

Министерство образования Республики Беларусь

**Учреждение образования
«Гомельский государственный технический
университет имени П. О. Сухого»**

Кафедра «Электроснабжение»

Т. В. Алферова, О. М. Попова

МАТЕМАТИЧЕСКИЕ ЗАДАЧИ ЭЛЕКТРОЭНЕРГЕТИКИ

КУРС ЛЕКЦИЙ

**по одноименной дисциплине
для студентов специальности 1-43 01 03
«Электроснабжение (по отраслям)»
дневной и заочной форм обучения**

Гомель 2013

УДК 31в631(075.8)
ББК 621.31:51-7я73
А53

*Рекомендовано научно-методическим советом
энергетического факультета ГГТУ им. П. О. Сухого
(протокол № 10 от 26.06.2012 г.)*

Рецензент: канд. техн. наук, доц. каф. «Информационные технологии»
ГГТУ им. П. О. Сухого *В. И. Токочаков*

Алферова, Т. В.
А53 Математические задачи электроэнергетики : курс лекций по одноим. дисциплине для студентов специальности 1-43 01 03 «Электроснабжение (по отраслям)» днев. и заоч. форм обучения / Т. В. Алферова, О. М. Попова. – Гомель : ГГТУ им. П. О. Сухого, 2013. – 107 с. – Систем. требования: PC не ниже Intel Celeron 300 МГц; 32 Mb RAM; свободное место на HDD 16 Mb; Windows 98 и выше; Adobe Acrobat Reader. – Режим доступа: <http://alis.gstu.by/StartEK/>. – Загл. с титул. экрана.

Составлен в соответствии с рабочей программой дисциплины «Математические задачи электроэнергетики» путем подбора и систематизации материала ряда литературных источников.

Для студентов специальности 1-43 01 03 «Электроснабжение (по отраслям)» дневной и заочной форм обучения.

**УДК 31в631(075.8)
ББК 621.31:51-7я73**

© Учреждение образования «Гомельский
государственный технический университет
имени П. О. Сухого», 2013

СОДЕРЖАНИЕ

Введение.....	4
1. Обработка статистической информации и получение статистических характеристик.....	5
1.1 Обработка экспериментальных данных.....	5
1.2 Выравнивание статистических рядов.....	16
1.3 Применение критериев согласия при проверке гипотез.....	23
2. Корреляционный анализ.....	32
2.1 Связи функциональные и статистические.....	32
2.2 Основные положения корреляционного анализа.....	33
2.3 Свойства коэффициента корреляции.....	35
2.4 Поле корреляции. Вычисление оценок параметров двумерной модели.....	37
2.5 Корреляционное отношение.....	41
3. Регрессионный анализ.....	44
3.1 Определение формы связи.....	44
3.2 Основные положения регрессионного анализа.....	46
3.3 Линейная регрессия.....	49
3.4 Нелинейная регрессия.....	52
4. Оптимизационные задачи электроснабжения.....	61
4.1 Задачи электроснабжения, требующие поиска оптимальных решений.....	61
4.2 Понятие об управлении. Принципы исследование операций и основные понятия.....	62
4.3 Модели, применяемые для решения оптимизационных задач.....	64
4.4 Классификация методов оптимизации.....	65
5. Задачи линейного программирования.....	69
6. Задачи нелинейного программирования.....	85
7. Динамическое программирование.....	96
Литература.....	107

ВВЕДЕНИЕ

Изучение дисциплины «Математические задачи электроэнергетики» имеет целью усвоение теоретических знаний и приобретение практических навыков, необходимых для самостоятельного проведения прикладных математических исследований в области получаемой в вузе специальности.

Основные задачи дисциплины при ее ориентации на специальность 1-42 01 03 «Электроснабжение (по отраслям)» могут быть сформулированы, исходя из требований к функционированию систем электроснабжения — обеспечение бесперебойного снабжения потребителей электроэнергией установленного качества.

В данном курсе лекций приведены разделы, касающиеся обработки экспериментальных данных, корреляционного и регрессионного анализа, методов линейного, нелинейного и динамического программирования.

Курс лекций составлен в соответствии с рабочей программой дисциплины «Математические задачи электроэнергетики» путем подбора и систематизации материала ряда литературных источников и предназначен для студентов специальности 1-43 01 03 «Электроснабжение (по отраслям)» дневной и заочной формы обучения.

Для успешного усвоения материала кратко излагаются основные теоретические положения и приводятся конкретные примеры.

1. ОБРАБОТКА СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ И ПОЛУЧЕНИЕ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК

1.1 Обработка экспериментальных данных

1.1.1 Вариационные ряды

Установление закономерностей, которым подчиняются массовые случайные явления, основано на изучении статистических данных – сведений о том, какие значения принял в результате наблюдений интересующий исследователя признак.

