Министерство образования Российской федерации

Нижегородский государственный университет
им. Н.И.Лобачевского

В.Л. Тарасов

Экономико-математические методы и модели

Учебное пособие
для студентов сокращенных форм обучения специальностей
«Финансы и кредит» и «Бухгалтерский учет, анализ и аудит»

Нижний Новгород

2003

Содержание

Предисловие

Настоящее учебное пособие предназначено студентам специальностей 060400 – «Финансы и кредит» и 060500 – «Бухгалтерский учет, анализ и аудит», проходящих обучение по сокращенным формам. В связи со спецификой данной формы обучения, а именно, сокращенным числом очных лекционных и практических занятий, в курс вошли наиболее важные, по мнению автора, вопросы: линейное программирование, динамическое программирование, теория игр, теория массового обслуживания, балансовые модели. Согласно действующему Государственному образовательному стандарту высшего профессионального образования для специальностей 060400 и 060500 в состав дисциплины Математика кроме упомянутых выше тем, входит еще ряд вопросов, относящихся к проблематике экономико-математических методов и моделей (сетевые графики; функции полезности, спроса, производственные функции, модели конкуренции, равновесия), которые не рассматриваются в данном пособии, однако большинство из них обсуждается в общеэкономических дисциплинах.

Глава 1. Балансовые модели

1. Основные понятия

Балансовые модели сопоставляют источники поступления какого-либо блага с направлениями его использования. Например, если говорят о зерновом балансе на какой-либо период, то имеют в виду сравнение поступления зерна из различных источников (урожай, резервы, импорт) с его расходованием по различным направлениям (снабжение населения, животноводство, пополнение запасов, экспорт). Государственный бюджет также является балансом, в котором указываются источники поступления денежных средств и направления их расходования. Если расходы превышают доходы, имеется дефицит и должен быть указан источник его покрытия, если доходы превышают расходы, имеется профицит, который можно использовать на пополнение резервов.

Для анализа экономики страны в целом, составляют *межотраслевые балансы производства и распределения продукции*. Народное хозяйство представляют совокупностью n отраслей. Количество отраслей, представленных в балансе, может быть различным в зависимости от степени детализации, например, можно рассмотреть только три отрасли: промышленность, сельское хозяйство, строительство. В реальных экономических исследованиях количество отраслей может составлять несколько сотен.

Балансовый метод получил большое развитие в работах Василия Васильевича Леонтьева (5.08.1906–8.02.1999), который представил его в математической форме и выполнил на его основе много конкретных экономических исследований. Леонтьев окончил Ленинградский университет в 1924г., в 1925г. уехал в Германию для продолжения образования, с 1931г. жил и работал в США. При изучении Леонтьевым американской экономики она представлялась различным количеством отраслей, до 450. В одной из работ В.В.Леонтьев исследовал экономику США с 8 отраслевой моделью. Для примера приведем список этих 8 отраслей:

1. Продукты питания и лекарства;
2. Ткани, одежда, мебель;
3. Оборудование и машины;
4. Транспортные средства и бытовая техника;
5. Строительство;
6. Металлы;
7. Энергия;
8. Химические продукты.

Каждая отрасль, производя свою продукцию, использует для производства продукцию других отраслей. В свою очередь, продукция, произведенная данной отраслью, поступает в другие отрасли для использования в дальнейшем процессе производства, и на непроизводственное или конечное потребление.

2. Схема межотраслевого баланса

Схема межотраслевого баланса производства и распределения продукции приведена в табл.1. Каждая отрасль фигурирует в балансе и как производящая и как потребляющая.

Таблица 1. Схема межотраслевого баланса

		Потребляющие отрасли				Конечный продукт	Валовой продукт
		1	2	…	n		
Производящие отрасли	1	x_{11}	x_{12}	…	x_{1n}	Y_1	X_1
	2	x_{21}	x_{22}	…	x_{2n}	Y_2	X_2
	.	.	.	.	.	.	.
	.	.	.	I	.	II	.
	.	.	.	.	.	.	.
	n	x_{n1}	x_{n2}	…	x_{nn}	Y_n	X_n
Амортизация		c_1	c_2	…	c_n	IV	$\sum Z_i$
Оплата труда		v_1	v_2	III	v_n		
Чистый доход		M_1	m_2	…	m_n		
Валовой продукт		X_1	X_2	…	X_n	$\sum Y_i$	

Выделяют четыре части или *квадранта* межотраслевого баланса (МОБ), которые имею различное экономическое содержание.

I. Первый квадрант – это таблица межотраслевых связей. Величина x_{ij} равна объему продукции i-й отрасли, использованной в качестве материальных затрат j-й отраслью в ее производстве. Сумма элементов i-й строки $\sum_{j=1}^{n} x_{ij}$ представляет *промежуточный продукт* i-й отрасли, который поступает в дальнейшее производственное использование. Сумма элементов j-го столбца $\sum_{i=1}^{n} x_{ij}$ представляет *материальные издержки* j-й отрасли.

II. Второй квадрант содержит *конечную продукцию* отраслей. Под конечной продукцией понимается часть продукции, выходящая за пределы текущего производственного потребления. Она используется на личное и общественное потребление, пополнение запасов, накопление, возмещение потерь, экспорт за минусом импорта. Второй квадрант характеризует отраслевую структуру *национального дохода*.

III. Третий квадрант характеризует национальный доход со стороны его стоимостного состава. Сумма чистого дохода (прибыли) отраслей и оплаты труда называется *чистой продукцией* или *добавленной стоимостью*. Сумма амортизации и добавленной стоимости обозначается Z_j и называется *условно-чистой продукцией*

$$Z_j = c_j + v_j + m_j,\ j=1, 2, \ldots, n$$

IV. Четвертый квадрант находится на пересечении столбца конечной продукции и строк третьего квадранта. При его подробном раскрытии он отражает конечное распределение национального дохода на доходы населения, предприятий и государства и показывает направления его использования на потребление и накопление.

Валовая продукция состоит из промежуточного и конечного продукта. Суммируя показатели таблицы по строкам, получим

$$\begin{cases} x_{11} + x_{12} + \ldots + x_{1j} + \ldots + x_{1n} + Y_1 = X_1, \\ x_{21} + x_{22} + \ldots + x_{2j} + \ldots + x_{2n} + Y_2 = X_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{i1} + x_{i2} + \ldots + x_{ij} + \ldots + x_{in} + Y_i = X_i, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{n1} + x_{n2} + \ldots + x_{nj} + \ldots + x_{nn} + Y_n = X_n, \end{cases} \tag{1}$$

или в кратком виде

$$\sum_{j=1}^{n} x_{ij} + Y_i = X_i, i = \overline{1,n}. \qquad (2)$$

С другой стороны, валовая продукция включает материальные издержки и условно-чистую продукцию, поэтому, суммируя столбцы таблицы МОБ, получим

$$\sum_{i=1}^{n} x_{ij} + Z_i = X_i, j = \overline{1,n}. \qquad (3)$$

Сложим n уравнений системы (2):

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} + \sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} X_i.$$

Сложим также n уравнений системы (3):

$$\sum_{j=1}^{n}\sum_{i=1}^{n} x_{ij} + \sum_{j=1}^{n} Z_j = \sum_{j=1}^{n} X_j.$$

Правые части последних двух равенств одинаковы, значит, одинаковы и левые, отсюда

$$\sum_{i=1}^{n} Y_i = \sum_{j=1}^{n} Z_j, \qquad (4)$$

т.е. *суммарный конечный продукт народного хозяйства равен суммарной условно-чистой продукции*.

3. Математическая модель межотраслевого баланса

Таблицы МОБ составляются на основе обработки статистических данных о работе народного хозяйства за определенный промежуток времени. Так в 1925г. ЦСУ СССР составило первый в мире баланс народного хозяйства СССР за 1923/24г., включающий шахматный баланс межотраслевых производственных связей по 14 отраслям промышленности, 4 отраслям сельского хозяйства и 4 прочим отраслям народного хозяйства. Аналогичные балансы для американской экономики составлялись при участии В.В.Леонтьева в 1930-х годах на основе данных американской статистики. Были построены таблицы размерности 41×41 для 1919 и 1929г. Анализ таблиц показал, что хотя межотраслевые поставки x_{ij} изменяются от года к году, относительные величины

$$a_{ij} = \frac{x_{ij}}{X_j}, \; i, j = \overline{1,n} \qquad (5)$$

меняются незначительно, являются устойчивыми. Коэффициенты a_{ij} называются *коэффициентами прямых материальных затрат*. Величина a_{ij} показывает, какое количество продукции i-й отрасли необходимо, чтобы j-я отрасль произвела единицу валовой продукции.

Устойчивость коэффициентов прямых материальных затрат дает возможность, составив отчетный баланс о работе экономики за какой-либо год и рассчитав коэффициенты a_{ij}, использовать их для целей планирования или для прогнозирования развития экономики в будущем.

Коэффициенты a_{ij} определяются сложившейся технологией производства, поэтому их называют технологическими коэффициентами. Они образуют матрицу

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

которая называется технологической.

С использованием коэффициентов прямых материальных затрат из (5) можно выразить величины межотраслевых поставок:

$$x_{ij} = a_{ij} X_j, \; i, j = \overline{1,n}. \tag{6}$$

С использование (6) система (1) запишется:

$$\begin{cases} a_{11}X_1 + a_{12}X_2 + \ldots + a_{1j}X_j + \ldots + a_{1n}X_n + Y_1 = X_1, \\ a_{21}X_1 + a_{22}X_2 + \ldots + a_{2j}X_j + \ldots + a_{2n}X_n + Y_2 = X_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{i1}X_1 + a_{i2}X_2 + \ldots + a_{ij}X_j + \ldots + a_{in}X_n + Y_i = X_i, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}X_1 + a_{n2}X_2 + \ldots + a_{nj}X_j + \ldots + a_{nn}X_n + Y_n = X_n. \end{cases} \tag{7}$$

Введем вектора-столбцы валовой и конечной продукции:

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}, \quad \boldsymbol{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix},$$

тогда систему (7) можно записать в виде

$$\boldsymbol{AX} + \boldsymbol{Y} = \boldsymbol{X}. \tag{8}$$

Система (7) или в матричном виде (8) называется экономико-математической моделью межотраслевого баланса или моделью «затраты – выпуск» (моделью Леонтьева).

Используя модель «затраты – выпуск» можно решать две основные задачи:

1. Пусть известна валовая продукция отраслей, т.е. вектор $\boldsymbol{X}$ и требуется найти конечную продукцию, т.е. вектор $\boldsymbol{Y}$. Конечная продукция легко находится из (8):

$$\boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{AX}. \tag{9}$$

2. Задана конечная продукция отраслей, т.е. вектор $\boldsymbol{Y}$, требуется найти валовую продукцию $\boldsymbol{X}$, выпуск которой обеспечит заданный уровень конечной продукции. Даная задача возникает при планировании работы отраслей для достижения необходимых целей, выраженных в объемах конечной продукции. Для решения этой задачи перепишем систему (8) в виде:

$$\boldsymbol{X} - \boldsymbol{AX} = \boldsymbol{Y},$$

или

$$(\boldsymbol{E} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{Y}, \tag{10}$$

где

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

единичная матрица.

Из (10) находим:

$$\boldsymbol{X} = (\boldsymbol{E} - \boldsymbol{A})^{-1}\boldsymbol{Y} = \boldsymbol{BY}\ . \qquad (10)$$

Здесь для обратной матрицы введено обозначение

$$\boldsymbol{B} = (\boldsymbol{E} - \boldsymbol{A})^{-1} \qquad (11)$$

Рассмотрим экономический смысл матрицы $\boldsymbol{B}$. Пусть требуется произвести конечную продукцию в следующих объемах:

$$Y_1 = 1,\ Y_2 = Y_3 = \ldots = Y_n = 0.$$

Подставим вектор $\boldsymbol{Y}$ с указанными компонентами в (10) и запишем выражение подробно:

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Отсюда

$$X_1 = b_{11},\ X_2 = b_{21},\ \ldots\ X_n = b_{n1}.$$

Таким образом, элементы первого столбца матрицы $\boldsymbol{B}$ равны валовой продукции отраслей, которую необходимо произвести, чтобы первая отрасль произвела единицу своей конечной продукции. Аналогичные выводы можно сделать относительно всех столбцов матрицы $\boldsymbol{B}$. Элементы матрицы $\boldsymbol{B}$ называются *коэффициентами полных материальных затрат*, а сама матрица – *матрицей полных материальных затрат*.

3.1. Свойства матрицы прямых материальных затрат

По экономическому смыслу матрицы $\boldsymbol{A}$ все ее элементы неотрицательны:

$$a_{ij} \geq 0, \quad i, j = \overline{1,n}\ .$$

Так как коэффициент a_{ij} представляет долю валовой продукции i-й отрасли, то

$$a_{ij} \leq 1, \quad i, j = \overline{1,n}\ .$$

Из (8)

$$\boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{AX} \geq 0\ ,$$

так как конечная продукция неотрицательна. Отсюда

$$\boldsymbol{X} \geq \boldsymbol{AX} \qquad (12)$$

Матрица $\boldsymbol{A}$, удовлетворяющая свойству (12), называется *продуктивной*.

Для того чтобы матрица была продуктивна, достаточно, чтобы суммы элементов каждого столбца были меньше единицы.

4. Межотраслевой баланс труда

Балансовый метод позволяет изучать различные экономические показатели. В данном разделе анализируется распределение затрат живого и овеществленного труда в отраслях.

Обозначим *затраты живого труда* в j-й отрасли L_j. Тогда *прямые затраты* труда на производство единицы валовой продукции (*коэффициенты прямой трудоемкости*) составят

$$t_j = \frac{L_j}{X_j}, \quad j = \overline{1,n}. \tag{13}$$

Введем понятие *полных затрат труда* как суммы прямых затрат живого труда и затрат *овеществленного труда*, перенесенных на продукт через израсходованные средства производства. Обозначим T_i полные затраты труда на производство единицы валовой продукции i-го вида. Произведение $a_{ij}T_i$ равно полным затратам овеществленного труда, заключенного в продукции i-й отрасли, которая использована для производства единицы валовой продукции j-й отрасли. Тогда полные трудовые затраты на единицу валовой продукции j-й отрасли (*коэффициенты полной трудоемкости*) будут равны:

$$T_j = t_j + \sum_{i=1}^{n} a_{ij} T_i, \quad j = \overline{1,n} \tag{14}$$

или в развернутом виде:

$$\begin{cases} T_1 = t_1 + a_{11}T_1 + a_{21}T_2 + \ldots + a_{i1}T_i + \ldots + a_{n1}T_n, \\ T_2 = t_2 + a_{12}T_1 + a_{22}T_2 + \ldots + a_{i2}T_i + \ldots + a_{n2}T_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ T_j = t_j + a_{1j}t_1 + a_{2j}t_2 + \ldots + a_{ij}t_i + \ldots + a_{nj}t_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ T_n = t_n + a_{1n}t_1 + a_{2n}t_2 + \ldots + a_{in}t_i + \ldots + a_{nn}t_n. \end{cases} \tag{15}$$

Введем вектор-строку

$$\boldsymbol{t} = (t_1, t_2, \ldots, t_n)$$

коэффициентов прямой трудоемкости и вектор-строку

$$\boldsymbol{T} = (T_1, T_2, \ldots, T_n)$$

коэффициентов полной трудоемкости. В матричных обозначениях система (15) запишется:

$$\boldsymbol{T} = \boldsymbol{t} + \boldsymbol{TA}. \tag{16}$$

Коэффициенты прямой трудоемкости могут быть получены из статистических данных о числе работающих в каждой отрасли, то есть вектор $\boldsymbol{t}$ может считаться известным. Вектор коэффициентов полной трудоемкости выражается через вектор коэффициентов прямой трудоемкости путем решения системы (16):

$$\boldsymbol{T} - \boldsymbol{TA} = \boldsymbol{t}, \quad \boldsymbol{T}(\boldsymbol{E} - \boldsymbol{A}) = \boldsymbol{t}, \quad \boldsymbol{T} = \boldsymbol{t}(\boldsymbol{E} - \boldsymbol{A})^{-1},$$

а с учетом обозначения (11)

$$\boldsymbol{T} = \boldsymbol{tB} \tag{17}$$

Обозначим через L общие затраты живого труда во всех отраслях. Имеем

$$L = \sum_{j=1}^{n} L_j = \sum_{j=1}^{n} t_j X_j = \boldsymbol{tX}.$$

Согласно (10) $\boldsymbol{X} = \boldsymbol{BY}$, поэтому

$$\boldsymbol{tX} = \boldsymbol{tBY},$$

а с учетом (17) получаем:

$$\boldsymbol{tX} = \boldsymbol{TY}. \tag{18}$$

Равенство (18) означает, что стоимость конечной продукции, оцененной по полным затратам труда, равна совокупным затратам живого труда.

Обратим внимание на *парадокс Леонтьева*. Изучая американскую экономику, Леонтьев пришел к заключению, что в американском экспорте труд преобладает над капиталом, то есть в общей стоимости экспорта стоимость труда превышает стоимость сырья, материалов, энергии. С этим можно согласиться, если обратить внимание на то, что в американском экспорте большую долю занимают кинофильмы, программные продукты, то есть товары, на создание которых затрачивается, в основном, труд. Импортируют же США много энергоресурсов, другого сырья. Такое соотношение экспорта и импорта выгодно для США, так как в большей степени экспортируются возобновляемые ресурсы.

5. Межотраслевой баланс фондов

Под фондами понимается капитал, занятый в отраслях. Фонды подразделяются на основные (здания, сооружения, оборудование и т.п.) и оборотные (сырье, материалы, финансы). Здесь не будем проводить детализацию и рассмотрим фонды как единую величину.

Пусть в j-й отрасли задействованы фонды Φ_j. Тогда на производство единицы валовой продукции приходятся фонды в количестве

$$f_j = \frac{\Phi_j}{X_j}, \quad j = \overline{1,n}. \tag{18}$$

Величины f_j называются *коэффициентами прямой фондоемкости*.

Коэффициенты полной фондоемкости F_i равны объему фондов во всех отраслях, который необходим для выпуска единицы валовой продукции i-й отрасли. Рассмотрим j-й столбец таблицы межотраслевого баланса и предположим, что валовая продукция этой отрасли равна 1. Тогда во всех отраслях необходимы фонды F_j, которые складываются из фондов f_j, непосредственно занятых в j-й отрасли и фондов во всех отраслях, необходимых для производства продукции в количествах a_{ij}, $i = \overline{1,n}$ и можно записать

$$F_j = f_j + \sum_{i=1}^{n} a_{ij} F_i, \quad j = \overline{1,n}. \tag{19}$$

Здесь произведение $a_{ij}F_i$ равно полному расходу фондов i-й отрасли, необходимому, для производства единицы валовой продукции j-й отрасли. В развернутом виде система (19) выглядит аналогично системе (15).

Введем вектор-строку

$$\boldsymbol{f} = (f_1, f_2, \ldots, f_n)$$

коэффициентов прямой фондоемкости и вектор-строку

$$\boldsymbol{F} = (F_1, F_2, \ldots, F_n)$$

коэффициентов полной фондоемкости. В матричных обозначениях система (19) запишется:

$$\boldsymbol{F} = \boldsymbol{f} + \boldsymbol{FA}. \tag{20}$$

Так как объем фондов, занятых в каждой отрасли, может быть найден по результатам статистической отчетности, вектор коэффициентов прямой фондоемкости $\boldsymbol{f}$ будем считать известным, тогда вектор коэффициентов полной фондоемкости находится как решение системы (20):

$$\boldsymbol{F} = \boldsymbol{f}(\boldsymbol{E} - \boldsymbol{A})^{-1} = \boldsymbol{f}\boldsymbol{B}. \qquad (21)$$

Использование коэффициентов фондоемкости позволяет увязать планируемый объем выпуска продукции с имеющимися производственными мощностями.

6. Пример расчетов с использованием балансовой модели

Пусть заданы: матрица прямых материальных затрат

$$\boldsymbol{A} = \begin{pmatrix} 0.1 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.2 \\ 0.5 & 0.1 & 0.3 \end{pmatrix},$$

вектор конечной продукции

$$\boldsymbol{Y} = \begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix},$$

вектор коэффициентов прямой трудоемкости

$$\boldsymbol{t} = (1.2, 1.4, 1.9)$$

и вектор коэффициентов прямых затрат фондов

$$\boldsymbol{f} = (1.5, 2.0, 1.2).$$

Требуется составить межотраслевой баланс производства и распределения продукции, баланс труда и баланс фондов.

6.1. Расчет межотраслевого баланса производства и распределения продукции

Для построения межотраслевого баланса производства и распределения продукции нужно найти конечную продукцию отраслей по формуле

$$\boldsymbol{X} = (\boldsymbol{E} - \boldsymbol{A})^{-1}\boldsymbol{Y}.$$

Предварительно проверяем, является ли матрица $\boldsymbol{A}$ продуктивной. Достаточным условием продуктивности является требование, чтобы суммы элементов столбцов матрицы были меньше 1. Непосредственной проверкой устанавливаем, что для заданной матрицы это требование выполняется.

