, Δ^{-1} $(\beta^{(q+1)})$ и, вычислив $A(\beta^{(q+1)}) = \Delta^{-1}$ $(\beta^{(q+1)})A$, делать выводы о его оптимальности. Однако, учитывая, что $\beta^{(q+1)}$ отличается от $\beta^{(q)}$ всего лишь одним столбцом, с точки зрения техники вычислений представляется рациональным непосредственно переходить от A $(\beta^{(q)})$ и b $(\beta^{(q)})$ к A $(\beta^{(q+1)})$ и b $(\beta^{(q+1)})$. Дело в том, что у матриц типа $\bar{A}$ $(\beta^{(q)})$ столбцы, соответствующие базисным векторам, состоят из нулей, за исключением одного элемента, равного единице. Позиция этого ненулевого элемента определяется порядковым номером базисного столбца в $N(\beta^{(q)})$. Поэтому для получения матрицы A $(\beta^{(q+1)})$ достаточно с помощью линейных операций над строками матрицы A $(\beta^{(q)})$ привести ее столбец, соответствующий вводимому в базис вектору, к «базисному» виду.

Для это применяется *преобразование Жордана—Гаусса* (так называемый метод полного исключения). В данном случае оно состоит в том, что мы должны «заработать» единицу на месте элемента $a_{r,l}(\beta^{(q)})$ (он обычно называется *ведущим*)* и нули на месте остальных элементов столбца $a^{l}(\beta^{(q)})$. Первое достигается посредством деления r-й строки на ведущий элемент, второе — путем прибавления вновь полученной r-й строки, умноженной на подходящий коэффициент, к остальным строкам матрицы $A(\beta^{(q)})$.

* Напомним, что l — номер столбца, вводимого в базис, а r — номер строки в симплекс-таблице, определяющей номер столбца, выводимого из базиса.

Формально результат выполнения данного преобразования над элементами $A(\beta^{(q)})$ и $b(\beta^{(q)})$ может быть выражен в следующем виде

$$a_{r,j}(\beta^{(q+1)}) = \frac{a_{r,j}(\beta^{(q)})}{a_{r,l}(\beta^{(q)})},\ j \in 1:n; \tag{1.28}$$

$$b_r(\beta^{(q+1)}) = \frac{b_r(\beta^{(q)})}{a_{r,l}(\beta^{(q)})}; \tag{1.29}$$

$$a_{i,j}(\beta^{(q+1)}) = a_{i,j}(\beta^{(q)}) - a_{i,l}(\beta^{(q)})\frac{a_{r,j}(\beta^{(q)})}{a_{r,l}(\beta^{(q)})},\ i \in 0:m,$$

$$i \neq r, j \in 1:n; \tag{1.30}$$

$$b_i(\beta^{(q+1)}) = b_i(\beta^{(q)}) - a_{i,l}(\beta^{(q)})\frac{b_r(\beta^{(q)})}{a_{r,l}(\beta^{(q)})},\ i \in 0:m,\ i \neq r. \tag{1.31}$$

Следует особо отметить смысл элементов вектора $b(\beta^{(q)})$. Его нулевой компонент $b_0(\beta^{(q)})$ в соответствии с построением содержит *значение целевой функции*, достигаемое ею на текущем плане

$$f(x(\beta^{(q)})) = \bar{b}_0(\beta^{(q)}), \tag{1.32}$$

а остальные элементы — *ненулевые компоненты этого плана:*

$$x_{N_i(\beta^{(q)})}(\beta^{(q)}) = \bar{b}_i(\beta^{(q)}),\ i \in 1:m. \tag{1.33}$$

Название метода произошло от понятия симплекса. Напомним, что *m-симплексом* называют выпуклый многогранник, аффинная оболочка* которого есть аффинное множество размерности m. В данном случае можно считать, что система расширенных базисных столбцов $\{a^{j1}, a^{j2}, \dots, a^{jm}\}$, рассматриваемых как точки в R^{m+1}, порождает $(m$ -1)-мерный симплекс в пространстве R^{m+1}.

