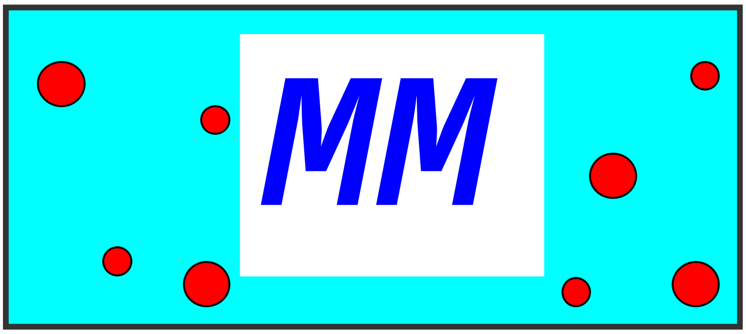
Федеральное агентство железнодорожного транспорта
Уральский государственный университет путей сообщения
Кафедра «Высшая математика»

И. Н. Пирогова
П. П. Скачков

МАТЕМАТИЧЕСКИЕ МОДЕЛИ



Екатеринбург
2009

Федеральное агентство железнодорожного транспорта
Уральский государственный университет путей сообщения
Кафедра «Высшая математика»

И. Н. Пирогова
П. П. Скачков

МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Методические указания
по методике проведения практических занятий
и самостоятельной работы для студентов всех специальностей
заочной формы обучения

Екатеринбург
2009

УДК 51
П 33

Пирогова, И. Н., Скачков, П. П.
П33 Математические модели : метод. указания. − Екатеринбург : УрГУПС, 2009. – 44 с.

Методические указания предназначены для проведения практических занятий и для самостоятельной работы по курсам «Линейное программирование», «Массовое обслуживание», «Математические модели» для студентов всех специальностей заочной формы обучения. При создании руководства использованы материалы кафедры «Высшая математика» УрГУПС.

Работа содержит краткие теоретические сведения и примеры решения задач по рассмотренным разделам. Предложенные задания могут быть использованы как темы курсовых или лабораторных работ студентов.

Работа состоит из двух частей. Первая часть посвящена вопросам, связанным с задачами линейного программирования и содержит следующие темы: обзор основных задач линейного программирования, методы решения задач: геометрический и симплекс-метод. Во второй части рассмотрены Марковские процессы и простейшие модели массового обслуживания.

Одобрено и рекомендовано к изданию на заседании кафедры «Высшая математика» (протокол № 1 от 10.06.2009 г.).

Авторы: И. Н. Пирогова, ст. преподаватель кафедры «Высшая математика», УрГУПС
П. П. Скачков, доцент кафедры «Высшая математика», канд. физ.-мат. наук, УрГУПС

Рецензент: Г. А. Тимофеева, профессор, зав. кафедрой «Высшая математика», д-р физ.-мат. наук, УрГУПС

ОГЛАВЛЕНИЕ

1. Введение

Решение современных экономических проблем и анализ экономических ситуаций невозможен без использования математических моделей. Даже формулировка основного принципа экономики – максимум эффективности при минимуме затрат ресурсов – подразумевает активное участие математического моделирования.

Математическую модель можно определить как внутренне непротиворечивую замкнутую систему математических соотношений, предназначенную для воспроизведения определенного качества изучаемого реального объекта или явления. Математические модели представляют собой основу компьютерного моделирования и обработки информации. К тому же необходимость решения актуальных экономических проблем часто инициирует развитие математического аппарата. Например, появление класса продуктивных матриц в линейной алгебре обусловлено исследованием моделей межотраслевого баланса; математическое программирование в своей основе имеет сугубо экономический аспект оптимального планирования распределения ограниченных ресурсов.

Можно отметить два наиболее важных аспекта применения математических методов: это теория массового обслуживания, которая применяется в сфере обслуживания (АТС, АЗС, магазины, кассы и т. п.), в современных высоких технологиях, а также линейное программирование в финансово-экономической сфере.

2. Основные понятия линейного программирования

Линейное программирование сформировалось как отдельный раздел прикладной математики в 40–50 гг. XX века благодаря работам советского ученого, академика Л. В. Канторовича. Методам линейного программирования посвящено много работ американских математиков. Так, математик Д. Данцинг ввел понятие линейного программирования и предложил в 1949 г. алгоритм решения задач линейного программирования, получивший название «симплекс-метод». Термин программирование введен в связи с тем, что неизвестные переменные, которые находятся в процессе решения задачи, обычно определяют программу или план работы некоторого экономического объекта. По оценкам специалистов примерно 80–85% всех решаемых на практике задач оптимизации относится к задачам линейного программирования.

Линейное программирование – это область математического программирования, являющегося разделом математики и изучающего методы нахождения экстремальных (наибольших и наименьших) значений линейной функции конечного числа переменных, причем на неизвестные переменные наложены линейные ограничения. Такая линейная функция называется *функцией цели*, а ограничения, которые представляют количественные соотношения между переменными, выражающие условия и требования экономической задачи и которые математически записываются в виде линейных уравнений и неравенств, называются *системой ограничений*. Математическое выражение целевой функции и ее ограничений называется *математической моделью экономической задачи.*

Математические модели задач линейного программирования (ЛП) различаются между собой по виду ограничений.

Определение 1. Математическая модель задачи ЛП, состоящая в определении максимального (минимального) значения функции цели, называется *стандартной*, если все ограничения в ней представлены в виде линейных неравенств.

Например, пусть в модели два неизвестных: $x_1,\ x_2$. Надо найти среди решений системы неравенств такие, которые дают целевой функции наибольшее (наименьшее) значение. Такая модель имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \le b_1, \\ a_{21}x_1 + a_{22}x_2 \le b_2, \\ \dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 \le b_m. \end{cases} \tag{1}$$

$$f(X) = c_1x_1 + c_2x_2 + c_0 \to \max(\min), \tag{2}$$

где $X = (x_1, x_2),\ (x_1 \ge 0;\ x_2 \ge 0)$.

Определение 2. Математическая модель задачи ЛП называется *основной*, если ограничения в ней представлены в виде уравнений при условии неотрицательности переменных. Такая модель имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{3}$$

$$x_j \ge 0\ ,\ j = 1, 2\ldots, n$$

$$f(X) = c_1x_1 + c_2x_2 + \ldots + c_nx_n + c_0 \to \max(\min), \tag{4}$$

где $X = (x_1, x_2, ..., x_n)$.

Система уравнений (3) обычно имеет конечное множество решений. Целевая функция $f(x)$ − линейная, поэтому частные производные постоянны, а это значит, что внутри области решений системы (3) экстремальных точек нет. Если целевая функция имеет оптимальное решение, то оно достигается в точках границы области.

Таким образом, в линейном программировании исследователя интересуют вершины многоугольника решений системы (3).

Определение 3. Математическая модель называется *канонической*, если ее система ограничений представлена в виде системы $\boldsymbol{m}$ линейно независимых уравнений (ранг системы $r = m$), в системе выделен единичный базис, свободные переменные и целевая функция выражены через свободные переменные.

Запишем *каноническую* модель задачи ЛП:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1s}x_s + x_{s+1} = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2s}x_s + x_{s+2} = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rs}x_s + x_{s+r} = b_r. \end{cases} \tag{5}$$

$$x_j \ge 0, j = 1,2\ldots,n;\ n \quad s+r;$$

$$f(x) = c_0 + c_1x_1 + \ldots + c_nx_n \to \min. \tag{6}$$

Определение 4. Будем называть переменные, входящие в одно из уравнений системы с коэффициентом один и отсутствующие в других уравнениях, *базисными* неизвестными, а все другие – *свободными*.

В нашей модели переменные $x_1,\ldots,x_s$ – свободные, и если их принять равными нулю, то из системы (5) легко получить значения базисных переменных $x_{s+1}, x_{s+2},\ldots, x_{s+r}$.

Определение 5. Решение системы называется *базисным*, если в нем свободные переменные равны нулю, и оно имеет вид

$$x_{\text{баз}} = (0,\ldots 0, b_1, \ldots b_r), \quad f(x_{\text{баз}}) = c_0.$$

Базисное решение является *угловой* точкой множества решений системы (5), или, можно сказать, определяет вершину многоугольника решений модели. Среди таких решений находится и то, при котором целевая функция принимает *оптимальное* значение.

Определение 6. Базисное решение называется *опорным*, если оно допустимо, то есть все правые части уравнений (неравенств) положительны $b_i \ge 0, \quad i = \overline{1,m}$.

Теорема 1. Базисное опорное решение в канонической модели является *оптимальным*, если все коэффициенты в целевой функции не положительны, то есть $c_{ij} \le 0, \quad j = \overline{1,n}$.

2.1. Правила перехода от одной модели к другой

2.1.1. Переход от стандартной модели задачи ЛП к канонической

При этом переходе требуется преобразовать систему ограничений-неравенств в систему уравнений.

Возьмем стандартную модель:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \le b_1, \\ a_{21}x_1 + a_{22}x_2 \le b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 \le b_m. \end{cases}$$

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

$$f(x) = c_1x_1 + c_2x_2 + c_0 \to \max(\min).$$

Введем дополнительные неизвестные $x_3 \geq 0, \ldots, x_{m+2} \geq 0$ такие, чтобы они уравняли левую и правую части неравенств. Получим новую модель:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + x_4 = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + x_{m+2} = b_m. \end{cases}$$

$$x_j \geq 0\ ,\ j \quad 1,2.=,n,n+1,\ldots m$$

$$f(x) = c_1x_1 + c_2x_2 + c_0 \to \max \quad \text{или} \quad f_1(x) = -c_1x_1 - c_2x_2 - c_0 \to \min$$

Это каноническая модель задачи ЛП.

2.1.2. Переход от канонической модели задачи ЛП к стандартной

Если нам задана каноническая модель ЗЛП, то можно перейти к стандартной модели, если просто отбросить базисные переменные, при этом равенства переходят в неравенства.

Пусть дана каноническая модель:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1s}x_s + x_{s+1} = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + \ldots + a_{rs}x_s + x_{s+r} = b_r. \end{cases}$$

$$x_j \geq 0\ ,\ j = 1, 2, \ldots n,$$

$$f(x) = c_1x_1 + \ldots + c_sx_s + c_0 \to \min$$

Так как переменные: x_{s+1}, ... , x_{s+r} неотрицательны, то стандартная модель будет иметь вид:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1s}x_s \leq b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + \ldots + a_{rs}x_s \leq b_r. \end{cases}$$

$$x_j \geq 0\ ,\ \ j = 1, \ldots s$$

$$f(x) = c_1x_1 + \ldots + c_sx_s + c_0 \to \min$$

3. Геометрическая иллюстрация решения задач ЛП

Пусть задана стандартная математическая модель задачи с двумя неизвестными:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1s}x_s \leq b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + \ldots + a_{rs}x_s \leq b_r. \end{cases} \tag{7}$$

$$x_j \geq 0\ ,\ j = 1, \ldots, s$$

$$f(x) = c_1x_1 + \ldots + c_sx_s + c_0 \to \min \tag{8}$$

Нахождение решения этой модели на основе ее геометрической интерпретации включает следующие этапы.

1. В плоскости $x_1 0 x_2$ строят прямые, уравнения которых получаются в результате замены в ограничениях (7) модели знаков неравенств на знаки точных равенств.

2. Находят полуплоскости, определенные каждым неравенством системы.

3. Находят выпуклый многоугольник решений всей системы (7).

4. Строят нормальный вектор целевой функции $\vec{n}=(c_1,\ c_2)$, причем, начало вектора совмещают с началом координат.

5. Строят прямую $c_1 x_1 + c_2 x_2 = 0$.

6. Передвигают эту прямую в направлении вектора $\vec{n}$, тогда в результате либо находят вершину или отрезок, в которой целевая функция принимает наибольшее значение, либо устанавливают неограниченность сверху этой функции на множестве допустимых решений.

7. Если функция ограничена, то определяют $x_{\max} = (x_1, x_2)$ и вычисляют значение функции в этой точке $f_{\max} = f\left(x_{\max}\right)$.

При геометрической интерпретации задач ЛП могут встретиться случаи, изображенные на рис. 1−4.

Рис. 1. Задача ЛП имеет единственное решение $f_{\max} = f\left(x_{\max}\right)$.

Рис. 2. Задача ЛП имеет бесчисленное множество решений, т. к. целевая функция достигает наибольшего значения на отрезке [A; B].

Рис. 3. Задача ЛП не имеет решения, т. к. функция неограниченна сверху.

Рис. 4. Задача ЛП не имеет решения, т. к. система (7) несовместна.

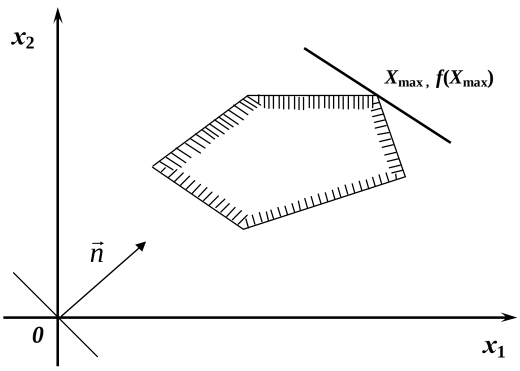


Рис. 1.