Вычисление обратной матрицы

$$\boldsymbol{B} = (\boldsymbol{E} - \boldsymbol{A})^{-1}$$

производим по формуле

$$(\boldsymbol{E} - \boldsymbol{A})^{-1} = \frac{\overline{(\boldsymbol{E} - \boldsymbol{A})}}{|\boldsymbol{E} - \boldsymbol{A}|}.$$

Здесь $|\boldsymbol{E} - \boldsymbol{A}|$ – определитель матрицы, $\overline{(\boldsymbol{E} - \boldsymbol{A})}$ – присоединенная матрица, которая образована алгебраическими дополнениями элементов транспонированной матрицы $(\boldsymbol{E} - \boldsymbol{A})^T$.

Формула для вычисления определителя третьего порядка имеет вид

$$|\mathbf{D}| = \begin{vmatrix} d_{11}d_{12}d_{13} \\ d_{21}d_{22}d_{23} \\ d_{31}d_{32}d_{33} \end{vmatrix} = d_{11}d_{22}d_{33} + d_{12}d_{23}d_{31} + d_{13}d_{21}d_{32} - d_{13}d_{22}d_{31} - d_{12}d_{21}d_{33} - d_{11}d_{23}d_{32}.$$

Алгебраическим дополнением некоторого элемента матрицы называется взятое со знаком плюс или минус значение определителя, полученного вычеркиванием из матрицы столбца и строки, на пересечении которых расположен данный элемент. Определитель берется со знаком плюс, если сумма номера строки и номера столбца четная и со знаком минус, если эта сумма нечетная. Например, алгебраическое дополнение элемента d_{12} матрицы $\boldsymbol{D}$ есть

$$D_{12} = (-1)^{1+2}\begin{vmatrix} d_{21} & d_{23} \\ d_{31} & d_{33} \end{vmatrix} = -(d_{21}d_{33} - d_{23}d_{31});$$

а алгебраическое дополнение элемента d_{31} есть

$$D_{31} = (-1)^{3+1}\begin{vmatrix} d_{12} & d_{13} \\ d_{22} & d_{23} \end{vmatrix} = d_{12}d_{23} - d_{13}d_{22}.$$

Вычисление матрицы $\boldsymbol{B}$ начинаем с нахождения матрицы $\boldsymbol{E}–\boldsymbol{A}$:

$$\boldsymbol{E} - \boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.1 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.2 \\ 0.5 & 0.1 & 0.3 \end{pmatrix} = \begin{pmatrix} 0.9 & -0.2 & -0.3 \\ -0.3 & 0.6 & -0.2 \\ -0.5 & -0.1 & 0.7 \end{pmatrix}.$$

Далее находим определитель $|\boldsymbol{E} - \boldsymbol{A}|$:

$$|\boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} 0.9 & -0.2 & -0.3 \\ -0.3 & 0.6 & -0.2 \\ -0.5 & -0.1 & 0.7 \end{vmatrix} =$$

$$0.9 \cdot 0.6 \cdot 0.7 + ((-0.2)\cdot(-0.2)\cdot(-0.5)) + ((-0.3)\cdot(-0.3)\cdot(-0.1)) -$$
$$((-0.3)\cdot 0.6(-0.5)) - ((-0.2)\cdot(-0.3)\cdot 0.7) - (0.9\cdot(-0.2)\cdot(-0.1)) =$$
$$0.378 - 0.02 - 0.009 - 0.09 - 0.042 - 0.018 = 0.199.$$

Транспонируем матрицу $(\boldsymbol{E} - \boldsymbol{A})$, то есть переставляем строки и столбцы:

$$(\boldsymbol{E} - \boldsymbol{A})^T = \begin{pmatrix} 0.9 & -0.3 & -0.5 \\ -0.2 & 0.6 & -0.1 \\ -0.3 & -0.2 & 0.7 \end{pmatrix}.$$

Находим алгебраические дополнения элементов матрицы $(\boldsymbol{E} - \boldsymbol{A})^T$:

$$A_{11} = (-1)^{1+1}\begin{vmatrix} 0.6 & -0.1 \\ -0.2 & 0.7 \end{vmatrix} = 0.4, \quad A_{12} = (-1)^{1+2}\begin{vmatrix} -0.2 & -0.1 \\ -0.3 & 0.7 \end{vmatrix} = 0.17,$$

$$A_{13} = (-1)^{1+3}\begin{vmatrix} -0.2 & 0.6 \\ -0.3 & -0.2 \end{vmatrix} = 0.22, \quad A_{21} = (-1)^{2+1}\begin{vmatrix} -0.3 & -0.5 \\ -0.2 & 0.7 \end{vmatrix} = 0.31,$$

$$A_{22} = (-1)^{2+2}\begin{vmatrix} 0.9 & -0.5 \\ -0.3 & 0.7 \end{vmatrix} = 0.48, \quad A_{23} = (-1)^{2+3}\begin{vmatrix} 0.9 & -0.3 \\ -0.3 & -0.2 \end{vmatrix} = 0.27,$$

$$A_{31} = (-1)^{3+1}\begin{vmatrix} -0.3 & -0.5 \\ 0.6 & -0.1 \end{vmatrix} = 0.33, \quad A_{32} = (-1)^{3+2}\begin{vmatrix} 0.9 & -0.5 \\ -0.2 & -0.1 \end{vmatrix} = 0.19,$$

$$A_{33} = (-1)^{3+3}\begin{vmatrix} 0.9 & -0.3 \\ -0.2 & 0.6 \end{vmatrix} = 0.48.$$

Получается следующая присоединенная матрица:

$$\overline{(\boldsymbol{E} - \boldsymbol{A})} = \begin{pmatrix} 0.40 & 0.17 & 0.22 \\ 0.31 & 0.48 & 0.27 \\ 0.33 & 0.19 & 0.48 \end{pmatrix}.$$

Делим все элементы этой матрицы на определитель и получаем матрицу коэффициентов полных материальных затрат

$$B = (E - A)^{-1} = \begin{pmatrix} 0.40 & 0.17 & 0.22 \\ 0.31 & 0.48 & 0.27 \\ 0.33 & 0.19 & 0.48 \end{pmatrix} \cdot \frac{1}{0.199} = \begin{pmatrix} 2.01 & 0.854 & 1.11 \\ 1.56 & 2.41 & 1.36 \\ 1.66 & 0.955 & 2.41 \end{pmatrix}.$$

Вычисления проводим с точностью три значащие цифры, проводя соответствующие округления.

Находим теперь вектор валовой продукции

$$\boldsymbol{X} = \boldsymbol{B} \cdot \boldsymbol{Y} = \begin{pmatrix} 2.01 & 0.854 & 1.11 \\ 1.56 & 2.41 & 1.36 \\ 1.66 & 0.955 & 2.41 \end{pmatrix} \cdot \begin{pmatrix} 100 \\ 200 \\ 300 \end{pmatrix} = \begin{pmatrix} 704 \\ 1045 \\ 1080 \end{pmatrix}.$$

Умножая столбцы матрицы $\boldsymbol{A}$ соответственно на X_1, X_2, X_3, получаем значения межотраслевых потоков продукции:

$$x_{11} = a_{11} X_1 = 0.1 \cdot 704 = 70.4; \quad x_{12} = a_{12} X_2 = 0.2 \cdot 1045 = 209; \quad x_{13} = a_{13} X_3 = 0.3 \cdot 1080 = 324;$$
$$x_{21} = a_{21} X_1 = 0.3 \cdot 704 = 211; \quad x_{22} = a_{22} X_2 = 0.4 \cdot 1045 = 418; \quad x_{23} = a_{23} X_3 = 0.2 \cdot 1080 = 216;$$
$$x_{31} = a_{31} X_1 = 0.5 \cdot 704 = 352; \quad x_{32} = a_{32} X_2 = 0.1 \cdot 1045 = 105; \quad x_{33} = a_{33} X_3 = 0.3 \cdot 1080 = 324.$$

На основе сделанных вычислений составим таблицу межотраслевого баланса (табл.2).

Таблица 2.Межотраслевой баланс производства и распределения продукции

Производящие отрасли	Потребляющие отрасли			Конечная продукция Y_i, i=1,2,3	Валовая продукция X_i, i=1,2,3
	1	2	3		
1 2 3	70.4 211 352	209 418 105	324 216 324	100 200 300	704 1045 1080
Материальные издержки	633	732	864	600	
Условно-чистая продукция Z_j, j=1,2,3	71.0	313	216	600	
Валовая продукция	704	1045	1080		

Из табл.2 видно, что

$$Z_1 + Z_2 + Z_3 = Y_1 + Y_2 + Y_3 = 600,$$

то есть соблюдается принцип единства стоимостного и материального состава национального дохода.

6.2. Расчет баланса труда

Затраты живого труда, необходимые для выпуска единицы продукции j-й отрасли называются коэффициентами прямой трудоемкости. Они образуют вектор-строку

$$\boldsymbol{t} = (t_1, t_2, t_3).$$

Коэффициенты полной трудоемкости равны затратам живого и овеществленного труда, использованного при производстве единицы продукции какой-либо отрасли.

Вектор коэффициентов полной трудоемкости

$$\boldsymbol{T} = (T_1, T_2, T_3)$$

связан с вектором коэффициентов прямой трудоемкости формулой

$$\boldsymbol{T} = \boldsymbol{tB}.$$

Вычислим компоненты вектора полной трудоемкости:

$$\boldsymbol{T} = (1.2, 1.4, 1.9)\begin{pmatrix} 2.01 & 0.854 & 1.11 \\ 1.56 & 2.41 & 1.36 \\ 1.66 & 0.955 & 2.41 \end{pmatrix};$$

$$T_1 = 1.2 \cdot 2.01 + 1.4 \cdot 1.56 + 1.9 \cdot 1.66 = 7.74,$$

$$T_2 = 1.2 \cdot 0.854 + 1.4 \cdot 2.41 + 1.9 \cdot 0.955 = 6.22,$$

$$T_3 = 1.2 \cdot 1.11 + 1.4 \cdot 1.36 + 1.9 \cdot 2.41 = 7.81.$$

Полные затраты живого труда в народном хозяйстве находим по формуле

$$L = \boldsymbol{tX} = t_1 X_1 + t_2 X_2 + t_3 X_3 =$$
$$1.2 \cdot 704 + 1.4 \cdot 1045 + 1.9 \cdot 1080 = 4360.$$

Убедимся, что эта величина равна стоимости конечной продукции, оцененной по полным затратам труда:

$$\boldsymbol{TY} = T_1 Y_1 + T_2 Y_2 + T_3 Y_3 =$$
$$7.74 \cdot 100 + 6.22 \cdot 200 + 7.81 \cdot 300 = 4361.$$

Совпадение говорит о правильности вычислений. Небольшое отклонение вызвано ошибками округления при приближенных вычислениях.

Умножая строки 1, 2, 3 межотраслевого баланса (табл.2) на коэффициенты прямой трудоемкости t_1, t_2, t_3 соответственно, получаем межотраслевой баланс труда (в трудовых измерителях).

Таблица 3. Межотраслевой баланс труда

		Потребляющие отрасли. Межотраслевые затраты овеществленного труда			Затраты труда на конечную продукцию	Затраты труда в отраслях (трудовые ресурсы)
		1	2	3		
Производящие отрасли	1	84.5	251	389	120	845
	2	295	585	302	280	1463
	3	669	200	616	570	2052

6.3. Расчет баланса фондов

Коэффициенты прямой фондоемкости показывают величину производственных фондов, непосредственно занятых в производстве данной отрасли, в расчете на единицу ее валовой продукции. Они обозначаются

$$\boldsymbol{f} = (f_1, f_2, f_3).$$

Коэффициенты полной фондоемкости

$$\boldsymbol{F} = (F_1, F_2, F_3)$$

отражают объем фондов, необходимых во всех отраслях для выпуска единицы продукции данной отрасли. Согласно (21)

$$\boldsymbol{F} = \boldsymbol{fB}.$$

Вычислим по этой формуле коэффициенты полной фондоемкости:

$$\boldsymbol{F} = (1.5,\ 2.0,\ 1.2)\begin{pmatrix} 2.01 & 0.854 & 1.11 \\ 1.56 & 2.41 & 1.36 \\ 1.66 & 0.955 & 2.41 \end{pmatrix};$$

$$\begin{aligned} F_1 &= 1.5\cdot 2.01 + 2.0\cdot 1.56 + 1.2\cdot 1.66 = 8.12, \\ F_2 &= 1.5\cdot 0.854 + 2.0\cdot 2.41 + 1.2\cdot 0.955 = 7.25, \\ F_3 &= 1.5\cdot 1.11 + 2.0\cdot 1.36 + 1.2\cdot 2.41 = 7.27. \end{aligned}$$

Умножая строки 1, 2, 3 межотраслевого баланса (табл.2) на коэффициенты прямой фондоемкости f_1, f_2, f_3 соответственно, получаем межотраслевой баланс фондов.

Таблица 4. Межотраслевой баланс фондов

Производящие отрасли	Потребляющие отрасли. Овеществленные в межотраслевых поставках фонды			Фонды, овеществленные в конечной продукции	Фонды, овеществленные в валовой продукции
	1	2	3		
1	106	314	486	150	1056
2	422	836	432	400	2090
3	422	126	389	360	1296

Глава 2. Линейное программирование

1. Введение

Во многих экономических задачах возникает проблема выбора каких-либо параметров с целью получения наилучшего значения некоторого критерия. Типичным примером такой задачи является проблема, стоящая перед предприятием, которое должно выбрать номенклатуру и объемы выпуска продукции, дающие максимальную прибыль, при условии, что используются имеющиеся ограниченные ресурсы. В случае, когда прибыль (*целевая функция*) и ограничивающие условия (*ограничения*) линейно зависят от варьируемых параметров, математическая модель задачи описывается линейными соотношениями и соответствующая задача оптимизации называется задачей линейного программирования.

Сам термин *линейное программирование* является буквальным переводом английского термина *linear programming*, более точный перевод которого означает *линейное планирование* и который применялся для обозначения задач оптимального планирования ресурсов, математическая модель которых линейна. В рассматриваемом контексте термин *программирование* не имеет отношения к разработке программ для компьютеров.

Впервые сформулировал и решил в 1939г. задачи, позднее названные задачами линейного программирования, советский математик Леонид Витальевич Канторович (1912-1986), рассматривая вопрос об оптимальной работе предприятия по производству фанеры. В 1975 "...за вклад в теорию оптимального распределения ресурсов" Канторовичу была присуждена Нобелевская премия (совместно с американским экономистом Т.Ч. Купмансом). История открытия Л.В. Канторовичем линейного программирования и сложный путь к признанию его работ описаны в статье [5], опубликованной в Интернете.

В 1947г. независимо от работ Канторовича линейное программирование было разработано американским математиком Дж. Б. Данцигом в связи с задачами военного планирования. Данциг предложил общий метод решения задач линейного программирования, получивший название *симплекс-метод*. Воспоминания Данцига об истории открытия линейного программирования напечатаны в статье [6], перевод которой доступен также в Интернете.

2. Задача о выборе оптимального плана производства

Пусть на предприятии изготавливаются два вида продукции P_1 и P_2. При этом используются три вида сырья S_1, S_2, S_3 (например, труд, сырье и оборудование). Числовые данные приведены в табл.2, там же указана прибыль, получаемая от реализации единицы готового изделия каждого вида.

Таблица 1. Исходные данные для планирования

Ресурсы	*Запасы ресурсов*	*Затраты ресурсов на производство единицы готовой продукции*	
		P_1	$P_{2.}$
S_1 (*Труд*)	140	7	6
S_2 (*Сырье*)	64	4	1
S_3 (*Оборудование*)	64	1	4
Прибыль от реализации единицы готовой продукции		5	6

Требуется выбрать такие объемы выпуска продукции двух видов x_1 и x_2, чтобы общая прибыль была максимальной

$$f(x_1, x_2) = 5x_1 + 6x_2 \Rightarrow \max, \tag{1}$$

выполнялись ограничения на ресурсы:

$$
\begin{aligned}
7x_1 + 6x_2 &\le 140,\\
4x_1 + x_2 &\le 64, \qquad (2)\\
x_1 + 4x_2 &\le 64,
\end{aligned}
$$

и объемы выпуска были неотрицательны:

$$x_1 \ge 0,\ x_2 \ge 0. \qquad (3)$$

Задача (1), (2), (3) является задачей линейного программирования (ЗЛП). Так называются задачи, в которых требуется найти максимум или минимум линейной функции, при ограничениях на переменные в виде линейных равенств или неравенств.

3. Геометрический способ решения ЗЛП

Задаче линейного программирования с двумя независимыми переменными можно дать наглядную геометрическую интерпретацию (рис.1). Каждое из ограничений (2) делит плоскость x_1, x_2 на две полуплоскости. В одной полуплоскости находятся допустимые по отношению к данному ограничению точки, а в другой полуплоскости – недопустимые. Границами являются прямые, уравнения которых получаются из неравенств (2), если в них знака неравенства заменить знаком равенства. На рис.1 показаны эти прямые, стрелки показывают допустимую по отношению к данному ограничению полуплоскость.

Прямые проще всего строить по двум их точкам. Например, для построения первого ограничения

$$7x_1 + 6x_2 = 140,$$

положим сначала в нем x_1=0, получим $6x_2$ = 140, x_2=140/6=23.7; затем положим x_2=0, получим $7x_1$ = 140, x_1=140/7=20. Соединяем точки (0, 23.7) и (20, 0). Для определения допустимой полуплоскости берем какую-нибудь контрольную точку, например (1,1), подставляем в ограничение и получаем

$$7 \cdot 1 = 6 \cdot 1 = 13 < 140\,,$$

значит, допустимой является полуплоскость, содержащая точку (1, 1). Аналогично строим другие ограничения.

Областью допустимых значений является многоугольник *OABCD*. В этом многоугольнике следует найти точку, в которой целевая функция принимает максимальное значение.

Известно, что направление наискорейшего возрастания функции указывает вектор градиента этой функции. Для нашего случая градиент целевой функции равен

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) = (5{,}6).$$

На рисунке градиент изображен жирной стрелкой, отложенной от начала координат. Для его построения следует отложить составляющие по осям координат: 5 – вдоль оси x_1, 6 – вдоль x_2.

Для характеристики поведения целевой функции на рисунке показаны ее линии уровня. Так называются линии, на которых целевая функция имеет постоянное значение. Взяв, например, $f(x_1, x_2)$=60, из (1) получаем уравнение прямой

$$5x_1 + 6x_2 = 60.$$

Известно, что градиент перпендикулярен линиям уровня, рисунок подтверждает это

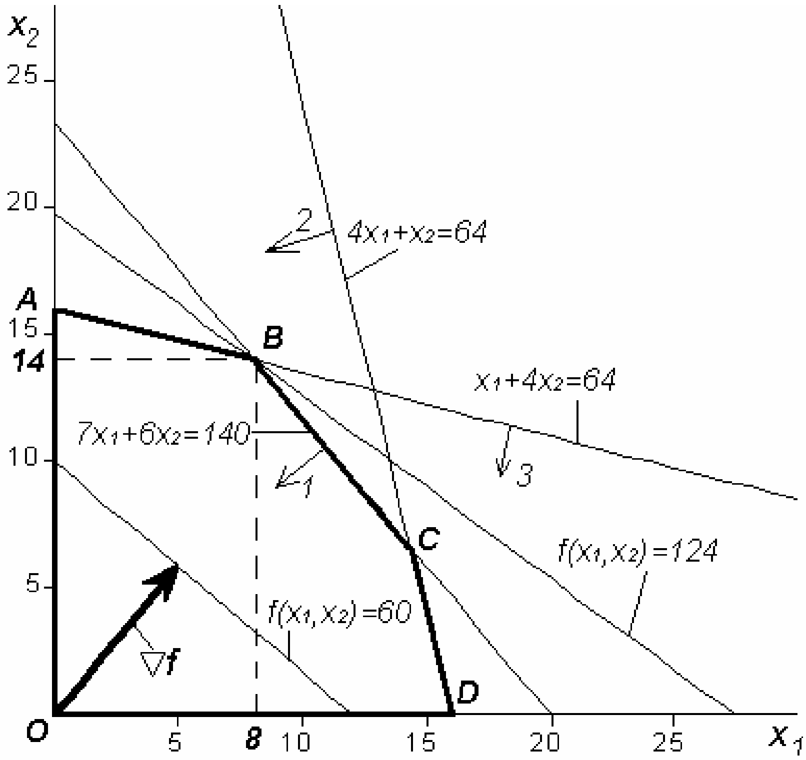

Рис.1. Геометрическое решение ЗЛП

Ясно, что значение целевой функции $f = 60$ не является максимальным, так как в допустимой области есть точки, для которых целевая функция имеет большее значение. Перемещая линию уровня $f = 60$ параллельно самой себе по направлению градиента, получаем, что максимальное значение целевая функция достигнет в точке B. Это точка пересечения первого и третьего ограничений. Рассматривая эти ограничения как равенства, получим систему

$$\begin{aligned} 7x_1 + 6x_2 &= 140, \\ x_1 + 4x_2 &= \ \ 64, \end{aligned}$$

решая которую, находим: $x_1 = 8$, $x_2 = 14$. При этих значениях искомых переменных целевая функция имеет значение

$$f(x_1, x_1) = 5 \cdot 8 + 6 \cdot 14 = 124\,.$$

Конечно, так просто геометрическим способом ЗЛП решается только в случае двух независимых переменных. Однако, анализируя рис.1, можно сделать важные выводы: допустимой областью является выпуклый многоугольник *OABCD*; оптимальным решением всегда является угловая точка допустимой области.

