МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Томский государственный университет систем управления и радиоэлектроники
(ТУСУР)

Кафедра моделирования и системного анализа (МиСА)

В.Г. Баранник, Е.В. Истигечева

СИСТЕМНЫЙ АНАЛИЗ, ОПТИМИЗАЦИЯ И ПРИНЯТИЕ РЕШЕНИЙ

Учебное пособие

Томск 2014

Баранник В.Г., Истигечева Е.В. Системный анализ, оптимизация и принятие решений / Учебное пособие – Томск: Томский государственный университет систем управления и радиоэлектроники. Кафедра моделирования и системного анализа, 2014. – 99 с.

Учебное пособие рассматривает основные методы оптимизации при принятии решений: сиплекс-метод, методы искусственного базиза, методы целочисленного, нелинейного и динамического программирования, метод потенциалов, методы выбора решений в условиях риска и неопределенности. Также изложены основы теории массового обслуживания применительно к области системного анализа и принятия решений.

СОДЕРЖАНИЕ

Введение

Процессы принятия решений лежат в основе большинства задач целенаправленной деятельности человека. Особенно, в части задач управления. Изучением и развитием методов принятия решений занимаются дисциплины направления «Системный анализ». Для помощи специалистам, вынужденным по роду своей деятельности принимать решения в сложных ситуациях разрабатываются специализированные системы поддержки принятия решений. Близкие по смыслу и назначению задачи решает теория управления, которая, однако, сосредотачивается на изучении методов управления динамическими системами различной природы, информация о которых имеет достаточно структурированный вид. Подавляющее число экономических и управленческих задач имеет такой характер, когда уже заведомо можно сказать, что мы имеем дело с большими системами. Системный анализ предусматривает специальные приемы, с помощью которых большую систему, трудную для рассмотрения исследователем, можно было бы разделить на ряд малых взаимодействующих систем или подсистем. Таким образом, изучение методов принятия решений в сложных ситуациях, выглядит обязательным для любых специалистов, вынужденных делать выбор, определять развитие систем и их модернизацию/изменение.

1. Постановка задач системного анализа и принятия решений

1.1. Математическая модель задач линейного программирования (ЛП)

Любое описание некоторой решаемой задачи в виде формул, уравнений, алгоритмов называется математической моделью этой задачи.

Линейное программирование рассматривает задачи с линейной математической моделью. Это задачи, в которых требуется найти такие значения переменных $x_1, x_2, ..., x_n$, при которых некоторая линейная функция от этих переменных принимает максимальное или минимальное значение:

$$E = c_1x_1 + c_2x_2 + ... + c_nx_n \to \max/\min \tag{1.1}$$

при выполнении ограничений на переменные $x_1, x_2, ..., x_n$, заданных линейными уравнениями или неравенствами:

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \geq b_1 \\
&a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \geq b_2 \\
&... \\
&a_{p1}x_1 + a_{p2}x_2 + ... + a_{pn}x_n \geq b_p \\
&a_{(p+1)1}x_1 + a_{(p+1)2}x_2 + ... + a_{(p+1)n}x_n = b_{p+1} \\
&a_{(p+2)1}x_1 + a_{(p+2)2}x_2 + ... + a_{(p+2)n}x_n = b_{p+2} \\
&... \\
&a_{(p+q)1}x_1 + a_{(p+q)2}x_2 + ... + a_{(p+q)n}x_n = b_{p+q} \\
&a_{(p+q+1)1}x_1 + a_{(p+q+1)2}x_2 + ... + a_{(p+q+1)n}x_n \leq b_{p+q+1} \\
&a_{(p+q+2)1}x_1 + a_{(p+q+2)2}x_2 + ... + a_{(p+q+2)n}x_n \leq b_{p+q+2} \\
&... \\
&a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \leq b_m
\end{aligned} \tag{1.2}$$

Здесь p – количество ограничений, имеющих вид «не меньше»;
q – количество ограничений «равно»;
m – общее количество ограничений.

Функция (1.1) называется целевой функцией или критерием оптимизации, а соотношения типа (1.2) – системой ограничений задачи.

Если по условию задачи требуется найти такие значения переменных $x_1, x_2, ..., x_n$, при которых целевая функция (1.1) будет иметь максимальное значение, то говорят, что целевая функция подлежит максимизации (или направлена на максимум). Если требуется, чтобы целевая функция принимала минимальное значение, то говорят, что она подлежит минимизации (направлена на минимум).

В большинстве задач при этом требуется, чтобы переменные $x_1, x_2, ..., x_n$ принимали неотрицательные значения (ограничение на неотрицательность):

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

В других задачах требуется, чтобы переменные принимали целочисленные значения (ограничение на целочисленность).

Любые значения переменных $x_1, x_2, ..., x_n$, удовлетворяющие ограничениям задачи (1.2), называются допустимыми решениями (независимо от того, какое значение при этом принимает целевая функция).

Большинство задач имеет бесконечно много допустимых решений. Все множество допустимых решений представляет собой область допустимых решений (ОДР).

Допустимые значения переменных $x_1, x_2, ..., x_n$, при которых целевая функция принимает максимальное или минимальное значение (в зависимости от постановки задачи), представляют собой оптимальное решение.

1.2. Примеры задач линейного программирования

Пример 1.1. Предприятие химической промышленности выпускает соляную и серную кислоту. Выпуск одной тонны соляной кислоты приносит предприятию прибыль в размере 25 ден. ед., выпуск одной тонны серной кислоты – 40 ден. ед. Для выполнения государственного заказа необходимо выпустить не менее 200 т соляной кислоты и не менее 100 т серной кислоты. Кроме того, необходимо учитывать, что выпуск кислот связан с образованием опасных отходов. При выпуске одной тонны соляной кислоты образуется 0,5 т опасных отходов, при выпуске одной тонны серной кислоты – 1,2 т опасных отходов. Общее количество опасных отходов не должно превышать 600 т, так как превышение этого ограничения приведет к выплате предприятием крупного штрафа.

Требуется определить, сколько соляной и серной кислоты должно выпустить предприятие, чтобы получить максимальную прибыль.

Составим математическую модель задачи.

Пусть x_1 – количество выпускаемой соляной кислоты (в тоннах);

x_2 – количество выпускаемой серной кислоты.

Тогда ограничения, связанные с необходимостью выполнения государственного заказа в виде производства не менее 200 т соляной кислоты можно записать следующим образом:

$$x_1 \geq 200 .$$

Аналогично составляется ограничение на производство не менее 100 т серной кислоты:

$$x_2 \geq 100 .$$

Так как при выпуске одной тонны соляной кислоты образуется 0,5 т опасных отходов, следовательно общее количество опасных отходов при выпуске соляной кислоты определяется величиной $0,5x_1$ т и, соответственно, при выпуске серной кислоты образуется $1,2x_2$ т опасных отходов.

Таким образом, общее количество опасных отходов определяется величиной

$$0,5\,x_1 + 1,2\,x_2 \text{ т.}$$

Указанная величина по условию задачи не должна превышать 600 т. В связи с чем, можно записать следующее ограничение:

$$0,5x_1 + 1,2x_2 \leq 600$$

Кроме того, переменные x_1 и x_2 по своему физическому смыслу обозначают количества выпускаемых кислот и, следовательно, не могут принимать отрицательных значений, что эквивалентно двум условиям неотрицательности:

$$x_1 \geq 0; x_2 \geq 0 .$$

В данной задаче требуется определить выпуск кислот, при котором прибыль будет максимальной. Прибыль от выпуска одной тонны соляной кислоты составляет 25 ден. ед.; значит, прибыль от выпуска соляной кислоты составит $25x_1$ ден. ед. Прибыль от выпуска серной кислоты составит $40x_2$ ден. ед.

Таким образом, общая прибыль от выпуска кислот составит $25x_1 + 40x_2$ ден. ед. Требуется найти такие значения переменных x_1 и x_2, при которых эта величина будет максимальной. Таким образом, целевая функция для данной задачи будет иметь следующий вид:

$$E = 25x_1 + 40x_2 \to \max .$$

Приведем полную математическую модель рассматриваемой задачи:

$$x_1 \ge 200$$

$$x_2 \ge 100 \qquad (1.3)$$

$$0,5x_1 + 1,2x_2 \le 600$$

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

$$E = 25x_1 + 40x_2 \to \max . \qquad (1.4)$$

В этой задаче имеется два ограничения «больше или равно» и одно ограничение «меньше или равно». Целевая функция подлежит максимизации.

Пример 1.2. Пусть в условиях примера 1.1 из-за ужесточения требований к экологической безопасности требуется:

- свести к минимуму количество опасных отходов;
- обеспечить экономическую целесообразность производства в виде требуемой величины прибыли не менее 20 000 ден. ед.

Очевидно, математическая модель такой задачи имеет следующий вид:

$$E = 0,6x_1 + 1,2x_2 \to \min \qquad (1.5)$$

$$25x_1 + 40x_2 \ge 20000 ; \qquad (1.6)$$

$$x_1 \ge 200 ;$$

$$x_2 \ge 100 ;$$

$$x_1 \ge 0;$$

$$x_2 \ge 0.$$

В этой модели производства целевая функция (1.5) представляет собой количество опасных отходов. При этом, минимизация целевой функции осуществляется с учетом ограничения (1.6) на величину прибыли от выпуска кислот не менее 20 тыс. ден. ед. Другие четыре ограничения сохраняют в этом случае свою форму.

1.3. Графический метод решения задач ЛП

Графический метод применяется для решения задач, в которых имеются только две переменные. Для таких задач имеется возможность графически изобразить область допустимых решений (ОДР).

Примечание. Графический метод может применяться также для решения задач с любым количеством переменных, если возможно выразить все переменные задачи через какие-либо две переменные.

Как отмечено выше, ОДР – это множество значений переменных $x_1, x_2, ..., x_n$, удовлетворяющих ограничениям (1.2). Таким образом, для задач с двумя переменными ОДР представляет собой множество точек $(x_1; x_2)$ т.е. некоторую область на плоскости (обычно многоугольник). Для задач с тремя переменными ОДР представляет собой многогранник в пространстве, для задач с большим количеством переменных – некоторую область многомерного пространства. Можно доказать, что экстремум (минимум или максимум) целевой функции всегда достигается при значениях переменных $x_1, x_2, ..., x_n$, соответствующих одной из угловых точек ОДР. Другими словами, оптимальное решение всегда находится в угловой точке ОДР. Поэтому задачу линейного программирования с двумя переменными можно решить следующим образом: построить ОДР на плоскости в системе координат $(x_1; x_2)$, определить все угловые точки ОДР, вычислить значения целевой функции в этих точках и выбрать оптимальное решение. Решим графическим методом задачу из примера 1.1 Решение показано на рис. 1.1.

Ограничение $x_1 \geq 200$ задается вертикальной линией $x_1 = 200$. Все точки $(x_1; x_2)$, расположенные справа от этой линии, удовлетворяют ограничению $x_1 \geq 200$, расположенные слева — не удовлетворяют.

Ограничение $x_2 \geq 100$ задается горизонтальной линией $x_2 = 100$. Все точки, расположенные сверху от этой линии, удовлетворяют ограничению $x_2 \geq 100$, расположенные снизу — не удовлетворяют.

Для построения линии, задающей ограничение $0,5x_1 + 1,2x_2 \leq 600$, удобно записать его в виде равенства:

$$0,5x_1 + 1,2x_2 = 600.$$

Выразим одну из переменных через другую: $x_2 = -0,417x_1 + 500$. Это уравнение прямой. Построим эту прямую. Она разбивает координатную плоскость на две полуплоскости. В одной из этих полуплоскостей находятся точки, удовлетворяющие ограничению, в другой — не удовлетворяющие. Чтобы найти полуплоскость, удовлетворяющую ограничению $0,5x_1 + 1,2x_2 \leq 600$, подставим в него координаты любой точки, например, (0; 0). Для этой точки ограничение выполняется. Значит, она находится в полуплоскости, удовлетворяющей ограничению.

Пересечение всех полуплоскостей, удовлетворяющих ограничениям задачи, представляет собой ОДР. На рис. 1.1 она выделена серым цветом.

Оптимальное решение находится в одной из угловых точек ОДР (на рис. 1.1 они обозначены как A, B, C). Эти точки можно найти из построенного графика или путем решения соответствующих систем из двух уравнений. Найдем значения целевой функции (1.4) в этих точках:

E(A) = 25·200 + 40·100 = 9000;
E(B) = 25·200 + 40·416,67 = 21 666,8;
E(C) = 25·960 + 40·100 = 28000.

Таким образом, оптимальное решение находится в точке:

C = (960; 100).

Это означает, что производство характеризуется следующими показателями:

- выпуск соляной кислоты – 960 т ;
- выпуск серной кислоты – 100 т ;
- величина прибыли – 28 000 ден. ед. ;
- общее количество опасных отходов в объеме 0,5·960 + 1,2·100 = 600 т.

Очевидно, ОДР задачи из примера 1.2 имеет только две угловые точки (A и B). Найдем для них значения целевой функции (1.5):

E(A) = 0,5·200 + 1,2·375 = 550;
E(B) = 0,5·640 + 1,2·100 = 440.

Таким образом, оптимальное решение находится в точке B = (640; 100).

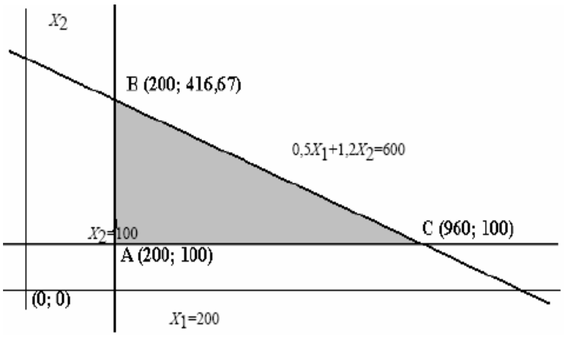

Рис. 1.1. Решение примера 1.1

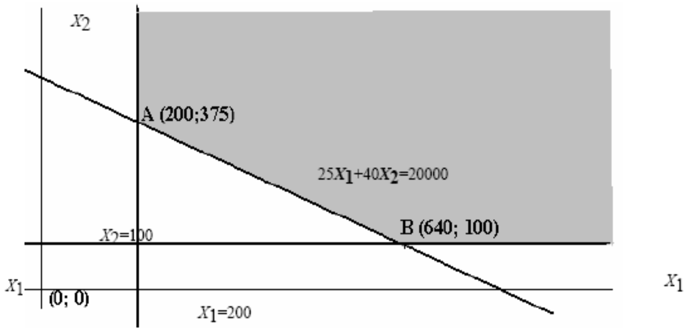

Рис. 1.2. Решение примера 1.2

Это означает, что предприятию следует выпустить только 640 т соляной кислоты и только 100 т серной кислоты. При этом образуется 440 т опасных отходов, а величина прибыли от производства кислот оценивается как $25 \cdot 640 + 40 \cdot 100 = 20\ 000$ ден. ед.

1.4. Преобразование задач ЛП к канонической форме

Для большинства методов решения задач ЛП требуется предварительно привести задачу к стандартной форме. Задача (или ее математическая модель) представлена в стандартной форме, если она соответствует следующим условиям:

- целевая функция подлежит максимизации;
- все ограничения имеют вид равенств;
- на все переменные накладываются ограничения неотрицательности.

Если целевая функция задачи подлежит минимизации, то для перехода к целевой функции, подлежащей максимизации, необходимо умножить исходную целевую функцию на —1. Доказано, что максимальное значение любой функции Е всегда равно минимальному значению функции —Е, взятому с обратным знаком.

Для преобразования ограничения «больше или равно» в равенство (т.е. в ограничение «равно») необходимо вычесть из левой части ограничения дополнительную переменную. Для преобразования ограничения «меньше или равно» в равенство необходимо прибавить к левой части ограничения дополнительную переменную.

На все переменные, используемые для приведения задачи к стандартной форме, накладываются ограничения неотрицательности.

Переменные, вычитаемые из ограничений «больше или равно» для их приведения к стандартной форме, называются избыточными.

Переменные, прибавляемые к ограничениям «меньше или равно» — остаточными. Если на какую-либо переменную не накладывается ограничение неотрицательности, то она заменяется на разность двух переменных, каждая из которых должна быть неотрицательной. Таким образом, если некоторая переменная x_j по своему физическому смыслу может принимать как положительные, так и отрицательные значения, то во всех ограничениях и в целевой функции ее следует заменить на разность двух переменных: $x'_j - x''_j$. На эти переменные накладываются ограничения неотрицательности: $x'_j \geq 0, x''_j \geq 0$.

Приведем к стандартной форме задачу из примера 1.1. Из ограничений «больше или равно» необходимо вычесть избыточные переменные, к ограничению «меньше или равно» — прибавить остаточную переменную. Целевая функция задачи подлежит максимизации, и на все переменные накладывается ограничение неотрицательности; поэтому никаких других преобразований не требуется. Математическая модель задачи в стандартной форме будет иметь следующий вид:

$$x_1 - x_3 = 200$$
$$x_2 - x_4 = 100$$
$$0,5x_1 + 1,2x_2 + x_5 = 600$$
$$x_j \geq 0,\ j = 1,...,5.$$
$$E = 25x_1 + 40x_2 \to \max.$$

Здесь переменные x_3 и x_4 — избыточные, x_5 — остаточная.

Примечание. Все переменные, которые вводятся в математическую модель для ее приведения к стандартной форме, имеют физический смысл. Так, в рассмотренном примере переменные x_3 и x_4 обозначают количество кислот, которое будет выпущено сверх государственного заказа. Переменная x_5 обозначает, насколько количество опасных отходов, образующихся при производстве кислот, будет меньше максимально допустимой величины

(600 т).

Приведем к стандартной форме задачу из примера 1.2. В ней имеются три ограничения «больше или равно». В каждое из них необходимо ввести избыточную переменную. Целевая функция задачи подлежит минимизации; ее необходимо умножить на —1, чтобы перейти к целевой функции, подлежащей максимизации. Математическая модель задачи в стандартной форме будет иметь следующий вид:

$$x_1 - x_3 = 200$$

$$x_2 - x_4 = 100$$

$$25x_1 + 40x_2 - x_5 = 20000$$

$$x_j \geq 0,\ j = 1,...,5.$$

$$-E = -0,5x_1 - 1,2x_2 \to \max.$$

2. Симплекс – метод решения задач ЛП

2.1. Задача планирования производства при ограниченных ресурсах.

Основным методом решения задач ЛП является симплексный метод (симплекс – метод). Вычислительная схема алгоритма симплекс–метода может быть проиллюстрирована на примере решения следующей задачи.

Пример 2.1. Один из цехов машиностроительного предприятия выпускает изделия двух видов: корпуса и задвижки. Для производства этих изделий требуются три вида сырья: алюминий, сталь и пластмасса. На выпуск одного корпуса расходуется 20 кг алюминия, 10 кг стали и 5 кг пластмассы. На выпуск одной задвижки расходуется 5 кг алюминия, 5 кг стали и 20 кг пластмассы. Запасы ресурсов ограничены: за рабочую смену цех может израсходовать не более 200 кг алюминия, 250 кг стали и 500 кг пластмассы. Выпуск одного корпуса приносит предприятию прибыль в размере 100 (ден. ед.), одной задвижки – 300 ден. ед. Требуется составить оптимальный план работы цеха, т.е. найти, сколько корпусов и задвижек требуется выпускать, чтобы получить максимальную прибыль (при соблюдении ограничений на ресурсы).

Пусть x_1 – количество выпускаемых корпусов;

x_2 – количество выпускаемых задвижек.

Так как на выпуск одного корпуса идет 20 кг алюминия, то расход алюминия на выпуск всех корпусов составит $20x_1$ кг.

Очевидно, что на выпуск задвижек будет израсходовано $5x_2$ кг алюминия.

Таким образом, общий расход алюминия составит

$$20x_1 + 5x_2 \text{ кг.}$$

В связи с тем, что указанная величина не должна превышать 200 кг, то ограничения на расход алюминия за смену можно записать в следующем виде:

$$20x_1 + 5x_2 \leq 200.$$

Аналогично составляются ограничения на расход стали:

$$10x_1 + 5x_2 \leq 250,$$

а также ограничения на расход пластмассы:

$$5x_1 + 20x_2 \leq 500.$$

Кроме того, переменные x_1 и x_2 по своему физическому смыслу не могут принимать отрицательных значений, так как они обозначают количество изделий. Поэтому необходимо указать ограничения неотрицательности:

$$x_1 \geq 0, x_2 \geq 0.$$

В данной задаче требуется определить количество выпускаемых изделий, при котором величина прибыли от их производства будет максимальной. Так как прибыль от выпуска одного корпуса составляет 100 (ден.ед.), то прибыль от выпуска корпусов составит $100\,x_1$ (ден.ед.) Прибыль от выпуска задвижек составит $300\,x_2$ (ден.ед.) Таким образом, общая прибыль от выпуска всех изделий составит $100\,x_1$ + $300\,x_2$ (ден.ед.) Требуется найти такие значения переменных x_1 и x_2, при которых величина прибыли будет максимальной.
Таким образом, целевая функция для данной задачи будет иметь следующий вид:

$$E = 100\,x_1 + 300\,x_2 \to \max.$$

Приведем полную математическую модель рассматриваемой задачи:

$$\begin{aligned} &20x_1 + 5x_2 \le 200\,; \\ &10x_1 + 5x_2 \le 250\,; \\ &5x_1 + 20x_2 \le 500\,; \\ &x_1 > 0\,; \\ &x_2 > 0\,; \end{aligned} \qquad (2.1)$$

$$E = 100\,x_1 + 300\,x_2 \to \max \qquad (2.2)$$

Примечание. В данной задаче на переменные x_1 и x_2 накладывается также ограничение целочисленности: они должны принимать только целые значения, так как обозначают количество изделий.

Первый способ решения задачи из Примера 2.1 сравнительно легко можно реализовать на основе графического метода (Рис. 2.1).

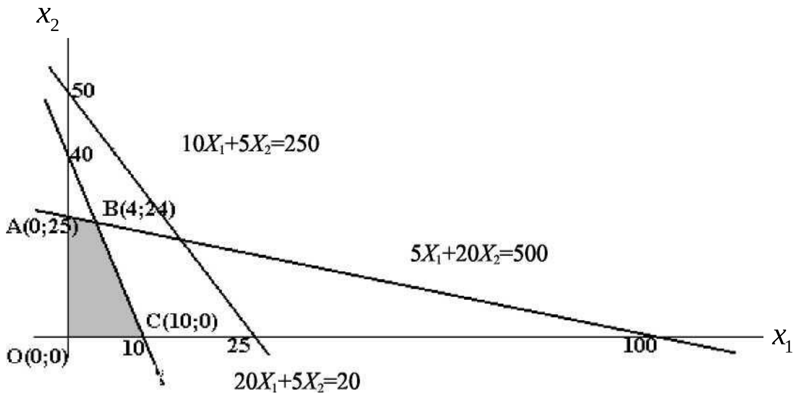

Рис. 2.1. Решение задачи из Примера 2.1

Найдем значения целевой функции для угловых точек ОДР:

$$E(O) = 100 \cdot 0 + 300 \cdot 0 \ = 0$$

$$E(A) = 100 \cdot 0 + 300 \cdot 25 = 7500,$$

$$E(B) = 100 \cdot 4 + 300 \cdot 24 = 7600,$$

$$E(C) = 100 \cdot 10 + 300 \cdot 0 \ = 1000.$$

Таким образом, оптимальное решение находится в точке B = (4; 24). Это означает, что цех должен выпускать за смену 4 корпуса и 24 задвижки. Прибыль при этом составит 7600 ден. ед.

Второй способ решения этой же задачи может быть реализован на основе симплекс-метода, позволяющего решать задачи с любым количеством переменных.

Для решения задачи симплекс-методом требуется привести ее к стандартной форме, как показано в п.1.4. Все ограничения задачи имеют вид «меньше или равно». Их необходимо преобразовать в равенства. Для этого требуется добавить в каждое ограничение дополнительную (остаточную) переменную. Математическая модель задачи в стандартной форме будет иметь следующий вид:

$$20x_1 + 5x_2 + x_3 = 200 \qquad (2.3)$$

$$10x_1 + 5x_2 + x_4 = 250$$

$$5x_1 + 20x_2 + x_5 = 500$$

$$x_j \geq 0,\ j = 1,...,5.$$

$$E = 100\,x_1 + 300\,x_2 \rightarrow \max \qquad (2.4)$$

Здесь x_3, x_4, x_5— остаточные переменные. Их физический смысл будет показан в п. 2.4.

2.2. Принцип работы симплекс–метода

Симплекс-метод позволяет решать задачи ЛП любой размерности, т.е. с любым количеством переменных. Решение задач ЛП на основе симплекс–метода состоит в целенаправленном переборе угловых точек ОДР в направлении улучшения значения целевой функции.

Известно, что экстремум (минимум или максимум) целевой функции всегда достигается при значениях переменных $x_1, x_2, ..., x_n$, соответствующих одной из угловых точек ОДР. Другими словами, оптимальное решение всегда находится в угловой точке ОДР.

Принцип работы симплекс–метода состоит в следующем. Находится какое-либо допустимое решение, соответствующее одной из угловых точек ОДР. Проверяются смежные с ней угловые точки ОДР. Под смежной здесь понимается угловая точка, расположенная на той же границе ОДР, что и текущая угловая точка (для двухмерной ОДР – на той же стороне многоугольника, для трехмерной – на том же ребре многогранника, и т.д.). Если ни в одной из смежных угловых точек значение целевой функции не улучшается, то решение задачи завершается; текущая угловая точка ОДР соответствует оптимальному решению задачи. Если имеются смежные угловые точки ОДР, для которых значение целевой функции улучшается, то выполняется переход в ту из них, для которой достигается наиболее быстрое улучшение значения целевой функции. Для новой угловой точки ОДР процесс повторяется, т.е. проверяются смежные угловые точки. Перебор угловых точек происходит до тех пор, пока не будет найдено оптимальное решение, т.е. пока не будет достигнута угловая точка ОДР, для которой ни в одной из смежных точек значение целевой функции не улучшается.

Поиск решения на основе симплекс–метода реализуется с помощью симплекс–таблиц.

Основные этапы реализации симплекс–метода следующие.

1. Задача линейного программирования приводится к стандартной форме.
2. Определяется начальное допустимое решение (начальная угловая точка ОДР).
3. Строится исходная симплекс–таблица. Выполняются преобразования симплекс–таблиц, соответствующие перебору угловых точек ОДР, до получения оптимального решения.

Реализация симплекс–метода существенно различается в зависимости от вида математической модели задачи. В данном разделе рассматривается реализация симплекс–метода для случая, когда математическая модель задачи состоит только из ограничений «меньше или равно», и целевая функция подлежит максимизации (как в примере 2.1). Реализация симплекс–метода для задач с математической моделью любого вида рассматривается в разделе 3.

2.3. Определение начального допустимого решения

В задаче, представленной в стандартной форме, количество переменных обычно больше, чем количество ограничений. Так, в стандартной форме для примера 2.1 имеются три ограничения (m = 3) и пять переменных (k = 5). Для определения начального решения (k – m) переменных принимаются равными нулю. Тогда в системе из m равенств остается m переменных (неизвестных); в этом случае их значения можно определить однозначно. Эти значения используются в качестве начального решения задачи. Переменные, значения которых принимаются равными нулю, называются небазисными, остальные – базисными. Количество базисных переменных (переменных, составляющих базис), всегда равно количеству ограничений (m).

Начальный базис легко определить, если в каждом ограничении имеется переменная, входящая в это ограничение с коэффициентом, равным единице, и при этом не входящая ни в одно из других ограничений. Эти переменные принимаются в качестве базисных. Все остальные (небазисные) переменные принимаются равными нулю. Таким образом, базисные переменные принимают значения, равные правым частям ограничений.

Такой способ определения начального базиса наиболее удобен при решении задач, для которых математическая модель состоит только из ограничений «меньше или равно». В таких задачах после приведения к стандартной форме в каждом ограничении имеется переменная, входящая в данное ограничение с коэффициентом, равным единице, и не входящая ни в одно из других ограничений. Эти переменные – остаточные переменные, введенные при приведении задачи к стандартной форме. Эти переменные принимаются в качестве базисных. Все остальные переменные (т.е. переменные, входившие в исходную математическую модель задачи), принимаются в качестве небазисных, т.е. равных нулю. Таким образом, в качестве начального допустимого решения (начальной угловой точки ОДР) принимается начало координат, т.е. решение, в котором все исходные переменные математической модели равны нулю: $x_1 = x_2 = ... = x_n = 0$.

Если базисные переменные присутствуют не во всех ограничениях задачи, то решение

$$x_1 = x_2 = ... = x_n = 0$$

обычно оказывается недопустимым (не соответствует системе ограничений). В этом случае начало координат не может использоваться в качестве начального допустимого решения (начальной угловой точки ОДР). Для поиска начального допустимого решения в таких случаях используются специальные методы (см. раздел 3). Обычно это требуется для задач, в которых имеются ограничения «не меньше» или «равно». Найдем начальное допустимое решение для примера 2.1. Для задачи, приведенной к стандартной форме (2.3), (2.4), в качестве базисных переменных следует выбрать переменные x_3, x_4, x_5, так как каждая из них входит только в одно ограничение с коэффициентом, равным единице, и не входит в другие ограничения. Базисные переменные имеются во всех ограничениях задачи. Переменные x_1 и x_2 принимаются равными нулю, т.е. небазисными. Таким образом, начальное решение задачи следующее:

$$x_1 = 0, x_2 = 0, x_3 = 200, x_4 = 250, x_5 = 500.$$

Это решение является допустимым, так как значения $x_1 = x_2 = 0$ соответствуют системе ограничений (2.1). Таким образом, в качестве начальной угловой точки ОДР выбрано начало координат.

Выбранное решение явно не является оптимальным, так как целевая функция $E = 100x_1 + 300x_2$ при этом равна нулю. По своему смыслу это решение означает, что никакие изделия не выпускаются.

2.4. Определение оптимального решения на основе симплекс–таблиц

Как отмечено выше, поиск оптимального решения на основе симплекс–метода состоит в целенаправленном переборе смежных угловых точек ОДР в направлении улучшения значения целевой функции. Можно доказать, что переход из одной угловой точки ОДР в другую (смежную) соответствует замене одной переменной в базисе. Такая замена означает, что одна из небазисных переменных (имевшая нулевое значение) включается в базис, т.е. увеличивается, а одна из базисных переменных уменьшается до нуля, т.е. исключается из базиса. Выбор таких переменных выполняется по определенным правилам, обеспечивающим максимально быстрое увеличение целевой функции.

Рассмотрим алгоритм поиска оптимального решения на основе симплекс–таблиц для примера 2.1.

1. Строится исходная симплекс-таблица. Общий вид симплекс–таблицы показан в табл. 2.1.

Таблица 2.1

Базис	x	x_2	$\dots$	x_n	x_{n+1}	x_{n+2}	$\dots$	x_k	Решение
E	$-c_1$	$-c_2$	$\dots$	$-c_n$	0	0	0	0	0
x_{n+1}	a_{11}	a_{12}	$\dots$	a_{1n}	1	0	0	0	b_1
x_{n+2}	a_{21}	a_{22}	$\dots$	a_{2n}	0	1	0	0	b_2
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_k	a_{m1}	a_{m2}	$\dots$	a_{mn}	0	0	0	1	b_m

Симплекс-таблица строится по следующим правилам:

- в первой строке перечисляются все переменные задачи, как исходные $(x_1, x_2, ..., x_n)$, так и дополнительные, введенные при приведении к стандартной форме $(x_{n+1}, x_{n+2}, ..., x_k)$. Для задач, содержащих только ограничения «меньше или равно», дополнительные переменные
$$x_{n+1}, x_{n+2}, ..., x_k \text{ –}$$
это остаточные переменные;
- в первой колонке таблицы («Базис») перечисляются переменные, составляющие начальный базис задачи. Их количество всегда равно количеству ограничений. Для задач, содержащих только ограничения «меньше или равно», начальный базис состоит из остаточных переменных $x_{n+1}, x_{n+2}, ..., x_k$. В этой же колонке указывается обозначение целевой функции E;
- в строке целевой функции указываются коэффициенты целевой функции с обратным знаком. Для переменных, не входящих в целевую функцию (например, для остаточных переменных $x_{n+1}, x_{n+2}, ..., x_k$), указываются нули;
- в строках базисных переменных указываются коэффициенты ограничений, в которые входят эти переменные. Для переменных, не входящих в ограничения, указываются нулевые коэффициенты;
- в последнем столбце («Решение») указываются значения базисных переменных (они равны правым частям ограничений), а также начальное значение целевой функции (0).

Если таблица построена правильно, то в столбце каждой базисной переменной должна присутствовать только одна единица (в строке ограничения, в которое входит эта переменная); остальные коэффициенты – нулевые.

Исходная симплекс–таблица для примера 2.1 приведена в табл. 2.2.

2. Проверяется условие окончания решения задачи. Если в строке целевой функции (Е–строке) все коэффициенты неотрицательны, это означает, что оптимальное решение найдено. В противном случае выполняется следующий шаг.

Таблица 2.2

Базис	x_1	x_2	x_3	x_4	x_5	Решение
Е	-100	-300	0	0	0	0
x_3	20	5	1	0	0	200
x_4	10	5	0	1	0	250
x_5	5	20	0	0	1	500

3. Определяется переменная для включения в базис. В качестве такой переменной выбирается переменная, которой соответствует максимальный по модулю отрицательный коэффициент в Е–строке. Включение в базис (т.е. увеличение) такой переменной приводит к наиболее быстрому росту целевой функции. Столбец переменной, выбранной для включения в базис, называется ведущим (разрешающим).

Для рассматриваемого примера в базис необходимо включить переменную Х% так как ей соответствует максимальный по модулю отрицательный коэффициент Е–строки (– 300). Это означает увеличение выпуска задвижек. Из условия задачи и целевой функции видно, что увеличение выпуска задвижек приводит к более быстрому росту целевой функции, чем увеличение выпуска корпусов: выпуск каждой задвижки увеличивает целевую функцию (прибыль) на 300 (ден.ед.), а выпуск каждого корпуса – только на 100 (ден.ед.)

Примечание. Если в Е–строке имеется несколько максимальных по модулю отрицательных элементов (равных между собой), то для включения в базис можно выбирать любую из соответствующих переменных.

4. Определяется переменная для исключения из базиса. Для этого вычисляются отношения значений базисных переменных (указанных в столбце «Решение») к соответствующим элементам ведущего столбца. Такие отношения (называемые симплексными отношениями) вычисляются только для положительных коэффициентов ведущего столбца. Переменная, которой соответствует минимальное симплексное отношение, исключается из базиса.

Строка переменной, выбранной для исключения из базиса, называется ведущей (разрешающей). Элемент на пересечении ведущей строки и столбца называется ведущим (разрешающим) элементом.

Найдем симплексные отношения для рассматриваемого примера:

$$200/5 = 40; \quad 250/5 = 50; \quad 500/20 = 25.$$

Минимальное симплексное отношение получено для последней строки, соответствующей базисной переменной x_5. Значит, эта переменная исключается из базиса (становится равной нулю).

Такой способ определения переменной, исключаемой из базиса, имеет следующее обоснование. При включении в базис новой переменной ее значение увеличивается. Чтобы по-прежнему соблюдались ограничения (2.1), необходимо в каждом из ограничений уменьшать базисные переменные. Увеличение переменной, включаемой в базис, возможно только до тех пор, пока одна из базисных переменных не станет равной нулю. Минимальное симплексное отношение показывает, какая из базисных переменных первой уменьшается до нуля (т.е. исключается из базиса). Так, для рассматриваемого примера, при увеличении переменной x_2 (т.е. при ее включении в базис) для соблюдения ограничений (2.1) необходимо уменьшать переменные x_3, x_4, x_5. Например, при каждом увеличении переменной x_2 на единицу необходимо уменьшать переменную x_3 на 5, x_4 - также на 5, x_5 -

на 20. Минимальное симплексное отношение показывает, что при увеличении x_2 переменная x_5 первой достигнет нулевого значения. Смысл симплексных отношений для данной задачи следующий: они показывают, что имеющегося запаса алюминия (200 кг) хватит на выпуск 40 задвижек, запаса стали (250 кг) - на 50 задвижек, запаса пластмассы (500 кг) - на 25 задвижек. Таким образом, запас пластмассы будет израсходован первым, поэтому из базиса исключается переменная x_5. Ниже будет показано, что переменная x_5 обозначает остаток запаса пластмассы.

Примечания:

1. Если имеется несколько минимальных симплексных отношений (равных между собой), то для исключения из базиса можно выбирать любую из соответствующих переменных.
2. Если все элементы ведущего столбца оказываются отрицательными или равными нулю (т.е. невозможно вычислить ни одного симплексного отношения), это означает, что переменную, включаемую в базис, можно увеличивать на любую величину, не нарушая ни одного из ограничений задачи. Целевая функция при этом также может увеличиваться бесконечно. Обычно такой случай означает, что допущена ошибка в постановке задачи или в математической модели (не учтено некоторое ограничение).

5. Выполняется преобразование симплекс-таблицы по следующим правилам:

- в столбце «Базис» вместо переменной, исключенной из базиса на шаге 4, указывается переменная, включенная в базис на шаге 3;
- все элементы ведущей строки делятся на ведущий элемент;
- все элементы ведущего столбца (кроме ведущего элемента) заменяются нулями;
- все остальные элементы таблицы (включая Е-строку и столбец «Решение») пересчитываются по «правилу прямоугольника».

Этот пересчет выполняется следующим образом:

- ведущий и пересчитываемый элемент образуют диагональ прямоугольника;
- находится произведение ведущего и пересчитываемого прямоугольника;
- из этого произведения вычитается произведение элементов, образующих противоположную диагональ прямоугольника;
- результат делится на ведущий элемент.

Выполним пересчет симплекс-таблицы, приведенной в табл. 2.2. В столбце «Базис» x_5 заменяется на x_2. Все элементы ведущей строки делятся на ведущий элемент, равный 20. Ведущий столбец (x_2) заполняется нулями. Все остальные элементы пересчитываются по «правилу прямоугольника». Например, коэффициент на пересечении Е–строки и столбца x_1 пересчитывается следующим образом:

$$(20\cdot(-100) - (-300)\cdot 5)/20 = -25.$$

Коэффициент на пересечении строки X3 и столбца X5 пересчитывается следующим образом: $(20\cdot 0 - 5\cdot 1)/20 = -0{,}25$.

Расчет этих элементов иллюстрируется в табл. 2.2.

Полученная симплекс-таблица приведена в табл. 2.3.