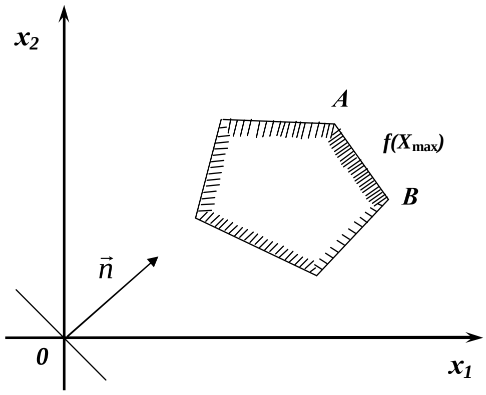


Рис. 2.

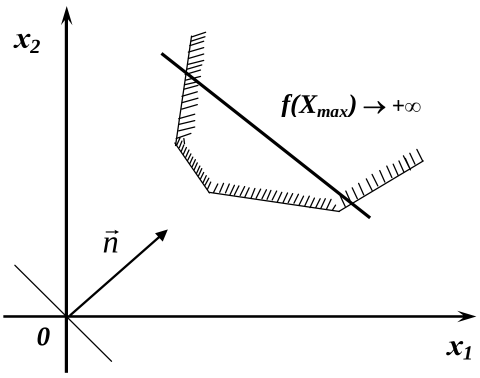


Рис. 3.

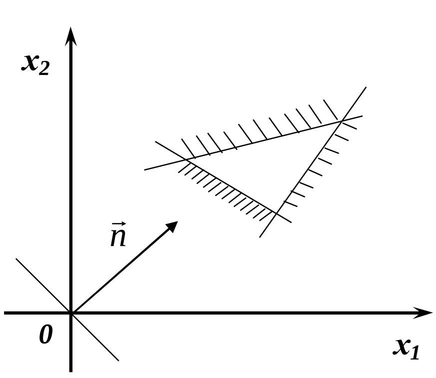


Рис. 4.

Пример 1. Для производства двух видов изделий *A* и *B* предприятие использует три вида сырья S_1, S_2, S_3. Нормы расхода сырья каждого вида на изготовление единицы продукции данного вида приведены в табл.

Сырье	*A*	*B*	Запасы
S_1	$a_{11} = 12$	$a_{12} = 4$	$b_1 = 300$
S_2	$a_{21} = 4$	$a_{22} = 4$	$b_2 = 120$
S_3	$a_{31} = 3$	$a_{32} = 12$	$b_3 = 252$
Прибыль	$c_1 = 30$	$c_2 = 40$	

Прибыль от реализации одного изделия каждого вида равна c_1 и c_2, а общее количество сырья вида S_i равно b_i, $i = 1, 2, 3$. Считая, что изделия *A* и *B* могут производиться в любых соотношениях (сбыт обеспечен), требуется составить такой план их выпуска, при котором прибыль предприятия от реализации всех изделий будет максимальной.

Решение. Обозначим через x_1 и x_2 количество изделий первого и второго вида в плане предприятия. Поскольку производство продукции ограничено только сырьем каждого типа S_i, то получим условия:

$$\begin{cases} 12x_1 + 4x_2 = 300, \\ 4x_1 + 4x_2 = 120, \\ 3x_1 + 12x_2 = 252. \end{cases}$$

Переменные x_1 и x_2 не могут быть отрицательными по смыслу задачи

$$x_j \geq 0, \quad j = 1, 2.$$

Вычислим прибыль от реализации продукции и получим:

$$f(X) = 30x_1 + 40x_2 \to \max, \text{ где } X = (x_1, x_2).$$

Итак, мы получили стандартную модель с двумя переменными.

Решим задачу линейного программирования геометрически, придерживаясь плана, приведенного ранее.

1. Строим прямые l_1, l_2, l_3:

$l_1 : 12x_1 + 4x_2 = 300$, по двум точкам A_1 (25; 0) и B_1 (0; 75).

$l_2 : 4x_1 + 4x_2 = 120$, по двум точкам A_2 (30; 0), B_2 (0; 30).

$l_3 : 3x_1 + 12x_2 = 252$, по двум точкам A_3 (0; 21), B_3 (40; 11).

2. Обратимся к неравенствам системы ограничений. Отметим те полуплоскости, которые им удовлетворяют.

3. Учтем на чертеже неотрицательность переменных x_1 и x_2, и получим многоугольник решений данной системы неравенств $OA_3EB_1A_1$ (рис. 5).

4. Построим нормальный вектор $\vec{n} = \{30; 40\}$

5. Построим прямую $30x_1 + 40x_2 = 0$ (l).

6. Передвигая прямую (l) в направлении вектора $\vec{n}$, находим, что в точке

E (12, 18) целевая функция будет иметь наибольшее значение. Координаты этой точки находим как координаты точки пересечения прямых (l_2) и (l_3), решая систему уравнений

$$\begin{cases} 4x_1 + 4x_2 = 120 \\ 3x_1 + 12x_2 = 252 \end{cases}.$$

7. Запишем окончательный ответ:

$$X_{max} = (12, 18),\ f_{max} = f(X_{max}) = 1080$$

Наибольшая прибыль будет равна 1080 (у.е).

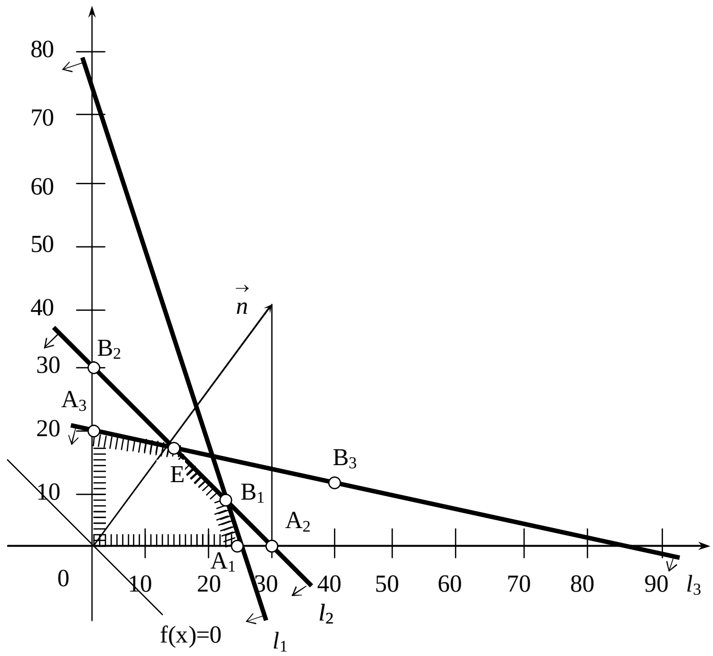


Рис. 5.

Индивидуальные задания

Задание 1

Предприятие выпускает два вида продукции A_1 и A_2, используя при этом три вида сырья S_1, S_2 и S_3. Известны запасы сырья равные b_1, b_2 и b_3 соответственно. Расход сырья вида S_i на производство единицы продукции A_j равен $a_{i,j}$. Доход от реализации единицы продукции A_j составляет c_j условных единиц. Требуется составить такой план производства продукции, при котором доход будет максимальным. Решить задачу графическим методом. Если при изучении темы рассматривались другие методы решения задач ЛП, то в качестве дополнительного задания может быть предложено решение той же задачи симплекс-методом [3,6].

Замечание. При построении графика выбрать масштаб таким образом, чтобы величине $\max(b_i / a_{i\,j})$ приблизительно соответствовало 15 см.

Вариант 1.			
	A_1	A_2	b_i
B_1	$a_{11}=1$	$a_{12}=3$	27
B_2	$a_{21}=3$	$a_{22}=2$	30
B_3	$a_{31}=2$	$a_{32}=3$	30
c_j	2	2	

Вариант 2.			
	A_1	A_2	b_i
B_1	$a_{11}=1$	$a_{12}=1$	10
B_2	$a_{21}=1$	$a_{22}=4$	28
B_3	$a_{31}=3$	$a_{32}=1$	24
c_j	4	8	

Вариант 3.			
	A_1	A_2	b_i
B_1	$a_{11}=3$	$a_{12}=2$	36
B_2	$a_{21}=6$	$a_{22}=7$	90
B_3	$a_{31}=1$	$a_{32}=2$	24
c_j	3	3	

Вариант 4.			
	A_1	A_2	b_i
B_1	$a_{11}=5$	$a_{12}=4$	40
B_2	$a_{21}=5$	$a_{22}=2$	35
B_3	$a_{31}=1$	$a_{32}=4$	24
c_j	4	6	

Вариант 5.			
	A_1	A_2	b_i
B_1	$a_{11}=2$	$a_{12}=1$	16
B_2	$a_{21}=2$	$a_{22}=2$	18
B_3	$a_{31}=1$	$a_{32}=4$	24
c_j	2	4	

Вариант 6.			
	A_1	A_2	b_i
B_1	$a_{11}=1$	$a_{12}=3$	21
B_2	$a_{21}=2$	$a_{22}=3$	24
B_3	$a_{31}=2$	$a_{32}=1$	20
c_j	4	3	

Вариант 7.			
	A_1	A_2	b_i
B_1	$a_{11}=3$	$a_{12}=1$	27
B_2	$a_{21}=1$	$a_{22}=4$	28
B_3	$a_{31}=3$	$a_{32}=4$	36
c_j	3	2	

Вариант 8.			
	A_1	A_2	b_i
B_1	$a_{11}=1$	$a_{12}=5$	50
B_2	$a_{21}=3$	$a_{22}=5$	60
B_3	$a_{31}=3$	$a_{32}=1$	48
c_j	5	10	

Вариант 9.			
	A_1	A_2	b_i
B_1	$a_{11}=5$	$a_{12}=2$	40
B_2	$a_{21}=1$	$a_{22}=3$	30
B_3	$a_{31}=4$	$a_{32}=3$	39
c_j	2	3	

Вариант 10.			
	A_1	A_2	b_i
B_1	$a_{11}=2$	$a_{12}=5$	50
B_2	$a_{21}=4$	$a_{22}=3$	44
B_3	$a_{31}=4$	$a_{32}=1$	36
c_j	2	1	

4. Транспортная задача линейного программирования

4.1 Постановка задачи

Пусть имеется ***m*** поставщиков некоторого однородного груза, сосредоточенного на станциях A_1, А$_2$, …, A_m и имеется ***n*** потребителей этого продукта, расположенные на станциях B_1, B_2, …, B_n. Известны также запасы этого груза $(a_1, a_2, \ldots ,a_m)$, которые должны быть вывезены за фиксированный период времени. Потребности получателей груза за этот же период времени составляют $(b_1, b_2, \ldots,b_n)$. Пусть запасы равны потребностям, т. е.

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j \, . \qquad (9)$$

При выполнении условия (9) транспортная задача называется **закрытой.**

Кроме того, известны затраты $c_{i\,j}$ на перевозку единицы груза с любой станции A_i на каждую станцию B_j .

Требуется составить такой план перевозок, чтобы весь груз был вывезен, все потребности были удовлетворены, а суммарные затраты на перевозки были минимальны.

4.2. Математическая модель

Обозначим величину перевозки со станции A_i на станцию B_j как $x_{i\,j}$. Тогда условия полного вывоза груза со всех станций A_i можно записать в виде системы уравнений:

$$\begin{cases} x_{11} + x_{12} + x_{13} + \ldots + x_{1n} = a_1, \\ x_{21} + x_{22} + x_{23} + \ldots + x_{2n} = a_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{m1} + x_{m2} + x_{m3} + \ldots + x_{mn} = a_m. \end{cases} \qquad (10)$$

С другой стороны, станции B_j должны получить необходимое количество груза, что тоже записывается в виде системы уравнений:

$$\begin{cases} x_{11} + x_{21} + x_{31} + \ldots + x_{m1} = b_1, \\ x_{12} + x_{22} + x_{32} + \ldots + x_{m2} = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_{1n} + x_{2n} + x_{3n} + \ldots + x_{mn} = b_n. \end{cases} \qquad (11)$$

При этом выполняется условие неотрицательности для переменных $x_{i\,j}$,

$$x_{i\,j} \geq 0, (i = 1,2,\ldots m;\ j = 1,2,\ldots n) .$$

Функция цели задачи – по критерию минимума суммарных затрат –

$$F(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{i\,j} x_{i\,j} \to \min \qquad (12)$$

Таким образом, транспортная задача представляет собой основную задачу линейного программирования.

Методом искусственного базиса можно привести эту задачу к каноническому виду и затем решить ее симплекс-методом. Ввиду того, что число переменных транспортной задачи велико и при этом на любом плане число свободных переменных также велико, то симплекс метод является чрезмерно громоздким и малоэффективным. В связи с этим были разработаны специальные методы решения транспортных задач.

Вся исходная информация для транспортной задачи организуется в виде матрицы или таблицы (здесь для краткости рассмотрена таблица задачи размерности 3×5), в которую заносится опорный план задачи. В этой же таблице проводится оптимизация плана и записывается решение задачи.