4. Симплекс–метод

4.1. Числовой пример

Говорят, что задача линейного программирования приведена к *каноническому виду*, если все ее ограничения имеют вид равенств. Ограничения–неравенства приводятся к виду равенств путем введения дополнительных неизвестных. Так, обозначив разницу между правой и левой частями первого неравенства системы (2) через x_3, будем иметь:

$$140 - 7x_1 - 6x_2 = x_3 \geq 0.$$

Отсюда

$$7x_1 + 6x_2 + x_3 = 140.$$

Вводя таким же образом дополнительные неизвестные x_4, x_5, получим вместо (2) систему

$$\begin{aligned} 7x_1 + 6x_2 + x_3 \qquad &= 140, \\ 4x_1 + \ \ x_2 + \ \ x_4 \quad &= \ 64, \\ x_1 + 4x_2 + \qquad x_5 &= \ 64, \end{aligned} \tag{4}$$

Как основные, так и дополнительные неизвестные неотрицательны:

$$x_j \geq 0, \quad j = \overline{1,5}. \tag{5}$$

Ограничения (4) представляют собой систему из трех уравнений с пятью неизвестными. Любой набор значений неизвестных, удовлетворяющий ограничениям (4), (5), называется *планом*.

Из системы трех уравнений (5) можно найти три неизвестные, если двум другим дать произвольное значение. Неизвестные, которые получают произвольное значение, называются *свободными*. Остальные неизвестные, находимые из системы ограничений, называются *базисными*. Проще всего давать свободным неизвестным нулевое значение. Если при этом значения базисных неизвестных окажутся неотрицательными, то полученный набор значений неизвестных образует допустимое решение ЗЛП. Оно соответствует одной из угловых точек допустимой области. Выбирая всевозможные комбинации свободных неизвестных, и решая систему ограничений, можно найти все угловые точки допустимой области и среди них выбрать точки с максимальным значением целевой функции. Этот метод перебора не годится для задач со многими неизвестными из-за большого числа вариантов, которые придется перебирать. Так даже для рассматриваемого примера число вариантов выбора двух свободных неизвестных из общего их числа, которое равно пяти, составит

$$C_4^2 = 5 \cdot 4 / 2 = 10,$$

то есть нужно решить 10 систем из трех уравнений.

Симплекс–метод дает способ перехода от одного набора свободных неизвестных к другому (или, что то же самое, от одного базиса к другому) с ростом целевой функции.

Чтобы лучше понять содержание вычислений по симплексному методу, запишем нашу задачу в буквенных обозначениях.

4.2. Симплекс-метод в общем виде

Требуется найти максимум целевой функции

$$f = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 \tag{6}$$

при ограничениях

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + x_3 \qquad &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \quad x_4 \quad &= b_2, \\ a_{31}x_1 + a_{32}x_2 + \qquad x_5 &= b_3, \end{aligned} \tag{7}$$

и условиях неотрицательности (5).

Из системы (7) видно, что если за свободные неизвестные принять x_1, x_2 и положить их равными нулю, то базисные неизвестные x_3, x_4, x_5 будут равны правым частям системы. В результате получается план

$$x_1 = 0, \ x_2 = 0, \ x_3 = b_1, \ x_4 = b_2, \ x_5 = b_3, \tag{8}$$

для которого целевая функция равна

$$f = c_3b_1 + c_4b_2 + c_5b_3. \tag{9}$$

Выразим из (7) базисные неизвестные x_3, x_4, x_5 через свободные x_1, x_2:

$$\begin{aligned} x_3 &= b_1 - a_{11}x_1 - a_{12}x_2, \\ x_4 &= b_2 - a_{21}x_1 - a_{22}x_2, \\ x_5 &= b_3 - a_{31}x_1 - a_{32}x_2. \end{aligned} \tag{10}$$

Подставим (10) в целевую функцию (6) и после приведения подобных членов будем иметь

$$f = (c_1 - c_3a_{11} - c_4a_{21} - c_5a_{31})x_1 + (c_2 - c_3a_{12} - c_4a_{22} - c_5a_{32})x_2 + c_3b_1 + c_4b_2 + c_5b_3. \tag{11}$$

Если коэффициенты перед x_1, x_2 в (11) окажутся отрицательными, то целевую функцию нельзя увеличить, переводя эти свободные неизвестные в базисные, то есть давая им какое-то положительное значение. Отрицательность коэффициентов перед неизвестными в (11) есть признак того, что найдено оптимальное решение. Коэффициенты перед свободными неизвестными в выражении для целевой функции принято называть *оценками*. Введем обозначения:

$$Z_1 = c_3a_{11} + c_4a_{21} + c_5a_{31}, \quad Z_2 = c_3a_{12} + c_4a_{22} + c_5a_{32}. \tag{12}$$

В этих обозначениях условия оптимальности запишутся:

$$c_1 - Z_1 < 0, \;\; c_2 - Z_2 < 0, \tag{13}$$

Пусть условия оптимальности (13) не выполнены, а именно пусть, для определенности, коэффициент перед x_1 в (11) положителен, тогда, увеличивая x_1, мы увеличим целевую функцию по сравнению с ее значением (9). Посмотрим, насколько может быть увеличена переменная x_1 по сравнению с нулем, при условии, что x_2 остается свободной, то есть будет иметь неизменное значение x_2=0. Из (10) видно, что увеличение x_1 будет вести к уменьшению x_3, x_4, x_5, если коэффициенты перед x_1 в (10) отрицательны. Увеличивать x_1 можно лишь до значения, при котором первая из неизвестных x_3, x_4, x_5 обратится в нуль. Положив нулю левые части (10), получим три уравнения:

$$0 = b_1 - a_{11}x_1, \quad 0 = b_2 - a_{21}x_1, \quad 0 = b_3 - a_{31}x_1.$$

Если положить

$$x_1 = min\left(\frac{b_1}{a_{11}}, \frac{b_2}{a_{21}}, \frac{b_3}{a_{31}}\right), \tag{14}$$

то только одна неизвестная из x_3, x_4, x_5 обратится в нуль, то есть станет свободной, а остальные останутся положительными. Предположим, что это будет x_3, тогда будем иметь в качестве свободных неизвестных x_2, x_3 и в качестве базисных x_1, x_4, x_5.

Для удобства дальнейших вычислений желательно преобразовать систему ограничений к виду (7), который характерен тем, что столбцы коэффициентов при базисных неизвестных образуют единичную матрицу. Так как старая базисная переменная x_3 заменяется новой базисной переменной x_1, то в первом уравнении коэффициент перед x_1 должен быть равен 1. Это достигается делением первого уравнения на a_{11}. Далее следует исключить неизвестную x_1 из второго и третьего уравнений. Для этого преобразованное первое уравнение умножим сначала на a_{21} и вычтем из второго, затем на a_{31} и вычтем из третьего. В результате этих преобразований система (7) примет вид

$$\begin{aligned} x_1 + a^1_{12}x_2 + a^1_{13}x_3 \qquad\qquad &= b^1_1, \\ a^1_{22}x_2 + a^1_{23}x_3 + x_4 \qquad &= b^1_2, \\ a^1_{32}x_2 + a^1_{33}x_3 + \qquad x_5 &= b^1_3. \end{aligned} \tag{15}$$

Здесь через a_{ij}^{1} обозначены новые значения коэффициентов. Заметим, что подобные преобразования выполняются в методе Гаусса при решении систем линейных уравнений.

Описанные выше действия по работе с системой ограничений (7) следует повторить для системы (15) и т.д., до тех пор, пока не будет выполнены условия оптимальности.

Для удобства и компактности вычислений по симплекс-методу применяются симплексные таблицы (табл.2).

4.3. Табличная форма симплекс-метода

В столбцы с «x_1» по «x_5» таблицы заносятся коэффициенты системы ограничений (4), в столбец «b_i» записываются их правые части. Элементы столбца «b_i» равны значениям базисных неизвестных текущего плана. В столбце «*Базис*» записываются базисные неизвестные, в столбец «c_j *базиса*» – коэффициенты перед базисными неизвестными в целевой функции. В верхней строке таблицы указываются коэффициенты целевой функции. В четвертой строке размещаются оценки.

Таблица 2. Вычисления по симплекс-методу

№ таблицы		*Базис*	c_j *базиса*	$c_1=5$	$c_2=6$	$c_3=0$	$c_4=0$	$c_5=0$	b_i	$\frac{b_i}{a_{ik}}$
				x_1	x_2	x_3	x_4	x_5		
1	1	x_3	0	7	6	1	0	0	140	23.3
	2	x_4	0	4	1	0	1	0	64	64
	3	x_5	0	1	$\boxed{4}$	0	0	1	64	$\underline{16}$
	4	c_j-Z_j		5	$\underline{6}$	0	0	0	0	
2	1	x_3	0	$\boxed{5.5}$	0	1	0	-1.5	44	$\underline{8}$
	2	x_4	0	3.75	0	0	1	-0.25	48	12.8
	3	x_2	6	0.25	1	0	0	0.25	16	64
	4	c_j-Z_j		$\underline{3.5}$	0	0	0	-1.5	96	
3	1	x_1	5	1	0	0.182	0	-0.273	8	
	2	x_4	0	0	0	-0.682	1	0.773	18	
	3	x_2	6	0	1	-0.045	0	0.318	14	
	4	c_j-Z_j		0	0	-0.636	0	-0.545	124	

Расчеты начинаются с вычисления оценок по формулам (12), (13). Для этого перемножаются элементы столбца «c_j *базиса*» на элементы столбцов «x_1»,…, «x_5», произведения суммируются, полученные суммы вычитаются из соответствующих c_j. Произведение столбца «c_j *базиса*» на столбец «b_i» дает значение целевой функции согласно (9). Оно записывается в оценочной строке столбца «b_i».

Итак, из первой симплексной таблицы имеем:

$$x_1=0,\ x_2=0,\ x_3=140,\ x_4=64.\ x_5=64,\ f=0.$$

В оценочной строке первой симплексной таблицы есть два положительных значения в столбцах «x_1» и «x_2». Любую из этих неизвестных можно перевести в базис, что приведет к росту целевой функции. Для определенности будем переводить в базис неизвестную с большей оценкой. В нашем примере это x_2, у которой оценка имеет значение 6 (подчеркнуто).

Столбец «x_2» называется *направляющим*.

Чтобы выяснить, какую неизвестную следует исключить из базиса, вычисляем отношения $\frac{b_i}{a_{i2}}$ для $a_{i2}>0$, записываем их в последний столбец таблицы. Подчеркнуто минимальное значение этого отношения 16, оно соответствует x_5, значит, эту переменную следует исключить из базиса и сделать свободной.

Строка «x_5» называется *направляющей*.

Элемент, расположенный на пересечение направляющей строки и направляющего столбца, называется *разрешающим*.

Вторую симплексную таблицу начинаем заполнять с указания измененного базиса и коэффициентов в целевой функции при базисных неизвестных.

Делим элементы третьей направляющей строки первой таблицы на разрешающий элемент и записываем во вторую симплексную таблицу также как третью строку. Далее умножаем ее на элементы направляющего столбца и вычитаем из соответствующих строк первой таблицы. Результат записываем во второй таблице. Конкретно, строка 3 второй таблицы умножается сначала на 6 и вычитается из строки 1 первой таблицы, затем умножается на 1 и вычитается из строки 2. Далее вычисляются элементы оценочной строки.

Из второй симплексной таблицы имеем:

$$x_1=0,\ x_2=16,\ x_3=44,\ x_4=48,\ x_5=0,\ f=96.$$

В оценочной строке второй таблицы есть одно положительное значение в столбце x_1. Эта неизвестная переводится в базис. В последний столбец помещаем отношения $\frac{b_i}{a_{i1}}$ для $a_{i1}>0$, находим, что минимум этого отношения 8 соответствует x_3. Эта неизвестная переводится в разряд свободных. Направляющей строкой является первая строка, направляющим столбцом является столбец «x_1», разрешающий элемент имеет значение 5.5. По изложенному выше алгоритму заполняем третью симплексную таблицу.

В третьей симплексной таблице все оценки отрицательны, значит последнее решение

$$x_1=8,\ x_2=14,\ x_3=0,\ x_4=18,\ x_5=0,\ f=124$$

является оптимальным. Геометрическим методом найдено то же решение.

Итак, наибольшая прибыль в количестве 124 единицы может быть получена при выпуске 8 единиц продукции первого вида и 14 единиц продукции второго вида.

5. Двойственная задача

5.1. Экономическая формулировка двойственной задачи

Пусть у предприятия есть возможность продать ресурсы по ценам y_1, y_2, y_3 за единицу каждого их вида. Это будет выгодно, если выручка от продажи ресурсов, идущих на изготовление единицы каждого вида продукции, будет больше, чем прибыль от реализации готовых изделий. Выручка от реализации ресурсов, затрачиваемых на изготовление одного готового изделия, получается сложением произведений норм расхода ресурсов, приведенных в табл.2 на цены. Сравнивая эту выручку с прибылью от реализации готовых изделий, получим, что продажа ресурсов целесообразна, если

$$\begin{aligned} 7y_1+4y_2+\ \ y_3 &\ge 5, \\ 6y_1+1y_2+4y_3 &\ge 6. \end{aligned} \tag{16}$$

Цены неотрицательны:

$$y_1 \ge 0,\ y_2 \ge 0,\ y_3 \ge 0. \tag{17}$$

Потенциальный покупатель стремится установить такие цены на ресурсы y_1, y_2, y_3, чтобы общие затраты на их приобретение были минимальны, то есть

$$g = 140y_1 + 64y_2 + 64y_3 \Rightarrow \min. \tag{18}$$

Задача (18), (16), (17) называется двойственной по отношению к задаче (1), (2), (3). Сама задача (1), (2), (3) называется при этом исходной или прямой.

5.2. Двойственная задача в общем виде

Пусть прямая ЗЛП состоит в максимизации линейной целевой функции n переменных

$$f = c_1x_1+c_2x_2+\ldots+c_nx_n \tag{19}$$

при наличии m ограничений на переменные

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m, \end{aligned} \tag{20}$$

и их неотрицательности

$$x_1 \ge 0,\ x_2 \ge 0,\ \ldots,\ x_n \ge 0. \tag{21}$$

Двойственная задача состоит в минимизации целевой функции

$$g = b_1y_1 + b_2y_2 + \ldots + b_my_m \tag{22}$$

при ограничениях

$$\begin{aligned} &a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_m \ge c_1, \\ &a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_m \ge c_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m \ge c_n, \end{aligned} \tag{23}$$

и неотрицательности переменных

$$y_1 \ge 0,\ y_2 \ge 0,\ \ldots,\ y_m \ge 0. \tag{24}$$

В двойственной задаче (22), (23), (24) число неизвестных равно числу ограничений прямой задачи, коэффициенты целевой функции равны правым частям ограничений прямой задачи (запасам ресурсов), правые части ограничений двойственной задачи равны коэффициентам целевой функции прямой задачи.

Если ввести матрицу коэффициентов ограничений прямой задачи

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \tag{25}$$

то, сравнивая (20) и (23), замечаем, что матрица коэффициентов ограничений двойственной задачи получается транспонированием матрицы коэффициентов ограничений прямой задачи.

Ниже приведена схема связи прямой и двойственной задач.

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="4">Неизвестные прямой задачи</td><td colspan="2" rowspan="2">Ограничения прямой задачи</td><td rowspan="2">Целевая функция двойственной задачи</td></tr>
<tr><td>x_1</td><td>x_2</td><td>…</td><td>x_n</td></tr>
<tr><td rowspan="4">Неизвестные двойственной задачи</td><td>y_1</td><td>a_{11}</td><td>a_{12}</td><td>…</td><td>a_{1n}</td><td>$\leq$</td><td>b_1</td><td rowspan="4">$g=b_1y_1+b_2y_2+\ldots+b_my_n \to \min$</td></tr>
<tr><td>y_2</td><td>a_{21}</td><td>a_{22}</td><td>…</td><td>a_{2n}</td><td>$\leq$</td><td>b_2</td></tr>
<tr><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td></tr>
<tr><td>y_m</td><td>a_{m1}</td><td>a_{m2}</td><td>…</td><td>a_{mn}</td><td>$\leq$</td><td>b_m</td></tr>
<tr><td colspan="2" rowspan="2">Ограничения двойственной задачи</td><td>$\geq$</td><td>$\geq$</td><td>…</td><td>$\geq$</td></tr>
<tr><td>c_1</td><td>c_2</td><td>…</td><td>c_n</td></tr>
<tr><td colspan="2">Целевая функция прямой задачи</td><td colspan="4">$f=c_1x_1+c_2x_2+\ldots c_nx_n \to \max$</td></tr>
</table>

Рис. 2. Схема связи прямой и двойственной задач линейного программирования

5.3. Теоремы двойственности

Связь прямой и двойственной задач выражается тремя теоремами двойственности.

Первая теорема двойственности. *Если одна из взаимно двойственных задач имеет оптимальное решение, то его имеет и другая, причем оптимальные значения их целевых функций равны:*

$$f_{max}=g_{min}.$$

Если целевая функция одной из задач не ограничена, то вторая задача не имеет допустимых решений.

Вторая теорема двойственности. *Пусть* $\boldsymbol{X}^*=\left(x_1^*, x_2^*,\ldots, x_n^*\right)$ *– допустимое решение прямой задачи* (19)–(21), *а* $\boldsymbol{Y}^*=\left(y_1^*, y_2^*,\ldots, y_m^*\right)$ *– допустимое решение двойственной задачи* (22)–(24). *Для того, чтобы* $\boldsymbol{X}^*$, $\boldsymbol{Y}^*$ *были оптимальными решениями прямой и двойственной задач, необходимо и достаточно, чтобы выполнялись соотношения:*

$$x_j^*\left(\sum_{i=1}^{m} a_{ij}y_i^*-c_j\right)=0,\ j=\overline{1,n}, \tag{26}$$

$$y_i^*\left(\sum_{j=1}^{n} a_{ij}x_j^*-b_i\right)=0,\ i=\overline{1,m}. \tag{27}$$

Прежде чем сформулировать третью теорему двойственности, заметим, что, если у прямой задачи (19)–(21) есть оптимальное решение, оно зависит от правых частей ограничений b_i, $i=\overline{1,m}$ (от запасов ресурсов), то есть можно записать

$$f_{max}=f_{max}(b_1, b_2, \ldots, b_m).$$

Третья теорема двойственности. *Компоненты оптимального решения двойственной задачи равны частным производным оптимального значения целевой функции прямой задачи по соответствующим аргументам, т.е.*

$$\frac{\partial f_{max}}{\partial b_i} = y_i^*, \quad i = \overline{1,m}. \qquad (28)$$

Из (28) вытекает, что изменение правой части какого–либо ограничения на Δb_i приведет к изменению максимального значения целевой функции на

$$\Delta f_{max} = \Delta b_i y_i^*. \qquad (29)$$

Следует иметь в виду, что (29) справедливо лишь при небольших вариациях Δb_i, которые не приводят к изменению решения двойственной задачи.

6. Экономический анализ оптимальных решений

Инструментом анализа являются компоненты оптимального решения двойственной задачи. Согласно Л.В. Канторовичу их принято называть *объективно обусловленными оценками*, просто *двойственными оценками* или по зарубежной терминологии *теневыми ценами*.

6.1. Решение двойственной задачи

Двойственная задача представляет собой обычную задачу линейного программирования и может быть решена симплексным методом. Однако, если известно оптимальное решение прямой задачи, оптимальное решение двойственной задачи можно найти более просто, а именно, используя вторую теорему двойственности.

Симплексным методом получено оптимальное решение прямой задачи $x_1^* = 8,\ x_2^* = 14,\ f_{max} = 124$. Согласно (26) если какая–либо компонента оптимального решения прямой задачи положительна, то соответствующее ограничение двойственной задачи выполняется как строгое равенство. В рассматриваемом примере $x_1^* > 0,\ x_2^* > 0$, поэтому ограничения (16) выполняются как равенства

$$\begin{aligned} 7y_1 + 4y_2 + y_3 &= 5, \\ 6y_1 + 1y_2 + 4y_3 &= 6. \end{aligned} \qquad (30)$$

Система двух уравнений (30) содержит три неизвестные. Из (27) видно, что если какое–либо ограничение прямой задачи выполняется как строгое неравенство, то соответствующая компонента оптимального решения двойственной задачи равна нулю. Подставим оптимальное решение прямой задачи $x_1^* = 8,\ x_2^* = 14$ в ограничения (2), получим

$$\begin{aligned} 7x_1^* + 6x_2^* &= 7\cdot 8 + 6\cdot 14 = 140 = 140, \\ 4x_1^* + 1x_2^* &= 4\cdot 8 + 1\cdot 14 = \ \ 46 < 64, \\ 1x_1^* + 4x_2^* &= 1\cdot 8 + 4\cdot 14 = \ \ 64 = \ \ 64. \end{aligned}$$

Второе ограничение прямой задачи для оптимального решения выполняется как строгое неравенство (46<64), поэтому в оптимальном решении двойственной задачи вторая неизвестная равна нулю: $y_2^* = 0$. Учитывая это, получим из (30) систему

$$\begin{aligned} 7y_1 + \ \ y_3 &= 5, \\ 6y_1 + 4y_3 &= 6, \end{aligned}$$

решая которую, найдем

$$y_1^* = 7/11 = 0.636, \quad y_3^* = 6/11 = 0.545.$$

Если прямая задача решалась с помощью симплексного метода, то решение двойственной задачи можно получить непосредственно из симплексной таблицы, используя следующее правило:

Двойственные оценки равны модулям элементов оценочной строки последнего шага симплексной таблицы, расположенных в столбцах, соответствующих дополнительным неизвестным.