Табл. 2.2. Примеры расчетов по «правилу прямоугольника»

Базис	x_1	x_2	x_3	x_4	x_5	Решение
E	-100	-300	0	0	0	0
x_3	20	5	1	0	0	200
x_4	10	5	0	1	0	250
x_5	5	20	0	0	1	500

Таблица 2.3

Базис	x_1	x_2	x_3	x_4	x_5	Решение
E	-25,00	0,00	0,00	0,00	15,00	7500,00
x_3	18,75	0,00	1,0	0,00	-0,25	75,00
x_4	8,75	0,00	0,00	1,00	-0,25	125,00
x_5	0,25	1,00	0,0	0,0	0,05	25,00

6. Выполняется возврат к шагу 2.

Для рассматриваемого примера на шаге 5 получено следующее решение (табл. 2.3): $x_2 = 25; x_3 = 75; x_4 = 125; x_1 = x_2 = 0$ (так как переменные x_1 и x_3 – небазисные). Это решение соответствует угловой точке ОДР, обозначенной на рис.2.1 как A $(x_1 = 0; x_2 = 25)$. Видно, что полученное решение (табл. 2.3) не является оптимальным, так как в Е–строке имеется отрицательный элемент (– 25). Поэтому алгоритм продолжается. Определяется переменная для включения в базис (шаг 3). Это переменная x_1, так как только для этой переменной в Е–строке содержится отрицательный элемент. Определяется переменная, исключаемая из базиса (шаг 4). Для этого вычисляются симплексные отношения:

75/18,75 = 4; 125/8,75 = 14,29; 25/0,25 = 100.

Минимальное симплексное отношение соответствует переменной x_3; она исключается из базиса. Таким образом, ведущий столбец $-x_1$, ведущая строка $-x_3$, ведущий элемент равен 18,75. Симплекс–таблица преобразуется по правилам, приведенным на шаге 5. Полученная симплекс-таблица показана в табл. 2.4.

Таблица 2.4

Базис	x_1	x_2	x_3	x_4	x_5	Решение
E	0,00	0,00	1,33	0,00	14,67	7600,00
x_1	1,00	0,00	0,05	0,00	– 0,01	4,00
x_4	0,00	0,00	– 0,47	1,00	– 0,13	90,00
x_2	0,00	1,00	– 0,01	0,010	0,05	24,00

Выполняется возврат к шагу 2 (проверка оптимальности полученного решения). Решение, полученное в табл. 2.4, оптимально, так как в Е–строке нет отрицательных элементов.

Полученное решение соответствует угловой точке ОДР, обозначенной на рис. 2.1 как B $(x_1 = 4; x_2 = 24)$.

По окончании алгоритма в столбце «Решение» находятся оптимальные значения базисных переменных, а также значение целевой функции, соответствующее полученному решению. Оптимальные значения небазисных переменных равны нулю.

Таким образом, для задачи из примера 2.1 получено следующее оптимальное решение:

$$x_1 = 4; x_2 = 24; x_3 = 0; x_4 = 90; x_5 = 0; E = 7600.$$

Значения переменных $x_1 = 4, x_2 = 24$ показывают, что цех должен выпускать за смену 4 корпуса и 24 задвижки. В этом случае будет получена максимальная прибыль в размере 7600 ден. ед. (значение целевой функции).

Остаточные переменные x_3, x_4, x_5 также имеют физический смысл. Например, из системы ограничений (2.1) видно, что величина $20x_1 + 5x_2$ представляет собой расход алюминия на выпуск всех изделий, а величина 200 (правая часть ограничения) — имеющийся запас алюминия. Переменная x_3 представляет собой разность этих величин, т.е. неизрасходованный остаток запаса алюминия. Так как $x_3 = 0$, значит, весь запас алюминия (200 кг) расходуется на выпуск изделий. Аналогично можно показать, что переменная x_4 представляет собой неизрасходованный остаток стали, а x_5 – пластмассы. Таким образом, остается неизрасходованным 90 кг стали (расход стали на выпуск всех изделий составит 250 – 90 = 160 кг). Неизрасходованный остаток пластмассы равен нулю, значит, все 500 кг пластмассы расходуются на выпуск изделий.

Примечания:

1. Смысл остаточных переменных (как и остальных переменных, используемых в математической модели) полностью зависит от постановки задачи.

2. По результатам решения задачи переменные x_1 и x_2 приняли целочисленные значения. Если бы они оказались дробными, то потребовалось бы использовать специальные методы (см. раздел 4) для получения целочисленного решения, так как эти переменные обозначают количество изделий.

2.5. Решение задач линейного программирования средствами Excel

Табличный процессор Excel имеет развитые средства, позволяющие решать разнообразные задачи оптимизации, в том числе задачи линейного и нелинейного математического программирования. Решим задачу из примера 2.1, используя табличный процессор Excel.

Предположим, что желательно получить результаты (значения переменных x_1 и x_2) в ячейках B2 и C2 (конечно, можно использовать и любые другие ячейки). В ячейках B3 и C3 введем коэффициенты целевой функции (100 и 300). В ячейке D3 введем формулу целевой функции: =СУММПРОИЗВ(B2:C3;B2:C2)

В ячейках B4 и C4 введем коэффициенты первого ограничения (на расход алюминия): 20 и 5. В ячейке D4 введем формулу этого ограничения: =СУММПРОИЗВ(B4:C4;B2:C2)

В ячейке F4 введем правую часть этого ограничения: 200.

Аналогично в ячейках B5 и C5 введем коэффициенты ограничения на расход стали (20 и 5), в ячейке D5 — формулу этого ограничения (=СУММПРО- ИЗВ(B5:C5;B2:C2)), в ячейке F5 - правую часть (250). В ячейках B6 и C6 введем коэффициенты ограничения на расход пластмассы (5 и 20), в ячейке D6 — формулу этого ограничения (=СУММПРОИЗВ(B6:C6;B2:C2)), в ячейке F6 - правую часть (500).

Примечание. Вводить описание математической модели в рабочий лист Excel можно и по-другому. Например, для ввода целевой функции достаточно в любой ячейке указать формулу: = 100 · B2 + 300 · C2. Для ввода первого ограничения достаточно в одной из ячеек указать формулу = 20 · B2 + 5 · C2, а в другой - правую часть ограничения (200). Однако показанный выше способ позволяет при необходимости легко внести изменения в постановку задачи.

Укажем также некоторые поясняющие надписи и обозначения (хотя это и необязательно). Рабочий лист будет иметь примерно такой вид, как показано на рис. 2.3.

	A	B	C	D	E		
1		x_1	x_2				
2	Решение						
3	Целевая	100	300	0	->	max	

	функция					
4	Ограничения	20	5	0	<=	200
5		10	5	0	<=	250
i		5	20	0	<=	500

Рис. 2.3. Рабочий лист Excel для решения примера 2.1

Примечание. Подписи и обозначения на рабочем листе (x_1, x_2 , ->, >= и т.д.), показанные на рис. 2.3, необязательны. Значения 0 в ячейках D3 -D6 получены автоматически для начальных значений переменных, равных нулю.

Для решения задачи из меню «Сервис» выберем элемент «Поиск решения». В поле «Установить целевую ячейку» указывается ячейка D3, где находится формула целевой функции. Используя переключатели, необходимо указать, что требуется установить ячейку D3 «равной максимальному значению» (так как целевая функция в этой задаче подлежит максимизации). В поле «Изменяя ячейки» указываются ячейки, в которых должны находиться значения переменных: B2:C2.

В области «Ограничения» указываются ограничения. Для начала их ввода требуется нажать кнопку «Добавить». На экран выводится окно «Добавление ограничения». В этом окне в поле «Ссылка на ячейку» указывается ячейка, в которой находится левая часть (формула) ограничения, а в поле «Ограничение» — правая часть ограничения (число или ссылка на ячейку, где находится правая часть ограничения). Чтобы задать первое из ограничений (на расход алюминия), требуется в поле «Ссылка на ячейку» указать ячейку D4, выбрать знак ограничения (<=), а в поле «Ограничение» указать ячейку F4. Для ввода ограничения требуется нажать кнопку «Добавить». Аналогично вводятся остальные ограничения. Для ввода ограничения на расход стали требуется в поле «Ссылка на ячейку» ввести D5, в поле знака ограничения — знак <=, в поле «Ограничение» — F5. Для ввода ограничения на расход пластмассы требуется в поле «Ссылка на ячейку» ввести D6, в поле знака ограничения — знак <=, в поле «Ограничение» — F6. Кроме того, требуется ввести ограничение на неотрицательность всех переменных: B2:C2>=0. Необходимо также указать, что переменные, определяемые в задаче, должны принимать целочисленные значения (так как они обозначают количество изделий). Для этого необходимо в поле «Ссылка на ячейку» указать B2:C2, а в поле знака ограничения выбрать отметку «цел». По окончании ввода всех ограничений требуется нажать OK.

Для решения задачи следует нажать кнопку «Выполнить». После появления окна с сообщением о том, что решение найдено, следует установить переключатель «Сохранить найденное решение» и нажать OK. Рабочий лист с результатами решения будет иметь примерно такой вид, как показано на рис. 2.4.

	A	B	C	D	E	F
1		x_1	x_2			
2	Решение	4	24			
3	Целевая функция	100	300	7600	->	так
4	Ограничения	20	5	200	<=	200
5		10	5	160	<=	250
6		5	20	500	<=	500
7						

Рис. 2.4. Рабочий лист Excel с результатами решения примера 2.1

В ячейках B2 и C2 получены оптимальные значения переменных, в ячейке D3 —

оптимальное значение целевой функции. Таким образом, за смену цех должен выпускать 4 корпуса и 24 задвижки. Прибыль составит 7600 ден. ед. Эти результаты совпадают с результатами, полученными с помощью графического метода и симплекс-метода.

В ячейках D4, D5 и D6 находятся значения левых частей ограничений (т.е. величины, получаемые при подстановке оптимальных значений переменных в ограничения). В данной задаче эти величины обозначают расход ресурсов. Таким образом, видно, что на выпуск 4 корпусов и 24 задвижек будет израсходовано 200 кг алюминия, 160 кг стали и 500 кг пластмассы.

2.6. Анализ оптимального решения на чувствительность

Анализ решения на чувствительность — это анализ влияния изменений в постановке задачи (запасов ресурсов, величин прибыли от выпуска изделий и т.д.) на оптимальное решение. Во многих случаях анализ на чувствительность позволяет, не решая задачу заново, найти новое оптимальное решение задачи при изменениях в ее постановке.
Примечание. Рассматриваемые методы анализа на чувствительность применимы в случаях, когда в постановке задачи изменяется только одна из величин (например, запас одного из ресурсов, прибыль от одного из изделий и т.д.). Если в постановке задачи изменяется сразу несколько величин, то для получения нового оптимального решения следует решать задачу заново.

Задачи, решаемые методами анализа на чувствительность, очень разнообразны. Рассмотрим возможности анализа на чувствительность для задач, аналогичных примеру 2.1 (т.е. для задач, где требуется определить оптимальные объемы производства нескольких изделий при ограничениях на ресурсы).

Статус ресурсов

По статусу ресурсы делятся на дефицитные и недефицитные. Если для реализации оптимального решения ресурс расходуется полностью, то он называется дефицитным, если не полностью — недефицитным. Статус ресурсов определяется по значениям остаточных переменных. В примере 2.1 алюминий и пластмасса являются дефицитными ресурсами, так как они расходуются полностью (см. подраздел 2.4). Сталь - недефицитный ресурс, так как 90 кг стали остаются неизрасходованными (x_4 = 90). Увеличение запасов дефицитных ресурсов позволяет увеличить целевую функцию (прибыль). Снижение запасов дефицитных ресурсов приводит к снижению прибыли. Увеличение запасов недефицитных ресурсов всегда нецелесообразно, так как оно приводит только к увеличению неизрасходованных остатков. Запас недефицитного ресурса можно снизить на величину его остатка; это никаким образом не влияет на оптимальное решение (в том числе на оптимальные объемы производства и на прибыль), уменьшается только неизрасходованный остаток ресурса. Если запас недефицитного ресурса снизится на величину, превышающую его остаток, то для определения нового оптимального плана производства необходимо решать задачу заново.

В примере 2.1 увеличение запасов алюминия и пластмассы позволит увеличить прибыль. Запас стали можно снизить на 90 кг (т.е. до 160 кг); эти 90 кг стали предприятие может, например, продать или использовать в другом цехе. Например, если запас стали составит не 250, а только 200 кг, то оптимальное решение задачи будет следующим:

$$x_1 = 4; x_2 = 24; x_3 = 0; x_4 = 50; x_5 = 0; E = 7600 \text{ (ден.ед.)}$$

Таким образом, оптимальное решение не изменится (кроме снижения неизрасходованного остатка стали).

Если запас стали снизится более чем на 90 кг (т.е. составит менее 160 кг), то для определения нового оптимального плана производства необходимо решать задачу заново. Для нового оптимального решения изменятся не только значения переменных, но и состав переменных в оптимальном базисе (т.е. в оптимальный базис будут входить не переменные

x_1, x_2 и x_5, а другие переменные). Значение целевой функции при этом снизится, т.е. составит менее 7600 ден. ед.

Ценность ресурсов

Ценность ресурса - это увеличение значения целевой функции (прибыли) при увеличении запаса ресурса на единицу (или, соответственно, снижение целевой функции при уменьшении запаса ресурса на единицу).

Примечание. Вместо названия «ценность ресурса» используются также названия «теневая цена», «скрытая цена».

Ценности ресурсов определяются по симплекс-таблице, соответствующей оптимальному решению. Ценности ресурсов представляют собой коэффициенты Е-строки при остаточных переменных, соответствующих остаткам ресурсов.

Для примера 2.1 ценность алюминия равна 1,33 ден. ед./кг, ценность пластмассы — 14,67 ден. ед./кг. Это означает, например, что увеличение запаса алюминия на единицу (т.е. на 1 кг) приводит к увеличению прибыли предприятия в среднем на 1,33 ден. ед. Например, если запас алюминия увеличится на 100 кг (т.е. составит 300 кг), то прибыль составит примерно 7600 + 1,33·100 = = 7733 ден. ед. Снижение запаса алюминия приведет к соответствующему снижению прибыли.

Примечание. Расчеты ожидаемой прибыли при изменении запаса ресурса, выполняемые с использованием ценности ресурса, могут оказаться приближенными, так как в большинстве случаев (в том числе и в рассматриваемом примере) количество выпускаемых изделий может представлять собой только целые числа. Поэтому, например, увеличение запаса алюминия или пластмассы на небольшую величину (например, на 1 кг) может и не привести к увеличению прибыли, так как этого недостаточно для увеличения выпуска изделий хотя бы на единицу.

Ценность недефицитного ресурса всегда равна нулю. В данном примере ценность стали равна нулю, так как увеличение ее запаса не приводит к увеличению прибыли, а снижение (не более чем на 90 кг) — не приводит к снижению прибыли.

Ценность ресурса показывает максимальную (предельную) цену, по которой выгодно закупать ресурсы. Например, в рассматриваемой задаче предприятию выгодно закупать алюминий по цене не более 1,33 ден. ед./кг, пластмассу — по цене не более 14,67 ден. ед./кг. Закупка ресурса по цене, превышающей его ценность, означает, что затраты предприятия на закупку ресурса превышают прибыль от его использования.

Анализ на чувствительность к изменениям запасов ресурсов

Изменение запаса ресурса соответствует изменению правой части соответствующего ограничения в математической модели. Для анализа влияния таких изменений на оптимальное решение используются коэффициенты из столбца остаточной переменной, входящей в изменившееся ограничение.

Выполним анализ на чувствительность к изменению запаса пластмассы для примера 2.1. Пусть запас изменился на d кг и составляет не 500, а 500 + d кг. Величина d может быть как положительной (запас ресурса увеличился), так и отрицательной (запас уменьшился). Если изменение запаса ресурса не выходит за некоторый диапазон (определение этого диапазона будет показано ниже), то новое оптимальное решение можно найти следующим образом:

$$\begin{aligned} x_1 &= 4 - 0,01d \\ x_2 &= 24 + 0,05d \\ x_4 &= 90 - 0,13d \\ E &= 7600 + 14,67d \end{aligned} \qquad (2.5)$$

Здесь коэффициенты —0,01; 0,05; —0,13 и 14,67 взяты из столбца переменной x_5 в окончательной симплекс-таблице (табл. 2.4).

Пусть, например, запас пластмассы составляет не 500, а 600 кг. Хотя постановка задачи изменилась, для поиска нового оптимального решения не требуется решать задачу заново. Достаточно подставить в уравнения (2.5) величину d = 100 (так как запас пластмассы увеличился по сравнению с первоначальной постановкой задачи на 100 кг). Новое оптимальное решение оказывается следующим: $x_1 = 3; x_2 = 29; x_4 = 77; E = 9067$. Это означает, что в новых условиях (при запасе пластмассы 600 кг) цеху следует выпускать за смену 3 корпуса и 29 задвижек. Неизрасходованный остаток стали составит примерно 77 кг. Прибыль составит примерно 9067 ден. ед. Алюминий и пластмасса будут израсходованы полностью (переменные x_3 и x_5 остаются небазисными, т.е. равными нулю).

Примечание. Так как прибыль от выпуска одного корпуса составляет 100 ден. ед., а от одной задвижки - 300 ден. ед., можно подсчитать точную величину прибыли: 100·3+300·29=9000 ден. ед. Можно также определить точную величину расхода стали на выпуск изделий: 10·3+5·29=175 кг; значит, остаток составит 250 - 175 = 75 кг. Незначительные неточности в результатах, полученных на основе анализа на чувствительность, связаны с округлениями, допускавшимися при расчетах в симплекс-таблицах.

Пусть запас пластмассы составляет не 500, а 400 кг. Для поиска нового оптимального решения достаточно подставить в уравнения (2.5) величину d = -100 (так как запас пластмассы уменьшился по сравнению с первоначальной постановкой задачи на 100 кг).
Новое оптимальное решение следующее: $x_1 = 5; x_2 = 19; x_4 = 103; E = 6133$. Это означает, что в новых условиях (при запасе пластмассы 400 кг) цеху следует выпускать за смену 5 корпусов и 19 задвижек. Неизрасходованный остаток стали составит примерно 103 кг. Прибыль составит примерно 6133 ден. ед. Алюминий и пластмасса будут израсходованы полностью (переменные x_3 и x_5 остаются небазисными, т.е. равными нулю).

Примечание. Точная величина прибыли для данного решения составляет

$$100 \cdot 5 + 300 \cdot 19 = 6200 \text{ (ден.ед.)}$$

Точная величина остатка стали составит

$$250 – 10 \cdot 5 – 5 \cdot 19 = 105 \text{ кг.}$$

Можно также определить диапазон изменений запаса ресурса, при котором состав переменных в оптимальном базисе остается прежним. Этот диапазон находится из условия неотрицательности всех переменных. Так, для примера 2.1 диапазон допустимых изменений запаса пластмассы, не приводящий к изменению состава переменных в оптимальном базисе, находится из следующих условий:

$$x_1 = 4 - 0,01 \cdot d \geq 0$$

$$x_2 = 24 + 0,05 \cdot d \geq 0$$

$$x_4 = 90 - 0,13 \cdot d \geq 0$$

Решив эту систему неравенств, получим: —480 ≤ d ≤ 400. Это означает, что базис оптимального решения будет состоять из переменных x_1, x_2, x_4, если запас пластмассы, заданный в постановке задачи, будет составлять от 500-480 до 500 + 400 кг, т.е. от 20 до 900 кг. Для любой величины запаса пластмассы, входящей в этот диапазон, новое оптимальное решение можно найти из уравнений (2.5).

Аналогично можно найти, что базис оптимального решения будет состоять из переменных x_1, x_2, x_4, если запас алюминия будет составлять от 120 до 391,5 кг. Для определения этого диапазона потребуется использовать коэффициенты из столбца переменной x_3.

Если запас ресурса выходит за найденный диапазон, то для получения нового оптимального решения необходимо решить задачу заново, используя симплекс-метод. Новое оптимальное решение будет отличаться от прежнего не только значениями, но и составом переменных в оптимальном базисе.

Например, уравнения (2.5) нельзя использовать для определения нового оптимального решения, если запас пластмассы составит 1100 кг (т.е. увеличится на 600 кг). Если подставить величину d = 600 в уравнения (2.5), то переменная x_1 (количество корпусов) примет отрицательное значение, что не имеет смысла. Чтобы получить оптимальный план выпуска изделий, необходимо решить задачу заново, изменив ограничение на запас пластмассы следующим образом: $5x_1 + 20x_2 \le 1100$.

Примечание. Аналогично выполняется анализ на чувствительность к изменению любых ограничений «меньше или равно», независимо от того, что они означают: ограничения на запасы ресурсов или какие-либо другие величины.

Анализ на чувствительность к изменениям коэффициентов целевой функции

В задачах, аналогичных примеру 2.1 (определение оптимальных объемов производства нескольких изделий при ограничениях на ресурсы), изменение коэффициентов целевой функции соответствует изменению прибыли от выпуска изделия. Для анализа влияния таких изменений на оптимальное решение используются коэффициенты из строки переменной, для которой изменился коэффициент целевой функции.

Можно доказать, что если изменение коэффициента целевой функции не выходит за некоторый диапазон (определение этого диапазона будет рассмотрено ниже), то оптимальное решение задачи не изменяется. Изменяется только значение целевой функции, а также коэффициенты Е–строки при небазисных переменных в окончательной симплекс–таблице. Будем обозначать коэффициенты Е–строки в окончательной симплекс–таблице как $F_j,\ j = 1,...,k$ (где k – общее количество переменных в задаче).

Выполним анализ на чувствительность к изменению прибыли от выпуска одного корпуса для примера 2.1. Пусть эта прибыль изменилась на d (ден.ед.) и составляет не 100, а 100 + d ден. ед. Величина d может быть как положительной (прибыль от выпуска одного корпуса увеличилась), так и отрицательной (прибыль снизилась). Если изменение прибыли не выходит за некоторый диапазон (определяемый ниже), то новые значения коэффициентов Е-строки при небазисных переменных для окончательной симплекс–таблицы, а также новое оптимальное значение целевой функции можно найти следующим образом:

$$
\begin{aligned}
&F_3 = 1,33 + 0,05d \\
&F_5 = 14,67 - 0,01d \qquad (2,6) \\
&E = 7600 + 4d,
\end{aligned}
$$

где F_3, F_5 – новые значения коэффициентов Е–строки при небазисных переменных в окончательной симплекс–таблице.

Величины 0,05; – 0,01 и 4 взяты из строки переменной x_1 в окончательной симплекс-таблице (табл. 2.4). Пусть, например, прибыль от выпуска одного корпуса составляет не 100, а 120 ден. ед. Подставив в уравнения (2.6) величину d = 20 (так как прибыль увеличилась по сравнению с первоначальной постановкой задачи на 20 ден. ед.), найдем новые значения коэффициентов Е-строки в окончательной симплекс-таблице: $F_3 = 2,33; F_5 = 14,47$. Новое оптимальное значение целевой функции составит Е = 7680 ден. ед. Значения коэффициентов Е- строки при небазисных переменных (F_1, F_2 и F_4) останутся равными нулю. Так как все коэффициенты Е-строки остались неотрицательными, оптимальное решение задачи не изменяется: $x_1 = 4; x_2 = 24; x_3 = 0; x_4 = 90; x_5 = 0$. Это означает, что в новых условиях (при увеличении прибыли от выпуска одного корпуса до 120 ден. ед.) цеху по-прежнему следует выпускать за смену 4 корпуса и 24 задвижки. Прибыль от их выпуска составит 7680 (ден.ед.) Анализ на чувствительность к изменению коэффициентов целевой функции позволяет выяснить диапазоны изменений этих коэффициентов, для которых найденное решение задачи остается оптимальным. Признаком оптимальности решения являются неотрицательные значения всех коэффициентов Е–строки (см. подраздел 2.4).

Найдем диапазон изменения прибыли от выпуска одного корпуса, для которого найденное решение задачи $(x_1 = 4; x_2 = 24; x_3 = 0; x_4 = 90; x_5 = 0)$ останется оптимальным. Для этого необходимо, чтобы все коэффициенты Е–строки оставались неотрицательными:

$$F_3 = 1,33 + 0,05d \geq 0$$

$$F_5 = 14,67 - 0,01d \geq 0.$$

Решив эту систему неравенств, получим: – 26,6 ≤ d ≤ 1467. Это означает, что найденное для задачи решение $(x_1 = 4; x_2 = 24; x_3 = 0; x_4 = 90; x_5 = 0)$ оптимально, если прибыль от выпуска одного корпуса будет составлять от 100 – 26,6 до 100 + 1467 ден. ед., т.е. от 73,4 до 1567 ден. ед. Для любой величины прибыли, входящей в этот диапазон, новые значения коэффициентов Е-строки и целевой функции можно найти из уравнений (2.6).

Аналогично можно определить, что оптимальное решение задачи не изменится, если прибыль от выпуска одной задвижки будет составлять от 6,6 до 433 ден. ед. Для определения этого диапазона потребуется использовать коэффициенты из строки переменной x_2 (табл. 2.4).

Если коэффициент целевой функции выходит за найденный диапазон, то для получения оптимального решения необходимо решить задачу заново, используя симплекс–метод. Новое оптимальное решение будет отличаться от прежнего не только значениями, но и составом переменных в оптимальном базисе. При этом прежнее решение (т.е. оптимальное решение исходной задачи) уже не будет оптимальным, но останется допустимым, так как оно удовлетворяет ограничениям задачи.

Например, если прибыль от выпуска одного корпуса составит 1600 (ден.ед.), т.е. увеличится на 1500 ден. ед., то для получения оптимального плана выпуска изделий необходимо решить задачу заново, изменив целевую функцию следующим образом:

$$E = 1600x_1 + 300x_2 \to \max .$$

Прежнее оптимальное решение

$$(x_1 = 4; x_2 = 24; x_3 = 0; x_4 = 90; x_5 = 0)$$

уже не является оптимальным. В этом легко убедиться, подставив величину d = 1500 в систему уравнений (2.6): коэффициент Е-строки при переменнойX5 принимает значение —0,33, т.е. становится отрицательным. В то же время прежнее решение остается допустимым, так как значения $x_1 = 4$ и $x_2 = 24$ удовлетворяют ограничениям задачи.

3. Решение задачи ЛП на основе методов искусственного базиса

3.1. Назначение и принцип работы методов искусственного базиса

Методы искусственного базиса предназначены для решения задач линейного программирования, содержащих ограничения различных видов: «больше или равно», «меньше или равно», «равно».

При решении задачи линейного программирования для построения начального базиса необходимо, чтобы в каждом ограничении присутствовала базисная переменная, т.е. переменная, входящая в данное ограничение с коэффициентом, равным единице, и не входящая ни в одно из других ограничений. В ограничениях «меньше или равно» в качестве таких переменных используются остаточные переменные, добавляемые в ограничение при его приведении к стандартной форме (см. подраздел 2.3). Для приведения к стандартной форме ограничений «больше или равно» вводятся избыточные переменные со знаком «минус». В ограничения «равно» не требуется вводить никаких дополнительных переменных, так как такие ограничения уже соответствуют стандартной форме. Поэтому в задачах, содержащих ограничения «больше или равно» или «равно», после приведения к стандартной форме обычно невозможно построить начальный базис, так как базисные переменные имеются не во всех ограничениях. Для задач, содержащих ограничения «не меньше» или «равно», обычно нельзя использовать в качестве начального допустимого решения (начальной угловой точки ОДР) начало координат, т.е. решение, в котором все исходные переменные математической модели равны нулю: $x_1 = x_2 = ... = x_n = 0$. Такое решение, как правило, оказывается недопустимым (не соответствует ограничениям).

Методы искусственного базиса применяются во всех случаях, когда базисные переменные имеются не во всех ограничениях задачи, приведенной к стандартной форме. Принцип работы всех методов искусственного базиса следующий. Во все ограничения, не содержащие базисных переменных, вводятся искусственные переменные (по одной в каждое ограничение), используемые для построения начального базиса. После этого выполняется поиск оптимального решения на основе обычных процедур симплекс-метода.

В окончательном (оптимальном) решении задачи все искусственные переменные должны быть равны нулю. Если в оптимальном решении какая-либо из искусственных переменных оказывается ненулевой, это означает, что задача не имеет допустимых решений. Причиной может быть ошибка в математической модели или противоречия в постановке задачи (например, количество изделий, которое требуется выпустить, не может быть выпущено из-за ограничений на ресурсы).

На искусственные переменные, как и на все остальные переменные в задаче, накладывается требование неотрицательности.

Искусственные переменные не имеют никакого физического смысла: их нельзя интерпретировать как количество изделий, запасы ресурсов и т.д. Они требуются только для построения начального базиса.

Основные методы искусственного базиса - двухэтапный метод, рассматриваемый ниже, и метод больших штрафов [1]. Поиск решения на основе этих методов выполняется с использованием симплекс-таблиц.

3.2. Двухэтапный метод

Основные этапы реализации двухэтапного метода (как и других методов искусственного базиса) следующие.

1. Строится искусственный базис. Находится начальное недопустимое решение. Выполняется переход от начального недопустимого решения к некоторому допустимому решению. Этот переход реализуется путем минимизации (сведения к нулю) искусственной целевой функции, представляющей собой сумму искусственных переменных.

2. Выполняется переход от начального допустимого решения к оптимальному решению.

Реализацию двухэтапного метода рассмотрим на следующем примере.

Пример 3.1. Изделия трех видов (A, B, C) вырезаются из стальных листов. Предприятие имеет 150 стальных листов. Каждый лист можно раскроить одним из трех способов. Количество изделий, получаемых из одного листа, и величины отходов для каждого способа раскроя приведены в табл. 3.1.

Таблица 3.1

Количество изделий	Способы раскроя		
	1	2	3
A	4	5	2
B	1	1	4
C	2	1	1
2 Отходы, см^2	20	25	17

Предприятию необходимо раскроить листы таким образом, чтобы отходы были минимальны. При этом необходимо выпустить не менее 400 изделий А, не менее 250 изделий В и не более 300 изделий С (последнее требование связано с тем, что спрос на изделия С ограничен).

Составим математическую модель задачи. Обозначим через x_1, x_2 и x_3 количество стальных листов, раскраиваемых первым, вторым и третьим способом соответственно. Математическая модель будет иметь следующий вид:

$$\begin{aligned} &4x_1 + 5x_2 + 2x_3 \ge 400 \\ &x_1 + x_2 + 4x_3 \ge 250 \\ &x_1 + x_2 + x_3 = 150 \\ &2x_1 + x_2 + x_3 \le 300 \\ &x_i \ge 0, i = 1,...,3. \end{aligned} \tag{3.1}$$

$$E = 20x_1 + 25x_2 + 17x_3 \to \min. \tag{3.2}$$

Здесь первое и второе ограничения устанавливают, что необходимо выпустить не менее 400 изделий А и не менее 250 изделий В. Третье ограничение указывает, что общее количество листов, раскроенных всеми тремя способами, должно составлять 150 (т.е. необходимо раскроить все листы). Четвертое ограничение указывает, что количество изделий С не должно превышать 300. Целевая функция представляет собой общее количество отходов.

Приведем задачу к стандартной форме, как показано в подразделе 1.4. Все ограничения требуется преобразовать в равенства. Для этого в ограничения «больше или равно» (первое и второе) необходимо ввести избыточные переменные. В ограничение «меньше или равно» (четвертое) добавляется остаточная переменная. В ограничение «равно» не требуется вводить никаких дополнительных переменных. Кроме того, требуется перейти к целевой функции, подлежащей максимизации. Для этого целевая функция Е умножается на —1. Математическая модель задачи в стандартной форме имеет следующий вид:

$$\begin{aligned} &4x_1 + 5x_2 + 2x_3 - x_4 = 400 \\ &x_1 + x_2 + 4x_3 - x_5 = 250 \\ &x_1 + x_2 + x_3 = 150 \\ &2x_1 + x_2 + x_3 + x_6 = 300 \\ &x_i \ge 0, i = 1,...,6. \end{aligned} \tag{3.3}$$

$$-E = -20x_1 - 25x_2 - 17x_3 \to \max. \tag{3.4}$$

В полученной системе уравнений базисная переменная имеется только в четвертом уравнении; это переменная x_6. Поэтому для решения задачи требуется использовать методы искусственного базиса.

Первый этап (поиск допустимого решения)

1. Во все ограничения, где нет базисных переменных, вводятся искусственные базисные переменные. В данной задаче их требуется ввести в первое, второе и третье ограничения. Добавлять искусственную переменную в четвертое ограничение не требуется, так как оно уже содержит базисную переменную x_6. Система ограничений с искусственными базисными переменными будет иметь следующий вид:

$$\begin{aligned}&4x_1+5x_2+2x_3-x_4+x_7=400\\&x_1+x_2+4x_3-x_5+x_8=250\\&x_1+x_2+x_3+x_9=150\\&2x_1+x_2+x_3+x_6=300\\&x_i\geq 0, i,...,9.\end{aligned}\tag{3.5}$$

Таким образом, начальный базис будет состоять из искусственных переменных x_7, x_8, x_9, а также остаточной переменной x_6.

2. Составляется искусственная целевая функция - сумма всех искусственных переменных:

$$W = x_7 + x_8 + x_9 \to \min. \tag{3.6}$$

Эта целевая функция подлежит минимизации, так как для определения начального допустимого решения необходимо, чтобы все искусственные переменные приняли нулевые значения.

Примечание. Искусственная целевая функция всегда (в любой задаче) подлежит минимизации.

3. Искусственная целевая функция выражается через небазисные переменные. Для этого сначала требуется выразить искусственные переменные через небазисные:

$$\begin{aligned}&x_7=400-4x_1-5x_2-2x_3+x_4\\&x_8=250-x_1-x_2-4x_3+x_5\\&x_9=150-x_1-x_2-x_3.\end{aligned}\tag{3.7}$$

Выраженные таким образом искусственные переменные подставляются в искусственную целевую функцию:

$$W=-6x_1-7x_2-7x_3+x_4+x_5+800\to\min. \tag{3.8}$$

4. Для приведения всей задачи к стандартной форме выполняется переход к искусственной целевой функции, подлежащей максимизации. Для этого она умножается на -1:

$$-W=6x_1+7x_2+7x_3-x_4-x_5-800\to\max. \tag{3.9}$$

Приведем полную математическую модель задачи, приведенную к стандартной форме:

$$\begin{aligned}&4x_1+5x_2+2x_3-x_4+x_7=400\\&x_1+x_2+4x_3-x_5+x_8=250\\&x_1+x_2+x_3+x_9=150\\&2x_1+x_2+x_3+x_6=300\\&x_i\geq 0, i=1,...,9.\end{aligned}\tag{3.10}$$

$$-E=-20x_1-25x_2-17x_3\to\max. \tag{3.11}$$

$$-W=6x_1+7x_2+7x_3-x_4-x_5-800\to\max. \tag{3.12}$$

5. Определяется начальное решение. Все исходные, а также избыточные переменные задачи являются небазисными, т.е. принимаются равными нулю. Искусственные, а также остаточные переменные образуют начальный базис: они равны правым частям ограничений. Для рассматриваемой задачи начальное решение следующее: $x_1=x_2=x_3=x_4=x_5=0, x_7=400, x_8=250, x_9=150$. Это решение является недопустимым: значения переменных $x_1=x_2=x_3=0$ не удовлетворяют постановке задачи.

Начальное значение целевой функции задачи $E = 20x_1 + 25x_2 + 17x_3$ равно нулю. Начальное значение искусственной целевой функции $-W = 6x_1 + 7x_2 + 7x_3 - x_4 - x_5 - 800$ равно -800.

6. Составляется исходная симплекс-таблица. Она отличается от симплекс-таблицы, используемой для обычного симплекс-метода (см. подраздел 2.4) только тем, что в нее добавляется строка искусственной целевой функции. В этой строке указываются коэффициенты искусственной целевой функции (приведенной к стандартной форме, т.е. подлежащей максимизации) с обратными знаками, как и для обычной целевой функции. Исходная симплекс-таблица для рассматриваемого примера приведена в табл. 3.2.

Таблица 3.2

Базис	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	Решение
-E	20	25	17	0	0	0	0	0	0	0
-W	-6	-7	-7	1	1	0	0	0	0	-800
x_7	4	5	2	-1	0	0	1	0	0	400
x_8	1	1	4	0	-1	0	0	1	0	250
x_9	1	1	1	0	0	0	0	0	1	150
x_6	2	1	1	0	0	1	0	0	0	300

7. Выполняется переход от начального недопустимого решения, содержащегося в исходной симплекс-таблице, к некоторому допустимому решению. Для этого с помощью обычных процедур симплекс-метода выполняется минимизация искусственной целевой функции W(или, то же самое, максимизация функции —W). При этом переменные, включаемые в базис, выбираются по строке искусственной целевой функции. Все остальные действия выполняются точно так же, как в обычном симплекс-методе. В результате минимизации искусственная целевая функция — W должна принять нулевое значение. Все искусственные переменные при этом также становятся равными нулю (исключаются из базиса), так как искусственная целевая функция представляет собой их сумму.

В табл. 3.2 в строке искусственной целевой функции содержатся два одинаковых коэффициента, имеющих максимальные по модулю (в этой строке) отрицательные значения, равные —7. Это коэффициенты при переменных x_2 и x_3. Для включения в базис можно выбрать любую из этих переменных. Выберем x_2. Столбец переменной x_2 становится ведущим.

Для определения переменной, исключаемой из базиса, найдем симплексные отношения: 400/5 = 80; 250/1 = 250; 150/1 = 150; 300/1 = 300. Минимальное симплексное отношение получено в строке переменной x_7; значит, эта переменная исключается из базиса. Строка переменной x_7 становится ведущей.

Выполняются преобразования таблицы по правилам симплекс-метода. Ведущая строка x_7 делится на ведущий элемент (в данном примере он равен 5). Ведущий столбец x_2 заполняется нулями. Все остальные элементы таблицы (включая строки основной и искусственной целевых функций, а также столбец решений) пересчитываются по «правилу прямоугольника». Полученная симплекс-таблица приведена в табл. 3.3.

Таблица 3.3

Базис	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	Решение
-E	0,00	0,00	7,00	5,00	0,00	0,00	-5,00	0,00	0,00	-2000,00
-W	-0,40	0,00	-4,20	-0,40	1,00	0,00	0,00	0,00	0,00	-240,00

x_2	0,80	1,00	0,40	-0,20	0,00	0,00	0,20	0,00	0,00	80,00
x_8	0,20	0,00	3,60	0,20	-1,00	0,00	-0,20	1,00	0,00	170,00
x_9	0,20	0,00	0,60	0,20	0,00	0,00	-0,20	0,00	1,00	70,00
x_6	1,20	0,00	0,60	0220	0,00	1,200	-0,20	0,00	0,00	220,00

Полученное решение еще не является допустимым: в базисе есть искусственные переменные, и искусственная целевая функция не равна нулю.

Примечание. В том, что решение недопустимо, легко убедиться, подставив значения исходных переменных задачи $(x_1 = 0, x_2 = 80, x_3 = 0)$ в математическую модель (3.1).

