2,8

Министерство образования Республики Беларусь

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Кафедра автоматизированных систем управления

С.С.Смородинский, Н.В.Батин

МЕТОДЫ И АЛГОРИТМЫ ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Учебно-методическое пособие
по курсу "Системный анализ и исследование операций"

для студентов специальности
"Автоматизированные системы обработки информации и управления"

Часть 1

Минск 1995

УДК 519.816

Смородинский С.С., Батин Н.В. Методы и алгоритмы для решения оптимизационных задач линейного программирования: Учебно-методическое пособие по курсу "Системный анализ и исследование операций" для студентов специальности "Автоматизированные системы обработки информации и управления". Ч.I.-Мн.:БГУИР, 1995.- 91с.

В пособии рассматривается методология решения оптимизационных задач линейного программирования: вычислительная процедура симплекс-метода и технология анализа модели на чувствительность, методы искусственного базиса, методы и алгоритмы для решения задач целочисленного программирования, методы и алгоритмы для решения транспортных задач, задачи о назначениях.

Пособие рекомендуется студентам специальности АСОИУ, изучающим проблематику моделирования и оптимизации решений на основе аппарата математического программирования. Пособие целесообразно использовать при освоении практической части курса "Системный анализ и исследование операций", а также в курсовом и дипломном проектировании. Пособие представляет интерес для специалистов, решающих задачи оптимизации в различных направлениях науки, техники и экономики.

Ил. 3, табл. 18, список лит. - 38 назв.

"Прежде всего системный анализ предназначен для решения слабоструктуризованных проблем, в которых преобладают качественные, малоизвестные и неопределенные стороны".

Ларичев О. И., 1979

"Когда используют термин "исследование операций" то почти всегда имеют в виду применение математических методов для моделирования систем и анализа их характеристик".

Таха Х., 1985

ВВЕДЕНИЕ

Пособие связано с проблематикой курса "Системный анализ и исследование операций". Цель курса - освоение системной методологии решения сложных проблем, которые возникают в процессе создания и развития целенаправленных организационно-технических систем различных уровней и назначения. Курс относится к числу дисциплин, образующих научную "подкладку" для разработки математического и программного обеспечения новых информационных технологий в автоматизированных системах обработки информации.

Основные задачи курса

1. Изучение методологических основ системного анализа и исследования операций для решения проблем с различной степенью структуризации.

2. Изучение принципов подготовки и принятия решений в условиях многовариантности, многокритериальности и неопределенности.

3. Изучение принципов аналитического моделирования систем, операций и процессов для задач прогнозирования, планирования, проектирования и управления.

Аналитическое моделирование возможно осуществлять на основе аппарата математического программирования, которое рассматривает задачи оптимизации, включая модели, методы и алгоритмы решения экстремальных задач при наличии ограничений.

Сущность задач оптимизации: определить значения переменных $X_1, X_2, \ldots, X_n$, которые обеспечивают экстремум целевой функции E с учетом ограничений, наложенных на аргументы функции $X_1, X_2, \ldots, X_n$ (элементы решения). При этом сложность решения задач зависит: 1) от вида функциональной зависимости, т.е. от связи функции E с элементами решения; 2) от размерности задачи, т.е. от количества элементов решения; 3) от вида и количества ограничений, наложенных на элементы решения.

Среди задач математического программирования наиболее изучены задачи линейного программирования (ЛП), в которых целевая функция E зависит линейно от элементов решения X_1, $X_2, \ldots, X_n$, а ограничения, накладываемые на элементы решения, имеют вид линейных равенств или неравенств относительно X_1, $X_2, \ldots, X_n$. Таким образом, в задачах ЛП находится экстремум линейной целевой функции при линейных ограничениях.

Термин "линейное программирование" появился впервые в 1951 г. в работах американских ученых Дж. Данцига и Т. Купманса, а первые исследования по ЛП были проведены в конце 30-х годов в Ленинградском университете Л.В. Канторовичем (основные задачи и приложения, критерий оптимальности, экономическая интерпретация, методы решения, геометрическая интерпретация результатов решения).

К истории ЛП имеет отношение Ж. Фурье, который в 20-х годах 19 столетия предложил способ решения систем линейных неравенств и описал в геометрической форме метод решения оптимизационной задачи ЛП [24, с. 3]. Через столетие этот метод получил свое развитие и был назван симплекс-методом. Дж. Данциг выполнил алгебраизацию симплекс-метода и внедрил его в практику. Оформление симплекс-метода в самостоятельную дисциплину стимулировалось успехами в области операционных исследований и появлением ЭВМ, так как с их помощью стало возможным решать задачи большой размерности.

"В настоящее время ЛП с точки зрения уровня теоретических разработок, сферы приложений и реализации вычислительных методов является одним из наиболее развитых направлений в области решения оптимизационных задач. Достижения в области ЛП содействовали прогрессу в разработке алгоритмов решения других задач математического программирования, в том числе в области целочисленного, стохастического и нелинейного программирования" (Таха Х., см. [25, с. 81-82]).

Модели, методы и вычислительные процедуры ЛП стали уже традиционными при решении задач планирования производства, управленческих задач, в практике создания и развития высокоэффективных технических систем и технологических процессов.

Учебное пособие включает материал по моделям, методам и алгоритмам ЛП.

В разделе 1 изучается основная вычислительная процедура для решения задач оптимизации симплекс-методом, рассматривается технология анализа модели на чувствительность на основе симплекс-таблицы для оптимального решения.

В разделе 2 изучаются методы искусственного базиса, позволяющие найти начальное допустимое базисное решение при наличии ограничений-равенств и "не меньше" (двухэтапный метод, метод больших штрафов).

В разделе 3 изучаются методы и алгоритмы для решения задач линейного целочисленного программирования - метод ветвей и границ (комбинаторный метод) и метод Р. Гомори (метод отсечений) для полностью или частично целочисленных задач.

В разделе 4 рассматриваются некоторые особые классы задач ЛП - задачи транспортного типа и задача о назначениях. Транспортные задачи решаются по двухэтапной схеме - определение начального опорного плана (метод северо-западного угла, метод минимального элемента, метод Фогеля) и определение оптимального плана в результате итерационного улучшения опорного плана (метод потенциалов). Задача о назначениях в ее классической постановке решается методом Мака.

Рекомендуемая схема решения оптимизационных задач ЛП: 1) постановка задачи оптимизации; 2) построение аналитической модели; 3) обоснование вычислительной процедуры; 4) решение задачи на основе симплекс-таблиц; 5) анализ модели на чувствительность; 6) определение оптимального целочисленного решения (в случае необходимости); 7) интерпретация результатов оптимизации.

Рекомендуемая студентам литература включает источники по методологии системного анализа, исследования операций и экспертных оценок [1-38].

Сущность задач оптимизации: определить значения переменных $X_1, X_2, \ldots, X_n$, которые обеспечивают экстремум целевой функции E с учетом ограничений, наложенных на аргументы функции $X_1, X_2, \ldots, X_n$ (элементы решения). При этом сложность решения задач зависит: 1) от вида функциональной зависимости, т.е. от связи функции E с элементами решения; 2) от размерности задачи, т.е. от количества элементов решения; 3) от вида и количества ограничений, наложенных на элементы решения.

Среди задач математического программирования наиболее изучены задачи линейного программирования (ЛП), в которых целевая функция E зависит линейно от элементов решения X_1, $X_2, \ldots, X_n$, а ограничения, накладываемые на элементы решения, имеют вид линейных равенств или неравенств относительно X_1, $X_2, \ldots, X_n$. Таким образом, в задачах ЛП находится экстремум линейной целевой функции при линейных ограничениях.

Термин "линейное программирование" появился впервые в 1951 г. в работах американских ученых Дж. Данцига и Т. Купманса, а первые исследования по ЛП были проведены в конце 30-х годов в Ленинградском университете Л.В. Канторовичем (основные задачи и приложения, критерий оптимальности, экономическая интерпретация, методы решения, геометрическая интерпретация результатов решения).

К истории ЛП имеет отношение Ж. Фурье, который в 20-х годах 19 столетия предложил способ решения систем линейных неравенств и описал в геометрической форме метод решения оптимизационной задачи ЛП [24, с.3]. Через столетие этот метод получил свое развитие и был назван симплекс-методом. Дж. Данциг выполнил алгебраизацию симплекс-метода и внедрил его в практику. Оформление симплекс-метода в самостоятельную дисциплину стимулировалось успехами в области операционных исследований и появлением ЭВМ, так как с их помощью стало возможным решать задачи большой размерности.

"В настоящее время ЛП с точки зрения уровня теоретических разработок, сферы приложений и реализации вычислительных методов является одним из наиболее развитых направлений в области решения оптимизационных задач. Достижения в области ЛП содействовали прогрессу в разработке алгоритмов решения других задач математического программирования, в том числе в области целочисленного, стохастического и нелинейного программирования" (Таха X., см. [25, с. 81-82]).

Модели, методы и вычислительные процедуры ЛП стали уже традиционными при решении задач планирования производства, управленческих задач, в практике создания и развития высокоэффективных технических систем и технологических процессов.

Учебное пособие включает материал по моделям, методам и алгоритмам ЛП.

В разделе 1 изучается основная вычислительная процедура для решения задач оптимизации симплекс-методом, рассматривается технология анализа модели на чувствительность на основе симплекс-таблицы для оптимального решения.

В разделе 2 изучаются методы искусственного базиса, позволяющие найти начальное допустимое базисное решение при наличии ограничений-равенств и "не меньше" (двухэтапный метод, метод больших штрафов).

В разделе 3 изучаются методы и алгоритмы для решения задач линейного целочисленного программирования - метод ветвей и границ (комбинаторный метод) и метод Р. Гомори (метод отсечений) для полностью или частично целочисленных задач.

В разделе 4 рассматриваются некоторые особые классы задач ЛП - задачи транспортного типа и задача о назначениях. Транспортные задачи решаются по двухэтапной схеме - определение начального опорного плана (метод северо-западного угла, метод минимального элемента, метод Фогеля) и определение оптимального плана в результате итерационного улучшения опорного плана (метод потенциалов). Задача о назначениях в ее классической постановке решается методом Мака.

Рекомендуемая схема решения оптимизационных задач ЛП: 1) постановка задачи оптимизации; 2) построение аналитической модели; 3) обоснование вычислительной процедуры; 4) решение задачи на основе симплекс-таблиц; 5) анализ модели на чувствительность; 6) определение оптимального целочисленного решения (в случае необходимости); 7) интерпретация результатов оптимизации.

Рекомендуемая студентам литература включает источники по методологии системного анализа, исследования операций и экспертных оценок [1-38].

1. ВЫЧИСЛИТЕЛЬНЫЕ ПРОЦЕДУРЫ ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

1.1. Постановка основной задачи линейного программирования

Математическое программирование рассматривает задачи оптимизации. Сущность таких задач можно выразить следующим образом: определить значения переменных $X_1, X_2, \ldots, X_n$, которые обеспечивают экстремум функции E $(X_1, X_2, \ldots, X_n)$ с учетом ограничений, наложенных на аргументы этой функции - переменные $X_1, X_2, \ldots, X_n$ (элементы решения). Функция E, для которой определяется экстремум, называется целевой функцией.

Сложность решения таких задач зависит от ряда факторов: от вида функциональной зависимости, т. е. от связи функции E с элементами решения; от размерности задачи, т. е. от количества элементов решения; от вида и количества ограничений, наложенных на элементы решения, и т. д.

Среди задач математического программирования наиболее изучены задачи линейного программирования, в которых целевая функция E линейно зависит от элементов решения $X_1, X_2, \ldots, X_n$, а ограничения, накладываемые на элементы решения, имеют вид линейных равенств или неравенств относительно X_1, $X_2, \ldots, X_n$. Таким образом, в задачах линейного программирования находится экстремум линейной целевой функции при линейных ограничениях. Задача линейного программирования в общей форме формулируется следующим образом: найти значения переменных X_1, $X_2, \ldots,$ X_n, удовлетворяющих ограничениям в виде равенств или неравенств:

$$
\begin{array}{llllll}
A_{11}*X_1 & + A_{12}*X_2 & + \ldots & + A_{1n}*X_n & >= B_1 & \\
\multicolumn{4}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & & \\
A_{p1}*X_1 & + A_{p2}*X_2 & + \ldots & + A_{pn}*X_n & >= B_p & \\
A_{(p+1)1}*X_1 & + A_{(p+1)2}*X_2 & + \ldots & + A_{(p+1)n}*X_n & = B_{p+1} & \\
\multicolumn{4}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & & (1.1) \\
A_{(p+q)1}*X_1 & + A_{(p+q)2}*X_2 & + \ldots & + A_{(p+q)n}*X_n & = B_{p+q} & \\
A_{(p+q+1)1}*X_1 & + A_{(p+q+1)1}*X_2 & + \ldots & + A_{(p+q+1)1}*X_n & <= B_{p+q+1} & \\
\multicolumn{4}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & & \\
A_{m1}*X_1 & + A_{m2}*X_2 & + \ldots & + A_{mn}*X_n & <= B_m &
\end{array}
$$

и обращающих в максимум или минимум линейную функцию этих переменных:

$$E = C_1 * X_1 + C_2 * X_2 + \ldots + C_n * X_n \to max/min. \qquad (1.2)$$

Здесь p - количество ограничений вида "не меньше", q - количество ограничений-равенств, r - количество ограничений вида "не больше", p+q+r=m, где m- общее количество ограничений. n - количество переменных.

Обычно на некоторые переменные, используемые в задаче, накладывается требование неотрицательности: $X_j \geq 0$, $j=1,\ldots,n_1$, $n_1 <= n$. В большинстве задач требуется неотрицательность всех переменных, т.е. $n_1 = n$.

Формулировка задачи (1.1),(1.2) представляет собой математическую модель задачи.

Допустимое решение задачи линейного программирования - это набор значений переменных $X_1, X_2, \ldots, X_n$, удовлетворяющих условиям (1.1). Множество всех допустимых решений называется областью допустимых решений (ОДР). Оптимальное решение - это то из допустимых решений, которое максимизирует (или минимизирует) функцию (1.2). Обычно задача линейного программирования имеет бесконечное множество допустимых и единственное оптимальное решение.

Решение задачи линейного программирования состоит в нахождении оптимального решения.

1.2. Пример задачи линейного программирования: задача распределения ресурсов

Задача распределения ресурсов является одной из типовых задач линейного программирования. Она формулируется следующим образом.

Имеется m видов ресурсов (под ресурсами понимается сырье, энергия, рабочее время, потребительский спрос и т.д.). Известны запасы каждого ресурса: B_i, $i=1,\ldots,m$. Ресурсы можно использовать для выпуска n видов продукции. Известна прибыль от выпуска единицы каждого вида продукции: C_j, $j=1,\ldots,n$. Известны расходы каждого ресурса на единицу каждого вида продукции: A_{ij}, $i=1,\dots,m$, $j=1,\dots,n$; здесь A_{ij} – расход i-го ресурса на выпуск единицы j-го вида продукции (возможно, что некоторые коэффициенты A_{ij} равны 0, если i-й ресурс для выпуска j-й продукции не требуется). Требуется составить оптимальный план производства, т.е. найти, сколько продукции каждого вида требуется выпускать, чтобы получить максимальную прибыль (при ограничениях на ресурсы).

Сформулировать эту задачу как задачу линейного программирования (построить математическую модель задачи) можно следующим образом:

$$\begin{aligned} &A_{11}*X_1 + A_{12}*X_2 + \dots + A_{1n}*X_n <= B_1 \\ &A_{21}*X_1 + A_{22}*X_2 + \dots + A_{2n}*X_n <= B_2 \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &A_{m1}*X_1 + A_{m2}*X_2 + \dots + A_{mn}*X_n <= B_m \\ &X_j >= 0, \quad j=1,\dots,n, \end{aligned} \tag{1.3}$$

$$E = C_1 * X_1 + C_2 * X_2 + \dots + C_n * X_n \;{-}{>}\; \max. \tag{1.4}$$

Здесь (1.3) – ограничения на ресурсы, (1.4) – целевая функция (прибыль). Через X_j обозначены подлежащие определению объемы выпуска каждого вида продукции; $j=1,\dots,n$. Очевидно, что в такой задаче все переменные по своему смыслу неотрицательны.

1.3. Графический метод решения задач линейного программирования

Графический метод может использоваться для задач, в исходной постановке которых имеется только 2 переменные, а также для задач произвольной размерности, в которых все переменные могут быть выражены через какие-либо две переменные. Для таких задач возможно графическое представление ограничений, ОДР и целевой функции. Таким образом, задача, для которой возможно решение графическим методом, имеет следующую формулировку:

$$
\begin{aligned}
&A_{11}*X_1 &&+ A_{12}*X_2 >= B_1 \\
&\ldots\ldots\ldots\ldots \\
&A_{p1}*X_1 &&+ A_{p2}*X_2 >= B_p \\
&A_{(p+1)1}*X_1 &&+ A_{(p+1)2}*X_2 <= B_{p+1} \\
&\ldots\ldots\ldots\ldots \\
&A_{m1}*X_1 &&+ A_{m2}*X_2 <= B_m
\end{aligned}
\qquad (1.5)
$$

$$E = C_1 * X_1 + C_2 * X_2 \to max/min. \qquad (1.6)$$

Здесь p - количество ограничений вида "не меньше", r - количество ограничений вида "не больше", p+r=m, где m- общее количество ограничений. Количество переменных равно двум. Ограничений-равенств в таких задачах, как правило, нет.

Решение задачи графическим методом выполняется в следующем порядке.

1. Строится система координат (X_1 - абсцисса, X_2 - ордината).

2. В системе координат строится графическое представление области допустимых решений. С этой целью для каждого ограничения из (1.5) выполняются следующие действия.

2.1. Одна из переменных выражается через другую. Например, если есть ограничение

$$A_{i1}*X_1 + A_{i2}*X_2 <= B_i,$$

то можно выразить X_2 следующим образом:

$$X_2 <= - \frac{A_{i1}}{A_{i2}}*X_1 + \frac{B_i}{A_{i2}}.$$

2.2. Строится прямая, задаваемая уравнением

$$X_2 = - \frac{A_{i1}}{A_{i2}}*X_1 + \frac{B_i}{A_{i2}}.$$

Эта прямая разбивает координатную плоскость на две полуплоскости, одна из которых соответствует ограничению.

2.3. Определяется полуплоскость, соответствующая ограничению (в зависимости от вида конкретного ограничения).

Операции 2.1 - 2.3 повторяются для всех m ограничений. Пересечение полуплоскостей, соответствующих каждому из ограничений, представляет собой ОДР. В большинстве задач - это выпуклый многоугольник (кроме некоторых особых случаев; см. ниже).

3. Строится вектор с координатами (C_1,C_2), где C_1, C_2 - коэффициенты целевой функции (1.6). Можно показать, что он представляет собой перпендикуляр к линиям уровня целевой функции. Линия уровня - это набор точек (значений переменных X_1 и X_2), при подстановке которых в целевую функцию она принимает некоторое постоянное значение. Перпендикулярно вектору (C_1,C_2) строится линия уровня целевой функции, проходящая через начало координат (т.е. соответствующая значению E=0).

4. Линия уровня целевой функции перемещается в направлении перпендикуляра; это соответствует возрастанию целевой функции. Если целевая функция (1.6) направлена на минимум, то перемещение выполняется до первой точки касания линии уровня с ОДР, если на максимум - то через всю ОДР до последней точки касания.

Найденная точка соответствует оптимальному решению. Это всегда одна из угловых точек ОДР. Значения переменных X_1 и X_2 в этой точке представляют собой оптимальное решение задачи. Значение целевой функции определяется подстановкой X_1 и X_2 в (1.6).

Приведем пример решения задачи графическим методом. Пусть некоторая задача сформулирована следующим образом:

$$10*X_1 + 5*X_2 >= 55$$
$$X_1 + X_2 >= 8$$
$$X_1, X_2 >= 0$$
$$2*X_1 + 5*X_2 \to \min.$$

Решение задачи приведено на рис. 1.1. Задача решается в следующем порядке.

1. Строится система координат (X_1,X_2).

2. Строятся прямые, заданные уравнениями $X_2 = -2*X_1+11$ (из 1-го ограничения) и $X_2 = -X_1+8$ (из 2-го). Каждая из них разбивает плоскость (X_1, X_2) на две полуплоскости; ограничениям соответствуют верхние полуплоскости. Пересечение этих полуплоскостей (с учетом неотрицательности обеих переменных) представляет собой ОДР; на рис. 1.1 она выделена штриховкой.

3. Строится вектор с координатами (2;5). Перпендикулярно ему строится линия уровня целевой функции $2*X_1+5*X_2=0$.

4. Линия уровня целевой функции перемещается в направлении вектора (2;5) до первого касания с ОДР (так как целевая функция в данной задаче подлежит минимизации). Эта точка имеет координаты (8;0).

Таким образом, оптимальное решение задачи: $X_1=8$, $X_2=0$. Значение целевой функции: $E = 2*8 + 5*0 = 16$.

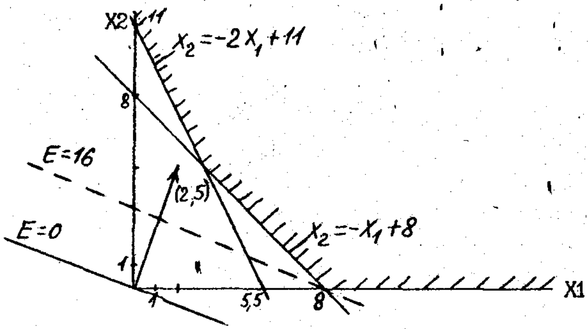


Рис. 1.1. Пример решения задачи линейного программирования графическим методом

При решении задачи графическим методом возможны следующие особые случаи.