Для нашего примера в последней строке симплексной таблицы в столбцах для дополнительных неизвестных x_3, x_4, x_5 получились значения –0.636, 0 и –0.545, значит $y_1^* = 0.636,\ y_2^* = 0,\ y_3^* = 0.545$, что совпадает с решением, полученным с помощью теоремы двойственности.

6.2. Двойственные оценки как мера дефицитности ресурсов

Ресурсы, полностью используемые в производстве, называются *дефицитными*. Признаком дефицитности ресурса является положительность его объективно обусловленной оценки.

В нашем примере таковыми являются труд и оборудование. Для этих ресурсов двойственные оценки положительны: $y_1^* = 0.636 > 0,\ y_3^* = 0.545 > 0$. Сырье дефицитным не является, у него нулевая двойственная оценка: $y_2^* = 0$. Вспоминая, что экономический смысл двойственных оценок состоит в том, что это цены ресурсов, можно сказать, что в данном примере труд более дефицитен, чем оборудование так как $y_1^* > y_3^*$.

6.3. Анализ эффективности планов производства

Из второй теоремы двойственности, а именно из (26) вытекает, что если j-й вид продукции вошел в оптимальный план производства, т.е. $x_j^* > 0$, то $\sum_{i=1}^{m} a_{ij} y_i^* = c_j$. Если же $\sum_{i=1}^{m} a_{ij} y_i^* > c_j$, то соответствующее $x_j^* = 0$. Таким образом, величина

$$\Delta_j = \sum_{i=1}^{m} a_{ij} y_i^* - c_j \qquad (31)$$

может служить критерием для оценки выгодности или невыгодности производства j-го вида продукции. Если $\Delta_j > 0$, то выручка от продажи ресурсов, идущих на изготовление единицы продукции больше, чем прибыль от реализации готового изделия, следовательно, производство данного вида продукции невыгодно. Если же $\Delta_j \le 0$, то затраты на ресурсы не превышают прибыль и производство данной продукции выгодно.

Таблица 3. Данные для трех видов продукции

Ресурсы	*Объективно обусловленные оценки*	*Затраты ресурсов на производство единицы готовой продукции*		
		P_3	P_4	P_5
S_1 (*Труд*)	0.636	4	7	6
S_2 (*Сырье*)	0	6	4	2
S_3 (*Оборудование*)	0.545	3	1	4
Прибыль от реализации единицы готовой продукции		4	5	5

Пусть нужно решить вопрос о выгодности производства трех видов продукции P_3, P_4, P_5. Затраты ресурсов на них представлены в табл.3. Вычисляем разности (31):

$$\Delta_3 = 4\cdot 0.636 + 6\cdot 0 + 3\cdot 0.545 - 4 = 4.18 - 4 = 0.18 > 0,$$

$$\Delta_4 = 7\cdot 0.636 + 4\cdot 0 + 1\cdot 0.545 - 6 = 5.0 - 6 = -1.0 < 0,$$

$$\Delta_5 = 6\cdot 0.636 + 2\cdot 0 + 4\cdot 0.545 - 5 = 6.0 - 5 = 1.0 > 0.$$

Так как $\Delta_3 > 0$, $\Delta_5 > 0$, то невыгодно производить продукцию P_3 и P_5, а продукцию P_4 производить выгодно, так как $\Delta_4 < 0$.

6.4. Анализ устойчивости двойственных оценок

В рассмотренном числовом примере дефицитными являются труд и оборудование. Если привлечь дополнительные трудовые ресурсы, то прибыль должна возрасти. Возникает вопрос, насколько можно увеличить запас этого ресурса, чтобы он не перешел в разряд недефицитных?

Приведем ЗЛП (19)–(21) к каноническому виду путем введения m дополнительных неизвестных $x_{n+1}, x_{n+2},\ldots,x_{n+m}$. Тогда ограничения (20) в матричном виде запишутся

$$\boldsymbol{AX} = \boldsymbol{b},$$

где

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 1 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & 0 & 0 & \ldots & 1 \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Пусть известен оптимальный план. Разобьем вектор $\boldsymbol{X}$ на два подвектора $\boldsymbol{X}^* > 0$ и $\boldsymbol{X}^0 = 0$. В первый подвектор включим неизвестные, вошедшие в базис оптимального плана, во второй подвектор – остальные неизвестные. Соответственно матрицу **A** разобьем на две. В матрицу $\boldsymbol{A}^*$ включим столбцы матрицы $\boldsymbol{A}$, соответствующие базисным неизвестным в оптимальном плане, в $\boldsymbol{A}^0$ поместим остальные столбцы. Тогда можно записать

$$\boldsymbol{A}^*\boldsymbol{X}^* + \boldsymbol{A}^0\boldsymbol{X}^0 = \boldsymbol{b}.$$

Так как $\boldsymbol{A}^0\boldsymbol{X}^0 = 0$, то

$$\boldsymbol{A}^*\boldsymbol{X}^* = \boldsymbol{b}. \tag{32}$$

Умножим обе части (32) на обратную матрицу $\mathbf{A}^{*-1}$, получим

$$\boldsymbol{X}^* = \boldsymbol{A}^{*-1}\boldsymbol{b} \tag{33}$$

Обозначим

$$\boldsymbol{A}^{*-1} = \boldsymbol{D}, \tag{34}$$

тогда (33) запишется:

$$\boldsymbol{X}^* = \boldsymbol{Db}. \tag{35}$$

Матрица $\boldsymbol{D}$ характеризует влияние запасов ресурсов $\boldsymbol{b}$ на величину оптимального выпуска продукции $\boldsymbol{X}^*$. Пусть запасы ресурсов изменились на $\Delta\boldsymbol{b}$ и составили $\boldsymbol{b}+\Delta\boldsymbol{b}$. Оптимальное решение изменится и будет равным $\boldsymbol{X}^*+\Delta\boldsymbol{X}^*$. Согласно (35)

$$\boldsymbol{X}^* + \Delta\boldsymbol{X}^* = \boldsymbol{D}(\boldsymbol{b} + \Delta\boldsymbol{b}),$$

отсюда, учитывая справедливость (35),

$$\Delta\,\boldsymbol{X}^* = \boldsymbol{D}\;\Delta\,\boldsymbol{b}.$$

Изменения правых частей системы ограничений может вызвать изменение решения и двойственной задачи, поэтому важно знать, какие изменения $\boldsymbol{b}$ не приводят к изменению двойственных оценок $\mathbf{y}$.

Обозначим через $\Delta b_j^{(-)}$ уменьшение b_j, не приводящее к изменению решения двойственной задачи, а через $\Delta b_j^{(+)}$ соответствующее максимально допустимое увеличение. Эти приращения находятся по формулам

$$\Delta b_j^{(-)} = \min_i \left(\frac{x_i^*}{d_{ij}} \right) \text{ для } d_{ij} > 0, \tag{36}$$

$$\Delta b_j^{(+)} = \left| \max_i \left(\frac{x_i^*}{d_{ij}} \right) \right| \text{ для } d_{ij} < 0. \tag{37}$$

Здесь d_{ij} – элементы матрицы $\boldsymbol{D}$.

Итак, изменение i–го ресурса в пределах от $b_j - \Delta b_j^{(-)}$ до $b_j + \Delta b_j^{(+)}$ не меняет двойственные оценки.

Матрица $\boldsymbol{D}$ получается автоматически при решении ЗЛП симплекс–методом. Столбцами $\boldsymbol{D}$ являются столбцы последнего шага симплексной таблицы, соответствующие исходному базису. При этом строки следует переставить так, чтобы номера базисных неизвестных шли в порядке возрастания.

Вернемся к числовому примеру. Вычисления по симплекс-методу показаны в табл.3. Исходный базис образуют неизвестные x_3. x_4, x_5. В табл.5 приведен фрагмент последнего шага симплексной таблицы.

Таблица 5. Последний шаг симплекс-метода

Базис	x_1	x_2	x_3	x_4	x_5
x_4	0	0	-0.682	1	0.773
x_1	1	0	0.182	0	-0.273
x_2	0	1	-0.045	0	0.318

Матрицу $\boldsymbol{D}$ образуют три последних столбца, но строки следует переставить так, чтобы неизвестные в столбце «*Базис»* шли в порядке x_1, x_2, x_4. В результате получаем

$$\boldsymbol{D} = \begin{pmatrix} 0.182 & 0 & -0.273 \\ -0.045 & 0 & 0.318 \\ -0.682 & 1 & 0.773 \end{pmatrix} \tag{38}$$

Матрицей $\boldsymbol{A}$ является матрица системы (4)

$$\boldsymbol{A} = \begin{pmatrix} 7 & 6 & 1 & 0 & 0 \\ 4 & 1 & 0 & 1 & 0 \\ 1 & 4 & 0 & 0 & 1 \end{pmatrix} \tag{39}$$

Так как оптимальное решение есть $x_1^*=8$, $x_2^*=14$, $x_3^*=0$, $x_4^*=18$. $x_5^*=0$, то $\boldsymbol{X}^*=(x_1, x_2, x_4)^T$ и $\boldsymbol{X}^0 = (x_3, x_5)^T$, а матрицу $\boldsymbol{A}^*$ образуют столбцы 1, 2, 4 исходной матрицы ограничений (39):

$$\boldsymbol{A}^* = \begin{pmatrix} 7 & 6 & 0 \\ 4 & 1 & 1 \\ 1 & 4 & 0 \end{pmatrix}.$$

Убедимся, что $\boldsymbol{D}=\boldsymbol{A}^{*-1}$. Для этого умножим $\boldsymbol{D}$ на $\boldsymbol{A}^*$:

$$DA^* = \begin{pmatrix} 0.182 & 0 & -0.273 \\ -0.045 & 0 & 0.318 \\ -0.682 & 1 & 0.773 \end{pmatrix} \begin{pmatrix} 7 & 6 & 0 \\ 4 & 1 & 1 \\ 1 & 4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Произведение $\boldsymbol{DA}^*$ равно единичной матрице, значит, матрица $\boldsymbol{D}$ выписана правильно.

С помощью элементов первого столбца матрицы $\boldsymbol{D}$ находятся пределы изменения для b_1, с помощью второго столбца – для b_2 и т.д. Строкам матрицы соответствуют базисные неизвестные в оптимальном плане. Для упрощения вычислений по формулам (36), (37), запишем данные для них в виде

$$\begin{array}{cccc} b_1 & b_2 & b_3 & \\ \end{array}$$
$$\begin{pmatrix} 0.182 & 0 & -0.273 \\ -0.045 & 0 & 0.318 \\ -0.682 & 1 & 0.773 \end{pmatrix} \begin{array}{l} x_1^* = 8 \\ x_2^* = 14 \\ x_4^* = 18 \end{array}$$

В первом столбце матрицы $\boldsymbol{D}$ только один положительный элемент, поэтому

$$\Delta b_1^{(-)} = min\left(\frac{x_1^*}{d_{11}}\right) = \frac{8}{0.182} = 44.$$

В первом столбце $\boldsymbol{D}$ два отрицательных элемента, поэтому

$$\Delta b_1^{(+)} = \left|max\left(\frac{x_2^*}{d_{21}}, \frac{x_4^*}{d_{31}}\right)\right| = \left|max\left(\frac{14}{-0.045}, \frac{18}{-0.682}\right)\right| =$$
$$|max(-308, -26.4)| = |-26.4| = 26.4.$$

Аналогично вычисляем

$$\Delta b_2^{(-)} = min\left(\frac{x_4^*}{d_{32}}\right) = min\left(\frac{18}{1}\right) = 18.$$

Так как отрицательных элементов во втором столбце нет, то

$$\Delta b_2^{(+)} = +\infty,$$

и сколь угодно большое увеличение запаса второго ресурса не приведет к изменению решения двойственной задачи, так как он является недефицитным и любое увеличение запаса этого ресурса не изменит ситуацию.

Находим пределы изменения запасов третьего ресурса

$$\Delta b_3^{(-)} = min\left(\frac{x_2^*}{d_{23}}, \frac{x_4^*}{d_{33}}\right) = min\left(\frac{14}{0.318}, \frac{18}{0.773}\right) =$$
$$min(44, 23.3) = 23.3,$$

$$\Delta b_3^{(+)} = \left|max\left(\frac{x_1^*}{d_{13}}\right)\right| = \left|max\left(\frac{8}{-0.273}\right)\right| = |max(-29.3)| = 29.3.$$

Итак, интервалы устойчивости двойственных оценок будут следующими:
для первого ресурса

$$\left(b_1 - \Delta b_1^{(-)},\ b_1 + \Delta b_1^{(+)}\right) = (140 - 44,\ 140 + 26.4) = (96,\ 166.4),$$

для второго ресурса $(64 - 18,\ 64 + \infty) = (46,\ +\infty)$,
для третьего ресурса – $(64 - 23.3,\ 64 + 29.3) = (40.7,\ 93.3)$.

6.5. Оценки как мера влияния ограничений на целевую функцию

Величина объективно обусловленной оценки того или иного ресурса показывает, насколько возросло бы максимальное значение целевой функции при увеличении объема данного ресурса на единицу. Это следует из третьей теоремы двойственности.

Пусть удалось увеличить объем первого ресурса на 11 условных единиц, его запас составит тогда 140+11=151. Этот объем не выходит за границы интервала устойчивости первого ресурса (96, 166.4), значит, такое изменение не повлечет изменения двойственных оценок. Увеличение прибыли составит

$$\Delta f = y_1^* \cdot \Delta b_1 = 0.636 \cdot 11 = 7.$$

Новый оптимальный план найдем по формуле (35):

$$\begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} = \boldsymbol{D} \begin{pmatrix} b_1 + \Delta b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0.182 & 0 & -0.273 \\ -0.045 & 0 & 0.318 \\ -0.682 & 1 & 0.773 \end{pmatrix} \begin{pmatrix} 151 \\ 64 \\ 64 \end{pmatrix} = \begin{pmatrix} 10.0 \\ 13.5 \\ 10.5 \end{pmatrix}$$

Таким образом, знание двойственных оценок позволяет получить оптимальное решение при измененных объемах ресурсов без нового решения задачи.

6.6. Оценки как инструмент балансирования суммарных затрат и прибыли

Найдем оптимальное значение целевой функции (18) двойственной задачи

$$g_{min} = g(\boldsymbol{Y}^*) = 140 y_1^* + 64 y_2^* + 64 y_3^* =$$
$$140 \cdot 7/11 + 64 \cdot 0 + 64 \cdot 6/11 = 124.$$

Это значение равно максимальной прибыли f_{max}=124, как того и требует первая теорема двойственности.

Смысл равенства

$$f_{max} = g_{min}$$

можно интерпретировать так: предприятию безразлично, производить ли продукцию по оптимальному плану и получить максимальную прибыль f_{max} или продавать ресурсы по оптимальным ценам и возместить от их продажи равные прибыли минимальные затраты на ресурсы g_{min}.

Глава 3. Теория игр

1. Основные понятия

На практике часто встречаются ситуации, когда приходится принимать решение в условиях конфликта интересов нескольких участников события. Такие ситуации возникают, например, в карточных играх, шахматах и т.п. В экономике конфликтные ситуации возникают при взаимодействии покупателя и продавца, банка и клиента, поставщика и потребителя и т.д. Особенностью подобных ситуаций является неопределенность, вызванная неизвестным заранее поведением участников конфликта, которые стремятся добиться максимальной реализации своих целей. Математическим описанием конфликтных ситуаций занимается *теория игр*. Ее цель – выработка рекомендаций по разумному поведению участников конфликта.

Каждая конкретная конфликтная ситуация очень сложна. Чтобы сделать возможным математический анализ конфликта, отбрасываются несущественные факторы и строится математическая модель, которую называют *игрой*. Стороны, участвующие в конфликте, называются *игроками*, исход игры называется *выигрышем*. От реального конфликта игра отличается наличием формальных *правил*. Правила устанавливают возможные варианты действий игроков и исход игры, то есть выигрыш или проигрыш участников в зависимости от сложившейся обстановки. Будем считать, что выигрыши участников имеют количественное выражение.

Игра называется *парной*, если в ней участвуют два игрока и *множественной*, если число участников более двух.

Игра называется игрой с *нулевой суммой*, если выигрыш одной стороны равен проигрышу другой.

Игра состоит из отдельных *ходов* участников. Различают *личные* и *случайные* ходы. Личный ход – это сознательный выбор игроком одного из вариантов действий. Случайный ход определяется не волей игрока, а случайно, например, подбрасыванием монеты. Игры, в которых возможны личные ходы, называются *стратегическими*. *Стратегией* игрока называется совокупность правил, определяющих выбор личных ходов игрока в зависимости от сложившейся ситуации.

Игра называется *конечной*, если у каждого игрока имеется конечное число стратегий и бесконечной, в противном случае, например, в шахматах число ходов конечно, хотя и весьма велико, а дуэль, в которой участники могут стрелять друг в друга в любой момент данного отрезка времени, является игрой с бесконечным числом ходов.

Оптимальной называется стратегия игрока, обеспечивающая ему максимальный выигрыш. Если игра повторяется многократно и содержит кроме личных и случайные ходы, максимизируется средний выигрыш. При нахождении оптимальных стратегий предполагается, что противник ведет себя разумно и стремится минимизировать свой проигрыш. Решить игру – это значит выбрать стратегию каждого игрока, удовлетворяющую *условию оптимальности*, то есть один игрок должен получить максимальный выигрыш, в то же время второй игрок должен иметь минимальный проигрыш.

2. Платежная матрица игры

Рассмотрим парную конечную игру, в которой участвуют игроки A и B. Пусть игрок A имеет m личных стратегий $A_1, A_2, \ldots, A_m$, а у игрока B есть n личных стратегий $B_1, B_2, \ldots, B_n$. Обозначим a_{ij} выигрыш игрока A, если он выбрал стратегию A_i, в то время как игрок B выбрал стратегию B_j. Таблица выигрышей образует платежную матрицу

Таблица 1. Платежная матрица парной игры

		Стратегии игрока *B*			
		B_1	B_2	…	B_n
Стратегии игрока *A*	A_1	a_{11}	a_{12}	…	a_{1n}
	A_2	a_{21}	a_{22}	…	a_{2n}
	…	…	…	…	…
	A_m	a_{m1}	a_{m2}	…	a_{mn}

Приведем простейший пример игры. Игрок *A* выбирает одну из сторон монеты. Второй игрок *B*, не зная выбора первого, также выбирает одну из ее сторон. После того как оба игрока сделали свой выбор, *A* платит *B* 1, если стороны совпали, и получает 1 в противном случае. В табл. 2 перечислены стратегии и выигрыши данной игры.

Таблица 2. Игра в выбор стороны монеты

		Стратегии игрока *B*	
		B_1 – «Орел»	B_2 – «Решка»
Стратегии игрока *A*	A_1 – «Орел»	-1	1
	A_2 – «Решка»	1	-1

Данная игра полностью определяется следующей платежной матрицей

$$\boldsymbol{P} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

3. Цена игры

Рассмотрим игру $m \times n$ с матрицей

$$\boldsymbol{P} = (a_{ij}), \quad i = \overline{1,m};\ j = \overline{1,n}.$$

Обозначим через α_i наименьший выигрыш игрока *A* при выборе им стратегии A_i

$$\alpha_i = \min_{j=\overline{1,n}} a_{ij}.$$

Эта величина является минимальным элементом *i*-й строки платежной матрицы. Именно такой выигрыш получит игрок *A*, если постоянно будет придерживаться стратегии A_i, так как игрок *B* стремится минимизировать свой проигрыш и соответствующим образом среагирует на действия игрока *A*. Среди всех α_i выберем наибольшее значение

$$\alpha = \max_{i=\overline{1,m}} \alpha_i = \max_{i=\overline{1,m}} \min_{j=\overline{1,n}} a_{ij}.$$

Если игрок *A* будет придерживаться осторожной стратегии и поступать так, чтобы при наихудших для него действиях противника получить максимальный выигрыш, то этот выигрыш составит не менее α. Эта величина представляет собой максимальный среди минимальных выигрышей или *максимин* и называется *нижней ценой игры*.

Игрок *B* стремится минимизировать свой проигрыш. Обозначим через β_j наибольший проигрыш игрока *B* при выборе им стратегии B_j

$$\beta_j = \max_{i=\overline{1,m}} a_{ij}.$$

Величина β_j есть максимальный элемент j-го столбца платежной матрицы. Исходя из принципа осторожности (минимизации своего проигрыша) игрок B должен выбрать стратегию, при которой его максимальный проигрыш минимален:

$$\beta = \min_{i=\overline{1,n}} \beta_j = \min_{i=\overline{1,n}} \max_{j=\overline{1,m}} a_{ij}.$$

Величина β называется *минимаксом* или *верхней ценой* игры. При осторожном подходе игрок B не может проиграть больше β.