Минимизация искусственной целевой функции продолжается. Для включения в базис выбирается переменная x_3, так как ей соответствует максимальный по модулю отрицательный коэффициент в строке искусственной целевой функции. Для выбора переменной, исключаемой из базиса, найдем симплексные отношения: 80/0,4 = 200; 170/3,6 = 47,2; 70/0,6 = 116,67; 220/0,6 = 366,67.

Минимальное симплексное отношение получено в строке переменной x_8; она исключается из базиса. Выполняются преобразования по правилам симплекс-метода. Полученная симплекс-таблица приведена в табл. 3.4.

Таблица 3.4

Базис	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	Решение
-E	-0,39	0,00	0,00	4,61	1,94	0,00	-4,61	-1,94	0,00	-2330,56
-W	-0,17	0,00	0,00	-0,17	-0,17	0,00	1,17	1,17	0,00	-41,67
x_2	0,78	1,00	0,00	-0,22	1,11	0,00	0,22	20,11	0,00	61,11
x_3	0,06	0,00	1,00	0,06	-0,28	0,00	-0,06	0,28	0,00	47,22
x_9	0,17	0,0	0,0	0,17	0,17	0,0	-0,17	-0,17	1,00	41,67
x_6	1,17	0,00	0,00	0,17	0,17	0,00	-0,17	-0,17	0,00	191,67

В базисе еще остаются искусственные переменные, поэтому поиск допустимого решения продолжается. В базис можно включить одну из переменных x_1, x_4 или x_5 (коэффициенты в строке искусственной целевой функции при этих переменных имеют одинаковые отрицательные значения). Выберем переменную x_1. Найдем симплексные отношения: 61,11/0,78 = 78,35; 47,22/0,06 = 787; 41,67/0,17 = 245,12; 191,67/1,17 = 163,82. Из базиса исключается переменная x_2. Новая симплекс-таблица приведена в табл. 3.5.

Таблица 3.5

Базис	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	Решение
-E	0,00	0,50	0,00	4,50	2,00	0,00	-4,5	00,00	0,00	-2300,00
-W	0,00	01,21	0,00	-0,21	-0,14	0,00	1,21	1,14	0,00	-28,57
x_1	1,00	1,29	0,00	-0,29	0,14	0,00	0,29	-0,14	0,00	78,57
x_3	0,00	-0,07	1,00	0,07	-0,29	0,00	-0,07	0,29	0,00	42,86
x_9	0,00	0,21	0,00	0,21	0,14	0,00	-0,21	-0,14	1,00	28,57
x_6	0,00	-1,50	0,00	1,50	0,00	1,500	00,50	0,00	0,00	100,00

В базисе есть искусственная переменная, поэтому поиск допустимого решения продолжается. В базис включается переменная x_4, имеющая максимальный по модулю отрицательный коэффициент в строке искусственной целевой функции. Чтобы определить переменную, исключаемую из базиса, найдем симплексные отношения: 42,86/0,07 = 612,29; 28,57/0,21 = 136,05; 100/0,5 = 200 (для строки переменной x_1 симплексное отношение не определяется, так как соответствующий коэффициент ведущего столбца имеет отрицательное значение). Из базиса исключается переменная x_9. Новая симплекс-таблица приведена в табл. 3.6.

Таблица 3.6

Базис	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	Решение
-E	0,00	5,00	0,00	0,00	-1,00	0,00	0,00	1,00	-21,0	-2900,00
-W	0,00	0,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00	0,00
x_1	1,00	1,00	0,00	0,00	0,33	0,00	0,00	-0,33	1,33	116,67
x_3	0,00	0,00	1,00	0,00	-0,33	0,00	0,00	0,33	-0,33	33,33
x_9	0,00	-1,00	0,00	1,00	0,67	0,00	-1,00	-0,67	4,67	133,33
x_6	0,00	-1,00	0,00	0,00	-0,33	1,00	0,00	0,33	-2,33	33,33

Как видно из табл. 3.6, искусственная целевая функция равна нулю, и все искусственные переменные исключены из базиса. Получено допустимое решение. Таким образом, первый этап двухэтапного метода завершен. Искусственная целевая функция и искусственные переменные исключаются из симплекс-таблицы (табл. 3.7).

Примечание. В том, что получено допустимое решение, легко убедиться, подставив значения исходных переменных задачи $(x_1 = 116,67, x_2 = 0, x_3 = 33,33)$ в математическую модель (3.1).

Таблица 3.7

Базис	x_1	x_2	x_3	x_4	x_5	x_6	Решение
-E	0,00	5,00	0,00	0,00	-1,00	0,00	-2900,00
x_1	1,00	1,00	0,00	0,00	0,33	0,00	116,67
x_3	0,00	0,00	1,00	0,00	-0,33	0,00	33,33
x_5	0,00	-1,00	0,00	1,00	0,67	0,00	133,33
x_6	0,00	-1,00	0,00	0,00	-0,33	1,00	33,33

Второй этап (поиск оптимального решения)

Решение, полученное по результатам первого этапа (табл. 3.7), не является оптимальным: в строке целевой функции имеется отрицательный элемент. Поиск оптимального решения выполняется по обычным правилам симплекс-метода. В базис включается переменная x_5. Для определения переменной, исключаемой из базиса, найдем симплексные отношения: 116,67/0,33 = 350, 133,33/0,67 = 200. Из базиса исключается переменная x_4 После преобразований по правилам симплекс-метода будет получена новая симплекс-таблица (табл. 3.8).

Таблица 3.8

Базис	x_1	x_2	x_3	x_4	x_5	x_6	Решение
-E	0,00	3,50	0,00	1,50	0,00	0,00	-2700,00
x_1	1,00	1,50	0,00	-0,50	0,00	0,00	50,00
x_3	0,00	-0,50	1,00	0,50	0,00	0,00	100,00

x_5	0,00	-1,50	0,00	1,50	1,500	0,00	200,00
x_6	0,00	-1,50	0,00	1,50	0,00	1,500	100,00

Получено оптимальное решение (признак его оптимальности — отсутствие отрицательных элементов в строке целевой функции). Основные переменные задачи приняли следующие значения: $x_1 = 50, x_2 = 0$ (эта переменная — небазисная), $x_3 = 100$. Это означает, что необходимо раскроить 50 стальных листов первым способом, 100 листов — третьим способом. Второй способ раскроя использовать не следует. Значение целевой функции E = 2 2700 показывает, что отходы при таком раскрое составят 2700 см .

Избыточная переменная $x_4 = 0$ означает, что изделий A будет выпущено не больше минимально необходимого количества, т.е. ровно 400. Избыточная переменная $x_5 = 200$ означает, что количество выпущенных изделий B будет на 200 больше, чем минимально необходимое количество, т.е. 250 + 200 = 450. Остаточная переменная $x_6 = 100$ означает, что количество выпущенных изделий C будет на 100 изделий меньше, чем максимально допустимое количество, т.е. 300 — 100 = 200. Таким образом, будет выпущено 400 изделий A, 450 изделий B и 200 изделий C.

3.3. Анализ оптимального решения на чувствительность

Задачи линейного программирования, для решения которых применяются методы искусственного базиса, очень разнообразны по своему содержанию. Методы анализа на чувствительность, используемые для таких задач, и интерпретация результатов полностью зависят от постановки задачи.

Анализ на чувствительность к изменениям правых частей ограничений «меньше или равно»

Для анализа влияния таких изменений на оптимальное решение используются коэффициенты из столбца остаточной переменной, входящей в изменившееся ограничение. Анализ выполняется точно так же, как показано в п. 2.6.3.

Рассмотрим анализ на чувствительность к изменению ограничения на выпуск изделий C для примера 3.1. Предположим, что спрос на изделия C изменился, в результате чего максимально допустимый выпуск этих изделий составляет не 300, а 300 + d штук. Для составления уравнений, позволяющих найти новое оптимальное решение, необходимо использовать коэффициенты из столбца переменной x_6. Новое оптимальное решение можно найти следующим образом:

$$\begin{aligned}
x_1 &= 50 + 0d \\
x_3 &= 100 + 0d \\
x_5 &= 200 + 0d \qquad (3.13) \\
x_6 &= 100 + 1d \\
-E &= -2700 + 0d
\end{aligned}$$

Из этих уравнений видно, что изменения ограничения на выпуск изделий C (если эти изменения не выходят за определенный диапазон) не приведут к каким-либо изменениям в решении задачи. Количество листов, раскраиваемых различными способами, а также количество отходов (целевая функция) останутся без изменений. Будет изменяться только переменная x_6, т.е. разность между фактическим и максимально допустимым выпуском изделий C.

Можно также определить диапазон изменений ограничения, при котором состав переменных в оптимальном базисе остается прежним. Этот диапазон находится из условия неотрицательности всех переменных:

$$
\begin{aligned}
x_1 &= 50 + 0d \ge 0 \\
x_3 &= 100 + 0d \ge 0 \\
x_5 &= 200 + 0d \ge 0 \\
x_6 &= 100 + 1d \ge 0
\end{aligned}
\qquad (3.14)
$$

Решив эту систему неравенств, получим: $-100 \le d \le \infty$. Это означает, что базис оптимального решения будет состоять из переменных x_1, x_3, x_5, x_6, если ограничение на выпуск изделий С будет составлять не менее 200 шт. Если это ограничение составит менее 200 штук, то для получения нового оптимального решения потребуется решить задачу заново. В этом случае новое оптимальное решение будет отличаться от прежнего не только значениями, но и составом переменных в оптимальном базисе.

Анализ на чувствительность к изменениям правых частей ограничений «больше или равно».

Для анализа влияния таких изменений на оптимальное решение используются коэффициенты из столбца избыточной переменной, входящей в изменившееся ограничение, причем эти коэффициенты используются с обратными знаками. В остальном анализ выполняется так же, как и для ограничений «меньше или равно».

Рассмотрим анализ на чувствительность к изменению ограничения на выпуск изделий А для примера 3.1. Предположим, что минимально необходимый выпуск этих изделий составляет не 400, а 400 + d штук. Такое изменение может быть связано, например, с заключением новых контрактов на поставку изделий А или, наоборот, со снижением спроса на эти изделия. Для составления уравнений, позволяющих найти новое оптимальное решение, необходимо использовать коэффициенты из столбца переменной x_5, взятые с обратными знаками. Новое оптимальное решение можно найти из следующих уравнений:

$$
\begin{aligned}
x_1 &= 50 + 0,5d \\
x_3 &= 100 - 0,5d \\
x_5 &= 200 - 1,5d \\
x_6 &= 100 - 0,5d \\
-E &= -2700 - 1,5d
\end{aligned}
\qquad (3.15)
$$

Пусть, например, предприятию необходимо выпустить не менее 420 изделий А. Для определения нового оптимального решения достаточно подставить в уравнения (3.15) величину d = 20 (так как ограничение на выпуск изделий А увеличилось по сравнению с первоначальной постановкой задачи на 20 штук). Новое оптимальное решение оказывается следующим: $x_1 = 60, x_3 = 90, x_5 = 170, x_6 = 90, -E = -2730$. Это означает, что в новых условиях (при потребности в выпуске не менее 420 изделий А) предприятию необходимо раскроить 60 листов первым способом и 90 листов — третьим способом. Отходы составят 2730 см^2 . Изделий А будет выпущено ровно столько, сколько необходимо, т.е. 420 (избыточная переменная x_4, означающая выпуск изделий А сверх минимально необходимого количества, остается небазисной, т.е. равна нулю). Изделий В будет выпущено на 170 штук больше минимально необходимого количества, т.е. 250 + 170 = 420. Выпуск изделий С составит на 90 штук меньше максимально допустимого количества, т.е. 300 - 90 = 210.

Диапазон изменений ограничения, при котором состав переменных в оптимальном базисе остается прежним, определяется из условия неотрицательности всех переменных.

Решив эту систему неравенств, получим: —100 ≤ d ≤ 133,33. Это означает, что базис оптимального решения будет состоять из переменных x_1, x_3, x_5, x_6, если ограничение, устанавливающее минимально необходимый выпуск изделий А, будет составлять от 400 - 100 до 400 + 133,33, т.е. от 300 до 533 штук (это ограничение не может быть дробным числом). Если это ограничение составит менее 300 штук или более 533 штук, то для получения нового оптимального решения потребуется решить задачу заново.

Анализ на чувствительность к изменениям коэффициентов целевой функции

Для анализа влияния таких изменений на оптимальное решение используются коэффициенты из строки переменной, для которой изменился коэффициент целевой функции. Если целевая функция в постановке задачи подлежит максимизации, то анализ выполняется точно так же, как показано в п. 2.6.4. Если целевая функция подлежит минимизации, то анализ выполняется аналогично, однако коэффициенты из строки переменной используются с обратными знаками.

Как указано в п. 2.6.4, изменение коэффициента целевой функции (в пределах определенного диапазона) не приводит к изменениям в оптимальном решении задачи. Изменяется только значение целевой функции, а также коэффициенты Е-строки при небазисных переменных в окончательной симплекс-таблице.

Будем обозначать коэффициенты Е-строки в окончательной симплекс-таблице как Fj,j = 1,..., k (где k— общее количество переменных в задаче).

Выполним анализ на чувствительность к изменению отходов от раскроя листа первым способом. Предположим, что величина отходов составляет не 20, а 20+d см . Такое изменение может быть связано, например, с использованием листов другого размера или с изменением самого способа раскроя. Величина d может быть как положительной, так и отрицательной. Чтобы составить уравнения, позволяющие найти новые значения элементов Е-строки для окончательной симплекс-таблицы, необходимо использовать коэффициенты из строки переменной x_1, взятые с обратными знаками:

$$\begin{aligned} &F_2 = 3,5 - 1,5d \\ &F_4 = 1,5 + 0,5d \\ &-E = -2700 - 50d, \end{aligned} \qquad (3.17)$$

где F_2, F_4 — новые значения коэффициентов Е-строки при небазисных переменных в окончательной симплекс-таблице.

Пусть, например, отходы от раскроя одного листа первым способом составляют не 20, а 22 см . Подставив в уравнения (3.17) величину d = 2, найдем новые значения коэффициентов Е-строки в окончательной симплекс-таблице: $F_2 = 0,5, F_4 = 2,5$. Новое оптимальное значение целевой функции 2800 см . Значения коэффициентов Е-строки при небазисных переменных (F_1, F_3, F_5, F_6) останутся равными нулю. Так как все коэффициенты Е-строки остаются неотрицательными, оптимальное решение задачи не изменяется: $x_1 = 50, x_2 = 0, x_3 = 100$. Это означает, что в новых условиях (когда отходы от раскроя одного листа первым способом составляют 2 22 см) предприятию по-прежнему следует раскраивать 50 листов первым способом и 100 листов — третьим способом. Выпуск изделий также не изменится: изделий А будет выпущено 400 штук, изделий В — 450 штук, 2 изделий С — 200 штук. Отходы составят 2800 см .

Найдем диапазон величины отходов от раскроя одного листа, для которого найденное решение задачи $(x_1 = 50, x_2 = 0, x_3 = 100, x_4 = 0, x_5 = 200, x_6 = 100)$ останется оптимальным. Этот диапазон определяется из условия неотрицательности всех коэффициентов Е-строки:

$$\begin{aligned} &F_2 = 3,5 - 1,5d \geq 0 \\ &F_4 = 1,5 + 0,5d \geq 0 \end{aligned} \qquad (3.18)$$

Решив эту систему неравенств, получим: —3 ≤ d ≤ 2,33. Это означает, что найденное для задачи решение $(x_1 = 50, x_2 = 0, x_3 = 100, x_4 = 0, x_5 = 200, x_6 = 100)$ оптимально, если отходы от раскроя одного листа первым способом составляют от 20 - 3 до 20 + 2,33 см^2 , т.е. от 17 до 22,33 см^2 .

Для любой величины отходов, входящей в этот диапазон, новые значения коэффициентов Е-строки и целевой функции можно найти из уравнений (3.17).

Аналогично можно определить, что оптимальное решение задачи не изменится, если отходы от раскроя одного листа третьим способом будут составлять от 10 до 20 см^2 . Для определения этого диапазона потребуется использовать коэффициенты из строки переменной x_3. Если изменится величина отходов от раскроя листа вторым способом, то для определения оптимального решения необходимо решать задачу заново (так как переменная x_2 не входит в базис, и анализ на чувствительность для коэффициента целевой функции при этой переменной невозможен).

Если коэффициент целевой функции выходит за найденный диапазон, то для получения оптимального решения необходимо решить задачу заново, используя симплекс-метод. Новое оптимальное решение будет отличаться от прежнего не только значениями, но и составом переменных в оптимальном базисе. При этом прежнее решение (т.е. оптимальное решение исходной задачи) уже не будет оптимальным, но останется допустимым, так как оно удовлетворяет ограничениям задачи.

Например, если отходы при раскрое одного листа первым способом составят 24 см^2, то для получения оптимального плана раскроя листов необходимо решить задачу заново, изменив целевую функцию следующим образом: $E = 24 \cdot x_1 + 25 \cdot x_2 + 17 \cdot x_3 \to \min$. Прежнее оптимальное решение $(x_1 = 50, x_2 = 0, x_3 = 100)$ уже не является оптимальным. В этом легко убедиться, подставив величину d = 4 в систему уравнений (3.17): коэффициент Е-строки при переменной x_2 принимает значение —2,5, т.е. становится отрицательным, что является признаком неоптимальности решения. В то же время прежнее решение остается допустимым, так как значения $x_1 = 50, x_2 = 0, x_3 = 100$ удовлетворяют ограничениям задачи (3.1).

4. Оптимизация на основе целочисленного программирования

4.1. Назначение методов целочисленного программирования

Методы целочисленного программирования предназначены для решения задач, в которых некоторые (или все) переменные по своему физическому смыслу должны принимать только целочисленные значения.

Если требование целочисленности накладывается на все переменные, имеющиеся в математической модели задачи (включая все остаточные и избыточные переменные, вводимые в модель при ее приведении к стандартной форме), то такая задача называется полностью целочисленной. Если требование целочисленности накладывается лишь на некоторые переменные, то задача называется частично целочисленной.

Например, задача, рассмотренная в примере 2.1, была частично целочисленной. На переменные x_1 и x_2, обозначавшие количество изделий, накладывалось требование целочисленности, а на переменные x_3, x_4 и x_5, обозначавшие неиспользованные остатки алюминия, стали и пластмассы, такое требование не накладывалось.

Задача из примера 3.1 была полностью целочисленной, так как основные переменные задачи (x_1, x_2 и x_3) обозначали количество стальных листов, а избыточные и остаточные переменные (x_4, x_5 и x_6) — количество изделий. Очевидно, что все эти величины по своему смыслу должны принимать только целочисленные значения.

Принцип работы всех методов линейного целочисленного программирования следующий. Сначала задача решается симплекс-методом без учета требований целочисленности. Если все переменные, которые по своему смыслу должны быть целочисленными, принимают в оптимальном решении целочисленные значения, то решение задачи на этом завершается. Если хотя бы одна из этих переменных принимает дробное значение, то в задачу вводятся дополнительные ограничения, исключающие полученное оптимальное (но нецелочисленное) решение из области допустимых решений. Новые задачи (одна или несколько), включающие дополнительные ограничения, также решаются симплекс-методом. Процесс продолжается до получения оптимального целочисленного решения.

Основные методы решения задач целочисленного программирования — метод ветвей и границ и метод Гомори [1].

4.2. Метод ветвей и границ

Принцип работы метода ветвей и границ основан на использовании списка решаемых задач (задач-кандидатов). Если в оптимальном решении задачи какая-либо из переменных, которая по своему смыслу должна быть целочисленной, приняла дробное значение, то в список решаемых задач включаются новые задачи с дополнительными ограничениями, исключающими полученное оптимальное (но нецелочисленное) решение из области допустимых решений, т.е. делающими это решение недопустимым. Для каждой задачи определяется также ее оценка — граница значения целевой функции (способ определения оценки будет показан ниже). Под оценкой задачи понимается величина, о которой точно известно, что оптимальное значение целевой функции не будет лучше этой величины. Из списка решаемых задач выбирается очередная задача (имеется несколько вариантов метода ветвей и границ, где правила выбора такой задачи могут быть разными). Выбранная задача решается симплекс-методом. Если решение выбранной задачи оказывается нецелочисленным, то в нее также вводятся дополнительные ограничения, и полученные задачи включаются в список решаемых задач. Если решение оказывается целочисленным, то все задачи-кандидаты, оценка которых хуже, чем полученное целочисленное решение, исключаются из списка решаемых задач, так как поиск их решения не имеет смысла (целевая функция этих задач в любом случае будет иметь худшее значение, чем в полученном целочисленном решении). Процесс продолжается, пока не будут решены все задачи, входящие в список решаемых задач. Лучшее из полученных целочисленных решений и является оптимальным решением.

В ходе реализации метода ветвей и границ может быть получено несколько целочисленных решений задачи. Будем называть лучшее из них «текущим наилучшим решением» (ТНР). Значение целевой функции, соответствующее этому решению, будем обозначать как ЕТНР. В начале решения задачи ТНР не существует (еще не найдено). Существует несколько алгоритмов реализации метода ветвей и границ. Рассмотрим один из них.

Примечание. В приведенном ниже описании предполагается, что целевая функция решаемой задачи подлежит максимизации.

1. Исходная задача решается симплекс-методом без учета требований целочисленности. Если все переменные, которые по своему смыслу должны быть целочисленными, принимают в оптимальном решении целочисленные значения, то решение задачи на этом завершается. В противном случае выполняется следующий шаг.

2. Пусть некоторая переменная x_i, которая по своему смыслу должна быть целочисленной, приняла дробное значение R_i (если таких переменных несколько, то выбирается любая из них). Пусть значение целевой функции в оптимальном (но нецелочисленном) решении равно Е. В список решаемых задач включаются две задачи. В одну из них добавляется ограничение $x_i \ge \left[R_i + 1\right]$, в другую - ограничение $x_i \ge \left[R_i\right]$, где [ ] - целая часть. Величина Е используется в качестве оценки этих задач.

3. Из списка решаемых задач выбирается задача с максимальным значением оценки (если имеется несколько задач с одинаковыми оценками, то выбирается любая из них). Эта задача решается симплекс-методом.

4. В зависимости от результатов решения задачи выполняется одно из следующих действий:

- если задача не имеет решений, то выполняется переход к шагу 5;
- если решение задачи оказалось нецелочисленным, и при этом еще не существует ТНР (еще не найдено ни одного целочисленного решения), то выполняется возврат к шагу 2;

- если решение задачи оказалось нецелочисленным, и при этом уже существует ТНР, то значение целевой функции, полученное в задаче (Е), сравнивается с E_{THP}. Если $E > E_{THP}$, то выполняется возврат к шагу 2; если $E \le E_{THP}$, то выполняется переход к шагу 5;
- если решение задачи оказалось целочисленным, и при этом еще не существует ТНР, то найденное решение принимается в качестве ТНР, и выполняется переход к шагу 5;
- если решение задачи оказалось целочисленным, и при этом уже существует ТНР, то значение целевой функции, полученное в задаче (Е), сравнивается с E_{THP}. Если $E > E_{THP}$, то найденное решение принимается в качестве ТНР и выполняется переход к шагу 5; если $E \le E_{THP}$, то выполняется переход к шагу 5.

5. Если в списке решаемых задач имеется хотя бы одна задача, то выполняется возврат к шагу 3. Если в списке нет ни одной задачи, то процесс решения завершен. Если по окончании решения существует ТНР, то оно является оптимальным целочисленным решением задачи. Если ТНР не существует, то задача не имеет целочисленных решений.

Применение метода ветвей и границ рассмотрим на следующем примере.

Пример 4.1. Предприятие выпускает панели для пультов управления двух видов: стандартные (для работы в обычных условиях) и специального исполнения (для работы при повышенных температурах). При изготовлении панелей используется пластмасса и алюминий. Для изготовления одной стандартной панели требуется 12,5 кг пластмассы и 8 кг алюминия; для изготовления одной панели специального исполнения - 7,2 кг пластмассы и 14,5 кг алюминия. Предприятие имеет 300 кг пластмассы и 400 кг алюминия. Прибыль предприятия от выпуска одной стандартной панели составляет 7 ден. ед., прибыль от выпуска одной панели специального исполнения - 9 ден. ед. Требуется определить, сколько панелей каждого вида должно выпускать предприятие, чтобы получить максимальную прибыль.

Обозначим количество выпускаемых стандартных панелей через x_1, количество выпускаемых панелей специального исполнения - через x_2. Математическая модель задачи будет иметь следующий вид:

$$12,5x_1 + 7,2x_2 \le 300$$

$$8x_1 + 14,5x_2 \le 400$$

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

$$x_1, x_2 \quad \text{-целые.}$$

$$E = 7x_1 + 9x_2 \to \max.$$

Приведем задачу к стандартной форме:

$$12,5x_1 + 7,2x_2 + x_3 = 300$$

$$8x_1 + 14,5x_2 + x_4 = 400$$

$$x_i \ge 0, i = 1,..,4.$$

$$x_1, x_2 \text{ -целые.}$$

$$E = 7x_1 + 9x_2 \to \max.$$

Остаточные переменные x_3 и x_4 обозначают, соответственно, неиспользованный остаток пластмассы и алюминия (в килограммах). Эти переменные могут принимать как целые, так и дробные значения. Таким образом, задача является частично целочисленной.

Для решения задачи воспользуемся методом ветвей и границ. Ход решения показан на рис. 4.1. Номера задач соответствуют порядку их включения в список решаемых задач.

Рамкой выделены целочисленные решения. Перечеркнуты постановки задач, для которых не потребовалось определять решение.

Примечание. Следует обратить внимание, что номера задач обозначают именно порядок их включения в список решаемых задач, а не порядок решения.

Исходная задача (на рис. 4.1 она обозначена как задача 1) решается симплекс-методом. Решение оказалось нецелочисленным: $x_1 = 11,89, x_2 = 21,03$. По условию задачи обе переменные должны принимать целочисленные значения. Для продолжения поиска оптимального решения выбирается любая из этих переменных, например, x_1. В список решаемых задач включаются задачи 2 и 3. В задачу 2 входит ограничение $x_1 \le 11$, а в задачу 3 - ограничение $x_1 \ge 12$. Смысл этих ограничений следующий:

- эти ограничения исключают из ОДР оптимальное, но нецелочисленное решение $x_1 = 11,89, x_2 = 21,03$, так как значение $x_1 = 11,89$ не соответствует ни ограничению $x_1 \le 11$, ни ограничению $x_1 \ge 12$. Таким образом, в ходе дальнейшего решения задачи исключается возврат к оптимальному, но нецелочисленному решению;
- эти ограничения не исключают из ОДР ни одного допустимого целочисленного решения, так как любое целочисленное значение x_1 соответствует либо ограничению $x_1 \le 11$, либо ограничению $x_1 \ge 12$.

Оценкой задач 2 и 3 является величина E = 272,46. Эта величина соответствует оптимальному (но нецелочисленному) решению задачи 1 $x_1 = 11,89, x_2 = 21,03$.В задачах 2 и 3 из-за включения дополнительных ограничений ($x_1 \le 11$ для задачи 2, $x_1 \ge 12$ для задачи 3) будет получено какое-то другое решение. Поэтому значения целевых функций в задачах 2 и 3 не могут быть больше, чем 272,46.

Так как задачи 2 и 3 имеют одинаковые оценки, для решения можно выбрать любую из них. Пусть выбрана задача 2. Она решается симплекс-методом. Решение также оказывается нецелочисленным. В список решаемых задач включаются задачи 4 и 5 с оценкой 270,66. Для составления этих задач вводятся ограничения $x_2 \le 21$ и $x_2 \ge 22$. Эти ограничения имеют тот же смысл, что и ограничения, используемые при составлении задач 2 и 3 (см. выше).

Таким образом, в списке решаемых задач находятся три задачи: 3 (с оценкой 272,46), 4 и 5 (с оценкой 270,66). Для решения выбирается задача с максимальной оценкой. В данном случае это задача 3. При ее решении максимальное значение целевой функции может составить не более чем 272,46, а для задач 4 или 5 - не более 270,66.

Примечание. Важно обратить внимание, что оценка задачи показывает именно максимально возможное значение целевой функции этой задачи. Оценка задачи известна до того, как она будет решена. Действительное оптимальное решение задачи можно определить, только решив ее. Поэтому в рассматриваемом примере выбор задачи 3 не означает, что ее решение обязательно будет лучше, чем у задач 4 или 5.

Решение задачи 3 также оказывается нецелочисленным. В список решаемых задач включаются задачи 6 и 7 с оценкой 271,5.

В списке решаемых задач находятся четыре задачи: 4 и 5 (с оценкой 270,66), 6 и 7 (с оценкой 271,5). Для решения можно выбрать любую из задач 6 или 7. Пусть выбрана задача 6. Ее решение оказалось нецелочисленным. В список решаемых задач включаются задачи 8 и 9 с оценкой 267,36.

В списке решаемых задач оказалось пять задач: 4 и 5 (с оценкой 270,66), 7 (с оценкой 271,5), 8 и 9 (с оценкой 267,36). Для решения выбирается задача 7. Она не имеет допустимых решений.

В списке решаемых задач осталось четыре задачи: 4 и 5 (с оценкой 270,66), 8 и 9 (с оценкой 267,36). Для решения можно выбрать задачу 4 или 5. Пусть выбрана задача 4. Ее решение оказалось целочисленным. Оно становится текущим наилучшим решением. Оценки задач 5, 8 и 9, находящихся в списке решаемых задач, сравниваются со значением целевой функции, полученным для целочисленного решения $E_{THP} = 266$. Оценки всех задач лучше, чем величина E_{THP}. Это значит, что при решении этих задач могут быть получены лучшие решения, чем ТНР. Поэтому ни одна из задач не исключается.

В списке решаемых задач осталось три задачи: 5 (с оценкой 270,66), 8 и 9 (с оценкой 267,36). Для решения выбирается задача 5. Ее решение нецелочисленно. Значение целевой функции (E = 268,88) лучше, чем E_{THP} = 266. Поэтому в список решаемых задач включаются две новые задачи (10 и 11) с оценкой 268,66.

Таким образом, в списке решаемых задач находятся четыре задачи: 8 и 9 (с оценкой 267,36), 10 и 11 (с оценкой 268,88). Для решения необходимо выбрать задачу 10 или 11. Пусть выбрана задача 10. Ее решение нецелочисленно. Значение целевой функции (E = 268,62) лучше, чем $E_{THP} = 266$. Поэтому в список решаемых задач включаются две новые задачи (12 и 13) с оценкой 268,62.

В списке решаемых задач находятся пять задач: 8 и 9 (с оценкой 267,36), 11 (с оценкой 268,88), 12 и 13 (с оценкой 268,62). Для решения выбирается задача 11. Она не имеет допустимых решений.

В списке решаемых задач осталось четыре задачи: 8 и 9 (с оценкой 267,36), 12 и 13 (с оценкой 268,62). Для решения можно выбрать задачу 12 или 13. Пусть выбрана задача 12. Ее решение оказалось целочисленным. Значение целевой функции (E = 268) лучше, чем $E_{THP} = 266$. Таким образом, получено лучшее решение, чем ТНР, найденное ранее в задаче 4. Полученное решение X1 = 10, X2 = 22 становится текущим наилучшим решением, а величина E_{THP} принимает значение 268. Кроме того, необходимо сравнить с новым значением E_{THP} оценки всех задач, имеющихся в списке решаемых задач. Оценки задач 8 и 9 (267,36) хуже, чем E_{THP}. Поэтому решать эти задачи не имеет смысла, так как их решение будет хуже, чем уже полученное целочисленное решение (ТНР). Задачи 8 и 9 исключаются из списка решаемых задач. Исключить задачу 13 нельзя, так как ее оценка (268,62) лучше, чем E_{THP}.

В списке решаемых задач остается одна задача: это задача 13. Она решается симплекс-методом. Ее решение нецелочисленно. Значение целевой функции (E = 265,19) хуже, чем ЕТНР. Поэтому включать в список решаемых задач новые задачи не требуется.

В списке решаемых задач не осталось ни одной задачи. Таким образом, получено оптимальное решение: $x_1 = 10, x_2 = 22, E = 268$. Это значит, что предприятию следует выпустить 10 стандартных панелей и 22 панели специального исполнения. Прибыль предприятия в этом случае будет максимальной и составит 268 ден. ед.

Можно также определить, что неизрасходованный остаток запаса пластмассы составит 16,6 кг, а алюминия - 1 кг.

Примечание. Математические модели некоторых задач, составляемых в ходе поиска решения по методу ветвей и границ, можно упростить. Например, в задаче 10 можно заменить ограничения $x_1 \le 11$ и $x_1 \le 10$ на одно ограничение: $x_1 \le 10$. В задаче 12 можно заменить ограничения $x_2 \ge 22$ и $x_2 \le 22$ на $x_2 = 22$. Аналогичные упрощения возможны и в задачах 8, 9, 11, 13.

$$12,5x_1 + 7,2x_2 \le 300$$

$$8x_1 + 14,5x_2 \le 400$$

$$x_1 = 11,89, x_2 = 21,03, E = 272,46$$

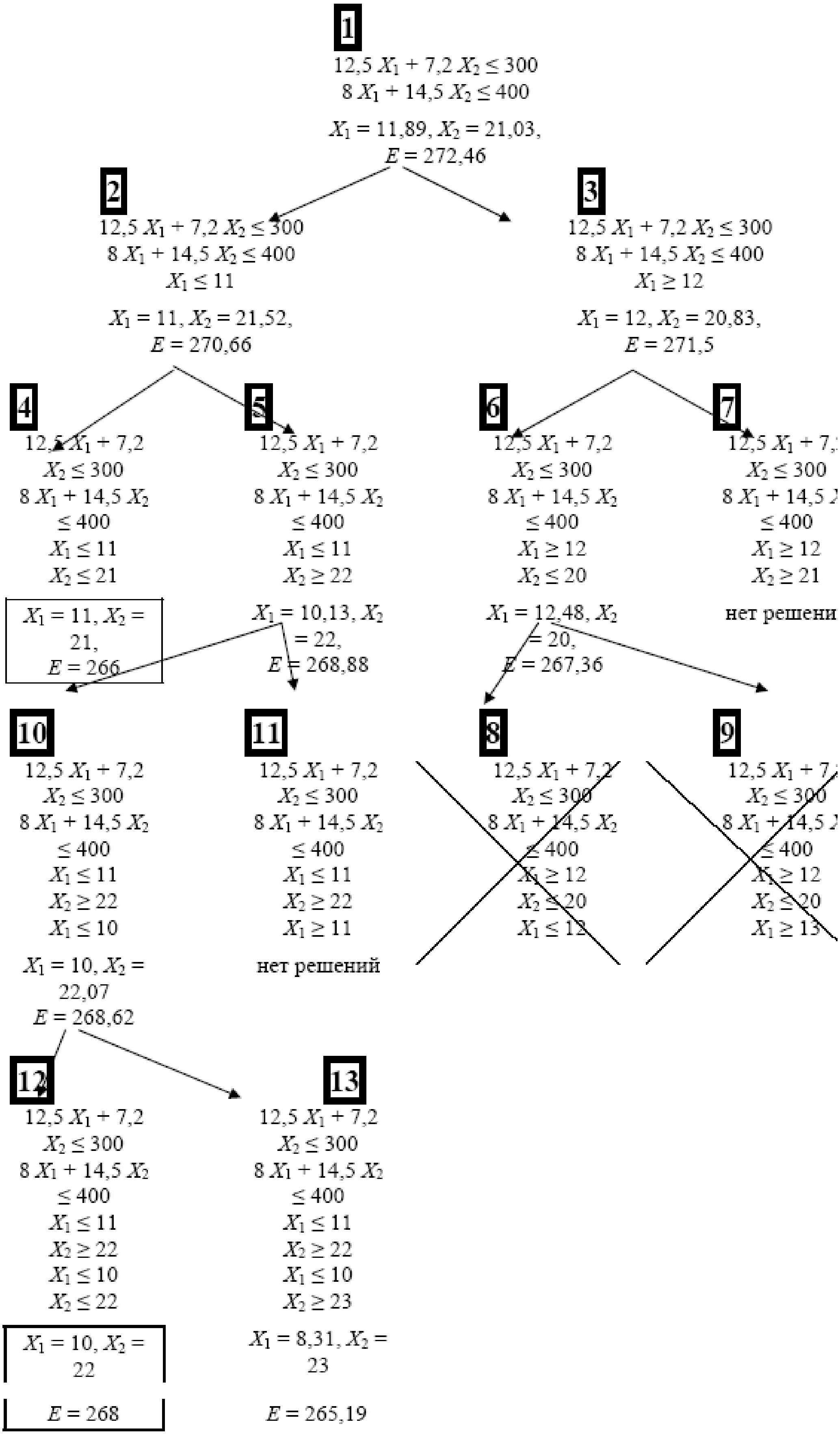

Рис. 4.1. Решение примера 4.1 методом ветвей и границ

5. Оптимизация в задачах транспортного типа

5.1. Постановка задачи

Транспортные задачи представляют собой особый класс задач линейного программирования. Эти задачи, как и любые задачи линейного программирования, могут решаться с использованием симплекс-метода. Однако для решения транспортных задач существуют специальные, более простые методы. Общая постановка транспортной задачи следующая. Имеются m поставщиков некоторого товара. Количество товара, имеющееся у поставщиков, составляет $a_1, a_2, ..., a_m$ единиц. Имеются n потребителей этого товара; их спрос составляет $b_1, b_2, ..., b_n$ единиц. Сумма запасов товара, имеющихся у поставщиков, равна сумме величин спроса всех потребителей:

Известны затраты на перевозку единицы товара от каждого поставщика каждому потребителю (стоимости перевозок):

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j \ .$$

Требуется составить оптимальный план перевозок, т.е. определить, сколько товара каждый поставщик должен доставлять каждому из потребителей, чтобы общие затраты на перевозки были минимальными. При этом, конечно, каждому потребителю должно быть доставлено необходимое количество товара.

Пример 5.1. С четырех складов (СК$_1$, СК$_2$, СК$_3$, СК$_4$) доставляется товар в три магазина (МГ1, МГ2, МГ3). На складе СК1 имеется 40 тонн товара, на складе СК2 — 50 тонн, на складе СК3 — 60 тонн, на складе СК4 — 30 тонн. Магазину МГ1 требуется 60 тонн товара, магазину МГ2 — 80 тонн, магазину МГ3 — 40 тонн. Затраты (в ден. ед.), связанные с перевозкой одной тонны товара с каждого склада в каждый магазин, приведены в табл. 5.1.

Таблица 5.1

Склады	Магазины		
	МГ	МГ2	МГ3
СК1	4	3	5
СК2	6	2	1
СК3	10	4	7
СК4	8	6	3

Требуется определить, сколько товара необходимо перевезти с каждого склада в каждый магазин, чтобы доставить всем магазинам необходимое количество товара с минимальными затратами.