	B_1	B_2	B_3	B_4	B_5	Запасы
A_1	c_{11} x_{11}	c_{12} x_{12}	c_{13} x_{13}	c_{14} x_{14}	c_{15} x_{15}	a_1
A_2	c_{21} x_{21}	c_{22} x_{22}	c_{23} x_{23}	c_{24} x_{24}	c_{25} x_{25}	a_2
A_3	c_{31} x_{31}	c_{32} x_{32}	c_{33} x_{33}	c_{34} x_{34}	c_{35} x_{35}	a_3
Потребности	b_1	b_2	b_3	b_4	b_5	

Для нахождения начального опорного решения существуют различные методы. Рассмотрим основные из этих методов.

4.3. Методы определения первого опорного плана

4.3.1. Метод северо-западного угла

В соответствии с этим методом, в верхнюю левую (северо-западную) клетку таблицы заносится максимально допустимая перевозка. При этом либо вывозится весь груз со станции A_1 , тогда все остальные клетки первой строки вычеркиваются, либо потребности первого потребителя B_1 полностью удовлетворены, тогда все клетки первого столбца вычеркиваются. После этого самой северо-западной клеткой становится клетка A_1-B_2 или A_2-B_1. Указанный алгоритм применяется до заполнения всей таблицы.

Рассмотрим применение этого метода на примере.

	B_1	B_2	B_3	B_4	B_5	
A_1	4	3	4	11	9	40
A_2	3	4	7	15	8	30
A_3	7	4	2	8	15	10
	20	21	7	24	8	

Ставим первую перевозку в клетку A_1-B_1. Она будет равна $\min\{20, 40\} = 20$. Закрыт первый столбец. Идем по строке. В клетку A_1-B_2 ставим перевозку, равную $\min\{21, 20\} = 20$. Закрыли первую строку (в закрытой строке или столбце будем ставить прочерки для того, чтобы не поставить туда перевозку). Спускаемся по второму столбцу. В клетку A_2-B_2 ставим перевозку $\min\{1, 30\} = 1$ и закрываем второй столбец. Идем по второй строке и ставим перевозку в клетку A_2-B_3. Она равна $\min\{7, 29\} = 7$. Закрыли третий столбец. Продолжаем идти по третьей строке. В клетку A_2-B_4 ставим перевозку $\min\{24, 22\} = 22$. Теперь закрыли вторую строку и спускаемся по четвертому столбцу. В клетку A_3-B_4 ставим перевозку $\min\{2, 10\} = 2$. Закрыв этот столбец, идем по третьей строке. Перевозка в клетке A_3-B_5 равна $\min\{8, 8\} = 8$. Таблица исчерпана. Все грузы вывезены, все потребности удовлетворены.

Недостатком данного метода является то, что он не учитывает стоимость перевозок и поэтому, как правило, получаемый план далек от оптимального.

	B_1	B_2	B_3	B_4	B_5	
A_1	20 [4]	20 [3]	– [4]	– [11]	– [9]	40
A_2	– [3]	1 [4]	7 [7]	22 [15]	– [8]	30
A_3	– [7]	– [4]	– [2]	2 [8]	8 [15]	10
	20	21	7	24	8	

Вычислим стоимость всех перевозок. Она равна

$$F(x) = 20\cdot 4 + 20\cdot 3 + 1\cdot 4 + 7\cdot 7 + 22\cdot 15 + 2\cdot 8 + 8\cdot 15 = 659$$

Теперь на этом же примере рассмотрим другой метод.

4.3.2. Метод наименьшей стоимости

Метод наименьшей стоимости учитывает затраты на перевозки и строится следующим образом: рассматривается матрица, находится клетка с наименьшей стоимостью, в которую заносится максимально допустимая перевозка. При этом вычеркивается либо строка, либо столбец, затем в остающейся части таблицы находится клетка с наименьшей стоимостью и т. д.

Опишем последовательность действий для данного метода на примере.

Выбираем в таблице клетку с наименьшей стоимостью (A_3-B_3) и ставим в нее перевозку $\min\{7, 10\} = 7$, закрыли третий столбец. В пункте A_3 осталось 10−7=3 ед.груза. В оставшихся клетках выбираем наименьшую стоимость. Таких клеток две: A_2-B_1 и A_1-B_2. Случайным образом выбираем сначала клетку A_1-B_2: ставим сюда перевозку, равную $\min\{21, 40\} = 21$. Мы закрыли второй столбец, а в пункте A_1 осталось 40−21=19 ед.груза. Далее ставим перевозку в клетку A_2-B_1. Она равна $\min\{20, 30\} = 20$. Закрыли первый столбец, а в пункте A_2 осталось 30−20=10 ед.груза. Минимальная стоимость в оставшихся клетках равна 8 для клетки A_3-B_4. Ставим туда перевозку $\min\{24, 3\} = 3$. Заполнена третья строка. Пункту B_4 необходимо еще 24−3=21 ед. груза. Такая же стоимость и в клетке A_2-B_5. Ставим туда перевозку $\min\{8, 10\} = 8$. Закрыли пятый столбец, а в пункте A_2 осталось 10−8=2 ед. груза. Во второй строке осталась свободной только клетка A_2-B_4. Ставим туда перевозку $\min\{21, 2\} = 2$. Вторая строка заполнена, пункту B_4 необходимо 21−2=19 ед. груза. Заполняем оставшуюся клетку A_1-B_4. Ставим сюда перевозку 19 ед. груза. Заполнена вся таблица. В большинстве случаев этот метод дает план, который ближе к оптимальному, однако, во всех случаях требуется сравнивать величины функции цели на получаемых планах и выбирать тот, для которого функция принимает наименьшее значение.

	B_1	B_2	B_3	B_4	B_5	
A_1	– 4	21 3	– 4	19 11	– 9	40
A_2	20 3	– 4	– 7	2 15	8 8	30
A_3	– 7	– 4	7 2	3 8	– 15	10
	20	21	7	24	8	

Вычислим общую стоимость перевозок.

$$F(x) = 20\cdot 3 + 21\cdot 3 + 2\cdot 7 + 19\cdot 11 + 2\cdot 15 + 3\cdot 8 + 8\cdot 8 = 464.$$

Стоимость перевозок, полученная методом наименьшей стоимости, меньше чем в методе северо-западного угла, следовательно, для данного примера, он дает лучший план (ближе к оптимальному плану).

4.4. Метод потенциалов

Метод потенциалов позволяет оценить опорный план и методом последовательного улучшения найти оптимальное решение.

Теоремы метода

Теорема 1: Если опорный план $X=(x_{i\,j})$ является оптимальным, то существует система из $(m+n)$ чисел, называемых потенциалами, U_i, V_j, такая, что:

а) $U_i+V_j= C_{i\,j}$, для $x_{i\,j}> 0$ (базисные переменные);

б) $U_i+V_j=C_{i\,j}$, для $x_{i\,j}= 0$ (свободные переменные).

Таким образом, для оптимальности опорного плана необходимо выполнение следующих условий:

а) для каждой занятой клетки сумма потенциалов равна стоимости перевозки единицы груза, стоящей в этой клетке:

$$U_i+V_j = C_{i\,j} \qquad (13)$$

б) для каждой свободной клетки сумма потенциалов меньше или равна стоимости перевозки единицы груза, стоящей в этой клетке:

$$U_i+V_j \le C_{i\,j} \qquad (14)$$

Примечание

Система (13) содержит $(m+n)$ неизвестных и $(m+n$-1) линейно независимых уравнений. Такая система имеет бесчисленное множество решений, которые можно получить, придавая одной из неизвестных конкретное значение. Это значение выбирается произвольно, например, можно придать U_1 значение равное 0, тогда другие потенциалы вычисляются из системы (13).

Теорема 2. Любая закрытая транспортная задача имеет решение, которое достигается за конечное число шагов метода потенциалов.

4.4.1. Построение цикла и определение величины перераспределения груза

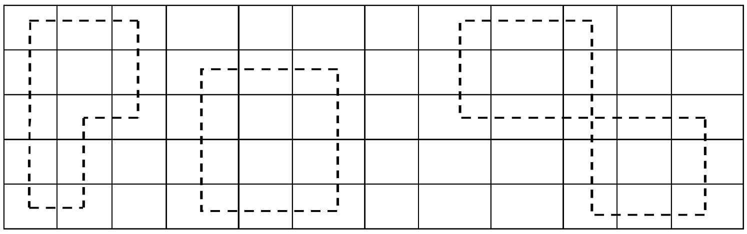
Циклом в таблице перевозок называется ломаная с вершинами в клетках и ребрами, расположенными вдоль строк или столбцов, удовлетворяющая двум требованиям:

а) ломаная должна быть связной, т. е. из любой ее вершины можно попасть в другую вершину, двигаясь по ребрам;

б) в каждой вершине цикла сходятся ровно два ребра − одно по строке, другое по столбцу.

Замечание: в цикле возможны самопересечения, но они происходят только не в вершинах цикла.

Примеры циклов



Теорема 3. Если в таблице перевозок m строк и n столбцов, и перевозками заполнено (m+n−1) клеток, то существует цикл, одна из вершин которого расположена в свободной клетке, а все остальные вершины в занятых клетках. Такой цикл называется ***циклом пересчета свободной клетки***.

Теорема 4. В таблице перевозок для каждой свободной клетки существует единственный цикл пересчета.

Клетки, в которых расположены в вершинах цикла, нумеруются в произвольном направлении так, что номер один присваивается пересчитываемой свободной клетке. ***Величиной сдвига по циклу*** θ называется минимальная из перевозок, стоящих в четных вершинах цикла. При пересчете плана величина θ вычитается из перевозок в четных вершинах цикла и прибавляется к перевозкам в нечетных вершинах (сдвиг по циклу).

Получить оптимальный план задачи можно при помощи ***метода потенциалов***. Сформулируем этапы этого метода: 1) найдем потенциалы поставщиков и потребителей как это указано в *примечании* к теореме 1; 2) проверим условие оптимальности (14) для всех свободных клеток. Если оно выполняется, то данный план оптимальный. Если в некоторых клетках условие (14) нарушено, то выделим клетку, в которой это нарушение $\Delta = C_{ij} - U_i - V_j > 0$ максимально и построим для этой клетки цикл пересчета; 3) проведем сдвиг по циклу на величину θ; 4) вернемся к этапу 1.

Метод потенциалов позволяет определять оптимальность плана и улучшать его с помощью цикла пересчета свободной клетки.

Пример.

Продолжим решение задачи, на примере которой рассмотрены методы нахождения первого опорного плана.

Напомним, что за опорное мы принимаем решение, полученное при расчете методом наименьшей стоимости.

С помощью метода потенциалов найдем оптимальное решение. Придадим потенциалу U_2 значение 0, а остальные определим из системы уравнений (13) для занятых клеток.

$$\begin{aligned} &U_2 + V_1 = 3 \qquad && U_1 + V_4 = 11 \\ &U_2 + V_4 = 15 && U_1 + V_2 = 3 \quad U_3 + V_3 = 2 \\ &U_2 + V_5 = 8 && U_3 + V_4 = 8 \end{aligned}$$

Из этой системы, подставляя в нее значение $U_2 = 0$, найдем остальные потенциалы и запишем их в таблицу. Например, $V_1 = 3 - 0 = 3$, $V_4 = 15 - 0 = 15$ и т. д.

Получим следующую таблицу.

	$V_1 = 3$	$V_2 = 4$	$V_3 = 6$	$V_4 = 12$	$V_5 = 8$	Запасы
$U_1 = -1$	4	3 19	4 +	11 21 –	9	40
$U_2 = 0$	3 20	4 2	7	15	8 8	30
$U_3 = -4$	7	4	2 7 –	8 3 +	15	10
Потребности	20	21	7	24	8	

Проверим выполнение условия оптимальности плана (14)

$$\begin{aligned} &U_1+V_1= -1 < 4 \qquad && U_2+V_3=9 > 7,\ \Delta = 2 \\ &U_1+V_3= 5 > 4,\ \Delta = 1 && U_3+V_1=-4 < 7 \\ &U_1+V_5= 4 < 9 && U_3+V_2= 0 < 4 \\ &U_2+V_2= 7 > 4,\ \Delta = 3 && U_3+V_5= 1 < 15. \end{aligned}$$

План не является оптимальным, так как имеются нарушения условий (14) в клетках A_1B_3, $A_2\,B_2$ и $A_2\,B_3$. Наибольшая невязка (нарушение) равна $\Delta=3$, следовательно, пересчет начинаем с клетки $A_2\,B_2$. Этой клетке присваивается номер 1 и строится цикл пересчета по правилам изложенным выше. Все клетки, в которых лежат вершины цикла, последовательно нумеруются, четным вершинам приписывается знак – , а нечетным + . Величина сдвига по циклу находится как минимальное значение из всех перевозок в четных вершинах цикла и равна $\theta = \min\{21;\ 2\} = 2$. Она вычитается из перевозок в четных вершинах цикла и прибавляется к перевозкам в нечетных. Получаем новый план, для которого вновь находим потенциалы и проверяем условия (14).