Если верхняя и нижняя цена игры совпадают, то их общее значение

$$\alpha = \beta = \nu$$

называется ценой игры. Минимаксные стратегии, соответствующие цене игры, называются *оптимальными чистыми стратегиями*, а их совокупность – оптимальным решением. Про такую игру говорят, что она *решается в чистых стратегиях*.

В табл. 3 приведена платежная матрица некоторой игры, в дополнительных столбцах указаны значения α_i и β_j.

Таблица 3. Платежная матрица с седловой точкой

	B_1	B_2	B_3	B_4	α_i
A_1	2	4	7	5	2
A_2	7	$\boxed{6}$	8	7	$\boxed{6}$
A_3	5	3	4	1	1
β_j	7	$\boxed{6}$	8	7	

Если игрок A будет постоянно придерживаться стратегии A_2, то будет иметь гарантированный выигрыш $\alpha = 6$. В этих условиях игроку B ничего не остается, как придерживаться стратегии B_2, которая ведет к наименьшему его проигрышу $\beta = 6$. В свою очередь, игрок B, стремясь минимизировать свой гарантированный проигрыш, выберет стратегию B_2, а игрок A, максимизируя в этих условиях свой выигрыш, выберет стратегию A_2, таким образом, здесь пара стратегий A_2, B_2 является решением игры, цена игры равна 6.

Элемент платежной матрицы $a_{22} = 6$ имеет минимальное значение во второй строке и максимальное значение во втором столбце. Такой элемент называется *седловой точкой* матрицы. Итак, если платежная матрица игры имеет седловую точку, то игра имеет решение в чистых стратегия.

Рассмотрим теперь игру платежная матрица которой приведена в табл. 4.

Таблица 4. Платежная матрица без седловой точки

	B_1	B_2	B_3	B_4	B_5	α_i
A_1	3	4	5	2	3	2
A_2	1	8	4	3	4	1
A_3	10	3	1	7	6	1
A_4	4	5	3	4	8	$\boxed{3}$
β_j	10	8	$\boxed{5}$	7	8	

Игрок A, пытаясь максимизировать гарантированный выигрыш, выберет стратегию A_4 и в этих условиях игрок B, естественно, выберет стратегию B_3. Выигрыш A и проигрыш B составит 3. Но игрок A заметив, что игрок B придерживается постоянно стратегии B_3, может выбрать стратегию A_1 и увеличить свой выигрыш до 5. В свою очередь игрок B в ответ на это может выбрать стратегию B_4 и снизить свой проигрыш до 2 и так далее. Таким образом, в данной игре не существует решения в чистых стратегиях.

4. Упрощение платежной матрицы

Перед тем, как решать игру $m \times n$ нужно попытаться упростить ее. Это делается путем отбрасывания стратегий, которые заведомо невыгодны игрокам.

Стратегия A_i игрока A называется *доминирующей* над стратегией A_k, если все выигрыши в i-й строке не меньше соответствующих выигрышей в k-й строке и, по крайней мере, один из них действительно больше

$$a_{ij} \geq a_{kj},\ j = \overline{1,n}.$$

Стратегия A_i называется *дублирующей* стратегию A_k, если все выигрыши в i-й строке равны соответствующим выигрышам в k-й строке

$$a_{ij} = a_{kj},\ j = \overline{1,n}.$$

Стратегия B_j игрока B называется *доминирующей* над стратегией B_p, если все выигрыши в j-м столбце не больше соответствующих выигрышей в p-м столбце и, по крайней мере, один из них действительно меньше

$$a_{ij} \leq a_{ip},\ i = \overline{1,m}.$$

Стратегия B_j называется *дублирующей* стратегию B_p, если все выигрыши в j-м столбце равны соответствующим выигрышам в p-м столбце

$$a_{ij} = a_{ip},\ i = \overline{1,m}.$$

Если у какой-то стратегии есть доминирующая или дублирующая, то эту стратегию можно отбросить.

Пусть имеется игра 5×5 с матрицей, приведенной в табл. 5.

Таблица 5

	B_1	B_2	B_3	B_4	B_5
A_1	4	7	2	3	4
A_2	3	5	6	8	9
A_3	4	4	2	2	8
A_4	3	6	1	2	4
A_5	3	5	6	8	9

Замечаем, стратегия A_5 дублирует стратегию A_2, поэтому любую из них можно отбросить. Сравнивая, далее, A_1 и A_4 видим, что A_1 доминирует над A_4 и A_4 можно отбросить. Отбросив стратегии A_5 и A_4, получаем игру 3×5, табл. 6.

Таблица 6

	B_1	B_2	B_3	B_4	B_5
A_1	4	7	2	3	4
A_2	3	5	6	8	9
A_5	3	5	6	8	9

Изучая табл. 6, устанавливаем, что стратегия B_1 доминирует над стратегией B_2, так как все проигрыши игрока B, соответствующие стратегии B_1 меньше проигрышей, соответствующих стратегии B_2. Кроме того, стратегия B_3 доминирует над стратегиями B_4 и B_5. Отбросив стратегии, у которых есть доминирующие, то есть B_2, B_4 и B_5, получим игру 3×2, табл. 7.

Таблица 7

	B_1	B_3
A_1	4	2
A_2	3	6
A_5	3	6

Теперь видно, что стратегии A_2 и A_5 дублируют друг друга и, отбросив стратегию A_5, окончательно получаем игру 2×2, табл. 8.

Таблица 7

	B_1	B_3
A_1	4	2
A_2	3	6

5. Решение игр в смешанных стратегиях

Если игра не имеет седловой точки, то применение чистых стратегий не дает оптимального решения игры. В этом случае можно получить оптимальное решение случайным образом чередуя чистые стратегии.

Смешанной стратегией S_A игрока A называется применение чистых стратегий A_1, A_2, … ,A_m с вероятностями p_1, p_2, … , p_m, причем

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

Обозначим

$$S_A=(p_1, p_2, \ldots , p_m).$$

Смешанную стратегию игрока B обозначим

$$S_B=(q_1, q_2, \ldots , q_n)$$

На основании принципа минимакса определяется оптимальное решение. Это пара смешанных стратегий S_A^*, S_A^* такая, что если один из игроков придерживается своей оптимальной стратегии, то другому не может быть выгодно отступать от своей. Выигрыш, соответствующий оптимальному решению, называется ценой игры ν. Цена игры удовлетворяет неравенству

$$\alpha \le \nu \le \beta .$$

Приведем без доказательства основную теорему теории игр.

Теорема Неймана (основная теорема теории игр)
Каждая конечная игра имеет, по крайней мере, одно оптимальное решение, возможно среди смешанных стратегий.

Пусть $S_A^*=(p_1, p_2, \ldots , p_m)$, $S_B^*=(q_1, q_2, \ldots , q_n)$ – пара оптимальных стратегий. Если чистая стратегия входит в смешанную с ненулевой вероятностью, то она называется *активной*.

Активные стратегии обладают свойством, выражаемым следующей теоремой.

Теорема об активных стратегиях

Если один из игроков придерживается своей оптимальной смешанной стратегии, то выигрыш остается неизменным и равным цене игры ν, если второй игрок не выходит за пределы своих активных стратегий.

6. Приведение матричной игры к задаче линейного программирования

Пусть игра задана платежной матрицей (a_{ij}). Найдем оптимальную стратегию игрока A

$$S_A^*=(p_1, p_2, \dots , p_m),$$

причем

$$p_1 + p_2 + \dots + p_m=1. \quad (1)$$

Оптимальная стратегия S_A^* обеспечивает игроку A выигрыш не меньший чем цена игры ν и выигрыш равный ν, если игрок B придерживается своей оптимальной стратегии.

Будем считать, что все выигрыши a_{ij} положительны. Этого всегда можно добиться, добавляя ко всем членам матрицы достаточно большое число M; от этого цена игры увеличится на M, а решение S_A^*, S_B^* не изменится. Если все a_{ij} положительны, то и цена игры, то есть средний выигрыш при оптимальной стратегии тоже положителен: $\nu > 0$.

Пусть игрок B использует чистую стратегию B_1, а игрок A использует оптимальную стратегию S_A^*, то есть выбирает чистые стратегии $A_1, A_2, \dots, A_m$ с вероятностями $p_1, p_2, \dots, p_m$. Средний выигрыш игрока A (математическое ожидание выигрыша)

$$a_{11}p_1 + a_{21}p_2 +\dots+ a_{m1}p_m$$

не может быть меньше цены игры ν и это справедливо для каждой из чистых стратегий B_1, $B_2,\dots, B_n$ игрока B т.е.

$$\left.\begin{array}{l} a_{11}p_1 + a_{21}p_2 + \dots + a_{m1}p_m \ge \nu, \\ a_{12}p_1 + a_{22}p_2 + \dots + a_{m2}p_m \ge \nu, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{1n}p_1 + a_{2n}p_2 + \dots + a_{mn}p_m \ge \nu. \end{array}\right\} \quad (2)$$

Каждое из неравенств поделим на ν и введем обозначения

$$y_1 = p_1/\nu, \quad y_2 = p_2/\nu, \dots, y_m = p_m/\nu. \quad (3)$$

В этих обозначениях система (2) запишется

$$\left.\begin{array}{l} a_{11}y_1 + a_{21}y_2 + \dots + a_{m1}y_m \ge 1, \\ a_{12}y_1 + a_{22}y_2 + \dots + a_{m2}y_m \ge 1, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{1n}y_1 + a_{2n}y_2 + \dots + a_{mn}y_m \ge 1. \end{array}\right\} \quad (4)$$

Поделим также (1) на ν, получим

$$y_1 + y_2 + \dots + y_m = 1/\nu = g. \quad (5)$$

Игрок A стремится максимизировать значение ν или, что то же самое, минимизировать значение $g = 1/\nu$. Таким образом, приходим к следующей задаче: *нужно подобрать такие значения неотрицательных переменных* $y_1, y_2,\dots, y_m$, *которые удовлетворяют ограничениям* (4) *и дают минимальное значение функции* (5). Такая задача, как известно, является задачей линейного программирования.

Оптимальная стратегия игрока B ищется аналогично. Пусть игрок B придерживается своей оптимальной стратегии $S_B^*=(q_1, q_2,\dots, q_n)$, а игрок A выбирает чистые стратегии, тогда средний проигрыш игрока B не может быть больше цены игры ν, т.е.

$$\left.\begin{array}{l} a_{11}q_1 + a_{12}q_2 + \ldots + a_{1n}q_n \le \nu, \\ a_{21}q_1 + a_{22}q_2 + \ldots + a_{2n}q_n \le \nu, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}q_1 + a_{m2}q_2 + \ldots + a_{mn}q_n \le \nu. \end{array}\right\} \quad (6)$$

Вероятности $q_1, q_2, \dots , q_n$ удовлетворяют условию

$$q_1 + q_2 + \dots + q_n=1. \quad (7)$$

Поделив (6) и (7) на ν и введя обозначения

$$x_1 = q_1/\nu, \quad x_2 = q_2/\nu, \dots, x_n = q_n/\nu. \quad (8)$$

получим

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le 1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le 1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le 1, \end{array}\right\} \quad (9)$$

$$x_1 + x_2 + \ldots + x_n = 1/\nu = f. \quad (10)$$

Цель игрока B состоит в минимизации проигрыша ν или в максимизации величины $f= 1/\nu$. Таким образом, задача состоит в нахождении неотрицательных переменных $x_1, x_2,\dots, x_n$, дающих максимальное значение функции (10) при выполнении ограничений (9).

Сравнивая две задачи линейного программирования (9), (10) и (4), (5) замечаем, что для записи ограничений используется одна и та же матрица

$$P = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

только в системе (9) используются строки матрицы P, а в системе (4) столбцы. Составим схему соответствия задач линейного программирования для игроков A и B, рис. 1.

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="4">Переменные задачи для игрока B</td><td colspan="2" rowspan="2">Ограничения задачи игрока B</td><td rowspan="2">Целевая функция задачи игрока A</td></tr>
<tr><td>x_1</td><td>x_2</td><td>…</td><td>x_n</td></tr>
<tr><td rowspan="4">Переменные задачи для игрока A</td><td>y_1</td><td>a_{11}</td><td>a_{12}</td><td>…</td><td>a_{1n}</td><td>$\le$</td><td>1</td><td rowspan="4">$g=y_1+y_2+\ldots+y_n\to\min$</td></tr>
<tr><td>y_2</td><td>a_{21}</td><td>a_{22}</td><td>…</td><td>a_{2n}</td><td>$\le$</td><td>1</td></tr>
<tr><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td></tr>
<tr><td>y_m</td><td>a_{m1}</td><td>a_{m2}</td><td>…</td><td>a_{mn}</td><td>$\le$</td><td>1</td></tr>
<tr><td colspan="2" rowspan="2">Ограничения задачи для игрока B</td><td>$\ge$</td><td>$\ge$</td><td>…</td><td>$\ge$</td></tr>
<tr><td>1</td><td>1</td><td>…</td><td>1</td></tr>
<tr><td colspan="2">Целевая функция задачи игрока B</td><td colspan="4">$f=x_1+x_2+\ldots+x_n\to\max$</td></tr>
</table>

Рис.1. Связь задач линейного программирования для двух игроков

Из рисунка видно, что задачи линейного программирования, сформулированные для двух игроков, являются взаимно двойственными. Решив одну из них, используя связь двойственных задач, можно найти решение другой. Для решения следует выбирать ту,которая проще.

7. Пример

Предприятие может выпускать три вида продукции A_1, A_2, A_3. Спрос на продукцию предприятия зависит от многих факторов и может находиться в одном из состояний B_1, B_2, B_3, B_4. Известна прибыль, которую приносит реализация каждого вида продукции при различных состояниях спроса, табл.8.

Таблица 8. Прибыль при различных состояниях спроса

		Состояния спроса			
		B_1	B_2	B_3	B_4
Виды продукции	A_1	3	3	6	8
	A_2	9	10	4	2
	A_3	7	7	5	4

Эта экономическая задача сводится к игре предприятия (игрок A), которое должно определить вид выпускаемой продукции против спроса (игрок B)

Анализируя стратегии игроков, замечаем, что стратегия B_1 доминирует над стратегией B_2, так как все элементы столбца B_1 меньше соответствующих элементов столбца B_2, поэтому столбец B_2 можно исключить из рассмотрения, в результате получаем платежную матрицу 3×3

$$P=\begin{pmatrix}3 & 6 & 8\\ 9 & 4 & 2\\ 7 & 5 & 4\end{pmatrix}$$

Определим нижнюю и верхнюю цену игры, табл.9.

Таблица 9. Цена игры

	B_1	B_2	B_3	α_i
A_1	3	6	8	$\boxed{3}$
A_2	9	4	2	2
A_3	7	5	4	4
β_j	9	$\boxed{6}$	8	

Нижняя и верхняя цены игры не совпадают:

$$\alpha = 3 < \beta = 6,$$

поэтому матрица не имеет седловой точки, а игра не имеет решения в чистых стратегиях. Найдем смешанные стратегии

$$S_A^*=(p_1, p_2, p_3,) \text{ и } S_B^*=(q_1, q_2, q_3,).$$

Введем обозначения $y_i = p_i/\nu$, i=1, 2, 3; $x_j = q_j/\nu$, j=1, 2, 3 и, пользуясь схемой на рис.1, составим две задачи линейного программирования.

Игрок *A* (предприятие)

$$g=y_1+ y_2+y_3\to\min,$$

$$\left.\begin{aligned} 3y_1+9y_2+7y_3\ge 1,\\ 6y_1+4y_2+5y_3\ge 1,\\ 8y_1+2y_2+4y_3\ge 1,\end{aligned}\right\}$$

$$y_i \ge 0,\ i = 1, 2, 3.$$

Игрок *B* (спрос)

$$f=x_1+x_2+x_3\to\max$$

$$\left.\begin{aligned} 3x_1+6x_2+8x_3\le 1,\\ 9x_1+4x_2+2x_3\le 1,\\ 7x_1+5x_2+4x_3\le 1,\end{aligned}\right\}$$

$$x_j \ge 0,\ j = 1, 2, 3.$$

Приведем ограничения этих задач к виду равенств, вводя дополнительные неизвестные:

$$\left.\begin{aligned} 3y_1+9y_2+7y_3-y_4 &=1,\\ 6y_1+4y_2+5y_3-y_5 &=1,\\ 8y_1+2y_2+4y_3-y_6 &=1,\end{aligned}\right\},\quad y_i\ge 0,\quad i=4,5,6; \tag{11}$$

$$\left.\begin{aligned} 3x_1+6x_2+8x_3+x_4 &=1,\\ 9x_1+4x_2+2x_3+x_5 &=1,\\ 7x_1+5x_2+4x_3+x_6 &=1,\end{aligned}\right\}\quad x_j\ge 0,\quad j=4,5,6. \tag{12}$$

Применение симплекс метода к решению задачи линейного программирования начинается с отыскания допустимого базисного решения. Это решение, обычно, находится с помощью зануления основных переменных задачи. Положим в (11)

$$y_1 = y_2 = y_3 = 0,$$

получим

$$y_4 = -1,\ y_5 = -1,\ y_6 = -1,$$

то есть данное базисное решение не является допустимым. Зануляя в (12) основные переменные, видим, что получается допустимое решение

$$x_1 = x_2 = x_3 = 0,\ x_4 = 1,\ x_5 = 1,\ x_6 = 1,$$

поэтому задача для игрока *B* решается проще. Решение ее симплекс методом приведено в табл.10.

Таблица 10. Симплексная таблица для игрока *B*

№ шага		Ба-зис	c_j ба-зиса	c_1=1	c_2=1	c_3=1	c_4=0	c_5=0	c_6=0	b_i	$\frac{b_i}{a_{ik}}$
				x_1	x_2	x_3	x_4	x_5	x_6		
1	1	x_4	0	3	$\boxed{6}$	8	1	0	0	1	<u>1/6</u>
	2	x_5	0	9	4	2	0	1	0	1	1/4
	3	x_6	0	7	5	4	0	0	1	1	1/5
	4	c_j–Z_j		1	<u>1</u>	1	0	0	0	0	
2	1	x_2	1	1/2	1	4/3	1/6	0	0	1/6	1/3
	2	x_5	0	7	0	-10/3	-2/3	1	0	1/3	1/7
	3	x_6	0	$\boxed{9/2}$	0	-8/3	-5/6	0	1	1/6	<u>1/27</u>
	4	c_j–Z_j		<u>1/2</u>	0	-4/3	-1/6	0	0	1/6	
3	1	x_2	1	0	1	44/27	7/27	0	-1/9	4/27	1/7
	2	x_5	0	0	0	22/27	7/27	1	-14/9	2/27	3/10
	3	x_1	1	1	0	-16/27	-5/27	0	2/9	1/27	<u>0</u>
	4	c_j–Z_j		0	0	-18/27	-2/27	0	-1/9	5/27	
							$-y_1$	$-y_2$	$-y_3$		

Итак, имеем следующее решение задачи линейного программирования, определяющее оптимальную смешанную игрока *B*:

$$f = 1/\nu = 5/27,\ x_1 = 1/27,\ x_2 = 4/27,\ x_3 = 0.$$

Отсюда

$$\nu = 1/f = 27/5,\ q_1 = x_1\,\nu = 1/5,\ \ q_2 = x_2\,\nu = 4/5,\ \ q_3 = x_3\,\nu = 0.$$

Решение двойственной задачи равно взятым с противоположным знаком значениям оценок на завершающем шаге симплекс метода. Таким образом, для игрока *A* решение следующее:

$$y_1 = 2/27,\ \ y_2 = 0,\ y_1 = 1/9.$$

Отсюда

$$p_1 = y_1\,\nu = 2/5,\ \ p_2 = y_2\,\nu = 0,\ \ p_3 = y_3\,\nu = 3/5.$$

Выполняется равенство

$$p_1 + p_2 + p_3 = 2/5 + 0 + 3/5 = 1.$$

Итак, продукцию A_2 производить не следует, доля продукции A_1 должна составлять в среднем 2/5, а продукции A_3 3/5 от общего объема выпуска, средняя прибыль составит 27/5 = 5.4

Глава 4. Динамическое программирование

1. Метод динамического программирования

Динамическое программирование есть особый метод оптимизации, приспособленный для нахождения оптимальных управлений в *многошаговых* (или многоэтапных) процессах. Например, деятельность некоторой отрасли производства в течение ряда лет естественным образом распадается на временные этапы. В других случаях, например, при распределении ограниченного объема капиталовложений между несколькими предприятиями разбиение на этапы вводится искусственно.

Пусть управляемый процесс состоит из m шагов. Эффективность процесса характеризуется показателем W – *выигрышем*. Пусть общий выигрыш складывается из выигрышей на отдельных шагах

$$W = \sum_{i=1}^{m} w_i \,, \tag{1}$$

где w_i – выигрыш на i–м шаге. Ход и исход каждого шага зависит от *параметра управления* x_i. Обозначим

$$\boldsymbol{x} = \left(x_1, x_2, \ldots, x_m\right) \tag{2}$$

совокупность управлений. Требуется найти такое управление $\boldsymbol{x}$, чтобы выигрыш был максимальным.