Данную задачу можно представить как задачу линейного программирования. Для построения математической модели этой задачи введем переменные $x_{ij}, i = 1, ..., 4, \ j = 1, ..., 3$, обозначающие количество товара, перевозимого с i-го склада в j-й магазин.

На складах имеется 180 единиц товара; магазинам требуется также 180 единиц товара. Поэтому для удовлетворения спроса всех магазинов потребуется вывезти со складов весь товар. Ограничения, выражающие это требование, имеют следующий вид:

$$x_{11} + x_{12} + x_{13} = 40$$

$$x_{21} + x_{22} + x_{23} = 50$$

$$x_{31} + x_{32} + x_{33} = 60$$

$$x_{41} + x_{42} + x_{43} = 30.$$

Каждый магазин должен получить ровно столько товара, сколько ему требуется. Ограничения, выражающие это условие, следующие:

$$x_{11}+x_{21}+x_{31}+x_{41}=60$$

$$x_{12}+x_{22}+x_{32}+x_{42}=80$$

$$x_{13}+x_{23}+x_{33}+x_{43}=40.$$

Так как переменные обозначают количество перевозимого товара, на них накладывается требование неотрицательности:

$$x_{ij}\geq 0, i=1,...,4, j=1,...,3.$$

Целевая функция представляет собой затраты на выполнение всех перевозок:

$$E=4x_{11}+3x_{12}+5x_{13}+6x_{21}+2x_{22}+1x_{23}+10x_{31}+4x_{32}+7x_{33}+8x_{41}+6x_{42}+3x_{43}\rightarrow \min.$$

Такую задачу можно решить симплекс-методом, как и любую задачу линейного программирования. Однако такое решение окажется достаточно сложным из-за большого количества переменных и ограничений, входящих в математическую модель задачи. Для решения задач такого вида существуют специальные, более простые методы.

При решении транспортной задачи удобно пользоваться расчетной таблицей, содержащей стоимости перевозок, запасы товара у поставщиков и величины спроса потребителей. По ходу решения задачи в нее заносятся величины перевозок (значения переменных x_{ij}), а также вспомогательные величины, используемые для решения задачи.

Расчетная таблица для примера 5.1 показана в табл. 5.2.

Таблица 5.2

Склады	Магазины		
	МГ	МГ2	мгз
СК1	4	3	5
СК2	6	2	1
СК3	10	4	7
СК4	8	6	3

Решение транспортной задачи включает два этапа:

• поиск допустимого решения, т.е. плана перевозок, при котором каждый потребитель получит весь необходимый товар, однако затраты на такие перевозки могут не быть минимальными;

• поиск оптимального решения, т.е. плана перевозок с минимальными затратами.

5.2. Допустимое решение

Имеется несколько методов поиска допустимого решения: метод северо-западного угла, метод минимального элемента, метод Фогеля [1,4].

Рассмотрим поиск допустимого решения на основе метода минимального элемента. Принцип работы этого метода состоит в том, что в первую очередь планируются перевозки, требующие минимальных затрат. Метод минимального элемента реализуется следующим образом. Выбирается минимальная стоимость перевозки единицы груза, т.е. минимальный из коэффициентов целевой функции $c_{ij}, i=1,...,m, j=1,...,n$. Если запас i-го поставщика превосходит спрос j-го потребителя $(a_i > b_j)$, то спрос j-го потребителя полностью удовлетворяется за счет перевозок от i-го поставщика: $x_{ij} = b_j$. При этом запас i-го поставщика уменьшается на величину $b_j (a_i = a_i - b_j)$, а j-й потребитель исключается из дальнейшего рассмотрения, так как он уже получил необходимое количество товара (j-й столбец вычеркивается из расчетной таблицы). Если, наоборот, спрос j-го потребителя превосходит запас, имеющийся у i-го поставщика $(a_i < b_j)$, то i-й поставщик перевозит весь имеющийся у него товар j-му потребителю: $x_{ij} = a_i$. При этом спрос j-го потребителя уменьшается на величину $a_i (b_j = b_j - a_i)$, а i-й поставщик исключается из дальнейшего рассмотрения, так как весь имеющийся у него запас израсходован (i-я строка вычеркивается из расчетной таблицы). В сокращенной расчетной таблице снова выбирается минимальная стоимость перевозок, и выполняются действия, аналогичные рассмотренным выше. Процесс продолжается, пока не будут распределены все запасы и удовлетворены все потребности.

Примечание. Если запас i-го поставщика оказывается равным спросу j-го потребителя $(a_i = b_j)$, то из рассмотрения исключается или поставщик, или потребитель (но не оба вместе). Если исключается поставщик, то предполагается, что неудовлетворенный спрос потребителя составляет 0 единиц товара. Если исключается потребитель, то предполагается, что у поставщика остается запас товара в размере 0 единиц. Дальнейшие действия выполняются обычным образом согласно алгоритму метода минимального элемента. Подробнее такой случай рассмотрен в подразделе 5.5.

Рассмотрим поиск допустимого плана для примера 5.1. Наиболее дешевыми являются перевозки от второго поставщика третьему потребителю, т.е. со склада СК2 в магазин МГ3. Запас товара на складе СК2 составляет 50 тонн, а спрос магазина МГ3 - 40 тонн. Поэтому склад СК1 поставляет магазину МГ3 40 тонн товара. Запас товара на складе СК1 уменьшается на 40 тонн (на складе остается 10 тонн), а магазин МГ3 исключается из рассмотрения, так как для него уже запланирована перевозка необходимого количества товара (табл. 5.3).

В сокращенной расчетной таблице (см. табл. 5.3) выбирается минимальный элемент. Наиболее дешевыми являются перевозки со склада СК2 в магазин МГ2. Запас товара на складе составляет 10 тонн (так как для этого склада уже запланирована перевозка 40 тонн товара в магазин МГ3), а спрос магазина - 80 тонн. Поэтому склад СК2 поставляет магазину МГ2 весь имеющийся у него товар. Склад СК2 исключается из рассмотрения, а спрос магазина МГ2 уменьшается на 10 тонн (табл. 5.4).

В сокращенной расчетной таблице (см. табл. 5.4) минимальную стоимость имеют перевозки со склада СК1 в магазин МГ2. Запас товара на складе СК1 составляет 40 тонн, а спрос магазина МГ2 - 70 тонн (так как для этого магазина уже запланирована перевозка 10 тонн товара со склада СК2). Поэтому склад СК1 поставляет магазину МГ2 весь имеющийся у него товар. Склад СК1 исключается из рассмотрения, а спрос магазина МГ2 уменьшается на 40 тонн (табл. 5.5).

В сокращенной расчетной таблице (см. табл. 5.5) наиболее дешевыми являются перевозки со склада СК3 в магазин МГ2. Запас товара на складе СК3 составляет 60 тонн, а спрос магазина МГ2 - 30 тонн. Поэтому склад СК3 поставляет магазину МГ2 30 тонн товара. Запас товара на складе СК3 уменьшается на 30 тонн (на складе остается еще 30 тонн), а магазин МГ2 исключается из рассмотрения, так как для него уже запланирована перевозка необходимого количества товара (табл. 5.6).

В сокращенной расчетной таблице (см. табл. 5.6) минимальный элемент соответствует перевозкам со склада СК4 в магазин МГ1. Запас товара на складе СК4 составляет 30 тонн, а спрос магазина МГ1 - 60 тонн. Поэтому склад СК4 поставляет магазину МГ1 весь имеющийся у него товар. Склад СК4 исключается из рассмотрения, а спрос магазина МГ1 уменьшается на 30 тонн (табл. 5.7).

Запас товара (30 тонн) остается только на складе СК3. Такое количество товара требуется магазину МГ1 (для всех остальных магазинов уже запланированы необходимые перевозки). Поэтому склад СК3 поставляет магазину МГ1 30 тонн товара. На этом поиск допустимого решения завершается.

Полученный допустимый план перевозок приведен в табл. 5.8. Склад СК1 поставляет 40 тонн товара магазину МГ2; склад СК2 поставляет 10 тонн товара магазину МГ2 и 40 тонн — магазину МГ3; склад СК3 поставляет 30 тонн товара магазину МГ1 и 30 тонн — магазину МГ2; склад СК4 поставляет 30 тонн товара магазину МГ1.

Другими словами, переменные приняли следующие значения: $x_{12}=40, x_{22}=10, x_{23}=40, x_{31}=30, x_{32}=30, x_{41}=30$; остальные переменные равны нулю. Общие затраты на перевозки составят 840 ден. ед. Переменные, принявшие ненулевые значения, называются базисными, остальные (нулевые) — небазисными. Если допустимый план перевозок составлен правильно, то количество базисных переменных всегда равно $m+n-1$, где m — количество поставщиков, n — количество потребителей.

Примечание. Иногда некоторые из базисных переменных также принимают нулевые значения (см. подраздел 5.5). Однако количество базисных переменных всегда должно быть равно $m+n-1$.

5.3. Метод потенциалов. Оптимальное решение

Оптимальное решение (план перевозок с минимальной стоимостью) определяется методом потенциалов на основе допустимого решения, полученного каким-либо из указанных выше способов. Общий вид расчетной таблицы, используемой при поиске оптимального решения на основе метода потенциалов, показан в табл. 5.9.

Смысл обозначений u_i, v_j, c_{ij} и т.д. будет показан ниже. Рассмотрим поиск оптимального решения для примера 5.1.

1. Составляются уравнения для определения вспомогательных величин u_i и $v_j, i=1,...,m, j=1,...,n: u_i+v_j=c_{ij}$, (5.1)

где c_{ij} — стоимости перевозок, соответствующие базисным переменным.

Количество таких уравнений равно $m+n-1$ (т.е. равно количеству базисных переменных). Величины u_i называются платежами (или потенциалами) поставщиков, а v_j — платежами (потенциалами) потребителей.

Система уравнений (5.1) для рассматриваемого примера имеет следующий вид:

$$u_1+v_2=3$$
$$u_2+v_2=2$$
$$u_2+v_3=1$$
$$u_3+v_1=10$$
$$u_3+v_2=4$$
$$u_4+v_1=8.$$

2. Из системы уравнений (5.1) определяются платежи. Так как количество неизвестных (платежей) равно $m+n$, а система состоит из $m+n-1$ уравнения, один из платежей (обычно u_1) принимается равным нулю.

Найдем платежи для рассматриваемого примера. Пусть $u_1=0$. Тогда $v_2=3$ (из первого уравнения). По значению $v_2=3$ из второго уравнения найдем $u_2=-1$. Из третьего уравнения найдем $v_3=2$. Из пятого уравнения по значению $v_2=3$ получим $u_3=1$. Продолжая расчеты аналогичным образом, получим: $v_1=9, u_4=-1$ (платежи указаны в порядке их вычисления). Для дальнейших расчетов удобно указать значения платежей возле соответствующих строк и столбцов расчетной таблицы (см. табл. 5.9, 5.10).

3. Для всех небазисных переменных находятся суммы платежей $\overline{c}_{ij} = u_i + v_j$ (псевдостоимости). Например, $\overline{c}_{11} = u_1 + v_1 = 9, \overline{c}_{13} = u_1 + v_3 = 2$, и т.д. Псевдостоимости удобно вычислять по расчетной таблице. Эти величины также следует занести в расчетную таблицу. В табл. 5.10 они указаны в левых верхних углах ячеек.

4. Для всех небазисных переменных находятся разности стоимостей и псевдостоимостей: $d_{ij} = c_{ij} - \overline{c}_{ij}$. Например, $d_{11} = c_{11} - \overline{c}_{11} = -5, d_{13} = c_{13} - \overline{c}_{13} = 3$, и т.д. В табл. 5.10 они указаны в правых нижних углах ячеек.

5. Если для всех величин d_{ij} выполняется условие $d_{ij} \geq 0$, то оптимальное решение найдено, и решение задачи завершается. Если имеются величины $d_{ij} < 0$, то выполняется следующий шаг.

6. Определяется переменная для включения в базис. В качестве такой переменной выбирается переменная, которой соответствует максимальная по модулю отрицательная величина d_{ij}. Включение переменной x_{ij} в базис означает, что должны выполняться перевозки от i-го поставщика к j-му потребителю. Соответствующая ячейка расчетной таблицы обозначается знаком «плюс».

В рассматриваемом примере имеются две отрицательные величины $d_{11} = -5$ и $d_{21} = -2$. Максимальная по модулю из этих величин - d_{11}. Значит, для включения в базис выбирается переменная x_{11}.

Примечание. Если имеется несколько максимальных по модулю отрицательных величин d_{ij} (равных между собой), то для включения в базис можно выбирать любую из соответствующих переменных.

7. Определяется переменная для исключения из базиса. Для этого строится цикл.

Рассмотрим построение цикла для примера 5.1. Включение в базис переменной x_{11} означает, что будут выполняться перевозки товара со склада СК1 в магазин МГ1. Так как запас товара на складе СК1 ограничен (составляет только 40 тонн), потребуется снизить поставки магазину МГ2, т.е. переменную x_{12}. Для того чтобы магазин МГ2 получил необходимое количество товара (80 тонн), необходимо, чтобы какой-либо из складов, уже поставляющих товар этому магазину, увеличил свои поставки. Магазину МГ2 поставляют товар два склада: СК2 и СК3. Однако склад СК2 не может увеличить поставки магазину МГ2, так как в таком случае этот склад должен был бы уменьшить поставки магазину МГ3, а этот магазин получает весь необходимый товар (40 тонн) только со склада СК2. Поэтому поставки магазину МГ2 увеличивает склад СК3 (переменная x_{32} увеличивается). Так как запас товара на складе СК3 ограничен, из-за увеличения поставок магазину МГ2 этот склад должен уменьшить поставки какому-либо другому магазину, в данном случае - магазину МГ1 (переменная x_{31} уменьшается). Для того чтобы магазин МГ1 получил необходимое количество товара, поставки в этот магазин будет выполнять склад СК1. Таким образом, цикл построен. Переменные, которые требуется увеличить, обозначаются знаком «плюс», а уменьшаемые - знаком «минус». Цикл показан в табл. 5.10.

Примечание. Если в расчетной таблице оказалась хотя бы одна величина $d_{ij} < 0$, то всегда можно построить цикл, причем только один.

Минимальная из переменных, отмеченных знаком «минус», исключается из базиса. Обозначим эту переменную как x_{rs}. В данном случае это переменная x_{31} Исключение переменной x_{rs} из базиса означает, что перевозки от r-го поставщика s-му потребителю прекращаются.

Примечание. Если несколько переменных, отмеченных знаком «минус», имеют одинаковое минимальное значение, то для исключения из базиса выбирается любая из них.

8. Определяются новые значения базисных переменных (новый план перевозок). Все

переменные, отмеченные знаком «плюс», увеличиваются на x_{rs}, а отмеченные знаком «минус» — уменьшаются на эту же величину. Значения остальных переменных не изменяются.

В данном примере $x_{rs} = x_{31} = 30$ (таким образом, перевозки со склада СК3 в магазин МГ1 выполняться не будут). Переменные, отмеченные в цикле знаком «плюс», увеличиваются на 30; это значит, что соответствующие перевозки увеличиваются на 30 тонн. Переменные, отмеченные знаком «минус», уменьшаются на 30. Новый план перевозок показан в табл. 5.11. Умножив величины перевозок (x_{ij} на соответствующие стоимости перевозки единицы груза (c_{ij}), найдем целевую функцию: E = 690.Таким образом, в результате перехода к новому решению затраты на перевозки снизились.

Примечание. Если несколько переменных принимают нулевые значения, то из базиса исключается только одна из них - переменная x_{rs}, выбранная на шаге 7. Остальные переменные остаются базисными, хотя и имеют нулевые значения. Подробнее такой случай рассмотрен в подразделе 5.5.

9. Выполняется возврат к шагу 1.

Продолжим поиск оптимального решения для примера 5.1. Для нового базиса составим систему уравнений (5.1), чтобы определить платежи:

$$u_1 + v_1 = 4$$
$$u_1 + v_2 = 3$$
$$u_2 + v_2 = 2$$
$$u_2 + v_3 = 1$$
$$u_3 + v_2 = 4$$
$$u_4 + v_1 = 8.$$

Принимая U1 = 0, найдем остальные платежи: $v_1 = 4, v_2 = 3, u_2 = -1, v_3 = 2, u_3 = 1, u_4 = 4$ (платежи указаны в порядке их вычисления). Затем определяются значения псевдостоимостей, далее — разности стоимостей и псевдостоимостей. Все эти величины приведены в табл. 5.11.

В таблице имеются отрицательные разности стоимостей и псевдостоимостей: это величины $d_{42} = -1$ и $d_{43} = -3$. Это означает, что полученное решение не является оптимальным. Переменная x_{43} включается в базис. Для определения переменной, исключаемой из базиса, строится цикл.

Включение в базис переменной x_{43} означает, что будут выполняться перевозки товара со склада СК4 в магазин МГ3. Поэтому требуется уменьшить перевозки со склада СК4 какому-либо магазину (так как запас товара на складе ограничен). В данном случае можно уменьшить только поставки магазину МГ1. Чтобы магазин МГ1 получил необходимое количество товара, требуется увеличение поставок со склада СК1, т.е. увеличение переменной x_{11}. Из-за увеличения поставок магазину МГ1 склад СК1 должен уменьшить поставки магазину МГ2 (уменьшается переменная x_{12}). Чтобы магазин МГ2 получил необходимое количество товара, необходимо, чтобы какой-либо из складов, уже поставляющих товар этому магазину, увеличил свои поставки. Магазину МГ2 поставляют товар склады СК2 и СК3; однако склад СК3 не может увеличить свои поставки, так как он поставляет магазину МГ2 весь имеющийся у него товар (60 тонн). Поэтому поставки магазину МГ2 увеличивает склад СК2 (увеличивается переменная x_{22}). Из-за этого склад СК2 должен снизить поставки магазину МГ3 (уменьшается переменная x_{23}). Для того чтобы магазин МГ3 получил необходимое количество товара, поставки в этот магазин будет выполнять склад СК4 (переменная x_{43} включается в базис). Таким образом, цикл построен

(см. табл. 5.11).
Из базиса исключается переменная $x_{12} = 10$, так как эта переменная — минимальная из отмеченных знаком «минус». Это значит, что перевозки со склада СК1 в магазин МГ2 выполняться не будут.

Определяется новый план перевозок: переменные, отмеченные в цикле знаком «плюс», увеличиваются на $x_{rs} = x_{12} = 10$, а отмеченные знаком «минус» — уменьшаются на 10. Новый план перевозок показан в табл. 5.12. Целевая функция (стоимость всех перевозок) для этого плана равна Е = 660 ден. ед.

Для полученного плана перевозок составим систему уравнений (5.1), чтобы определить платежи:

$$u_1 + v_1 = 4$$
$$u_2 + v_2 = 2$$
$$u_2 + v_3 = 1$$
$$u_3 + v_2 = 4$$
$$u_4 + v_1 = 8$$
$$u_4 + v_3 = 3.$$

В табл. 5.13 приведены платежи (u_i и v_j), псевдостоимости $(\overline{c}_{ij})$, разности стоимостей и псевдостоимостей (d_{ij}). Все величины d_{ij} неотрицательны. Это означает, что полученное решение оптимально.

Таким образом, оптимальный план перевозок состоит в следующем. Склад СК1 поставляет 40 тонн товара магазину МГ1; склад СК2 поставляет 20 тонн товара магазину МГ2 и 30 тонн — магазину МГ3; склад СК3 поставляет 60 тонн товара магазину МГ2; склад СК4 поставляет 20 тонн товара магазину МГ1 и 10 тонн — магазину МГ3. Другими словами, переменные приняли следующие значения: $x_{11} = 40, x_{22} = 20, x_{23} = 60, x_{41} = 20, x_{43} = 10$; остальные переменные равны нулю. Общие затраты на перевозки составят 660 ден. ед.

5.4. Транспортные задачи с неправильным балансом

Транспортные задачи с неправильным балансом — это задачи, в которых сумма запасов товара, имеющихся у поставщиков, неравна сумме величин спроса потребителей.

Все методы решения транспортных задач предназначены для задач с правильным балансом. Поэтому для решения задачи с неправильным балансом ее необходимо привести к правильному балансу, т.е. преобразовать в обычную транспортную задачу с правильным балансом. Способы такого преобразования зависят от постановки задачи. Полученная задача с правильным балансом решается обычными методами, как показано выше.

Пример 5.2. Пусть в условиях примера 5.1 на складе СК3 имеется только 45 тонн товара. Требуется составить план перевозок, при котором затраты на их выполнение будут минимальными.

В данной задаче сумма запасов товара на складах составляет 165 тонн, а потребности магазинов - 180 тонн. Поэтому при любом решении некоторые магазины получат меньше товара, чем им требуется. Задача состоит только в минимизации затрат на доставку имеющегося товара (165 тонн).

Для приведения этой задачи к правильному балансу необходимо добавить фиктивного поставщика. Его запас принимается равным разности между фактическими запасами, имеющимися на складах, и потребностями магазинов, т.е. 15 тонн. Обозначим фиктивный склад как СК5. Стоимости перевозки единицы товара с этого склада в любой магазин будем считать равными нулю. Расчетная таблица для этой задачи показана в табл. 5.14.

Таким образом, получена задача с правильным балансом. Она решается точно так же, как обычная транспортная задача.

Сначала требуется найти допустимый план перевозок. Воспользуемся для этого методом минимального элемента. Минимальные стоимости соответствуют перевозкам со склада СК5 (т.е. фиктивного склада) в любой магазин; эти стоимости равны нулю. Выберем перевозки со склада СК5 в магазин МГ1. Запас товара на складе СК5 считается равным 15 тонн, а спрос магазина МГ1 составляет 60 тонн. Поэтому поставки со склада СК5 магазину МГ1 принимаются равными 15 тонн. Склад СК5 исключается из рассмотрения (строка СК5 вычеркивается из расчетной таблицы), а спрос магазина МГ1 уменьшается на 15 тонн.

Аналогично выполняется дальнейшее построение допустимого плана по методу минимального элемента. Полученный допустимый план приведен в табл. 5.15. Значение целевой функции (стоимость перевозок) для этого плана составляет 690 ден. ед.

На основе допустимого плана перевозок определяется оптимальный план. Для этого используется метод потенциалов. Оптимальный план приведен в табл. 5.16.

Оптимальный план перевозок состоит в следующем. Склад СК1 поставляет 40 тонн товара магазину МГ1; склад СК2 поставляет 35 тонн товара магазину МГ2 и 15 тонн - магазину МГ3; склад СК3 поставляет 45 тонн товара магазину МГ2; склад СК4 поставляет 5 тонн товара магазину МГ1 и 25 тонн - магазину МГ3. Поставки с фиктивного склада, вошедшие в оптимальный базис, означают, что соответствующим потребителям будет поставлено меньше товара, чем им требуется. Таким образом, магазину МГ1 будет поставлено на 15 тонн товара меньше, чем ему требуется (т.е. всего 45 тонн вместо 60 необходимых). Общие затраты на перевозки составят 540 ден. ед.

Пример 5.3. В условиях примера 5.2 требуется составить план перевозок с минимальными затратами. При этом необходимо выполнить следующее дополнительное требование: спрос магазина МГ1 должен быть обязательно удовлетворен полностью.

Как и в примере 5.2, для приведения этой задачи к правильному балансу необходимо добавить фиктивного поставщика (склад СК5) с запасом товара в размере 15 тонн. Однако стоимость перевозки единицы товара от фиктивного поставщика в магазин МГ1 следует считать очень большим числом (например, 1000 ден. ед.). Стоимости перевозки единицы товара от фиктивного поставщика в другие магазины будем считать равными нулю. Расчетная таблица для этой задачи показана в табл. 5.17.

Получена задача с правильным балансом. Она решается точно так же, как обычная транспортная задача. Сначала по методу минимального элемента определяется допустимый план (табл. 5.18), затем по методу потенциалов - оптимальный план (табл. 5.19). Эти методы обеспечивают получение плана перевозок с минимальными затратами. Поэтому перевозки с фиктивного склада СК5 в магазин МГ1 не войдут в базис (т.е. примут нулевое значение), так как затраты на эти перевозки предполагаются очень большими.

Для данной задачи получен следующий оптимальный план перевозок. Склад СК1 поставляет 40 тонн товара магазину МГ1; склад СК2 поставляет 20 тонн товара магазину МГ2 и 30 тонн - магазину МГ3; склад СК3 поставляет 45 тонн товара магазину МГ2; склад СК4 поставляет 20 тонн товара магазину МГ1 и 10 тонн - магазину МГ3. Магазину МГ2 будет поставлено на 15 тонн товара меньше, чем ему требуется (т.е. всего 65 тонн вместо 80 необходимых). Общие затраты на перевозки составят 600 ден. ед.

Пример 5.4. В условиях примера 5.2 требуется составить план перевозок с минимальными затратами. При этом необходимо выполнить следующее дополнительное требование: недопоставка товара распределяется равномерно между всеми магазинами. Это означает, что каждый из магазинов должен получить несколько меньше товара, чем ему требуется.

Для приведения такой задачи к правильному балансу предполагается, что спрос всех потребителей снизился. Для этого все величины спроса умножаются на коэффициент $\sum_{j=1}^{n} a_i \Big/ \sum_{i=1}^{m} b_j$, т.е. на отношение суммы величин спроса к сумме запасов. Вводить фиктивного поставщика в этом случае не требуется.

В рассматриваемом примере сумма величин спроса равна 180, сумма запасов -165.

Значит, величины спроса умножаются на коэффициент, равный 165/180 = 0,917.

Можно сказать, что таким образом спрос магазинов корректируется с учетом возможностей поставщиков (складов). Уменьшенная величина спроса для магазина МГ1 составит 60·0,917 = 55,02 тонны, для МГ2 - 80·0,917 = 73,36 тонны, для МГ3 - 40·0,917 = 36,68 тонны. Округлим эти величины до целых чисел (будем считать, что величины спроса магазинов должны выражаться целыми числами). Расчетная таблица для данной задачи приведена в табл. 5.20.

Полученная задача с правильным балансом решается обычным образом. Оптимальный план для рассматриваемой задачи приведен в табл. 5.21.

Таким образом, получен следующий оптимальный план перевозок. Склад СК1 поставляет 40 тонн товара магазину МГ1; склад СК2 поставляет 28 тонн товара магазину МГ2 и 22 тонны - магазину МГ3; склад СК3 поставляет 45 тонн товара магазину МГ2; склад СК4 поставляет 15 тонн товара магазину МГ1 и 15 тонн - магазину МГ3. Недопоставка товара магазину МГ1 составит 60 - 55 = 5 тонн, магазину МГ2 - 80 -73 = 7 тонн, магазину МГ3 - 40-37 = 3 тонны. Общие затраты на перевозки составят 583 ден. ед.

Аналогично решаются транспортные задачи в случаях, когда сумма запасов товара у поставщиков больше, чем сумма величин спроса потребителей. При любом решении таких задач часть товара остается у поставщиков.

Если требуется только обеспечить минимальную стоимость перевозок (без каких-либо дополнительных условий), то для приведения задачи к правильному балансу добавляется фиктивный потребитель; его спрос принимается равным излишку запасов товара, т.е. величине $\sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j$. Стоимости перевозок к этому потребителю принимаются равными нулю. После этого задача решается обычным образом.

Если требуется также обеспечить вывоз всего товара у какого-либо поставщика, то стоимость перевозки товара от этого поставщика фиктивному потребителю принимается равной очень большому числу. Если требуется распределить излишек товара по всем поставщикам, то запасы товара у всех поставщиков искусственно уменьшаются. Для этого необходимо умножить все величины запасов на коэффициент $\sum_{j=1}^{n} b_j \Big/ \sum_{i=1}^{m} a_i$. Вводить фиктивного потребителя в этом случае не требуется.

5.5. Вырожденное решение

Для того чтобы иметь возможность решить систему уравнений (5.1) при поиске оптимального плана перевозок по методу потенциалов, необходимо, чтобы количество базисных переменных всегда было в точности равно $m+n-1$. Если некоторые из базисных переменных принимают нулевые значения, то решение называется вырожденным. Необходимо учитывать некоторые особенности поиска оптимального плана при появлении вырожденного решения. Такое решение возникает в следующих случаях:

- при поиске допустимого плана - в случае, если запас товара у поставщика оказывается в точности равен спросу потребителя;
- при поиске оптимального плана - в случае, если при выборе переменной для исключения из базиса оказывается несколько переменных, имеющих одинаковое минимальное значение.

Рассмотрим пример, в котором возникают оба случая вырожденного решения.

Пример 5.5. С двух карьеров (К1 и К2) поставляются стройматериалы на три стройки (С1, С2, С3). Возможности карьеров по поставке стройматериалов (тонны), потребности строек (тонны) и стоимости перевозок одной тонны стройматериалов с каждого карьера на каждую из строек (ден. ед.) приведены в табл. 5.22.

Требуется составить план перевозок стройматериалов с карьеров на стройки, при котором затраты на перевозки будут минимальными.

Решим эту задачу, как показано в подразделах 5.2 и 5.3.

Для определения допустимого решения воспользуемся методом минимального элемента. Наиболее дешевыми (6 ден. ед. за тонну) являются перевозки с карьера К2 на стройку С2. Карьер К2 поставляет стройке С2 15 тонн стройматериалов, так как карьер может поставить 45 тонн, а стройке требуется 15 тонн. Стройка С2 исключается из рассмотрения (второй столбец вычеркивается из расчетной таблицы), а в карьере К2 остается 30 тонн стройматериалов.

В сокращенной расчетной таблице минимальный элемент соответствует перевозкам с карьера К2 на стройку С1. Запас стройматериалов в карьере К2 в точности равен потребности стройки С1 (30 тонн). Карьер К2 поставляет стройке С1 30 тонн стройматериалов. Из рассмотрения можно исключить или поставщика (карьер К2), или потребителя (стройку С1), но не обоих вместе. Пусть исключается карьер К2 (вторая строка вычеркивается из расчетной таблицы). При этом считается, что спрос стройки С1 остается неудовлетворенным на 0 тонн стройматериалов.

В оставшейся части расчетной таблицы выбирается минимальный элемент. Он соответствует перевозкам с карьера К1 на стройку С1 (14 ден. ед. за тонну). Карьер может поставить 30 тонн стройматериалов, а стройке требуется 0 тонн. Поэтому величина поставок стройматериалов с карьера К1 на стройку

С1 принимается равной нулю, однако переменная $x_{11} = 0$ включается в базис.

Стройка С1 исключается из рассмотрения.

В карьере СК3 имеется запас стройматериалов в размере 30 тонн. Такое количество стройматериалов требуется стройке С3. Поэтому карьер К1 поставляет стройке С3 30 тонн стройматериалов. На этом поиск допустимого плана перевозок завершается. Он приведен в табл. 5.23.

Полученный допустимый план перевозок состоит в следующем. Карьер К1 поставляет 30 тонн стройматериалов стройке С3; карьер К2 поставляет 30 тонн стройматериалов стройке С1 и 15 тонн - стройке С2. Общие затраты на перевозки составят 1080 ден. ед.

Переменная $x_{11} = 0$ включена в базис только для того, чтобы количество базисных переменных было равно $m+n-1$ (в данном случае — 4). Никаких перевозок с карьера К1 на стройку С1 не выполняется.

Найдем оптимальный план перевозок, используя метод потенциалов. Составим систему уравнений для определения платежей:

$$u_1 + v_1 = 14$$
$$u_1 + v_3 = 25$$
$$u_2 + v_1 = 8$$
$$u_2 + v_2 = 6.$$

Принимая $u_1 = 0$, найдем платежи: $v_1 = 14, v_3 = 25, u_2 = -6, v_2 = 12$. Найдем псевдостоимости $\overline{c}_{ij} = u_i + v_j$, затем — разности стоимостей и псевдостоимостей $d_{ij} = c_{ij} - \overline{c}_{ij}$. Эти величины приведены в табл. 5.24. Величина d_{23} оказалась отрицательной. Это означает, что имеющийся план перевозок не является оптимальным. Переменная x_{23} включается в базис. Для определения нового плана перевозок строится цикл (см. табл. 5.24).

Обе переменные, обозначенные знаком «минус» (x_{13} и x_{21}), имеют одинаковые значения, равные 30. Любую из этих переменных (но только одну) можно исключить из базиса. Пусть исключается переменная x_{21}. Вычисляются новые значения базисных переменных: все переменные, отмеченные знаком «плюс», увеличиваются на 30 (т.е. на величину переменной, исключаемой из базиса), а переменные, отмеченные знаком «минус» — уменьшаются на ту же величину. Переменная x_{13} становится равной нулю, но остается в базисе (так как исключается из базиса только одна переменная - x_{21}). Новый план перевозок приведен в табл. 5.25.

Проверим полученный план на оптимальность. Составим систему уравнений для определения платежей:

$$u_1 + v_1 = 14$$
$$u_1 + v_3 = 25$$
$$u_2 + v_2 = 6$$
$$u_2 + v_3 = 10.$$

Найдем платежи: $u_1 = 0, v_1 = 14, v_3 = 25, u_2 = -15, v_2 = 21$. Найдем псевдостоимости, затем - разности стоимостей и псевдостоимостей. Эти величины приведены в табл. 5.26.

Все разности стоимостей и псевдостоимостей оказались положительными. Это означает, что получено оптимальное решение. Таким образом, план перевозок должен быть следующим. Карьер К1 поставляет 30 тонн стройматериалов стройке С1; карьер К2 поставляет 15 тонн стройматериалов стройке С2 и 30 тонн - стройке С3. Общие затраты на перевозки составят 810 ден. ед.

6. Оптимизация принятия решений на основе методов нелинейного программирования

6.1. Постановка задачи нелинейного программирования

Задачи нелинейного программирования — это задачи, в которых (как и в задачах линейного программирования) требуется найти значения переменных $x_1, x_2, ..., x_n$, обеспечивающие максимальное или минимальное значение некоторой целевой функции при соблюдении системы ограничений. При этом целевая функция и/или некоторые из ограничений являются нелинейными, т.е. содержат нелинейные составляющие, например, $x_1^2, x_1 x_2, \sqrt{x_1}$ и т.д.

Как и в задачах линейного программирования, любые значения переменных, $x_1, x_2, ..., x_n$ удовлетворяющие ограничениям задачи, называются допустимыми решениями, а все множество допустимых решений — областью допустимых решений (ОДР). Допустимые значения переменных $x_1, x_2, ..., x_n$, при которых целевая функция принимает экстремальное значение, представляют собой оптимальное решение. В задачах нелинейного программирования оптимальное решение может находиться как на границе, так и внутри ОДР.

Так как математические модели задач нелинейного программирования очень разнообразны, не существует универсальных методов, позволяющих решать любую задачу нелинейного программирования. Разработано большое количество методов, каждый из которых предназначен для решения определенного класса задач. Классификация методов нелинейного программирования приведена в табл. 6.1.

Таблица 6.1 Классификация методов нелинейного программирования

Признак классификации	Классы методов	Описание
Используемый математический аппарат	Аналитические (классические)	Поиск решения на основе использования необходимых и достаточных условий экстремумов функций (например, равенства нулю производных)
	Численные	Поиск решения на основе постепенного сужения диапазона, в котором ищется решение
	Покоординатные	Поиск решения на основе поочередного поиска экстремума целевой функции по каждой из

		переменных
	Градиентные	Поиск решения на основе использования градиента целевой функции
	Случайного поиска	Поиск решения на основе многократного случайного выбора возможных решений
Решаемые задачи	Методы решения задач без ограничений	Предназначены для решения задач, в которых требуется найти экстремум нелинейной функции без ограничений на значения переменных
	Методы квадратичного программирования	Предназначены для решения задач с линейными ограничениями и квадратичной целевой функцией, т.е. целевой функцией, содержащей квадраты переменных (например, x_1^2) и/или их попарные произведения (например, $x_1 x_2$)
	Методы сепарабельного программирования	Предназначены для решения задач, в которых ограничения и целевая функция являются сепарабельными, т.е. могут быть представлены в виде сумм функций одной переменной

6.2. Примеры задач нелинейного программирования

Пример 6.1. Для транспортировки некоторого химиката требуется изготовить контейнеры. Требования к контейнерам следующие: 1) емкость контейнера — 6м^3 ; 2) высота может составлять от 1 до 3 м; 3) основание контейнера должно быть квадратным. Дно и стенки контейнера, непосредственно соприкасающиеся с химикатом, должны быть изготовлены из более стойкого материала, чем крышка контейнера. Стоимость материала дна и стенок контейнера — 6 ден. ед./м^2 , стоимость материала крышки — 4 ден. ед./м^2 . Требуется найти габаритные размеры контейнера (размеры основания и высоту), при которых его стоимость будет минимальной. Обозначим высоту контейнера как H, а размеры его основания (длину и ширину) - как L. Тогда для данной задачи можно построить следующую математическую модель:

$$H \geq 1$$

$$H \leq 3$$

$$H \cdot L^2 = 6$$

$$E = 6 \cdot L^2 + 24 \cdot H \cdot L + 4 \cdot L^2 \to \min$$

Здесь первое и второе ограничения устанавливают, что высота контейнера должна составлять от 1 до 3 м; третье ограничение устанавливает, что емкость контейнера равна 6 м^3. Целевая функция E выражает стоимость контейнера (первое слагаемое - стоимость материала для основания, второе - стоимость материала для стенок, третье - для крышки). Данная задача представляет собой задачу нелинейного программирования: нелинейными здесь являются целевая функция и ограничение на емкость контейнера.

Пример 6.2. Предприятие выпускает электроприборы двух типов (А и В) и запасные части к ним. В комплект запасных частей, выпускаемых вместе с каждым прибором, может входить от трех до шести запасных частей, причем количество запасных частей для всех приборов одного типа должно быть одинаковым. Расход материалов на выпуск приборов и запасных частей следующий.

Таблица 6.2

Материал	Прибор А	Запасная часть к прибору А	Прибор В	Запасная часть к прибору В
Провод, см Пластмасса, г	7 12	3 2	10 8	2 1,5

Предприятие имеет возможность израсходовать на выпуск приборов не более 0,6 м провода и не более 0,5 кг пластмассы.

Прибыль предприятия от выпуска одного прибора А составляет 8 ден. ед., одной запасной части к прибору А - 2 ден. ед., одного прибора В - 9 ден. ед., одной запасной части к прибору В - 1,5 ден. ед.

Требуется определить, сколько приборов каждого типа и запасных частей к ним должно выпустить предприятие, чтобы получить максимальную прибыль.

Введем переменные: x_1 - количество приборов типа А; x_2 — количество запасных частей к ним; x_3 - количество запасных частей в комплекте к одному прибору типа А (соотношение количества запасных частей и приборов типа А); x_4 - количество приборов типа В; x_5 - количество запасных частей к ним; x_6 - количество запасных частей в комплекте к одному прибору типа В (соотношение количества запасных частей и приборов типа В).