$U_1+V_1=2<4$ $U_2+V_4=12<15$
$U_1+V_3=5>4,\ \Delta=1$ $U_3+V_1=-1<7$
$U_1+V_5=7<9$ $U_3+V_2=0<4$
$U_2+V_3=6<7$ $U_3+V_5=4<15.$

План не является оптимальным, так как условие (14) не выполняется в клетке A_1B_3 с невязкой $\Delta=1$, следовательно, пересчет начинаем с клетки A_1B_3. Величина сдвига по циклу $\theta=\min\{21;7\}=7$ вычитается из перевозок в четных вершинах цикла и прибавляется к перевозкам в нечетных.

Получаем новый план, для которого вновь находим потенциалы и проверяем условия оптимальности. Из приведенного примера видно, что фактически цикл пересчета освобождает «невыгодные» клетки с большой стоимостью или, по крайней мере, уменьшает величину перевозки в таких клетках. Проверяем условие оптимальности:

	$V_1=3$	$V_2=4$	$V_3=5$	$V_4=12$	$V_5=8$	Запасы
U_1 = -1	4	3 19	4 **7**	11 14	9	40
U_2 = 0	3 20	4 2	7	15	8 8	30
U_3 = -4	7	4	2	8 10	15	10
Потребности	20	21	7	24	8	

$U_1+V_1=2<4$ $U_3+V_3=1<2$
$U_1+V_5=7<9$ $U_3+V_1=-1<7$
$U_2+V_3=5<7$ $U_3+V_2=0<4$
$U_2+V_4=12<15$ $U_3+V_5=4<15.$

Опорный план оптимален, так как все невязки равны нулю.
Вычисляем функцию цели

$$F(X_{\text{опт}})=19\cdot 3=7\cdot 4+14\cdot 11+20\cdot 3+2\cdot 4+8\cdot 8+10\cdot 8=451.$$

Ответ представим в виде

$$X_{\text{опт}}=\begin{pmatrix} 0 & 19 & 7 & 14 & 0 \\ 20 & 2 & 0 & 0 & 8 \\ 0 & 0 & 0 & 10 & 0 \end{pmatrix};\ F(X_{\text{опт}})=451.$$

Индивидуальные задания

Задание 2

На станциях A_i (i=1,2,3) сосредоточен однородный груз в количестве a_i единиц груза, который требуется перевезти на станции назначения B_j ($j=1,..5$) в соответствии с потребностями каждой станции в b_j единиц груза. Известны затраты c_{ij} на перевозку единицы груза с любой станции A_i на любую станцию B_j. Требуется составить такой план перевозок, чтобы весь груз был вывезен, все потребности были бы удовлетворены, а суммарные затраты были бы минимальны. Исходные данные приведены в табл. 4.1 и 4.2.

Таблица 4.1

Запасы и потребности на станциях − участниках процесса перевозок

Вар.	Запасы			Потребности				
	A_1	A_2	A_3	B_1	B_2	B_3	B_4	B_5
1	85	15	50	45	45	10	25	25
2	53	17	30	20	20	20	13	27
3	40	30	10	20	21	7	24	8
4	300	280	220	180	140	190	120	170
5	80	70	10	40	49	7	56	8
6	72	36	22	30	10	20	35	35
7	120	80	100	85	65	90	30	30
8	230	120	150	70	110	80	110	130
9	70	60	10	35	42	7	48	8
10	150	200	150	160	70	90	80	100

Таблица 4.2

Стоимости перевозки единицы груза

Вар.	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}
1	14	15	2	7	9	8	6	9	13	4	5	11	10	6	23
2	13	9	15	3	18	7	8	6	10	9	16	4	10	11	29
3	7	4	5	11	8	3	3	4	15	8	5	2	2	8	15
4	12	21	9	10	16	13	15	11	13	21	19	26	12	17	20
5	15	4	9	11	8	3	7	8	15	8	13	10	2	12	15
6	8	3	15	6	14	9	12	16	7	9	14	12	5	11	22
7	7	4	12	9	6	11	2	7	3	15	4	5	12	8	5
8	7	20	3	9	35	3	14	10	12	46	15	25	11	16	48
9	13	4	8	11	8	3	6	7	15	8	11	8	2	11	15
10	9	20	7	11	16	4	14	12	15	17	15	22	11	12	19

5. Основные понятия теории массового обслуживания (цепи Маркова)

Теория систем массового обслуживания стала развиваться с начала XX века. Лохансен в 1907 г. сформулировал основные предпосылки новой теории. В 1909 г. шведский математик Эрланг применил теорию вероятностей к исследованию зависимости обслуживания телефонных вызовов от числа поступивших на телефонную станцию вызовов. В нашей стране известный математик Хинчин систематизировал основные положения систем массового обслуживания в монографии «Теория очередей». Именно так называется за рубежом теория массового обслуживания. Данная теория в качестве математического аппарата использует понятия теории вероятностей случайных величин, а также некоторые другие понятия, которые будут введены в следующих разделах.

5.1. Марковские цепи с дискретным временем и конечным числом состояний

Пусть физическая система S находится в одном из состояний, образующих конечное множество $\{S_1,...,S_n\}$, и переходит из одного состояния в другое $S_i \to S_j$ случайным образом только в фиксированные моменты времени $t_1, t_2, \ldots$.

Тогда говорят, что в этой системе происходит *случайный процесс*.

Важным классом случайных процессов являются *марковские процессы* (или процессы без последействия), введенные в научный обиход математиком А. А. Марковым.

На интуитивном уровне это значит, что *при известном состоянии системы в данный момент времени прогноз о ее будущем поведении не зависит от состояний, в которых находилась эта система в прошлом*.

Поскольку множество состояний $\{S_1,...,S_n\}$ конечно, а множество моментов времени $\{t_1, t_2, ...\}$ состоит из изолированных точек, такой случайный процесс называется *марковской цепью с дискретным временем и конечным числом состояний.*

Как мы убедимся позже, такие цепи и некоторые их модификации могут служить удобной математической моделью для описания многих интересных и важных технических (в частности, транспортных) объектов.

Приведем математическое описание марковских цепей с дискретным временем, полагая для упрощения выкладок, что система может находиться только в трех состояниях S_1, S_2, S_3, т.е. принимая $n = 3$.

Обозначим символом p_{ij}, $i, j = 1, 2, 3$ *вероятность перехода* системы из состояния S_i в состояние S_j за один шаг. Такие вероятности образуют *матрицу P вероятностей переходов* за один шаг:

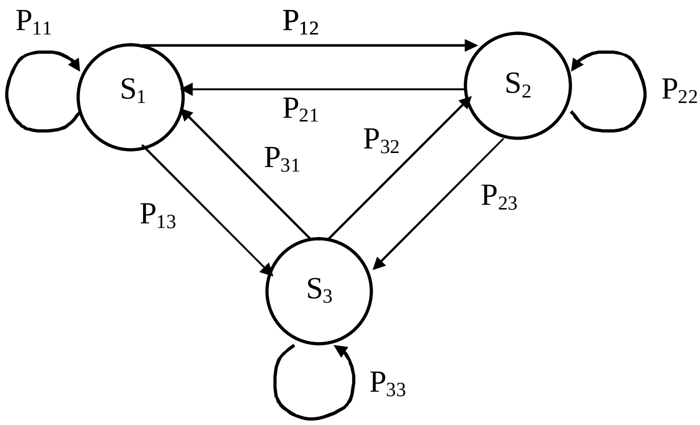
$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}. \qquad (15)$$

Будем считать, что вероятности p_{ij} не зависят от номера шага, на котором осуществляется переход $S_i \to S_j$, и называть такие цепи *однородными марковскими цепями.* В дальнейшем будем рассматривать только однородные марковские цепи.

Элементы матрицы (15) обладают следующими свойствами:

$$0 \le p_{ij} \le 1,\ i, j = 1, 2, 3;\ \sum_{j=1}^{3} p_{ij} = 1. \qquad (16)$$

Однородную марковскую цепь принято изображать в виде *размеченного графа состояний*.



Рассмотрим матрицу-строку $Q(k) = (p_1(k), p_2(k), p_3(k))$, где $p_i(k)$– вероятность того, что после k шагов система находится в состоянии S_i, $i = 1, 2, 3$. Для элементов этой матрицы при любом $k \ge 0$ выполняется равенство $\sum_{i=1}^{3} p_i(k) = 1$.

При $k = 0$ имеем матрицу $Q(0) = (p_1(0), p_2(0), p_3(0))$ – *начальное распределение вероятностей состояний.*

Матрицы $P, Q(k), Q(0)$ удовлетворяют следующим матричным уравнениям:

$$Q(k) = Q(k-1)P \text{ или } Q(k) = Q(0)P^k. \qquad (17)$$

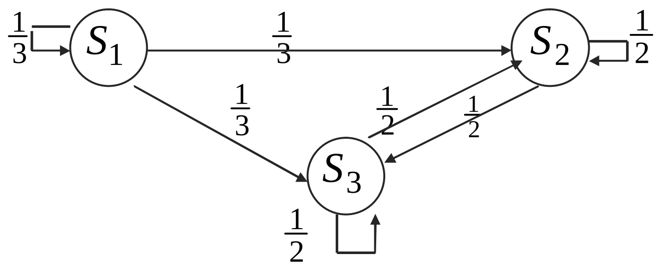
Эти уравнения позволяют определить вероятности состояний после k-го шага по известному начальному распределению и заданной матрице P вероятностей переходов за один шаг.

Введем следующие понятия.

Определение 1. Состояние s_i называется *существенным,* если, выйдя из этого состояния, система может в него вернуться за один или несколько шагов : $p_{ij}(n) > 0$ и существует k такое, что $p_{ji}(k) > 0$.

Состояние S_i называется *несущественным*, если, выйдя из S_i, система не может вернуться в него: $p_{ij}(n) > 0$, $p_{ji}(k) = 0$, для любого k.

Так, марковская цепь, граф которой приведен ниже, имеет одно несущественное состояние S_1.



Определение 2. Распределение вероятностей состояний марковской цепи называется *стационарным*, если оно не изменяется во времени.

Стационарное распределение Q удовлетворяет матричному уравнению

$$Q = Q \cdot P\ , \tag{18}$$

или в координатной форме
$$\begin{cases} p_1 = p_1 p_{11} + p_2 p_{21} + p_3 p_{31}, \\ p_2 = p_1 p_{12} + p_2 p_{22} + p_3 p_{32}, \\ p_3 = p_1 p_{13} + p_2 p_{23} + p_3 p_{33}. \end{cases} \tag{19}$$

Для получения единственного решения к уравнениям (18) или (19) всегда необходимо добавить условие нормировки:

$$p_1 + p_2 + p_3 = 1. \tag{20}$$

Определение 3. Марковская цепь называется *регулярной*, если из любого существенного состояния можно попасть в любое другое существенное состояние за конечное число шагов.

Определение 4. Вероятности $\tilde{p}_i = \lim\limits_{n \to \infty} p_i(n)$ называются *предельными* (или *финальными*) вероятностями состояний системы.

Теорема 1. Если марковская цепь регулярна, то предельные вероятности состояний системы совпадают со стационарными вероятностями.

Пример 1. в городе *N* каждый житель имеет одну из трех профессий: *A*, *B*, *C*. Дети отцов, имеющих профессии *A*, *B*, *C*, сохраняют профессии отцов с вероятностями $\frac{3}{5}$, $\frac{2}{3}$, $\frac{1}{4}$ соответственно, а если не сохраняют, то с равными вероятностями выбирают любую из двух других профессий.

Найти:

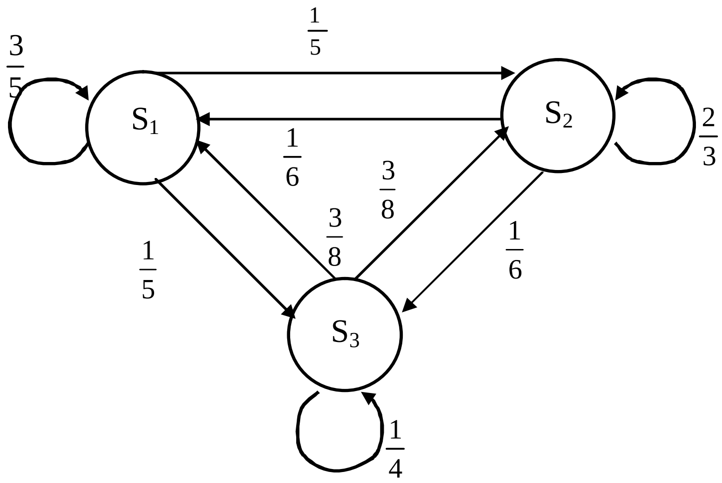
1) распределение по профессиям в следующем поколении, если в данном поколении профессию *A* имело 20 % жителей, *B* – 30 %, *C* – 50 % ;

2) предельное распределение по профессиям, когда число поколений растет;

3) распределение по профессиям, не меняющееся при смене поколений.

Решение

1. Составим граф состояний.