Например, владелец автомашины эксплуатирует ее в течение m лет. В начале каждого года он может принять одно из трех решения:

1) Продать машину и заменить ее новой;
2) Отремонтировать и продолжить эксплуатацию;
3) Продолжить эксплуатацию без ремонта.

Здесь выигрыш – сумма затрат на эксплуатацию

$$W = \sum_{i=1}^{m} w_i \,,$$

где w_i – затраты в i–м году. Величину W следует минимизировать.

Многошаговые задачи проще решать, находя не сразу все неизвестные управления, а шаг за шагом, находя отдельные решения для каждого шага. Планируя многошаговый процесс надо выбирать управление на каждом шаге с учетом всех его будущих последствий на еще предстоящих шагах. Управление на i–м шаге выбирается не так, чтобы выигрыш именно на этом шаге был максимальным, а так, чтобы была максимальна сумма выигрышей на всех оставшихся до конца процесса шагах плюс данный шаг.

Среди всех шагов последний может планироваться сам по себе, без оглядки на остальные шаги, так как их нет. При планировании последнего шага надо рассмотреть все возможные результаты предпоследнего шага и для каждого из них найти *условное оптимальное управление* для последнего m–го шага. После этого, рассматривая возможные исходы m-2–го шага, нужно найти условные оптимальные управления для m-1–го шага. Так, двигаясь в обратном направлении от конца процесса к его началу, можно построить совокупность условных оптимальных управлений. Затем надо двигаться от начального состояния системы S_0, выбирая определенные ранее условные оптимальные управления.

Таким образом, метод динамического программирования предполагает два этапа – условная оптимизация от конца к началу и переход от начала к концу с выбором оптимального решения.

2. Пример. Задача о наборе высоты и скорости

В качестве конкретного примера, рассмотрение которого поможет понять суть метода динамического программирования, разберем задачу о наборе самолетом необходимой высоты и скорости полета. Пусть самолет, находящийся на высоте H_0 и имеющий скорость V_0 должен подняться на высоту H_k и приобрести скорость V_k. Известен расход горючего при подъеме с любой высоты H_1 на любую высоту H_2 при постоянной скорости, а также расход горючего на увеличение скорости от V_1 до V_2 при известной высоте.

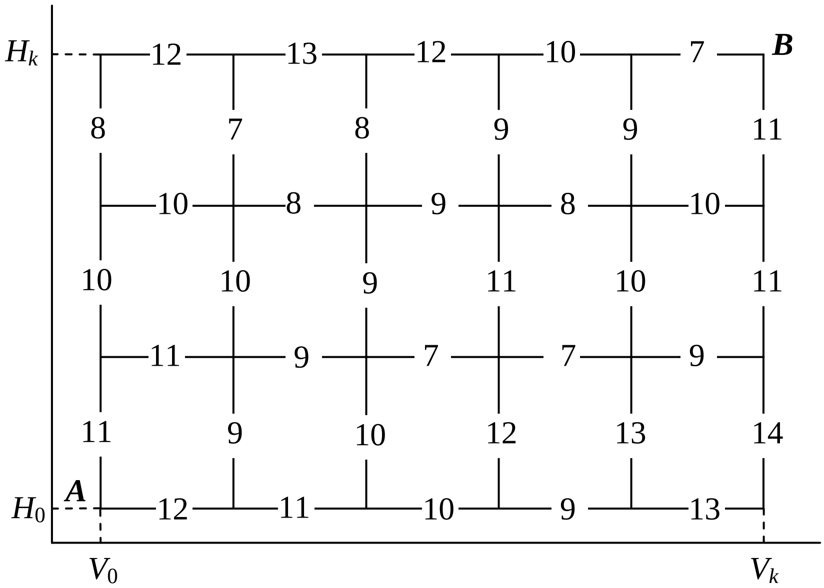

Рис. 1. Расход горючего при наборе высоты и скорости

Согласно законам аэродинамики лучше разгоняться на большей высоте, где плотность воздуха и, следовательно, сопротивление меньше. Набирать же высоту лучше при большей скорости, так как подъемная сила крыла при этом больше. Таким образом, высота и скорость – противоречивые цели.

Пусть скорость может увеличиваться пятью шагами, а высота – тремя (рис.1). На горизонтальных линиях рисунка укажем затраты горючего на увеличение скорости, а на вертикальных – затраты горючего на набор высоты.

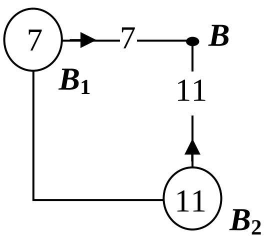

Рис.2. Последний шаг

Переход от начального состояния ***A*** к конечному состоянию ***B*** можно сделать многими путями. Нужно выбрать такой путь, которому соответствует минимальный расход топлива.

Процесс перехода состоит из 8 шагов: 5 по горизонтали и 3 по вертикали. Рассмотрим последний, 8–й шаг. В положение ***B*** можно перейти из двух узлов $\boldsymbol{B_1}$ и $\boldsymbol{B_2}$ (рис.2). Если 8-й шаг будет из $\boldsymbol{B_1}$, то затраты составят 7, а если из $\boldsymbol{B_2}$, то 11. Эти значения запишем в кружках и стрелками покажем направления движения из узлов.

Предпоследний, 7–й шаг может выполниться из трех узлов $\boldsymbol{C_1}$, $\boldsymbol{C_2}$, $\boldsymbol{C_3}$ (рис.3). В кружках поместим минимальные затраты при переходе из каждого из узлов. Существует единственный вариант перехода из узлов $\boldsymbol{C_1}$ и $\boldsymbol{C_3}$ в узел ***B***. Для узла $\boldsymbol{C_2}$ есть два варианта, поэтому выбираем 7–й шаг так, чтобы вместе с 8–м шагом получился минимальный расход топлива. Путь из $\boldsymbol{C_2}$ через $\boldsymbol{B_1}$ потребует затрат в количестве 9+7=16, а через $\boldsymbol{B_2}$: 10+11=21. Оптимальным вариантом перехода из $\boldsymbol{C_2}$ в ***B*** оказывается путь через $\boldsymbol{B_1}$.

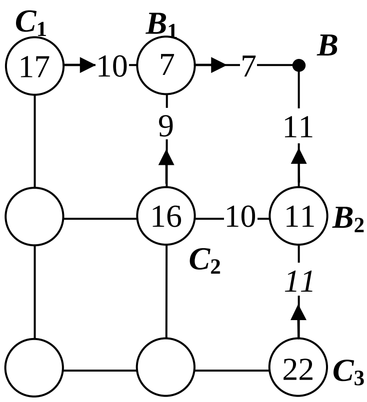

Рис.3. Предпоследний шаг

То же самое проделаем со всеми узлами. Результат представлен на рис.4. Когда дойдем до узла ***A***, становится понятно, что минимальный расход топлива составит 70 и ясен оптимальный путь из ***A*** в ***B***. На рис.4 оптимальный путь отмечен двойными кружками.

Если же непосредственно двигаться из ***A*** в ***B***, выбирая при каждом шаге минимальный

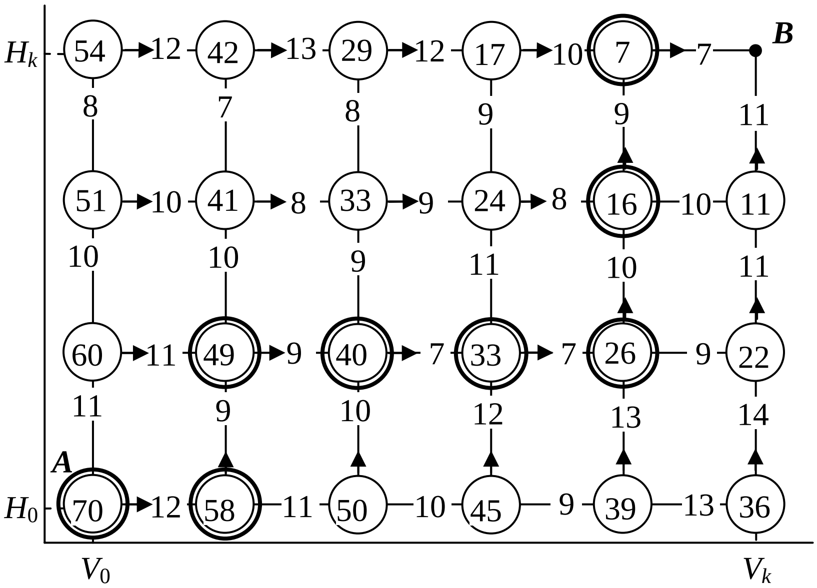

Рис.4. Оптимальное решение

расход, получили бы 11+10+8+12+13+12+10+7=83.

Итак, решение задачи оптимизации методом динамического программирования состоит из двух этапов. На первом этапе процесс рассматривается от его конца к началу, и строятся условные оптимальные управления. На втором этапе процесс рассматривается от начала к концу и ищется безусловное оптимальное управление.

3. Задача о распределении ресурсов

Пусть имеется запас ресурсов K, который нужно распределить между m предприятиями $П_1$, $П_2$, …, $П_m$. Каждое из предприятий $П_i$ при вложении в него средств x приносит доход $\varphi_i(x)$.Следует так распределить средства K между предприятиями, чтобы получить наибольший доход.

Здесь задача не разделяется естественным образом на последовательно выполняемые шаги, поэтому введем такое разделение искусственно.

Будем считать первым шагом процесса вложение средств в $П_1$, вторым – в $П_2$ и т.д.

Управляемая система в данном случае – средства, которые распределяются. Состояние системы перед каждым шагом характеризуется одним числом S – наличным запасом еще не вложенных средств. Шаговыми управлениями являются средства x_1, x_2, …, x_m, выделяемые предприятиям. Требуется найти оптимальное управление, т.е. совокупность чисел x_1, x_2, …, x_m, при которой суммарный доход максимален

$$W=\sum_{i=1}^{m}\varphi_i(x_i)\Rightarrow max\,. \qquad (3)$$

Найдем для каждого i–го шага условный оптимальный выигрыш (от этого шага и до конца), если в начале i–го шага был запас средств S. Обозначим условный оптимальный выигрыш $W_i(S)$, а соответствующее ему условное оптимальное управление, то есть средства, вкладываемые в i–е предприятие, $x_i(S)$.

На последнем m–м шаге

$$x_m(S) = S,$$

т.е. все оставшиеся нераспределенными средства вкладываются в последнее предприятие. Условный оптимальный выигрыш равен просто прибыли, получаемой последним предприятием

$$W_m(S) = \varphi_m(S).$$

Перейдем к предпоследнему, $(m-1)$–му шагу. Пусть мы подошли к нему с запасом средств S. Обозначим $W_{m-1}(S)$ условный оптимальный выигрыш на двух последних шагах: $(m-1)$–м и m–м. Если на $(m-1)$–м шаге $(m-1)$–му предприятию будут выделены средства x, то на последний шаг останется $S-x$. Выигрыш на двух последних шагах составит

$$\varphi_{m-1}(x) + W_m(S-x).$$

Нужно найти такое x, при котором этот выигрыш максимален

$$W_{m-1}(S) = \max_{x \le S}\{\varphi_{m-1}(x) + W_m(S-x)\}. \qquad (4)$$

Далее оптимизируем $(m-2)$–й, $(m-3)$–й и т.д. шаги. Для любого i–го шага будем находить условный оптимальный выигрыш за все шаги с этого шага и до конца по формуле

$$W_i(S) = \max_{x \le S}\{\varphi_i(x) + W_{i+1}(S-x)\} \qquad (5)$$

и соответствующее ему оптимальное управление $x_i(S)$– то значение x, при котором этот максимум достигается.

Продолжая анализ шагов таким образом, дойдем до первого предприятия $П_1$. Здесь нам не нужно варьировать значения S, эта величина задана и равна K. Оптимальный выигрыш составит

$$W^* = W_1(K) = \max_{x \le K}\{\varphi_1(x) + W_2(K-x)\} \qquad (6)$$

То значение x, при котором достигается максимум (6), и есть оптимальное управление x_1^* на первом шаге. После вложения средств x_1^* в первое предприятие их останется $K - x_1^*$. Используя найденную ранее зависимость условно оптимального управления для второго шага $x_2(S)$, находим безусловное оптимальное управление для второго шага,

$$x_2^* = x_2\left(K - x_1^*\right),$$

затем, аналогично, для третьего шага

$$x_3^* = x_3\left(K - x_1^* - x_2^*\right),$$

и т.д.

4. Принцип оптимальности

Общая идея метода динамического программирования состоит в том, что в одном направлении проводится *пошаговая условная оптимизация*, а в другом направлении – *безусловная оптимизация*. Критерий выбора условно оптимальных решений состоит в следующем *принципе оптимальности*, сформулированным американским математиком Р. Беллманом (1920-1985), автором метода динамического программирования:

Каково бы ни было состояние системы S перед очередным шагом, надо выбирать управление на этом шаге так, чтобы выигрыш на данном шаге плюс оптимальный выигрыш на всех последующих шагах был максимальным.

Математически данный принцип выражается уравнением (5). Уравнение (5) называется уравнением Беллмана.

5. *Числовой пример для задачи о распределении ресурсов*

Пусть исходный запас средств K=7 и требуется его распределить между четырьмя предприятиями (m=4). Количество средств, которое можно вложить в каждое предприятие кратно $\Delta x = 1$, то есть можно вложить только целое количество средств, кроме того объем средств, вкладываемых в каждое предприятие не должен превышать $d = 5$ единиц. Функции дохода заданы в таблице 1.

Таблица 1.

x	$\varphi_1(x)$	$\varphi_2(x)$	$\varphi_3(x)$	$\varphi_4(x)$
1	0.1	0.6	0.3	1.0
2	0.5	1.1	0.6	1.2
3	1.2	1.2	1.3	1.3
4	1.8	1.4	1.4	1.3
5	2.5	1.6	1.5	1.3

5.1. Математическая модель задачи

Требуется найти такие объемы средств, выделяемые предприятиям x_1, x_2, x_3, x_4, чтобы весь запас средств был распределен:

$$x_1 + x_2 + x_3 + x_4 = K,$$

а общий выигрыш был максимален

$$W^* = max\{\varphi_1(x_1) + \varphi_2(x_2) + \varphi_3(x_3) + \varphi_4(x_4)\}.$$

5.2. Построение условно оптимальных решений

Проведем условную оптимизацию, начиная с последнего 4-го шага. Надо рассмотреть все возможные значения S и найти соответствующие $x_4(S)$ и $W_4(S)$. К последнему шагу мы должны подойти с запасом средств $S \le d = 5$. На последнем шаге все средства выделяются четвертому предприятию, поэтому

$$x_4(S)=S, \quad W_4(S)= \varphi_4(S),$$

то есть условный оптимальный выигрыш на 4-м шаге равен доходу 4-го предприятия от вложенных в него средств.

Рассмотрим третий шаг. Пусть объем распределяемых средств на третьем шаге S=1. Количество средств x, вкладываемых в третье предприятие, может иметь значения 0 и 1. Согласно (4)

$$W_3(1) = \max_{x \le 1}\{\varphi_3(x) + W_4(1-x)\}. \tag{7}$$

Вычислим значения выражения в фигурных скобках:

$$\text{при } x=0, \ \varphi_3(0)+W_4(1) = 0+1 = 1;$$

$$\text{при } x=1, \ \varphi_3(1)+W_4(0) = 0.3+0 = 0.3.$$

Отсюда видно, что правая часть (7) максимальна при x=0, Значит $W_3(1)$=1, $x_3(1)$=0. Для построения условного оптимального решения на третьем шаге следует рассмотреть все возможные значения для S, то есть 0, 1, …, 7. Соответствующие расчеты разместим в табл.2.

Таблица 2. Условно оптимальные решения для третьего шага

S	x	$S–x$	$\varphi_3(x)$	$W_4(S–x)$	$\varphi_3(x)+ W_4(S–x)$	Условно оптимальное решение	
						$x_3(S)$	$W_3(S)$
1	2	3	4	5	6	7	8
1	***0***	1	0	1.0	***1.0***	0	1.0
	1	0	0.3	0	0.3		
2	0	2	0	1.2	1.2		
	1	1	0.3	1.0	***1.3***	1	1.3
	2	0	0.6	0	0.6		
3	0	3	0	1.3	1.3		
	1	2	0.3	1.2	1.5		
	2	1	0.6	1.0	***1.6***	2	1.6
	3	0	1.3	0	1.3		
4	0	4	0	1.3	1.3		
	1	3	0.3	1.3	1.6		
	2	2	0.6	1.2	1.8		
	3	1	1.3	1.0	***2.3***	3	2.3
	4	0	1.4	0	1.4		
5	0	5	0	1.3	1.3		
	1	4	0.3	1.3	1.6		
	2	3	0.6	1.3	1.9		
	3	2	1.3	1.2	***2.5***	3	2.5
	4	1	1.4	1.0	2.4		
	5	0	1.5	0	1.5		
6	1	5	0.3	1.3	1.6		
	2	4	0.6	1.3	1.9		
	3	3	1.3	1.3	***2.6***		2.6
	4	2	1.4	1.2	***2.6***	3, 4	
	5	1	1.5	1.0	2.5		
7	2	5	0.6	1.3	1.9		
	3	4	1.3	1.3	2.6		
	4	3	1.4	1.3	***2.7***	4, 5	2.7
	5	2	1.5	1.2	***2.7***		

Поиск условно оптимальных решений состоит в нахождении максимального значения в столбце 6 табл.2 для каждого S. Эти максимальные значения выделены, выделены также и соответствующие им значения x. Найденные условно оптимальные решения выписаны в столбцах 7 и 8.

Отметим, что в некоторых случаях максимум достигается при нескольких значениях x, однако для динамического программирования это не является большой проблемой, так как важно лишь значение максимума.

При построении условно оптимальных решений для S = 6 и 7 учитываем ограничение, что объем ресурсов, выделяемых одному предприятию, не превышает d = 5. Это значит, что

$$x \le 5 \quad \text{и} \quad S - x \le 5.$$

Вычисления по нахождению условно оптимального решения на втором шаге поместим в табл.3. Значения выигрыша на втором шаге $\varphi_2(x)$ берем из исходной табл.1, а условно оптимальный выигрыш на третьем шаге $W_3(S–x)$ – из табл. 2.

Таблица 3.Условно оптимальные решения для второго шага

S	x	$S–x$	$\varphi_2(x)$	$W_3(S–x)$	$\varphi_2(x)+ W_3(S–x)$	Условно оптимальное решение	
						$x_2(S)$	$W_2(S)$
1	2	3	4	5	6	7	8
1	***0***	1	0	1.0	***1.0***	0	1.0
	1	0	0.6	0	0.6		
2	0	2	0	1.3	1.3		
	1	1	0.6	1.0	***1.6***	1	1.6
	2	0	1.1	0	1.1		
3	0	3	0	1.6	1.6		
	1	2	0.6	1.3	1.9		
	2	1	1.1	1.0	***2.1***	2	2.1
	3	0	1.4	0	1.4		
4	0	4	0	2.3	2.3		
	1	3	0.6	1.6	2.2		
	2	2	1.1	1.3	***2.4***	2	2.4
	3	1	1.2	1.0	2.2		
	4	0	1.4	0	1.4		
5	0	5	0	2.5	2.5		
	1	4	0.6	2.3	***2.9***	1	2.9
	2	3	1.1	1.6	2.7		
	3	2	1.2	1.3	2.5		
	4	1	1.4	1.0	2.4		
	5	0	1.6	0	1.6		
6	0	6	0	2.6	2.6		
	1	5	0.6	2.5	3.1		
	2	4	1.1	2.3	***3.4***	2	3.4
	3	3	1.2	1.6	2.8		
	4	2	1.4	1.3	2.7		
	5	1	1.6	1.0	2.6		
7	0	7	0	2.7	2.7		
	1	6	0.6	2.6	3.2		
	2	5	1.1	2.5	***3.6***	2	3.6
	3	4	1.2	2.3	3.5		
	4	3	1.4	1.6	3.0		
	5	2	1.6	1.3	2.9		

При построении условного оптимального управления на первом шаге рассматриваем только один случай

$$S = K = 7.$$

Таблица 4. Условно оптимальные решения для первого шага

x	K-x	$\varphi_1(x)$	$W_2(K–x)$	$\varphi_1(x)+W_2(K–x)$	Оптимальное решение	
					x_1^*	W^*
0	7	0	3.6	3.6	5	4.1
1	6	0.1	3.4	3.5		
2	5	0.5	2.9	3.4		
3	4	1.2	2.4	3.6		
4	3	1.8	2.1	3.9		
5	2	2.5	1.6	***4.1***		

5.3. Нахождение безусловного оптимального управления

Результаты оптимизации по всем шагам сведем в табл.5.