Составим математическую модель задачи:

$$x_2 = x_1 x_3$$
$$x_3 \le 6$$
$$x_3 \ge 3$$
$$x_5 = x_4 x_6$$
$$x_6 \le 6$$
$$x_6 \ge 3$$
$$7x_1 + 3x_2 + 10x_4 + 2x_5 \le 600$$
$$12x_1 + 2x_2 + 8x_4 + 1,5x_5 \le 500$$
$$x_i - целые, i = 1,...,6$$
$$x_i \ge 0, i = 1,...,6$$
$$E = 8x_1 + 2x_2 + 9x_4 + 1,5x_5 \to \max.$$

Здесь первое ограничение устанавливает, что общее количество запасных частей к приборам типа А (x_2) должно быть равно произведению количества этих приборов (x_1) на количество запасных частей в одном комплекте (x_3). Второе и третье ограничения устанавливают, что количество запасных частей к одному прибору типа А должно составлять не менее трех и не более шести. Четвертое, пятое и шестое ограничения имеют тот же смысл, но для приборов типа В. Седьмое и восьмое ограничения устанавливают предельный расход провода и пластмассы. Целевая функция выражает прибыль от выпуска приборов и запасных частей. В этой задаче нелинейными являются первое и четвертое ограничения.

Пример 6.3. Предприятие выпускает электронные изделия двух типов (изделия А и В). На выпуск изделий расходуется платина и палладий. На одно изделие А требуется 13 г платины и 8 г палладия, на одно изделие В — 6 г платины и 11 г палладия. Предприятие имеет возможность использовать не более 90 г платины и не более 88 г палладия. Изделия А продаются по цене 12 тыс. ден. ед., изделия В - по 10 тыс. ден. ед.

Величины себестоимости изделий (т.е. затраты на их выпуск) зависят от объема их производства и приближенно описываются следующими формулами:

• себестоимость одного изделия А: $7+0,2x_1$, где x_1 - объем производства изделий А;

• себестоимость одного изделия В: $8+0,2x_2$, где x_2 — объем производства изделий В. Требуется составить план производства, обеспечивающий предприятию максимальную прибыль.

Как отмечено выше, объемы производства изделий А и В обозначены через переменные x_1 и x_2. Составим целевую функцию задачи, выражающую прибыль от производства изделий. Будем считать, что прибыль от продажи одного изделия представляет собой разность его цены и себестоимости. Прибыль от продажи одного изделия А можно выразить следующей формулой: $12-(7+0,2x_1)=5-0,2x_1$. Аналогично выразим прибыль от продажи одного изделия В: $10-(8+0,2x_2)=2-0,2x_2$. Таким образом, целевая функция задачи (прибыль от продажи всех изделий А и В) имеет следующий вид:

$$E=(5-0,2x_1)x_1+(2-0,2x_2)x_2=5x_1-0,2x_1^2+2x_2-0,2x_2^2\to\max.$$

Приведем математическую модель задачи:

$$\begin{aligned}&13x_1+6x_2\le 90\\&8x_1+11x_2\le 88\\&x_1,x_2\ge 0\\&x_1,x_2-\textit{целые}\\&E=5x_1-0,2x_1^2+2x_2-0,2x_2^2\to\max.\end{aligned}$$

Здесь ограничения устанавливают предельный расход платины и палладия.

В этой задаче система ограничений линейная, а целевая функция — нелинейная (квадратичная). Таким образом, данная задача представляет собой задачу нелинейного квадратичного программирования.

6.3. Решение задач нелинейного программирования. Метод Франка-Вульфа

Наиболее универсальными из методов нелинейного программирования являются градиентные методы. Эти методы основаны на использовании градиента целевой функции. Напомним, что градиент любой функции нескольких переменных $F(x_1,x_2,...,x_n)$ — это вектор, координаты которого представляют собой частные производные этой функции:

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

Из высшей математики известно важное свойство градиента: он указывает направление наискорейшего возрастания функции. Вектор —grad F называется антиградиентом функции F и указывает направление ее наискорейшего убывания.

Принцип работы всех градиентных методов заключается в пошаговом переходе от некоторого начального решения к новым решениям в направлении градиента (для задач, в которых требуется максимизация целевой функции) или антиградиента (для задач минимизации). В большинстве случаев решение задачи на основе градиентных методов включает следующие основные этапы.

1. Определяется некоторое начальное допустимое решение задачи.

2. Выполняется переход от текущего решения к новому решению в направлении градиента (или антиградиента). Величина этого перехода определяется по-разному в зависимости от условий задачи и используемого метода.

3. Определяется разность значений целевой функции для нового и предыдущего решений. Если эта разность мала (не превышает некоторой заданной точности), то найденное решение принимается в качестве оптимального. В противном случае выполняется возврат к

шагу 2.

Рассмотрим один из градиентных методов - метод Франка-Вульфа. Этот метод предназначен для решения задач с линейной системой ограничений и нелинейной целевой функцией. Метод наилучшим образом подходит для решения задач с квадратичной целевой функцией, хотя может применяться и для решения задач с целевыми функциями другого вида. Решим задачу из примера 6.3, используя метод Франка—Вульфа.

Предварительно найдем градиент целевой функции:

$$gradE\left(\frac{\partial E}{\partial x_1}, \frac{\partial E}{\partial x_2}\right) = (5-0,4x_1, 2-0,4x_2).$$

Найдем также начальное допустимое решение. В качестве такого решения можно использовать любой набор значений x_1 и x_2, удовлетворяющий системе ограничений. Начальное допустимое решение можно найти, например, следующим образом: исключить из целевой функции все нелинейные элементы и решить симплекс-методом полученную задачу линейного программирования. Для рассматриваемого примера такая задача будет следующей:

$$13x_1 + 6x_2 \le 90$$
$$8x_1 + 11x_2 \le 88$$
$$x_1, x_2 \ge 0$$
$$E = 5x_1 + 2x_2 \to \max.$$

Решив эту задачу, получим начальное допустимое решение: $x_1^{(0)} = 6,923, x_2^{(0)} = 0$. Найдем значение целевой функции для этого решения:

$$E^{(0)} = 5 \cdot 6,923 - 0,2 \cdot 6,923^2 + 2 \cdot 0 - 0,2 \cdot 0^2 = 25,03.$$

Необходимо также задать требуемую точность решения задачи (ε). Зададим ее равной 500 ден. ед., т.е. будем считать, что решение найдено, если переход к новому решению приводит к увеличению целевой функции не более чем на 500 ден. ед. В данной задаче целевая функция выражается в тысячах ден. ед., поэтому ε = 0,5. Решим задачу, используя итерационный алгоритм на основе метода Франка-Вульфа.

Итерация 1

1. Определяется градиент целевой функции в точке ОДР, соответствующей текущему решению:

$gradE(x^{(0)}) = (5-0,4 \cdot 6,923; 2-0,4 \cdot 0) = (2,2; 2).$

2. Определяется угловая точка ОДР, соответствующая предельно допустимому (без нарушения ограничений) перемещению от текущего решения в направлении градиента. Для этого решается задача линейного программирования с исходной системой ограничений и целевой функцией, коэффициентами которой являются координаты градиента:

$$13x_1 + 6x_2 \le 90$$
$$8x_1 + 11x_2 \le 88$$
$$x_1, x_2 \ge 0$$
$$W = 2,2x_1 + 2x_2 \to \max.$$

Решение этой задачи следующее: $x_1^* = 4,863, x_2^* = 4,463$. Это означает, что поиск нового решения будет выполняться в направлении от точки $x^{(0)} = (6,923; 0)$ к точке $x^* = (4,863; 4,463)$.

3. Составляются уравнения для перехода к новому решению:

где λ - коэффициент, задающий величину перемещения от текущего решения к новому решению в направлении точки x^*. Этот коэффициент определяется на следующем шаге.

Для рассматриваемого примера эти уравнения имеют следующий вид:

$$x_1^{(1)} = 6,923 + \lambda(4,863 - 6,923) = 6,923 - 2,06\lambda;$$
$$x_2^{(1)} = 0 + \lambda(4,463 - 0) = 4,463\lambda.$$

4. Определяется коэффициент λ. Этот коэффициент находится таким образом, чтобы переход к новому решению обеспечивал максимальное значение целевой функции. С этой целью уравнения для перехода к новому решению, построенные на шаге 3, подставляются в целевую функцию E. В результате целевая функция представляется как функция от коэффициента λ:

$$E = 5(6,923 - 2,0\lambda) - 0,2(6,923 - 2,06\lambda)^2 + 2\cdot 4,463\lambda - 0,2(4,463\lambda)^2 = -4,8\lambda^2 + 4,3\lambda + 25.$$

Значение λ находится из условия экстремума целевой функции, т.е. из условия

$$dE/d\lambda = 0 : dE/d\lambda = -9,6\lambda + 4,3 = 0.$$
$$\lambda = 0,448.$$

Примечания:

1. Если значение λ, найденное из уравнения $dE/d\lambda = 0$, оказывается больше единицы, то принимается λ = 1. Это означает, что экстремум целевой функции в направлении, задаваемом градиентом, находится за пределами ОДР. В этом случае новое решение должно находиться на границе ОДР (т.е. в точке x^*), но не выходить за нее.

2. Если уравнение $dE/d\lambda = 0$ не имеет решений (например, не зависит от λ), то также принимается λ = 1.

3. Из уравнений, составленных на шаге 3, определяется новое решение:

$$x_1^{(1)} = 6,923 - 2,06\cdot 0,448 = 6;$$
$$x_2^{(1)} = 4,463\cdot 0,448 = 2.$$

Определяется значение целевой функции для полученного решения:

$$E^{(1)} = 5\cdot 6 - 0,2\cdot 6^2 + 2\cdot 2 - 0,2\cdot 2^2 = 26.$$

4. Проверяется условие окончания поиска решения. Для этого определяется разность значений целевой функции для нового и предыдущего решения:

$$\Delta E = \left|E^{(1)} - E^{(0)}\right| = |26 - 25,03| = 0,97.$$

Эта величина сравнивается с заданной точностью ε. Если ΔE ≤ ε, то текущее решение принимается в качестве оптимального. В данном случае ΔE = 0,97, ε = 0,5. Таким образом, условие окончания поиска решения не выполняется. Требуется следующая итерация.

Итерация 2

1. Определяется градиент целевой функции в точке ОДР, соответствующей текущему решению:

$$gradE(x^{(1)}) = (5 - 0,4\cdot 6; 2 - 0,4\cdot 2) = (2,6;1,2).$$

2. Определяется угловая точка ОДР, соответствующая предельно допустимому перемещению от текущего решения в направлении градиента. Для этого решается следующая задача линейного программирования:

$$13x_1 + 6x_2 \le 90$$
$$8x_1 + 11x_2 \le 88$$
$$x_1, x_2 \ge 0$$
$$W = 2,6x_1 + 1,2x_2 \to \max.$$

Решение этой задачи следующее: $x_1^* = 4,863, x_2^* = 4,463.$

3. Составляются уравнения для перехода к новому решению:

$$x_1^{(2)} = 6 + \lambda(4,863 - 6) = 6 - 1,137\lambda;$$
$$x_2^{(2)} = 2 + \lambda(4,463 - 2) = 2 + 2,463\lambda.$$

4. Определяется коэффициент λ из условия экстремума целевой функции:

$E = 5(6-1,137\lambda) - 0,2(6-1,137\lambda)^2 + 2(2+2,463\lambda) - 0,2(2+2,463\lambda)^2 = -1,5\lambda^2 + 26.$

$dE/d\lambda = -3\lambda = 0.$

$\lambda = 0.$

5. Из уравнений, составленных на шаге 3, определяется новое решение:

$$x_1^{(2)} = 6 - 1,137 \cdot 0 = 6;$$

$$x_2^{(2)} = 2 - 2,463 \cdot 0 = 2.$$

Определяется значение целевой функции для полученного решения: $E^{(2)} = 5 \cdot 6 - 0,2 \cdot 6^2 + 2 \cdot 2 - 0,2 \cdot 2^2 = 26.$

6. Проверяется условие окончания поиска решения. Определяется разность значений целевой функции для нового и предыдущего решения: $\Delta E = \left| E^{(2)} - E^{(2)} \right| = 0$.

Так как $\Delta E \le \varepsilon$, оптимальное решение найдено: $x_1 = 6, x_2 = 2$. Оптимальное значение целевой функции E = 26.

Таким образом, предприятию следует выпустить 6 изделий типа А и 2 изделия типа В. Такой план производства обеспечит предприятию максимальную прибыль в размере 26 тыс. ден. ед.

Примечания:

1. В данном случае решение оказалось целочисленным, как и требуется по смыслу задачи. Если бы оно оказалось дробным, то для поиска оптимального целочисленного решения потребовалось бы применять специальные методы нелинейного целочисленного программирования. Эти методы достаточно сложны и не рассматриваются в данном пособии.

2. В рассмотренном примере решалась задача с целевой функцией, подлежащей максимизации. Если требуется минимизация целевой функции, то задача решается точно так же. Единственное отличие состоит в том, что целевая функция W в задаче, решаемой на шаге 2, также подлежит минимизации.

6.4. Решение задач нелинейного программирования средствами Excel

Решение задач нелинейного программирования в Excel в основном аналогично решению задач линейного программирования (см. подраздел 2.5). Рассмотрим решение примера 6.3 средствами Excel.

Предположим, что желательно получить результаты (значения переменных x_1 и x_2) в ячейках B2 и C2. В ячейке B3 введем формулу целевой функции: = 5·B2 - 0,2·B2^2 + 2·C2 - 0,2·C2^2

В ячейке B4 введем формулу первого ограничения (на расход платины): = 13·B2 + 6·C2

В ячейке D4 введем правую часть этого ограничения: 90.

Аналогично в ячейке B5 введем формулу ограничения на расход палладия (= 8·B2 + 11·C2), в ячейке D5 — правую часть этого ограничения (88).

Укажем также некоторые поясняющие надписи и обозначения (хотя это и необязательно). Рабочий лист будет иметь примерно такой вид, как показано на рис. 6.1.

Примечание. Значения 0 в ячейках B3 -B5 получены автоматически для начальных значений переменных, равных нулю.

Для решения задачи из меню «Сервис» выберем элемент «Поиск решения». В поле «Установить целевую ячейку» указывается ячейка B3, где находится формула целевой функции. Используя переключатели, необходимо указать, что требуется установить ячейку B3 «равной максимальному значению» (так как целевая функция в этой задаче подлежит максимизации). В поле «Изменяя ячейки» указываются ячейки, в которых должны находиться значения переменных: B2:C2.

В области «Ограничения» указываются ограничения, как показано в подразделе 2.5.

Необходимо ввести ограничения на расход платины и палладия, заданные на рабочем листе, а также ограничение на неотрицательность всех переменных (B2:C2>=0) и ограничение на их целочисленность (в поле «Ссылка на ячейку» указать B2:C2, а в поле знака ограничения выбрать отметку «цел»).

Для решения задачи следует нажать кнопку «Выполнить». Рабочий лист с результатами решения показан на рис. 6.2.

В ячейках B2 и C2 получены оптимальные значения переменных, в ячейке B3 — оптимальное значение целевой функции. Эти величины совпадают с результатами, полученными по методу Франка—Вульфа. В ячейках B4 и B5 находятся значения левых частей ограничений. Видно, что на выпуск изделий будет израсходовано 90 г платины и 70 г палладия.

Примечания:

1. В табличном процессоре Excel для решения задач оптимизации (элемент меню «Поиск решения») используются именно градиентные методы.

2. В некоторых случаях табличный процессор Excel не находит решения задачи при нулевых начальных значениях переменных. Обычно это происходит в случаях, когда в целевой функции или в каком-либо ограничении используется операция деления. При нулевых значениях переменных происходит деление на ноль и выводится сообщение об ошибке. В таких случаях в ячейках, в которых определяются значения переменных, перед началом решения задачи следует указать произвольные начальные значения (например, единицы).

7. Динамическое программирование

7.1. Постановка задачи. Принцип работы метода динамического программирования

Метод динамического программирования предназначен для задач, решение которых может быть представлено как некоторая многошаговая операция, т.е. последовательность однотипных шагов. Решение на каждом шаге принимается с учетом результатов предыдущих шагов, а также с учетом последствий принимаемого решения для последующих шагов.

К числу задач, для которых может применяться метод динамического программирования, относится большинство задач планирования на несколько периодов времени (например, на несколько лет). Шагом в таких задачах является один плановый период (например, один год). Метод динамического программирования применяется также для многих задач, в которых имеется возможность искусственно представить процесс принятия решения как последовательность из нескольких однотипных шагов.

Общая постановка задачи, решаемой методом динамического программирования, следующая. Имеется некоторая операция, находящаяся в начальном состоянии S_0. Операция реализуется за N шагов. На каждом шаге принимается некоторое решение U_k, где k — номер шага (k = 1,..., N). Выбор каждого решения Uk вызывает переход операции из состояния S_{k-1} в новое состояние S_k, а также обеспечивает некоторое значение критерия эффективности Z_k. Требуется определить последовательность решений $U_1, U_2, ..., U_k$, обеспечивающих экстремальное (максимальное или минимальное) значение общего критерия эффективности E, зависящего от значений критерия эффективности на отдельных шагах $Z_1, Z_2, ..., Z_k$.

Примечание. В литературе по динамическому программированию вместо термина «решение» обычно используется термин «управление».

Основной принцип решения задач на основе метода динамического программирования (принцип оптимальности, или принцип Беллмана) состоит в следующем: решение на каждом шаге выбирается таким образом, чтобы обеспечить максимальную эффективность на данном шаге и на всех последующих шагах.

Задача, представленная в виде многошаговой операции, может быть решена методом динамического программирования, если она удовлетворяет следующим свойствам:

- отсутствие последействия: состояние операции по окончании каждого шага (S_k) и критерий эффективности на каждом шаге (Z_k) зависят только от решения, принятого на данном шаге (U_k), и от состояния операции в начале данного шага (S_{k-1}), и не зависят от того, каким образом операция перешла в состояние S_{k-1};
- аддитивность или мультипликативность критерия эффективности: общий критерий эффективности представляет собой сумму критериев эффективности на отдельных шагах ($E = Z_1 + Z_2 + ... + Z_N$) или их произведение ($E = Z_1 \cdot Z_2 \cdot ... \cdot Z_N$).

Решение задач динамического программирования обычно включает два цикла: сначала - от последнего шага к первому (обратная прогонка, или условная оптимизация), затем - от первого шага к последнему (прямая прогонка, или безусловная оптимизация).

В цикле условной оптимизации для каждого шага находится множество возможных состояний операции в начале данного шага. Для каждого из этих состояний находится условно оптимальное решение, т.е. решение, оптимальное для данного состояния. Поиск условно оптимальных решений начинается с последнего (N-го) шага, так как на этом шаге имеется возможность выбирать решение только с учетом эффективности на данном шаге (последующих шагов нет). Затем на других шагах (N—1-м, N—2-м,..., первом) условно оптимальные решения выбираются согласно принципу оптимальности, т.е. с учетом эффективности на данном шаге и на последующих шагах. На всех шагах от N-го до второго определяется несколько условно оптимальных решений — по одному для каждого возможного состояния. Для первого шага начальное состояние (S_0) обычно известно точно, поэтому для этого шага находится только одно (безусловно оптимальное) решение U_1^*.

В цикле безусловной оптимизации для каждого шага определяется безусловно оптимальное решение. Поиск безусловно оптимальных решений начинается с первого шага, так как для него известно начальное состояние 50, поэтому можно определить единственное (безусловно оптимальное) решение U_1^*. Определяется состояние 51, в которое переходит операция из состояния S_0 в результате решения U_1^* , т.е. состояние в начале второго шага. Для него в цикле условной оптимизации уже найдено оптимальное решение U_1^*. Определяется состояние операции в начале третьего шага - состояние S_2, в которое операция переходит в результате решения U_2^*.

Для этого состояния в цикле условной оптимизации также найдено оптимальное решение U_3^*. Аналогично определяются безусловно оптимальные решения для последующих шагов.

Важно отметить, что для метода динамического программирования не существует вычислительной процедуры, одинаковой для всех задач (в отличие, например, от симплекс-метода). Это означает, что правила вычислений, составления таблиц и т.д. полностью зависят от конкретной задачи. Общими являются лишь основные принципы решения: принцип оптимальности, решение задачи с использованием условной и безусловной оптимизации.

7.2. Примеры решения задач на основе метода динамического программирования

При рассмотрении примеров будем использовать следующие обозначения: S_k — состояние в конце k-го шага (или, другими словами, состояние в начале k+1-го шага); U_k — любое возможное (допустимое) решение на k-м шаге; U_k^* - оптимальное решение на k-м шаге; Z_k - критерий эффективности на k-м шаге; E_k - суммарный критерий эффективности на всех шагах, начиная с k-го (т.е. на шагах от k-го до N-го); E_k^* - оптимальный (в рассматриваемых примерах — максимальный) суммарный критерий эффективности на всех шагах, начиная с k-го.

Пример 7.1. Денежные средства в размере 60 млн ден. ед. распределяются между четырьмя предприятиями (П1, П2, П3, П4), принадлежащими одной крупной фирме. Денежные средства выделяются в размерах, кратных 20 млн ден. ед. Каждым предприятием разработаны планы использования денежных средств на развитие производства. Определена прибыль, которую получит каждое предприятие в результате использования выделенных средств (табл. 7.1).

Таблица 7.1

Выделенные средства, млн ден. ед.	Прибыль предприятий, млн ден. ед.			
	П1	П2	П3	П4
0	0	0	0	0
20	9	10	6	12
40	16	18	12	17
60	22	20	25	20

Например, если предприятию Ш не будут выделены средства, то оно не получит никакой прибыли. Если этому предприятию будет выделено 20 млн ден. ед., то его прибыль от использования этих средств составит 9 млн ден. ед. Если будет выделено 40 млн ден. ед., то прибыль составит 16 млн ден. ед., а при выделении 60 млн -22 млн ден. ед.
Требуется распределить имеющиеся средства (60 млн ден. ед.) между предприятиями таким образом, чтобы общая прибыль фирмы была максимальной.

В данной задаче в качестве шагов будем рассматривать выделение средств предприятиям: первый шаг — выделение средств предприятию П1, второй — П2, и т.д. (всего 4 шага). Таким образом, распределение средств между предприятиями можно рассматривать как многошаговую операцию. В качестве состояния этой операции будем использовать величину имеющихся средств, которые требуется распределить. Начальное состояние $S_0 = 60$. Решение на каждом шаге — это денежные средства, выделяемые предприятию ($U_k, k = 1,...,4$). Критерий эффективности для каждого шага — прибыль, полученная предприятием ($Z_k, k = 1,...,4$). Общий критерий эффективности — это прибыль фирмы, т.е. сумма прибылей всех предприятий: $E = Z_1 + Z_2 + Z_3 + Z_4$. Задача удовлетворяет свойству отсутствия последействия. Состояние по окончании каждого шага (т.е. имеющаяся сумма средств, подлежащая распределению, S_k) зависит только от суммы, имевшейся в начале шага (S_{k-1}) и от решения, принятого на данном шаге (т.е. от выделенной на данном шаге денежной суммы U_k): $S_k = S_{k-1} - U_k$. Критерий эффективности на каждом шаге (т.е. прибыль предприятия Z_k) зависит только от решения на данном шаге, т.е. от выделенной предприятию суммы U_k, и не зависит от того, сколько средств выделено другим предприятиям.

Задача удовлетворяет также свойству аддитивности критерия эффективности: общий критерий эффективности (прибыль фирмы) равен сумме критериев эффективности на отдельных шагах (прибылей предприятий). Задача решается в два цикла.

Цикл условной оптимизации

Шаг 4 (выделение средств предприятию П4)

Определяются все возможные состояния S_3 к началу шага 4 (или к концу шага 3), т.е. все возможные значения остатка денежных средств после их выделения предприятиям П1, П2 и П3. Этот остаток может составлять 0 ден. ед. (если все средства выделяются предприятиям П1, П2 и П3), 20 млн ден. ед. (если предприятиям П1, П2, П3 выделяется 40 млн ден. ед.), 40 млн ден. ед. (если предприятиям П1, П2, П3 выделяется 20 млн ден. ед.) или 60 млн ден. ед. (если предприятиям П1, П2, П3 средства вообще не выделяются). Для каждого из возможных состояний определяется условно оптимальное решение, т.е. решение, оптимальное при условии, что остаток денежных средств равен S_3. Так как предприятие П4 — последнее (предполагается, что другим предприятиям средства уже выделены), оптимальное решение состоит в выделении предприятию П4 всех оставшихся средств.

Возможные состояния в начале четвертого шага S_3, соответствующие им условно оптимальные решения U_4^* и значения критерия эффективности (прибыль предприятия П4) E_4^* приведены в табл. 7.2.

Важно обратить внимание, что по результатам четвертого шага не выяснено, сколько средств требуется выделять предприятию П4. Это пока невозможно, так как неизвестно начальное состояние четвертого шага S_3. Поэтому найдены условно оптимальные решения для всех возможных состояний.

Шаг 3 (выделение средств предприятиям П3 и П4)

Все расчеты для шага 3 приведены в табл. 7.3.

В таблице использованы следующие обозначения: S_2 - возможные суммы денежных средств, распределяемые между предприятиями П3 и П4 (т.е. оставшиеся после выделения средств предприятиям П1 и П2); U_3 - возможные варианты выделения средств предприятию П3; Z_3- прибыль предприятия П3 от выделения средств в размере U_3; S_3 - остаток денежных средств после их выделения предприятиям П1, П2 и П3 (т.е. средства, выделяемые предприятию П4); E_4^*- прибыль предприятия П4 от выделенных ему средств в размере S_3; E_3 -суммарная прибыль предприятий П3 и П4 (сумма величин из столбцов Z_3 и E_4^*); U_3^* - условно оптимальное решение для состояния S_2 (денежные средства, которые следует выделить предприятию П3 при наличии суммы S_2); E_3^* - условно оптимальный критерий эффективности для предприятий П3 и П4, т.е. прибыль, получаемая этими предприятиями в результате решения U_3^*. Рассмотрим порядок решения для шага 3.

Определяются все возможные состояния S_2 к началу шага 3 (или к концу шага 2), т.е. все возможные значения денежных средств, распределяемых между предприятиями П3 и П4. Этот остаток может составлять 0 ден. ед. (если все средства выделяются предприятиям Ш и П2), 20 млн ден. ед. (если предприятиям Ш и П2 выделено 40 млн ден. ед.), 40 млн ден. ед. (если предприятиям П1 и П2 выделено 20 млн ден. ед.) или 60 млн ден. ед. (если предприятиям П1 и П2 средства вообще не выделяются). Для каждого из возможных состояний определяется условно оптимальное решение, т.е. решение, оптимальное при условии, что остаток денежных средств равен S_2. Средства для предприятия П3 должны выделяться таким образом, чтобы обеспечить максимальную суммарную прибыль для П3 и П4.

Предположим, что денежные средства, распределяемые между предприятиями П3 и П4, составляют 20 млн ден. ед. ($S_2 = 20$). Эти средства можно оставить для предприятия П4 (тогда предприятию П3 средства не выделяются, $U_3 = 0$) или выделить их предприятию П3 ($U_3 = 20$). Если $U_3 = 0$, то предприятие П3 не получит прибыли ($Z_3 = 0$). В этом случае остаток средств (состояние) в конце третьего шага составит $S_3 = 20$ млн ден. ед. Эти средства будут выделены предприятию П4, и его прибыль составит $E_4 = 12$ млн ден. ед. Суммарная прибыль предприятий П3 и П4 составит $E_3 = 0 + 12 = 12$ млн ден. ед. Если

$U_3 = 20$, то предприятие П3 получит прибыль $Z_3 = 6$ млн ден. ед. Остаток средств (состояние) в конце третьего шага составит $S_3 = 0$. Предприятию П4 не будет выделено никаких средств, и оно не получит прибыли ($E_4 = 0$). Суммарная прибыль предприятий П3 и П4 составит $E_3 = 6 + 0 = 6$ млн ден. ед. Таким образом, если между предприятиями П3 и П4 распределяется сумма в размере 20 млн ден. ед., то эти средства не следует выделять предприятию П3; их следует выделить предприятию П4, так как общая прибыль в этом случае будет максимальной. Другими словами, для состояния $S_2 = 20$ условно оптимальное решение $U_3^* = 0$, условно оптимальный критерий эффективности $E_3^* = 12$.

Предположим, что денежные средства, распределяемые между предприятиями П3 и П4, составляют 40 млн ден. ед. ($S_2 = 40$). Предприятию П3 можно выделить 0, 20 или 40 млн ден. ед. ($U_3 = 0$, 20 или 40). Если $U_3 = 0$, то предприятие П3 не получит прибыли ($Z_3 = 0$). В этом случае остаток средств (состояние) в конце третьего шага составит $S_3 = 40$ млн ден. ед. Эти средства будут выделены предприятию П4, и его прибыль составит $E_4^* = 17$ млн ден. ед. Суммарная прибыль предприятий П3 и П4 составит $E_3 = 0 + 17 = 17$ млн ден. ед. Аналогично можно определить, что при выделении предприятию П3 суммы в размере 20 млн ден. ед. (т.е. при $U_3 = 20$) суммарная прибыль предприятий П3 и П4 составит $E_3 = 6 + 12 = 18$ млн ден. ед. При $U_3 = 40$ суммарная прибыль предприятий П3 и П4 составит $E_3 = 12 + 0 = 12$ млн ден. ед. Таким образом, для состояния $S_2 = 40$ условно оптимальное решение $U_3^* = 20$, условно оптимальный критерий эффективности $E_3^* = 18$ млн ден. ед. Это означает, что при распределении между предприятиями П3 и П4 средств в размере 40 млн ден. ед. предприятию П3 следует выделить 20 млн ден. ед.

Аналогично можно определить, что при распределении между предприятиями П3 и П4 средств в размере 60 млн ден. ед. предприятию П3 следует выделить все 60 млн ден. ед. (для состояния $S_2 = 60$ условно оптимальное решение $U_3^* = 60$, условно оптимальный критерий эффективности $E_3^* = 25$ млн ден. ед.).

Шаг 2 (выделение средств предприятиям П2, П3 и П4)

Обозначения в таблице: S_1 - возможные суммы денежных средств, распределяемые между предприятиями П2, П3 и П4 (т.е. оставшиеся после выделения средств предприятию П1); U_2 - возможные варианты выделения средств предприятию П2; Z_2 - прибыль предприятия П2 от выделения средств в размере U_2; S_2 - остаток денежных средств после их выделения предприятиям П1 и П2 (т.е. средства, выделяемые предприятиям П3 и П4); E_3^* - максимальная суммарная прибыль предприятий П3 и П4 от выделенных им средств в размере S_2 (определяется из табл. 7.3); E_2 - суммарная прибыль предприятий П2, П3 и П4 (сумма величин из столбцов Z_2 и E_3^*); U_2^* – условно оптимальное решение для состояния S_1 (денежные средства, которые следует выделить предприятию П2 при наличии суммы S_1); E_2^* - условно оптимальный критерий эффективности для предприятий П2, П3 и П4, т.е. прибыль, получаемая этими предприятиями в результате решения U_2^*.Рассмотрим порядок решения для шага 2.

Определяются все возможные состояния S_1 к началу шага 2, т.е. все возможные значения денежных средств, распределяемых между предприятиями П2, П3 и П4. Этот остаток может составлять 0; 20; 40 или 60 млн ден. ед. (в зависимости от того, сколько средств выделяется предприятию П1).

Для каждого из возможных состояний S_1 определяется условно оптимальное решение, т.е. оптимальная денежная сумма, выделяемая предприятию П2 при условии, что имеются денежные средства в размере S_1. Средства для предприятия П2 должны выделяться таким образом, чтобы обеспечить максимальную суммарную прибыль предприятий П2, П3 и П4.

Предположим, что денежные средства, распределяемые между предприятиями П2, П3 и П4, составляют 20 млн ден. ед. ($S_1 = 20$). Предприятию П2 можно выделить 0 или 20 млн ден. ед. ($U_2 = 0$ или $U_2 = 20$). Если $U_2 = 0$, то предприятие П2 не получит прибыли ($Z_2 = 0$). В этом случае остаток средств
(состояние) в конце второго шага составит $S_2 = 20$ млн ден. ед. Эти средства будут распределены между предприятиями П3 и П4. Из табл. 7.3 видно, что при оптимальном распределении таких средств между предприятиями П3 и П4 максимальная суммарная прибыль этих предприятий составит $E_3^* = 12$ млн ден. ед. Суммарная прибыль предприятий П2, П3 и П4 составит $E_2 = 0 + 12 = 12$ млн ден. ед. Если $U_2 = 20$, то предприятие П2 получит прибыль $Z_2 = 10$ млн ден. ед.

Остаток средств (состояние) в конце третьего шага составит $S_3 = 0$. Предприятиям П3 и П4 не будет выделено никаких средств, и они не получат прибыли ($E_3^* = 0$), а суммарная прибыль предприятий П2, П3 и П4 составит $E_2 = 10 + 0 = 10$ млн ден. ед. Таким образом, для состояния $S_1 = 20$ условно оптимальное решение $U_2^* = 0$, условно оптимальный критерий эффективности $E_2^* = 12$ млн ден. ед. Это означает, что при распределении средств в размере 20 млн ден. ед. между предприятиями П2, П3 и П4, предприятию П2 не следует выделять средства; все имеющиеся средства следует распределить между предприятиями П3 иП4.

Предположим, что денежные средства, распределяемые между предприятиями П2, П3 и П4, составляют 40 млн ден. ед. ($S_1 = 40$). Предприятию П2 можно выделить 0, 20 или 40 млн ден. ед. ($U_2 = 0$, 20 или 40). Если U2 = 0, то предприятие П2 не получит прибыли ($Z_2 = 0$). Остаток средств в конце второго шага составит $S_2 = 40$ млн ден. ед. Эти средства будут распределены между предприятиями П3 и П4. Из табл. 7.3 видно, что максимальная прибыль этих предприятий от использования таких средств составит $E_3^* = 18$ млн ден. ед. Суммарная прибыль предприятий П2, П3 и П4 составит $E_2 = 0 + 18 = 18$ млн ден. ед. Аналогично можно определить, что при $U_2 = 20$ суммарная прибыль предприятий П2, П3 и П4 составит $E_2 = 10 + 12 = 22$ млн ден. ед. При $U_2 = 40$ суммарная прибыль предприятий П2, П3 и П4 составит $E_2 = 18 + 0 = 18$ млн ден. ед. Таким образом, для состояния $S_1 = 40$ условно оптимальное решение $U_2 = 20$, условно оптимальный критерий эффективности $E_2^* = 22$ млн ден. ед.

Это означает, что при распределении средств в размере 40 млн ден. ед. между предприятиями П2, П3 и П4, предприятию П2 следует выделить 20 млн ден. ед. Аналогично можно определить, что при распределении между предприятиями П2, П3 и П4 средств в размере 60 млн ден. ед. предприятию П2 следует выделить 40 млн ден. ед.

Шаг 1 (выделение средств предприятиям П1, П2, П3 и П4)

Все расчеты для шага 1 приведены в табл. 7.5. Обозначения в таблице: S_0 - начальная сумма денежных средств, распределяемых между всеми предприятиями; U_1 - возможные варианты выделения средств предприятию П1;

Z_1 - прибыль предприятия П1 от выделения средств в размере U_1; S_1 — остаток денежных средств после их выделения предприятию П1 (т.е. средства, выделяемые предприятиям П2, П3 и П4); E_2^* - максимальная суммарная прибыль предприятий П2, П3 и

П4 от выделенных им средств в размере S_1 (определяется из табл. 7.4); E_1 - суммарная прибыль предприятий П1, П2, П3 и П4, т.е всех предприятий (сумма величин из столбцов Z_1 и E_2^*); U_1^* - безусловно оптимальное решение для состояния S_0 (денежные средства, которые следует выделить предприятию П1 при наличии суммы S_0); E_1^* - безусловно оптимальный критерий эффективности для предприятий П1, П2, П3 и П4, т.е. прибыль, получаемая всеми предприятиями в результате решения U_1^* .

Начальная сумма денежных средств (состояние S_0) известна: $S_0 = 60$. Требуется определить, сколько средств необходимо выделить предприятию П1, чтобы обеспечить максимальную суммарную прибыль предприятий П1, П2, П3 и П4, т.е. всех предприятий. Так как начальное состояние на этом шаге известно точно (в отличие от других шагов), будет найдено безусловно оптимальное решение.

Таблица 7.5

S0	U1	Z1	S1	E*	E1		*
60	0 20 40 60	0 9 16 22	60 40 20 0	30 22 12 0	30 31 28 22	20	31

Предприятию П1 можно выделить 0, 20, 40 или 60 млн ден. ед. ($U_1 = 0$, 20, 40 или 60). В зависимости от выделенных средств прибыль предприятия П1 (Z_1) может составлять 0, 9, 16 или 22 млн ден. ед. Остаток средств в конце первого шага S_1 (сумма, выделяемая предприятиям П2, П3 и П4) может составлять 60, 40, 20 или 0 млн ден. ед. Из табл. 7.4 определяется максимальная прибыль предприятий П2, П3 и П4 (E_2^*) от использования средств в размере S1: она может составлять 30, 22, 12 или 0 млн ден. ед. Для всех случаев определяется суммарная прибыль предприятий П1, П2, П3 и П4 (E1): она может составлять 30, 31, 28 или 22 млн ден. ед. Таким образом, максимальная прибыль предприятий П1, П2, П3 и П4 (т.е. всех предприятий) достигается, если выделить предприятию П1 20 млн ден. ед. (при условии, что для остальных предприятий средства также будут распределяться оптимальным образом). Это означает, что оптимальным решением является выделение предприятию П1 средств в размере 20 млн ден. ед.: $U_1^* = 20$. Прибыль всех предприятий в этом случае составит 31 млн ден. ед.

Цикл безусловной оптимизации

Для первого шага (выделение средств предприятию П1) получено безусловно оптимальное решение: $U_1^* = 20$ млн ден. ед. Для предприятий П2, П3 и П4 остается 40 млн ден. ед. Таким образом, состояние в начале второго шага $S_1 = 40$. Из табл. 7.4 для этого состояния определяется оптимальное решение: $U_2^* = 20$ (предприятию П2 выделяется 20 млн ден. ед.). Для предприятий П3 и П4 остается 20 млн ден. ед. (состояние в начале третьего шага $S_2 = 20$). Из табл. 7.3 для этого состояния определяется оптимальное решение: $U_3^* = 0$ (предприятию П3 средства не выделяются). Для предприятия П4 остается 20 млн ден. ед. ($S_3 = 20$). Поэтому $U_4^* = 20$.