По графу видно, что все состояния системы существенны, поэтому цепь регулярна. Запишем матрицу вероятностей переходов за один шаг

$$P=\begin{pmatrix}\frac{3}{5}&\frac{1}{5}&\frac{1}{5}\\ \frac{1}{6}&\frac{2}{3}&\frac{1}{6}\\ \frac{3}{8}&\frac{3}{8}&\frac{1}{4}\end{pmatrix}$$

2. Найдем матрицу вероятностей переходов за два шага $P(2)=P(1)\cdot P(1)$:

$$P(2)=\begin{pmatrix}\frac{3}{5}&\frac{1}{5}&\frac{1}{5}\\ \frac{1}{6}&\frac{2}{3}&\frac{1}{6}\\ \frac{3}{8}&\frac{3}{8}&\frac{1}{4}\end{pmatrix}\begin{pmatrix}\frac{3}{5}&\frac{1}{5}&\frac{1}{5}\\ \frac{1}{6}&\frac{2}{3}&\frac{1}{6}\\ \frac{3}{8}&\frac{3}{8}&\frac{1}{4}\end{pmatrix}=\begin{pmatrix}\frac{281}{600}&\frac{197}{600}&\frac{122}{600}\\ \frac{197}{720}&\frac{389}{720}&\frac{67}{360}\\ \frac{61}{160}&\frac{201}{480}&\frac{1}{5}\end{pmatrix}.$$

Заметим, что вероятность перехода из состояния S_2 в состояние S_1 за один шаг $p_{21}(1)=\frac{1}{6}$, а за два шага $p_{21}(2)=\frac{197}{720}$, т. е. она увеличилась.

3. По условию задачи $P_1(0)=0{,}2;\ P_2(0)=0{,}3;\ P_3(0)=0{,}5$, т. е. $Q(0)=(0{,}2;\ 0{,}3;\ 0{,}5)$. Найдем распределение вероятностей состояний за один и за два шага, $Q(1)$:

$$Q(1)=(0{,}2;\ 0{,}3;\ 0{,}5)\cdot\begin{pmatrix}\frac{3}{5}&\frac{1}{5}&\frac{1}{5}\\ \frac{1}{6}&\frac{2}{3}&\frac{1}{6}\\ \frac{3}{8}&\frac{3}{8}&\frac{1}{4}\end{pmatrix}=\left(\frac{143}{400};\ \frac{171}{400};\ \frac{86}{400}\right).$$

4. По формулам (18) найдем стационарное распределение вероятностей:

$$\left(p_1, p_2, p_3\right)\cdot\begin{pmatrix}\frac{3}{5} & \frac{1}{5} & \frac{1}{5}\\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6}\\ \frac{3}{8} & \frac{3}{8} & \frac{1}{4}\end{pmatrix} = \left(p_1, p_2, p_3\right)\cdot$$

$$\begin{cases}\frac{3}{5}p_1 + \frac{1}{6}p_2 + \frac{3}{8}p_3 = p_1\\ \frac{1}{5}p_1 + \frac{2}{3}p_2 + \frac{3}{8}p_3 = p_2\\ \frac{1}{5}p_1 + \frac{1}{6}p_2 + \frac{1}{4}p_3 = p_3\end{cases}.$$

Данная система, следующая из системы (19), линейно зависима, так как должно выполняться очевидное условие для неизвестных $p_1 + p_2 + p_3 = 1$. Следовательно, при решении необходимо отбросить одно из первых трех уравнений и добавить указанное условие. Приведя все уравнения к стандартному виду, получим

$$\begin{cases}-\frac{2}{5}p_1 + \frac{1}{6}p_2 + \frac{3}{8}p_3 = 0,\\ \frac{1}{5}p_1 - \frac{1}{3}p_2 + \frac{3}{8}p_3 = 0,\\ p_1 + p_2 + p_3 = 1.\end{cases}$$

Тогда $Q_{\text{стац.}} = \left(\frac{15}{41}; \frac{18}{41}; \frac{8}{41}\right)$

5. Марковская цепь регулярна, поэтому предельные вероятности совпадают со стационарными. Получим $\tilde{p}_1 = \frac{15}{41}; \tilde{p}_2 = \frac{18}{41}; \tilde{p}_3 = \frac{8}{41}$.

Отметим, что вероятности всех трех состояний меняются на каждом шаге процесса, но сумма $p_1 + p_2 + p_3 = 1$ остается неизменной.

5.2. Однородные марковские цепи с непрерывным временем

Физическая система S может находиться в одном из состояний $\{S_1, S_2, \dots, S_n\}$, однако, в отличие от предыдущего будем считать, что переход из одного состояния в другое возможен *в любой случайный момент времени t*, причем множество таких моментов перехода является непрерывным. Если, как и ранее, выполняется марковское свойство (независимость будущего от прошлого при известном настоящем), то будем называть такой случайный процесс *марковской цепью с конечным числом состояний и непрерывным временем*.

Обозначим символом $Q(t)=(p_1(t),\ldots,p_n(t))$ *матрицу вероятностей состояний системы S в момент времени t*, $p_i(t)$ – вероятность того, что в момент времени *t* система находится в состоянии S_i. Ясно, что $\sum_{i=1}^{n} p_i(t)=1$. Для того чтобы найти эти вероятности, надо знать характеристики процесса, аналогичные вероятностям перехода p_{ij} за один шаг.

Определение 1. *Интенсивностью* λ_{ij} *перехода системы* $S_i \to S_j$ *называется предел* ,

$$\lambda_{ij}=\lim_{\Delta t\to 0}\frac{p_{ij}(\Delta t)}{\Delta t}, \qquad (21)$$

где $p_{ij}(\Delta t)$–вероятность перехода $S_i \to S_j\,(i\neq j)$ на интервале времени $[t, t+\Delta t]$.

Величины λ_{ij} при $i\neq j$ в общем случае зависят от расположения интервала $[t, t+\Delta t]$ на оси времени (являются функциями времени). Если, однако, $\lambda_{ij}=\text{const}$ при $i\neq j$, то *марковская цепь* называется *однородной*. Далее рассматриваются только такие марковские цепи.

Равенство (21) можно с точностью до бесконечно малых более высокого порядка, чем Δt заменить следующим приближенным равенством:

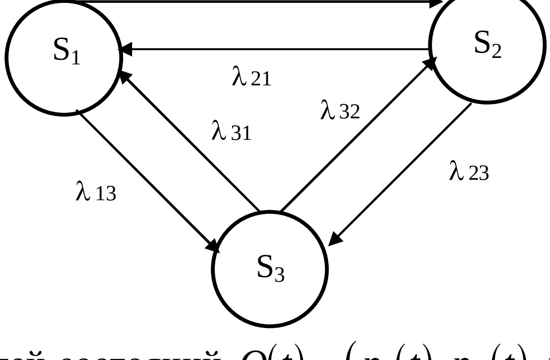
$$p_{ij}(\Delta t)\approx \lambda_{ij}\Delta t. \qquad (22)$$

Положим, как и ранее, для упрощения выкладок *n* = 3 и рассмотрим следующую *матрицу интенсивностей переходов:*

$$\Lambda=\begin{pmatrix}\lambda_{11} & \lambda_{12} & \lambda_{13}\\ \lambda_{21} & \lambda_{22} & \lambda_{23}\\ \lambda_{31} & \lambda_{32} & \lambda_{33}\end{pmatrix}. \qquad (23)$$

Элементы матрицы интенсивностей Λ удовлетворяют условиям: $\lambda_{ij}\geq 0$ при $i\neq j$; $\lambda_{ii}<0$ при i = 1, 2, 3; $\sum_{j=1}^{3}\lambda_{ij}=0$ при i = 1, 2, 3. (24)

Марковскую цепь с непрерывным временем можно изображать *размеченным графом состояний*



Матрица вероятностей состояний $Q(t)=(p_1(t), p_2(t), p_3(t))$ удовлетворяет *системе дифференциальных уравнений Колмогорова.* Эта система может быть записана в матричной форме

$$Q'(t) = Q(t)\Lambda \qquad (25)$$

или в координатной форме

$$\begin{aligned} p_1'(t) &= \lambda_{11}p_1(t) + \lambda_{21}p_2(t) + \lambda_{31}p_3(t) \\ p_2'(t) &= \lambda_{12}p_1(t) + \lambda_{22}p_2(t) + \lambda_{32}p_3(t) \\ p_3'(t) &= \lambda_{13}p_1(t) + \lambda_{23}p_2(t) + \lambda_{33}p_3(t). \end{aligned} \qquad (26)$$

Здесь $Q'(t) = \left(p_1'(t), p_2'(t), p_3'(t)\right)$ – матрица, составленная из производных $p_i'(t)$, (i = 1, 2, 3) вероятностей состояний в момент времени t. Для решения системы (25) или (26) необходимо, как обычно, задать начальные условия:

$$p_1(0) = p_1^0,\ p_2(0) = p_2^0,\ p_3(0) = p_3^0, \qquad (27)$$

где p_i^0, i = 1, 2, 3 –заданные числа, причем $p_1^0 + p_2^0 + p_3^0 = 1$.

Определение 2. Распределение вероятностей называется *стационарным*, если вероятности p_i состояний S_i не зависят от времени $p_1(t) \equiv p_1$, $p_2(t) \equiv p_2$, $p_3(t) \equiv p_3$ или $Q(t) \equiv Q = \text{const}$.

Для стационарного распределения вероятностей системы $Q'(t) \equiv 0$, тогда из (25) получаем систему алгебраических уравнений

$$Q \cdot \Lambda = 0,$$

или в координатной форме:

$$\begin{cases} \lambda_{11}p_1 + \lambda_{21}p_2 + \lambda_{31}p_3 = 0 \\ \lambda_{12}p_1 + \lambda_{22}p_2 + \lambda_{32}p_3 = 0 \\ \lambda_{13}p_1 + \lambda_{23}p_2 + \lambda_{33}p_3 = 0. \end{cases} \qquad (28)$$

Уравнения для стационарного случая можно составить непосредственно по графу: сумма произведений $\lambda_{ij}p_i$ для дуг, выходящих из состояния S_i, равна сумме произведений $\lambda_{ji}p_j$ для дуг, входящих в состояние S_i.

Для многих практических задач важно знать, как ведут себя вероятности состояний системы при неограниченном возрастании времени. Обозначим $\tilde{p}_i = \lim\limits_{t\to\infty} p_i(t)$, $\tilde{p}_i$ называются *предельными (финальными)* вероятностями системы.

Понятия существенного (несущественного) состояния, регулярности марковской цепи аналогичны тем, что даны для марковской цепи с дискретным временем.

Если марковская цепь регулярна, то предельные вероятности состояний совпадают со стационарными.

Система, для которой существуют предельные вероятности состояний, называется *эргодической* и возникающий в ней случайный процесс – *эргодическим*.

Пример 1. В прибор входят два устройства *A* и *B*. Интенсивность отказов устройства *A* – $\lambda_1 = 1$, устройства *B* – $\lambda_2 = 2$. После выхода из строя устройство сразу начинают ремонтировать. Интенсивность ремонта устройства *A* – $\mu_1 = 2$, устройства *B* – $\mu_2 = 3$.

1. Описать состояния системы и составить граф состояний марковского процесса.

2. Записать дифференциальные уравнения вероятностей состояний системы.

3. Найти стационарное распределение и финальные вероятности.

4. Известно, что второе устройство более современно и имеет производительность вдвое больше, чем первое. Первое устройство приносит доход 5 условных единиц в единицу времени, а второе – 10 условных единиц. Вычислить доход, который приносят эти устройства.

5. (Задача о рационализации.) Предлагается произвести рационализацию процесса, в результате которой удастся сократить время ремонта вдвое. Но мы можем (из-за финансовых трудностей) применить ее только к одному из устройств. Показать расчетами, к какому из устройств надо применить рационализацию.

Решение

1. Опишем состояния системы:

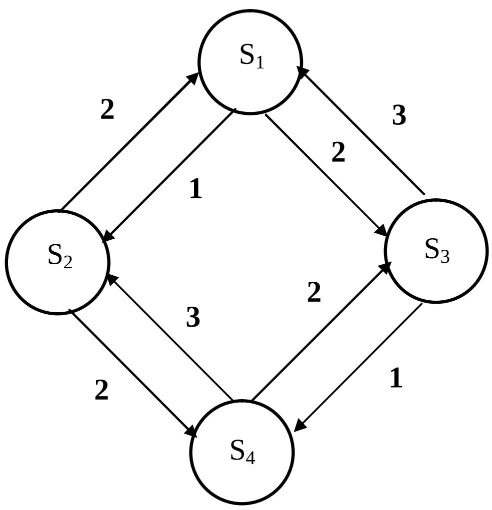
S_1 – оба устройства работают;

S_2 – первое устройство ремонтируется, второе – работает;

S_3 – второе устройство ремонтируется, первое – работает;

S_4 – оба устройства ремонтируются.