1. Область допустимых решений не существует, т.е. полуплоскости, найденные на шаге 2, не пересекаются (рис.1.2,а).

Это значит, что задача не имеет решений, т.е. система ограничений противоречива.

2. Область допустимых решений неограничена сверху (для задач на максимизацию) или снизу (для задач на минимизацию) (рис. 1.2,б). Обычно это указывает на ошибку в постановке задачи, например, на то, что некоторое ограничение не учтено.

3. Линия уровня целевой функции на шаге 4 совпадает с одной из границ ОДР (рис. 1.2, в). Это значит, что задача имеет бесконечно много оптимальных решений; в качестве такого решения можно взять любую точку соответствующей границы ОДР.

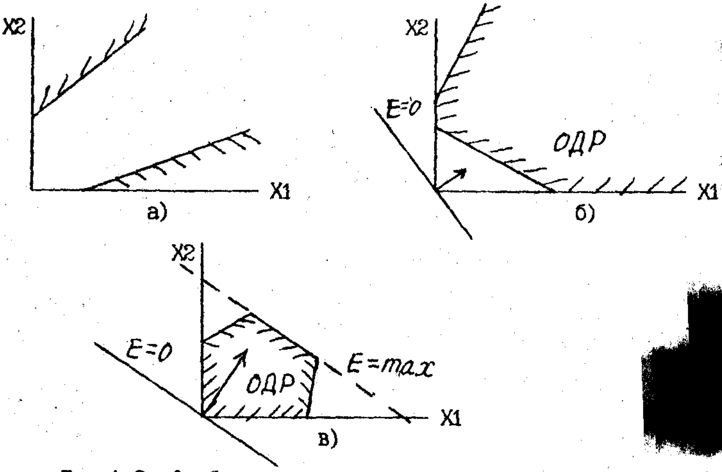


Рис. 1.2. Особые случаи решения:

а) отсутствие решения (ОДР не существует);

б) неограниченное решение (в задаче на максимизацию);

в) бесконечное множество решений

1.4. Приведение задач линейного программирования к стандартной форме

Любая задача линейного программирования приводится к стандартной (канонической) форме основной задачи линейного программирования (ОЗЛП), которая формулируется следующим образом: найти неотрицательные значения переменных X_1, X_2,..., X_n, удовлетворяющих ограничениям в виде равенств

$$\begin{array}{l} A_{11}*X_1 + A_{12}*X_2 + \ldots + A_{1n}*X_n = B_1 \\ A_{21}*X_1 + A_{22}*X_2 + \ldots + A_{2n}*X_n = B_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_{m1}*X_1 + A_{m2}*X_2 + \ldots + A_{mn}*X_n = B_m \end{array} \qquad (1.7)$$

$X_j >= 0, \quad j=1,\ldots,n,$

и обращающих в максимум линейную функцию этих переменных

$$E = C_1 * X_1 + C_2 * X_2 + \dots + C_n * X_n \rightarrow max. \qquad (1.8)$$

При этом также требуется, чтобы правые части равенств были неотрицательны, т.е. должно соблюдаться условие: $B_i >= 0$, $i = 1, \dots, m$. При сведении задачи, заданной в общей (произвольной) форме (1.1),(1.2), к стандартной форме (1.7),(1.8), количество переменных, как правило, изменяется, поэтому значения n для (1.1),(1.2) и (1.7),(1.8) обычно разные.

Приведение к стандартной форме необходимо, так как большинство методов решения задач линейного программирования разработано именно для стандартной формы.

Для приведения к стандартной форме произвольной задачи линейного программирования может потребоваться выполнить следующие действия:

- перейти от задачи на минимизацию целевой функции к задаче на ее максимизацию;
- изменить знаки правых частей ограничений;
- перейти от ограничений-неравенств к равенствам;
- избавиться от переменных, не имеющих ограничений на знак.

Для перехода от задачи на минимизацию целевой функции к задаче на ее максимизацию следует умножить целевую функцию на -1, так как минимальное значение функции E равно максимальному значению -E, взятому с обратным знаком: min E = = -max (-E).

Для изменения знака правой части ограничения также следует умножить это ограничение на -1.

Для перехода от ограничений-неравенств к равенствам вводятся дополнительные (балансовые) неотрицательные переменные. От ограничения в виде неравенства "не больше"

$$A_{11}*X_1 + A_{12}*X_2 + \dots + A_{1n}*X_n <= B_1$$

можно перейти к ограничению в виде равенства

$$A_{11}*X_1 + A_{12}*X_2 + \dots + A_{1n}*X_n + X_{n+1} = B_1.$$

Аналогично от ограничения в виде неравенства "не меньше"

$$A_{11}*X_1 + A_{12}*X_2 + \dots + A_{1n}*X_n >= B_1$$

можно перейти к ограничению в виде равенства

$$A_{11}*X_1 + A_{12}*X_2 + \dots + A_{1n}*X_n - X_{n+1} = B_1.$$

Переменные, используемые для перехода к равенствам от неравенств "не больше", называются остаточными, а для неравенств "не меньше" - избыточными.

Переход от переменных, не ограниченных в знаке, к неотрицательным переменным выполняется следующим образом. Пусть переменная X_j может быть как положительной, так и отрицательной. Ее можно заменить на разность двух неотрицательных переменных, т.е. в выражениях (1.1),(1.2) заменить X_j на $X_{j1}-X_{j2}$, где $X_{j1}>=0$, $X_{j2}>=0$. Так как других органичений на значения X_{j1} и X_{j2} не накладывается, их разность может быть как положительной, так и отрицательной. В практических задачах переменные, неограниченные в знаке, встречаются сравнительно редко.

Рассмотрим приведение к стандартной форме задачи о распределении ресурсов (1.3), (1.4). Для этого в каждое ограничение включается остаточная переменная:

$$\begin{array}{l}A_{11}*X_1 + A_{12}*X_2 + \dots + A_{1n}*X_n + X_{n+1} = B_1\\ A_{21}*X_1 + A_{22}*X_2 + \dots + A_{2n}*X_n + X_{n+2} = B_2\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ A_{m1}*X_1 + A_{m2}*X_2 + \dots + A_{mn}*X_n + X_{n+m} = B_m\end{array} \qquad (1.9)$$

$$X_j>=0, \qquad j=1,\dots,n+m$$

$$E = C_1 * X_1 + C_2 * X_2 + \dots + C_n * X_n \to max. \qquad (1.10)$$

Как видно, здесь для приведения задачи к стандартной форме потребовалось ввести m остаточных переменных: X_{n+1}, $X_{n+2},\dots,X_{n+m}$.

1.5. Решение задач линейного программирования симплекс-методом

1.5.1. Основная идея симплекс-метода

Симплекс-метод, называемый также методом последовательного улучшения плана, предназначен для решения задач линейного программирования любой размерности.

В теории линейного программирования доказывается, что экстремум целевой функции всегда достигается в угловых точках области допустимых решений (для задач с двумя переменными ОДР представляет собой многоугольник на плоскости, для трех переменных - многогранник, для многомерных задач ОДР также многомерна). Поэтому в принципе задача может решаться следующим образом: определяются все угловые точки ОДР, вычисляются значения целевой функции во всех этих точках и выбирается точка с экстремальным значением целевой функции. Однако для практических целей такой метод нецелесообразен в силу его трудоемкости.

Симплекс-метод реализует перебор угловых точек ОДР в направлении улучшения значения целевой функции. Основная идея этого метода следующая.

Прежде всего находится какое-либо допустимое начальное (опорное) решение, т.е. какая-либо угловая точка ОДР. Процедура метода позволяет ответить на вопрос, является ли это решение оптимальным. Если "да", то задача решена. Если "нет", то выполняется переход к смежной угловой точке ОДР, где значение целевой функции улучшается, т.е. к нехудшему допустимому решению. Под "смежной" здесь понимается угловая точка, расположенная на той же границе ОДР, что и прежняя точка (для двумерной ОДР - на той же стороне многоугольника, для трехмерной - на том же ребре многогранника, и т.д.). Если некоторая угловая точка имеет несколько смежных, то вычислительная процедура метода обеспечивает переход к той из них, для которой улучшение целевой функции будет наибольшим. Процесс перебора угловых точек ОДР повторяется, пока не будет найдена точка, которой соответствует экстремум целевой функции.

1.5.2. Определение начального допустимого базисного решения

Для задачи, представленной в стандартной форме, количество переменных обычно больше, чем количество ограничений. Поэтому для нахождения начального решения задачи (соответствующего начальной угловой точке ОДР) требуется выразить m переменных (т.е. количество переменных, равное количеству уравнений) через остальные $n-m$ переменных, принять эти $n-m$ переменных равными нулю и, таким образом, найти значения m переменных. Переменные, значения которых принимаются равными нулю, называются небазисными, а остальные m переменных-базисными. Значения базисных переменных неотрицательны (некоторые из них могут оказаться равными нулю). Количество базисных переменных всегда равно количеству ограничений. Найденное таким образом решение называется начальным допустимым базисным решением. Оно соответствует всем ограничениям (т.е. является допустимым).

Проще всего найти начальное решение в случае, когда в каждом ограничении есть переменная, которая входит в него с коэффициентом 1 и при этом отсутствует в других ограничениях. Такие переменные принимаются в качестве базисных (образуют начальный базис задачи). Остальные (небазисные) переменные принимаются равными нулю. Таким образом, базисные переменные принимают значения, равные правым частям ограничений.

Таким образом, для построения начального базиса задачи требуется, чтобы в каждое ограничение задачи (1.7),(1.8) входила переменная, имеющая коэффициент 1 и не входящая в другие уравнения (базисная переменная). Иногда такие переменные появляются в каждом ограничении в ходе приведения задачи к стандартной форме. В случаях, когда в системе нет m базисных переменных, их создают с помощью специальных способов (это называется построением искусственного базиса). Искусственный базис обычно требуется строить, если в исходной системе ограничений (1.1) имеются ограничения - равенства и "не меньше". При использовании искусственного базиса начальное решение оказывается недопустимым; от него по определенным алгоритмам выполняется переход к начальному допустимому решению (см. раздел 2).

Например, для задачи распределения ресурсов (1.3), (1.4), приведенной к виду (1.9),(1.10), в качестве базисных переменных выбираются остаточные переменные X_{n+1}, $X_{n+2}, \ldots, X_{n+m}$. Остальные переменные (X_1, $X_2, \ldots, X_n$) считаются небазисными (принимаются равными нулю). Отсюда находятся значения базисных переменных (начальное решение): $X_{n+i}=B_i$, $i=1,\ldots,m$. Такое решение допустимо, но неоптимально (целевая функция равна нулю); по содержательному смыслу оно соответствует случаю, когда никакие изделия не выпускаются.

Начальное решение для задачи вида (1.3), (1.4) соответствует угловой точке ОДР, заданной ограничениями на неотрицательность переменных (т.е. ограничениями $X_j>=0$, $j=1,\ldots,n$). Таким образом, угловой точкой ОДР, соответствующей начальному решению, здесь является начало координат. Переход к смежной угловой точке ОДР соответствует замене одной переменной в базисе, т.е. удалению из базиса одной переменной и вводу в него другой. Процедура симплекс-метода обеспечивает изменение базиса, приводящее к максимальному росту целевой функции.

1.5.3. Алгоритм реализации симплекс-метода

Решение задач с помощью симплекс-метода выполняется с использованием симплекс-таблиц. Алгоритм решения при этом реализуется в следующем порядке.

1. Задача линейного программирования сводится к стандартной форме (1.7),(1.8).

2. Определяется начальное решение. При этом, если не во всех уравнениях стандартной формы задачи имеются базисные переменные, то для решения задачи следует применить один из методов построения искусственного базиса. В данном описании алгоритма предполагается, что стандартная форма задачи имеет вид (1.9),(1.10), т.е. в исходной постановке задачи использовались только ограничения "не больше".

3. Целевая функция записывается в виде уравнения

$$E - C_1*X_1 - C_2*X_2 - \ldots - C_n*X_n = 0.$$

2 4. По данным задачи составляется исходная симплекс-таблица:

Таблица 1.1

Базисные переменные	X_1	X_2	...	X_s	...	X_{n+m}	Решение
E	$-C_1$	$-C_2$	...	$-C_s$	...	$-C_{n+m}$	0
X_{n+1}	A_{11}	A_{12}	...	A_{1s}	...	$A_{1(n+m)}$	B_1
X_{n+2}	A_{21}	A_{22}	...	A_{2s}	...	$A_{2(n+m)}$	B_2
....							
X_r	A_{r1}	A_{r2}	...	A_{rs}	...	$A_{r(n+m)}$	B_r
....							
X_{n+m}	A_{m1}	A_{m2}	...	A_{ms}	...	$A_{m(n+m)}$	B_m

Как видно, в первом столбце таблицы указаны базисные переменные, в последнем столбце ("Решение") - их значения, а также значение целевой функции. В заголовке таблицы перечисляются все используемые переменные (включая остаточные, избыточные и искусственные, если они есть); в данном описании предполагается, что в задаче есть только исходные и остаточные переменные. В строках таблицы указываются коэффициенты ограничений задачи.

Здесь $C_{n+1}=0,\dots,C_{n+m}=0$ (так как остаточные переменные не присутствуют в целевой функции). В столбцах, соответствующих базисным переменным, содержится только по одной единице (для X_{n+i} - в i-й строке), остальные - нули, т.е. $A_{ij}=0$, $j>n$, $j\neq n+i$.

5. Если все коэффициенты в E-строке неотрицательны, это означает, что оптимальное решение найдено, и алгоритм завершается. Иначе - выполняется следующий шаг.

6. Из числа текущих небазисных переменных выбирается переменная, включаемая в новый базис (это включение должно приводить к улучшению значения целевой функции). В качестве такой переменной выбирается переменная, которой соответствует максимальный по модулю отрицательный коэффициент в E-строке. Такой способ выбора объясняется тем, что отрицательный коэффициент в E-строке соответствует положительному

коэффициенту самой целевой функции (так как эти коэффициенты записываются в E-строке с обратным знаком); значит, увеличивая соответствующую переменную, можно увеличить значение E. Выбор максимального по модулю отрицательного элемента означает включение в базис переменной, увеличение которой приводит к максимальному росту целевой функции.

Пусть для включения в базис выбрана переменная X_s. Столбец с номером s будем называть ведущим (разрешающим).

7. Выбирается переменная, исключаемая из базиса. Для этого вычисляются отношения элементов текущего решения к элементам ведущего столбца: B_i/A_{is} (эти отношения называются симплексными отношениями). Симплексные отношения вычисляются только для положительных элементов разрешающего столбца ($A_{is}>0$, $i=1,\dots,m$). Переменная, которой соответствует минимальное симплексное отношение, исключается из базиса. Строку, соответствующую исключаемой переменной (обозначим ее номером r), называют ведущей (разрешающей) строкой, а элемент на пересечении ведущей строки и столбца (A_{rs}) – ведущим (разрешающим) элементом.

по минимальному симплекс-ному отношению

Такой способ выбора переменной для исключения из базиса имеет следующее обоснование. При включении в базис новой переменной ее значение (которое раньше было нулевым) увеличивается. Чтобы по-прежнему соблюдались все ограничения (1.9), в каждом из ограничений должна уменьшаться соответствующая базисная переменная (небазисные переменные не могут уменьшаться, так как они равны нулю, а на все переменные налагается требование неотрицательности). Увеличение значения переменной X_s возможно до тех пор, пока одна из базисных переменных не станет равной нулю. Симплексное отношение позволяет определить, какая из базисных переменных первой становится нулевой (т.е. выводится из базиса).

8. Выполняется преобразование симплекс-таблицы по следующим правилам:

- все элементы ведущей строки делятся на ведущий элемент

$$A'_{rj} = A_{rj}/A_{rs}, \quad B'_r = B_r/A_{rs}; \tag{1.11}$$

- остальные элементы таблицы преобразуются по формулам

ведущий столбец, кро... та
и нулями

все остальные эл-ты пересчит.
по прав. пр...

$$A'_{ij} = \frac{A_{ij} * A_{rs} - A_{rj} * A_{is}}{A_{rs}},$$

$$C'_j = \frac{C_j * A_{rs} - A_{rj} * C_s}{A_{rs}}, \quad E'_0 = \frac{E_0 * A_{rs} - B_r * C_s}{A_{rs}}, \qquad (1.12)$$

$$B'_i = \frac{B_i * A_{rs} - B_r * A_{is}}{A_{rs}}.$$

Здесь E_0 - значение целевой функции (в исходной симплекс-таблице E_0=0).

Как видно из формул, элементы E-строки и столбца решений подвергаются тем же преобразованиям, что и остальные элементы симплекс-таблицы. В результате находится новое решение (т.е. значения новых базисных переменных) и новое значение целевой функции. Элемент A_{rs} оказывается равным 1, а все элементы ведущего столбца (кроме A_{rs}) - равными 0. Таким образом, смысл формул (1.11),(1.12) в том, что в результате их применения новая базисная переменная присутствует только в одном ограничении с коэффициентом 1, а в других - отсутствует. Преобразование E-строки соответствует выражению целевой функции через новые небазисные переменные.

По результатам преобразований строится новая симплекс-таблица:

Таблица 1.2

Базисные переменные	X_1	X_2	...	X_s	...	X_{n+m}	Решение
E	C'_1	C'_2	...	0	...	C'_{n+m}	E'_0
X'_{i_1}	A'_{11}	A'_{12}	...	0	...	$A'_{1(n+m)}$	B'_1
X_{i_2}	A'_{21}	A'_{22}	...	0	...	$A'_{2(n+m)}$	B'_2
....							
X_s	A'_{r1}	A'_{r2}	...	1	...	$A'_{r(n+m)}$	B'_r
....							
X_{i_m}	A'_{m1}	A'_{m2}	...	0	...	$A'_{m(n+m)}$	B'_m

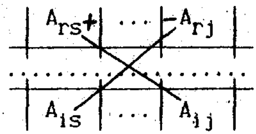
Здесь $X_{i_1}, X_{i_2}, \ldots X_s, \ldots, X_{i_m}$ - переменные, составляющие базис на данном цикле решения (т.е. некоторые из переменных $X_1, X_2, \ldots, X_{n+m}$; индексы $i_1, i_2, \ldots, i_m$ принимают значения из диапазона 1,2,...,n+m). На каждом цикле в базисе заменяется одна переменная.

9. Выполняется возврат к шагу 3

По окончании реализации алгоритма в столбце "Решение" находятся значения переменных, вошедших в оптимальный базис, а также значение целевой функции, соответствующее оптимальному решению. Переменные, не вошедшие в оптимальный базис, в оптимальном решении равны нулю.

Примечания.

1. При пересчете симплекс-таблицы целесообразно пользоваться следующим правилом (правило прямоугольника):



т.е. пересчитываемый и ведущий элемент образуют прямоугольник; из произведения пересчитываемого и ~~разрешающего~~ элемента вычитается произведение элементов, расположенных на другой диагонали этого прямоугольника; результат делится на ведущий элемент.

2. Если на шаге 7 (выбор переменной, исключаемой из базиса) оказывается, что все элементы ведущего столбца отрицательны или равны нулю (A_{is}<=0, i=1,...,m), это означает, что значение новой базисной переменной, а значит, и значение целевой функции можно увеличивать на сколь угодно большую величину, не нарушая ни одного из ограничений задачи. Такой случай называется неограниченным решением. В реальных задачах это обычно указывает на ошибку в постановке задачи, например, на то, что некоторое ограничение не учтено.

1.6. Основные задачи анализа оптимального решения на чувствительность

В окончательной симплекс-таблице, содержащей оптимальное решение, содержится не только само оптимальное решение, но и другая информация. На основе последней симплекс-таблицы решаются задачи анализа на чувствительность - определение влияния изменений в исходных данных задачи на оптимальное решение. Интерпретация симплекс-таблицы и анализ на чувствительность полностью зависят от содержательного смысла конкретной задачи.

Ниже рассматриваются вопросы интерпретации симплекс-таблицы и анализа на чувствительность для задачи о распределении ресурсов (1.3),(1.4). Предполагается, что задача решается на основе стандартной формы (1.9),(1.10).

1.6.1. Статус ресурсов

По статусу ресурсы делятся на дефицитные и недефицитные. Если некоторый ресурс при реализации оптимального плана расходуется полностью, он называется дефицитным, если не полностью - недефицитным.

Статус ресурсов определяется по значениям остаточных переменных $X_{n+1}, X_{n+2}, \ldots, X_{n+m}$, введенных в систему ограничений, для приведения ее к стандартной форме. Из системы (1.9) видно, что эти переменные обозначают остатки ресурсов при реализации оптимального плана. Если некоторая остаточная переменная X_{n+i} входит в оптимальное решение, по содержательному смыслу это означает, что соответствующий (i-й) ресурс при реализации оптимального плана расходуется не полностью, т.е. является недефицитным. Значение переменной X_{n+i} равно остатку соответствующего ресурса. Если остаточная переменная не входит в оптимальный базис (т.е. равна нулю), это означает, что остаток ресурса равен нулю; ресурс расходуется полностью, т.е. он дефицитен.

Увеличение запасов дефицитных ресурсов позволяет увеличить значение целевой функции, а снижение этих запасов приводит к уменьшению целевой функции. Увеличение запасов недефицитных ресурсов всегда нецелесообразно (оно не влияет на значение целевой функции). Запас недефицитного ресурса можно уменьшить на величину его остатка (т.е. значения соответствующей переменной X_{n+i}) без каких-либо изменений оптимального плана.