* *Аффинной оболочкой* множества называют наименьшее аффинное

множество, в котором содержится данное множество.

В заключение настоящего пункта обобщим изложенные вопросы и приведем *схему алгоритма симплекс-метода для решения задачи максимизации*. Она включает однократно выполняемый 0-этап и повторяемый конечное число раз I-этап (стандартную итерацию).

0-этап. Нахождение допустимого базисного плана (см. п. 1.4.5). Результатом 0-этапа является допустимый базисный план $x(\beta^{(1)})$, а также соответствующие ему матрица $A(\beta^{(1)})$ и вектор $b(\beta^{(1)})$, которые будут использованы на первой итерации. Полагаем номер текущей итерации q равным 1 и переходим к I-этапу.

I-этап. Стандартная итерация алгоритма — выполняется для очередного базисного плана $x(\beta^{(q)})$.

1°. *Проверка оптимальности* текущего базисного плана: осуществляется просмотр строки оценок $a_0(\beta^{(q)})$. Возможны два варианта:

1′. $a_0(\beta^{(q)})\geq 0$ — план, соответствующий текущему базису задачи, **оптимален**. Вычислительный процесс закончен. Согласно формулам (1.33) и (1.32) выписываются оптимальный план задачи $x^* = x(\beta^{(q)})$ и значение целевой функции $f(x^*) = f(x(\beta^{(q)}))$.

1″. В строке оценок $a_0(\beta^{(q)})$ существует по меньшей мере один элемент $a_{0,j}(\beta^{(q)})<0$, т. е. имеющий отрицательную оценку. Следовательно, план $x(\beta^{(q)})$ — **неоптимален.** Выбирается столбец с номером l, имеющий минимальную (максимальную по абсолютной величине) отрицательную оценку

$$a_{0,l}(\beta^{(q)}) = \min_{j:a_{0,j}(\beta^{(q)})<0} \{a_{0,j}(\beta^{(q)})\}.$$

Он называется *ведущим* и *должен быть введен в очередной базис*. Переходим к пункту 2° алгоритма.

2°. *Определение столбца, выводимого из базиса*. Рассматривается ведущий столбец $a^l(\beta^{(q)})$ Возможны два варианта:

2'. Для всех $i\in 1{:}m$ $a_{i,l}(\beta^{(q)})\leq 0$. Делается вывод о *неограниченности целевой* функции и завершается вычислительный процесс.

2''. Существует по крайней мере одна строка с номером $i\in 1{:}m$, для которой $a_{i,l}(\beta^{(q)})>0$. Согласно правилу (1.27) определяются место r и номер $N_r(\beta^{(q)}) = j_r$ для столбца, выводимого из базиса. Переходим к пункту 3° алгоритма.

3°. *Пересчет элементов матрицы A и столбца* b относительно нового базиса. В соответствии с формулами (1.28)-(1.31) осуществляем расчет элементов матрицы задачи A и вектора ограничений b относительно базиса $\beta^{(q+1)}$, который получается после введения в базис $\beta^{(q)}$ столбца a^l и выводом из него столбца a^r. Полагаем номер текущей итерации $q := q+1$ и переходим к первому пункту алгоритма.

$$T^{(q)} = \begin{array}{|c|c|c|} \hline & b_0(\beta^{(q)}) & a_0(\beta^{(q)}) \\ \hline N(\beta^{(q)}) & b(\beta^{(q)}) & A(\beta^{(q)}) \\ \hline \end{array}$$

Рис.1.5

1.4.2. Табличная реализация симплекс-метода. С точки зрения обеспечения рациональности и наглядности вычислительного процесса выполнение алгоритма симплекс-метода удобно оформлять в виде последовательности таблиц. В различных источниках приводятся разные модификации симплекс-таблиц, отличающиеся друг от друга расположением отдельных элементов. Однако это не должно вызывать смущения у читателя, так как все они базируются на одних и тех же принципах. Остановимся на структуре таблицы, показанной на рис. 1.5.*

---

* Настоящая структура симплекс-таблиц строится на идеях и принципах их организации, предложенных в [1].