Граф данной марковской цепи изображен на рисунке



Матрица интенсивностей переходов имеет вид:

$$\Lambda = \begin{pmatrix} -3 & 1 & 2 & 0 \\ 2 & -4 & 0 & 2 \\ 3 & 0 & -4 & 1 \\ 0 & 3 & 2 & -5 \end{pmatrix}.$$

2. Составим систему дифференциальных уравнений Колмогорова:

$$\left(p_1'(t), p_2'(t), p_3'(t), p_4'(t)\right) = \left(p_1(t), p_2(t), p_3(t), p_4(t)\right) \cdot \begin{pmatrix} -3 & 1 & 2 & 0 \\ 2 & -4 & 0 & 2 \\ 3 & 0 & -4 & 1 \\ 0 & 3 & 2 & -5 \end{pmatrix}.$$

Тогда получим:

$$p_1' = -3p_1 + 2p_2 + 3p_3$$
$$p_2' = p_1 - 4p_2 + 3p_4$$
$$p_3' = 2p_1 + 2p_4 - 4p_3$$
$$p_4' = 2p_2 + p_3 - 5p_4$$

(при расчете мы использовали правило умножения двух матриц).

3. Найдем стационарные вероятности системы, полагая, что $p_1' = 0$, $p_2' = 0$, $p_3' = 0$, $p_4' = 0$. Получим систему линейных алгебраических уравнений:

$$\begin{cases} -3p_1 + 2p_2 + 3p_3 = 0 \\ p_1 - 4p_2 + 3p_4 = 0 \\ 2p_1 + 2p_4 - 4p_3 = 0 \\ 2p_2 + p_3 - 5p_4 = 0 \end{cases}.$$

Четвертое уравнение этой системы получается из первых трех. Его можно отбросить и присоединив к системе уравнений условие нормировки $p_1 + p_2 + p_3 + p_4 = 1$, решить систему:

$$\begin{cases} -3p_1 + 2p_2 + 3p_3 = 0 \\ p_1 - 4p_2 + 3p_4 = 0 \\ 2p_1 - 4p_3 + 2p_4 = 0 \\ p_1 + p_2 + p_3 + p_4 = 1 \end{cases}.$$

Решая систему, получим:

$$p_1 = \frac{6}{15};\ p_2 = \frac{3}{15};\ p_3 = \frac{4}{15};\ p_4 = \frac{2}{15},\ Q \approx (0{,}4;\ 0{,}2;\ 0{,}27;\ 0{,}13).$$

Таким образом, система находится 40 % времени в состоянии S_1, 20 % времени в состоянии S_2, 27 % времени в состоянии S_3 и оставшиеся 13 % времени в четвертом состоянии.

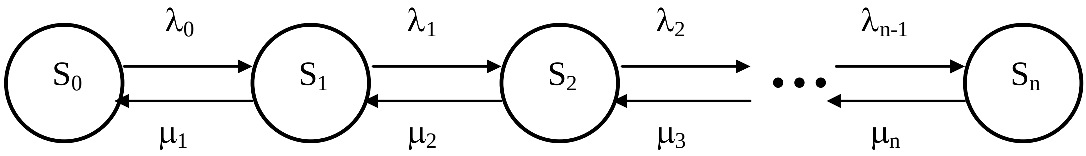
4. Подсчитаем доход, который приносят эти устройства

$$D = 0{,}4(5+10)+0{,}2\cdot 10+0{,}27\cdot 5 = 9{,}35.$$

5. Далее в задаче просят провести рационализацию одного из устройств. При этом сокращается время ремонта и тогда в первом случае изменяются две величины $p_{21} = p_{43} = 4$, а во втором – две другие величины $p_{31} = p_{42} = 6$. Предлагаем провести новые расчеты и вычислить доходы при этих рационализациях. Нужно будет выбрать ту рационализацию, которая дает наибольший доход.

5.3. Процесс гибели и размножения

Процессом гибели и размножения называется однородная марковская цепь с непрерывным временем, размеченный граф состояний которой имеет вид:



Термин этот заимствован из математической биологии, а его *суть состоит в том, что за малое время* Δt *система практически может перейти только в соседнее состояние* (или остаться на месте), но не может перескочить через одно или несколько состояний.

Найдем для этого процесса стационарное распределение. Составляя по графу систему алгебраических уравнений, получим для стационарных вероятностей $p_0, \ldots, p_n$ следующую систему:

$$\begin{aligned} &\lambda_0 p_0 = \mu_1 p_1 \\ &(\lambda_1 + \mu_1) p_1 = \lambda_0 p_0 + \mu_2 p_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\lambda_{n-1} p_{n-1} = \mu_n p_n. \end{aligned} \qquad (29)$$

Выразим из этой системы $p_1, \ldots, p_n$ через p_0 и подставим в условие нормировки $p_0 + p_1 + \ldots + p_n = 1$.

После несложных преобразований получим окончательно

$$p_0 = \left(1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} + \cdots + \frac{\lambda_0 \ldots \lambda_{n-1}}{\mu_1 \ldots \mu_n}\right)^{-1},$$

$$p_i = \frac{\lambda_0 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i} p_0, \; i = 1, \ldots, n. \qquad (30)$$

Эти формулы широко используются в следующих разделах, поскольку процессы, изучаемые в теории массового обслуживания, как правило, являются процессами размножения и гибели.

Отметим также, что этот процесс является регулярным, поэтому стационарное распределение (30) совпадает с предельным.

Индивидуальные задания

Задание 3

Дана матрица переходных вероятностей марковской цепи с дискретным временем, требуется

а) составить граф марковской цепи;

б) найти вероятности переходов из одного состояния в другое за два шага;

в) определить стационарные и финальные вероятности состояний системы.

1.1.
$$P=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix}$$

1.2
$$P=\begin{pmatrix}\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{3}&0&\frac{2}{3}\\\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\end{pmatrix}$$

1.3.
$$P=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}&0\\\frac{2}{3}&\frac{1}{3}&0\\\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\end{pmatrix}$$

1.4
$$P=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\\\frac{2}{3}&0&\frac{1}{3}\end{pmatrix}$$

1.5
$$P=\begin{pmatrix}\frac{3}{4}&0&\frac{1}{4}\\\frac{2}{3}&\frac{1}{3}&0\\\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\end{pmatrix}$$

1.6
$$P=\begin{pmatrix}\frac{5}{6}&\frac{1}{6}&0\\\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\end{pmatrix}$$

1.7
$$P=\begin{pmatrix}\frac{1}{4}&\frac{1}{4}&\frac{1}{2}\\\frac{3}{4}&0&\frac{1}{4}\\0&\frac{2}{3}&\frac{1}{3}\end{pmatrix}$$

1.8
$$P=\begin{pmatrix}0&\frac{2}{3}&\frac{1}{3}\\\frac{2}{3}&\frac{1}{3}&0\\\frac{1}{4}&\frac{1}{4}&\frac{1}{2}\end{pmatrix}$$

1.9
$$P=\begin{pmatrix}\frac{4}{6}&\frac{1}{6}&\frac{1}{6}\\\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\\frac{1}{2}&\frac{1}{2}&0\end{pmatrix}$$

1.10
$$P=\begin{pmatrix}\frac{3}{4}&0&\frac{1}{4}\\\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\\\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\end{pmatrix}$$

1.11
$$P=\begin{pmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{3}\\\frac{3}{4}&\frac{1}{4}&0\\\frac{1}{3}&0&\frac{2}{3}\end{pmatrix}$$

1.12
$$P=\begin{pmatrix}\frac{1}{6}&\frac{2}{3}&\frac{1}{6}\\\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\\\frac{1}{2}&\frac{2}{3}&0\end{pmatrix}$$

Задание 4

Задана матрица интенсивностей переходов непрерывной цепи Маркова Λ, требуется:

а) составить размеченный граф состояний системы, соответствующий этой матрице;

б) записать систему дифференциальных уравнений Колмогорова;

в) найти предельное и финальное распределение вероятностей состояний.

2.1
$$\Lambda = \begin{pmatrix} -2 & 0 & 2 \\ 2 & -3 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

2.2
$$\Lambda = \begin{pmatrix} -4 & 2 & 2 \\ 0 & -3 & 3 \\ 1 & 1 & -2 \end{pmatrix}$$

2.3
$$\Lambda = \begin{pmatrix} -4 & 1 & 3 \\ 0 & -3 & 3 \\ 2 & 0 & -2 \end{pmatrix}$$

2.4
$$\Lambda = \begin{pmatrix} -6 & 1 & 5 \\ 3 & -4 & 1 \\ 2 & 0 & -2 \end{pmatrix}$$

2.5
$$\Lambda = \begin{pmatrix} -2 & 2 & 0 \\ 2 & -5 & 3 \\ 1 & 3 & -4 \end{pmatrix}$$

2.6
$$\Lambda = \begin{pmatrix} -3 & 1 & 2 \\ 1 & -3 & 2 \\ 1 & 1 & -2 \end{pmatrix}$$

2.7
$$\Lambda = \begin{pmatrix} -5 & 3 & 2 \\ 1 & -5 & 4 \\ 0 & 1 & -1 \end{pmatrix}$$

2.8
$$\Lambda = \begin{pmatrix} -4 & 2 & 2 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

2.9
$$\Lambda = \begin{pmatrix} -5 & 0 & 5 \\ 2 & -3 & 1 \\ 1 & 3 & -4 \end{pmatrix}$$

2.10
$$\Lambda = \begin{pmatrix} -7 & 2 & 5 \\ 1 & -5 & 4 \\ 0 & 3 & -3 \end{pmatrix}$$

2.11
$$\Lambda = \begin{pmatrix} -5 & 3 & 2 \\ 1 & -2 & 1 \\ 3 & 2 & -5 \end{pmatrix}$$

2.12
$$\Lambda = \begin{pmatrix} -8 & 5 & 3 \\ 2 & -5 & 3 \\ 1 & 7 & -8 \end{pmatrix}$$

6. Системы массового обслуживания

6.1. Основные понятия и классификация систем массового обслуживания

При исследовании транспортных процессов приходится сталкиваться с работой своеобразных систем. Примерами могут служить билетные кассы, АТС, справочные бюро, пункты экипировки локомотивов и другие.

Математический аппарат, изучающий закономерности функционирования систем, удовлетворяющих массовый спрос, в том числе образование очередей в такого рода системах, называется *теорией массового обслуживания*.

Методы теории массового обслуживания все более широко применяются на транспорте для расчета рациональной организации подачи вагонов под разгрузку и погрузку, для расчета мощностей пунктов текущего ремонта вагонов, для выбора рационального числа кассовых аппаратов на вокзалах и станциях метрополитена. Поэтому современному инженеру железнодорожного транспорта необходимо знать основы теории массового обслуживания.

Заявкой (требованием) называется спрос на удовлетворение какой-либо потребности. Далее примем, что все заявки однотипные. Удовлетворение спроса называется *обслуживанием заявки*.

Системой массового обслуживания (СМО) называется любая система, предназначенная для обслуживания каких-либо заявок, поступающих в нее в случайные моменты времени. Устройство, непосредственно обслуживающее заявку, называется *каналом обслуживания*. СМО может содержать одно такое устройство, тогда она называется *одноканальной*. Если СМО содержит несколько обслуживающих устройств, то она называется *многоканальной*.

Поступление заявки в СМО назовем *событием*. Последовательность событий, состоящих в поступлении заявок в СМО, назовем *входящим потоком заявок*. Последовательность событий, состоящих в выходе заявок из СМО, назовем *выходящим потоком заявок*.

В зависимости от поведения заявки в СМО различают СМО *с отказами* и СМО *с очередью* (или с ожиданием). В СМО *с отказами* заявка, поступившая в момент занятости всех каналов, получает отказ и покидает СМО. В СМО *с очередью* (или ожиданием) заявка, поступившая в момент занятости всех каналов, становится в очередь и ожидает освобождения одного из каналов обслуживания.

Возможны СМО *смешанного типа*. Например, СМО с ограниченной очередью. В такой СМО заявка становится в очередь при занятости всех каналов, если очередь невелика и, скажем, не достигла длины m. Если все m мест в очереди заняты, заявка покидает СМО. К СМО смешанного типа относятся СМО с ограниченным временем ожидания. Заявка, поступившая в момент занятости всех каналов, становится в очередь, но уходит из СМО необслуженной, если время ожидания слишком велико.

СМО могут быть открытого и замкнутого типа. В *открытых* СМО интенсивность поступающего в нее потока заявок не зависит от состояния самой СМО, так как круг «клиентов» (поступающих заявок) практически не ограничен. Примерами таких СМО являются вокзальные кассы, метрополитен, телевизионные ателье больших городов и т. д. В СМО *замкнутого* типа обслуживается ограниченный круг «клиентов», поэтому интенсивность потока заявок существенно зависит от состояния системы. Примерами таких СМО являются различные ремонтные системы в автопарках, цехах и т. д.

СМО с очередью и смешанного типа различаются также по *дисциплине обслуживания*: обслуживаются ли заявки в порядке поступления или в случайном порядке, или есть заявки, которые обслуживаются вне очереди (*СМО с приоритетом*).

6.2. Простейший поток и его свойства

Рассмотрим входящий поток в СМО как последовательность точек $t_0, t_1, t_2, ..., t_i, ...$ моментов поступления заявок на оси времени Ot. Здесь t_0 – начальный момент времени.