1.6.2. Ценность ресурсов

Ценность ресурса - это величина увеличения значения целевой функции при увеличении запасов данного ресурса на единицу (или соответственно величина уменьшения целевой функции при снижении запаса ресурса). Другие названия этой величины - теневая (скрытая) цена. В симплекс-таблице, соответствующей оптимальному решению, теневые цены содержатся в E-строке и представляют собой коэффициенты при остаточных переменных, соответствующим остаткам ресурсов. Таким образом, если переменная X_{n+i} обозначает остаток i-го ресурса, то ценность этого ресурса равна коэффициенту E-строки при переменной X_{n+i}.

Остаточные переменные, соответствующие недефицитным ресурсам, входят в базис оптимального решения, поэтому коэффициенты при них в E-строке таблицы оптимального решения равны нулю. Это значит, что ценность этих ресурсов равна нулю (увеличение их запаса не увеличивает прибыль).

Хотя ценность ресурса имеет ту же размерность, что и коэффициенты целевой функции (обычно - размерность денежной единицы), это не то же самое, что цена, по которой производится закупка ресурса.

1.6.3. Анализ на чувствительность к изменениям запасов ресурсов

Изменение запасов ресурсов соответствует изменению правых частей ограничений в постановке задачи.

Пусть в результате решения задачи (1.3),(1.4) получен оптимальный базис, состоящий из переменных $Y_1, Y_2, \ldots, Y_m$ (это могут быть любые из переменных $X_1, X_2, \ldots, X_{n+m}$). Пусть базисные переменные приняли значения соответственно $R_1, R_2, \ldots, R_m$, а целевая функция - значение Z.

Предположим, что в постановке задачи (1.3),(1.4) запас некоторого (i-го) дефицитного ресурса изменен на величину d. Эта величина (вариация запаса ресурса) может быть как положительной (если запас ресурса увеличивается), так и отрицательной (запас уменьшается). Если изменение ресурса не выходит за допустимый диапазон (определение этого диапазона будет рассмотрено ниже), то для измененной задачи оптимальные значения базисных переменных (т.е. оптимальное решение) следующие:

$$Y_k = R_k + A_{k(n+i)}*d, \qquad k=1,\ldots,m, \qquad (1.13)$$

где R_k, $k=1,\ldots m$, - значения базисных переменных в оптимальном решении исходной задачи (без изменения запаса ресурса);

$A_{k(n+i)}$ - коэффициенты столбца остаточной переменной X_{n+i} в окончательной симплекс-таблице для исходной задачи.

Значение целевой функции для оптимального решения измененной задачи следующее:

$$E = Z + F_{n+i}*d, \qquad (1.14)$$

где Z - значение целевой функции для оптимального решения исходной задачи;

F_{n+i} - коэффициент E-строки при остаточной переменной X_{n+i} в окончательной симплекс-таблице (ценность i-го ресурса).

По содержательному смыслу формулы (1.13),(1.14) означают изменение объемов производства, остатков недефицитных ресурсов и величины прибыли при изменении запаса дефицитного ресурса.

Таким образом, для исследования влияния изменения запаса ресурса на оптимальное решение нет необходимости решать задачу заново (с новым ограничением). Для нахождения оптимального решения достаточно по окончательной симплекс-таблице исходной задачи составить уравнения (1.13), (1.14) и подставить в них величину изменения запаса ресурса (значение d).

Изменение запасов ресурсов (т.е. правых частей ограничений) может привести к недопустимости оптимального базиса, найденного для исходной задачи. Так как на все переменные, используемые в задаче, накладывается требование неотрицательности, допустимый диапазон изменения запаса i-го ресурса (т.е. диапазон допустимых значений d) находят из системы неравенств:

$$\forall Y_k = R_k + A_{k(n+i)}*d > 0 \qquad k=1,\ldots,m. \qquad (1.15)$$

Все элементы этой системы неравенств имеют тот же смысл, что и в (1.13).

Пусть для вариации i-го ресурса найден диапазон (D_1, D_2), при котором все неравенства (1.15) выполняются. Это означает, что при любом изменении запаса i-го ресурса на величину d, такую, что $D_1 < d < D_2$, состав переменных в оптимальном базисе остается таким же, что и без такого изменения. Оптимальные значения базисных переменных и целевой функции при таких изменениях находятся по формулам (1.13),(1.14).

Выход значения d из диапазона, заданного системой (1.15), приводит к недопустимости базиса, найденного для исходной задачи, так как одна из базисных переменных оказывается отрицательной. При этом для решения сохраняется признак оптимальности (так как коэффициенты Е-строки не изменяются, т.е. остаются неотрицательными). В этом случае необходимо искать новый базис, т.е. решать задачу заново с самого начала или использовать специальные методы перехода к допустимому решению (например, двойственный симплекс-метод, рассматриваемый в разделе 1.7).

Приведем обоснование рассмотренного метода анализа оптимального решения на чувствительность к изменениям правых частей ограничений. Пусть запас i-го ресурса изменен на величину d. Правая часть i-го ограничения в (1.3) будет иметь вид B_i+d; другие ограничения не изменятся (т.е. их можно представить как B_k+0*d, $k=1,\dots,m, k\neq i$). Введем в симплекс-таблицу дополнительный столбец для коэффициентов при d. Начальная симплекс-таблица будет иметь следующий вид:

Таблица 1.3

Базисные переменные	X_1	X_2	...	X_{n+m}	Решение	Коэффициенты при d
E	$-C_1$	$-C_2$	...	$-C_{n+m}$	0	0
X_{n+1}	A_{11}	A_{12}	...	$A_{1(n+m)}$	B_1	0
X_{n+2}	A_{21}	A_{22}	...	$A_{2(n+m)}$	B_2	0
....						
X_{n+i}	A_{i1}	A_{i2}	...	$A_{i(n+m)}$	B_i	1
....						
X_{n+m}	A_{m1}	A_{m2}	...	$A_{m(n+m)}$	B_m	0

Таким образом, столбец коэффициентов при вариации d в начальной симплекс-таблице будет содержать одну единицу, остальные - нули. Такой же вид будет иметь столбец, соответствующий остаточной переменной (X_{n+i}), вводимой в ограничение на i-й ресурс и входящей в начальный базис. Так как элементы столбца решений (а значит, и элементы дополнительного столбца) никогда не используются в качестве ведущих, элементы дополнительного столбца подвергаются тем же преобразованиям, что и столбец переменной X_{n+i}. Поэтому в окончательной симплекс-таблице (содержащей оптимальное решение) столбец коэффициентов при d и столбец переменной X_{n+i} также будут одинаковыми. Из этого видно, что оптимальные значения переменных и целевой функции при изменении правой части i-го ограничения на величину d могут быть найдены по формулам (1.13),(1.14). Этот же вывод является обоснованием интерпретации коэффициентов E-строки при переменных $X_{n+1}, X_{n+2}, \ldots, X_{n+m}$ как ценностей ресурсов.

Если некоторый (i-й) ресурс является недефицитным, т.е. расходуется не полностью, то увеличение его запаса не имеет смысла. Запас ресурса может быть уменьшен на величину переменной X_{n+i} в оптимальном решении (т.е. на величину остатка этого ресурса). Это не приводит к каким-либо изменениям в оптимальном решении (изменяется только значение самой переменной X_{n+i}). Дальнейшее уменьшение запаса i-го ресурса делает решение недопустимым, так как переменная X_{n+i} принимает отрицательное значение; в этом случае необходимо искать новый базис.

1.6.4. Анализ на чувствительность к изменениям коэффициентов целевой функции

Изменение коэффициента целевой функции по содержательному смыслу задачи может отражать, например, изменение цены выпускаемого изделия.

Можно показать, что если изменение коэффициента целевой функции не выходит за некоторый диапазон (определение этого диапазона будет рассмотрено ниже), то для измененной задачи состав оптимального базиса и значения переменных в этом базисе остаются такими же, что и для исходной задачи (т.е. задачи без изменения коэффициента целевой функции), а изменяется только строка целевой функции.

Пусть в результате решения задачи (1.3),(1.4) получен оптимальный базис, состоящий из переменных $Y_1, Y_2, \ldots, Y_m$ (это могут быть любые из переменных $X_1, X_2, \ldots, X_{n+m}$). Остальные переменные являются небазисными (т.е. равны нулю); обозначим их как $W_1, W_2, \ldots, W_n$. Пусть в окончательной симплекс-таблице коэффициенты при небазисных переменных приняли значения F_1, $F_2, \ldots, F_n$; коэффициенты при базисных переменных равны нулю. Значение целевой функции обозначим как Z.

Предположим, что в постановке задачи (1.3), (1.4) коэффициент целевой функции C_j при некоторой базисной переменной X_j изменен на величину d. Эта вариация коэффициента целевой функции может быть как положительной, так и отрицательной. Значения коэффициентов E-строки для окончательной симплекс-таблицы измененной задачи можно найти следующим образом:

$$F'_k = F_k \pm A_{qk} * d, \qquad (1.16)$$

где k - номера небазисных переменных в оптимальном решении исходной задачи;

F_k - коэффициенты E-строки при небазисных переменных в окончательной симплекс-таблице исходной задачи;

q - номер строки, соответствующей переменной X_j в окончательной симплекс-таблице исходной задачи;

A_{qk} -коэффициенты строки переменной X_j в окончательной симплекс-таблице исходной задачи.

Значение целевой функции для оптимального решения измененной задачи следующее:

$$E = Z \pm R_j * d, \qquad (1.17)$$

где Z - значение целевой функции для оптимального решения исходной задачи;

R_j - значение переменной X_j в оптимальном решении исходной задачи

По содержательному смыслу формулы (1.16),(1.17) могут означать, например, изменение ценностей ресурсов и величины прибыли в результате изменения цены на выпускаемое изделие.

Таким образом, для исследования влияния изменения коэффициента целевой функции на результаты решения задачи нет необходимости решать ее заново. Значения переменных в оптимальном решении остаются без изменений. Для нахождения новых ценностей ресурсов и значения целевой функции достаточно по окончательной симплекс-таблице исходной задачи составить уравнения (1.16), (1.17) и подставить в них величину изменения коэффициента целевой функции (т.е. значение d).

Изменение коэффициента целевой функции может привести к неоптимальности базиса, найденного для исходной задачи. Признаком оптимального решения является неотрицательность всех коэффициентов Е-строки. Поэтому диапазон изменений коэффициента целевой функции при переменной X_j, не приводящих к неоптимальности найденного базиса (т.е. диапазон допустимых значений d), находят из системы неравенств

$$F'_k = F_k + A_{qk} * d > 0. \qquad (1.18)$$

Все элементы этой системы неравенств имеют тот же смысл, что и в (1.16).

Пусть для коэффициента C_j найден диапазон (D_1, D_2), при котором все неравенства (1.18) выполняются. Это означает, что при любом изменении коэффициента целевой функции при X_j на величину d, такую, что $D_1 < d < D_2$, состав переменных в оптимальном базисе и их значения остаются такими же, что и без такого изменения. Значения элементов Е-строки при таких изменениях находятся по формулам (1.16),(1.17).

Выход значения d из диапазона, заданного системой (1.18), приводит к неоптимальности базиса, найденного для исходной задачи, так как один из коэффициентов Е-строки оказывается отрицательным. Решение при этом остается допустимым, так как значения базисных переменных не изменяются. Для поиска оптимального решения в этом случае необходимо определять новый базис, т.е. продолжать решение задачи обычным симплекс-методом.

1.7. Двойственный симплекс-метод

Двойственный симплекс-метод обычно применяется для поиска нового оптимального решения в тех случаях, когда к задаче, для которой уже имеется оптимальное решение, добавляется ограничение, которое делает это решение недопустимым. Это требуется при решении следующих задач.

1. Анализ оптимального решения на чувствительность к добавлению нового органичения (т.е. нахождение нового решения в случае, когда в формулировку задачи, для которой уже найдено оптимальное решение, добавляется новое ограничение).

2. Поиск целочисленного решения (см. раздел 3).

Двойственный симплекс-метод может также применяться для решения некоторых обычных задач линейного программирования.

Приведем основные отличия двойственного симплекс-метода от обычного.

При решении задачи обычным симплекс-методом в качестве начального используется некоторое допустимое, но не оптимальное решение. Затем выполняется переход от исходного к оптимальному решению (последовательное улучшение плана). При этом все промежуточные решения являются допустимыми. При каждом переходе к новому решению значение целевой функции улучшается.

В отличие от обычного, двойственный симплекс-метод обеспечивает переход от недопустимого решения, для которого в то же время выполняются признаки оптимальности, к допустимому (и дейсствительно оптимальному) решению. Исходное решение для двойственного симплекс-метода иногда называют "лучшим, чем оптимальное". При реализации двойственного симплекс-метода все промежуточные решения являются недопустимыми, но признаки оптимальности для них сохраняются. При каждом переходе к новому решению значение целевой функции ухудшается.

Рассмотрим использование двойственного симплекс-метода для решения задачи поиска нового оптимального решения при добавлении нового ограничения, делающего имеющееся оптимальное решение недопустимым. Пусть в оптимальном решении некоторой задачи переменные $Y_1, Y_2, \ldots, Y_m$ являются базисными и имеют значения R_1, R_2,..., R_m, а остальные переменные W_1, W_2,..., W_n - небазисные (нулевые). Пусть к задаче добавляется ограничение, которому не соответствуют какие-либо из значений $R_1, R_2, \ldots, R_m$ (т.е. оптимальное решение становится недопустимым). Двойственный симплекс-метод для поиска нового оптимального решения реализуется в следующем порядке.

1. Новое ограничение выражается через небазисные переменные $W_1, W_2, \ldots, W_n$.

2. Новое ограничение приводится к стандартной форме (т.е. к равенству). Для этого в него добавляется новая переменная (остаточная или избыточная). Эта переменная становится базисной для введенного ограничения.

3. К симплекс-таблице, содержащей оптимальное решение, добавляется строка, соответствующая новому ограничению, и столбец базисной переменной для этого ограничения. Если вводится несколько новых ограничений, то к симплекс-таблице добавляется столько же строк и столбцов.

4. Если все элементы столбца решений неотрицательны (все переменные в задаче имеют неотрицательные значе-ния), это означает, что допустимое (и одновременно - оптимальное) решение найдено; в этом случае алгоритм завершается.

В противном случае решение, содержащееся в симплекс-таблице, недопустимо. Выбирается переменная, исключаемая из базиса. Для исключения из базиса выбирается переменная с максимальным по модулю отрицательным значением, т.е. переменная, которой соответствует максимальный по модулю отрицательный элемент стоблца решений. Обозначим переменную, выбранную для исключения из базиса, как X_r; соответствующую ей строку будем называть ведущей (разрешающей).

5. Выбирается переменная, включаемая в базис. Для этого вычисляются отношения коэффициентов строки целевой функции к элементам ведущей строки: C_j/A_{rj}. Эти отношения вычисляются только для отрицательных элементов ведущей строки ($A_{rj}<0$). Переменная, которой соответствует минимальное по модулю отношение, включается в базис. Столбец, соответствующий включаемой переменной (обозначим ее как X_s), называют ведущим (разрешающим) столбцом, а элемент на пересечении ведущей строки и столбца (A_{rs}) - ведущим (разрешающим) элементом.

6. Выполняется пересчет симплекс-таблицы по тем же правилам, что и в обычном симплекс-методе.

7. Выполняется возврат к шагу 4.

Двойственный симплекс-метод может также применяться для решения обычных задач линейного программирования. Его применение возможно, если для задачи удается найти начальное недопустимое решение, для которого выполняются условия оптимальности. Обычно такое начальное решение можно найти для задач, у которых в системе ограничений нет ограничений-равенств, а целевая функция подлежит минимизации.

Пусть для некоторой задачи построена следующая математическая модель:

$$
\begin{array}{llll}
A_{11}*X_1 & + A_{12}*X_2 + \ldots & + A_{1n}*X_n & >= B_1 \\
A_{21}*X_1 & + A_{22}*X_2 + \ldots & + A_{2n}*X_n & >= B_2 \\
\ldots\ldots\ldots\ldots & & & \\
A_{p1}*X_1 & + A_{p2}*X_2 & + \ldots + A_{pn}*X_n & >= B_p \\
A_{(p+1)1}*X_1 & + A_{(p+1)2}*X_2 & + \ldots + A_{(p+1)n}*X_n & <= B_{p+1} \\
A_{(p+2)1}*X_1 & + A_{(p+2)2}*X_2 & + \ldots + A_{(p+2)n}*X_n & <= B_{p+2} \\
\ldots\ldots\ldots\ldots & & & \\
A_{m1}*X_1 & + A_{m2}*X_2 & + \ldots + A_{mn}*X_n & <= B_m
\end{array}
$$

$X_j>=0, \ i=1,\ldots,n.$

$$E = C_1 * X_1 + C_2 * X_2 + \ldots + C_n * X_n \ -> \min.$$

Пусть все коэффициенты целевой функции положительны: $C_j>=0, \ j=1,\ldots,n$.

Задача имеет m ограничений, из них p ограничений вида "не меньше", m-p ограничений "не больше". Для приведения такой задачи к стандартной форме требуется ввести p избыточных и m-p остаточных переменных, а также перейти к целевой функции, направленной на максимум. Стандартная форма задачи имеет следующий вид:

$$
\begin{array}{llllll}
A_{11}*X_1 & +\ A_{12}*X_2 & +\ \dots\ +\ A_{1n}*X_n & -\ X_{n+1} & = B_1 \\
A_{21}*X_1 & +\ A_{22}*X_2 & +\ \dots\ +\ A_{2n}*X_n & -\ X_{n+2} & = B_2 \\
\dots\dots & & & & \\
A_{p1}*X_1 & +\ A_{p2}*X_2 & +\ \dots\ +\ A_{pn}*X_n & -\ X_{n+p} & = B_p \\
A_{(p+1)1}*X_1 & +\ A_{(p+1)2}*X_2 & +\ \dots\ +\ A_{(p+1)n}*X_n & +\ X_{n+p+1} & = B_{p+1} \\
A_{(p+2)1}*X_1 & +\ A_{(p+2)2}*X_2 & +\ \dots\ +\ A_{(p+2)n}*X_n & +\ X_{n+p+2} & = B_{p+2} \\
\dots\dots & & & & \\
A_{m1}*X_1 & +\ A_{m2}*X_2 & +\ \dots\ +\ A_{mn}*X_n & +\ X_{n+m} & = B_m
\end{array}
$$

$X_j>=0,\ i=1,\dots,n+m$

$$-E = -C_1 * X_1 - C_2 * X_2 - \dots - C_n * X_n \ -> \max.$$

Составим начальный базис из избыточных и остаточных переменных X_{n+i}, $i=1,\dots,m$, а все переменные исходной задачи будем считать небазисными: $X_j=0$, $j=1,\dots,n$. Целевую функцию представим в виде уравнения

$$-E + C_1*X_1 + C_2*X_2 + \dots + C_n*X_n = 0.$$

Построим симплекс-таблицу для данного решения:

Таблица 1.4

Базис	X_1	X_2	..	X_{n+1}	..	X_{n+p}	X_{n+p+1}	..	X_{n+m}	Решение
$-E$	C_1	C_2	..	0	..	0	0	..	0	0
X_{n+1}	A_{11}	A_{12}	..	-1	..	0	0	..	0	$-B_1$
X_{n+2}	A_{21}	A_{22}	..	0	..	0	0	..	0	$-B_2$
....										
X_{n+p}	A_{p1}	A_{p2}	..	0	..	-1	0	..	0	$-B_p$
X_{n+p+1}	$A_{(p+1)1}$	$A_{(p+1)2}$	..	0	..	0	1	..	0	B_{p+1}
X_{n+p+2}	$A_{(p+2)1}$	$A_{(p+2)2}$	..	0	..	0	0	..	0	B_{p+2}
....										
X_{n+m}	A_{m1}	A_{m2}	..	0	..	0	0	..	1	B_m

Как видно, в полученном базисе положительными являются только остаточные переменные ($X_{n+i}=B_i$, $i=p+1,\dots,m$). Все избыточные переменные принимают отрицательные значения ($X_{n+i}=-B_i$, $i=1,\dots,p$); значит, данное решение недопустимо. При этом в строке целевой функции симплекс-таблицы все коэффициенты неотрицательны, т.е. решение соответствует условию оптимальности. Данная задача может быть решена двойственным симплекс-методом.

2. МЕТОДЫ ИСКУССТВЕННОГО БАЗИСА В ЛИНЕЙНОМ ПРОГРАММИРОВАНИИ

Двухэтапный метод

2.1. Назначение методов искусственного базиса

Методы искусственного базиса предназначены для построения начального базиса (т.е. для получения начального решения) в случаях, когда его построение непосредственно на основе стандартной формы (1.7),(1.8) невозможно.

Как показано ранее (см. раздел 1.5.2), для построения начального базиса необходимо, чтобы в каждом уравнении стандартной формы присутствовала переменная, входящая в данное уравнение с коэффициентом 1 и не входящая в другие уравнения. Из этих переменных и составляется начальный базис (количество переменных в базисе равно количеству ограничений). Если базисные переменные имеются не во всех уравнениях, они создаются с помощью методов искусственного базиса. Это требуется практически всегда, когда в исходной постановке задачи имеются ограничения "равно" и/или "не меньше".