Таким образом, оптимальное решение задачи следующее. Предприятию Ш следует выделить 20 млн ден. ед., предприятию П2 — также 20 млн ден. ед., предприятию П3 — не выделять средства, предприятию П4 — выделить 20 млн ден. ед. Общая прибыль составит 31 млн ден. ед., в том числе прибыль предприятия П1 — 9 млн ден. ед., П2 — 10 млн ден. ед., П3 — 0, П4 — 12 млн ден. ед.

Пример 7.2. Фирма владеет двумя предприятиями (П1 и П2). В связи с тем, что спрос на продукцию этих предприятий имеет сезонный характер, прибыль от вложения средств в производство продукции на этих предприятиях различна в разные периоды года. Прибыль (в процентах) для различных периодов года приведена в табл. 7.6.

В конце каждого квартала выручка каждого предприятия распределяется следующим образом: 20% выплачивается акционерам фирмы, 80% - перераспределяется между предприятиями.

В начале года для вложения в производство выделена сумма в размере 5 млн ден. ед. Требуется составить план распределения средств в течение года таким образом, чтобы сумма, выплачиваемая акционерам в течение года, была максимальной.

Таблица 7.6

Предприятие	Прибыль, %			
	Январь-март	Апрель-июнь	Июль-сентябрь	Октябрь-декабрь
Ш	80	50	50	70
П2	40	120	80	40

Величины в таблице обозначают следующее: если, например, предприятию П в начале января будет выделен 1 млн ден. ед., то прибыль предприятия к концу марта составит 800 тыс. ден. ед. (80% от выделенной суммы). Таким образом, выручка предприятия за квартал составит 1 млн 800 тыс. ден. ед.

В данной задаче в качестве шагов будем рассматривать выделение средств предприятиям в начале каждого квартала: первый шаг - первый квартал, и т.д.

В качестве состояния операции будем использовать величину имеющихся средств, которые требуется распределить. Начальное состояние $S_0 = 5$. Состояние в начале k-го шага будем обозначать как S_{k-1}. Решение на каждом шаге - это денежные средства, выделяемые каждому из предприятий. Будем обозначать средства, выделяемые предприятию П1, как U_k, а средства, выделяемые предприятию П2 — как $S_{k-1} - U_k, k = 1,...,4$. Критерий эффективности для каждого шага — сумма выплат акционерам ($Z_k, k = 1,...,4$). Общий критерий эффективности — это выплаты акционерам в течение года: $E = Z_1 + Z_2 + Z_3 + Z_4$.

Цикл условной оптимизации

Шаг 4 (распределение средств в четвертом квартале)

Пусть в конце третьего квартала между предприятиями распределяется сумма в размере S_3. Пусть предприятию Ш выделяется сумма в размере U_4, а предприятию П2$-S_3 - U_4$. Тогда выручка предприятий к концу четвертого квартала составит $1,7 \cdot U_4 + 1,4 \cdot (S_3 - U_4)$. Из этой суммы акционерам будет выплачено 20%. Таким образом, выплаты акционерам в конце четвертого квартала определяются следующим образом:

$E_4 = Z_4 = 0,2 \cdot (1,7 \cdot U_4 + 1,4 \cdot (S_3 - U_4)) = 0,06 \cdot U_4 + 0,28 \cdot S_3.$

Видно, что максимальные выплаты акционерам обеспечиваются при максимальном значении U_4. Это означает, что предприятию Ш в четвертом квартале следует выделить всю имеющуюся сумму: $U_4^* = S_3$. Здесь U_4^*–условно оптимальное решение для четвертого шага. Условно оптимальное значение критерия эффективности (выплаты акционерам) на четвертом шаге определяется следующим образом: $E_4^* = 0,06 \cdot S_3 + 0,28 \cdot S_3 = 0,34 \cdot S_3.$

Шаг 3 (распределение средств в третьем и четвертом кварталах)

Пусть в конце второго квартала между предприятиями распределяется сумма в размере S_2. Пусть предприятию П1 выделяется сумма в размере U_3, а предприятию П2 $-S_2-U_3$. Тогда выручка предприятий к концу третьего квартала составит $1{,}5\cdot U_3+1{,}8\cdot(S_2-U_3)$. Из этой суммы 20% будет выплачено акционерам, а 80% - распределено между предприятиями. Выплаты акционерам за третий и четвертый кварталы определяются следующим образом:

$E_3=Z_3+E_4^*=0{,}2\cdot(1{,}5\cdot U_3+1{,}8\cdot(S_2-U_3))+E_4^*.$

Выразим величину E_3 через U_3 и S_2. Как показано выше, $E_4^*=0{,}34\cdot S_3$ где S_3 - сумма средств, распределяемых между предприятиями в начале четвертого квартала. Эта сумма составляет 80% от выручки предприятий в третьем квартале: $S_3=0{,}8\cdot(1{,}5\cdot U_3+1{,}8\cdot(S_2-U_3))$.

Таким образом, $E_4^*=0{,}34\cdot 0{,}8\cdot(1{,}5\cdot U_3+1{,}8\cdot(S_2-U_3))=0{,}49\cdot S_2-0{,}08\cdot U_3$.

Подставляя эту величину в уравнение для E_3, получим:

$$E_3=0{,}2\cdot(1{,}5\cdot U_3+1{,}8\cdot(S_2-U_3))+0{,}49\cdot S_2-0{,}08\cdot U_3=0{,}85\cdot S_2=0{,}14\cdot U_3.$$

Видно, что максимальные выплаты акционерам обеспечиваются при минимальном значении U3 (так как коэффициент при U3 отрицательный). Таким образом, условно оптимальное решение для третьего квартала состоит в том, чтобы не выделять средства предприятию П1: $U_3^*=0$. Подставляя $U_3=0$, получим выражение для условно оптимального значения критерия эффективности на третьем и четвертом шагах: $E_3^*=0{,}85\cdot S_2$.

Шаг 2 (распределение средств во втором - четвертом кварталах)

Пусть в конце первого квартала между предприятиями распределяется сумма в размере S_1. Пусть предприятию П1 выделяется сумма в размере U_2, а предприятию П2 - сумма в размере S_1-U_2. Тогда выручка предприятий к концу второго квартала составит $1{,}5\cdot U_2+2{,}2\cdot(S_1-U_2)$. Из этой суммы 20% будет выплачено акционерам, а 80% — распределено между предприятиями. Выплаты акционерам за второй — четвертый кварталы определяются следующим образом:

$$E_2=Z_2+E_3^*=0{,}2\cdot(1{,}5\cdot U_2+2{,}2\cdot(S_1-U_2))+E_3^*.$$

Выполнив расчеты, аналогичные приведенным на шаге 3, выразим величину E_2 через U_2 и S_1:

$$E_2=0{,}2\cdot(1{,}5\cdot U_2+2{,}2\cdot(S_1-U_2))+E_3^*=0{,}2\cdot(1{,}5\cdot U_2+2{,}2\cdot(S_1-U_2))+0{,}85\cdot S_2=$$
$$=0{,}2\cdot(1{,}5\cdot U_2+2{,}2\cdot(S_1-U_2))+0{,}85\cdot 0{,}8\cdot(1{,}5\cdot U_2+2{,}2\cdot(S_1-U_2))=1{,}94\cdot S_1-0{,}62\cdot U_2.$$

Видно, что максимальные выплаты акционерам обеспечиваются при минимальном значении U_2 (так как коэффициент при U_2 отрицательный). Таким образом, условно оптимальное решение для второго квартала состоит в том, чтобы не выделять средства предприятию П1: $U_2^*=0$. Подставляя $U_2=0$ в выражение для E_2, получим выражение для условно оптимального значения критерия эффективности на втором — четвертом шагах: $E_2^*=1{,}94\cdot S_1$.

Шаг 1 (распределение средств в первом — четвертом кварталах)

В начале первого квартала между предприятиями распределяется сумма в размере $S_0=5$. Пусть предприятию П1 выделяется сумма в размере U_1, а предприятию П2 — сумма в размере S_0-U_1. Выручка предприятий к концу первого квартала составит $1{,}8\cdot U_1+1{,}4\cdot(S_0-U_1)$. Как и в другие кварталы, 20% из этой суммы будет выплачено акционерам, а 80% — распределено между предприятиями. Выплаты акционерам за первый — четвертый кварталы определяются следующим образом:

$$E_1=Z_1+E_2^*=0{,}2\cdot(1{,}8\cdot U_1+1{,}4\cdot(S_0-U_1))+E_2^*.$$

Выполнив расчеты, аналогичные приведенным на предыдущих шагах, выразим величину E_1 через U_1 и S_0:

$$E_1 = 0,2\cdot(1,8\cdot U_1 + 1,4\cdot(S_0 - U_1)) + E_2^* = 0,2\cdot(1,8\cdot U_1 + 1,4\cdot(S_0 - U_1)) + 1,94\cdot S_1 =$$
$$= 0,2\cdot(1,8\cdot U_1 + 1,4\cdot(S_0 - U_1)) + 1,94\cdot 0,8\cdot(1,8\cdot U_1 + 1,4\cdot(S_0 - U_1)) = 2,45\cdot S_0 + 0,7\cdot U_1.$$

Видно, что максимальные выплаты акционерам за весь год обеспечиваются при максимальном значении U_1. Таким образом, безусловно оптимальное решение для первого квартала состоит в том, чтобы выделить предприятию П1 все имеющиеся средства: $U_1^* = S_0$. Для первого шага (в отличие от остальных шагов) начальное состояние известно: $S_0 = 5$ млн ден. ед. Итак, в первом квартале предприятию П1 следует выделить 5 млн ден. ед.

Цикл безусловной оптимизации

Шаг 1 (распределение средств в первом квартале)

Безусловно оптимальное решение для первого квартала найдено выше: $U_1^* = 5$ (все имеющиеся средства выделяются предприятию П1).

Выручка предприятия П1 к концу первого квартала составит $1,8 \cdot 5$ = 9 млн ден. ед. Найдем сумму выплат акционерам в первом квартале: $Z_1 = 0,2\cdot 9 = 1,8$ млн ден. ед. Найдем состояние в конце первого квартала, т.е. сумму средств, распределяемых между предприятиями в начале второго квартала:

$$S_1 = 0,8\cdot 9 = 7,2 \text{ млн ден. ед.}$$

Шаг 2 (распределение средств во втором квартале)

Состояние в начале второго шага $S_1 = 7,2$ млн ден. ед. В ходе условной оптимизации выяснено, что все эти средства следует выделить предприятию П2 ($U_2^* = 0$).

Выручка предприятия П2 к концу второго квартала составит $2,2 \cdot 7,2$ = = 15,84 млн ден. ед. Выплаты акционерам во втором квартале составят $Z_2 = 0,2\cdot 15,84 = 3,17$ млн ден. ед. Состояние в конце второго шага (т.е. сумма средств, распределяемых между предприятиями в начале третьего квартала) следующее: $S_2 = 0,8\cdot 15,84 = 12,67$ млн ден. ед.

Шаг 3 (распределение средств в третьем квартале)

Состояние в начале третьего шага $S_2 = 12,67$ млн ден. ед. В ходе условной оптимизации выяснено, что все эти средства следует выделить предприятию П2 ($U_3^* = 0$).

Выручка предприятия П2 к концу третьего квартала составит $1,8 \cdot 12,67$ = = 22,81 млн ден. ед. Выплаты акционерам в третьем квартале составят $Z_3 = 0,2\cdot 22,81 = 4,56$ млн ден. ед. Состояние в конце третьего шага (т.е. сумма средств, распределяемых между предприятиями в начале четвертого квартала) следующее: $S_3 = 0,8\cdot 22,81 = 18,25$ млн ден. ед.

Шаг 4 (распределение средств в четвертом квартале)

Состояние в начале четвертого шага $S_3 = 18,25$ млн ден. ед. В ходе условной оптимизации выяснено, что все эти средства следует выделить предприятию П1 ($U_3^* = S_3 = 18,25$).

Выручка предприятия П1 к концу четвертого квартала составит $1,7 \cdot 18,25$ = 31,03 млн ден. ед. Выплаты акционерам в четвертом квартале составят $Z_3 = 0,2\cdot 31,03 = 6,21$ млн ден. ед.

Таким образом, оптимальное решение задачи следующее. В первом квартале следует выделить предприятию П1 5 млн ден.ед, во втором квартале выделить предприятию П2 7,2 млн ден. ед., в третьем квартале выделить предприятию П2 12,67 млн ден. ед., в четвертом квартале выделить предприятию П1 18,25 млн ден. ед. Выплаты акционерам в первом квартале составят 1,8 млн ден. ед., во втором — 3,17, в третьем — 4,56, в четвертом — 6,21 млн ден. ед. Таким образом, выплаты акционерам в течение года составят 15,74 млн ден. ед.

8. Оптимизация на основе моделей массового обслуживания

8.1. Понятие системы массового обслуживания

Система массового обслуживания (СМО) - это любая система, предназначенная для обслуживания поступающих в нее заявок.

Примеры СМО - предприятие, выполняющее заказы; станок, обрабатывающий детали; компьютер, решающий задачи; магазин, обслуживающий покупателей, и т.д.

Заявки, поступающие на обслуживание в СМО (заказы, детали, задачи, покупатели и т.д.), образуют поток заявок. Элементы СМО, обслуживающие заявки, называются каналами обслуживания.

В большинстве случаев интервалы времени между моментами поступления заявок и/или времена обслуживания заявок в СМО представляют собой случайные величины. Другими словами, в большинстве случаев заранее точно неизвестно, когда поступит очередная заявка и сколько времени займет ее обслуживание. Поэтому теория систем массового обслуживания основана на математическом аппарате теории вероятностей и математической статистики.

8.2. Потоки заявок в СМО. Законы распределения интервалов времени между заявками и времени обслуживания

Поток Пальма. Для расчета характеристик СМО требуется формальное описание потока заявок, поступающих в нее. Как правило, достаточно точный расчет характеристик СМО возможен только в случаях, когда поток заявок представляет собой поток Пальма. Потоком Пальма называется поток событий (под событием в данном случае понимается поступление заявки), обладающий свойствами стационарности, ординарности, отсутствия последействия.

Стационарность. Поток событий является стационарным, если количество событий на любом интервале времени зависит только от длительности этого интервала, но не зависит от его расположения на оси времени. Например, поток деталей, поступающих по конвейеру на станок, можно считать стационарным. Поток пассажиров в метро в течение дня нельзя считать стационарным, так как количество пассажиров, обслуживаемых в течение некоторого интервала времени (например, 20 минут или одного часа) в «час пик», существенно отличается от количества пассажиров, обслуживаемых за такой же интервал времени в середине дня. В то же время, если рассматривать отдельно потоки пассажиров в «час пик» и в промежутке между «часами пик», то такие потоки можно считать стационарными.

Ординарность. Поток событий является ординарным, если вероятность появления нескольких (двух или более) событий за элементарный (т.е. очень короткий, близкий к нулевому) интервал времени очень мала по сравнению с вероятностью появления за этот же период одного события. Другими словами, поток заявок является ординарным, если заявки поступают на обслуживание не группами, а по одной.

Ограниченность последействия. Поток событий является потоком с ограниченным последействием, если интервалы времени между событиями представляют собой независимые случайные величины, распределенные по одному и тому же закону.

Пуассоновский поток. Наиболее точный расчет характеристик возможен для СМО, в которых поток заявок является пуассоновским (простейшим). Пуассоновским называется поток заявок, обладающий свойствами стационарности, ординарности и отсутствия последействия. Поток обладает свойством отсутствия последействия, если количество событий на любом интервале времени не зависит от количества событий на любом другом интервале времени. Другими словами, поток заявок обладает свойством отсутствия последействия, если моменты поступления заявок никак не зависят от моментов поступления предыдущих заявок. Заявки поступают на обслуживание независимо друг от друга.

Поясним разницу между ограниченностью последействия и отсутствием последействия на следующем примере. Пусть известно, что интервалы времени между моментами поступления деталей на станок могут составлять от 5 до 10 минут. Это значит, что время между моментами поступления деталей представляет собой случайную величину, распределенную по равномерному закону в диапазоне от 5 до 10. Поток деталей в этом случае обладает свойством ограниченности последействия, так как интервалы времени между моментами поступления деталей распределены по одному и тому же закону (равномерному) и не зависят друг от друга: интервал между любыми двумя деталями может представлять собой любую величину в пределах от 5 до 10 минут, независимо от того, какими были интервалы между предыдущими деталями. В то же время такой поток не обладает свойством отсутствия последействия, так как имеется зависимость между моментами поступления деталей. Если в некоторый момент на обработку поступила деталь, то следующая деталь поступит не раньше, чем через 5 минут и не позже, чем через 10 минут. Таким образом, поток деталей в этом случае является потоком Пальма, но не является пуассоновским.

В пуассоновском потоке интервалы времени между моментами поступления заявок распределены по экспоненциальному (показательному) закону. Упрощенно говоря, это означает, что интервалы между заявками могут быть как очень короткими, так и очень длительными.

Количество заявок в пуассоновском потоке, поступающих на обслуживание за некоторый интервал времени, представляет собой пуассоновскую случайную величину. Это означает, что вероятность поступления ровно k заявок за некоторый интервал времени t определяется по формуле:

$$P_i = \frac{(\lambda t)^i}{i!} e^{-\lambda t},$$

где λ — интенсивность потока заявок, т.е. среднее количество заявок, поступающих в единицу времени.

Пуассоновский поток представляет собой частный случай потока Пальма, так как обладает свойствами стационарности, ординарности и ограниченности последействия. Свойство ограниченности последействия для пуассоновского потока выполняется, так как интервалы времени между моментами поступления заявок в таком потоке представляют собой независимые случайные величины, распределенные по одному и тому же (экспоненциальному) закону.

Законы распределения интервалов времени между заявками и времени обслуживания в СМО. Интервалы времени между моментами поступления заявок и времена обслуживания заявок в СМО обычно представляют собой случайные величины. Во многих случаях эти величины описываются следующими законами распределения:

- экспоненциальный закон - если интервал времени между заявками или время их обслуживания может быть как очень коротким, так и очень длительным;
- равномерный закон - если интервал времени между заявками или время их обслуживания всегда принимает значение в пределах некоторого диапазона;
- гауссовский (нормальный) закон - если интервал времени между заявками или время их обслуживания в значительном большинстве случаев принимает значения, близкие к некоторой средней величине;
- закон Эрланга k-го порядка - если интервал времени между заявками или время их обслуживания представляет собой сумму k случайных величин, распределенных по экспоненциальному закону.

В некоторых случаях интервал времени между заявками или время их обслуживания может быть точно известным заранее, т.е. представляет собой не случайную, а детерминированную величину.

Поток заявок, в котором интервалы времени между заявками распределены по закону Эрланга k-го порядка, называется потоком Эрланга k-го порядка. Поток заявок, в котором интервалы времени между заявками представляют собой детерминированные величины, называется регулярным. Эти потоки представляют собой частные случаи потока Пальма.

Коэффициент вариации. Для расчета характеристик СМО обычно требуется знать коэффициенты вариации интервалов времени между заявками и времени обслуживания заявок. Коэффициент вариации любой случайной величины определяется по формуле

$$V = \frac{\sigma}{\overline{X}}, \qquad (8.2)$$

где σ - стандартное отклонение случайной величины;
X - математическое ожидание (среднее значение) случайной величины.
Физический смысл коэффициента вариации следующий: чем он больше, тем больше разброс возможных значений случайной величины, т.е. отклонение ее отдельных значений от среднего значения.

В табл. 8.1 приведены коэффициенты вариации для некоторых величин, часто применяемых для описания СМО.

Таблица 8.1

Распределение	Коэффициент вариации
Экспоненциальное	1
Распределение Эрланга k-го порядка	$1/\sqrt{k}$
Равномерное	$\frac{a-b}{(a+b)\sqrt{3}}$, где a и b - границы возможных значений величины
Детерминированная величина	0

В дальнейшем будем обозначать коэффициент вариации интервалов времени между заявками как ν, а коэффициент вариации времени обслуживания - как ε.

8.3. Типовой узел СМО. Классификация СМО

Любая СМО может быть представлена в виде одного или нескольких типовых узлов.

Примечание. Типовой узел, представляет собой многоканальную СМО. В такой СМО заявка направляется на любой из каналов обслуживания, оказавшийся свободным.

Имеется несколько вариантов классификации СМО (табл. 8.2).

Для описания СМО может использоваться обозначение $A/B/m$ - d, где A - обозначение закона распределения интервалов времени между заявками; B — обозначение закона распределения времени обслуживания заявок; m — количество каналов; d- обозначение дисциплины обслуживания. В качестве A и B обычно используются следующие обозначения: M - экспоненциальное распределение, G - любое другое. Для некоторых распределений используются специальные обозначения, например, D — детерминированная величина, Ek - распределение Эрланга k-го порядка, и т.д. Примеры таких обозначений будут приведены ниже.

Таблица 8.2 Классификация СМО

Признак классификации	Типы СМО	Описание
Законы распределения интервалов времени	Марковские	Интервалы времени между моментами поступления заявок и времена обслуживания заявок распределены по экспоненциальному закону
	Немарковские	Закон распределения интервалов времени между моментами поступления заявок и/или времен

		обслуживания заявок отличается от экспоненциального
Количество каналов	Одноканальные	В любой момент времени в СМО может обслуживаться не более одной заявки
	Многоканальные	В СМО одновременно может обслуживаться несколько заявок
Ограничения на очередь	Без ограничений на очередь	Если заявка поступает на обслуживание в момент, когда все каналы заняты, то она становится в очередь. Ограничений на количество заявок в очереди и на время пребывания в ней нет
	С ограничениями на очередь	Имеются ограничения на количество заявок в очереди и/или время их пребывания в очереди. Если заявка поступает на обслуживание в момент, когда в очереди уже имеется предельное количество заявок, или время пребывания заявки в очереди превышает допустимое, то она покидает СМО (не обслуживается)
	Без очереди	Если заявка поступает на обслуживание в момент, когда все каналы заняты, то она покидает СМО (не обслуживается). Очередь в такой СМО не образуется
Количество заявок	Замкнутые	Имеется фиксированное количество заявок, периодически требующих обслуживания.
	Разомкнутые	Количество заявок, которые могут поступать в СМО, не ограничено
	Смешанные	Имеется фиксированное количество заявок, периодически требующих обслуживания. При этом возможно поступление дополнительных заявок
Количество узлов СМО и связь между ними	Однофазные (одиночные)	Один типовой узел
	Многофазные	Последовательность типовых узлов. Все заявки, обслуженные в одном узле, направляются в следующий узел
	Сеть СМО	СМО, состоящая из нескольких типовых узлов. Количество узлов, в которых требуется обслуживание, и порядок их прохождения могут быть различными для разных заявок
Дисциплина обслуживания (порядок обслуживания заявок из очереди)	FIFO	«Первым пришел - первым обслужен» (обслуживание в порядке поступления)
	LIFO	«Первым пришел - последним обслужен»
	С относительными приоритетами	Первыми из очереди выбираются заявки с более высоким приоритетом. Если обслуживание заявки началось, то оно всегда доводится до конца, даже если в это время поступает заявка с более высоким приоритетом
	С абсолютными приоритетами	Первыми из очереди выбираются заявки с более высоким приоритетом. Обслуживание заявки прерывается, если поступает заявка с более высоким приоритетом

	Квантованное обслуживание	На обслуживание каждой заявки выделяется определенное время. Если за это время обслуживание не завершается, то заявка возвращается в очередь, и обслуживается следующая заявка
	По необходимому времени обслуживания	Первыми обслуживаются заявки, для которых требуется меньше времени

8.4. Параметры и характеристики СМО

Под параметрами СМО будем понимать величины, описывающие поток заявок СМО и каналы обслуживания.

Основным параметром потока заявок является его интенсивность (λ) — среднее количество заявок, поступающих в СМО в единицу времени.

Основные параметры каналов обслуживания - количество каналов (m), среднее время обслуживания заявки в канале ($\bar{x}$). В расчетах вместо величины x часто используется интенсивность обслуживания заявок $\mu = 1/\bar{x}$. Эта величина представляет собой среднее количество заявок, которое может быть обслужено одним каналом СМО в единицу времени. Другими словами, интенсивность обслуживания - это количество заявок, обслуживаемых каналом в единицу времени при условии, что канал никогда не простаивает из-за отсутствия заявок.

Параметром СМО с ограничением на количество заявок в очереди является также максимальное (предельно допустимое) количество заявок в очереди (n).

Под характеристиками СМО будем понимать величины, по которым можно оценивать эффективность работы СМО и выбирать лучший из нескольких вариантов СМО. В качестве характеристик СМО обычно используются следующие величины:

P_0 - вероятность простоя СМО. Эта величина показывает, какую часть от общего времени работы СМО все ее каналы свободны, т.е. простаивают из-за отсутствия заявок;

$P_{отк}$ — вероятность отказа. Эта величина показывает, какая доля всех поступающих заявок не обслуживается системой из-за занятости ее каналов или большого количества заявок в очереди. Для СМО без ограничений на очередь $P_{отк} = 0$;

$P_{обсл}$ – вероятность обслуживания. Эта величина показывает, какая доля всех поступающих заявок обслуживается системой. Очевидно, что $P_{обсл} = 1 - P_{отк}$.

Для СМО без отказов $P_{обсл} = 1$;

U - коэффициент загрузки СМО. Эта величина показывает, какую часть от общего времени своей работы СМО выполняет обслуживание заявок;

$\bar{q}$ - среднее число заявок в очереди (средняя длина очереди);

$\bar{S}$ - среднее число заявок на обслуживании (в каналах), или среднее число занятых каналов;

$\bar{k}$ - среднее число заявок в СМО, т.е. на обслуживании и в очереди;

$\bar{w}$ — среднее время пребывания заявки в очереди (среднее время ожидания обслуживания);

$\bar{t}$ — среднее время пребывания заявки в СМО, т.е. в очереди и на обслуживании;

γ - пропускная способность (среднее количество заявок, обслуживаемых в единицу времени).

Величины P_0, U и $\bar{S}$ характеризуют степень загрузки СМО. Эти величины представляют интерес с точки зрения стороны, осуществляющей эксплуатацию СМО. Например, если в качестве СМО рассматривается предприятие, выполняющее некоторые заказы, то эти величины представляют интерес для владельцев предприятия. Обычно желательно, чтобы коэффициент загрузки СМО имел значение на уровне 0,75 - 0,85. Значения U < 0,75 указывают, что СМО простаивает значительную часть времени, т.е. используется нерационально. Значения U > 0,85 указывают на перегрузку СМО.

Величины $P_{отк}$, $P_{обсл}$, $\bar{w}$ и $\bar{t}$ характеризуют качество обслуживания заявок. Они представляют интерес с точки зрения пользователей СМО. Желательна минимизация значений $P_{отк}$, $\bar{w}$, $\bar{t}$ и максимизация $P_{обсл}$.

Величина γ представляет собой среднее количество заявок, обслуживаемых СМО в единицу времени. Эта величина представляет интерес с точки зрения стороны, осуществляющей эксплуатацию СМО. Обычно желательна максимизация этой величины, особенно в случаях, когда обслуживание каждой заявки обеспечивает получение определенной прибыли.

Величины $\bar{q}$ и $\bar{k}$ обычно используются в качестве вспомогательных для расчета других характеристик СМО.

При расчете характеристик СМО используется следующая величина, называемая нагрузкой на СМО:

$$\rho = \frac{\lambda}{m\mu}. \qquad (8.3)$$

Величина ρ представляет собой отношение интенсивности потока заявок к интенсивности, с которой СМО может их обслуживать. Любая СМО без ограничений на очередь может нормально работать (т.е. обслуживать все поступающие заявки) только при условии, что $\rho < 1$. Величина $\rho > 1$ означает, что количество заявок, поступающих в СМО в единицу времени (λ), превышает количество заявок, которые СМО может обслужить в единицу времени ($m\mu$). В таких условиях в СМО без ограничений на очередь количество заявок, ожидающих обслуживания, будет постоянно возрастать, так как заявки будут поступать в СМО быстрее, чем она может их обслуживать. Для СМО с ограничениями на очередь и без очереди возможны любые значения ρ, так как в таких СМО часть заявок получает отказ, т.е. не допускается в СМО.

Приведем некоторые соотношения, которые могут применяться для расчета характеристик любой разомкнутой СМО.

Коэффициент загрузки:

$$U = \rho(1 - P_{отк}). \qquad (8.4)$$

Среднее число заявок на обслуживании (среднее число занятых каналов):

$$\bar{S} = mU. \qquad (8.5)$$

Среднее число заявок в СМО:

$$\bar{k} = \bar{q} + \bar{S}. \qquad (8.6)$$

Пропускная способность СМО:

$$\lambda = \mu\bar{S}, \qquad (8.7)$$

или

$$\gamma = \lambda(1 - P_{отк}). \qquad (8.8)$$

Среднее время пребывания заявки в очереди (формула Литтла):

$$\bar{w} = \frac{\bar{q}}{\gamma}. \qquad (8.9)$$

Среднее время пребывания заявки в СМО:

$$\bar{t} = \bar{w} + \bar{x}, \qquad (8.10)$$

или

$$\bar{t} = \frac{\bar{k}}{\gamma}. \qquad (8.11)$$

Формулы (8.4)—(8.11) могут применяться для расчета характеристик любых разомкнутых СМО, независимо от количества каналов, потока заявок, закона распределения времени обслуживания и т.д.

Для разомкнутых СМО без ограничений на очередь верны следующие формулы:

- коэффициент загрузки: $U = \rho$;
- пропускная способность: $\gamma = \lambda$.

Эти формулы представляют собой частные случаи формул (8.4) и (8.8) для $P_{отк} = 0$.

Вероятность простоя (P_0), вероятность отказа ($P_{отк}$) и средняя длина очереди ($\overline{q}$) рассчитываются по-разному в зависимости от типа СМО.

Точный расчет характеристик возможен только для марковских СМО. Для немарковских СМО без ограничений на очередь возможен приближенный расчет характеристик. Для определения характеристик СМО других типов применяются специальные методы (например, методы имитационного моделирования), не рассматриваемые в данном пособии.

На основе рассмотренных характеристик СМО могут рассчитываться другие показатели, характеризующие эффективность ее работы.

8.5. Вероятности состояний СМО

Вероятности состояний СМО - это вероятности пребывания в СМО определенного количества заявок. Обычно при вычислении вероятностей состояний требуется определять величины P_j - вероятности пребывания в СМО ровно j заявок. Например, P_2 — вероятность того, что в СМО (т.е. на обслуживании и в очереди) находятся ровно две заявки. Если, например, $P_2 = 0{,}15$, это означает, что в течение 15% времени (от всего времени работы СМО) в ней находятся ровно две заявки. В течение остального времени (85%) количество заявок в СМО составляет менее двух (одну или ни одной) или более двух (три или больше).

Величины P_j вычисляются по-разному для разных типов СМО. Точный расчет P_j возможен только для марковских СМО, т.е. СМО типа M/M/m.

Определив вероятности P_j, можно, используя формулы теории вероятностей, найти вероятности других состояний СМО, например, следующие:

- вероятность пребывания в СМО не более R заявок:

$$P(j \le R) = \sum_{j=0}^{R} P_j; \qquad (8.12)$$

- вероятность пребывания в СМО более R заявок:

$$P(j > R) = 1 - \sum_{j=0}^{R} P_j. \qquad (8.13)$$

Вероятности состояний обычно требуются в качестве промежуточных величин для вычисления других характеристик СМО.

8.6. Экономические характеристики СМО

Под экономическими характеристиками будем понимать величины, выражающие прибыль от работы СМО, затраты на обслуживание заявок и т.д. Расчет таких характеристик зависит от постановки задачи. Приведем несколько общих формул, применимых в большинстве задач.

Выручка от обслуживания заявок в СМО в течение времени Т:

$$V = \gamma C T, \qquad (8.14)$$

где γ - пропускная способность СМО;

C - выручка от обслуживания одной заявки.
Затраты, связанные с обслуживанием заявок в СМО в течение времени Т:

$$Z_{обс} = \gamma C_{обс} T, \qquad (8.15)$$

где $C_{обс}$ - затраты, связанные с обслуживанием одной заявки.

Затраты, связанные с эксплуатацией СМО в течение времени Т:

$$Z_{эксп} = (\bar{S} C_{раб} + (m - \bar{S}) C_{пр}) T, \qquad (8.16)$$

где m — количество каналов в СМО;
$\bar{S}$ — среднее число заявок на обслуживании (в каналах), или среднее число занятых каналов;
$C_{раб}$ - затраты, связанные с работой одного канала в течение единицы времени;
$C_{пр}$ – затраты, связанные с простоем одного канала в течение единицы времени. Убытки, связанные с отказами в обслуживании за время Т:

$$Z_{отк} = \lambda C_{отк} P_{jnr} T, \qquad (8.17)$$

где λ - интенсивность потока заявок;
$C_{отк}$ - убытки, связанные с отказом в обслуживании одной заявки;
$P_{отк}$ - вероятность отказа.

Убытки за время Т, связанные с пребыванием заявок в СМО (как в очереди, так и на обслуживании):

$$z_{пр} = \bar{k} C_{пр} T, \qquad (8.18)$$

где $\bar{k}$ - среднее число заявок в СМО;
$C_{пр}$ - убытки, связанные с пребыванием заявки в СМО в течение единицы времени.

8.7. Одноканальные СМО без ограничений на очередь

Для расчета характеристик таких СМО применяются следующие формулы. Вероятность простоя:

$$P_0 = 1 - \rho. \qquad (8.19)$$

Средняя длина очереди:

$$\bar{q} = \frac{\rho^2 (v^2 + \varepsilon^2)}{2(1-\rho)}, \qquad (8.20)$$

где ν - коэффициент вариации интервалов времени между заявками;
ε - коэффициент вариации времени обслуживания.

Формула (8.20) позволяет точно рассчитать среднюю длину очереди только для СМО типа $M/M/1$, т.е. для марковских СМО. Для других СМО величина $\bar{q}$, найденная по формуле (8.20), является приближенной.

Для СМО типа $M/M/1$ можно также определить следующие величины. Вероятности пребывания в СМО j заявок:

$$P_j = \rho^j (1-\rho), \; j = 1, 2, ... \qquad (8.21)$$

Вероятность того, что время пребывания заявки в СМО превысит некоторую заданную величину Т:

$$P(t > T) = e^{-\mu(1-\rho)T}. \qquad (8.22)$$

Пример 8.1. В небольшой мастерской имеется один стенд для ремонта и наладки некоторых приборов. В течение рабочего дня в мастерскую в среднем поступает 10 заявок на ремонт приборов; поток заявок можно считать пуассоновским. За рабочий день на стенде можно отремонтировать 12 приборов. Время ремонта одного прибора - случайная величина, которую можно приближенно считать экспоненциальной. Приборы, ожидающие ремонта, размещаются на специальных стеллажах, составляемых из стандартных секций. Одна секция вмещает 6 приборов. Если на стеллажах нет места, то прибор приходится размещать на полу, что нежелательно.

За ремонт одного прибора мастерская берет плату в размере 40 ден. ед., если заказ на ремонт выполняется в течение одного дня, и 30 ден. ед. - если выполнение заказа занимает более одного дня. Затраты, связанные с ремонтом одного прибора, составляют 25 ден. ед.

Найти характеристики работы мастерской. Найти среднюю прибыль мастерской за один рабочий день. Определить необходимую емкость стеллажей для приборов, ожидающих ремонта, чтобы вероятность их переполнения не превышала 5%.

Поток заявок (приборов) является пуассоновским, время обслуживания распределено по экспоненциальному закону, и имеется один канал обслуживания. Поэтому мастерскую можно представить как СМО типа $M/M/1$. В этой СМО $\lambda = 10$ заявок/день, $\mu = 12$ заявок/день, $\bar{x} = 1/12 = 0,083$ дня. Число каналов m = 1.

Найдем нагрузку на СМО: $\rho = \lambda/(m\mu) = 0,83$.

Найдем вероятность простоя по формуле (8.19): $P_0 = 0,17$.

Найдем среднюю длину очереди по формуле (8.20). Из табл. 8.1 найдем коэффициенты вариации. Так как поток заявок пуассоновский, интервалы времени между заявками - экспоненциальные случайные величины. Поэтому $\nu = 1$. Время обслуживания заявок (т.е. ремонта приборов) - также экспоненциальная случайная величина, поэтому $\varepsilon = 1$. По формуле (8.20) получим: $\bar{q} = 4,17$ приборов.

Мастерская принимает на обслуживание все поступающие приборы (отказов в обслуживании нет). Поэтому $P_{отк} = 0, P_{обсл} = 1$.

Найдем остальные характеристики СМО по формулам (8.4)—(8.11): $U = 0,83$; $\bar{S} = 0,83$ прибора; $\bar{k} = 5$ приборов; $\gamma = 10$ приборов/день; $\bar{w} = 0,417$ дня; $\bar{t} = 0,5$ дня.

Проанализируем полученные характеристики СМО.

Мастерская загружена на 83%, т.е. занята ремонтом приборов в течение 83% всего времени своей работы. В течение 17% времени мастерская простаивает из-за отсутствия заказов. Таким образом, загрузка мастерской достаточно высока. Такую загрузку можно считать нормальной. Однако дальнейшее увеличение загрузки нежелательно. Поэтому в случае увеличения потока приборов, поступающих на ремонт, потребуется оснащение мастерской дополнительными стендами.

В среднем в очереди находится 4,17 прибора, а в мастерской (т.е. в очереди и на обслуживании) - 5 приборов. Мастерская обслуживает в среднем 10 приборов в день, т.е. все поступающие приборы. Время от поступления прибора в мастерскую до начала его ремонта (т.е. время пребывания прибора в очереди) составляет в среднем 0,417 дня. Время от поступления прибора в мастерскую до окончания его ремонта (время пребывания прибора в мастерской) составляет в среднем 0,5 дня.

Найдем прибыль от работы мастерской за один рабочий день. При этом необходимо учесть, что выручка мастерской от ремонта одного прибора может быть разной. Она составляет 40 ден. ед., если заказ на ремонт выполняется в течение одного дня, и 30 ден. ед. - в противном случае. По формуле (8.22) найдем долю заказов, время выполнения которых превышает один день: $P(t > T) = e^{-12\cdot(1-0,83)\cdot 1} = 0,13$. Значит, в 13% случаев выручка от ремонта одного прибора составляет 30 ден. ед., в остальных случаях - 40 ден. ед. Можно считать, что средняя выручка от ремонта одного прибора составляет 30·0,13 + 40·0,87 = 38,7 ден. ед. Мастерская ремонтирует в среднем 10 приборов в день, затрачивая на ремонт каждого прибора 25 ден. ед. Таким образом, прибыль мастерской за один день составит в среднем (38,7 - 25)·10 =127 ден. ед.