Поток заявок назовем *простейшим*, если он удовлетворяет трем условиям:

1. *Отсутствие последействия*. Это условие означает, что заявки поступают в СМО независимо друг от друга, т. е. поступление заявки после момента времени t не зависит от того, когда и в каком количестве появлялись заявки до момента t.

2. *Стационарность*. Это условие означает, что вероятность поступления некоторого числа заявок в СМО за время Δt зависит лишь от длины интервала $\Delta t = (t + \Delta t) - t$ и не зависит от точки t отсчета этого интервала на оси времени Ot. Если выполнено условие стационарности, то можно говорить о среднем числе заявок, поступающих в СМО за единицу времени, например за один час, не указывая, за какой именно час.

3. *Ординарность*. Это условие означает, что одновременное поступление в СМО двух и более заявок маловероятно, т. е. вероятность появления за бесконечно малое время Δt более чем одной заявки есть бесконечно малая более высокого порядка малости, чем Δt.

Таким образом, если поток заявок простейший, то случайные моменты времени t_i (i = 1, 2, ...) поступления заявок в СМО распределены на оси времени со средней плотностью λ (стационарность); эти точки попадают в непересекающиеся интервалы независимо друг от друга (нет последействия); заявки поступают в СМО поодиночке (ординарность). Величина λ называется *интенсивностью* потока заявок и представляет собой *среднее число (математическое ожидание числа) заявок, поступающих в систему за единицу времени*.

Можно показать, что для простейшего потока вероятность $p_i(t)$ поступления в СМО ровно i заявок за время t вычисляется по формуле

$$p_i(t) = \frac{(\lambda t)^i}{i!} e^{-\lambda t}, \quad (i \geq 0), \tag{31}$$

т. е. вероятности $p_i(t)$ распределены по закону Пуассона с параметром λt. В связи с этим простейший поток часто называют *пуассоновским* потоком.

Обозначим через T интервал времени между поступлениями двух последовательных заявок. Найдем функцию распределения случайной величины T

$$F(t) = P(T < t) = 1 - p_0(t),$$

где $P(T < t)$ вероятность того, что случайная величина T примет значение, меньшее, чем t; p_0 – вероятность противоположного события (т. е. за время t в СМО не поступила ни одна заявка). По формуле (31) имеем

$$p_0(t) = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t},$$

откуда

$$F(t) = 1 - e^{-\lambda t}, \quad (t > 0). \tag{32}$$

Найдем плотность распределения случайной величины T:

$$f(t) = F'(t) = \lambda e^{-\lambda t}, \quad (t > 0).$$

Определяя математическое ожидание и дисперсию случайной величины T, получим $M[T] = \frac{1}{\lambda}, \quad D[T] = \frac{1}{\lambda^2}, \quad \sigma = \sqrt{D[T]} = \frac{1}{\lambda}.$ (33)

6.3. Марковские системы массового обслуживания

Для задания СМО необходимо задать вероятностные характеристики времени обслуживания одной заявки. Обозначим это время через $T_{обсл}$. Величина $T_{обсл}$ является случайной. Во многих задачах теории массового обслуживания закон распределения времени обслуживания предполагается показательным, т.е.

$$F(t) = P\{T_{обсл} < t\} = 1 - e^{-\mu t}. \tag{34}$$

Параметр этого распределения μ есть величина, обратная среднему времени обслуживания, т. е.

$$\mu = \frac{1}{M[T_{обсл.}]}. \tag{35}$$

Часто μ называют *интенсивностью потока обслуживания, которая означает среднее число заявок, обслуживаемых одним каналом в единицу времени*. При этом под потоком обслуживания понимается поток заявок, обслуживаемых друг за другом одним непрерывно занятым каналом. Если $T_{обсл}$ представляет собой случайную величину, имеющую показательное распределение, то поток обслуживания является *простейшим*.

Если входящий поток и все потоки обслуживания простейшие, то процесс, протекающий в СМО, является *марковским случайным процессом (цепью) с дискретными состояниями и непрерывным временем*. Поэтому СМО, в которой все потоки простейшие, называют *марковской СМО*.

Таким образом, предположение о показательном законе распределения времени обслуживания и интервала времени между двумя последовательными поступлениями заявок играет исключительную роль в теории массового обслуживания, так как упрощает аналитическое исследование СМО, сводя его к исследованию цепей Маркова.

6.4. Показатели эффективности систем массового обслуживания

Обычно практический интерес в теории массового обслуживания вызывают предельные средние характеристики системы, которые называются *показателями эффективности СМО*. В качестве показателей эффективности для стационарного режима могут рассматриваться следующие:

A – среднее число заявок, обслуживаемое СМО в единицу времени. Эту характеристику называют *абсолютной пропускной способностью СМО;*

Q – вероятность обслуживания поступившей заявки, или *относительная пропускная способность СМО.* Очевидно,

$$Q = \frac{A}{\lambda};$$

$P_{\text{отк}}$ – вероятность отказа, т. е. вероятность того, что поступившая заявка не будет обслужена, $P_{\text{отк}} = 1 - Q$ *;*

$\bar{z}$ – среднее число заявок в СМО (имеются в виду все заявки, как обслуживаемые, так и ожидающие очереди, если она есть);

$\bar{r}_{\text{о}}$ – среднее число заявок в очереди, если она есть;

$\bar{t}_{\text{сист}} = \dfrac{z_{\text{сист}}}{\lambda}$ – среднее время пребывания заявки в СМО, как в очереди, если она есть, так и в момент обслуживания;

$\bar{t}_{\text{о}} = \dfrac{r_{\text{о}}}{\lambda}$ – среднее время пребывания заявки в очереди;

$\bar{k} = \dfrac{A}{\mu}$ – среднее число занятых каналов.

Выбор показателей эффективности СМО зависит от типа СМО. Например, абсолютная пропускная способность *A*, являясь основной характеристикой обслуживания в СМО с отказами, совпадает с интенсивностью потока для СМО с неограниченной очередью.

6.5. Многоканальная СМО с отказами (задача Эрланга)

Задача Эрланга хронологически является одной из первых, классических задач массового обслуживания. Возникла она из нужд телефонии и была решена датским математиком Эрлангом.

Пусть имеется n каналов обслуживания (линий связи), на которые поступает поток заявок с интенсивностью λ. Поток обслуживания имеет интенсивность μ. Потоки в СМО – простейшие.

Найти предельные вероятности состояний СМО, вычислить характеристики эффективности ее работы:

$A-$ абсолютную пропускную способность;

$Q-$ относительную пропускную способность;

$P_{\text{отк}}$ – вероятность того, что заявка не будет обслужена;

$\bar{k}-$ среднее число занятых каналов.

Решение.

Обозначим $S_0, S_1, ..., S_n$ состояния системы массового обслуживания:

$S_0 -$ в системе нет заявок,

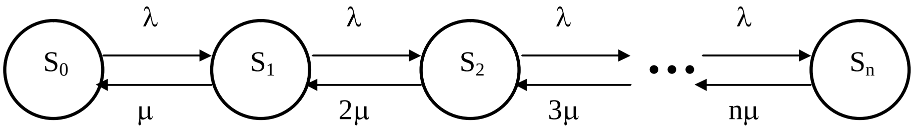
$S_1 -$ в системе одна заявка,

....

$S_n -$ в системе n заявок.

Следующая заявка, пришедшая в систему, получит отказ. p_n – вероятность отказа нашей системы: $p_{\text{отк}} = p_n$.

Граф состояний представлен на рисунке



Входящий поток простейший с интенсивностью λ, поэтому переход из состояния S_i в S_{i+1} происходит с одной и той же интенсивностью, так как входной поток формируется вне СМО и не зависит от состояния системы.

Переходы из состояния S_k в соседнее S_{k-1} происходит с интенсивностью $k\mu$ –(кратное количеству занятых каналов). Граф на рисунке соответствует графу процесса гибели и размножения. Согласно формулам схемы гибели и размножения получим

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2!\mu^2} + \frac{\lambda^3}{3!\mu^3} + \ldots + \frac{\lambda^n}{n!\mu^n}\right)^{-1} \quad (36)$$

$$p_i = \frac{\lambda^i}{i!\mu^i} \cdot p_0, \; i = 1,2,\ldots,n\ .$$

Обозначим $\frac{\lambda}{\mu} = \rho$, тогда получим:

$$p_0 = \left(1 + \rho + \frac{\rho^2}{2!} + \ldots + \frac{\rho^n}{n!}\right)^{-1}; \; p_i = \frac{\rho^i}{i!} \cdot p_0, \; i = 1,2,\ldots n \quad (37)$$

Величина ρ определяет среднее число заявок, приходящее за среднее время обслуживания одной заявки. Его называют коэффициентом *загрузки системы.*

Вычислим коэффициенты эффективности работы системы массового обслуживания с отказами:

$$P_{\text{отк}} = p_n = \frac{\rho^n}{n!} \cdot p_0\ ; \; Q = 1 - P_{\text{отк}} = 1 - \frac{\rho^n}{n!} \cdot p_0\ ; \; A = Q \cdot \lambda. \quad (38)$$

Найдем среднее число занятых каналов

k	0	1	...	n
p_i	p_0	p_1	...	p_n

$$M[k] = \overline{k} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \ldots + n \cdot p_n,$$

$$\bar{k} = \left(\rho + 2\cdot\frac{\rho^2}{2!} + 3\frac{\rho^3}{3!} + \ldots + n\cdot\frac{\rho^n}{n!}\right)\cdot p_0 \cdot \qquad (39)$$

Можно найти среднее число занятых каналов с помощью таких рассуждений: A – абсолютная пропускная способность СМО, μ – интенсивность потока обслуженных системой заявок. Каждый канал в единицу времени обслуживает в среднем μ заявок. Значит, среднее число занятых каналов равно

$$\bar{k} = \frac{A}{\mu} = \rho\left(1 - \frac{\rho^n}{n!}p_0\right). \qquad (40)$$

Пример 6.5.1. Простейшая СМО представляет собой автоматическую телефонную станцию с пятью линиями связи (число каналов – n = 5).

Если заявка приходит в момент, когда все каналы заняты, то она получает отказ. Интенсивность потока заявок $\lambda = 0{,}7$ заявки в минуту, среднее время обслуживания одной заявки $t_{\text{обсл}} = 3{,}4$ мин.

1. Описать состояния СМО, построить ее граф.

2. Найти предельные вероятности состояний и характеристики эффективности работы АТС. Оценить ее работу.

3. Найти зависимость среднего числа занятых каналов и пропускной способности АТС от интенсивности входного потока, сделать выводы.

4. Сколько всего потребуется каналов для того, чтобы удовлетворить не менее 99 % поступающих заявок? Какая доля каналов при этом будет простаивать?

5. Содержание каждого канала в течение месяца обходится в 10 тысяч усл. ед. Каждая обслуженная заявка приносит доход в 1,5 усл. ед. Определить приносит ли АТС доход от обслуживания всех заявок.

Решение

1. Обозначим $S_0, S_1, S_2, ..., S_5$ – состояния СМО:

S_0 – в системе все каналы свободны;

S_1 – один канал занят, 4 – свободны;

S_2 – два канала заняты, 3 – свободны;

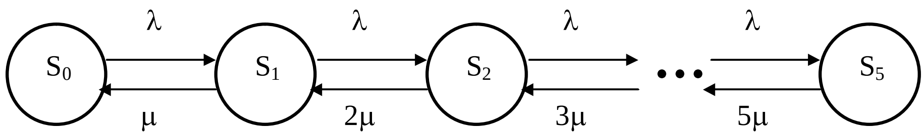
S_3 – три канала заняты, 2 – свободны;

S_4 – четыре канала заняты, 1 – свободен;

S_5 – все пять каналов заняты.

Следующая заявка, поступившая в СМО, получает отказ.

Граф системы массового обслуживания имеет вид:



2. Из условия задачи известно:

$\lambda = 0{,}7\,{}^{1}\!/_{\text{мин}}$; $\bar{t}_{\text{обсл}} = 3{,}4$ мин; тогда $\mu = \dfrac{1}{t_{\text{обсл}}} = \dfrac{1}{3{,}4} = 0{,}29\,{}^{1}\!/_{\text{мин}}$;

$\rho = 0{,}7\,{}^{1}\!/_{\text{мин}} \cdot 3{,}4 \text{ мин} = 2{,}38$.

По формулам (37−40) найдем нужные вероятности и характеристики:

$$p_0 = \left(1 + 2{,}38 + \frac{(2{,}38)^2}{2!} + \frac{(2{,}38)^3}{3!} + \frac{(2{,}38)^4}{4!} + \frac{(2{,}38)^5}{5!}\right)^{-1} = 0{,}096;$$

$p_i = p_0 \cdot \dfrac{\rho^i}{i!}$; $i = 1, 2, 3, 4, 5$;

$p_1 = 0{,}228$; . . . $p_5 = 0{,}061$.

Тогда $P_{\text{отк}} = p_5 = 0{,}061$.

Это значит, что в нашей АТС 6 % абонентов получают отказ:

$Q = 1 - P_{\text{отк}} = 0{,}939$; $A = Q \cdot \lambda = 0{,}657\,{}^{1}\!/_{\text{мин}}$.