Использ двухэта метод

Пусть имеется задача линейного программирования в общей форме

$$
\begin{array}{lllll}
A_{11}*X_1 & +\ A_{12}*X_2 & +\ \dots\ +\ A_{1n}*X_n & >= B_1 & \\
\dots\dots & & & & \\
A_{p1}*X_1 & +\ A_{p2}*X_2 & +\ \dots\ +\ A_{pn}*X_n & >= B_p & \\
A_{(p+1)1}*X_1 & +\ A_{(p+1)2}*X_2 & +\ \dots\ +\ A_{(p+1)n}*X_n & = B_{p+1} & \\
\dots\dots & & & & (2.1) \\
A_{(p+q)1}*X_1 & +\ A_{(p+q)2}*X_2 & +\ \dots\ +\ A_{(p+q)n}*X_n & = B_{p+q} & \\
A_{(p+q+1)1}*X_1 & +\ A_{(p+q+1)1}*X_2 & +\ \dots\ +\ A_{(p+q+1)1}*X_n & <= B_{p+q+1} & \\
\dots\dots & & & & \\
A_{m1}*X_1 & +\ A_{m2}*X_2 & +\ \dots\ +\ A_{mn}*X_n & <= B_m &
\end{array}
$$

$$E = C_1 * X_1 + C_2 * X_2 + \dots + C_n * X_n \ -> \max/\min. \qquad (2.2)$$

Здесь p - количество ограничений вида "не меньше", q - количество ограничений-равенств, r - количество ограничений вида "не больше", p+q+r=m, где m- общее количество ограничений, n - количество переменных.

Будем считать, что для всех переменных задачи установлено ограничение неотрицательности (или уже выполнен переход к неотрицательным переменным), т.е. $X_j>=0$, $j=1,\dots,n$.

Приведем задачу к стандартной форме

$$
\begin{array}{lllll}
A_{11}*X_1 & +A_{12}*X_2 & +\dots+A_{1n}*X_n & -X_{n+1} & =B_1 \\
\dots\dots\dots & & & & \\
A_{p1}*X_1 & +A_{p2}*X_2 & +\dots+A_{pn}*X_n & -X_{n+p} & =B_p \\
A_{(p+1)1}*X_1 & +A_{(p+1)2}*X_2 & +\dots+A_{(p+1)n}*X_n & & =B_{p+1} \\
\dots\dots\dots & & & & \\
A_{(p+q)1}*X_1 & +A_{(p+q)2}*X_2 & +\dots+A_{(p+q)n}*X_n & & =B_{p+q} \\
A_{(p+q+1)1}*X_1 & +A_{(p+q+1)1}*X_2 & +\dots+A_{(p+q+1)1}*X_n & +X_{n+p+1} & =B_{p+q+1} \\
\dots\dots\dots & & & & \\
A_{m1}*X_1 & +A_{m2}*X_2 & +\dots+A_{mn}*X_n & +X_{n+p+r} & =B_m
\end{array}
\qquad (2.3)
$$

$$X_j>=0,\ j=1,\dots,n+p+r$$

$$E=C_1*X_1+C_2*X_2+\dots+C_n*X_n\ ->\ \max. \qquad (2.4)$$

Здесь целевая функция (2.4) совпадает с (2.2), если функция (2.2) направлена на максимум, или представляет собой результат умножения (2.2) на -1, если она направлена на минимум.

Как видно, для перехода к равенствам от неравенств "не меньше" введено p избыточных переменных, а для неравенств "не больше" - r остаточных переменных. Для ограничений-равенств (их количество равно q) дополнительные переменные не вводятся. В данном случае базисные переменные (т.е. входящие в данное уравнение с коэффициентом 1 и не входящие ни в одно из других уравнений) есть только в уравнениях, полученных из ограничений "не больше". Использовать в качестве базисных избыточные переменные невозможно, так как они входят в уравнения с коэффициентами -1; в случае придания переменным X_1, ..., X_n нулевых значений эти избыточные переменные принимают отрицательные значения, т.е. решение оказывается недопустимым. В уравнениях, полученных из ограничений-равенств, вообще нет переменных, входящих только в данное уравнение. Таким образом, вместо необходимых m=p+q+r базисных переменных имеется только r таких переменных. Один из способов разрешения этой проблемы - использование методов искусственного базиса.

Искусственный базис строится следующим образом. В каждое уравнение стандартной формы, не содержащее базисных переменных (т.е. полученное из ограничения-равенства или "не меньше"), добавляется по одной искусственной переменной.

Для упрощения записей обозначим через N общее количество переменных в задаче, представленной в стандартной форме $N=n+p+r$. Для построения искусственного базиса требуется ввести K искусственных переменных, $K=p+q$ (или, то же самое, $K=m-r$). Система ограничений с добавленными искусственными переменными представляется в следующем виде:

$$
\begin{array}{lllllll}
A_{11}*X_1 & +A_{12}*X_2 & +\dots+A_{1n}*X_n & -X_{n+1} & +X_{N+1} & =B_1 & \\
\dots\dots & & & & & & \\
A_{p1}*X_1 & +A_{p2}*X_2 & +\dots+A_{pn}*X_n & -X_{n+p} & +X_{N+p} & =B_p & \\
A_{(p+1)1}*X_1 & +A_{(p+1)2}*X_2 & +\dots+A_{(p+1)n}*X_n & & +X_{N+p+1} & =B_{p+1} & \\
\dots\dots & & & & & & (2.5) \\
A_{(p+q)1}*X_1 & +A_{(p+q)2}*X_2 & +\dots+A_{(p+q)n}*X_n & & +X_{N+K} & =B_{p+q} & \\
A_{(p+q+1)1}*X_1 & +A_{(p+q+1)2}*X_2 & +\dots+A_{(p+q+1)1}*X_n & +X_{n+p+1} & & =B_{p+q+1} & \\
\dots\dots & & & & & & \\
A_{m1}*X_1 & +A_{m2}*X_2 & +\dots+A_{mn}*X_n & +X_{n+p+r} & & =B_m &
\end{array}
$$

$$X_j>=0,\quad j=1,\dots,N+K.$$

На все искусственные переменные (как и на остальные) накладывается требование неотрицательности.

После построения искусственного базиса, придав нулевые значения исходным переменным $X_1,\dots,X_n$, а также избыточным переменным $X_{n+1},\dots,X_{n+p}$, получим начальный базис: $X_{N+i}=B_i$, $i=1,\dots,K$ (искусственные переменные); $X_{n+p+i}=B_{K+i}$, $i=1,\dots,r$ (остаточные переменные). Всего в таком базисе имеется m переменных; их начальные значения равны правым частям ограничений.

Искусственные переменные используются только для получения начального базиса. Они не имеют какого-либо содержательного смысла, связанного с условием задачи (т.е. их нельзя интерпретировать как запасы ресурсов, объемы производства и т.д.). Всякое решение, содержащее искусственные переменные, является недопустимым, т.е. не соответствует исходной системе ограничений (2.1). Поэтому методы решения, использующие искусственный базис, должны обеспечивать, чтобы в окончательном решении задачи все искусственные переменные получали нулевые значения.

Основные этапы реализации всех методов, использующих искусственный базис, следующие.

1. Задача, представленная в общей форме, приводится к стандартной форме.

2. В уравнения, не содержащие базисных переменных, добавляются искусственные неотрицательные переменные (т.е. создается искусственный базис). Таким образом, находится начальное недопустимое решение.

3. С помощью обычных процедур симплекс-метода находится решение, в котором все искусственные переменные равны нулю, т.е. начальное допустимое решение. Если такое решение найти не удается, это означает, что исходная задача не имеет допустимых решений.

~~4. Найденное~~ на шаге 3 допустимое решение используется в качестве начального решения исходной задачи при последующем поиске оптимального решения.

Искусственные переменные можно рассматривать как меру недопустимости исходного (искусственного) решения. Когда эти переменные оказываются равными нулю, мера недопустимости решения оказывается нулевой, т.е. соответствующее решение является допустимым.

Ниже рассматриваются два основных метода решения задач линейного программирования на основе искусственного базиса: двухэтапный метод и М-метод (метод больших штрафов).

2.2. Двухэтапный метод

Название этого метода связано с тем, что задача с помощью этого метода решается в два этапа: сначала отыскивается допустимое решение, не содержащее искусственных переменных, а затем на основе найденного решения ищется оптимальное решение исходной задачи. Основные шаги реализации метода следующие.

1. Задача линейного программирования сводится к стандартной форме.

2. Строится искусственный базис.

3. Составляется искусственная целевая функция: сумма всех искусственных переменных.

4. Реализуется первый этап двухэтапного метода: с помощью обычных процедур симплекс-метода выполняется минимизация искусственной целевой функции. Если ее минимальное значение равно 0, то соответствующее решение является допустимым решением исходной задачи. Очевидно, что при нулевом значении искусственной целевой функции все искусственные переменные также нулевые (так как искусственная целевая функция-их сумма, и все они неотрицательны).

Если минимальное значение искусственной целевой функции оказывается отличным от нуля, это означает, что задача не имеет допустимых решений.

5. Реализуется второй этап двухэтапного метода: найденное на шаге 4 допустимое решение используется в качестве начального решения исходной задачи для поиска ее оптимального решения.

Рассмотрим реализацию метода более подробно. Пусть в задаче, представленной в стандартной форме, имеется всего N переменных (включая остаточные и избыточные). Базис состоит из m переменных (где m - количество уравнений), из них K - искусственных: X_{N+1}, $X_{N+2}, \ldots, X_{N+K}$ ($K<=m$). Метод реализуется в следующем порядке.

1. Задача сводится к стандартной форме.

2. Выполняется переход к целевой функции, направленной на максимум (если такой переход требуется), и она записывается в виде уравнения

$$E - C_1 * X_1 - C_2 * X_2 - \ldots - C_N * X_N = 0. \qquad (2.6)$$

Так как среди переменных X_j, $j=1,\ldots N$ есть остаточные и избыточные, некоторые из коэффициентов C_j равны нулю.

3. Строится искусственный базис, т.е. вводятся искусственные переменные X_{N+1}, $X_{N+2}, \ldots, X_{N+K}$.

4. Составляется искусственная целевая функция - сумма искусственных переменных:

$$W = X_{N+1} + X_{N+2} + \ldots + X_{N+K} \to \min. \qquad (2.7)$$

5. Так как в вычислительных процедурах симплекс-метода целевая функция должна быть выражена только через небазисные переменные, выполняется преобразование функции (2.7). Для этого искусственные переменные, составляющие начальный базис, выражаются через небазисные переменные. Например, если $N+j$-я искусственная переменная входит в i-е уравнение

$$A_{i1}*X_1 + A_{i2}*X_2 + \ldots + A_{iN}*X_N + X_{N+j} = B_i$$

то ее можно выразить следующим образом:

$$X_{N+j} = B_i - A_{i1}*X_1 - A_{i2}*X_2 - \ldots - A_{iN}*X_N.$$

Таким образом представляются все искусственные переменные X_{N+j}, $j=1,\ldots K$. Они подставляются в функцию (2.7). В результате получим искусственную целевую функцию, выраженную через небазисные переменные:

$$W = W_1*X_1 + W_2*X_2 + \ldots + W_N*X_N + W_0 \to \min. \qquad (2.8)$$

Здесь коэффициенты W_0 и W_i, $i=1,\ldots,N$, могут быть любыми (положительными, отрицательными или нулевыми).

6. Для приведения всей задачи к стандартной форме выполняется переход к искусственной целевой функции, направленной на максимум:

$$-W = - W_1*X_1 - W_2*X_2 - \ldots - W_N*X_N - W_0 \to \max. \qquad (2.9)$$

Искусственная целевая функция (как и основная) представляется в виде уравнения

$$-W + W_1*X_1 + W_2*X_2 + \ldots + W_N*X_N = - W_0. \qquad (2.10)$$

7. Определяется начальное (недопустимое) решение. Базис состоит из m переменных, из них К - искусственных, остальные

m-K - остаточные переменные. Если остаточных переменных нет (т.е. в задаче не было ограничений "не больше"), то обычно весь базис составляется из искусственных переменных. Базисные переменные принимают значения, равные ограничениям задачи. Все остальные переменные (входившие в исходную постановку задачи, а также избыточные) считаются равными нулю. В этом случае исходная целевая функция (2.4) принимает значение 0, искусственная целевая функция (2.7) -значение W_0.

8. Составляется исходная симплекс-таблица:

Таблица 2.1

Базис	X_1	X_2	...	X_N	X_{N+1}	X_{N+2}	...	X_{N+K}	Решение
E	C_1	C_2	...	C_N	0	0	...	0	0
-W	W_1	W_2	...	W_N	0	0	...	0	$-W_0$
X_{i_1}	A_{11}	A_{12}	...	A_{1N}	$A_{1(N+1)}$	$A_{1(N+2)}$	...	$A_{1(N+K)}$	B_1
X_{i_2}	A_{21}	A_{22}	...	A_{2N}	$A_{2(N+1)}$	$A_{2(N+2)}$	...	$A_{2(N+K)}$	B_2
....									
X_{i_m}	A_{m1}	A_{m2}	...	A_{mN}	$A_{m(N+1)}$	$A_{m(N+2)}$	...	$A_{m(N+K)}$	B_m

Здесь $X_{i_1}, X_{i_2}, \dots, X_{i_m}$ - переменные, составившие начальный базис. Среди них K искусственных, остальные - остаточные, т.е. индексы $i_1, i_2, \dots, i_m$ принимают значения из диапазонов $n+p+1, n+p+2, \dots, n+p+r$ (для остаточных переменных) и $N+1, N+2, \dots, N+K$ (для искусственных).

В строке основной целевой функции (E) только первые n коэффициентов могут быть ненулевыми. Остальные коэффициенты соответствуют остаточным, избыточным и искусственным переменным; эти коэффициенты всегда нулевые, т.е. $C_j=0$, $j=n+1, \dots, N+K$.

В строке искусственной целевой функции (W) ненулевыми могут быть первые N коэффициентов, так как искусственные переменные выражены через исходные и избыточные: $W_j=0$, $j=N+1, \dots, N+K$. Коэффициенты при остаточных переменных также нулевые.

Во всех строках, соответствующих ограничениям, первые n коэффициентов (соответствующие исходным переменным) могут быть ненулевыми. Другие коэффициенты A_{ij}, где i - номер строки, $j=n+1,\ldots,N+K$, принимают следующие значения:

- если i - номер строки, соответствующей ограничению вида "не меньше", то коэффициент в столбце избыточной переменной данного ограничения равен -1, а коэффициент в столбце искусственной переменной данного ограничения равен 1;

- если i - номер строки, соответствующей ограничению вида "равно", то коэффициент в столбце искусственной переменной данного ограничения равен 1;

- если i - номер строки, соответствующей ограничениям вида "не больше", то коэффициент в столбце остаточной переменной данного ограничения равен 1;

- остальные коэффициенты - нулевые.

9. Реализуется первый этап двухэтапного метода: с помощью обычных процедур симплекс-метода выполняется минимизация функции W (или, то же самое, максимизация -W). При этом переменные, включаемые в базис, выбираются по W-строке (т.е. на каждом цикле в базис включается переменная, которой соответствует максимальный по модулю отрицательный коэффициент в W-строке; столбец, соответствующий этой переменной, становится ведущим). Все преобразования выполняются по обычным формулам симплекс-метода или по "правилу прямоугольника". Преобразованиям подвергается вся симплекс-таблица, включая Е-строку, W-строку и столбец решений.

10. Выполняется анализ результатов минимизации искусственной целевой функции. Если найденное минимальное значение W равно 0 и все искусственные переменные оказались небазисными (т.е. равными 0), это означает, что найдено допустимое решение исходной задачи, т.е. решение, удовлетворяющее исходным ограничениям; выполняется переход к шагу 11. Если в результате минимизации W какие-либо из искусственных переменных приняли ненулевые значения (т.е. вошли в базис), это означает, что исходная задача не имеет допустимых решений.

11. Пусть на шаге 10 найдено допустимое решение исходной задачи. Оно находится в последней симплекс-таблице, полученной при минимизации W:

Таблица 2.2

Базис	X_1	X_2	...	X_N	X_{N+1}	X_{N+2}	...	X_{N+K}	Решение
E	F_1	F_2	...	F_N	F_{N+1}	F_{N+2}	...	F_{N+K}	E_0
-W	V_1	V_2	...	V_N	V_{N+1}	V_{N+2}	...	V_{N+K}	0
X_{k_1}	D_{11}	D_{12}	...	D_{1N}	$D_{1(N+1)}$	$D_{1(N+2)}$	...	$D_{1(N+K)}$	R_1
X_{k_2}	D_{21}	D_{22}	...	D_{2N}	$D_{2(N+1)}$	$D_{2(N+2)}$	...	$D_{2(N+K)}$	R_2
....									
X_{k_m}	D_{m1}	D_{m2}	...	D_{mN}	$D_{m(N+1)}$	$D_{m(N+2)}$	...	$D_{m(N+K)}$	R_m

Здесь $X_{k_1}, X_{k_2}, \ldots, X_{k_m}$ - новые базисные переменные. Среди них могут быть переменные исходной задачи, избыточные и остаточные (индексы $k_1, k_2, \ldots, k_m$ принимают значения из диапазона $1, 2, \ldots, N$). R_1, $R_2, \ldots, R_m$ - значения базисных переменных. Коэффициенты исходной целевой функции F_j, $j=1,\ldots,N+K$, искусственной целевой функции V_j, $j=1,\ldots, N+K$, и коэффициенты ограничений D_{ij}, $i=1,\ldots,m$, $j=1,\ldots,N+K$, получены в результате преобразований симплекс-метода, выполненных над коэффициентами C_j, W_j и A_{ij} соответственно. Очевидно, что некоторые из коэффициентов в табл. 2.2 равны нулю (какие конкретно - зависит от того, какие переменные вошли в базис).

В новом базисе нет ни одной искусственной переменной; таким образом, найдено допустимое решение исходной задачи. Первый этап двухэтапного метода завершен.

Так как все искусственные переменные обратились в 0, они отбрасываются. Найденное решение используется в качестве начального допустимого решения исходной задачи. Для поиска оптимального решения исходной задачи строится сокращенная симплекс-таблица (путем отбрасывания из табл. 2.2 строки искусственной целевой функции и всех столбцов, соответствующим искусственным переменным X_{N+j}, $j=1,\ldots,K$):

Таблица 2.3

Базис	X_1	X_2	...	X_N	Решение
E	F_1	F_2	...	F_N	E_0
X_{k_1}	D_{11}	D_{12}	...	D_{1N}	R_1
X_{k_2}	D_{21}	D_{22}	...	D_{2N}	R_2
...	...	...			
X_{k_m}	D_{m1}	D_{m2}	...	D_{mN}	R_m

12. Реализуется второй этап двухэтапного метода: над построенной симплекс-таблицей выполняются обычные процедуры симплекс-метода (т.е. в E-строке находится максимальный по модулю отрицательный элемент, соответствующая переменная включается в базис, выбирается переменная для исключения из базиса, и т.д.). В результате находится оптимальное решение исходной задачи.

2.3. Метод больших штрафов

Метод больших штрафов основан на создании искусственного базиса (как и в двухэтапном методе) и введении штрафа за использование искусственных переменных. Рассмотрим порядок реализации метода.

1. Задача представляется в стандартной форме. Выполняется переход к целевой функции, направленной на максимум (если исходная целевая функция направлена на минимум). Пусть в задаче, представленной в стандартной форме, имеется всего N переменных (включая исходные, остаточные и избыточные).

2. Создается искусственный базис (так же, как и для двухэтапного метода). Пусть базис состоит из m переменных, из них K - искусственных (K<=m, где m - количество уравнений). Таким образом, общее количество переменных равно N+K.

3. Выполняется преобразование целевой функции: вводится штраф за использование искусственных переменных:

$$E=C_1*X_1+C_2*X_2+\ldots+C_N*X_N-M*X_{N+1}-M*X_{N+2}-\ldots-M*X_{N+K}\to\max, \quad (2.11)$$

где X_j, $j=1,\ldots,N$ - переменные, имеющиеся в стандартной форме задачи (исходные, остаточные и избыточные);

X_j, $j=N+1,\ldots,N+K$ - искусственные переменные;

M - некоторые бесконечно большие положительные коэффициенты.

В (2.11) коэффициенты C_j при остаточных и избыточных переменных равны нулю.

Для штрафных коэффициентов M не задается какого-либо конкретного значения; считается, что они больше любого конкретного числа.

Как видно, наличие искусственных переменных в целевой функции резко уменьшает ее значение, так как эти переменные умножаются на большие числа M (отсюда название - M-метод); в этом состоит штраф за использование искусственных переменных. Таким образом, максимальное (т.е. оптимальное) значение целевой функции может быть достигнуто только при нулевых значениях всех искусственных переменных.

4. Так как в вычислительных процедурах симплекс-метода целевая функция должна быть выражена только через небазисные переменные, из нее исключаются переменные, составляющие искусственный базис. Для этого искусственные переменные, входящие в начальный базис, выражаются через небазисные (исходные и избыточные) переменные аналогично тому, как это делается в двухэтапном методе. Полученные для искусственных переменных выражения подставляются в (2.11). В результате получается целевая функция, выраженная только через небазисные переменные:

$$E=(D_1+S_1*M)*X_1+(D_2+S_2*M)*X_2+\ldots+(D_N+S_N*M)*X_N+R*M\to\max, \quad (2.12)$$

где D_j, S_j, R - некоторые числа (среди них могут быть положительные, отрицательные и нулевые); $j=1,\ldots,N$.

5. Целевая функция записывается в виде уравнения

$$E-(D_1+S_1*M)*X_1-(D_2+S_2*M)*X_2-\ldots-(D_N+S_N*M)*X_N=R*M, \quad (2.13)$$

6. Определяется начальный базис, состоящий из искусственных и остаточных переменных (так же, как в двухэтапном методе). При этом целевая функция принимает значение R*M.

7. Составляется исходная симплекс-таблица:

Таблица 2.4

БП	X_1	X_2	..	X_N	X_{N+1}	X_{N+2}	..	X_{N+K}	Реш.
E	$-D_1-S_1*M$	$-D_2-S_2*M$	..	$-C_N-S_N*M$	0	0	..	0	R*M
X_{i_1}	A_{11}	A_{12}	..	A_{1N}	$A_{1(N+1)}$	$A_{1(N+2)}$	..	$A_{1(N+K)}$	B_1
X_{i_2}	A_{21}	A_{22}	..	A_{2N}	$A_{2(N+1)}$	$A_{2(N+2)}$	..	$A_{2(N+K)}$	B_2
....									
X_{i_m}	A_{m1}	A_{m2}	..	A_{mN}	$A_{m(N+1)}$	$A_{m(N+2)}$	..	$A_{m(N+K)}$	B_m

Здесь $X_{i_1}, X_{i_2}, \ldots, X_{i_m}$ - переменные, составляющие начальный базис, т.е. искусственные, а также остаточные (если есть).