Симплекс-таблица $T^{(q)}$, изображенная на *рис. 1.5*, соответствует допустимому базису КЗЛП $\beta^{(q)}$), получаемому на q-й итерации. Столбец $N(\beta^{(q)})$ содержит номера базисных столбцов (в той последовательности, в которой они входят в базис), столбец $b(\beta^{(q)})$ —компоненты вектора ограничений относительно текущего базиса $\beta^{(q)}$, $A(\beta^{(q)})$ — компоненты матрицы задачи относительно текущего базиса $\beta^{(q)}$. Наконец, в строке $a_0(\beta^{(q)})$ находятся текущие оценки столбцов, а ячейка $b_0(\beta^{(q)})$ содержит значение целевой функции, достигаемое для текущего плана.

Безусловно, следует добавить, что табличная модификация симплекс-метода имеет важное практическое значение не столько как удобная форма организации ручного счета, сколько как основа для реализации данного алгоритма в рамках программного обеспечения ЭВМ.

1.4.3. Пример решения ЗЛП симплекс-методом. Рассмотрим на конкретном примере процесс решения КЗЛП табличным симплекс-методом. Пусть дана каноническая задача ЛП:

$$f(x) = 50x_1 - 10x_2 + 6x_3 + 40x_4 - 30x_5 \to \max, \qquad (1.34)$$

$$\begin{cases} x_1 + x_4 + 3x_5 = 12, \\ 2x_1 + x_2 + 3x_4 = 14, \\ -2x_1 + 3x_3 - 4x_4 = 17, \\ x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0,\ x_4 \ge 0,\ x_5 \ge 0 \end{cases}. \qquad (1.35)$$

Как видно, столбцы матрицы с номерами {5, 2, 3} являются линейно независимыми. И можно получить разложение по данным столбцам вектора ограничений с положительными коэффициентами. Последнее означает, что столбцы {5, 2, 3} образуют допустимый базис, с которого можно начать решение задачи. Из столбцов, входящих в базис, с учетом нулевых элементов формируется матрица $\Delta(\beta^{(1)})$ и обратная по отношению к ней $\Delta^{-1}(\beta^{(1)})$:

$$\bar{\Delta}(\beta^{(1)})=\left[\begin{array}{c|ccc}-1 & -30 & -10 & 6\\ \hline 0 & 3 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 3\end{array}\right],\ \bar{\Delta}^{-1}(\beta^{(1)})=\left[\begin{array}{c|ccc}-1 & -10 & -10 & 2\\ \hline 0 & 1/3 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1/3\end{array}\right].$$

Используя их, по формуле (1.26) получаем

$$\bar{A}(\beta^{(1)})=\bar{\Delta}^{-1}(\beta^{(1)})\bar{A}=$$

$$\left[\begin{array}{c|ccc}-1 & -10 & -10 & 2\\ \hline 0 & 1/3 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1/3\end{array}\right]\times\left[\begin{array}{ccccc}50 & -10 & 6 & 40 & -30\\ \hline 1 & 0 & 0 & 1 & 3\\ 2 & 1 & 0 & 3 & 0\\ -2 & 0 & 3 & -4 & 0\end{array}\right]$$

$$=\left|\begin{array}{ccccc}-84 & 0 & 0 & -88 & 0\\ 1/3 & 0 & 0 & 1/3 & 1\\ 2 & 1 & 0 & 3 & 0\\ -2/3 & 0 & 1 & -4/3 & 0\end{array}\right|;$$

$$\bar{b}(\beta^{(1)})=\bar{\Delta}^{-1}(\beta^{(1)})\times\bar{b}=\left[\begin{array}{c|ccc}-1 & -10 & -10 & 2\\ \hline 0 & 1/3 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1/3\end{array}\right]\times\left[\begin{array}{c}0\\ \hline 12\\ 14\\ 17\end{array}\right]=\left[\begin{array}{c}-226\\ \hline 4\\ 14\\ 17/3\end{array}\right]$$