Таблица 5. Результаты условной оптимизации. *K=7*

S	i=4		i=3		i=2		i=1	
	$x_4(S)$	$W_4(S)$	$x_3(S)$	$W_3(S)$	$x_2(S)$	$W_2(S)$	$x_1(S)$	$W_1(S)$
1	***1***	1.0	***0***	1.0	0	1.0		
2	2	1.2	1	1.3	***1***	1.6		
3	3	1.3	2	1.6	2	2.1		
4	4	1.3	3	2.3	2	2.4		
5	5	1.3	3	2.5	1	2.9		
6			3, 4	2.6	2	3.4		
7			4, 5	2.7	2	3.6	***5***	***4.1***

В табл.5 S означает запас ресурсов, с которым мы приходим к очередному шагу. Из таблицы видно, что безусловный оптимальный выигрыш есть

$$W^*=4.1,$$

при этом первому предприятию следует выделить 5 единиц ресурсов из 7. Значит, в начале второго шага мы будем иметь 7 – 5 = 2 единицы ресурсов. Из строки таблицы для S = 2 находим, что второму предприятию следует выделить 1 единицу. К началу третьего шага останется 2 – 1 = 1 единица ресурсов. При S = 1 третье предприятие не получает ничего. К началу четвертого шага остается 1 единица ресурсов, которая выделяется четвертому предприятию.

Итак, оптимальное управление есть

$$x_1=5,\ x_2=1,\ x_3=0,\ x_4=1.$$

В табл.5 эти управления выделены курсивом и подчеркиванием.

5.4. Решение при увеличенном запасе средств

Пусть исходный запас средств увеличен на две единицы: ΔK=2. Общий их запас составит $K + \Delta K$ = 7 + 2 = 9. В этом случае в таблицу 5 следует добавить две дополнительные строки, соответствующие условным оптимальным управлениям для S = 8 и S = 9. Чтобы получить эти две строки составим дополнительно таблицы 2.1, 3.1, 4.1.

Таблица 2.1. Условно оптимальные решения для третьего шага

S	x	$S–x$	$\varphi_3(x)$	$W_4(S–x)$	$\varphi_3(x)+ W_4(S–x)$	Условно оптимальное решение	
						$x_3(S)$	$W_3(S)$
8	3	5	1.3	1.3	2.6		
	4	4	1.4	1.3	2.7		
	5	3	1.5	1.3	***2.8***	5	2.8
9	4	5	1.4	1.3	2.7		
	5	4	1.5	1.3	***2.8***	5	2.8

Таблица 3.1. Условно оптимальные решения для второго шага

S	x	$S–x$	$\varphi_2(x)$	$W_3(S–x)$	$\varphi_2(x)+W_3(S–x)$	Условно-оптимальное решение	
						$x_2(S)$	$W_2(S)$
8	0	8	0	2.8	2.8		
	1	7	0.6	2.7	3.3		
	2	6	1.1	2.6	***3.7***	2, 3, 4	3.7
	3	5	1.2	2.5	***3.7***		
	4	4	1.4	2.3	***3.7***		
	5	3	1.6	1.6	3.2		
9	0	9	0	2.8	2.8		
	1	8	0.6	2.8	3.4		
	2	7	1.1	2.7	3.8		
	3	6	1.2	2.6	3.8		
	4	5	1.4	2.5	***3.9***	4, 5	3.9
	5	4	1.6	2.3	***3.9***		

При составлении табл.4.1, как и при составлении табл.4 рассматриваем только один случай $S = K = 9$, так как на первом шаге доступен весь имеющийся запас средств.

Таблица 4.1. Условно оптимальные решения для первого шага. K=9

x	$K–x$	$\varphi_1(x)$	$W_2(K–x)$	$\varphi_1(x)+W_2(K–x)$	Оптимальное решение	
					x_1^*	W^*
0	9	0	3.9	3.9		
1	8	0.1	3.7	3.8		
2	7	0.5	3.6	4.1		
3	6	1.2	3.4	4.6		
4	5	1.8	2.9	4.7		
5	4	2.5	2.4	***4.9***	5	4.9

Пополним табл.5 результатами из табл.2.1, 3.1, 4.1, получим табл.5.1.

Таблица 5.1. Результаты условной оптимизации. K=9

S	i=4		i=3		i=2		i=1	
	$x_4(S)$	$W_4(S)$	$x_3(S)$	$W_3(S)$	$x_2(S)$	$W_2(S)$	$x_1(S)$	$W_1(S)$
1	***1***	1.0	0	1.0	0	1.0		
2	2	1.2	***1***	1.3	1	1.6		
3	3	1.3	2	1.6	2	2.1		
4	4	1.3	3	2.3	***2***	2.4		
5	5	1.3	3	2.5	1	2.9		
6			3, 4	2.6	2	3.4		
7			4, 5	2.7	2	3.6		
8			5	2.8	2, 3, 4	3.7		
9			5	2.8	4, 5	3.9	***5***	***4.9***

Из табл.5.1 видно, что максимальный выигрыш составляет 4.9, а оптимальное распределение ресурсов между предприятиями имеет вид

$$x_1=5,\ x_2=2,\ x_3=1,\ x_4=1.$$

5.5. Решение при уменьшенном запасе средств

Пусть общий запас средств составляет не 7, а 5 единиц, а все остальные условия задачи остаются прежними. Для получения решения нужно пересчитать условно оптимальные управления на первом шаге для $K = 5$. Результаты приведены в табл.4.2. Использовано очевидное свойство $W_2(0)=0$.

Таблица 4.2. Условно оптимальные решения для первого шага. K=5

x	K-x	$\varphi_1(x)$	$W_2(K–x)$	$\varphi_1(x)+ W_2(K–x)$	Оптимальное решение	
					x_1^*	W^*
0	5	0	2.9	***2.9***	0	2.9
1	4	0.1	2.4	2.5		
2	3	0.5	2.1	2.6		
3	2	1.2	1.6	2.8		
4	1	1.8	1.0	2.8		
5	0	2.5	0.0	2.5		

Таблица 5.2. Результаты условной оптимизации. K=5

S	i=4		i=3		i=2		i=1	
	$x_4(s)$	$W_4(S)$	$x_3(S)$	$W_3(S)$	$x_2(S)$	$W_2(S)$	$x_1(S)$	$W_1(S)$
1	***1***	1.0	0	1.0	0	1.0		
2	2	1.2	1	1.3	1	1.6		
3	3	1.3	2	1.6	2	2.1		
4	4	1.3	***3***	2.3	2	2.4		
5	5	1.3	3	2.5	***1***	2.9	***0***	2.9

Из табл.5.2 видно, что оптимальный выигрыш есть W^*=2.9, а оптимальное распределение ресурсов таково:

$$x_1=0,\ x_2=1,\ x_3=3,\ x_4=1.$$

5.6. Решение для меньшего числа предприятий

Пусть требуется распределить средства в количестве K = 7 между вторым, третьим и четвертым предприятиями. Это значит, что второй шаг начинается с запасом средств S = 7. Из табл.5 находим следующее оптимальное решение

$$x_2 = 2,\quad x_3 = 3,\quad x_4 = 2.$$

Максимальный выигрыш составит W^* = 3.6.

Глава 5. Теория массового обслуживания

1. Основные понятия

Теория массового обслуживания или *теория очередей* изучает модели разного рода реальных *систем массового обслуживания* (СМО). Примерами СМО могут служить телефонные станции, билетные кассы, магазины, парикмахерские и т.п.

Каждая СМО состоит из определенного числа обслуживающих единиц или *каналов* обслуживания, например, в супермаркете каналом обслуживания является касса. СМО могут быть *одноканальными* и *многоканальными*.

Заявки на обслуживание поступают в СМО нерегулярно, в случайные моменты времени, образуя *случайный поток заявок*. Обслуживание заявок продолжается в течение случайного промежутка времени. Случайный характер работы СМО приводит к тому, что в какие-то периоды времени могут возникать очереди заявок на обслуживание, а в другие периоды СМО может простаивать из-за отсутствия заявок. С точки зрения владельца СМО ее каналы должны быть постоянно загружены, а с точки зрения пользователей СМО минимальным должно быть время обслуживания.

Теория массового обслуживания изучает модели СМО, связывающие заданные условия работы (число каналов, их производительность, характер потока заявок) с показателями эффективности СМО. Такими показателями могут быть среднее число заявок, обслуживаемых в единицу времени, среднее число занятых каналов, среднее число заявок в очереди и среднее время обслуживания.

СМО делятся на два основных типа: СМО *с отказами* и СМО *с очередью*. В СМО с отказами заявка, поступившая в момент, когда все каналы заняты, покидает СМО необслуженной, например, заявка на телефонный разговор получает отказ, если все телефонные линии заняты. В СМО с очередью при отсутствии свободных каналов заявка помещается в очередь.

СМО с очередью могут использовать различную *дисциплину обслуживания*. Дисциплина *«первый пришел – первый вышел»* обеспечивает обслуживание заявок из очереди в порядке их поступления. Дисциплина *«последний пришел – первый вышел»* может применяться, например, при обработке изделий, предварительно поступающих на склад, когда проще извлечь последнее поступившее изделие. Может быть организована *очередь с приоритетами,* когда заявки в очереди ранжируются по некоторому принципу.

СМО называется *открытой*, если характеристики потока заявок не зависят от состояния самой СМО, например, от числа занятых каналов. В *замкнутой* СМО поток заявок зависит от состояния СМО, например, если рабочий обслуживает группу станков, иногда требующих наладки, то интенсивность «требований» от станков зависит от того, сколько станков неисправно и ждет наладки.

2. Понятие марковского случайного процесса

Под *случайным процессом* понимается процесс изменения состояния некоторой системы с течением времени заранее неизвестным, случайным образом. Рассмотренный выше процесс работы СМО является случайным процессом.

Процесс называется *процессом с дискретными состояниями*, если его возможные состояния $S_1, S_2, S_3, \ldots$ можно заранее перечислить, а переход системы из состояние в состояние происходит скачком. Процесс называется *процессом с непрерывным временем*, если моменты возможных переходов из состояния в состояние не фиксированы заранее, а случайны.

Процесс работы СМО представляет собой случайный процесс с дискретными состояниями и непрерывным временем, так как состояние СМО меняется скачком в неопределенные моменты времени, например, в момент поступления новой заявки.

Случайный процесс, протекающий в системе, называется *марковским*, если для любого момента времени t_0 вероятностные характеристики процесса в будущем зависят только от его состояния в данный момент t_0 и не зависят от того, когда и как система пришла в это состояние.

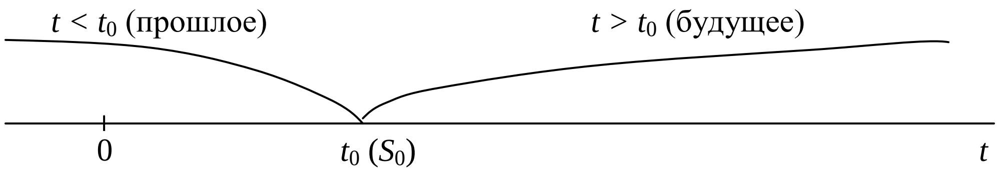

Рис 1. К понятию марковского процесса

Например, показание S счетчика в такси в какой-либо момент времени t зависят только от показаний S_0 счетчика в начальный момент времени t_0 и от пути, пройденного такси после этого момента, что является случайной величиной, и не зависит от того, как изменялись показания счетчика до момента t_0.

Приведем пример *немарковского* случайного процесса. Рассмотрим автомобиль, который может находиться в двух состояниях: «исправен» и «не исправен». Если в данный момент автомобиль находится в состоянии «исправен», то его состояние в будущем зависит не только от начального состояния но и от того, как он перешел в это состояние, например, автомобиль стал исправным после капитального ремонта, тогда следует ожидать, что он будет в исправном состоянии еще длительное время, если же автомобиль долго не ремонтировался и пока исправен, то следует ожидать скорой поломки.

Марковские процессы важны, так как более просты для изучения. В ряде случаев предысторией процессов можно пренебречь и приближенно считать их марковскими.

При анализе случайных процессов с дискретными состояниями и непрерывным временем удобно пользоваться геометрической схемой – так называемым *графом состояний*. Состояния системы изображаются прямоугольниками или кружками, а возможные переходы между состояниями – стрелками, соединяющими состояния. Построим для примера граф состояний технической системы S, состоящей из двух узлов, которые могут выходить из строя (отказывать) в случайные моменты времени. После отказа сразу начинается ремонт узла, который может продолжаться заранее неизвестное, случайное время. Возможные состояния системы можно перечислить:

S_0 – оба узла исправны,

S_1 – первый узел ремонтируется, второй исправен,

S_2 – второй узел ремонтируется, первый исправен,

S_3 – оба узла ремонтируются.

Граф состояний системы приведен на рис.2.

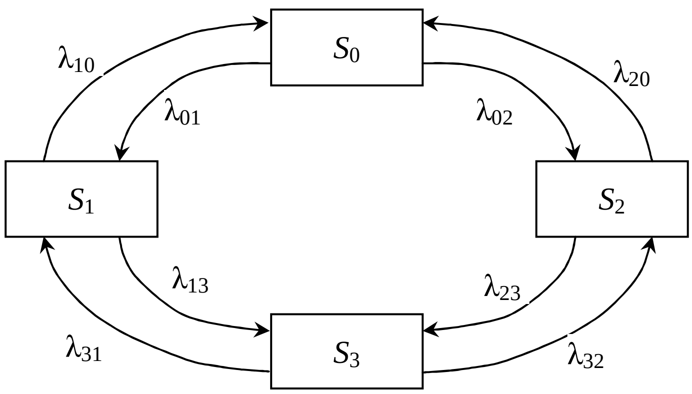

Рис.2. Граф состояний системы из двух узлов

На графе нет стрелок, связывающих состояния S_0 и S_3, так как вероятностью одновременного выхода из строя обоих узлов, так же как и вероятностью одновременного их восстановления пренебрегаем.

3. Потоки событий

Потоком событий называется последовательность однородных событий, следующих одно за другим в случайные моменты времени. Например, поток пассажиров в метро, поток сообщений на пейджер, поток автомобилей на дороге. Заявки, поступающие в СМО, образуют поток событий.

Интенсивностью потока λ называется среднее число событий происходящих в единицу времени.

Поток событий называется *регулярным*, если события следуют одно за другим через равные промежутки времени.

Поток событий называется *стационарным*, если его вероятностные характеристики не зависят от времени. В частности, интенсивность потока λ должна быть постоянной. Например, поток автомобилей на дороге в течение суток не является стационарным, так как его интенсивность днем больше, чем ночью, но в определенный период, например в час пик поток автомобилей может считаться стационарным.

Поток событий называется *потоком без последействия*, если для любых двух непересекающихся участков времени τ_1 и τ_2 число событий, попадающих на один из них, не зависит от числа событий, попадающих на другой. Например, поток пассажиров, входящих в метро, практически не имеет последействия, а поток покупателей, отходящих от прилавка с покупками, имеет последействие, так как после того как произойдет одно событие (покупатель отошел от прилавка, сделав покупку), следующее событие может произойти не в произвольный момент времени, а только спустя минимальное время, необходимое на обслуживание отдельного покупателя, таким образом, здесь имеется определенная зависимость последующего события от предыдущего.

Поток событий называется *ординарным*, если события в нем появляются поодиночке, а не группами. Например, поток поездов, подходящих к станции ординарен, а поток вагонов нет. Если поток ординарен, то вероятностью попадания на малый промежуток времени Δt двух или более событий можно пренебречь.

Поток событий называется *простейшим*, (или *стационарным пуассоновским*), если он стационарен, ординарен и не имеет последействия. Простейшие потоки возникают при наложении достаточно большого числа независимых, стационарных и ординарных потоков. Регулярный поток не является простейшим, так как в равные отрезки времени происходит равное число событий, следовательно, в таком потоке есть последействие.

Выделим на оси времени Ot промежуток τ, рис.3, и рассмотрим случайную величину m – число событий, которое происходит за этот промежуток.

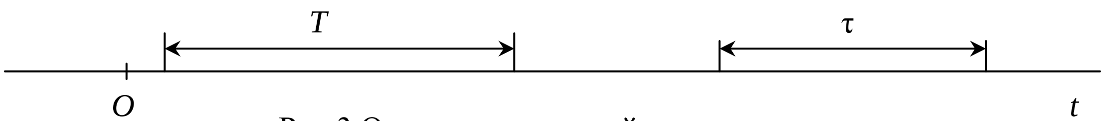

Рис.3 Ось времени случайного процесса

Для простейшего потока событий вероятность того, что в течение промежутка времени τ произойдет ровно m событий, находится по формуле

$$P_m(\tau) = \frac{(\lambda\tau)^m}{m!} e^{-\lambda\tau}. \qquad (1)$$

Говорят также, что случайная величина m имеет распределение Пуассона. Для случайной величины, распределенной по закону Пуассона математическое ожидание равно дисперсии

$$M(m) = m_m = \sigma_m^2 = \lambda\tau.$$

Здесь $M(m)$ или m_m математическое ожидание случайной величины m, σ_m^2 – дисперсия.

Вероятность того, что в течение времени τ не произойдет ни одного события

$$P_0(\tau) = e^{-\lambda\tau}. \qquad (2)$$

Пусть T случайная величина, равная промежутку времени между двумя событиями из потока событий. Найдем закон распределения этой случайной величины

$$F(t) = p(T < t).$$

$F(t)$ есть вероятность того, что интервал между событиями T меньше t, это значит, что в течение промежутка времени t произойдет хотя бы одно событие. Противоположное событие, состоящее в том, что в течение времени t не произойдет ни одного события в соответствии с (2) равно $e^{-\lambda t}$, значит

$$F(t) = p(T < t) = 1 - e^{-\lambda t}. \qquad (3)$$

Плотность распределения случайной величины равна производной от функции распределения, то есть

$$\varphi(t) = F'(t) = \lambda e^{-\lambda t}. \qquad (4)$$

Распределение, задаваемое функцией распределения (3) или плотностью распределения (4) называется показательным или экспоненциальным. Таким образом, интервал времени между двумя произвольными соседними событиями имеет показательное распределение. Математическое ожидание случайной величины с показательным распределением равно дисперсии, то есть

$$M(T) = m_T = \sigma_T^2 = \frac{1}{\lambda}. \qquad (5)$$

Для простейшего потока событий вероятность попадания на элементарный (малый) промежуток времени Δt равна в соответствии с (3)

$$p_{\Delta t} = p(T < \Delta t) = 1 - e^{-\lambda\Delta t} = 1 - \left(1 - \lambda\Delta t + \frac{(\lambda\Delta t)^2}{2} + \cdots\right) \approx \lambda\Delta t \qquad (6)$$

4. Уравнения Колмогорова. Предельные вероятности состояний

Рассмотрим систему, описываемую графом, изображенным на рис.1. Будем считать, что потоки событий, переводящие систему из состояния S_i в состояние S_j, являются простейшими с интенсивностями λ_{ij}, $i, j = 0, 1, 2, 3$. Например, λ_{01} это интенсивность потока отказов первого узла, λ_{10} – интенсивность потока ремонтов первого узла. Граф состояний системы с проставленными у стрелок интенсивностями потоков называется *размеченным*.

Вероятностью i-го состояния системы называется вероятность $p_i(t)$ того, что в момент времени t система находится в i-м состоянии S_i. Для любого момента времени t сумма вероятностей состояний равна единице:

$$\sum_{i=0}^{3} p_i(t) = 1. \qquad (7)$$

Рассмотрим систему в момент времени t и найдем вероятность того, что в момент времени $t+\Delta t$ система будет находиться в состоянии S_0. Это может быть достигнуто двумя способами.

1. *Система в момент времени t находилась в состоянии S_0 и за время Δt не вышла из него*. Найдем вероятность этого способа. Суммарный поток событий, выводящий систему из состояния S_0, имеет интенсивность $\lambda_{01} + \lambda_{02}$, значит вероятность того, что за время Δt система выйдет из состояния S_0, равна $(\lambda_{01} + \lambda_{02})\Delta t$, а вероятность того, что не выйдет: $1 - (\lambda_{01} + \lambda_{02})\Delta t$.

Вероятность того, что система была и осталась в состоянии S_0, находим по теореме умножения вероятностей:

$$p_0(t)[1 - (\lambda_{01} + \lambda_{02})\Delta t]$$

2. *Система в момент времени t находилась в состоянии S_1 или S_2 и за время Δt перешла в состояние S_0*. Найдем вероятность этого варианта. В момент времени t система находится в состоянии S_1 с вероятностью $p_1(t)$. За время Δt система перейдет из состояния S_1 в состояние S_0 с вероятностью $\lambda_{10}\Delta t$, значит вероятность того, что система была в состоянии S_1 и за время Δt перейдет в S_0 равна $p_1(t)\lambda_{10}\Delta t$. Аналогично вероятность перехода из состояния S_2 в S_0 равна $p_2(t)\,\lambda_{20}\Delta t$. Полная вероятность второго варианта составит:

$$p_1(t)\lambda_{10}\Delta t + p_2(t)\lambda_{20}\Delta t$$

По теореме сложения вероятностей

$$p_0(t+\Delta t) = p_1(t)\lambda_{10}\Delta t + p_2(t)\lambda_{20}\Delta t + p_0(t)[1 - (\lambda_{01} + \lambda_{02})\Delta t],$$

отсюда

$$\frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = p_1(t)\lambda_{10} + p_2(t)\lambda_{20} - p_0(t)(\lambda_{01} + \lambda_{02}).$$

Переходя к пределу, получим в левой части производную функции $p_0(t)$ и следующее дифференциальное уравнение

$$p_0'(t) = p_1(t)\lambda_{10} + p_2(t)\lambda_{20} - p_0(t)(\lambda_{01} + \lambda_{02}). \qquad (8)$$

Рассуждая аналогично, запишем для остальных состояний еще три дифференциальных уравнения. Присоединяя к ним уравнение (8) получим систему дифференциальных уравнений для вероятностей состояний:

$$\begin{aligned}
p_0' &= \lambda_{10}p_1 + \lambda_{20}p_2 - (\lambda_{01} + \lambda_{02})p_0,\\
p_1' &= \lambda_{01}p_0 + \lambda_{31}p_3 - (\lambda_{10} + \lambda_{13})p_1,\\
p_2' &= \lambda_{02}p_0 + \lambda_{32}p_3 - (\lambda_{20} + \lambda_{23})p_2,\\
p_3' &= \lambda_{13}p_1 + \lambda_{23}p_2 - (\lambda_{31} + \lambda_{32})p_3,
\end{aligned} \qquad (9)$$

которая называется системой Колмогорова.