Найдем необходимую емкость стеллажей для приборов, ожидающих ремонта. Предположим, что стеллажи состоят из одной стандартной секции, т.е. вмещают 6 приборов. Найдем вероятность их переполнения. Стеллажи будут переполнены, если в мастерской будет находиться более семи приборов, так как в этом случае один прибор будет находиться на обслуживании (на стенде), шесть - на стеллажах, остальные - на полу. Таким образом, вероятность переполнения стеллажей, вмещающих шесть приборов, равна вероятности пребывания в мастерской восьми и более приборов. Эта вероятность определяется по формуле (8.13): $P(j>7)=1-P_0-P_1-P_2-P_3-P_4-P_5-P_6-P_7$. Величина P_0 найдена выше: $P_0=0,17$. Остальные вероятности найдем по формуле (8.21): $P_1=0,141, P_2=0,117, P_3=0,097, P_4=0,081, P_5=0,067, P_6=0,056, P_7=0,046$. Здесь P0 – вероятность простоя (вероятность того, что в СМО нет ни одной заявки); P_1 - вероятность пребывания в СМО ровно одной заявки (вероятность того, что на обслуживании в СМО находится заявка, но в очереди заявок нет); P_2 -вероятность пребывания в СМО ровно двух заявок (вероятность того, что на обслуживании в СМО находится заявка, и еще одна заявка находится в очереди), и т.д. Вероятность переполнения стеллажей следующая:

$$P(j>7)=1-0,17-0,141-0,117-0,097-0,081-0,067-0,056-0,046=0,225.$$

Эта вероятность превышает 5%, поэтому одной секции недостаточно.

Найдем вероятности переполнения стеллажей большей емкости:

- из двух секций: $P(j>13)=0,074;$ $P(j>19)=0,024.$
- из трех секций: P(j > 19) = 0,024.

Таким образом, чтобы вероятность переполнения стеллажей не превышала 5%, они должны состоять не менее чем из трех секций, т.е. вмещать 18 приборов.

Пример 8.2. Станок-автомат используется для выпуска некоторых деталей. Интервалы между моментами поступления деталей на обработку составляют от 3 до 8 минут. Время обработки детали на станке-автомате - случайная величина, распределенная по гауссовскому закону со средним значением 4 мин и стандартным отклонением 30 с. Затраты на один час работы станка-автомата составляют 40 ден. ед., на один час простоя - 15 ден. ед. Материал, необходимый для изготовления детали, стоит 1,5 ден. ед. Готовые детали продаются по цене 10 ден. ед.

Найти характеристики работы станка и среднюю прибыль от его работы за одну рабочую смену (8 часов).

Интервалы времени между моментами поступления заявок (деталей) представляют собой случайные величины, распределенные по равномерному закону. Время обработки детали - гауссовская случайная величина. Имеется один канал обслуживания. Поэтому станок можно представить как СМО типа $G/G/1$. Интервал между деталями составляет в среднем 5,5 мин, поэтому $\lambda=1/5,5=0,182$ детали/мин. Обработка одной детали занимает в среднем 4 мин, поэтому $x=4$ мин, $m=0,25$ детали/мин. Число каналов $m=1$.

Найдем нагрузку на СМО: $\rho=\lambda/(m\mu)=0,728$.

Найдем вероятность простоя по формуле (8.19): $P_0=0,272$.

Найдем среднюю длину очереди по формуле (8.20). Коэффициент вариации для интервалов между моментами поступления деталей найдем из табл. 8.1 (для равномерного закона): $v=(8-3)/((8+3)\sqrt{3})=0,262$. Коэффициент вариации для времени обработки деталей найдем по формуле (8.2): $\varepsilon=0,5/4=0,125$. По формуле (8.20) получим: $\bar{q}=0,082$ деталей.

Станок обрабатывает все поступающие детали. Поэтому $P_{отк}=0, P_{обсл}=1$. Найдем остальные характеристики СМО по формулам (8.4)—(8.11): $U=0,728$, $\bar{S}=0,728$ детали, $\bar{k}=0,81$ детали, $\gamma=0,182$ детали/мин, $\bar{w}=0,45$ мин, $\bar{t}=4,45$ мин. Найдем прибыль от работы станка за рабочую смену. По формуле (8.14) найдем выручку от продажи деталей, выпускаемых за смену: V= 0,182·10·480 = 873,6 ден. ед. (здесь 480 - продолжительность рабочей смены в минутах). По формуле (8.15) найдем затраты на материал для изготовления деталей: $Z_{обсл}$ = =0,182·1,5·480= 131,04 ден. ед. По формуле (8.16) найдем затраты, связанные с эксплуатацией станка в течение рабочей смены: $Z_{эксп}$ = (0,728·40 + (1 - 0,728)·15)·8 = 265,6 ден. ед. (здесь 8 - продолжительность рабочей смены в часах, так как в формуле используются величины затрат за один час работы и простоя станка). Найдем прибыль от работы станка: 873,6 - 131,04- - 265,6 = 476,96 ден. ед.

8.8. Многоканальные СМО без ограничений на очередь

Для многоканальных СМО точный расчет характеристик возможен только при условии, что СМО является марковской, т.е. для СМО типа $M/M/m$. Для расчета характеристик таких СМО применяются следующие формулы.
Вероятность простоя:

$$P_0 = \left[\sum_{i=0}^{m-1} \frac{(m\rho)^i}{i!} + \frac{(m\rho)^m}{m!(1-\rho)} \right]^{-1}, \qquad (8.23)$$

где m — количество каналов (т.е. количество заявок, которые могут обслуживаться в СМО одновременно).

Средняя длина очереди:

$$\bar{q} = \frac{\rho(m\rho)^m}{m!(1-\rho)^2} P_0. \qquad (8.24)$$

Вероятности пребывания в СМО j заявок:

$$P_j = \begin{cases} \dfrac{(m\rho)^j}{j!} P_0, \; j = 1, ..., m, & (8.25) \\ \dfrac{(m\rho)^j}{m^{j-m} m!} P_0, \; j > m. & (8.26) \end{cases}$$

Формула (8.25) позволяет найти вероятности состояний СМО, при которых очередь отсутствует (количество заявок, обслуживаемых в СМО, не превышает количества каналов), а формула (8.26) - вероятности состояний при наличии очереди.

Примечание. Приведенные формулы могут применяться и для приближенного расчета характеристик немарковских многоканальных СМО (т.е. СМО типа $M/G/m$, $G/M/m$ или $G/G/m$).

Пример 8.3. В ремонтной службе предприятия выполняется наладка некоторых механизмов. На наладку поступает в среднем 10 механизмов в час (поток механизмов можно считать пуассоновским). Наладка одного механизма занимает в среднем 15 мин (время наладки инструмента можно считать экспоненциальной случайной величиной). В ремонтной службе работают три наладчика. Заработная плата наладчика составляет 30 ден. ед. в день.
В то время, когда механизм находится в ремонтной службе (т.е. налаживается или ожидает наладки), он не может использоваться для работы. Простой механизма в течение часа приносит предприятию убытки в размере 6 ден. ед.

Найти: а) характеристики работы ремонтной службы; б) потери предприятия в течение рабочей смены (8 часов), связанные с наладкой инструментов, включая затраты на содержание ремонтной службы и убытки от простоя механизмов; в) вероятность того, что наладка механизма начнется сразу же после его поступления (без ожидания в очереди); г) вероятность того, что количество механизмов, ожидающих наладки, окажется свыше пяти; д) определить, целесообразно ли уменьшить количество наладчиков до двух; е) определить, целесообразно ли увеличить количество наладчиков до четырех.

а) Ремонтную службу можно рассматривать как СМО типа M/M/3 без ограничений на очередь. В этой СМО $\lambda = 10$ механизмов/час = 0,167 механизма/мин, $\bar{x} = 15$ мин, $\mu = 0,067$ механизма/мин.

Найдем нагрузку на СМО: $\rho = \lambda/(m\mu)$.

Найдем вероятность простоя по формуле (8.23):

$$P_0 = \left[\frac{(3\cdot 0,833)^0}{0!} + \frac{(3\cdot 0,833)^1}{1!} + \frac{(3\cdot 0,833)^2}{2!} + \frac{(3\cdot 0,833)^3}{3!(1-0,833)}\right]^{-1} = 0,046.$$

По формуле (8.24) определяем среднюю длину очереди (т.е. среднее количество механизмов, ожидающих наладки): $\bar{q} = 3,43$ механизма.

Ремонтная служба выполняет наладку всех поступающих механизмов. Поэтому $P_{отк} = 0, P_{обсл} = 1$. Найдем остальные характеристики по формулам (8.4)—(8.11): $U = 0,833, \bar{S} = 2,49$ механизма, $\bar{k} = 5,92$ механизма, γ = 0,167 механизма/мин, $\bar{w}$ = 20,5 мин, $\bar{t}$ = 35,5 мин.

б) Затраты на содержание ремонтной службы составляют 30·3 = 90 ден. ед. Убытки предприятия, связанные с простоем механизмов, найдем по формуле (8.18): $Z_{пр}$ = 5,92·6·8 = 284,16 ден. ед. за смену. Таким образом, полные потери предприятия, связанные с наладкой механизмов, составляют 90 + 284,16 = = 374,16 ден. ед. за смену.

в) Найдем вероятность того, что наладка механизма начнется сразу же после его поступления. Это произойдет в случае, если в момент поступления механизма в ремонтную службу хотя бы один наладчик окажется свободным. Для этого требуется, чтобы количество механизмов, находящихся в ремонтной службе, не превышало двух.

Вероятность такого состояния находится по формуле (8.12): $P(j \le 2) = P_0 + P_1 + P_2$.

Вероятности P_1 и P_2 найдем по формуле (8.25): P_1 = 0,115, P_2 = 0,144. Таким образом, $P(j \le 2)$ = 0,046 + 0,115 + 0,144 = 0,305. Это значит, что примерно в 30,5% случаев механизм, доставленный в ремонтную службу, сразу же поступит к наладчику.

г) Найдем вероятность того, что количество механизмов, ожидающих наладки, окажется свыше пяти. Такое состояние означает, что количество механизмов, находящихся в ремонтной службе, превышает восемь (три из них - на наладке, остальные - в очереди). Вероятность такого состояния находится по формуле (8.13): $P(j > 8) = 1 - (P_0 + P_1 + ... + P_8)$. Вероятности P_1, P_2, P_3 найдем по формуле (8.25): P1 =0,115, P2 = 0,144, P3 = 0,12. Вероятности P4, P5, …,P 8 найдем по формуле (8.26): P4 = 0,1,P5 = 0,083, P6 = 0,069, P7 = 0,058, P8 = 0,048. Таким образом, P(j > 8) = 0,218.

д) Найдем нагрузку на СМО при $m = 2: \rho = \lambda/(m\mu)$. Величина ρ > 1 означает, что механизмы поступают в ремонтную службу с большей интенсивностью, чем она может их обслуживать. Таким образом, два наладчика «не справятся» с потоком заявок. Уменьшить количество наладчиков до двух нельзя.

е) Найдем характеристики СМО при m = 4: ρ = 0,623, P_0 = 0,074, $\bar{q}$ = 0,53 механизма, $P_{отк}$ = 0, $P_{обсл}$ = 1, U = 0,623, $\bar{S}$ = 2,49 механизма, $\bar{k}$ = 3,02 механизма, $\bar{w}$ = 3,14 мин, $\bar{t}$ = 18,14 мин, γ = 0,167 механизма/мин.

Затраты на содержание ремонтной службы составляют 30·4= 120 ден. ед. По формуле (8.18) найдем убытки предприятия, связанные с простоем механизмов: $Z_{пр}$ = 3,02·6·8 = 144,96 ден. ед. за смену. Таким образом, полные потери предприятия, связанные с наладкой инструментов, составляют 120 + 144,96 = 264,96 ден. ед. за смену. Эти потери меньше, чем для трех наладчиков. Поэтому увеличение количества наладчиков до четырех следует признать выгодным.

8.9. СМО с ограничением на длину очереди

В СМО такого типа в очереди может находиться не более n заявок. Если заявка поступает в СМО в момент, когда в очереди уже находятся n заявок, то она не обслуживается (не допускается в очередь). Для расчета характеристик таких СМО применяются следующие формулы.

Вероятность простоя:

$$P_0 = \left[\sum_{i=0}^{m} \frac{(m\rho)^i}{i!} + \frac{(m\rho)^{m+1}}{m \cdot m!} \cdot \frac{1-\rho^n}{1-\rho} \right]^{-1}, \tag{8.27}$$

где m — количество каналов СМО;
n — максимально допустимое количество заявок в очереди.

Вероятность отказа в обслуживании:

$$P_{отк} = \frac{(m\rho)^{m+n}}{m^n \cdot m!} P_0. \tag{8.28}$$

Средняя длина очереди:

$$\bar{q} = \frac{(m\rho)^{m+1} \cdot P_0}{m \cdot m!} \cdot \frac{1-(n+1)\rho^n + n\rho^{n+1}}{(1-\rho)^2}. \tag{8.29}$$

Вероятности пребывания в СМОj заявок:

$$P_j \begin{cases} \dfrac{(m\rho)^j}{j!} P_0, \ j = 1, \ldots, m, & (8.30) \\ \dfrac{(m\rho)^j}{m^{j-m} m!} P_0, \ j = m+1, \ldots, m+n. & (8.31) \end{cases}$$

Формула (8.30) позволяет найти вероятности состояний СМО, при которых очередь отсутствует (количество заявок, обслуживаемых в СМО, не превышает количества каналов), а формула (8.31) - вероятности состояний при наличии очереди.

При ρ = 1 расчет вероятности простоя и средней длины очереди выполняется по следующим формулам:

$$P_0 = \left[\sum_{i=0}^{m} \frac{(m\rho)^i}{i!} + \frac{n(m\rho)^m}{m!} \right]^{-1};$$

$$\bar{q} = \frac{(m\rho)^m \cdot P_0}{m!} \cdot \frac{n(n+1)}{2}.$$

Пример 8.4. Предприятие выполняет заказы на переводы с иностранных языков. В среднем на предприятие поступает 8 заказов в день (поток заказов можно считать пуассоновским). Средний размер перевода - 5 страниц (размер перевода можно считать случайной величиной, распределенной по экспоненциальному закону). На предприятии работают 4 переводчика. Норма для переводчика - 7 страниц в день. Чтобы исключить невыполнение заказов в срок, предприятие не принимает новые заказы, если уже имеется 6 переводов, ожидающих выполнения.

Переводчику выплачивается 2 ден. ед. за каждую переведенную страницу, плюс 100 ден. ед. в месяц. Заказчик платит предприятию 4 ден. ед. за каждую переведенную страницу. Найти а) характеристики работы предприятия; б) среднюю заработную плату переводчика за месяц (25 рабочих дней); в) среднюю прибыль предприятия за месяц; г) определить, выгодно ли предприятию принять на работу еще одного переводчика.

а) На предприятии работают 4 переводчика (m = 4). Поток заказов - пуассоновский, время их выполнения - экспоненциальная случайная величина. Поэтому предприятие можно рассматривать как СМО типа М/М/4 с ограничением на длину очереди (n = 6). В этой СМО λ = = 8 заказов/день. Так как переводчик может перевести в среднем 7 страниц в день, а средний размер перевода - 5 страниц, значит, производительность переводчика (интенсивность обслуживания заявок) составляет μ = 7/5 = 1,4 заказа/день. Среднее время работы переводчика над заказом $x = 1/\mu$ = 0,71 дня.

Найдем нагрузку на СМО: $\rho = \lambda/(m\mu) = 1,43$.

Найдем вероятность простоя по формуле (8.27):

$$P_0=\left[\frac{(4\cdot 1,43)^0}{0!}+\frac{(4\cdot 1,43)^1}{1!}+\frac{(4\cdot 1,43)^2}{2!}+\frac{(4\cdot 1,43)^3}{3!}+\frac{(4\cdot 1,43)^4}{4!}+\frac{(4\cdot 1,43)^5}{4\cdot 4!}+\frac{1-1,43^6}{1-1,43}\right]^{-1}=0,000827.$$

По формуле (8.28) определяем вероятность отказа в обслуживании: $P_{отк}$ = 0,31. Это означает, что примерно 31% заказчиков, обращающихся на предприятие, получают отказ из-за перегруженности переводчиков. Вероятность обслуживания составляет $P_{обсл}=1-P_{отк}=0,69.$

По формуле (8.29) найдем среднюю длину очереди (т.е. среднее количество заказов, ожидающих выполнения): $\bar{q}$ = 4,1 заказа.

Найдем остальные характеристики по формулам (8.4)—(8.11): U = 0,982, $\bar{S}$ = 3,93 заказа, $\bar{k}$ = 8,03 заказа, γ = 5,5 заказа/день, $\bar{w}$ = 0,75 дня, $\bar{t}$ =1,46 дня.

б) Найдем заработную плату переводчика за месяц (25 рабочих дней). Переводчики, работающие на предприятии, выполняют в среднем 5,5 заказа в день (γ = 5,5). Средний размер заказа - 5 страниц; за каждую страницу переводчику выплачивается 2 ден. ед. Таким образом, сумма, выплачиваемая всем переводчикам за выполнение заказов в течение месяца, составляет в среднем 5,5·5·2·25 = 1375 ден. ед. Заработная плата каждого из переводчиков составляет в среднем 1375/4 + 100 = 443,75 ден. ед. в месяц.

в) Найдем среднюю прибыль предприятия за месяц. Выручка предприятия от выполнения переводов за месяц составляет в среднем 5,5·5·4·25 = 2750 ден. ед. Переводчикам выплачивается 4·443,75 = 1775 ден. ед. Таким образом, прибыль предприятия составляет 2750 - 1775 = 975 ден. ед. в месяц.

г) Определим, выгодно ли предприятию принять на работу еще одного переводчика. Найдем характеристики работы предприятия для m = 5: ρ = 1,14, P_0 = 0,00154, $P_{отк}$ = 0,175, $P_{обсл}$ = = 0,825, $\bar{q}$ = 2,99 заказа, U = 0,943, $\bar{S}$ = 4,72 заказа, $\bar{k}$ = 7,71 заказа, γ = 6,6 заказа/день, $\bar{w}$= = 0,45 дня, $\bar{t}$ = 1,16 дня. Выполнив расчеты, как показано выше, найдем, что выручка предприятия за месяц составит 6,6·5·4·25 = 3300 ден. ед., заработная плата переводчика - (6,6·2·4·25)/5 + 100 = 430 ден. ед., прибыль предприятия - 3300 - 5·430 = 1150 ден. ед.

Таким образом, прием на работу еще одного переводчика приведет к росту прибыли предприятия с 975 до 1150 ден. ед., т.е. на 175 ден. ед. в месяц. Рост прибыли достигнут за счет уменьшения количества отказов, в результате чего увеличивается пропускная способность предприятия. Снижение количества отказов также является положительным результатом, так как улучшает репутацию предприятия. Еще одним положительным результатом является сокращение среднего срока выполнения заказа (с 1,46 до 1,16 дня). Отрицательным результатом является некоторое снижение заработной платы переводчика (с 443,75 до 430 ден. ед. в месяц). Однако предприятие имеет возможность избежать этого, использовав часть дополнительной прибыли на повышение заработной платы переводчиков. Например, если предприятие будет выплачивать переводчикам дополнительно по 15 ден. ед., то заработная плата переводчиков не снизится (и даже увеличится с 443,75 до 445 ден. ед.), а дополнительная прибыль предприятия составит 175-5·15 = 100 ден. ед. Таким образом, прием на работу еще одного переводчика следует признать выгодным.

8.10. СМО без очереди

В СМО такого типа заявки никогда не становятся в очередь. Если заявка поступает в СМО в момент, когда все каналы заняты, то она не обслуживается (не допускается в СМО). Для расчета характеристик таких СМО применяются следующие формулы.

Вероятность простоя:

$$P_0 = \left[\sum_{i=0}^{m} \frac{(m\rho)^i}{i!} \right]^{-1}. \qquad (8.32)$$

Вероятность отказа в обслуживании:

$$P_{отк} = \frac{(m\rho)^m}{m!} P_0. \qquad (8.33)$$

Вероятности пребывания в СМО j заявок:

$$P_j = \frac{(m\rho)^j}{j!} P_0,\ j = 1,...,m. \qquad (8.34)$$

Пример 8.5. В телефонную справочную службу поступает в среднем два запроса в минуту (поток запросов можно считать пуассоновским). Длительность разговора можно считать экспоненциальной случайной величиной; в среднем разговор занимает 1,5 минуты. Если в момент звонка в справочную службу все телефоны оказываются занятыми, то клиент вынужден отказаться от разговора (положить трубку).

Найти, сколько телефонов должна иметь справочная служба, чтобы обслуживать не менее 85% клиентов. Для выбранного количества телефонов найти характеристики справочной службы.

Справочную службу можно рассматривать как СМО типа $M/M/m$ без очереди. Величину m (количество каналов) требуется определить. В данной СМО λ = 2 звонка/мин, $\bar{x}$ = 1,5 мин, μ = 0,67 звонка/мин.

Предположим, что в справочной службе имеется только один телефон (m = 1). Тогда $\rho = \lambda/(m\mu) = 3$. По формуле (8.32) найдем вероятность простоя СМО:

$$P_0 = \left[\frac{(1\cdot 3)^0}{0!} + \frac{(1\cdot 3)^1}{1!} \right]^{-1} = 0,25.$$

По формуле (8.33) найдем вероятность отказа: $P_{отк}$ = 0,75. Таким образом, 75% клиентов, обратившихся в справочную службу, не будут обслужены из-за ее загруженности. Такой вариант справочной службы явно неприемлем.

Будем увеличивать количество телефонов (т.е. число каналов m) до тех пор, пока вероятность отказа не примет значение, не превышающее 0,15. Значения вероятности отказа для различного количества телефонов приведены в табл. 8.3.

Таблица 8.3

т	1	2	3	4	5
Р от к	0, 7 5	0, 5 3	0, 3 4	0, 2	0, 1 1

Таким образом, для обеспечения заданного уровня качества работы справочной службы она должна иметь пять телефонов.

Для выбранного варианта справочной службы вероятность простоя составляет P_0 = 0,06 (эта величина была необходима для вычисления Ротк). По формулам (8.4)-(8.11) найдем остальные характеристики: U = 0,53, $P_{обсл}$ = 0,89, $\bar{q}$ =0 (так как очередь не образуется), $\bar{S}$ = 2,66 звонка, $\bar{k}$ = 2,66 звонка, γ = 1,78 звонка/мин, $\bar{w}$ = 0, $\bar{t}$ = 1,5 мин. Таким образом, справочная служба обслуживает 89% клиентов. Недостатком выбранного варианта является низкая загрузка. Однако уменьшить количество телефонов для устранения этого недостатка нельзя, так как при этом увеличится количество отказов.

8.11. СМО с заявками с разным временем обслуживания

В таких СМО обслуживаются заявки нескольких типов, различающихся по времени обслуживания. Пусть в СМО обслуживается R типов заявок. Обозначим долю заявок i-го типа в потоке заявок как $P_i, i=1,...,R, P_1+P_2+...+P_R=1$. Времена обслуживания заявок разных типов представляют собой случайные (или детерминированные) величины; для расчета характеристик СМО законы распределения этих величин должны быть известны. Среднее время обслуживания заявки i-го типа обозначим как $\overline{x}_i, i=1,...,R.$

Для расчета характеристик таких СМО необходимо определить среднее время обслуживания и коэффициент вариации времени обслуживания всех заявок в СМО. Среднее время обслуживания заявок находится по формуле

$$\overline{x}=\sum_{i=1}^{R}P_i\overline{x}_i. \qquad (8.35)$$

Для определения коэффициента вариации времени обслуживания всех заявок применяются формулы, известные из теории вероятностей. Коэффициент вариации вычисляется следующим образом.

1. Находятся дисперсии времени обслуживания заявок каждого типа: $D_i, i=1,...,R$. Для этого должны быть известны законы распределения времени обслуживания заявок.

2. Находятся вторые начальные моменты времени обслуживания заявок каждого типа:

$$\alpha_i=D_i+\overline{x}_i^2, i=1,...,R. \qquad (8.36)$$

Примечание. Второй начальный момент случайной величины - это математическое ожидание (т.е. среднее значение) квадрата этой величины.

3. Находится второй начальный момент времени обслуживания всех заявок:

$$\alpha=\sum_{i=1}^{R}P_i\alpha_i. \qquad (8.37)$$

4. Находится дисперсия времени обслуживания всех заявок:

$$D=\alpha-\overline{x}^2. \qquad (8.38)$$

5. Находится коэффициент вариации времени обслуживания всех заявок:

$$\varepsilon=\frac{\sqrt{D}}{\overline{x}}. \qquad (8.39)$$

Дальнейший расчет характеристик СМО выполняется точно так же, как для любой СМО типа $M/G/1$, $G/G/1$, $M/G/m$ или $G/G/m$ (см. подразделы 8.7, 8.8).

Пример 8.6. Детали, при изготовлении которых допущен дефект, направляются для устранения дефекта на специальный станок. При изготовлении деталей возможны дефекты трех видов (обозначим их как A, B и C). Детали, имеющие несколько дефектов (два или три) одновременно, не подлежат ремонту; поэтому каждая деталь, поступающая на станок, имеет только один дефект. Детали, имеющие дефект A, поступают на станок в среднем через каждые 20 мин, с дефектом B — через каждые 10 мин, с дефектом C - через каждые 25 мин. Потоки деталей с дефектами каждого вида можно считать пуассоновскими. Устранение дефекта A занимает от 2 до 5 мин, устранение дефекта B — ровно 5 мин; время устранения дефекта C представляет собой гауссовскую случайную величину со средним значением 7 мин и стандартным отклонением 1,5 мин.

Требуется найти характеристики работы станка.

Так как потоки деталей с дефектами каждого вида можно считать пуассоновскими, поток всех деталей также можно считать пуассоновским (подробнее об этом см. в подразделе 8.13). Найдем интенсивности потоков деталей с дефектами каждого вида: λ_A = 1/20 = 0,05 детали/мин, λ_B = 1/10 = 0,1 детали/мин, λ_C = 1/25 = 0,04 детали/мин. Интенсивность потока всех деталей представляет собой сумму интенсивностей отдельных потоков:

$\lambda = \lambda_A + \lambda_B + \lambda_C = 0{,}19$ детали/мин.

Найдем доли деталей каждого вида в общем потоке деталей: P_A = 0,05/0,19 = 0,26; P_B = 0,1/0,19 = 0,53; P_C = 0,04/0,19 = 0,21.

Найдем среднее время обработки деталей всех видов. Средние времена обработки деталей каждого вида следующие: $\overline{x}_A$ = (2 + 5)/2 = 3,5 мин, $\overline{x}_B$ == 5 мин, $\overline{x}_C$ = 7 мин. Среднее время обработки деталей всех видов: x = 0,26·3,5 + 0,53·5 + 0,21·7 = 5,03 мин.

Найдем коэффициент вариации времени обработки всех деталей. Для этого необходимо определить дисперсии времен обработки деталей каждого вида.

Время обработки деталей с дефектом А - случайная величина, распределенная по равномерному закону. Из теории вероятностей известно, что дисперсия такой величины определяется по формуле: D = (b − a) /(23), где a и b - границы диапазона значений случайной величины. Таким образом, D_A = (5 − 2) /(23) = 0,87.

Время обработки деталей с дефектом В - детерминированная (точно известная) величина. Для таких величин дисперсия равна нулю: D_C = 0.

Время обработки деталей с дефектом С - случайная величина, распределенная по гауссовскому закону. Для нее известно стандартное отклонение: σ = 1,5 мин. Так как дисперсия любой случайной величины представляет собой квадрат ее стандартного отклонения, D_C = $=1{,}5^2$ = 2,25.

Найдем коэффициент вариации времени обслуживания всех заявок по формулам (8.36)–(8.39).

Вторые начальные моменты времен обработки деталей каждого вида:

$$\alpha_A = D_A + \overline{x}_A = 0{,}87 + 3{,}5^2 = 13{,}12;\ a_B = 25;\ \alpha_C = 51{,}25.$$

Второй начальный момент времени обслуживания всех заявок:

$$\alpha = 0{,}26 \cdot 13{,}12 + 0{,}53 \cdot 25 + 0{,}21 \cdot 51{,}25 = 27{,}42.$$

Дисперсия времени обслуживания всех заявок: D = 27,42 - $5{,}03^2$ =2,12.

Коэффициент вариации времени обслуживания всех заявок: $\varepsilon = \sqrt{2{,}12}/5{,}03$ = 0,29.

Таким образом, все параметры, необходимые для расчета характеристик станка, известны. Поток деталей, поступающих на обработку, является пуассоновским, время обслуживания распределено по некоторому произвольному закону, и имеется один канал обслуживания. Поэтому станок можно представить как СМО типа $M/G/1$. В этой СМО λ = 0,19 детали/мин, $\overline{x}$ = 5,03 мин, μ = 1/5,03 = 0,2 детали/мин, ν = 1, ε = 0,29.

Найдем характеристики СМО, как показано в подразделе 8.7: ρ = 0,95; P_0 = 0,05; $\overline{q}$ = 9,78 детали; $P_{отк}$ = 0; $P_{обсл}$ = 1; U = 0,95; $\overline{S}$ = 0,95 детали; $\overline{k}$ = 10,73 детали; γ = 0,19 детали/мин; $\overline{w}$ = 51,5 мин; $\overline{t}$ = 56,5 мин.

Из полученных характеристик видно, что станок явно перегружен: коэффициент загрузки U = 0,95. Это приводит к скоплению большого количества деталей, ожидающих обработки: в среднем ожидают обработки 9,78 детали. Среднее время пребывания деталей в очереди (т.е. время ожидания обработки) очень велико: оно составляет 51,5 мин, что более чем в 10 раз превышает время самой обработки. Для устранения этих недостатков можно рекомендовать установить еще один станок для ремонта дефектных деталей.

8.12. СМО с приоритетами

В таких СМО каждой заявке назначается некоторый приоритет. Если в очереди находятся заявки с разными приоритетами, то первыми на обслуживание поступают заявки с более высоким приоритетом.

Для такой дисциплины обслуживания заявок точный расчет характеристик возможен только при следующих условиях: поток заявок является пуассоновским; СМО является одноканальной; нет ограничений на очередь. Другими словами, тип СМО - $M/M/1$ или $M/G/1$ без ограничений на очередь.

Приоритеты заявок могут быть относительными или абсолютными.

В СМО с относительными приоритетами заявка, поступившая на обслуживание (т.е. в канал), всегда обслуживается до конца, даже если в это время поступает заявка с более высоким приоритетом.

В СМО с абсолютными приоритетами обслуживание заявки прерывается, если поступает заявка с более высоким приоритетом. Заявка, обслуживание которой было прервано, возвращается в очередь и поступает на дообслуживание только тогда, когда в очереди не останется ни одной заявки с более высоким приоритетом.

Пусть в СМО имеется R значений (уровней) приоритета. Будем обозначать номером 1 высший приоритет, а номером R - низший. Будем обозначать характеристики СМО для заявок с i-м уровнем приоритета индексом, обозначающим приоритет (например, $\overline{t_1}$ — среднее время пребывания в СМО заявок с первым уровнем приоритета). Средние характеристики СМО для заявок всех уровней приоритета будем указывать без индексов (например, $\overline{t}$ — среднее время пребывания в СМО всех заявок).

В расчетах характеристик СМО с приоритетами используются величины нагрузки на СМО, создаваемой заявками каждого уровня приоритета:

$$\rho_i = \frac{\lambda_i}{\mu_i}, \qquad (8.40)$$

где λ_i - интенсивность потока заявок с i-м уровнем приоритета;

i - интенсивность обслуживания заявок с i-м уровнем приоритета, определяемая как $\mu_i = 1/\overline{x_i}$, где $\overline{x_i}$ — среднее время обслуживания заявок с i-м уровнем приоритета.

Нагрузка на СМО, создаваемая всеми заявками, определяется следующим образом:

$$\rho = \sum_{i=1}^{R} \rho_i. \qquad (8.41)$$

Расчет характеристик СМО с приоритетами во многих случаях удобно начинать с вычисления среднего времени пребывания в очереди для заявок с различными уровнями приоритета. Для СМО с относительными приоритетами эти величины вычисляются следующим образом: для заявок с высшим приоритетом (с приоритетом 1):

$$\overline{w_i} = \frac{\sum_{j=1}^{R} \rho_j \overline{x}_j (1+\varepsilon_j^2)}{2(1-\rho_1)}, \qquad (8.42)$$

где ε_j - коэффициент вариации времени обслуживания заявок с j-м уровнем приоритета;

• для заявок с приоритетами 2,3,…, R:

$$\overline{w_i} = \frac{\sum_{j=1}^{R} \rho_j \overline{x}_j (1+\varepsilon_j^2)}{2(1-\sum_{j=1}^{i-1}\rho_j)(1-\sum_{j=1}^{i}\rho_j)}, i = 2,...,R. \quad (8.43)$$

Для СМО с абсолютными приоритетами средние времена пребывания заявок в очереди рассчитываются по следующим формулам:

• для заявок с высшим приоритетом (с приоритетом 1):

$$\overline{w_1} = \frac{\rho_1 \overline{x}_1 (1+\varepsilon_1^2)}{2(1-\rho_1)}; \qquad (8.44)$$

• для заявок с приоритетами 2,3,…, R:

$$\overline{w}_i = \frac{\overline{x}_i \sum_{j=1}^{i-1} \rho_j}{1-\sum_{j=1}^{i-1} \rho_j} + \frac{\sum_{j=1}^{i} \rho_j \overline{x}_j (1+\varepsilon_j^2)}{2(1-\sum_{j=1}^{i-1} \rho_j)(1-\sum_{j=1}^{i} \rho_j)}, i = 2,..., R. \quad (8.45)$$

Другие характеристики СМО определяются по следующим формулам (для СМО как с относительными, так и с абсолютными приоритетами).

Среднее время пребывания заявки в очереди: R

$$\overline{w} = \frac{\sum_{i=1}^{R} \lambda_i \overline{w}_i}{\lambda}, \quad (8.46)$$

или

$$\overline{w} = \sum_{i=1}^{R} P_j \overline{w}_i, \quad (8.47)$$

где λ - интенсивность потока всех заявок в СМО: $\lambda = \lambda_1 + \lambda_2 + ... + \lambda_R$. P_i - доля заявок с i-м уровнем приоритета в потоке заявок, поступающих в СМО, $P_1 + P_2 + ... + P_R = 1$.

Среднее время пребывания заявки в СМО:

$$\overline{t}_i = \overline{w}_i + \overline{x}_i, i = 1,..., R, \quad (8.48)$$

$$\overline{t} = \frac{\sum_{i=1}^{R} \lambda_i \overline{t}_i}{\lambda}, \quad (8.49)$$

или

$$\overline{t} = \sum_{i=1}^{R} P_i \overline{t}_i. \quad (8.50)$$

Среднее число заявок в СМО:

$$\overline{k}_i = \lambda_i \overline{t}_i, i = 1,..., R, \quad (8.51)$$

$$\overline{k} = \sum_{i=1}^{R} \overline{k}_i. \quad (8.52)$$

Среднее число заявок на обслуживании (среднее число занятых каналов):

$$\overline{S}_i = \rho_i, i = 1,..., R, \quad (8.53)$$

$$\overline{S} = \rho. \quad (8.54)$$

Среднее число заявок в очереди:

$$\overline{q}_i = \overline{k}_i - \overline{S}_i, i = 1,..., R, \quad (8.55)$$

$$\overline{q} = \sum_{i=1}^{R} \overline{q}_i. \quad (8.56)$$

Пропускная способность СМО:

$$\gamma_i = \lambda_i, i = 1,..., R, \quad (8.57)$$

$$\gamma = \lambda. \quad (8.58)$$

Вероятность простоя СМО:

$$P_0 = 1 - \rho. \quad (8.59)$$

Коэффициент загрузки СМО:

$$U_i = \rho_i, i = 1,..., R, \quad (8.60)$$

$$U = \rho. \quad (8.61)$$

Пример 8.7. В автоматизированной системе управления технологическим процессом (АСУТП) обрабатываются сигналы трех типов (сигналы А,В,С), поступающие от производственного оборудования. Сигналы типа А поступают в среднем через каждые 2 секунды. Сигналы типа В поступают со средней интенсивностью 3 сигнала в секунду, сигналы типа С - 6 сигналов в секунду. Обработка одного сигнала типа А занимает в среднем 20 мс, сигнала типа В - 50 мс, сигнала типа С - 100 мс. Интервалы времени между сигналами и время обработки сигналов можно считать случайными величинами, распределенными по экспоненциальному закону.

Предлагаются три варианта дисциплины обслуживания сигналов: а) в порядке поступления (дисциплина FIFO); б) с относительными приоритетами; в) с абсолютными приоритетами. При обслуживании с приоритетами более высокий приоритет должны иметь сигналы, требующие меньшего времени обработки (поэтому высший приоритет будут иметь сигналы А, менее высокий - сигналы В, самый низкий - сигналы С).

Требуется выбрать дисциплину обслуживания, обеспечивающую минимальное среднее время обработки всех сигналов.

Найдем некоторые величины, которые потребуются в дальнейших расчетах. Интенсивности потоков сигналов каждого типа известны: λ_A = 0,5 сигнала/с, λ_B = 3 сигнала/с, λ_C = 6 сигналов/с. Вычислим интенсивность потока всех сигналов: X = λA + λB + λC = 9,5 сигнала/с. Найдем доли сигналов каждого типа в общем потоке сигналов: P_A = 0,5/9,5 = 0,05; P_B = 3/9,5 = 0,32; P_C = 6/9,5 = 0,63.

Выполним расчет характеристик АСУТП для различных дисциплин обслуживания.

Дисциплина обслуживания FIFO. Так как потоки сигналов каждого типа являются пуассоновскими, поток всех сигналов также можно считать пуассоновским. Хотя время обслуживания сигналов каждого типа представляет собой случайную величину, распределенную по экспоненциальному закону, время обслуживания всех сигналов нельзя считать экспоненциальной случайной величиной, так как среднее время обработки сигналов разных типов различно. Поэтому АСУТП представляет собой СМО типа $M/G/1$ с заявками разных типов (сигналы А, В, С), для которых требуется разное время обслуживания: $\bar{x}_1$ = 0,02 с, $\bar{x}_2$ = 0,05 с, $\bar{x}_3$ = 0,1 с. Расчет характеристик такой СМО показан в подразделе 8.11.
Найдем среднее время обработки всех сигналов по формуле (8.35): $\bar{x}$ = 0,05·0,02 + 0,32·0,05 + 0,63·0,1 = 0,08 с.

Найдем коэффициент вариации времени обслуживания всех заявок, как показано в подразделе 8.11. Время обслуживания сигналов каждого типа - случайная величина, распределенная по экспоненциальному закону. Из теории вероятностей известно, что для таких случайных величин дисперсия определяется по формуле: $D = \bar{x}^2$, где $\bar{x}$ — математическое ожидание (среднее значение) случайной величины.