Используя полученные значения $\bar{A}(\beta^{(1)})$ и $\bar{b}(\beta^{(1)})$, заполняем симплекс-таблицу $T^{(1)}$, соответствующую первой итерации (q=1).

$$T^{(1)}=\begin{array}{|c|cccc|c|c|}\hline & -226 & -84 & 0 & 0 & -88 & 0\\ 5 & 4 & 1/3 & 0 & 0 & 1/3 & 1\\ \hline 2 & 14 & 2 & 1 & 0 & 3 & 0\\ \hline 3 & 17/3 & -2/3 & 0 & 1 & -4/3 & 0\\ \hline\end{array}$$

Поскольку строка оценок $a_0(\beta^{(1)})$ в первом и четвертом столбцах содержит отрицательные элементы ($a_{0,1}(\beta^{(1)})$ = -84, $a_{0,4}(\beta^{(1)})$ = -88), план $x(\beta^{(1)})$ = (0,14,17/3,0,4) не является оптимальным, и значение целевой функции $f(x(\beta^{(1)}))$ = -226 может быть улучшено. Следуя рекомендации п.1″ алгоритма симплекс-метода, полагаем номер вводимого в очередной базис столбца l = 4 (т.к. |-88| > |-84|). Рассматриваем *ведущий* столбец (выделен пунктирной рамкой)

$$a^4(\beta^{(1)})=\left[\begin{array}{r}1/3\\ 3\\ -4/3\end{array}\right]$$

Он содержит два положительных элемента. Применяя рекомендацию п. 2" алгоритма, получаем λ_1= 4/(1/3) =12, λ_2=14/3, и, стало быть r = 2. Следовательно, столбец с номером $N_2(\beta^{(1)})$ = 2 должен быть выведен из базиса. Таким образом, получаем очередной допустимый базис задачи $N(\beta^{(2)})$ = {5, 4, 3}. Элемент $a_{2,3}(\beta^{(1)})$ является ведущим (обведен кружком). Применив формулы (1.28)-(1.31), переходим к симплекс-таблице, соответствующей второй итерации $T^{(2)}$, и полагаем индекс текущей итерации q = 2.

$T^{(2)}$ =

	554/3	−76/3	88/3	0	0	0
5	22/9	1/9	−1/9	0	0	1
4	14/3	2/3	1/3	0	1	0
3	107/9	2/9	4/9	1	0	0

В ходе выполнения второй итерации опять-таки определяются вводимый столбец a^1, выводимый a^4 и ведущий элемент $a_{2,1}(\beta^{(2)})$. Перейдя к итерации q = 3, имеем таблицу $T^{(3)}$.

$T^{(3)}$ =

	362	0	42	0	38	0
5	5/3	0	−1/6	0	−1/6	1
1	7	1	1/2	0	3/2	0
3	31/3	0	1/3	1	−1/3	0

Как видно из $T^{(3)}$, строка оценок содержит только неотрицательные значения, поэтому достигнутый базис $N(\beta^{(3)})$ = {5, 1, 3} является оптимальным. В итоге мы получаем, что вектор x^* = (7, 0, 31/3, 0, 5/3) является оптимальным планом (точкой максимума) задачи (1.34)-(1.35), максимальное значение целевой функции равно $f^* = f(x^*)$=362, а решение КЗЛП имеет вид ((7, 0, 31/3, 0, 5/3); 362).

1.4.4. Сходимость симплекс-метода. Вырожденность в задачах ЛП. Важнейшим свойством любого вычислительного, алгоритма является *сходимость*, т. е. возможность получения в ходе его применения искомых результатов (с заданной точностью) за конечное число шагов (итераций).

Легко заметить, что проблемы со сходимостью симплекс-метода потенциально могут возникнуть на этапе выбора значения r (п. 2") в случае, когда одинаковые минимальные значения отношения

$$\lambda_i = \frac{b_i(\beta^{(q)})}{a_{i,l}(\beta^{(q)})} \quad (i \in \{1:m \mid a_{i,l}(\beta^{(q)}) > 0\})$$

будут достигнуты для нескольких строк таблицы $T^{(q)}$ одновременно. Тогда на следующей итерации столбец $b(\beta^{(q+1)})$ будет содержать нулевые элементы.