Данная таблица ничем не отличается от табл. 2.1 (построенной для двухэтапного метода), кроме вида Е-строки, а также отсутствия искусственной целевой функции.

8. Задача решается с помощью обычных процедур симплекс-метода. При этом учитывается, что коэффициенты М, присутствующие в коэффициентах целевой функции,- это некоторые очень большие числа.

Так как использование искусственных переменных приводит к резкому уменьшению значения целевой функции, то в оптимальном решении (которому соответствует максимум целевой функции) эти переменные должны обратиться в 0, т.е. оказываются небазисными. Если в полученном оптимальном решении некоторые искусственные переменные остаются в базисе, это означает, что исходная задача не имеет допустимого решения.

Приведем примеры операций над симплекс-таблицей с участием штрафных коэффициентов М. Пусть имеется симплекс-таблица:

Базисные переменные	X_1	X_2	X_3	X_4	X_5	X_6	Решение
E	-100	-20-M	-1-5M	10-3M	0	0	-200*M
X_5	10	15	5	10	1	0	60
X_6	7	8	7	2	0	1	140

Для включения в базис здесь следует выбрать переменную X_3, так как ей соответствует максимальный по модулю отрицательный коэффициент в E-строке (с учетом M). Для исключения из базиса выбирается переменная X_5, так как ей соответствует минимальное симплексное отношение (60/5=12, 140/7=20). Таким образом, в качестве ведущего столбца выбирается столбец X_3, в качестве ведущей строки - строка X_5; элемент 5 на их пересечении - ведущий элемент. Затем выполняется преобразование симплекс-таблицы по обычным правилам. Например, коэффициент E-строки при X_4, равный 10-3M, пересчитывается по "правилу прямоугольника" следующим образом:

$$\frac{5 * (10-3M) - 10 * (-1-5*M)}{5} = 12 + 7*M.$$

Точно так же пересчитываются все другие элементы симплекс-таблицы.

2.4. Анализ оптимального решения на чувствительность

Задачи, для решения которых требуется использовать методы искусственного базиса, очень разнообразны по своему содержанию; их нельзя свести, например, к задачам распределения ресурсов. Из-за этого можно привести лишь несколько указаний по интерпретации результатов и анализу на чувствительность, которые могут применяться для всех задач.

1. Все используемые в задаче переменные, кроме искусственных, имеют содержательный смысл. Для переменных, входящих в исходную формулировку задачи, их содержательный смысл полностью зависит от постановки задачи. Для остаточных переменных интерпретация во многих случаях примерно такая же, как в задачах распределения ресурсов, т.е. значение остаточной переменной представляет собой неизрасходованный запас ресурса, неиспользованное время и т.д. Значения избыточных переменных интерпретируются как превышение ограничения "не меньше".

Пусть, например, имеется ограничение: "необходимо выпустить не менее 1000 деталей типа 1". Его можно сформулировать следующим образом:

$X_1 >= 1000$.

Для приведения к стандартному виду в ограничение вводится избыточная переменная, например:

$X_1 - X_5 = 1000$.

Пусть в оптимальном решении эти переменные имеют следующие значения: $X_1=1200$, $X_5=200$. В этом случае значение переменной X_5 - это количество деталей типа 1, выпущенных сверх минимально необходимого.

2. Коэффициенты из столбцов окончательной симплекс-таблицы, соответствующих избыточным переменным, можно использовать для анализа решения на чувствительность к изменениям правых частей ограничений.

Пусть в i-е ограничение, имеющее вид "не меньше", введена избыточная переменная X_{n+i} (будем считать, что в задаче n исходных переменных и m ограничений):

$$A_{i1}*X_1 + A_{i2}*X_2 + \ldots + A_{in}*X_n - X_{n+i} = B_i.$$

Пусть в результате решения задачи (2.1),(2.2) получен оптимальный базис, состоящий из переменных $Y_1, Y_2, \ldots, Y_m$ (это могут быть любые из переменных $X_1, X_2, \ldots, X_N$, т.е. любые переменные, кроме искусственных). Пусть базисные переменные приняли значения соответственно $R_1, R_2, \ldots, R_m$, а целевая функция - значение Z.

Предположим, что в постановке задачи (2.1),(2.2) правая часть i-го ограничения (т.е. величина B_i) изменена на величину d. Если это изменение не выходит за допустимый диапазон, то для измененной задачи оптимальные значения базисных переменных (т.е. оптимальное решение) следующие:

$$Y_k = R_k - A_{k(n+i)}*d, \qquad k=1,\ldots,m, \qquad (2.14)$$

где R_k, $k=1,\ldots m$, – значения базисных переменных в оптимальном решении исходной задачи ~~(без изменения запаса ресурса);~~

$A_{k(n+i)}$ – коэффициенты столбца избыточной переменной X_{n+i} в окончательной симплекс-таблице для исходной задачи.

Значение целевой функции для оптимального решения измененной задачи следующее:

$$E = Z - F_{n+i} * d, \qquad (2.15)$$

где Z – значение целевой функции для оптимального решения исходной задачи;

F_{n+i} – коэффициент E-строки при избыточной переменной X_{n+i} в окончательной симплекс-таблице.

Допустимый диапазон изменения величины B_i, т.е. диапазон изменений, не приводящих к недопустимости найденного оптимального базиса, находится из системы неравенств

$$Y_k = R_k - A_{k(n+i)} * d > 0 \qquad k=1,\ldots,m. \qquad (2.16)$$

Все элементы этой системы неравенств имеют тот же смысл, что и в (2.14).

Как видно, анализ на чувствительность к изменениям правых частей ограничений выполняется в основном так же, как и для задачи о распределении ресурсов. Единственное отличие от формул (1.13)-(1.15) состоит в изменении знака перед коэффициентами $A_{k(n+i)}$ и F_{n+i} на противоположный. Обоснование данного метода анализа практически не отличается от приведенного для задачи распределения ресурсов.

3. Коэффициенты при избыточных переменных в E-строке симплекс-таблицы, содержащей оптимальное решение, во многих задачах можно интерпретировать как величину ухудшения целевой функции при увеличении на единицу правой части ограничения, в которое была введена данная избыточная переменная. Можно сказать, что коэффициенты E-строки при избыточных переменных имеют смысл, противоположный ценности ресурсов.

Пусть, например, в стандартной форме некоторой задачи ограничение: "необходимо выпустить не менее 1000 деталей типа 1" сформулировано следующим образом: $X_1-X_5=1000$, где X_1 - количество деталей, которое требуется выпустить, X_5 - избыточная переменная (т.е. выпуск сверх минимально необходимого). Пусть в этой задаче целевая функция выражает прибыль от выпуска деталей. Если в окончательной симплекс-таблице переменная X_5 оказывается небазисной ($X_5=0$), а коэффициент при этой переменной в Е-строке оказывается равным некоторому числу F, это означает, что в случае увеличения задания на выпуск деталей на единицу (т.е. на одну деталь) значение целевой функции уменьшится на F денежных единиц.

4. Если целевая функция в задаче (2.1),(2.2) направлена на максимум, то анализ на чувствительность к изменению коэффициентов целевой функции выполняется точно так же, как и для задачи распределения ресурсов (см. формулы (1.16)-(1.18)). Если исходная целевая функция направлена на минимум (т.е. при приведении задачи к стандартной форме она умножалась на -1), то расчеты также выполняются по формулам (1.16)-(1.18), но знак перед A_{qk} и R_j изменяется на противоположный.

2.5. Сравнение методов искусственного базиса

Количество итераций, требуемое для решения задачи как двухэтапным методом, так и методом больших штрафов, в большинстве случаев одинаково.

Достоинством метода больших штрафов является то, что он в малой степени отличается от обычного симплекс-метода (т.е. от той его реализации, которая используется, когда искусственный базис не требуется). В то же время метод больших штрафов имеет по сравнению с двухэтапным методом большой недостаток, связанный с использованием коэффициентов M. Так как эти коэффициенты - очень большие числа, они обычно значительно превышают любые коэффициенты, используемые в исходной постановке задачи. В результате ошибок округления (особенно при вычислениях на ЭВМ) процесс решения может стать нечувствительным к исходным коэффициентам задачи. Двухэтапный метод не имеет этого недостатка.

3. МЕТОДЫ РЕШЕНИЯ ЗАДАЧ ЛИНЕЙНОГО ЦЕЛОЧИСЛЕННОГО ПРОГРАММИРОВАНИЯ

3.1. Назначение методов целочисленного программирования

Целочисленное программирование ориентировано на решение задач, в которых все или некоторые переменные должны принимать только целочисленные значения. Если условие целочисленности наложено на все переменные, имеющиеся в задаче (в том числе переменные, введенные для приведения задачи к стандартному виду, т.е. остаточные и избыточные), то она называется полностью целочисленной. Примеры полностью целочисленных задач: распределение ресурсов для производства деталей, если в качестве ресурсов используются заготовки; распределение самолетов по авиалиниям, по каждой из которых требуется перевезти определенное количество пассажиров. Задачи, в которых требование целочисленности относится только к некоторым переменным, называются частично целочисленными. Примеры частично целочисленных задач: распределение материала (например, металла, пластмассы) или времени для производства неделимых изделий разных типов.

В некоторых случаях при получении дробного значения переменной, которая по своему содержательному смыслу должна быть целочисленной, допустимо ее округление. Однако при этом необходима осторожность, так как округленное решение может оказаться недопустимым (не удовлетворять ограничениям).

Рассмотрим пример. Пусть имеется ограничение

$$5*X_1 + 2*X_2 <= 53,$$

где X_1, X_2 - целочисленные переменные (например, количество выпускаемых деталей). Предположим, что получены оптимальные значения этих переменных: $X_1 = 7,8$, $X_2 = 6,9$. Округлить переменные до 8 и 7 нельзя, так как это приведет к невыполнению ограничения. В то же время решение $X_1=7$, $X_2=6$ может оказаться неоптимальным.

В некоторых случаях округление решения невозможно в принципе. Например, в задачах с ограничениями в виде равенств округленное решение никогда не соответствует ограничениям. Округление решения невозможно также в задачах с булевыми (т.е. принимающими только значения, равные 0 или 1) переменными. Примеры задач с булевыми переменными - задачи, в которых требуется принять решение о реализации одного из нескольких проектов: в таких задачах переменная X_i принимает значение 1, если i-й проект должен быть реализован, и 0 - в противном случае.

Имеются специальные методы решения целочисленных задач. Все они основаны на получении решения задачи без учета требований целочисленности и последующем уточнении решения.

В некоторых (но достаточно редких) случаях оказывается, что задача имеет допустимые нецелочисленные решения, но не имеет целочисленных решений.

Методы решения задач целочисленного программирования можно разделить на две группы:

- комбинаторные методы;
- методы отсечений.

3.2. Комбинаторные методы. Метод ветвей и границ

Комбинаторные методы основаны на переборе допустимых решений задачи. Эти методы могут применяться во всех случаях, когда множество решений задачи является дискретным; таким образом, область применения этих методов не ограничивается решением задач целочисленного программирования. Простейшим (но наименее эффективным) из комбинаторных методов является метод полного перебора.

Наиболее распространенным из комбинаторных методов является метод ветвей и границ. Он представляет собой метод целенаправленного перебора, т.е. для поиска решения выбирается наиболее перспективное направление. Метод может применяться для решения как полностью, так и частично целочисленных задач. Метод ветвей и границ предусматривает решение задачи без учета требований целочисленности и последующий переход к целочисленному решению.

Приведем некоторые основные термины, используемые в описании метода ветвей и границ.

1. Список решаемых задач - список задач, которые требуется решить для нахождения оптимального решения. Сначала в этом списке содержится только исходная задача; затем в него добавляются задачи, полученные из исходной путем добавления ограничений, обеспечивающих целочисленность решения.

2. Оценка задачи - граница наилучшего (минимального или максимального - в зависимости от конкретной задачи) значения целевой функции, которое может быть достигнуто при решении данной задачи. Важно заметить, что оценка - это не то же самое, что значение целевой функции. Оценка всегда лучше, чем оптимальное значение целевой функции. Оценка задачи известна до ее решения.

3. Текущее наилучшее решение - лучшее из найденных целочисленных решений. Оно включает лучшие целочисленные значения переменных задачи и соответствующее им значение целевой функции. В начале решения задачи оно не существует (еще не найдено).

В процессе решения задачи может оказаться удобным графическое представление процесса решения в виде дерева. Корнем (начальной вершиной) такого дерева является исходная задача; другие вершины соответствуют задачам, порождаемым в процессе решения. Возле вершин могут записываться оценки задач и результаты их решения.

Метод реализуется в следующем порядке.

1. Исходная задача заносится в список решаемых задач. Ее оценка принимается бесконечно большой (здесь предполагается, что целевая функция решаемой задачи подлежит максимизации).

2. Из списка решаемых задач выбирается задача с наилучшим значением оценки, т.е. наиболее перспективная задача. Если есть несколько задач с одинаковыми оценками, то выбирается любая из них; если в списке только одна задача, то она выбирается для решения. Обозначим выбранную задачу как Т.

3. Выбранная задача исключается из списка решаемых задач и решается обычным симплекс-методом без учета требований целочисленности.

4. В зависимости от результата решения выполняются следующие действия.

4.1. Решение существует и является целочисленным

Если уже существует текущее наилучшее решение, то полученное решение сравнивается с ним; если найденное решение лучше текущего, то оно принимается в качестве нового текущего наилучшего решения. Если текущее наилучшее решение еще не существует, то в качестве такого решения принимается найденное решение. Новые задачи в список решаемых задач не добавляются.

4.2. Решение не существует (задача Т не имеет решений)

Новые задачи в список решаемых задач не добавляются, и текущее наилучшее решение не изменяется.

4.3. Решение существует, но является нецелочисленным, и при этом текущее наилучшее решение не существует

Выбирается переменная X_i, принявшая дробное значение R_i (если таких переменных несколько, то выбирается любая из них). Допустимое целое значение X_i должно удовлетворять одному из неравенств: $X_i <= [R_i]$ или $X_i >= [R_i]+1$. Здесь $[R_i]$ - целая часть числа R_i, т.е. максимальное целое число, не превосходящее R_i (например, $[5,7]=5$, $[-3,2]=-4$).

Вместо задачи Т в список решаемых задач вносятся две новые задачи: T_1 и T_2. Все ограничения и целевая функция обеих задач те же, что и у задачи Т; кроме того, в одну из новых задач добавляется ограничение $X_i <= [R_i]$, а в другую - $X_i >= [R_i]+1$. При этом оптимальное нецелочисленное решение задачи Т становится недопустимым; в то же время ни одно из допустимых целочисленных значений X_i не исключается новыми ограничениями. Вместе с каждой из двух новых задач сохраняется ее оценка: значение целевой функции задачи Т, из которой были получены эти новые задачи. Это значение соответствует оптимальному, но нецелочисленному решению задачи Т. Оно является предельной оценкой (границей) целевой функции для каждой из новых задач; очевидно, что оптимальное значение целевой функции у задач T_1 и T_2 не может быть лучше, чем у задачи Т, так как решение T_1 и T_2 должно соответствовать всем тем же ограничениям, что и решение задачи Т, а также дополнительному ограничению. Оценка используется для выбора наиболее перспективной задачи.

Текущее наилучшее решение не изменяется.

4.4. Решение существует, является нецелочисленным, и при этом уже существует текущее наилучшее решение.

Выполняется сравнение оптимального значения целевой функции задачи T (соответствующего найденному оптимальному, но нецелочисленному решению) со значением целевой функции текущего наилучшего решения. Если значение целевой функции текущего наилучшего решения хуже, чем в оптимальном решении задачи T, то выполняются действия, указанные в 4.3. В противном случае новые задачи в список решаемых задач не вносятся, так как, если бы они и были рассмотрены, их решение не может быть лучше, чем решение задачи T (и тем более не лучше, чем текущее наилучшее решение).

5. Если существует текущее наилучшее решение, то выполняется просмотр списка решаемых задач. Задачи, у которых оценки хуже текущего наилучшего решения, исключаются из списка решаемых задач, так как для них уже нельзя найти лучшего решения, чем текущее наилучшее.

6. Если список решаемых задач пуст, то алгоритм завершается. Если при этом существует текущее наилучшее решение, то оно является оптимальным целочисленным решением исходной задачи. Если текущего наилучшего решения нет, это значит, что исходная задача не имеет целочисленных решений. Если список решаемых задач не пуст, выполняется возврат к шагу 2.

Название метода связано с разветвлением решаемой задачи на две новые задачи (если решение оказывается нецелочисленным) и с использованием оценок граничных значений целевой функции при выборе задачи для решения.

3.3. Методы отсечений

3.3.1. Общая схема реализации методов отсечений

Эти методы основаны на введении в задачу дополнительных ограничений, позволяющих учесть требование целочисленности. Основные этапы реализации всех методов отсечений следующие.

1. Задача решается без учета требований целочисленности (например, обычным симплекс-методом). Если все переменные,

которые по содержательному смыслу должны быть целочисленными, принимают целые значения, то алгоритм завершается; иначе—выполняется следующий шаг.

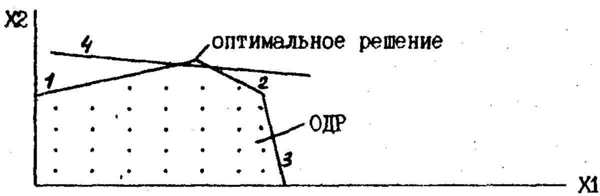
2. Вводится дополнительное ограничение, обладающее следующими свойствами.

1 Оно отсекает (т. е. делает недопустимым) найденное оптимальное, но нецелочисленное решение. Другими словами, вводится такое ограничение, которому не соответствует найденное решение.

2 Вводимое ограничение не должно отсекать ни одного из целочисленных решений, принадлежащих области допустимых решений исходной задачи.

3. Задача решается с учетом добавленного ограничения. Если решение оказывается целочисленным, то оно принимается в качестве оптимального, и алгоритм завершается; иначе - возврат к шагу 2.

На рис. 3.1 приведена графическая иллюстрация дополнительного ограничения, отсекающего нецелочисленное решение. Точками показаны допустимые целочисленные решения.



1,2,3 - границы ОДР (ограничения из постановки задачи); 4 дополнительное ограничение

Рис. 3.1. Дополнительное ограничение в методе отсечений

√ Из методов отсечений наиболее широко применяются методы Гомори. Другие названия этих методов - методы отсекающих плоскостей, дробные алгоритмы.

Примечание. В методах Гомори используются термины "целая часть" и "дробная часть". Под целой частью числа здесь понимается максимальное целое число, не превосходящее данного числа, а под дробной частью - разность самого числа и его целой части. Целая часть числа обозначается квадратными скобками, а дробная - фигурными. Например, если B=1,7, то [B]=1, {B}=B-[B]=0,7. Если B=-2,33, то [B]=-3, {B}=0,67. Таким образом, для дробной части всегда верно неравенство 0 <= {B} < 1.

3.3.2. Метод Гомори для полностью целочисленных задач

Для применения этого метода необходимо, чтобы все коэффициенты и правые части ограничений были целочисленными. Это связано с тем, что наличие дробных коэффициентов может привести к нарушению целочисленности дополнительных (остаточных и избыточных) переменных. Так как алгоритм ориентирован на решение полностью целочисленной задачи, нахождение целочисленного решения в этом случае может оказаться невозможным даже в случае, когда задача, записанная только в исходных переменных, в действительности имеет целочисленное решение.

Для получения органичений с целыми коэффициентами требуется умножение исходных ограничений на некоторые числа. Например, если имеется ограничение

$$X_1 + \frac{1}{3} * X_2 <= \frac{13}{2},$$

то его следует умножить на 6; будет получено ограничение

$$6*X_1 + 2*X_1 <= 39.$$

Метод реализуется в следующем порядке.

1. Задача решается без учета требований целочисленности (например, обычным симплекс-методом). Если все переменные в решении принимают целые значения, то алгоритм завершается.

2. Пусть решение исходной задачи оказалось нецелочисленным. Предположим, что последняя симплекс-таблица имеет следующий вид:

Таблица 3.1

БП	X_1	X_2	...	X_i	...	X_m	W_1	W_2	...	W_j	...	W_n	Решение
E	0	0	...	0	...	0	C_1	C_2	...	C_j	...	C_n	E_0
X_1	1	0	...	0	...	0	A_{11}	A_{12}	...	A_{1j}	...	A_{1n}	B_1
X_2	0	1	...	0	...	0	A_{21}	A_{22}	...	A_{2j}	...	A_{2n}	B_2
.....													
X_i	0	0	...	1	...	0	A_{i1}	A_{i2}	...	A_{ij}	...	A_{in}	B_i
.....													
X_m	0	0	...	0	...	1	A_{m1}	A_{m2}	...	A_{mj}	...	A_{mn}	B_m

Здесь через X_i, $i=1,\dots,m$, обозначены переменные, вошедшие в оптимальный базис; W_j, $j=1,\dots,n$ - небазисные переменные. Такие обозначения и расположение переменных (все базисные переменные - в левой части таблицы, небазисные - в правой) приняты только для удобства дальнейших объяснений; действительное расположение переменных и их обозначения могут быть любыми.