Среднее число занятых каналов − $\bar{k} = 2{,}23 = \bar{z}_{\text{сист}}$.

В системе заявка в среднем находится $\bar{t}_{\text{сист}} = 3{,}19$ мин (по формулам Литтла: $\bar{t}_{\text{сист}} = {}^{\bar{z}_{\text{сист}}}\!/_{\lambda}$).

Результаты расчетов позволяют сделать вывод, что при достаточно высокой вероятности отказа, загруженность АТС составляет около 45 % каналов.

Дальнейшая работа проводится с использованием компьютера.

3. Так как по условию мы можем менять лишь один параметр (число каналов), то, задавая значения n = 6,7…, получим значения для $P_{\text{отк}}$:

n	5	6	7
$P_{\text{отк}}$	0,061	0,023	0,008

Следовательно, для того чтобы $P_{\text{отк}} < 0{,}01$, достаточно использовать 7 каналов.

4. Расчет на ЭВМ при $n = 5$ дает:

λ	0.7	1.0	2.0	4.0	10.0	50.0	200
$\bar{k}$	2.23	2.91	4.04	4.62	4.84	4.97	4.99

λ	0.7	1.0	2.0	4.0	10.0	50.0	200
$\bar{A}$	0.66	0.85	1.17	1.34	1.42	1.46	1.47

Из таблиц следует, что обе рассматриваемые величины с ростом интенсивности входного потока возрастают, стремясь к некоторым предельным значениям. При этом среднее число занятых каналов стремится к полному числу каналов в СМО (n = 5), а пропускная способность − к некоторой величине, больше которой при данном числе каналов и среднем времени обслуживания АТС пропускать не может. Отметим, что фактически при больших λ СМО не работает из-за неприемлемых значений $p_{\text{отк}}$ (так при $\lambda = 200$, $p_{\text{отк}} = 0.99$).

5. На АТС обслуживается в среднем A = 0.66 (заяв./мин) с доходом d = 1.5 усл. ед., тогда прибыль за одну минуту, $D = d \cdot A = 0{,}99\,{}^{\text{усл.ед}}\!/_{\text{мин}}$, а суммарный доход за месяц, $Dm = D \cdot 60 \cdot 24 \cdot 30 = 42768$ усл.уд. На содержание 5 каналов требуется 50 000 усл. ед., значит предприятие не рентабельно.

6.6 Одноканальная СМО с неограниченной очередью

Из СМО с ожиданием самой простой является одноканальная СМО с неограниченной очередью. Такие СМО часто встречаются в практике: врач, обслуживающий пациентов; телефон-автомат с одной будкой; ЭВМ, выполняющая заказы пользователей.

Пусть имеется одноканальная СМО с очередью, на которую не наложено никаких ограничений. На эту СМО поступает поток заявок с интенсивностью λ; поток обслуживания простейший и имеет интенсивность μ, $\left(\overline{t}_{\text{обсл}} = \frac{1}{\mu}\right)$.

Найти предельные вероятности состояний СМО, а также характеристики эффективности ее работы.

Запишем состояния СМО (состояния пронумерованы по числу заявок в системе):

S_0 – канал свободен,

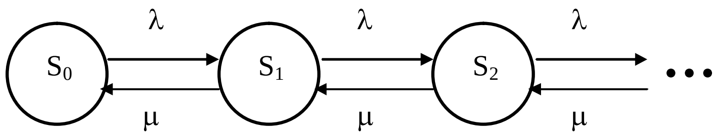
S_1 – канал занят, очереди нет,

S_2 – канал занят, одна заявка в очереди,

. . . .

S_n – канал занят; $n-1$ заявка в очереди и так далее.

Теоретически число состояний не ограничено (бесконечно). Граф состояний показан на рисунке



Существуют ли в этом случае предельные вероятности состояний, ведь число состояний системы бесконечно?

Формула для определения p_0 имеет вид:

$$p_0 = \left(1 + \rho + \rho^2 + \ldots + \rho^n + \ldots\right)^{-1}.$$

В скобке стоит сумма бесконечной геометрической прогрессии. Если знаменатель прогрессии $\rho = \frac{\lambda}{\mu} < 1$, то сумму можно вычислить по формуле

$$1+\rho+\rho^2+\ldots=\frac{1}{1-\rho}.$$

Если $\rho \geq 1$, то ряд расходится, сумма растет неограниченно, СМО не может работать из-за неограниченно возрастающей очереди. Дальнейшие формулы справедливы только при $\rho < 1$.

$$p_0 = 1-\rho; \;\; p_k = \rho^k\left(1-\rho\right), \;\; k = 1,2\ldots \tag{41}$$

Среднее число заявок в очереди одноканальной СМО и время ожидания обслуживания равно

$$\overline{r}_{\text{о}} = \frac{\rho^2}{1-\rho}; \quad \overline{t}_{\text{о}} = \frac{\rho^2}{\lambda\left(1-\rho\right)} \;. \tag{42}$$

Среднее число занятых каналов – $\overline{k} = \rho$. Найдем среднее число заявок в системе:

$$\overline{z}_{\text{сист}} = \frac{\rho^2}{1-\rho} + \rho = \frac{\rho}{1-\rho} \;, \tag{43}$$

и среднее время нахождения заявки в системе:

$$\overline{t}_{\text{сист}} = \overline{t}_{\text{обсл}} + \overline{t}_{\text{о}} = \frac{1}{\mu}\cdot\frac{1}{1-\rho}. \tag{44}$$

Пример 1. На железнодорожную сортировочную горку прибывают составы с интенсивностью $\lambda = 2$ состава в час.

Среднее время, в течение которого горка обрабатывает состав, равна 24 минутам. Составы, прибывшие в момент, когда горка занята, становятся в очередь и ожидают в парке прибытия, где имеются три запасных пути, на каждом из которых может ожидать один состав.

Состав, прибывший в момент, когда три запасных пути заняты, ставится на внешний путь. Станция платит штраф за пребывание состава на внешнем пути (*a* усл. ед. за 1 час).

1. Описать состояния системы и построить граф состояний.

2. Вычислить вероятности состояний СМО для стационарного случая и коэффициенты эффективности работы горки. Сделать выводы о качестве ее работы.

3. Вычислить штраф, который должна выплачивать станция за месяц из-за ожидания обслуживания на внешних путях.

Решение

1. Состояния СМО и ее граф приведены выше. Нас будут интересовать те состояния, при которых составы попадают на внешний путь:

S_5 – один состав на горке, три в очереди на запасных путях станции, один на внешних путях;

S_6 – один на горке, три на запасных путях, два на внешних путях и так далее.

2. Вычислим стационарные вероятности состояний СМО:

$\lambda = 2$ сост./ч; $t_{обсл} = 24 мин = 0,4 ч.$; $\mu = 2,5$ сост./ч;

$\rho = 2 \cdot 0,4 = 0,8 < 1 \Rightarrow p_0 = 1 - \rho = 0,2$

$p_1 = 0,16;\ p_2 = 0,128;\ p_3 = 0,102;\ p_4 = 0,082\ldots$

Теперь найдем вероятность того, что прибывающий состав попадает на внешний путь

$$p(A) = p_4 + p_5 + p_6 + p_7 + \ldots = \rho(1-\rho)\left(1+\rho+\rho^2+\ldots\right) = \frac{\rho^4(1-\rho)}{1-\rho} = \rho^4$$

$p(A) = \rho^4 \approx 0,41$

Таким образом, в 41% случаев состав попадает на внешний путь.

Вычислим коэффициенты эффективности работы горки.

Среднее число составов, ожидающих в очереди (как в парке, так и вне его)

$$\bar{r}_о = \frac{0,64}{1-0,8} = 3,2 \text{ состава, } \bar{t}_о = \frac{\bar{r}_о}{\lambda} = 1,6 \text{ ч.}$$

Среднее число составов в парке расформирования

$$\bar{z}_{сист} = \frac{0,8}{1-0,8} = 4 \text{ состава, } \bar{t}_{сист} = \frac{\bar{z}_{сист}}{\lambda} = 2 \text{ ч.}$$

Таким образом, составу приходится более 1,5 часов стоять в очереди.

3. Вычислим среднее время ожидания на внешних путях. Составим закон распределения случайной величины $r_{вн}$; $r_{вн}$ – число составов, поставленных на внешних путях.

$r_{вн}$	0	1	2	3	4	....
p_i	$p(0)$	p_5	p_6	p_7	p_8	....

$p(0) = p_0 + p_1 + p_2 + p_3 + p_4$

$$M[r_{вн}] = \bar{r}_{вн} = 1\cdot p_5 + 2\cdot p_6 + 3\cdot p_7 + \ldots = \rho^5(1-\rho)\left(1+2\rho+3\rho^2+\ldots\right) =$$

$$= \rho^5(1-\rho)\left(\rho+\rho^2+\rho^3+\ldots\right)' = \rho^5(1-\rho)\cdot\left(\frac{\rho}{1-\rho}\right)' = \rho^5(1-\rho)\cdot\frac{1}{(1-\rho)^2} = \frac{\rho^5}{1-\rho}$$

В нашем примере:

$$\bar{r}_{вн} = \frac{(0,8)^5}{1-0,8} = 0,82;\ t_{вн} = \frac{\bar{r}_{вн}}{\lambda} = 0,41$$

Вычислим штраф за сутки:

Ш = 2 сост/ч · 24 ч · 0,41 · *а* = 19,66 *а* усл. ед.

Индивидуальные задания

Задание 5. *Системы массового обслуживания с отказами*

АТС имеет 4 линии связи. Поток вызовов простейший с интенсивностью λ вызовов в минуту. Время переговоров распределено по показательному закону, среднее время составляет t мин ($\overline{t}_{\text{обсл}}$). Информация об исходных данных приведена в таблице.

1. Описать состояния СМО, построить граф состояний.

2. Найти предельные вероятности состояний системы. Найти показатели эффективности работы АТС, проанализировать эти показатели.

3. Изучить зависимость среднего числа занятых каналов и абсолютной пропускной способности АТС от интенсивности входного потока, зависимости представить в виде таблиц и графиков.

4. Определить, сколько линий должна иметь АТС, чтобы вероятность отказа не превышала 0,01.

5. Содержание каждого канала в месяц обходится в 10 тыс. усл. ед. Каждая обслуженная заявка приносит доход в 1,5 усл. ед. Определить, приносит ли АТС доход от обслуживания всех заявок? Каким должно быть число каналов, чтобы доход был максимальным?

Вар.	1	2	3	4	5	6	7	8	9	10
λ	0.9	0.7	0.8	1.3	0.7	0.8	0.9	1.2	1.3	1.1
t	2.9	2.4	0.9	2.6	2.5	2.4	2.5	2.6	2.8	2.3

Задание 6. *Системы массового обслуживания с очередями*

В приемно-отправочный парк станции поступает простейший поток поездов со средней интенсивностью λ составов в час.

Одна бригада осмотрщиков обрабатывает состав со средней продолжительностью t мин$(\overline{t}_{\text{обсл}})$. Время обработки распределено по показательному закону. Очередь не ограничена.

Исходные данные приведены в таблице

Вар.	1	2	3	4	5	6	7	8	9	10
λ	3.0	4.0	5.0	6.0	5.6	4.8	3.5	4.3	5.1	5.4
t	15	12	10	8	9	10	14	11	10	9

1. Описать состояния системы, построить граф состояний.

2. Найти вероятности состояний для стационарного случая и показатели эффективности работы бригады осмотрщиков. Оценить эффективность работы бригады.

3. Известно, что в летние месяцы интенсивность потока составов возрастает вдвое, время осмотра также возрастает в 1,5 раза за счет увеличения длины состава. Определить необходимое число бригад для нормальной работы станции.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Введение в исследование операций / Х. Таха. Т.1. – М. : Мир, 1985.
2. Линейное программирование: методическое руководство / А. И. Недвецкая, М. А. Толмачева. – Свердловск, 1985.
3. Экономико-математические модели управления : метод. руководство / Г. А.Тимофеева. − Екатеринбург, 2000.
4. Основы математики и ее приложения в экономическом образовании: учеб. для вузов / М. С. Красс, Б. П. Чупрынов. – М. : Дело, 2001.
5. Транспортная задача линейного программирования: метод. руководство / П. П. Скачков, И. Н. Пирогова. − Екатеринбург, 2004.
6. Линейное программирование : метод. указания / П. И. Гниломедов, И. Н. Пирогова, П. П. Скачков. − Екатеринбург, 2007.
7. Математические модели массового обслуживания: метод. руководство / Т. В. Величко, П. П. Скачков, Г. А. Тимофеева. – Екатеринбург, 2004.

Учебное издание

Ирина Николаевна Пирогова

Павел Павлович Скачков

Математические модели

Методические указания
по методике проведения практических занятий
и самостоятельной работы для студентов всех специальностей
заочной формы обучения

Редактор *С. И. Семухина*

Подписано в печать 04.09.2009. Формат 60 × 84/16
Бумага офсетная. Усл. печ. л. 2,8
Тираж 200 экз. Заказ № 230

Издательство УрГУПС
620034, Екатеринбург, ул. Колмогорова, 66