Напомним, что

☞ *допустимый базисный план канонической задачи ЛП, соответствующий базису* $\beta^{(q)}$, *называется* **вырожденным**, *если некоторые его базисные компоненты равны нулю, т. е. вектор* $b(\beta^{(q)})$ *содержит нулевые элементы.*

Задача ЛП, имеющая вырожденные планы, называется ***вырожденной***. При «выходе» на вырожденный план мы фактически получаем разложение столбца b по системе из меньшего, чем m, количества столбцов a^j и, следовательно, лишаемся возможности корректно определить, ввод какого столбца в базис приведет к росту значения целевой функции. Подобные ситуации, в конечном счете, могут привести к зацикливанию вычислительного процесса, т. е. бесконечному перебору одних и тех же базисов.

С точки зрения первой геометрической интерпретации ЗЛП ситуация вырожденности означает, что через некоторую угловую точку многогранного множества допустимых планов задачи, соответствующую текущему базисному плану, проходит более чем m гиперплоскостей ограничений задачи. Иными словами, одно или несколько ограничений в данной точке являются избыточными. Последнее предоставляет повод для размышлений об экономической стороне проблемы вырожденности как проблемы наличия избыточных ограничений и в некоторых случаях определяет пути ее решения.

С точки зрения второй геометрической интерпретации ЗЛП вырожденность означает «попадание» линии, проходящей через вершину вектора b параллельно оси аппликат, в ребро конуса, натянутого на систему расширенных векторов текущего базиса. Так, в случае, изображенном на *рис. 1.6,* эта линия попадает в ребро, определяемое вектором a^{j2} , т. е., несмотря на то что текущий базис содержит столбцы a^{j1} и a^{j2}, для представления вектора b достаточно только одного a^{j2}.

Из данной интерпретации вытекает и идея метода решения вырожденных задач ЛП, получившего названия *метода возмущений*: при выходе на вырожденный план осуществляется незначительный сдвиг вектора b и вырожденность (как это видно из иллюстрации) устраняется.

Необходимо, однако, сказать, что рассмотренная здесь проблема зацикливания для большинства практически значимых задач является достаточно редкой и маловероятной. Более того, она очень часто разрешается за счет ошибок округлений при выполнении расчетов на ЭВМ.

1.4.5. Нахождение допустимого базисного плана. В рассмотренном выше примере исходный базисный план, необходимый для начала вычислений по симплекс-методу, был подобран за счет особенностей матрицы условий. Действительно, данная матрица уже содержала необходимое количество «почти базисных» столбцов. Очевидно, что для подавляющего большинства задач ЛП невозможно подобным образом сразу и в явном виде указать исходный допустимый базисный план. Вообще говоря, существуют различные приемы решения данной задачи. Мы остановимся на одном из них, получившем название *метода минимизации невязок*. Его сильной стороной, безусловно, является универсальность. Хотя, в некоторых частных случаях, он может оказаться слишком громоздким.

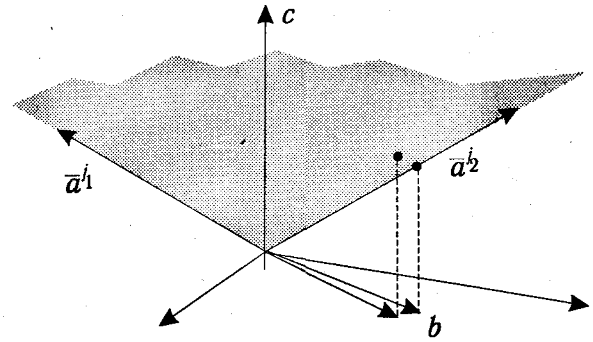

Рис. 1.6

Идея метода минимизации невязок состоит в построении соответствующей вспомогательной задачи, для которой можно в явном виде указать исходный базисный план, и решении ее с помощью процедуры симплекс-метода.