Пусть некоторые переменные приняли дробные значения, т.е. некоторые из чисел B_k, $k=1,\dots,m$ - дробные. Для дальнейшего рассмотрения выбирается переменная, значение которой имеет максимальную дробную часть. Пусть это переменная X_i, а ее значение - $B_i = [B_i]+\{B_i\}$; здесь $[B_i]$ - целая часть числа B_i, $\{B_i\}$ - дробная часть. Номер итерации алгоритма принимается равным 1: $r=1$.

3. Составляется дополнительное ограничение, которое делает полученное решение недопустимым. Для этого все коэффициенты строки, соответствующей выбранной переменной X_i, представляются в виде суммы целой и дробной части: $A_{ij} = [A_{ij}]+\{A_{ij}\}$, $j=1,\dots,n$. Правая часть также представляется в виде $B_i = [B_i]+\{B_i\}$. Новое ограничение записывается в следующем виде:

$$-\{A_{i1}\}*W_1 - \{A_{i2}\}*W_2 - \dots - \{A_{in}\}*W_n <= -\{B_i\}. \qquad (3.1)$$

Это ограничение называется отсечением Гомори. Можно показать, что такое ограничение следует из требования целочисленности всех переменных.

4. Полученное ограничение приводится к стандартному виду

$$-\{A_{i1}\}*W_1 - \{A_{i2}\}*W_2 - \dots - \{A_{in}\}*W_n + S_r = -\{B_i\}, \qquad (3.2)$$

где r - номер итерации алгоритма.

Здесь S_r - неотрицательная остаточная переменная. Так как в полученном ранее решении все переменные $W_j, j=1,\dots,n$ - небазисные (т.е. равные нулю), переменная S_r принимает значение $-\{B_i\}$, т.е. оказывается отрицательной. Таким образом, решение оказывается недопустимым.

5. К полученной ранее симплекс-таблице добавляются строка и столбец, соответствующие построенному ограничению (3.2) и новой базисной переменной S_r:

Таблица 3.2

БП	X_1	X_2	..	X_i	..	X_m	W_1	W_2	..	W_j	..	W_n	S_r	Решение
E	0	0	..	0	..	0	C_1	C_2	..	C_j	..	C_n	0	E_0
X_1	1	0	..	0	..	0	A_{11}	A_{12}	..	A_{1j}	..	A_{1n}	0	B_1
X_2	0	1	..	0	..	0	A_{21}	A_{22}	..	A_{2j}	..	A_{2n}	0	B_2
....														
X_i	0	0	..	1	..	0	A_{i1}	A_{i2}	..	A_{ij}	..	A_{in}	0	B_i
....														
X_m	0	0	..	0	..	1	A_{m1}	A_{m2}	..	A_{mj}	..	A_{mn}	0	B_m
S_r	0	0	..	0	..	0	$-\{A_{i1}\}$	$-\{A_{i2}\}$	..	$-\{A_{ij}\}$	..	$-\{A_{in}\}$	1	$-\{B_i\}$

Как видно, данная симплекс-таблица содержит недопустимое решение (переменная S_r имеет отрицательное значение). В то же время для данной симплекс-таблицы выполняется признак оптимальности решения (так как симплекс-таблица, построенная до добавления нового ограничения, соответствовала оптимальному нецелочисленному решению). Для поиска допустимого решения в этом случае используется двойственный симплекс-метод (см. раздел 1.7). ?

6. Новая задача (полученная при добавлении ограничения) решается двойственным симплекс-методом.

7. Если найденное решение целочисленное, то алгоритм завершается. В противном случае выбирается новая переменная X_i, значение которой имеет максимальную дробную часть. Происходит переход к очередной итерации алгоритма ($r=r+1$), и выполняется возврат к шагу 3.

Примечания.

1. При каждом повторении шага 5 в симплекс-таблицу добавляется одна строка и один столбец.

2. Переменные S_r, добавляемые при каждом выполнении шага 4, не имеют какого-либо содержательного смысла. В оптимальном целочисленном решении эти переменные всегда равны нулю.

3.3.3. Метод Гомори для частично целочисленных задач

Основные этапы реализации этого метода те же, что и для полностью целочисленных задач. Для этого метода необязательно, чтобы все коэффициенты задачи были целыми. Метод реализуется в следующем порядке:

1. Задача решается без учета требований целочисленности (например, обычным симплекс-методом). Если все переменные, которые по своему содержательному смыслу должны быть целочисленными, принимают целые значения, то найденное решение оптимально, и алгоритм завершается.

2. Пусть решение исходной задачи оказалось нецелочисленным. Предположим, что последняя симплекс-таблица имеет следующий вид.

Таблица 3.3

БП	X_1	X_2	...	X_i	...	X_m	W_1	W_2	...	W_j	...	W_n	Решение
E	0	0	...	0	...	0	C_1	C_2	...	C_j	...	C_n	E_0
X_1	1	0	...	0	...	0	A_{11}	A_{12}	...	A_{1j}	...	A_{1n}	B_1
X_2	0	1	...	0	...	0	A_{21}	A_{22}	...	A_{2j}	...	A_{2n}	B_2
....													
X_i	0	0	...	1	...	0	A_{i1}	A_{i2}	...	A_{ij}	...	A_{in}	B_i
....													
X_m	0	0	...	0	...	1	A_{m1}	A_{m2}	...	A_{mj}	...	A_{mn}	B_m

Здесь, как и для полностью целочисленной задачи, через X_i, $i=1,\dots,m$, обозначены переменные, вошедшие в оптимальный базис; W_j, $j=1,\dots,n$ - небазисные переменные. Такие обозначения и расположение переменных (все базисные переменные - в левой части таблицы, небазисные - в правой) приняты только для удобства дальнейших объяснений.

Пусть некоторые из переменных, которые должны быть целыми, приняли дробные значения. Для дальнейшего рассмотрения выбирается переменная, значение которой имеет максимальную дробную часть. Пусть это переменная X_i, а ее значение - $B_i = [B_i]+\{B_i\}$; здесь $[B_i]$ - целая часть числа B_i, $\{B_i\}$ - дробная часть. Номер итерации алгоритма принимается равным 1: $r=1$.

3. Составляется дополнительное ограничение, которое делает полученное решение недопустимым. Для этого все коэффициенты строки, соответствующей выбранной переменной X_i, представляются в виде суммы целой и дробной части: $A_{ij} = [A_{ij}]+\{A_{ij}\}$, $j=1,\dots,n$. Правая часть также представляется в виде $B_i = [B_i]+\{B_i\}$. Записывается новое ограничение:

$$L_1*W_1 + L_2*W_2 + \dots + L_n*W_n >= \{B_i\} \qquad (3.3)$$

где

$$L_j=\begin{cases} A_{ij}, & \text{если } A_{ij}>=0 \text{ и } W_j \text{ может быть дробной,}\\ \dfrac{\{B_i\}}{\{B_i\}-1}*A_{ij}, & \text{если } A_{ij}<0 \text{ и } W_j \text{ может быть дробной,}\\ \{A_{ij}\}, & \text{если } \{A_{ij}\}<=\{B_i\} \text{ и } W_j \text{ должна быть целой,}\\ \dfrac{\{B_i\}}{1-\{B_i\}}*(1-\{A_{ij}\}), & \text{если } \{A_{ij}\}>\{B_i\} \text{ и } W_j \text{ должна быть целой,}\end{cases}$$

$j=1,\dots,n$.

4. Полученное ограничение приводится к стандартному виду

$$-L_1*W_1 - L_2*W_2 - \dots - L_n*W_n + S_r = -\{B_i\}, \qquad (3.4)$$

где r - номер итерации алгоритма.

Здесь S_r - неотрицательная остаточная переменная, не имеющая никакого содержательного смысла; в оптимальном целочисленном решении эта переменная оказывается равной нулю. Так как в полученном ранее решении все переменные W_j, $j=1,..,n$ - небазисные (т.е. равные нулю), переменная S_r принимает значение $-\{B_i\}$, т.е. оказывается отрицательной. Таким образом, решение оказывается недопустимым.

5. К полученной ранее симплекс-таблице добавляются строка и столбец, соответствующие построенному ограничению (3.4) и новой базисной переменной S_r:

Таблица 3.4

БП	X_1	X_2	...	X_i	...	X_m	W_1	W_2	...	W_j	...	W_n	S_r	Решение
E	0	0	...	0	...	0	C_1	C_2	...	C_j	...	C_n	0	E_0
X_1	1	0	...	0	...	0	A_{11}	A_{12}	...	A_{1j}	...	A_{1n}	0	B_1
X_2	0	1	...	0	...	0	A_{21}	A_{22}	...	A_{2j}	...	A_{2n}	0	B_2
....														
X_i	0	0	...	1	...	0	A_{i1}	A_{i2}	...	A_{ij}	...	A_{in}	0	B_i
....														
X_m	0	0	...	0	...	1	A_{m1}	A_{m2}	...	A_{mj}	...	A_{mn}	0	B_m
S_r	0	0	...	0	...	0	$-L_1$	$-L_2$	...	$-L_j$	...	$-L_n$	1	$-\{B_i\}$

Как видно, данная симплекс-таблица содержит недопустимое решение (переменная S_r имеет отрицательное значение). В то же время для данной симплекс-таблицы выполняется признак оптимальности решения (так как симплекс-таблица, построенная до добавления нового ограничения, соответствовала оптимальному нецелочисленному решению). Для поиска допустимого решения в этом случае используется двойственный симплекс-метод (см. раздел 1.7).

6. Новая задача (полученная при добавлении ограничения) решается двойственным симплекс-методом.

7. Если найденное решение целочисленное, то алгоритм завершается. В противном случае выбирается новая переменная X_i, значение которой имеет максимальную дробную часть. Происходит переход к очередной итерации алгоритма ($r=r+1$), и выполняется возврат к шагу 3.

Примечание. Как и для полностью целочисленной задачи, при каждом повторении шага 5 в симплекс-таблицу добавляется одна строка и один столбец.

4. ОСОБЫЕ КЛАССЫ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ И МЕТОДЫ ИХ РЕШЕНИЯ

Среди задач линейного программирования, решаемых (в общем случае) симплекс-методом, существуют некоторые классы задач, которые из-за особенностей своей постановки могут быть решены более простыми методами. К ним относится, в частности, транспортная задача и задача о назначениях.

4.1. Транспортная задача

4.1.1. Постановка задачи

Транспортная задача формулируется следующим образом.

Имеется M пунктов отправления (поставщиков, производителей и т.д.), в которых имеются запасы некоторого товара (груза) в количествах A_1, $A_2, \dots, A_M$ единиц. Имеется также N пунктов назначения (потребителей), каждому из которых требуется $B_1, B_2, \dots, B_N$ единиц товара.

Предполагается, что сумма всех заявок равна сумме всех запасов:

$$\sum_{i=1}^{M} A_i = \sum_{j=1}^{N} B_j. \qquad (4.1)$$

Известна стоимость перевозки единицы груза от каждого пункта отправления до каждого пункта назначения. Эти стоимости задаются в виде матрицы стоимостей (тарифов) $C = (C_{ij})$, $i=1,\dots,M$, $j=1,\dots,N$.

Требуется составить такой план перевозок, при котором все заявки были бы выполнены, и при этом общая стоимость перевозок была бы минимальной. Здесь план перевозок - количество груза, поставляемого от каждого поставщика каждому потребителю.

Такую задачу можно сформулировать как задачу линейного программирования. Для этого вводится M*N переменных X_{ij}, обозначающих объем перевозок (количество перевозимых единиц груза) от i-го поставщика к j-му потребителю.

Система ограничений для такой задачи состоит из двух групп ограничений:

- ограничения на возможности поставки:

$$
\begin{aligned}
&X_{11} + X_{12} + \ldots + X_{1N} = A_1 \\
&X_{21} + X_{22} + \ldots + X_{2N} = A_2 \\
&\ldots\ldots \\
&X_{M1} + X_{M2} + \ldots + X_{MN} = A_M;
\end{aligned} \tag{4.2}
$$

- ограничения на спрос:

$$
\begin{aligned}
&X_{11} + X_{21} + \ldots + X_{M1} = B_1 \\
&X_{12} + X_{22} + \ldots + X_{M2} = B_2 \\
&\ldots\ldots \\
&X_{1N} + X_{2N} + \ldots + X_{MN} = B_N.
\end{aligned} \tag{4.3}
$$

Здесь система ограничений (4.2) устанавливает, что каждый поставщик поставляет потребителям ровно столько груза, сколько у него имеется. Система (4.3) показывает, что каждому потребителю доставляется столько груза, сколько ему требуется.

По содержанию задачи в системе (4.2) могли бы задаваться ограничения "не больше", а в (4.3) - "не меньше". Однако в соответствии с условием (4.1) используются именно ограничения-равенства.

Целевая функция:

$$E = \sum_{i=1}^{M} \sum_{j=1}^{N} C_{ij} * X_{ij} \to \min. \tag{4.4}$$

Целевая функция здесь выражает стоимость выполнения всех перевозок и подлежит минимизации.

Очевидно, что все переменные в такой задаче по своему смыслу неотрицательны, т.е. выполняется условие $X_{ij} >= 0$, $i=1,\ldots,M$, $j=1,\ldots,N$.

Требуется найти значения X_{ij}, $i=1,\ldots,M$, $j=1,\ldots,N$, минимизирующие целевую функцию (4.4) и удовлетворяющие ограничениям (4.2),(4.3). Их также удобно представлять в виде матрицы $X = (X_{ij})$; она называется матрицей перевозок. Некоторые переменные X_{ij} всегда оказываются равными нулю; по содержательному смыслу это означает, что перевозки от i-го поставщика j-му потребителю не выполняются (так как они невыгодны).

Любой набор значений переменных X_{ij}, удовлетворяющий ограничениям (4.2),(4.3), называется планом транспортной задачи (планом перевозок), а план, минимизирующий функцию (4.4), - оптимальным планом.

Такая задача может быть решена как обычная задача линейного программирования (т.е. симплекс-методом). Однако она имеет некоторые особенности, позволяющие решить ее более простыми методами. Все коэффициенты в ограничениях принимают значения 0 или 1, а каждая переменная входит только в 2 ограничения. Кроме того, сумма правых частей ограничений (4.2) равна сумме правых частей ограничений (4.3) (из условия (4.1)), поэтому система ограничений оказывается линейно зависимой, т.е. в ней содержится M+N-1 независимых ограничений; отсюда следует, что решение задачи будет содержать только M+N-1 базисных переменных (остальные - небазисные, т.е. нулевые).

Из содержательного смысла задачи ясно, что она всегда имеет допустимое решение (каким-то образом выполнить перевозки можно). Среди допустимых решений имеется оптимальное (возможно, не одно).

Важно отметить, что любая задача независимо от ее содержательного смысла может быть решена методами, разработанными для транспортной задачи, если ее можно представить в форме (4.2),(4.3),(4.4).

Для решения транспортной задачи ее удобно представлять в виде таблицы, содержащей матрицу стоимостей, запасы поставщиков и потребности потребителей (т.е. ограничения (4.1), (4.2), (4.3) и коэффициенты целевой функции):

Таблица 4.1

Поставщики \ Потребители	1	2		N	
1	C_{11}	C_{12}	...	C_{1N}	A_1
2	C_{21}	C_{22}		C_{2N}	A_2
....					
M	C_{M1}	C_{M2}		C_{MN}	A_M
	B_1	B_2		B_N	

В процессе решения задачи любым методом в эту таблицу заносится план перевозок (т.е. значения X_{ij}), а также другие данные, используемые в процессе решения задачи (в зависимости от используемого метода). Таким образом, в этой таблице совмещаются матрицы стоимостей и перевозок.

Транспортная задача решается в два этапа:

1) поиск допустимого решения (опорного плана);

2) поиск оптимального решения.

4.1.2. Методы поиска опорного плана

Опорным планом называется любой план, соответствующий системе ограничений (4.2),(4.3). В то же время существуют методы, позволяющие найти опорный план, близкий к оптимальному.

Все рассматриваемые методы позволяют найти опорный план $X=(X_{ij})$. Для нахождения стоимости перевозок (т.е. целевой функции) следует подставить найденные значения X_{ij} в (4.4).

Метод северо-западного угла

Этот метод является простейшим из всех методов поиска опорного плана. При его реализации заполнение матрицы перевозок X (т.е. составление плана) начинается с верхнего левого (северо-западного) угла таблицы 4.1. Величины перевозок X_{ij} заносятся в табл. 4.1 (таким образом, в каждой клетке указывается стоимость и объем перевозок). Если $A_1>B_1$ (запасы 1-го поставщика превышают потребности 1-го потребителя), то весь спрос 1-го потребителя удовлетворяется за счет 1-го поставщика (т.е. принимается $X_{11} = B_1$), и 1-й столбец таблицы, соответствующий 1-му потребителю, исключается из рассмотрения (вычеркивается); в противном случае все грузы от 1-го поставщика отправляются к 1-му потребителю ($X_{11}=A_1$), и из дальнейшего рассмотрения исключается 1-я строка таблицы (поставщик). Процесс повторяется для северо-западного угла сокращенной таблицы; решение заканчивается, когда достигается правый нижний угол таблицы (т.е. получает значение переменная X_{MN}).

Приведем алгоритм метода северо-западного угла.

1. Номера поставщика (i) и потребителя (j) принимаются равными 1: $i=1$, $j=1$.

2. Если $i=M$ и $j=N$, то $X_{ij}=B_j$, и алгоритм завершается. В противном случае выполняются следующие действия:

Если $A_i>B_j$, то $X_{ij}=B_j$, $A_i=A_i-B_j$, $j=j+1$, из рассмотрения исключается j-й столбец. Если $A_i<B_j$, то $X_{ij}=A_i$, $B_j=B_j-A_i$, $i=i+1$, исключается i-я строка. Если $A_i=B_j$, то см. раздел 4.1.4.

3. Переход к шагу 2.

Здесь (как и во всех других методах) исключение i-й строки означает, что запасы i-го поставщика полностью исчерпаны, а исключение j-го столбца, что спрос j-го потребителя полностью удовлетворен.

Метод северо-западного угла очень прост, легко реализуется вручную, но полученный опорный план обычно далек от оптимального, так как при его построении вообще не учитываются стоимости.

Метод минимального элемента

Метод минимального элемента позволяет при достаточно простой реализации учитывать стоимости перевозок. В соответствии с этим методом составление плана начинается с минимального элемента матрицы стоимостей, т.е. первыми планируются самые дешевые перевозки. Как и в методе северо-западного угла, из рассмотрения на каждом шаге исключается или поставщик, запасы которого исчерпаны, или потребитель, спрос которого удовлетворен. Процесс продолжается, пока не будут распределены все запасы (и одновременно - пока не будет удовлетворен весь спрос). Все действия выполняются на табл. 4.1.

Приведем алгоритм реализации метода минимального элемента.

1. Фиксируется количество поставщиков (M) и потребителей (N).

2. В матрице стоимостей находится минимальный элемент: $C_{ij}=\min C$.

3. Если $M=1$ и $N=1$, то $X_{ij}=B_j$, и алгоритм завершается. В противном случае выполняются следующие действия.

Если $A_i>B_j$, то $X_{ij}=B_j$, $A_i=A_i-B_j$, $N=N-1$, из рассмотрения исключается j-й столбец. Если $A_i<B_j$, то $X_{ij}=A_i$, $B_j=B_j-A_i$, $M=M-1$, исключается i-я строка. Если $A_i=B_j$, то см. раздел 4.1.4.

4. Переход к шагу 2.

Решения, полученные этим методом, практически всегда лучше, чем полученные по методу северо-западного угла.

Метод Фогеля

Основная идея метода Фогеля состоит в следующем. В каждой строке и в каждом столбце матрицы стоимостей находится разность между двумя минимальными элементами (т.е. между двумя минимальными стоимостями). В строке или в столбце, которому соответствует максимальная разность, находится минимальный элемент (минимальная стоимость) и для него планируются максимально возможные перевозки; при этом, как и в предыдущих методах, из рассмотрения исключается или поставщик, или потребитель. Процесс продолжается, пока не будут распределены все запасы (и одновременно - пока не будет удовлетворен весь спрос).

Приведем алгоритм реализации метода Фогеля.

1. Фиксируется количество поставщиков (M) и потребителей (N).

2. В каждой строке матрицы стоимостей находится модуль разности между двумя минимальными элементами этой строки: d_i, $i=1,\dots,M$.

3. В каждом столбце матрицы стоимостей также находится модуль разности между двумя минимальными элементами этого столбца: e_j, $j=1,\dots,N$.

4. Из найденных на шагах 3 и 4 разностей находится максимальная; в соответствующей строке или столбце находится минимальный элемент C_{ij}.

5. Если $M=1$ и $N=1$, то $X_{ij}=B_j$, и алгоритм завершается.

6. Если $A_i>B_j$, то $X_{ij}=B_j$, $A_i=A_i-B_j$, $N=N-1$, из рассмотрения исключается j-й столбец. Если $A_i<B_j$, то $X_{ij}=A_i$, $B_j=B_j-A_i$, $M=M-1$, исключается i-я строка. Если $A_i=B_j$, то см. раздел 4.1.4.

7. Переход к шагу 2.

Примечания.

1. Если при выполнении шагов 3 и 4 оказывается, что в матрице стоимостей осталась одна строка, то разности минимальных элементов всех столбцов принимаются равными нулю, так как в каждом столбце остается только по одному элементу: $e_j=0, j=1,\dots,N$. Аналогично, если остается только один столбец, то $d_i=0, i=1,\dots,M$.

2. Если в строке или столбце несколько элементов, равных минимальному, то на шаге 3 или 4 соответствующая разность находится между минимальным и следующим по величине элементом (т.е. эта разность не равна нулю). Например, если в i-й строке есть стоимости 5, 8, 5 и 12, то $d_i=8-5=3$.