Таким образом, D_A = 0,022 = 0,0004, D_B = 0,052 = 0,0025, D_C = 0,12 = 0,01. Выполнив расчеты по формулам (8.36)-(8.39), вычислим коэффициент вариации времени обслуживания всех заявок: ε = 1,107.
Дальнейший расчет выполняется согласно подразделу 8.7, т.е. для СМО типа $M/G/1$ без ограничений на очередь, где λ = 9,5 сигнала/с, $\bar{x}$ = 0,08 с, $\mu = 1/\bar{x}$ = 12,5 сигнала/с, ν = 1, ε = 1,107. Результаты приведены в табл. 8.4.

Обслуживание с относительными приоритетами. При такой дисциплине обслуживания сигналы А, В, С представляют собой заявки с первым, вторым и третьим уровнем приоритета соответственно.

По формуле (8.40) найдем нагрузку на СМО, создаваемую сигналами каждого
типа: ρ_A = 0,01; ρ_B = 0,15; ρ_C = 0,6. По формуле (8.41) вычислим общую нагрузку на СМО: ρ = 0,76.

Найдем среднее время пребывания в очереди для сигналов каждого типа. Для сигналов типа А это время вычисляется по формуле (8.42), для других сигналов - по формуле (8.43):

$$\bar{w}_A = \frac{0,01 \cdot 0,02 \cdot (1+1^2) + 0,15 \cdot 0,05 \cdot (1+1^2) + 0,6 \cdot 0,01 \cdot (1+1_2)}{2 \cdot (1-0,01)} = 0,0684\,\text{с};$$
$$\bar{w}_B = \frac{0,01 \cdot 0,02 \cdot (1+1^2) + 0,15 \cdot 0,05 \cdot (1+1^2) + 0,6 \cdot 0,01 \cdot (1+1_2)}{2 \cdot (1-0,01) \cdot (1-0,01-0,15)} = 0,0814\,\text{с};$$
$$\bar{w}_C = \frac{0,01 \cdot 0,02 \cdot (1+1^2) + 0,15 \cdot 0,05 \cdot (1+1^2) + 0,6 \cdot 0,01 \cdot (1+1_2)}{2 \cdot (1-0,01-0,15) \cdot (1-0,01-0,15-0,6)} = 0,3358\,\text{с}.$$

Примечание. В расчетах по формулам (8.42) и (8.43) использовались значения коэффициентов вариации $\varepsilon_j = 1,\ j = 1,...,3$, так как время обработки сигналов всех типов представляет собой случайную величину, распределенную по экспоненциальному закону.

Полученные величины означают, что для сигналов типа А среднее время пребывания в очереди (т.е. время от поступления сигнала в АСУТП до начала его обработки) составляет 0,0684 с, для сигналов типа В - 0,0814 с, для сигналов типа С -0,3358 с. Как и следовало ожидать, чем выше приоритет сигнала, тем меньше время его пребывания в очереди.

Остальные характеристики работы АСУТП найдем по формулам (8.46)-(8.61). Полученные характеристики для каждого типа сигналов, а также средние характеристики для сигналов всех типов приведены в табл. 8.4.

Обслуживание с абсолютными приоритетами. Найдем среднее время пребывания в очереди для сигналов каждого типа. Для сигналов типа А это время вычисляется по формуле (8.44), для других сигналов - по формуле (8.45):

$$\bar{w}_A = \frac{0,01 \cdot 0,02 \cdot (1+1^2)}{2 \cdot (1-0,01)} = 0,0002\,\text{с};$$
$$\bar{w}_B = \frac{0,05 \cdot 0,01}{1-0,01} + \frac{0,01 \cdot 0,02 \cdot (1+1^2) + 0,15 \cdot 0,05 \cdot (1+1^2)}{2 \cdot (1-0,01) \cdot (1-0,01-0,15)} = 0,0098\,\text{с};$$
$$\bar{w}_C = \frac{0,1 \cdot (0,01+0,15)}{1-(0,01+0,15)} + \frac{0,01 \cdot 0,02 \cdot (1+1^2) + 0,15 \cdot 0,05 \cdot (1+1^2) + 0,6 \cdot 0,1 \cdot (1+1^2)}{2 \cdot (1-0,01-0,15) \cdot (1-0,01-0,15-0,6)} = 0,3549\,\text{с}.$$

Для сигналов, имеющих высший приоритет (сигналов типа А), время пребывания в очереди оказалось очень малым, так как для таких сигналов ожидание в очереди требуется только в том случае, если в момент поступления такого сигнала в АСУТП обрабатывается сигнал такого же типа.

Остальные характеристики работы АСУТП найдем по формулам (8.46)-(8.61).

Из полученных результатов видно, что среднее время пребывания заявки в СМО $\bar{t}$ (т.е. среднее время от поступления сигнала в АСУТП до окончания его обработки) для дисциплины обслуживания FIFO составляет 0,362 с, для обслуживания с относительными приоритетами - 0,321 с, с абсолютными приоритетами - 0,307 с. Таким образом, для того, чтобы среднее время обработки сигналов в АСУТП было минимальным, следует использовать дисциплину обслуживания с абсолютными приоритетами.

8.13. Многофазные СМО. Сети СМО

Многофазные СМО и сети СМО состоят из нескольких типовых узлов (см. рис. 8.1), т.е. представляют собой совокупность нескольких СМО.

Многофазные СМО состоят из нескольких типовых узлов, расположенных последовательно. Все заявки, обслуженные в одном узле, направляются в следующий узел. Другими словами, выходной поток одного узла многофазной СМО является входным потоком для следующего.

Сеть СМО также состоит из нескольких типовых узлов СМО. Однако порядок прохождения узлов может быть различным для разных заявок. Для части заявок может требоваться обслуживание во всех узлах, а для других - только в некоторых из узлов.

Точный расчет характеристик таких СМО возможен только в случае, если все потоки заявок являются пуассоновскими, а все времена обслуживания - экспоненциальными случайными величинами. В других случаях возможен лишь приближенный расчет характеристик СМО.

При расчете характеристик многофазных СМО и сетей СМО необходимо учитывать следующее:

- если на вход СМО поступает несколько потоков заявок, то интенсивность полного потока заявок в этой СМО равна сумме интенсивностей отдельных потоков;
- если на вход СМО поступает часть заявок из некоторого потока, интенсивность которого равна λ, то интенсивность входного потока заявок в СМО можно определить по формуле $\lambda_{вх} = P\lambda$, где P — вероятность попадания заявки во входной поток;
- интенсивность выходного потока в СМО (т.е. потока обслуженных заявок) равна интенсивности входного потока.

Пример 8.8. На производственный участок У3 поступают для обработки детали с двух других участков (У 1 и У2).

Детали с участка У1 поступают в среднем через каждые 10 мин, с участка У2 -через каждые 15 мин. Все детали проходят обработку на станке СТ1. Обработка детали на этом станке занимает от 3 до 7 мин. Со станка СТ1 детали направляются на группу из трех одинаковых станков СТ2. Перед этими станками находится накопитель для деталей, ожидающих обработки (общий для всех трех станков). Деталь направляется на любой свободный станок СТ2. Обработка детали на станке СТ2 занимает в среднем 20 мин. Накопитель, расположенный перед станками СТ2, вмещает 8 деталей. При его заполнении детали направляются для обработки на станок СТ3. Обработка на этом станке занимает в среднем 25 мин.

При обработке деталей возможны дефекты. На станках СТ2 дефект допускается примерно в 8% случаев, на станке СТ3 - в 15% случаев. Все дефектные детали направляются на станок СТ4 для устранения дефекта; это занимает в среднем 10 мин.

Требуется найти характеристики работы всех станков, а также определить среднее время пребывания детали на участке У3.

Характеристики станков приведены в табл. 8.5. Их расчет рассматривается ниже.

Расчет характеристик станка СТ1. Чтобы выполнить этот расчет, будем считать потоки деталей из цехов У1 и У2 пуассоновскими. Тогда станок СТ1 можно рассматривать как одноканальную СМО, где поток заявок является пуассоновским, а время обслуживания распределено по равномерному закону, т.е. СМО типа $M/G/1$.

Найдем интенсивности потоков деталей с участков У1 и У2: λ_1 = 0,1 детали/мин, λ_2 = 0,07 детали/мин. Интенсивность входного потока для станка СТ1 равна сумме интенсивностей этих потоков: λ_{CT1} = 0,17 детали/мин. Среднее время обработки детали $\bar{x}_{CT1}$ = (3 + 7)/2 = 5 мин, интенсивность обработки деталей μ_{CT1} = 1/5 = 0,2 детали/мин. Из табл. 8.1 найдем коэффициенты вариации, необходимые для расчета средней длины очереди: ν = 1 (так как поток заявок - пуассоновский), ε = 0,23 (определяется по формуле для равномерного закона распределения). Дальнейший расчет выполняется, как показано в подразделе 8.7.

Расчет характеристик станков СТ2. Будем считать поток деталей, выходящих со станка СТ1, пуассоновским (хотя это не вполне точно, так как время обработки на станке СТ1 - не экспоненциальная, а равномерная случайная величина). Будем также считать, что время обработки деталей на станках СТ2 представляет собой экспоненциальную случайную величину. Тогда группу станков СТ2 можно рассматривать как трехканальную марковскую СМО ($M/M/3$) с ограничением на длину очереди (так как накопитель вмещает только 8 деталей, и при его заполнении детали направляются на другой станок). Здесь λ_{CT2} = 0,17 детали/мин (так как на станки СТ2 поступают все детали со станков СТ1), $\bar{x}_{CT2}$ = 20 мин, μ_{CT2} = 0,05 детали/мин, n= 8. Расчет характеристик станка СТ2 выполняется согласно подразделу 8.9.

Расчет характеристик станка СТ3. На этот станок поступают детали, не попавшие на станки СТ2 из-за заполнения накопителя (т.е. детали, получившие отказ в обслуживании). Поэтому интенсивность входного потока для станка СТ3 можно найти следующим образом: $\lambda_{CT3} = P_{отк}\lambda_{CT3}$ = 0,16·0,17 = 0,027 детали/мин (здесь $P_{отк}$ = 0,16 - вероятность отказа в обслуживании для группы станков СТ2; см. табл. 8.5). Будем считать поток деталей на станок СТ3 пуассоновским, а время обработки на этом станке - экспоненциальным. Тогда станок СТ3 можно рассматривать как одно-канальную марковскую СМО ($M/M/1$) без ограничений на очередь, где λ_{CT3} = 0,027 детали/мин, $\bar{x}_{CT3}$ = 25 мин, μ_{CT3} = 0,04 детали/мин.

Расчет характеристик станка СТ3 выполняется согласно подразделу 8.7.

Расчет характеристик станка СТ4. На этот станок поступает 8% деталей, обработанных на станках СТ2, и 15% деталей, обработанных на СТ3. Поэтому интенсивность входного потока для станка СТ3 можно найти следующим образом: $\lambda_{CT4} = 0,08 \cdot \gamma_{CT2} + 0,15 \cdot \gamma_{CT3}$= 0,08·0,14 + 0,15·0,027 = 0,015 детали/мин (здесь γ_{CT2}, γ_{CT3} – значения пропускной способности станков СТ2 и СТ3; см. табл. 8.5). Будем считать поток деталей на станок СТ4 пуассоновским, а время обработки на этом станке - экспоненциальным. Тогда станок СТ4 можно рассматривать как одноканальную марковскую СМО ($M/M/1$) без ограничений на очередь, где λ_{CT4} = 0,015 детали/мин, $\bar{x}_{CT4}$ = 10 мин, μ_{CT4} = 0,1 детали/мин. Расчет характеристик станка СТ4 выполняется согласно подразделу 8.7.

Таблица 8.5

Характеристики	СТ1	СТ2	СТ3	СТ4
ρ	0,85	1,13	0,68	0,15
P_0	0,15	0,009	0,32	0,85
$\bar{q}$, деталей	2,54	4,38	1,4	0,027
$P_{отк}$	0	0,16	0	0
$P_{обсл}$	1	0,84	1	1
U	0,85	0,95	0,68	0,15
$\bar{S}$, деталей	0,85	2,86	0,68	0,15
$\bar{k}$, деталей	3,39	7,24	2,08	0,177
γ, деталей/мин	0,17	0,14	0,027	0,015
$\bar{w}$, мин	14,9	31,29	51,9	1,76
$\bar{t}$, мин	19,9	51,29	76,9	11,8

Расчет среднего времени пребывания детали на участке У3. Все детали проходят обработку на станке СТ1. Среднее время пребывания детали на этом станке (включая время ожидания обработки и само время обработки) составляет 19,9 мин. Затем 84% деталей проходят обработку на одном из станков СТ2; среднее время пребывания детали на этих станках составляет 51,29 мин. Остальные 16% деталей обрабатываются на станке СТ3;

среднее время пребывания детали на этом станке - 76,9 мин. Кроме того, для 8% деталей, обрабатывавшихся на станках СТ2, и 15% деталей, обрабатывавшихся на станке СТ3, требуется устранение дефекта на станке СТ4. Это занимает в среднем 11,8 мин. Таким образом, среднее время пребывания детали на участке можно найти так:

$\overline{t}_{У3}$ = 19,9 + 0,84·51,29 + 0,16·76,9 + (0,84·0,08 + 0,16·0,15)·11,8 = 76,36 мин.

По результатам анализа характеристик станков можно выявить следующие недостатки в работе участка и предложить способы их устранения:

- перегрузка группы станков СТ2 и недостаточная загрузка станка СТ3. Для устранения этого недостатка можно предложить уменьшить размер накопителя перед станками СТ2. В таком случае большее количество деталей будет направляться на станок СТ3. В результате загрузка станков СТ2 снизится, а станка СТ3 - повысится;
- явная недогрузка станка СТ4: станок простаивает 85% рабочего времени. Для устранения этого недостатка можно предложить использовать станок СТ4 не только для устранения дефектов, но и для каких-либо других работ. Другой вариант - вообще отказаться от станка СТ4 и выполнять устранение дефектов на тех станках, где изделие изготавливается, т.е. на станках СТ2 и СТ3 (если это возможно).

8.14. Замкнутые СМО

Замкнутые СМО - это СМО с фиксированным количеством заявок, периодически требующих обслуживания.

Будем обозначать количество заявок в замкнутой СМО как N, а среднее время между окончанием обслуживания заявки и ее следующим обращением за обслуживанием (время между обращениями) - как T. Как и для других СМО, количество каналов будем обозначать как m, а среднее время обслуживания заявки - как $\overline{x}$.

Точный расчет характеристик замкнутых СМО возможен только в случае, если и время обслуживания заявки, и время между обращениями представляют собой случайные величины, распределенные по экспоненциальному закону.

Характеристики эффективности работы замкнутых СМО в основном те же, что и для разомкнутых СМО. Однако расчет характеристик для замкнутых и разомкнутых СМО существенно различается.

Для расчета характеристик замкнутых СМО применяются следующие формулы.

Вероятность простоя:

$$P_0\left[1+\sum_{j=1}^{m}\frac{N!(\overline{x}/T)^j}{(N-j)!\cdot j!}+\sum_{j=m+1}^{N}\frac{N!(\overline{x}/T)^j}{(N-j)!\cdot m!\cdot m^{j-m}}\right]^{-1}. \quad (8.62)$$

Среднее число заявок в очереди:

$$\overline{q}=P_0\sum_{j=m+1}^{N}(j-m)\frac{N!(\overline{x}/T)^j}{(N-j)!\cdot m!\cdot m^{j-m}}. \quad (8.63)$$

Среднее число заявок на обслуживании (среднее число занятых каналов):

$$\overline{S}=\begin{cases}1-P_0, m=1, & (8.64)\\ m-mP_0\sum_{j=1}^{m-1}(m-j)\frac{N!(\overline{x}/T)^j}{(N-j)!\cdot j!}, m>1. & (8.65)\end{cases}$$

Среднее число заявок в СМО:

$$\overline{k}=\overline{q}+\overline{S}. \quad (8.66)$$

Среднее время пребывания заявки в СМО:

$$\overline{t}=\frac{\overline{k}T}{N-\overline{k}}. \quad (8.67)$$

Среднее время пребывания заявки в очереди:

$$\bar{w} = \bar{t} - \bar{x}. \qquad (8.68)$$

Коэффициент загрузки:

$$U = \frac{\bar{S}}{m}. \qquad (8.69)$$

Пропускная способность:

$$\gamma = \frac{\bar{S}}{\bar{x}}. \qquad (8.70)$$

Вероятности пребывания в СМО j заявок:

$$P_j = \begin{cases} \dfrac{N!(\bar{x}/T)^j}{(N-j)!\cdot j!} P_0,\ j = 1,\dots,m, & (8.71) \\ \dfrac{N!(\bar{x}/T)^j}{(N-j)!\cdot m!\cdot m^{j-m}} P_0,\ j = m+1,\dots,N. & (8.72) \end{cases}$$

Примечание. Формула (8.71) позволяет найти вероятности состояний СМО, при которых очередь отсутствует (количество заявок, обслуживаемых в СМО, не превышает количества каналов), а формула (8.72) - вероятности состояний при наличии очереди.

Пример 8.9. В цехе имеется 10 установок для производства пластмассы. Заправка установки сырьем занимает в среднем 20 мин. После заправки установка работает в среднем 2 часа; затем требуется новая заправка. Производительность установки - 12 кг пластмассы в час. Прибыль предприятия от продажи одного килограмма пластмассы составляет 3 ден. ед. Оператор, обслуживающий установки, получает 70 ден. ед. за рабочую смену (8 часов). Требуется определить, сколько операторов должно обслуживать цех, чтобы прибыль предприятия от его работы была максимальной.

Операторы, выполняющие заправку установок, могут рассматриваться как замкнутая СМО. Заявками в такой СМО являются установки, для которых периодически требуется заправка. Здесь T = 120 мин, $\bar{x}$ = 20 мин, N = 10. Будем считать, что время работы установки (время между заправками) и время ее заправки - экспоненциальные случайные величины. Тогда операторы могут рассматриваться как замкнутая СМО типа $M/M/m$. Величину m (количество операторов) требуется выбрать по результатам решения задачи.

Пусть в цехе работает только один оператор (m = 1). Найдем характеристики работы оператора по формулам (8.62)-(8.70): P_0 = 0,043; $\bar{q}$ = 3,302 установки; $\bar{S}$ = 0,957 установки; $\bar{k}$ = 4,259 установки; $\bar{t}$ = 89,017 мин; $\bar{w}$ =69,017 мин; U = 0,957; γ = 0,048 установки/мин.

Например, вероятность простоя определяется по формуле (8.62) следующим образом:

$$P_0 = \left[1 + \frac{10!(20/120)^1}{9!\cdot 1!} + \frac{10!(20/120)^2}{8!\cdot 1!\cdot 1^0} + \frac{10!(20/120)^3}{7!\cdot 1!\cdot 1^1} + \frac{10!(20/120)^4}{6!\cdot 1!\cdot 1^2} + \frac{10!(20/120)^5}{5!\cdot 1!\cdot 1^3} + \right.$$

$$\left. + \frac{10!(20/120)^6}{4!\cdot 1!\cdot 1^4} + \frac{10!(20/120)^7}{3!\cdot 1!\cdot 1^5} + \frac{10!(20/120)^8}{2!\cdot 1!\cdot 1^6} + \frac{10!(20/120)^9}{1!\cdot 1!\cdot 1^7} + \frac{10!(20/120)^{10}}{0!\cdot 1!\cdot 1^8} \right] = 0,043.$$

Вычислим, какую долю рабочего времени составляют простои установок, связанные с заправкой и ожиданием заправки. Величина t = 89,017 мин означает, что простой установки (включающий ожидание заправки и саму заправку) занимает в среднем 89,017 мин. После заправки установка работает в среднем 120 мин, затем снова простаивает в течение 89,017 мин, и т.д. Таким образом, простои составляют 89,017/(120 + 89,017) = 0,426, или 42,6% рабочего времени. Установка работает только 100 - 42,6 = 57,4% рабочего времени.

Определим прибыль от работы цеха за рабочую смену (8 часов). В цехе имеется 10 установок. Каждая установка выпускает 12 кг пластмассы в час (когда установка ожидает заправки или заправляется, она не выпускает пластмассу). Каждый килограмм пластмассы приносит предприятию прибыль в размере 3 ден. ед. Из этой прибыли необходимо вычесть заработную плату оператора (70 ден. ед.). Таким образом, прибыль от работы цеха за 8 часов можно найти так: $10 \cdot 8 \cdot (1 - 0{,}426) \cdot 12 \cdot 3 - 70 = 1583$ ден. ед.

Видно, что данный вариант организации работы цеха имеет серьезные недостатки, основной из которых - значительные простои установок в ожидании заправки, что приводит к потерям прибыли. Еще один недостаток - явная перегрузка оператора (коэффициент загрузки составляет 0,957).

Выполним аналогичные расчеты для случаев, когда в цехе работают 2, 3 или 4 оператора (m = 2,3 или 4). Результаты приведены в табл. 8.6.

Таблица 8.6

m	1	2	3	4
P_0	0,043	0,178	0,209	0,213
$\bar{q}$, установок	3,302	0,564	0,092	0,013
$\bar{S}$, установок	0,957	1,348	1,415	1,427
$\bar{k}$, установок	4,259	1,912	1,508	1,44
$\bar{t}$, мин	89,017	28,36	21,305	20,187
$\bar{w}$, мин	69,017	8,361	1,305	0,187
U	0,957	0,674	0,472	0,357
γ, установок/мин	0,048	0,067	0,071	0,071
Простой установок, %	42,6	19,1	15,1	14,4
Прибыль, ден. ед./смену	1583	2190	2235	2185

Из табл. 8.6 видно, что максимальная прибыль от работы цеха (2235 ден. ед. за смену) достигается, если обслуживание установок выполняется тремя операторами (m = 3). При большем количестве операторов прибыль снижается, так как затраты, связанные с заработной платой операторов, превышают выигрыши от сокращения простоев установок. Таким образом, можно рекомендовать, чтобы установки для производства пластмассы обслуживали три оператора. Однако следует отметить, что этот вариант имеет и недостаток - низкую загрузку операторов (U = 0,472).

9. Оптимизация принятия решений в условиях риска, неадекватностей и неопределенностей

9.1. Понятия риска и неопределенности. Постановка задачи

Во многих случаях результат принятия решения зависит не только от самого решения, но и от некоторых внешних условий. Под внешними условиями понимаются любые факторы, на которые невозможно влиять (или возможность такого влияния ограничена): спрос на продукцию, действия конкурентов, природно-климатические факторы и т.д. Так как заранее точно неизвестны условия реализации решения, не могут быть заранее известны и его результаты: прибыль, затраты, сроки реализации решения и т.д.

Под неопределенностью понимается неполнота информации о внешних условиях, влияющих на результат принимаемого решения. Под риском понимается возможность каких-либо неблагоприятных последствий принятого решения: потери ресурсов, недополучения прибыли, возникновения дополнительных расходов, несвоевременного выполнения работ и т.д.

Задачи, связанные с принятием решений в условиях риска, возникают, например, при планировании производства. Результат принятого решения (например, прибыль от выпуска продукции) зависит не только от действий предприятия (т.е. от вида выпускаемой продукции, объема производства, качества продукции и т.д.), но и от внешних факторов (например, от спроса на продукцию, от наличия на рынке аналогичных видов продукции и т.д.). Очевидно, что внешние условия не могут быть точно известны заранее, и предприятие не может существенно влиять на них.

Имеется большое количество разнообразных формулировок задач, решаемых в условиях риска и неопределенности, и методов их решения.

Многие задачи, решаемые в условиях риска и неопределенности, могут быть сформулированы следующим образом. Требуется выбрать одно из M возможных решений (альтернатив): $A_1, A_2, ..., A_M$. Известно, что каждое из решений может быть реализовано в одном из N вариантов внешних условий: $B_1, B_2, ..., B_N$. Для каждого из решений известны его последствия (выигрыши стороны, принимающей решение) в каждом из вариантов внешних условий: $E_{ij}, i = 1, ..., M, j = 1, ..., N$. Эти выигрыши можно свести в таблицу, называемую матрицей выигрышей (или платежной матрицей). Такая матрица представляет собой математическую модель задачи. Общий вид матрицы выигрышей показан в табл. 9.1.

Метод построения матрицы выигрышей полностью зависит от конкретных условий задачи.

Требуется выбрать наиболее эффективный вариант решения, т.е. одно из решений $A_1, A_2, ..., A_M$.

Таблица 9.1

	B_1	B_2	...	B_N
A_1	E_{11}	E_{12}	...	E_{1N}
A_2	E_{21}	E_{22}	...	E_{2N}
…	…	…	…	…
A_M	E_{M1}	E_{M2}	…	E_{MN}

Примечания:

1. В матрице выигрышей могут быть отрицательные элементы, соответствующие убыткам.

2. В некоторых случаях вместо матрицы выигрышей используется матрица затрат. В этом случае элемент E_{ij} — это затраты, связанные с i-м решением в j-м варианте внешних условий.

Пример 9.1. Крупная фирма предполагает приобрести пакет акций одного из трех предприятий (П1, П2, П3). Прибыль, которую получит фирма от покупки акций, не может быть точно известна заранее, так как она зависит от того, как будет изменяться стоимость этих акций. Возможные величины прибыли фирмы от покупки акций предприятий (в млн ден. ед.) приведены в табл. 9.2.

Таблица 9.2

Пакет акций	Изменение стоимости акций		
	Рост	Стабильное состояние	Снижение
П1	10	6	-7
П2	6	4	-3
П3	8	3	-2

Величины в таблице обозначают следующее: если, например, фирма приобретет пакет акций предприятия П1, и их стоимость будет расти, то прибыль фирмы составит 10 млн ден. ед. Если стоимость акций предприятия П1 будет оставаться стабильной, то прибыль фирмы составит 6 млн ден. ед. В случае снижения стоимости акций фирма понесет убыток в размере 7 млн ден. ед.

Согласно имеющимся экспертным оценкам, возможны четыре сценария развития экономической ситуации (С1, С2, С3, С4):

- сценарий С1: стоимость акций предприятий П1 и П2 остается стабильной, стоимость акций предприятия П3 растет;
- сценарий С2: стоимость акций П1 снижается, П2 и П3 - растет;
- сценарий С3: стоимость акций П1 растет, П2 и П3 - снижается;
- сценарий С4: стоимость акций всех предприятий остается стабильной. Требуется определить, какой пакет акций следует приобрести фирме, чтобы получить максимальную прибыль.

Данная задача решается в условиях риска и неопределенности, так как прибыль фирмы зависит не только от ее решения (т.е. от того, какой пакет акций она купит), но и от внешних условий (от сценария развития экономической ситуации).

При решении этой задачи важно понимать, что при покупке акций еще неизвестно, по какому сценарию будет развиваться экономическая ситуация. Влиять на этот сценарий невозможно. Следует также обратить внимание, что фирма может приобрести пакет акций только одного предприятия.

Составим матрицу выигрышей. Для этого найдем, какой будет прибыль фирмы для различных решений (т.е. при покупке различных пакетов акций) в разных внешних условиях. Предположим, что фирма купит пакет акций предприятия П1. Если экономическая ситуация будет развиваться по сценарию С1, то фирма получит прибыль в размере 6 млн ден. ед., так как при таком сценарии стоимость акций предприятия Ш будет оставаться стабильной. Если экономическая ситуация будет развиваться по сценарию С2, то фирма понесет убыток в размере 7 млн ден. ед., так как при таком сценарии стоимость акций П1 снизится. При сценарии С3 фирма получит прибыль в размере 10 млн ден. ед., так как стоимость акций П1 возрастет. При сценарии С4 прибыль фирмы составит 6 млн ден. ед., так как стоимость акций П1 будет оставаться стабильной.

Выполнив аналогичные рассуждения для всех вариантов решения (покупка пакета акций П1, П2 или П3) и для всех вариантов внешних условий (сценарий С1, С2, С3 или С4), получим матрицу выигрышей (табл. 9.3).

Таблица 9.3

Приобретенный пакет акций	Сценарий			
	С1	С2	С3	С4
П1	6	-7	10	6
П2	4	6	-3	4
П3	8	8	-2	3

На основании этой матрицы требуется выбрать одно из решений, т.е. определить, какой пакет акций следует приобрести.

Примечание. Следует обратить внимание, что в данной задаче невозможно точно определить, какой будет прибыль фирмы. Например, если фирма купит пакет акций П1, то она может получить прибыль в размере 6 или 10 млн ден. ед., или понести убыток в размере 7 млн ден. ед. (в зависимости от того, по какому сценарию будет развиваться экономическая ситуация).

9.2. Методы выбора решений в условиях риска и неопределенности

Существует несколько методов (критериев) для выбора решений в условиях риска и неопределенности. Используемый метод зависит от имеющейся информации о внешних условиях, прежде всего - от того, имеется ли информация о вероятностях внешних условий.

Выбор решений при известных вероятностях внешних условий. Критерий Байеса

Если известны вероятности внешних условий, то для оценки и выбора решений применяется критерий Байеса. Он может использоваться в двух видах: как критерий максимума среднего выигрыша или как критерий минимума среднего риска.

Пусть известны вероятности вариантов внешних условий: $P_1, P_2, ..., P_N$.

Если решение выбирается по значениям выигрышей, то для каждого решения находится средняя оценка по всем вариантам внешних условий (средний выигрыш):

$$Z_i = \sum_{j=1}^{N} (E_{ij} P_j), i = 1, ..., M. \qquad (9.1)$$

где P_j — вероятности внешних условий.

Лучшим является решение с максимальной оценкой.

Решим задачу из примера 9.1, используя критерий Байеса, согласно имеющимся экспертным оценкам, наиболее вероятен сценарий С4: его вероятность составляет 50%. Менее вероятно изменение экономической ситуации по сценарию С1: вероятность этого сценария — 30%. Наименее вероятны сценарии С2 и С3: вероятность каждого из них — 10%. Найдем оценки решений для данной задачи по формуле (9.1):

Z_1 = 6·0,3 - 7·0,1 + 10·0,1 + 6·0,5 = 5,1 (оценка для пакета акций П1);

Z_2 = 4·0,3 + 6·0,1 - 3·0,1 + 4·0,5 = 3,5 (оценка для пакета акций П2);

Z_3 = 8·0,3 + 8·0,1 - 2·0,1 + 3·0,5 = 4,5 (оценка для пакета акций П3).

Таким образом, фирме рекомендуется приобрести пакет акций предприятия П1.

В некоторых случаях для выбора решения используется матрица рисков ($R_{ij}, i = 1,.., M, j = 1, ..., N$). Под риском понимается потерянный выигрыш: разность между выигрышем, максимально возможным для данного варианта внешних условий, и фактическим выигрышем. Для данной задачи матрица рисков приведена в табл. 9.4.

Таблица 9.4

Приобретенный пакет акций	Сценарий			
	С1	С2	С3	С4
П1	2	15	0	0
П2	4	2	13	2
П3	0	0	12	3

Здесь, например, для первого варианта внешних условий (сценарий С1) максимальная прибыль достигается при покупке пакета акций П3; эта прибыль составляет 8 млн ден. ед. При покупке пакета акций П1 прибыль будет меньше и составит только 6 млн ден. ед. Потерянный выигрыш (риск) определяется как 8-6 = 2 млн ден. ед. Аналогично находятся другие значения рисков.

Оценки решений по критерию минимума среднего риска находятся по следующей формуле:

$$Z_i = \sum_{j=1}^{N} (R_{ij} P_j), i = 1, ..., M. \qquad (9.2)$$

Лучшим является решение с минимальной оценкой. Оценки решений для данной задачи по формуле (9.2):

$$Z_1 = 2 \cdot 0,3 + 15 \cdot 0,1 + 0 \cdot 0,5 = 2,1;$$

$$Z_2 = 4 \cdot 0,3 + 2 \cdot 0,1 + 13 \cdot 0,1 + 2 \cdot 0,5 = 3,7;$$

$$Z_3 = 2,7.$$

Таким образом, фирме рекомендуется приобрести пакет акций предприятия П1.

Выбор решений при неизвестных вероятностях внешних условий

Если вероятности внешних условий неизвестны, то для оценки и выбора решений могут применяться следующие критерии.

Критерий Лапласа: применяется, если можно предполагать, что все варианты внешних условий одинаково вероятны. Для каждого решения находится средняя оценка по всем вариантам внешних условий (средний выигрыш):

$$Z_i = \frac{1}{N}\sum_{j=1}^{N} E_{ij} P_j, i = 1, ..., M. \qquad (9.3)$$

Лучшим является решение с максимальной оценкой.

Предположим, что для задачи из примера 9.1 вероятности всех четырех сценариев развития экономической ситуации примерно одинаковы. Найдем оценки по критерию Лапласа по формуле (9.3):

$$Z_1 = (6 - 7 + 10 + 6)/4 = 3,75;$$

$$Z_2 = (4 + 6 - 3 + 4)/3 = 2,75;$$

$$Z_3 = 4,25.$$

Таким образом, если есть основания предполагать, что все сценарии одинаково вероятны, то фирме следует приобрести пакет акций предприятия П3.

Критерий Вальда (критерий крайнего пессимизма, максиминный критерий): решение выбирается в расчете на наихудшие внешние условия. В качестве оценки каждого решения используется минимальный выигрыш, который можно получить при выборе этого решения:

$$Z_i = \min_j E_{ij}, i = 1, ..., M. \qquad (9.4)$$

Лучшим является решение с максимальной оценкой.

В примере 9.1 оценки решений по критерию Вальда следующие: Z_1 = min(6; -7; 10; 6) = -7; Z_2 = min(4; 6; -3; 4) = -3, Z_3 = -2. Другими словами, при покупке пакета акций Ш фирма в самом худшем случае понесет убыток в размере 7 млн ден. ед. (если экономическая ситуация будет развиваться по сценарию С2). При покупке пакета акций П2 убыток в самом худшем случае составит 3 млн ден. ед. (при сценарии С3), при покупке пакета акций П3 - 2 млн ден. ед. (при сценарии С3). Таким образом, фирме рекомендуется приобрести пакет акций предприятия П3.

Критерий Сэвиджа (критерий крайнего пессимизма, минимаксный критерий): решение принимается в расчете на наихудшие внешние условия (как и при использовании критерия Вальда), но для оценки решений используется матрица рисков. В качестве оценки используется максимальный риск (максимальный потерянный выигрыш), соответствующий данному решению:

$$Z_i = \max_j R_{ij}, i = 1, ..., M. \qquad (9.5)$$

Лучшим является решение с минимальной оценкой.

В примере 9.1 Z_1 = max(2; 15; 0; 0) = 15; Z_2 = max(4; 2; 13; 2) = 13; Z_3 = 12. Таким образом, рекомендуется приобрести пакет акций предприятия П3.

Критерий Гурвица: решение принимается с учетом того, что возможны как благоприятные, так и неблагоприятные внешние условия. При использовании этого критерия требуется указать «коэффициент пессимизма» — число в диапазоне от 0 до 1, представляющее собой субъективную (т.е. не рассчитанную, а указанную человеком) оценку возможности неблагоприятных внешних условий. Если есть основания предполагать, что внешние условия будут неблагоприятными, то коэффициент пессимизма назначается близким к единице. Если неблагоприятные внешние условия маловероятны, то используется коэффициент пессимизма, близкий к нулю. Оценки решений находятся по следующей формуле:

$$Z_i = a \cdot \min_j R_{ij} + (1 - a) \cdot \max_j R_{ij}, i = 1, ..., M, \qquad (9.6)$$

где a — коэффициент пессимизма.

Лучшим является решение с максимальной оценкой.

Предположим, что в задаче из примера 9.1 есть основания предполагать, что неблагоприятные условия (способствующие снижению стоимости акций) немного более вероятны, чем благоприятные. Для принятия решения по критерию Гурвица выберем коэффициент пессимизма a = 0,6.

Найдем оценки решений:

$$Z_1 = 0,6\cdot(-7)+0,4\cdot10=-0,2; Z_2 = 0,6\cdot(-3)+0,4\cdot6=0,6; Z_3 = 0,6\cdot(-2)+0,4\cdot8=2.$$

Таким образом, рекомендуется приобрести пакет акций предприятия П3.

С учетом всех использованных критериев, лучшим решением является покупка пакета акций предприятия П3. Это решение оказалось лучшим по всем критериям (по критериям Лапласа, Вальда, Сэвиджа и Гурвица).

Примечание. Задачи, связанные с принятием решений при известных вероятностях внешних условий, в литературе называются «задачами принятия решений в условиях риска», а задачи, связанные с принятием решений при неизвестных вероятностях внешних условий - «задачами принятия решений в условиях неопределенности».

Литература

1. Абчук В.А. Экономико-математические методы. СПб: Союз, 1999. - 320 с.
2. Анфилатов В.С. Системный анализ в управлении. М.: Финансы и статистика, 2002. - 368 с.
3. Банди Б. Основы линейного программирования. М.: Радио и связь, 1989. - 176 с.
4. Вентцель Е.С. Исследование операций: задачи, принципы, методология. М.: Высш. шк., 2001. - 208 с.
5. Вентцель Е.С., Овчаров Л.А. Прикладные задачи теории вероятностей. М.: Радио и связь, 1983. - 416 с.
6. Дубров А.М., Лагоша Б.А., Хрусталев Е.Ю. Моделирование рисковых ситуаций в экономике и бизнесе. М.: Финансы и статистика, 1999. - 176 с.
7. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы для экономистов и менеджеров. М.: Финансы и статистика, 1998. - 350 с.
8. Замков О.О. и др. Математические методы в экономике. - М.: Изд-во МГУ, 1997. - 368 с.
9. Кузнецов А.В., Сакович В.А., Холод Н.И. Высшая математика: Математическое программирование. – Мн.: Выш. шк., 2001. – 351 с.
10. Лабскер Л.Г., Бабешко Л.О. Игровые методы в управлении экономикой и бизнесом. М.: Дело, 2001. - 464 с.
11. Ларичев О.И. Теория и методы принятия решений. М.: Логос, 2000. - 296 с.
12. Сборник задач и упражнений по высшей математике: Математическое программирование / Под общей ред. А.В. Кузнецова и Р.А. Рутковского. – Мн.: Выш. шк., 2002. – 447 с.
13. Смородинский С.С., Батин Н.В. Методы анализа и принятия решений в слабоструктурированных задачах. Мн.: БГУИР, 2002. - 116 с.
14. Смородинский С.С., Батин Н.В. Оптимизация решений на основе методов и моделей мат. программирования: Учеб. пособие. – Мн.: БГУИР, 2003. – 136 с.
15. Спицнадель В.Н. Основы системного анализа. СПб.: Издательский дом «Бизнес-пресса», 2000. - 326 с.
16. Таха Х. Введение в исследование операций. – М.: Вильямс, 2001. – 912 с.
17. Экономико-математические методы и модели / Под ред. А.В. Кузнецова. – Мн.: БГЭУ, 1999. – 413 с.