Система уравнений Колмогорова позволяет найти, в принципе, вероятности состояний системы как функции времени. Представляет интерес поведение вероятностей $p(t)$ при $t\to\infty$. Если существуют пределы вероятностей, которые не зависят от начальных состояний системы, то они называются *предельными* или *финальными вероятностями*. В теории случайных процессов доказывается, что *если число состояний системы конечно и из каждого из них за конечное число шагов можно перейти в любое другое, то предельные вероятности существуют*.

Будем обозначать предельные вероятности теми же буквами:

$$\lim_{t\to\infty} p_i(t) = p_i,\ i = 0,1,2,3..$$

Так как предельные вероятности постоянны, их производные равны нулю и из (9) получаем следующую систему уравнений для предельных вероятностей:

$$\begin{aligned}
(\lambda_{01} + \lambda_{02})p_0 &= \lambda_{10}p_1 + \lambda_{20}p_2,\\
(\lambda_{10} + \lambda_{13})p_1 &= \lambda_{01}p_0 + \lambda_{31}p_3,\\
(\lambda_{20} + \lambda_{23})p_2 &= \lambda_{02}p_0 + \lambda_{32}p_3,\\
(\lambda_{31} + \lambda_{32})p_3 &= \lambda_{13}p_1 + \lambda_{23}p_2.
\end{aligned} \qquad (10)$$

Вид системы (10) позволяет сформулировать следующее ***правило составления уравнений*** для предельных вероятностей по графу состояния системы: *слева в уравнениях стоит предельная вероятность данного состояния* p_i, *умноженная на суммарную интенсивность всех потоков, ведущих из данного состояния, а справа – сумма произведений интенсивностей всех потоков, входящих в i-е состояние, на вероятности тех состояний, из которых эти потоки исходят*.

Хотя система (10) из четырех уравнений содержит 4 неизвестные p_1, p_2, p_3, p_4, она определяет решение только с точностью до произвольного множителя, так как является однородной. Однозначное решение получается с помощью условия

$$p_1 + p_2 + p_3 + p_4 = 1 \tag{11}$$

Предельная вероятность p_i состояния S_i равна средней доле времени, которое система находится в этом состоянии.

Пусть интенсивности потоков событий имеют следующие численные значения:

$$\lambda_{01} = 1,\ \lambda_{02} = 2,\ \lambda_{13} = 2,\ \lambda_{23} = 1,\ \lambda_{10} = 2,\ \lambda_{31} = 4,\ \lambda_{20} = 4,\ \lambda_{32} = 2,$$

тогда система (10) запишется:

$$\begin{aligned} 3p_0 &= 2p_1 + 4p_2, \\ 4p_1 &= p_0 + 4p_3, \\ 5p_2 &= 2p_0 + 2p_3, \\ 6p_3 &= 2p_1 + p_2. \end{aligned} \tag{12}$$

Отбросим здесь, например, первое уравнение, добавим к системе (12) уравнение (11) и запишем ее в форме, удобной для применения метода исключения Гаусса:

$$\begin{aligned} p_0 - 4p_1 \quad\quad\quad + 4p_3 &= 0, \\ 2p_0 \quad\quad - 5p_2 + 4p_3 &= 0, \\ 2p_1 + p_2 - 6p_3 &= 0, \\ p_0 + p_1 + p_2 + p_3 &= 1. \end{aligned} \tag{13}$$

Используя первое уравнение системы (13), исключим p_0 из остальных уравнений:

$$\begin{aligned} p_0 - 4p_1 \quad\quad + 4p_3 &= 0, \\ 8p_1 - 5p_2 - 6p_3 &= 0, \\ 2p_1 + p_2 - 6p_3 &= 0, \\ 5p_1 + p_2 - 3p_3 &= 1. \end{aligned}$$

Поделим третье уравнение на 2, переставим его со вторым и используем для исключения p_1:

$$\begin{aligned} p_0 - 4p_1 \quad\quad + 4p_3 &= 0, \\ p_1 + \frac{1}{2}p_2 - 3p_3 &= 0, \\ -9p_2 + 18p_3 &= 0, \\ -\frac{3}{2}p_2 + 12p_3 &= 1. \end{aligned}$$

Используем теперь третье уравнение для исключения p_2 из четвертого:

$$
\begin{aligned}
p_0 - 4p_1 \qquad\quad + \ 4p_3 &= 0,\\
p_1 + \frac{1}{2}p_2 - \ 3p_3 &= 0,\\
p_2 - \ 2p_3 &= 0,\\
9p_3 &= 1.
\end{aligned} \tag{14}
$$

Из четвертого уравнения системы (14) находим

$$p_3 = \frac{1}{9},$$

из третьего

$$p_2 = 2p_3 = \frac{2}{9},$$

из второго

$$p_1 = -\frac{1}{2}p_2 + 3p_3 = \frac{2}{9}$$

и из первого

$$p_0 = 4p_1 - 4p_3 = \frac{4}{9}.$$

Итак, доля времени, когда данное устройство исправно, составляет $p_0 = \frac{4}{9} = 0.44$ или 44%, остальное время устройство будет ремонтироваться.

5. Процесс размножения и гибели

Граф состояний многих систем массового обслуживания совпадает с графом состояния биологических систем, в которых происходят процессы размножения и гибели. Этот граф показан на рис. 4.

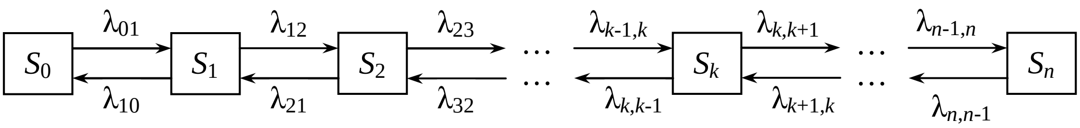

Рис.4. Граф состояний процесса гибели и размножения

Предположим, что все потоки событий, переводящие систему по стрелкам графа, простейшие с указанными у стрелок интенсивностями.

По графу состояний составим и решим систему для предельных вероятностей состояний, существование которых вытекает из того, что из любого состояния можно перейти в любое другое и число состояний конечно.

Для первого состояния S_0 имеем:

$$\lambda_{01}p_0 = \lambda_{10}\, p_1. \tag{15}$$

Для второго состояния S_1:

$$(\lambda_{12} + \lambda_{10})\, p_1 = \lambda_{01}\, p_0 + \lambda_{21}\, p_2.$$

В силу (15) последнее равенство приводится к виду

$$\lambda_{12}p_1 = \lambda_{21}\, p_2;$$

Далее, аналогично,

$$\lambda_{23} p_2 = \lambda_{32}\, p_3$$

и так далее. В результате получается следующая система:

$$\left.\begin{aligned} &\lambda_{01} p_0 = \lambda_{10} p_1, \\ &\lambda_{12} p_1 = \lambda_{21} p_2, \\ &.\ .\ .\ .\ .\ .\ .\ .\ . \\ &\lambda_{k-1,k}\, p_{k-1} = \lambda_{k,k-1} p_k, \\ &.\ .\ .\ .\ .\ .\ .\ .\ . \\ &\lambda_{n-1,n}\, p_{n-1} = \lambda_{n,n-1} p_n. \end{aligned}\right\} \qquad (16)$$

К системе (16) надо добавить еще нормировочное условие:

$$p_0 + p_1 + \ldots + p_n = 1. \qquad (17)$$

Решим эту систему. Из первого уравнения (16) выразим p_1 через p_0:

$$p_1 = \frac{\lambda_{01}}{\lambda_{10}} p_0. \qquad (18)$$

Из второго, с учетом (18), получим:

$$p_2 = \frac{\lambda_{12}}{\lambda_{21}} p_1 = \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} p_0. \qquad (19)$$

Для любого k от 1 до n:

$$p_k = \frac{\lambda_{k-1,k} \cdots \lambda_{12}\lambda_{01}}{\lambda_{k,k-1} \cdots \lambda_{21}\lambda_{10}} p_0. \qquad (20)$$

Подставим предельные вероятности, выраженные через p_0, в (17):

$$p_0\left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} + \ldots + \frac{\lambda_{n-1,n} \cdots \lambda_{12}\lambda_{01}}{\lambda_{n,n-1} \cdots \lambda_{21}\lambda_{10}}\right) = 1,$$

отсюда получим выражение для p_0:

$$p_0 = \left(1 + \frac{\lambda_{01}}{\lambda_{10}} + \frac{\lambda_{12}\lambda_{01}}{\lambda_{21}\lambda_{10}} + \ldots + \frac{\lambda_{n-1,n} \cdots \lambda_{12}\lambda_{01}}{\lambda_{n,n-1} \cdots \lambda_{21}\lambda_{10}}\right)^{-1}. \qquad (21)$$

Определив по формуле (21) p_0, остальные предельные вероятности вычисляем по формулам (20).

6. СМО с отказами

В качестве показателей эффективности СМО с отказами будем принимать:

A – *абсолютную пропускную способность СМО*, т.е. среднее число заявок, обслуживаемых в единицу времени;

Q – *относительную пропускную способность*, т.е. среднее долю заявок, обслуживаемых системой;

$P_{\text{отк.}}$ – *вероятность отказа*, т.е. того, что заявка покинет систему необслуженной;

$\bar{k}$ – среднее число занятых каналов (для многоканальной системы).

6.1. Одноканальная система с отказами

Имеется один канал обслуживания, на который поступает поток заявок с интенсивностью λ. Пусть полностью загруженный канал обслуживает в среднем μ заявок в единицу времени. Эта величина есть интенсивность обслуживания, она обратно пропорциональна среднему времени обслуживания

$$\bar{t}_{об.} = \frac{1}{\mu}. \tag{22}$$

Система может находиться в двух состояниях: S_0 – канал свободен и S_1 – канал занят. Размеченный граф состояний показан на рис.5.

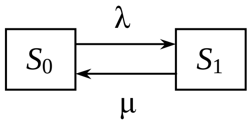

Рис.5. Граф состояний одноканальной системы с отказами

Составим, аналогично (10), систему уравнений для предельных вероятностей:

$$\left.\begin{aligned} \lambda p_0 &= \mu p_1, \\ \mu p_1 &= \lambda p_0. \end{aligned}\right\}$$

Видно, система выродилась в одно уравнение, но добавляя условие

$$p_0 + p_1 = 1,$$

находим предельные вероятности

$$p_0 = \frac{\mu}{\lambda+\mu}, \quad p_1 = \frac{\lambda}{\lambda+\mu}, \tag{23}$$

которые выражают среднее время нахождения системы в состояниях S_0 (канал свободен) и S_1 (канал занят).

Если канал занят, заявка покидает систему, вероятность этого p_1, т.е.

$$P_{\text{отк.}} = p_1 = \frac{\lambda}{\lambda+\mu} \tag{24}$$

Заявка будет обслужена, если при ее поступлении канал свободен, значит, доля обслуженных заявок совпадает с p_0, т.е.

$$\mathrm{Q} = p_0 = \frac{\mu}{\lambda+\mu}. \tag{25}$$

Абсолютную пропускную способность получим, умножив интенсивность потока заявок на вероятность того, что заявка будет обслужена, то есть на относительную пропускную способность:

$$A = Qp_0 = \frac{\lambda\mu}{\lambda+\mu}. \tag{26}$$

6.2. Многоканальная система с отказами (задача Эрланга)

Пусть в систему с n каналами обслуживания поступает поток заявок с интенсивностью λ. Интенсивность обслуживания одного канала равна μ. Система может находиться в состояниях: S_0 – все каналы свободны (занято 0 каналов), S_i, $i = \overline{1,n}$ – занято i каналов. Граф состояний системы показан на рис.6.

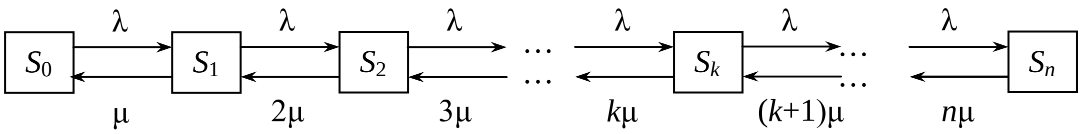

Рис.6 Граф состояний многоканальной СМО с отказами

Из состояния S_0 в S_1 систему переводит поток заявок с интенсивностью λ (как только приходит заявка, система перескакивает из S_0 в S_1). Тот же поток переводит систему из любого левого состояния в соседнее правое.

Рассмотрим интенсивности обслуживания. Если работает один канал (состояние S_1), то поток обслуживания имеет интенсивность μ, Если работают два канала (состояние S_2), то система перейдет в S_1, если закончит работу либо первый либо второй канал и суммарная интенсивность их потоков обслуживания составит 2μ. Суммарный поток обслуживаний, даваемый тремя каналами, составит 3μ, k каналами – kμ.

Граф состояний на рис.6 соответствует схеме гибели и размножения, поэтому применив формулу (21), получим

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{2\cdot 3\mu^3} + \ldots + \frac{\lambda^n}{n!\mu^n}\right)^{-1}. \tag{27}$$

Согласно формулам (18) – (20) предельные вероятности остальных состояний будут:

$$p_1 = \frac{\lambda}{\mu}p_0,\ p_2 = \frac{\lambda^2}{2\mu^2}p_0,\ p_3 = \frac{\lambda^3}{2\cdot 3\mu^3}p_0, \ldots, p_k = \frac{\lambda^k}{k!\mu^k}, \ldots, p_n = \frac{\lambda^n}{n!\mu^n} \tag{28}$$

Величина

$$\rho = \frac{\lambda}{\mu} \tag{29}$$

называется *приведенной интенсивностью потока заявок*. Ее смысл – среднее число заявок, приходящих за среднее время обслуживания одной заявки. Используя (29) перепишем (27), (28) в виде:

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2} + \frac{\rho^3}{2\cdot 3} + \ldots + \frac{\rho^n}{n!}\right)^{-1}, \tag{30}$$

$$p_1 = \rho p_0,\ p_2 = \frac{\rho^2}{2}p_0,\ p_3 = \frac{\rho^3}{2\cdot 3}p_0, \ldots, p_k = \frac{\rho^k}{k!}, \ldots, p_n = \frac{\rho^n}{n!}. \tag{31}$$

Формулы (30), (31) называются формулами Эрланга. (Агнер Краруп Эрланг, датский инженер и математик, родился 1.01.1878, умер 3.02.1929, занимался расчетом пропускной способности телефонных каналов).

Вероятность отказа СМО есть вероятность того, что все каналы заняты, значит

$$P_{отк.} = p_n = \frac{\rho^n}{n!}p_0. \tag{32}$$

Отсюда находим относительную пропускную способность – вероятность того, что заявка будет обслужена:

$$Q = 1 - P_{отк.} = 1 - \frac{\rho^n}{n!}p_0. \tag{32}$$

Абсолютная пропускная способность

$$A = \lambda Q = \lambda\left(1 - \frac{\rho^n}{n!}p_0\right). \tag{33}$$

Среднее число занятых каналов $\bar{k}$ можно найти, используя следующие рассуждения. Известна абсолютная пропускная способность системы A, т.е. число заявок, обслуживаемых всей системой в единицу времени. Каждый канал обслуживает за единицу времени в среднем μ заявок. Значит среднее число работающих или занятых каналов

$$\bar{k} = \frac{A}{\mu}, \tag{34}$$

а с учетом (33),

$$\bar{k} = \rho\left(1 - \frac{\rho^n}{n!} p_0\right). \qquad (35)$$

7. Пример

7.1. Одноканальная система с отказами

Кафе, в котором фирменным блюдом является шашлык, решило организовать доставку шашлыка заказчикам на дом и установило телефон для приема заказов. После рекламы новой услуги стало поступать в среднем 90 заявок в час. Средняя продолжительность разговора с клиентом 3 мин.

Требуется определить вероятность того, что телефон для приема заявки окажется занятым, долю заказчиков, которым удастся дозвониться с первого раза и среднее число заявок, принятых за час.

В данном случае имеем одноканальную систему обслуживания (телефон). Интенсивность потока заявок составляет

$$\lambda = 90\ 1/\textit{час},$$

среднее время обслуживания

$$\bar{t}_{об.} = 3\textit{мин} = 1/20\,\textit{час}.$$

Интенсивность потока обслуживания равна

$$\mu = \frac{1}{\bar{t}_{об.}} = 20\frac{1}{\textit{час}}.$$

Вероятность того, что клиент, набрав номер кафе, обнаружит телефон занятым – это вероятность отказа, находим ее по формуле (24):

$$P_{\text{отк.}} = \frac{\lambda}{\lambda + \mu} = \frac{90}{90 + 20} = \frac{9}{11} = 0.818.$$

Значит, 81.8% потенциальных клиентов не дозвонятся и, возможно, будут потеряны для кафе.

Средняя доля клиентов, дозвонившихся с первого раза – это относительная пропускная способность СМО, она находится по формуле (25):

$$Q = \frac{\mu}{\lambda + \mu} = \frac{20}{90 + 20} = \frac{2}{11} = 0.181,$$

таким образом, в среднем только 18.1% клиентов дозвонятся с первого раза.

Среднее число заявок, принятых за час – это абсолютная пропускная способность СМО, находим ее по формуле (26):

$$A = \frac{\lambda\mu}{\lambda + \mu} = \frac{90 \cdot 20}{90 + 20} = \frac{180}{11} = 16.3.$$

Проведенный анализ показывает, что одного канала (телефона) в данном случае слишком мало, так как слишком большая доля клиентов оказывается необслуженной, что ведет к потере потенциальной прибыли у данного предприятия.

7.2. Многоканальная система с отказами

Пусть условия, рассмотренной выше задачи сохраняются, но для приема заказов было поставлено 3 телефонных аппарата. Требуется рассчитать критерии эффективности такой СМО.

Количество каналов обслуживания составляет

$$n = 3.$$

Находим приведенную интенсивность потока заявок по формуле (29):

$$\rho = \frac{\lambda}{\mu} = \frac{90}{20} = 4.5.$$

Вероятность p_0 того, что все каналы свободны, находим по формуле (30):

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2} + \frac{\rho^3}{2 \cdot 3}\right)^{-1} = \left(1 + 4.5 + \frac{4.5^2}{2} + \frac{4.5^3}{2 \cdot 3}\right)^{-1} = 0.0325.$$

Вероятность отказа находим по формуле (32):

$$P_{отк.} = \frac{\rho^n}{n!} p_0 = \frac{4.5^3}{3!} 0.0325 = 0.493.$$

Относительная пропускная способность, согласно (33) составит

$$Q = 1 - P_{отк.} = 1 - \frac{\rho^n}{n!} p_0 = 1 - 0.493 = 0.507.$$

Абсолютную пропускную способность находим по формуле (33):

$$A = \lambda Q = \lambda\left(1 - \frac{\rho^n}{n!} p_0\right) = 90 \cdot 0.507 = 45.6.$$

Среднее число занятых каналов (телефонов) составит, в соответствии с (34)

$$\bar{k} = \frac{A}{\mu} = \frac{45.6}{20} = 2.28.$$

Несмотря на увеличение числа каналов до трех, данная СМО остается недостаточно эффективной, так как только половина клиентов получает обслуживание.

Литература

Основная

1. Экономико-математические методы и прикладные модели: Учебное пособие для вузов/ В.В. Федосеев, А.Н. Гармаш, Д.М. Дайитбегов и др.; Под ред. В.В. Федосеева.– М.: ЮНИТИ, 1999.– 391с.
2. Вентцель Е.С. Исследование операций: задачи, принципы, методология.– М.: Наука, Гл. ред. физ.-мат. лит., 1988.– 208с.
3. Исследование операций в экономике: Учебн. пособие для вузов/ Н.Ш. Кремер, Б.А. Бутко, М.Н. Фридман; Под ред. Н.Ш. Кремера.– М.: Банки и биржи, ЮНИТИ, 1997.– 407с.
4. Фомин Г.П. Математические методы и модели в коммерческой деятельности: Учебник.– М.: Финансы и статистика, 2001.– 544с.

Дополнительная

5. Вершик А.М. О Л. В. Канторовиче и линейном программировании.– http://www.mathsoc.spb.ru/pantheon/kantorov/vershik.html
6. Dantzig G.B. Reminiscences About the Origins of Linear Programming //Mathematical Programming: The State of the Art / Ed. by A. Bachem, M. Groetschel and B. Korte.- Berlin: Springer Verlag, 1983.– P. 79-86. Есть русский перевод: Данциг Дж. Б. Воспоминания о началах линейного программирования.– http://www.webcenter.ru/~zwb/origins.htm