Решения, полученные методом Фогеля, близки к оптимальным, а в некоторых случаях - оптимальные.

4.1.3. Поиск оптимального плана. Метод потенциалов

Основная идея метода потенциалов

Общую идею метода потенциалов можно выразить следующим образом. Пусть имеется некоторый допустимый план $X = (X_{ij})$, найденный одним из рассмотренных выше методов. Пусть значение целевой функции для этого плана равно Z. Можно доказать, что в этом случае целевая функция (4.4) может быть представлена в следующем виде:

$$E = \sum_{i=1}^{M}\sum_{j=1}^{N}(C_{ij} - U_i - V_j) * X_{ij} = \sum_{i=1}^{M}\sum_{j=1}^{N}(C_{ij} - \hat{C}_{ij}) * X_{ij} = Z - \sum_{i=1}^{M} U_i * A_i - \sum_{j=1}^{N} V_j * B_j, \tag{4.5}$$

где C_{ij} - коэффициенты целевой функции (4.4);

A_i - запасы груза у поставщиков, $i=1,\dots,M$;

B_j - величины спроса потребителей, $j=1,\dots,N$;

U_i - некоторые числа, называемые платежами поставщиков, $i=1,\dots,M$;

V_j - некоторые числа, называемые платежами потребителей, $j=1,\dots,N$;

$\hat{C}_{ij} = U_i + V_j$ - величина, называемая псевдостоимостью перевозки от i-го поставщика к j-му потребителю.

Известно, что коэффициенты целевой функции при базисных переменных равны нулю (см. симплекс-метод). Значит, если X_{ij} – базисная переменная, то $C_{ij}-U_i-V_j=0$, или $C_{ij}=U_i+V_j$. Всего имеется M+N-1 базисных переменных. Составив M+N-1 уравнений вида $U_i+V_j=C_{ij}$, можно найти значения платежей U_i и V_j, $i=1,\dots,M$, $j=1,\dots,N$. При этом, так как количество неизвестных (платежей) равно M+N, а количество уравнений - M+N-1, один из платежей (любую из величин U_i или V_j) необходимо приравнять к нулю.

По значениям платежей U_i и V_j определяются значения псевдостоимостей $\hat{C}_{ij}$ и разности C_{ij} - $\hat{C}_{ij}$ для всех небазисных переменных X_{ij}. Из (4.5) видно, что если $C_{ij} - \hat{C}_{ij}<0$, то увеличение соответствующей переменной X_{ij} приводит к уменьшению значения целевой функции. Значит, в этом случае для снижения стоимости перевозок следует увеличить значение X_{ij} на определенную величину (т.е. ввести ее в базис), исключив из базиса одну из находившихся в нем переменных и изменив соответствующим образом значения других базисных переменных (чтобы соблюдались все ограничения задачи). Затем для измененного базиса снова находятся платежи и псевдостоимости и т.д. Если выполняется условие $C_{ij}-\hat{C}_{ij}>=0$ (т.е. для всех небазисных переменных псевдостоимости не превосходят стоимости перевозок), это означает, что никакое увеличение любой из переменных не уменьшает значения целевой функции, т.е. план оптимален.

Алгоритм реализации метода потенциалов

Решение задач с помощью метода потенциалов выполняется на основе какого-либо допустимого плана, найденного одним из рассмотренных выше способов. Поиск оптимального плана методом потенциалов выполняется в следующем порядке.

1. Находится некоторый допустимый план. Этот план заносится в таблицу, т.е. составляется матрица перевозок (табл. 4.2). В допустимом плане имеется M+N-1 базисных переменных (некоторые из них в принципе могут быть и нулевыми), остальные - небазисные (нулевые).

Таблица 4.2

Поставщики \ Потребители	1	2	...	N	
1	C_{11} X_{11}	C_{12} X_{12}	...	C_{1N} X_{1N}	A_1
2	C_{21} X_{21}	C_{22} X_{22}	...	C_{2N} X_{2N}	A_2
....					
M	C_{M1} X_{M1}	C_{M2} X_{M2}	...	C_{MN} X_{MN}	A_M
	B_1	B_2		B_N	

2. Для всех заполненных (базисных) клеток матрицы перевозок составляются уравнения для нахождения платежей:

$$U_i + V_j = C_{ij}. \tag{4.6}$$

Здесь i, j - такие, что X_{ij} - базисные; всего составляется $M+N-1$ таких уравнений.

3. Из системы (4.6) находятся значения платежей U_i и V_j (одно из них предполагается равным нулю). Найденные значения удобно отметить у соответствующих строк и столбцов матрицы перевозок.

4. Для всех небазисных переменных (т.е. для свободных клеток матрицы перевозок) находятся псевдостоимости $\widehat{C}_{ij} = U_i + V_j$; их также следует занести в матрицу перевозок:

Таблица 4.3

Поставщики \ Потребители	1	2	...	N	
1	$\widehat{C}_{11}$ C_{11} X_{11}	$\widehat{C}_{12}$ C_{12} X_{12}	...	$\widehat{C}_{1N}$ C_{1N} X_{1N}	A_1 (U_1)
2	$\widehat{C}_{21}$ C_{21} X_{21}	$\widehat{C}_{22}$ C_{22} X_{22}	...	$\widehat{C}_{2N}$ C_{2N} X_{2N}	A_2 (U_2)
....					
M	$\widehat{C}_{M1}$ C_{M1} X_{M1}	$\widehat{C}_{M2}$ C_{M2} X_{M2}	...	$\widehat{C}_{MN}$ C_{MN} X_{MN}	A_M (U_M)
	B_1 (V_1)	B_2 (V_N)		B_N (V_N)	

5. Для всех небазисных переменных находятся разности стоимостей и псевдостоимостей: $D_{ij} = C_{ij} - \hat{C}_{ij}$.

6. Если все $D_{ij} >= 0$, то оптимальный план найден; алгоритм завершается. Значения переменных X_{ij} равны оптимальным объемам перевозок. Соответствующие оптимальному плану значения платежей U_i и V_j называются потенциалами поставщиков и потребителей соответственно. Если есть отрицательные разности $D_{ij} < 0$, то выполняется следующий шаг.

7. Переменная X_{ij}, соответствующая максимальному по модулю отрицательному значению D_{ij}, включается в базис. В матрице перевозок соответствующая клетка отмечается знаком "плюс".

8. Определяется переменная, исключаемая из базиса.

Так как включение переменной X_{ij} в базис соответствует увеличению перевозок от i-го поставщика j-му потребителю, для соблюдения ограничений (4.2) необходимо при этом уменьшить перевозки от этого же поставщика какому-то другому (k-му) потребителю, т.е. уменьшить X_{ik}; для удовлетворения его потребностей (ограничения (4.3)) необходимо увеличить перевозки ему от некоторого другого поставщика и т.д. Все эти изменения представляются в матрице перевозок в виде цикла

Таблица 4.4

	1		j		s		k		N	
1	$\hat{C}_{11}$ C_{11} X_{11}	..	$\hat{C}_{1j}$ C_{1j} X_{1j}	..	$\hat{C}_{1s}$ C_{1s} X_{1s}	..	$\hat{C}_{1k}$ C_{1k} X_{1k}	..	$\hat{C}_{1N}$ C_{1N} X_{1N}	A_1 (U_1)
....										
i	$\hat{C}_{i1}$ C_{i1} X_{i1}	..	$\hat{C}_{ij}$ C_{ij} X_{ij} +	..	$\hat{C}_{is}$ C_{is} X_{is}	..	$\hat{C}_{ik}$ C_{ik} X_{ik} −	..	$\hat{C}_{iN}$ C_{iN} X_{iN}	A_i (U_i)
....										
r	$\hat{C}_{r1}$ C_{r1} X_{r1}	..	$\hat{C}_{rj}$ C_{rj} X_{rj}	..	$\hat{C}_{rs}$ C_{rs} X_{rs} −	..	$\hat{C}_{rk}$ C_{rk} X_{rk} +	..	$\hat{C}_{rN}$ C_{rN} X_{rN}	A_r (U_r)
....										
q	$\hat{C}_{q1}$ C_{q1} X_{q1}	..	$\hat{C}_{qj}$ C_{qj} X_{qj} −	..	$\hat{C}_{qs}$ C_{qs} X_{qs} +	..	$\hat{C}_{qk}$ C_{qk} X_{qk}	..	$\hat{C}_{qN}$ C_{qN} X_{qN}	A_q (U_q)
....										
M	$\hat{C}_{M1}$ C_{M1} X_{M1}	..	$\hat{C}_{Mj}$ C_{Mj} X_{Mj}	..	$\hat{C}_{Ms}$ C_{Ms} X_{Ms}	..	$\hat{C}_{Mk}$ C_{Mk} X_{Mk}	..	$\hat{C}_{MN}$ C_{MN} X_{MN}	A_M (U_M)
	B_1 (V_1)		B_j (V_j)		B_s (V_s)		B_k (V_k)		B_N (V_N)	

Вершины цикла (базисные переменные), соответствующие уменьшению перевозок, отмечаются знаком "минус", а соответствующие увеличению - знаком "плюс". Переменная из числа отмеченных знаком "минус", имеющая минимальное значение (обозначим ее X_{rs}), исключается из базиса, т. е. становится равной нулю; по содержательному смыслу задачи это значит, что перевозки от r-го поставщика к s-му потребителю не должны выполняться.

9. Определяются новые значения базисных переменных. Все переменные, отмеченные знаком "минус", уменьшаются на величину X_{rs}, а все отмеченные знаком "плюс" (в том числе новая базисная переменная X_{ij}) - увеличиваются на X_{rs}.

10. Для нового базиса вычисляется значение целевой функции:

$$E = Z - |D_{ij}| * X_{rs}, \qquad (4.7)$$

где Z - значение целевой функции (стоимость перевозок) для предыдущего базиса;

D_{ij} - максимальная по модулю отрицательная разность стоимости и псевдостоимости, найденная на шаге 5; эта величина называется ценой цикла;

X_{rs} - значение переменной, исключенной из базиса.

Как видно из (4.7), в результате перехода к новому базису значение целевой функции уменьшается (т. е. стоимость перевозок снижается).

11. По результатам шага 9 строится новая матрица перевозок (в нее заносится новый план).

12. Возвращение к шагу 2.

Примечание. Если при пересчете базисных переменных (на шаге 9) оказывается, что одновременно несколько базисных переменных обращаются в ноль (т. е. решение оказывается вырожденным), то из базиса исключается только одна из них; остальные остаются в базисе (с нулевыми значениями). Это необходимо для того, чтобы на следующем этапе оптимизации сохранялась возможность построить систему уравнений (4.6) для нахождени.. платежей U_i и V_j. См. также раздел 4.1.4.

4.1.4. Вырожденное решение

Решение называется вырожденным, если в базисе оказывается менее, чем M+N-1 ненулевых переменных. Вырожденное решение может быть получено на этапе поиска как допустимого, так и оптимального решения. В базисе всегда должно быть ровно M+N-1 переменных (некоторые из них могут быть и нулевыми); это необходимо для того, чтобы при поиске оптимального решения всегда сохранялась возможность построения системы (4.6) из M+N-1 уравнений для нахождения платежей U_i и V_j. С этой целью необходимо соблюдать следующие правила.

1. При поиске допустимого решения (любым методом) на каждом шаге исключается или только строка, или только столбец (но не то и другое вместе). Если оказывается, что запасы поставщика и спрос потребителя совпадают (A_i=B_j), то принимается X_{ij}=B_j, но из рассмотрения исключается или только потребитель (тогда остаток запаса у поставщика предполагается равным 0), или только поставщик (тогда считается, что спрос потребителя остался неудовлетворенным на 0 единиц). Затем построение плана продолжается в соответствии с обычными правилами используемого метода. В результате будет получено решение, содержащее ровно M+N-1 базисных переменных, некоторые из которых будут нулевыми.

2. При поиске оптимального решения, если при пересчете базисных переменных оказывается, что несколько базисных переменных одновременно обращаются в ноль, то из базиса исключается только одна из них; остальные остаются в базисе с нулевыми значениями.

Рассмотрим пример вырожденного допустимого решения. Пусть требуется методом Фогеля найти допустимый план перевозок для транспортной задачи, заданной в табл. 4.5.

Таблица 4.5

Поставщики	Потребители 1	2	3	
1	7	8	12	60
2	6	5	10	40
	30	40	30	

Приведем ход решения задачи в соответствии с алгоритмом, описанным в разделе 4.1.2.3.

1. $M=2$, $N=3$.

2. $d_1=8-7=1$, $d_2=6-5=1$.

3. $e_1=7-6=1$, $e_2=8-5=3$, $e_3=12-10=2$.

4. Максимальная разность: $e_2=3$. Минимальный элемент (минимальная стоимость перевозки) ищется во 2-м столбце: $C_{22}=5$.

5. $M=2$, $N=3$; алгоритм не завершен.

6. $A_2=B_2$. Можно исключить из рассмотрения или поставщика (2-я строка), или потребителя (2-й столбец). Исключаем 2-й столбец. $X_{22}=40$, $A_2=40-40=0$, $N=3-1=2$.

2. $d_1=12-7=5$, $d_2=10-6=4$.

3. $e_1=7-6=1$, $e_3=12-10=2$ (e_2 не определяется, так как 2-й столбец исключен).

4. Максимальная разность: $d_1=5$. Минимальный элемент ищется в 1-й строке: $\min(C_{11},C_{13}) = C_{11} = 7$.

5. $M=2$, $N=2$; алгоритм не завершен.

6. $A_1>B_1$; $X_{11}=30$, $A_1=60-30=30$, $N=2-1=1$, исключается 1-й столбец.

2. $d_1=0$, $d_2=0$ (так как есть только один столбец).

3. $e_3=12-10=2$.

4. Максимальная разность: $e_3=2$. Минимальный элемент ищется в 3-м столбце: C_{23}.

5. $M=2$, $N=1$; алгоритм не завершен.

6. $A_2<B_3$; $X_{23}=0$, $B_3=30-0=30$, $M=2-1=1$, исключается 2-я строка.

2. $d_1=0$.

3. $e_3=0$.

4. Так как $d_1=e_3=0$, в качестве максимальной можно взять любую из этих разностей. Пусть максимальная разность: $d_1=0$. Минимальный элемент в 1-й строке: $C_{13}=12$ (единственный).

5. $M=1$, $N=1$; $X_{13}=30$. Конец алгоритма. Допустимый план:

Таблица 4.6

Поставщики \ Потребители	1	2	3	
1	30 (7)	(8)	30 (12)	60
2	(6)	40 (5)	0 (10)	40
	30	40	30	

$E = 30*7 + 30*12 + 40*5 + 0*10 = 770$ (стоимость перевозок).

Как видно, полученное решение является вырожденным (имеется только 3 ненулевых переменных). Однако в базисе имеется $M+N-1 = 4$ переменных; это позволяет найти оптимальное решение методом потенциалов.

4.1.5. Задачи с неправильным балансом

Задачи с неправильным балансом - это задачи, в которых нарушено условие (4.1), т.е. сумма запасов (предложение) не совпадает со спросом. Здесь вовможны два случая.

1. Предложение превышает спрос:

$$\sum_{i=1}^{M} A_i > \sum_{j=1}^{N} B_j. \qquad (4.8)$$

В этом случае сверх имеющихся N потребителей вводится еще один, фиктивный потребитель (N+1-й), спрос которого считается равным избытку запасов над заявками:

$$B_{N+1} = \sum_{i=1}^{M} A_i - \sum_{j=1}^{N} B_j.$$

Стоимости перевозок от всех поставщиков данному потребителю считаются равными нулю, т.е. $C_{i(N+1)} = 0$, $i=1,\dots,M$.

После этого задача решается как обычная транспортная задача с правильным балансом, с M поставщиками и N+1 потребителем. Если какие-либо из переменных $X_{i(N+1)}$ попадают в оптимальное решение, это означает, что соответствующее количество груза остается у i-го поставщика.

2. Спрос превышает предложение:

$$\sum_{i=1}^{M} A_i < \sum_{j=1}^{N} B_j. \qquad (4.9)$$

В этом случае сверх имеющихся M поставщиков вводится еще один, фиктивный поставщик (M+1-й), запас груза у которого считается равным избытку спроса над запасами:

$$A_{M+1} = \sum_{j=1}^{N} B_j - \sum_{i=1}^{M} A_i.$$

Стоимости перевозок от данного поставщика любому из потребителей считаются равными нулю, т.е. $C_{(M+1)j} = 0$, $j=1,\dots,N$.

После этого задача решается как обычная транспортная задача с правильным балансом, с M+1 поставщиком и N потребителями. Если какие-либо из переменных $X_{(M+1)j}$ попадают в оптимальное решение, это означает, что соответствующее количество груза не будет поставлено j-му потребителю.

Такой способ применим только в случае, если единственным требованием к плану является обеспечение минимальной стоимости перевозок. Однако в задачах с недостатком запасов возможны и дополнительные требования. Например, если требуется равномерно распределить недопоставки между всеми потребителями (т.е. удовлетворить спрос каждого потребителя в равной доле), то спрос каждого потребителя умножается на коэффициент, вычисляемый по формуле

$$k = \sum_{i=1}^{M} A_i \Big/ \sum_{j=1}^{N} B_j.$$

Как видно, $k<1$. Таким образом, спрос всех потребителей искусственно занижается. Задача сводится к правильному балансу и решается как обычная транспортная задача, без введения фиктивного поставщика.

Если в задаче с недостатком запасов требуется, чтобы спрос какого-либо (j-го) потребителя обязательно был удовлетворен полностью, то стоимость перевозки от фиктивного поставщика данному потребителю предполагается равной некоторому очень большому числу M. Таким образом, перевозки от фиктивного поставщика j-му потребителю становятся невыгодными; при нахождении оптимального плана они окажутся равными нулю.

4.2. Задача о назначениях

4.2.1. Постановка задачи

Задача о назначениях представляет собой частный случай задачи линейного программирования, а также транспортной задачи.

Формулировка задачи о назначениях следующая.

Необходимо выполнить N работ. Каждая из этих работ может быть выполнена каждым из N исполнителей. При этом каждый исполнитель может быть назначен только на одну (любую) работу. Известны затраты (времени, денежных средств или других ресурсов), связанные с выполнением каждым из исполнителей каждой работы. Эти затраты задаются в виде матрицы $C = (C_{ij})$, $i,j=1,\dots,N$; здесь C_{ij} - затраты на выполнение i-м исполнителем j-й работы. Матрица C называется матрицей затрат. Строки матрицы соответствуют исполнителям, столбцы - работам.

Требуется распределить исполнителей по работам таким образом, чтобы все работы были выполнены, и при этом общие затраты на их выполнение были минимальными.

Такую задачу можно сформулировать как задачу линейного программирования. Для этого вводится N*N переменных X_{ij}, принимающих следующие значения:

$$X_{ij} = \begin{cases} 1, & \text{если i-й исполнитель назначен на j-ю работу;} \\ 0, & \text{в противном случае.} \end{cases}$$

Система ограничений для такой задачи состоит из двух групп ограничений:

$$\begin{aligned} X_{11} + X_{12} + \dots + X_{1N} &= 1 \\ X_{21} + X_{22} + \dots + X_{2N} &= 1 \\ \dots\dots & \\ X_{N1} + X_{N2} + \dots + X_{NN} &= 1 \end{aligned} \qquad (4.10)$$

$$\begin{aligned} X_{11} + X_{21} + \dots + X_{N1} &= 1 \\ X_{12} + X_{22} + \dots + X_{N2} &= 1 \\ \dots\dots & \\ X_{1N} + X_{2N} + \dots + X_{NN} &= 1. \end{aligned} \qquad (4.11)$$

Целевая функция:

$$E = \sum_{i=1}^{N} \sum_{j=1}^{N} C_{ij} * X_{ij} \to \min. \qquad (4.12)$$

Здесь группа ограничений (4.10) устанавливает, что каждый из исполнителей может быть назначен только на одну работу, а группа (4.11) - что на каждую работу должен быть назначен в точности один исполнитель. Целевая функция выражает затраты на выполнение всех работ.

Задача состоит в определении значений X_{ij}, $i,j=1,\dots,N$, минимизирующих целевую функцию (4.12) и удовлетворяющих ограничениям (4.10),(4.11). Решение задачи удобно представлять в виде матрицы $X = (X_{ij})$. В этой матрице имеется ровно N элементов, равных 1 (по одному в каждой строке и в каждом столбце); остальные элементы матрицы - нулевые. Матрица X называется матрицей назначений.

Такая задача может быть решена как обычная задача линейного программирования (т.е. симплекс-методом) Ее можно также решить как транспортную задачу, в которой запасы всех поставщиков и потребности потребителей равны 1. Однако для решения задачи о назначениях имеются специальные методы. Эти методы могут применяться для решения любой задачи, независимо от ее содержательного смысла, если ее можно представить в форме (4.10),(4.11),(4.12).

Как и транспортная задача, задача о назначениях всегда имеет допустимое решение. Среди допустимых решений имеется оптимальное (возможно, не одно).

4.2.2. Особые случаи задачи о назначениях

Рассмотрим некоторые особые случаи, которые могут возникнуть при решении задачи о назначениях.

1. Если какой-либо (i-й) исполнитель не может выполнить некоторую (j-ю) работу, то стоимость выполнения этой работы данным исполнителем C_{ij} принимается равной очень большому числу. Таким образом назначение i-го исполнителя на j-ю работу исключается.

2. Если количество работ и исполнителей не совпадает, то вводятся фиктивные исполнители или фиктивные работы (аналогично транспортной задаче с неправильным балансом).