Не умаляя общности, можно считать, что вектор ограничений в исходной задаче неотрицателен, т. е. $b^i \geq 0$, $i \in 1:m$. Тогда для КЗЛП максимизации функции

$$f(x) = c_1x_1 \quad \dots \quad + c_nx_n \to \max \tag{1.36}$$

на множестве, определяемом ограничениями

$$\begin{aligned} &a_{1,1}x_1 \quad \dots \quad + a_{1,n}x_n = b_1, \\ &\qquad\quad \dots \\ &a_{m,1}x_1 \quad \dots \quad + a_{m,n}x_n = b_m, \\ &x_1 \geq 0, \qquad\quad x_n \geq 0. \end{aligned} \tag{1.37}$$

строится вспомогательная задача

$$\tilde{f}(x) = 0 \times x_1 \dots + 0 \times x_n - 1 \times x_{n+1} \dots - 1 \times x_{n+m} \to \max \tag{1.38}$$

$$\begin{aligned} &\tilde{D} = \{x \in R^{n+m} \mid \\ &a_{1,1}x_1 \dots + a_{1,n}x_n + 1 \times x_{n+1} \dots + 0 \times x_{n+m} = b_m, \\ &\qquad\qquad \dots \\ &a_{m,1}x_1 \dots + a_{m,n}x_n + 0 \times x_{n+1} \dots + 1 \times x_{n+m} = b_m \\ &x_1 \geq 0, \quad x_n \geq 0, \; x_{n+1} \geq 0, x_{n+m} \geq 0\}. \end{aligned} \tag{1.39}$$

Как видно из (1.36)-(1.39), задача $(\tilde{D}, \tilde{f})$ отличается от задачи (D, f) матрицей, получаемой в результате добавления к исходной матрице A единичной матрицы, которой соответствуют дополнительные (фиктивные) переменные $x_{n+1},...,x_{n+m}$. Данным переменным (собственно говоря, их и называют *невязками*) в целевой функции $\tilde{f}$ отвечают коэффициенты (-1), а переменным $x_1,...,x_n$, которые являются общими для обеих задач, — нулевые коэффициенты. Существенной особенностью $(\tilde{D}, \tilde{f})$ является наличие в ней очевидного исходного базиса, образуемого добавленными столбцами, и соответствующего плана с базисными компонентами $x_{n+i} = b_i \geq 0$, $i \in 1{:}\ m$. В силу структуры целевой функции $\tilde{f}$ в процесс поиска ее максимума процедурой симплекс-метода фиктивные переменные (невязки) x_{n+i} должны минимизироваться желательно до нуля, что может быть достигнуто за счет последовательного перевода их в небазисные компоненты плана. Если в оптимальном плане $\tilde{x}$, полученном в результате решения вспомогательной задачи, последние m компонент (т. е. невязки) равны нулю, то его начальные n компонент удовлетворяют системе ограничений, определяющих область D, и $\tilde{x}$ легко (простым отбрасыванием невязок) может быть преобразован в стартовый допустимый план основной задачи (D, f). При этом все строки симплекс-таблицы, полученной на последней итерации вспомогательной задачи (за исключением нулевой), могут быть перенесены в первую таблицу основной задачи. Элементы же нулевой строки для вновь формируемой симплекс-таблицы вычисляются по формулам:

$$a_{0,j}(\beta^{(1)}) = -c_j + c(\beta^{(*)})a^j(\beta^{(*)}), \qquad (1.40)$$

$$b_0(\beta^{(1)}) = c(\beta^{(*)})b(\beta^{(*)}), \qquad (1.41)$$

где $\beta^{(*)}$ — базис, полученный на последней итерации вспомогательной задачи.

В случае, когда оптимальный план вспомогательной задачи $\tilde{x}$ все же содержит не равные нулю фиктивные компоненты (т. е. $\tilde{f}(\tilde{x})<0$), мы приходим к заключению об *отсутствии допустимых планов* у исходной задачи (D, f), т. е. $D = \varnothing$.