3. Иногда в постановке задачи о назначениях вместо затрат задаются выигрыши, т.е. вместо матрицы затрат C задается

матрица выигрышей $V = (V_{ij})$, где V_{ij} - выигрыш от выполнения i-м исполнителем j-й работы, $i,j=1,\dots,N$. Такая задача также формулируется в форме (4.10),(4.11),(4.12), но целевая функция (4.12) подлежит максимизации. В этом случае необходимо перейти к задаче с целевой функцией, подлежащей минимизации. Для этого следует выбрать максимальный элемент матрицы выигрышей D и вычесть из него все элементы этой матрицы. Таким образом выполняется переход к матрице затрат C. Другой способ перехода - умножение матрицы выигрышей на -1.

4.2.3. Решение задачи о назначениях. Метод Мака

Все методы решения задачи о назначениях основаны на следующем простом утверждении: решение задачи не изменится, если к любой строке или столбцу матрицы затрат прибавить или вычесть любое число. По содержательному смыслу такая операция соответствует изменению затрат на выполнение всех работ для некоторого исполнителя (если изменяются элементы строки) или изменению затрат на выполнение некоторой работы для всех исполнителей (при изменении столбца).

Метод Мака основан на выборе в каждой строке матрицы затрат минимального элемента (т.е. нахождении для каждого исполнителя той работы, которую он может выполнить с минимальными затратами). Если найденные минимальные элементы оказываются распределенными по всем N столбцам матрицы, то решение задачи завершается; по содержанию задачи это означает, что на каждую работу назначен исполнитель. Если минимальные элементы находятся не во всех столбцах матрицы, то по определенному алгоритму выполняются сложения одного и того же числа с элементами некоторых столбцов; цель этих операций - распределение минимальных элементов строк по всем столбцам.

Перед реализацией алгоритма Мака выполняется подготовительный этап: в каждой строке матрицы затрат C находится минимальный элемент. Эти элементы подчеркиваются (набор подчеркнутых элементов называется базисом задачи о назначениях). Если подчеркнутые элементы имеются во всех столбцах матрицы C, то выполнение алгоритма не требуется: назначение исполнителей на работы задается положением подчеркнутых элементов. Если подчеркнутые элементы есть не во всех столбцах, это означает, что исполнители назначены не на все работы; в этом случае необходимы преобразования матрицы затрат по приведенному ниже алгоритму.

Метод Мака реализуется в следующем порядке.

1. Множество столбцов матрицы C разделяется на два множества: A и A'. Множество A называется выбранным, а A' - невыбранным. Множество A принимается пустым, а A' - содержащим все столбцы матрицы:

$$A = \varnothing$$
$$A' = \{j \mid j=1,\dots,N\}.$$

2. Из множества A' выбирается столбец, содержащий более одного подчеркнутого элемента (если таких столбцов несколько, то берется любой из них). Этот столбец исключается из множества A' и включается во множество A.

3. Находится множество строк, в которых имеются подчеркнутые элементы, содержащиеся в столбцах множества A. Это множество строк обозначим через B. Таким образом, i-я строка включается во множество B, если существует подчеркнутый элемент C_{ij}, такой, что $j \in A$.

4. Для каждой из строк, входящих во множество B, находится минимальный элемент из столбцов, не принадлежащих множеству A:

$$C_{ik} = \min_{j \notin A} C_{ij}, \qquad i \in B.$$

5. Для каждой из строк, входящих во множество B, находится разность:

$$d_i = C_{ik} - C_{ij}, \qquad i \in B,$$

где C_{ik} - элемент, найденный на шаге 4;

C_{ij} - подчеркнутый элемент строки i.

Здесь $k \notin A$, $j \in A$, $i \in B$. Разности d_i всегда неотрицательны (могут быть нулевыми).

6. Из найденных на шаге 5 разностей находится минимальная:

$$D = \min_{i \in B} d_i.$$

Пусть минимальная разность соответствует r-й строке. Пусть элемент r-й строки, найденный на шаге 4, находится в s - м столбце, а подчеркнутый элемент этой строки - в q-м столбце. Тогда $D = C_{rs} - C_{rq}$.

7. Все элементы стоблцов из множества A увеличиваются на величину D:

$$C_{ij} = C_{ij} + D, \qquad i=1,\ldots,N, \; j \in A.$$

8. Элемент C_{rs} отмечается точками внизу. Далее он будет называться "отмеченный точками элемент".

9. Если в столбце s имеется хотя бы один подчеркнутый элемент, то этот столбец исключается из множества A' и включается в A, и выполняется переход к шагу 3; иначе - выполняется следующий шаг.

10. Элемент C_{rs} подчеркивается, а с элемента C_{rq} снимается подчеркивание.

11. Если в столбце q (т.е. в столбце, содержащем элемент, с которого на шаге 10 снято подчеркивание) есть еще хотя бы один подчеркнутый элемент, то выполняется переход к шагу 15. Если в этом столбце нет ни одного подчеркнутого элемента, то в нем обязательно есть элемент, отмеченный точками. Пусть этот элемент находится в строке p (элемент C_{pq}). В строке p находится подчеркнутый элемент; пусть это элемент C_{pf}.

12. Выполняются присваивания: r=p, s=q. Таким образом, отмеченный точками элемент C_{pq} обозначается как C_{rs}.

13. Выполняется присваивание: q=f. Подчеркнутый элемент C_{pf} обозначается как C_{rq}.

14. Возврат к шагу 10.

15. Если подчеркнутые элементы есть во всех столбцах, то алгоритм завершается; иначе- выполняется возврат к шагу 1.

По окончании алгоритма в матрице затрат имеется ровно N подчеркнутых элементов, по одному в каждой строке и столбце. Оптимальное решение задачи о назначениях находится следующим образом: если C_{ij} - подчеркнутый элемент, то $X_{ij}=1$ (т.е. i-й исполнитель должен быть назначен на j-ю работу). Таким образом, ровно N переменных X_{ij} получают значение 1, остальные - значение 0.

Затраты на выполнение работ (значение целевой функции) вычисляются по формуле (4.12). При этом используются элементы C_{ij} из исходной матрицы затрат (а не из полученной в результате выполнения алгоритма).

4.2.4. Пример решения задачи о назначениях

Пусть требуется выполнить 4 работы. Каждая из них может быть выполнена любым из четырех исполнителей. Известны затраты, связанные с выполнением работ исполнителями (строки соответствуют исполнителям, столбцы - работам):

$$\begin{matrix} 4 & 6 & 9 & 7 \\ 13 & 10 & 14 & 14 \\ 9 & 9 & 16 & 13 \\ 12 & 10 & 12 & 10. \end{matrix}$$

Приведем последовательность решения задачи методом Мака.

Сначала в каждой строке матрицы затрат находится минимальный элемент (каждый исполнитель назначается на работу, которую он может выполнить с минимальными затратами). Эти элементы подчеркиваются:

$$\begin{matrix} \underline{4} & 6 & 9 & 7 \\ 13 & \underline{10} & 14 & 14 \\ \underline{9} & 9 & 16 & 13 \\ 12 & \underline{10} & 12 & 10. \end{matrix}$$

При этом оказывается, что на 3-ю и 4-ю работы не назначены исполнители. Значит, требуются преобразования матрицы затрат для поиска допустимого (и одновременно - оптимального) плана.

1. $A=\varnothing$, $A'=\{1,2,3,4\}$.

2. Во множестве A' есть два столбца, содержащих более одного подчеркнутого элемента: столбцы 1 и 2. Выбираем любой из них, например, 1. Тогда $A=\{1\}$, $A'=\{2,3,4\}$.

3. В столбце 1 есть подчеркнутые элементы из строк 1 и 3. Значит, $B=\{1,3\}$.

4. В каждой из строк 1 и 3 выбираем минимальный элемент из столбцов 2,3,4. Для строки 1 это элемент $C_{12}=6$, для строки 3 - $C_{32}=9$.

5. Для строк 1 и 3 находим разности элементов, найденных на шаге 4, и подчеркнутых элементов: $d_1=C_{12}-C_{11}=6-4=2$, $d_3=C_{32}-C_{31}=9-9=0$.

6. Находим минимальную разность: $D=\min\{2,0\}=0$. Таким образом, $D=C_{32}-C_{31}$. $C_{rs}=C_{32}$, $C_{rq}=C_{31}$, $r=3, s=2, q=1$.

7. Ко всем элементам столбца 1 прибавляется значение $D=0$: $C_{i1}=C_{i1}+0$, $i=1,\dots,4$. Матрица затрат не изменяется.

8. Элемент $C_{rs}=C_{32}$ отмечается точками:

$$\begin{array}{cccc} \underline{4} & 6 & 9 & 7 \\ 13 & \underline{10} & 14 & 14 \\ \underline{9} & \underset{\cdots}{9} & 16 & 13 \\ 12 & \underline{10} & 12 & 10. \end{array}$$

9. В столбце $s=2$ есть подчеркнутые элементы. Значит, он переносится во множество A: $A=\{1,2\}$, $A'=\{3,4\}$. Переход к шагу 3.

3. В столбцах 1 и 2 есть подчеркнутые элементы из всех строк, значит, $B=\{1,2,3,4\}$.

4. В каждой из строк 1,2,3,4 находится минимальный элемент из столбцов 3,4: $C_{14}=7$, $C_{23}=14$, $C_{34}=13$, $C_{44}=10$.

5. Находятся разности найденных на предыдущем шаге и подчеркнутых элементов: $d_1=7-4=3$, $d_2=14-10=4$, $d_3=13-9=4$, $d_4=10-10=0$.

6. $D=\min\{3,4,4,0\}=0$. $D=C_{44}-C_{42}$, $C_{rs}=C_{44}$, $C_{rq}=C_{42}$, $r=4$, $s=4$, $q=2$.

7. К элементам столбцов 1 и 2 прибавляется $D=0$: $C_{i1}=C_{i1}+0$, $C_{i2}=C_{i2}+0$, $i=1,\dots,4$. Матрица затрат не изменяется.

8. Элемент $C_{rs}=C_{44}$ отмечается точками:

$$\begin{array}{cccc} \underline{4} & 6 & 9 & 7 \\ 13 & \underline{10} & 14 & 14 \\ \underline{9} & \underset{\cdots}{9} & 16 & 13 \\ 12 & \underline{10} & 12 & \underset{\cdots}{10}. \end{array}$$

9. В столбце $s=4$ нет ни одного подчеркнутого элемента. Значит, выполняется следующий шаг.

10. Элемент $C_{rs}=C_{44}$ подчеркивается, а с $C_{rq}=C_{42}$ снимается подчеркивание:

$$\begin{array}{cccc} \underline{4} & 6 & 9 & 7 \\ 13 & \underline{10} & 14 & 14 \\ \underline{9} & \underset{\cdots}{9} & 16 & 13 \\ 12 & 10 & 12 & \underline{10}. \end{array}$$

11. В столбце $q=2$ есть подчеркнутый элемент. Выполняется переход к шагу 15.

15. Подчеркнутые элементы есть не во всех столбцах (т.е. исполнители назначены не на все работы). Возврат к шагу 1.

1. $A=\varnothing$, $A'=\{1,2,3,4\}$.

2. $A=\{1\}$, $A'=\{2,3,4\}$.

3. $B=\{1,3\}$.

4. В строках 1 и 3 выбираются минимальные элементы из столбцов 2,3,4: $C_{12}=6$, $C_{32}=9$.

5. $d_1=C_{12}-C_{11}=6-4=2$, $d_3=C_{32}-C_{31}=9-9=0$.

6. $D=\min\{2,0\}=0$. $D=C_{32}-C_{31}$. $C_{rs}=C_{32}, C_{rq}=C_{31}$, $r=3, s=2, q=1$.

7. Ко всем элементам столбца 1 прибавляется значение $D=0$: $C_{i1}=C_{i1}+0$, $i=1,\dots,4$. Матрица затрат не изменяется.

8. Элемент $C_{rs}=C_{32}$ отмечается точками:

$$\begin{array}{cccc} \underline{4} & 6 & 9 & 7 \\ 13 & \underline{10} & 14 & 14 \\ \underline{9} & \underset{\cdots}{9} & 16 & 13 \\ 12 & 10 & 12 & \underline{10}. \end{array}$$

9. В столбце $s=2$ есть подчеркнутый элемент. Значит, он переносится во множество A: $A=\{1,2\}$, $A'=\{3,4\}$. Переход к шагу 3.

3. В столбцах 1 и 2 есть подчеркнутые элементы из строк 1,2,3: $B=\{1,2,3\}$.

4. В строках 1, 2 и 3 выбираются минимальные элементы из столбцов 3,4: $C_{14}=7$, $C_{23}=14$, $C_{34}=13$.

5. $d_1=3$, $d_2=4$, $d_3=4$.

6. $D=\min\{3,4,4\}=3$. $D=C_{14}-C_{11}$. $C_{rs}=C_{14}$. $C_{rq}=C_{11}$. $r=1$, $s=4$, $q=1$.

7. Ко всем элементам столбцов 1 и 2 прибавляется значение $D=3$: $C_{i1}=C_{i1}+3$, $C_{i2}=C_{i2}+3$, $i=1,\ldots,4$. Матрица затрат:

$$\begin{array}{cccc} \underline{7} & 9 & 9 & 7 \\ 16 & \underline{13} & 14 & 14 \\ \underline{12} & 12 & 16 & 13 \\ 15 & \dot{13} & 12 & \underline{10} \end{array}.$$

8. Элемент $C_{rs}=C_{23}$ отмечается точками:

$$\begin{array}{cccc} \underline{8} & 10 & 9 & \underset{\cdots}{8} \\ 17 & \underline{14} & \underset{\cdots}{\underline{14}} & 15 \\ \underline{13} & 13 & 16 & 14 \\ 16 & \dot{14} & 12 & \underline{11} \end{array}.$$

9. В столбце $s=3$ нет подчеркнутых элементов; выполняестя следующий шаг.

10. $C_{rs}=C_{23}$ подчеркивается, а с $C_{rq}=C_{22}$ снимается подчеркивание:

$$\begin{array}{cccc} \underline{8} & 10 & 9 & \underset{\cdots}{8} \\ 17 & 14 & \underline{14} & 15 \\ \underline{13} & 13 & 16 & 14 \\ 16 & \dot{14} & 12 & \underline{11} \end{array}.$$

8. Элемент $C_{rs}=C_{14}$ отмечается точками:

$$\begin{array}{cccc} \underline{7} & 9 & 9 & 7 \\ 16 & \underline{13} & 14 & \dot{14} \\ \underline{12} & 12 & 16 & 13 \\ 15 & \dot{13} & 12 & \underline{10} \end{array}$$

9. В столбце s=4 есть подчеркнутый элемент: $A=\{1,2,4\}$, $A'=\{3\}$, переход к шагу 3.

3. $B=\{1,2,3,4\}$.

4. $C_{13}=9$, $C_{23}=14$, $C_{33}=16$, $C_{43}=12$.

5. $d_1=2$, $d_2=1$, $d_3=4$, $d_4=2$.

6. $D=\min\{2,1,4,2\}=1$. $D=C_{23}-C_{22}$. $C_{rs}=C_{23}$, $C_{rq}=C_{22}$. $r=2$, $s=3$, $q=2$.

7. Ко всем элементам столбцов 1,2,4 прибавляется значение D=1: $C_{ij}=C_{ij}+1$, $i=1,\ldots,4$, $j=1,2,4$. Матрица затрат:

$$\begin{array}{cccc} \underline{8} & 10 & 9 & 8 \\ 17 & \underline{14} & 14 & \dot{15} \\ \underline{13} & 13 & 16 & 14 \\ 16 & \dot{14} & 12 & \underline{11} \end{array} .$$

11. В столбце q=2 нет подчеркнутых элементов. В нем есть элемент, отмеченный точками: $C_{pq}=C_{32}$. Этот элемент находится в строке 3: p=3. Подчеркнутый элемент в строке 3: $C_{pf}=C_{31}$, f=1.

12. r=3, s=2, $C_{rs}=C_{32}$.

13. q=1, $C_{rq}=C_{31}$.

14. Переход к шагу 10.

10. Элемент $C_{rs}=C_{32}$ подчеркивается, а с $C_{rq}=C_{31}$ снимается подчеркивание:

$$\begin{array}{cccc} \underline{8} & 10 & 9 & 8 \\ 17 & 14 & \underline{14} & \dot{15} \\ 13 & \underline{13} & 16 & 14 \\ 16 & 14 & 12 & \underline{11} \end{array} .$$

11. В столбце q=1 есть подчеркнутый элемент. Значит, выполняется переход к шагу 15.

15. Подчеркнутые элементы есть во всех столбцах. Конец алгоритма.

Таким образом, 1-й исполнитель должен быть назначен на 1-ю работу, 2-й - на 3-ю, 3-й - на 2-ю, 4-й - на 4-ю. Затраты на выполнение работ: $E = C_{11}+C_{23}+C_{32}+C_{44} = 4+14+9+10 = 37$.

ЛИТЕРАТУРА

Основная литература

1. Перегудов Ф. И., Тарасенко Ф. П. Введение в системный анализ.- М.: Высш. школа, 1989.- 367 с.

2. Ларичев О. И. Наука и искусство принятия решений.- М.: Наука, 1979.- 200 с.

3. Таха Х. Введение в исследование операций. В 2-х кн.- М.: Мир, 1985.- 479 с. и 496 с.

4. Банди Б. Основы линейного программирования.- М.: Радио и связь, 1989.- 176 с.

5. Балашевич В. А. Основы математического программирования.- Мн: Вышэйшая школа, 1985.- 173 с.

6. Дегтярев Ю. И. Исследование операций.- М.: Высш. школа, 1986.- 320 с.

7. Бешелев С. Д., Гурвич Ф. Г. Математико-статистические методы экспертных оценок.- М.: Статистика, 1980.- 263 с.

8. Литвак Б. Г. Экспертная информация: методы получения и анализа.- М.: Радио и связь, 1982.- 182 с.

9. Подиновский В. В., Ногин В. Д. Парето-оптимальные решения многокритериальных задач.- М.: Наука, 1982.- 256 с.

10. Половинкин А. И. Основы инженерного творчества.- М.: Машиностроение, 1988.- 368 с.

11. Смородинский С. С., Танаевский В. Г. Методы и технология системного проектирования.- Мн.: : МРТИ, 1985.- 56 с.

12. Лорьер Ж.-Л. Системы искусственного интеллекта.- М.: Мир, 1991.- 568 с.

Дополнительная литература

13. Губанов В. А., Захаров В. В., Коваленко А. Н. Введение в системный анализ: Учеб. пособие/Под ред. Л. А. Петросяна.- Л.: Изд-во Ленинградского университета, 1988.- 232 с.

14. Моисеев Н. Н. Математические задачи системного анализа.- М.: Наука, 1981.- 488 с.

15. Акофф Р. Искусство решения проблем.- М.: Мир, 1982.- 224 с.

16. Ларичев О. И. Объективные модели и субъективные решения.– М.: Наука, 1987.- 143 с.

17. Грешилов А. А. Как принять наилучшее решение в реальных условиях.–М.: Радио и связь, 1991.- 320 с.

18. Вентцель Е. С. Исследование операций: задачи, принципы, методология.– М.: Наука, 1980.- 208 с.

19. Кудрявцев Е. М. Исследование операций в задачах, алгоритмах и программах.– М.: Радио и свяаь, 1984.- 184 с.

20. Морозов В. В., Сухарев А. Г., Федоров В. В. Исследование операций в задачах и упражнениях.– М.: Высш. школа, 1986.- 287 с.

21. Банди Б. Методы оптимизации. Вводный курс.– М.: Радио и связь, 1988.- 128 с.

22. Габасов Р., Кириллова Ф. М. Методы оптимизации.– Мн: Изд-во БГУ, 1981.- 350 с.

23. Мину М. Математическое программирование. Теория и алгоритмы.– М.: Наука, 1990.- 488 с.

24. Хачиян Л. Г. Сложность задач линейного программирования.– М.: Знание, 1987.- 32 с.

25. Исследование операций. Т. 1: Методологические основы и математические методы/ Под ред. Д. Моудера и С. Элмаграби. М.: Мир, 1981.- 712 с.

26. Исследование операций. Т. 2: Модели и применения /Под ред. Д. Моудера и С. Элмаграби.– М.: Мир, 1981.- 677 с.

27. Справочник по исследованию операций /Под ред. Ф. А. Матвейчука.– М.: Воениздат, 1979.- 368 с.

28. Филлипс Д., Гарсиа-Диас А. Методы анализа сетей.– М.: Мир, 1984.- 496 с.

29. Клейнрок Л. Теория массового обслуживания.– М.: Машиностроение, 1979.- 432 с.

30. Жожикашвили В. А., Вишневский В. М. Сети массового обслуживания. Теория и применение к сетям ЭВМ.– М.: Радио и связь, 1988.- 192 с.

31. Панкова Л. А., Петровский А. М., Шнейдерман М. В. Организация экспертизы и анализ экспертной информации.– М.: Наука, 1984.- 120 с.

32. Моисеева Н. К., Карпунин М. Г. Основы теории и практики функционально-стоимостного анализа.– М.: Высш. школа, 1988.- 192 с.

33. Кини Р. Л. Принятие решений при многих критериях: предпочтения и замещения.- М.: Радио и связь, 1981. - 560 с.

34. Брахман Т. Р. Многокритериальность и выбор альтернативы в технике.- М.: Радио и связь, 1984. - 288 с.

35. Новые области применения математики/Под ред. Д. Лайтхилла.-Мн.: Вышэйшая школа, 1981. - 494 с.

36. Кузнецов О. П., Адельсон-Вельский Г, М, Дискретная математика для инженера.- М.: Энергоатомиздат, 1988. - 480 с.

37. Котлер Ф. Основы маркетинга. — Новосибирск: Наука, 1992. - 736 с.

38. Мескон М. Х., Альберт М., Хедоури Ф. Основы менеджмента.- М.: Дело, 1992. - 702 с.

СОДЕРЖАНИЕ