Статистические ряды распределения характеризуют структуру изучаемого явления, позволяют судить об однородности (неоднородности) совокупности, границах ее изменения, закономерностях развития (рис. 1.1).



Рис 1.1

Рассмотрим данную классификацию. Статистические ряды подразделяются на ряды динамики и ряды распределения.

Ряды динамики отражают изменения процесса, явления во времени.

В качестве примера рассмотрим статистический ряд годового электропотребления (табл.1.1)

Таблица 1.1

Год	2005	2006	2007	2008	2009	2010	2011
Электропотребление, МВтч	30,5	31,7	32,3	35,7	35,9	36,2	37,3

Вид ряда – ряд динамики. В основе группировки лежит существенный, количественный признак.

Группировка простая, так как выполнена по одному признаку (по электропотреблению).

Ряды распределения показывают состояние исследуемого явления, его состав или структуру (табл. 1.2).

Таблица 1.2

Тарифный разряд	Количество рабочих, m_x	Доля рабочих, ω_x
1	4	0,04
2	6	0,06
3	12	0,12
4	16	0,16
5	44	0,44
6	18	0,18
Σ	100	1,00

Вид ряда - структурный ряд распределения; в основе группировки лежит существенный, количественный, непрерывный признак.

Группировка простая, поскольку выполнена по одному признаку (тарифный разряд).

Атрибутивные - ряды, построенные на основе качественных признаков.

Вариационные - ряды, построенные на основе количественных признаков.

В зависимости от прерывности вариации признака различают *ранжированные, дискретные (прерывные) и интервальные (непрерывные)* вариационные ряды.

Ранжированным считается ряд, в котором единицы совокупности расположены в порядке возрастания (убывания) изучаемого признака.

Вариационный ряд имеет два элемента:

-*варианту* - отдельное значение группировочного признака в вариационном ряду распределения (x_i);

-частоту - число повторений отдельных вариантов (f_i). Частоты, выраженные в долях единицы или процентах к итогу, называют *частотами*.

Если в основе группировки лежит количественный дискретный признак, т.е. признак, который принимает определенные (например, целые) значения, то число выделяемых групп будет равно числу вариант значений признака. Пример дискретного вариационного ряда представлен в табл. 1.3

Таблица 1.3

Успеваемость в баллах (x_i)	6	7	8	9	10
Число учеников (f_i)	5	6	7	10	2

Интервальные вариационные ряды значение варианты имеют в виде интервалов. Интервал – это разность между максимальным и минимальным значениями признака в каждой группе (нижней и верхней границами интервала). Интервалы могут быть открытые и закрытые, равные и неравные (табл. 1.4)

Таблица 1.4

Промышленные предприятия с численностью промышленно-производственного персонала, чел.	Число предприятий, %
До 100	42,5
101-200	25,5
201-500	19,5
501-1000	6,5
1001 и более	6,0

Интервалы, в которых указана лишь одна граница (верхняя или нижняя), называются открытыми. В приведенном примере это интервалы групп 1 и 5, что свидетельствует о широкой колеблемости признака.

Остальные интервалы имеют верхнюю и нижнюю границы и являются закрытыми. В группах 2, 3 и 4 использованы неравные интервалы, соответственно 100, 200 и 500 человек, которые применяются в случаях группировки больших неоднородных статистических совокупностей со значительной и неравномерной вариацией признаков.

Внутри типичных групп для характеристики количественных различий единиц группы применяют равные интервалы.

1.1.2 Построение интервальных рядов распределения

При построении интервальных рядов распределения возникает вопрос о численности групп. Число образуемых групп (n) может быть представлено программой исследования или получено на основе формулы Стерджесса:

$$n = 1 + 3.322 \lg \cdot N, \quad (1.1)$$

где N – число единиц совокупности

На основе этой формулы выявлено соотношение между численностью совокупности и оптимальным числом групп (табл. 1.5)

Таблица 1.5

N	15-24	24-44	44-89	89-179	179-359	359-719
n	5	6	7	8	9	10

Величину интервала (i) можно определить по формуле

$$i = \frac{x_{\max} - x_{\min}}{n}, \quad (1.2)$$

где x_{\max} и x_{\min} – максимальное и минимальное значение группировочного показателя в совокупности. Нижней границей первой группы принимается минимальное значение признака, а верхней – $x_{\min} + i$.

Верхняя граница первой группы является нижней границей второй группы, верхняя граница будет больше нижней на величину интервала. По этому принципу строятся границы всех групп. Верхняя граница последней группы равна x_{\max} .

Вторичная группировка, предусматривающая образование новых групп, производится двумя способами: изменением величины интервалов и долевой перегруппировкой.

1.1.3 Графическое представление статистических данных

График представляет собой наглядное средство выражения и анализа явлений посредством линий, точек, геометрических фигур, карт-схем.

По способу построения графики делят на диаграммы и статистические карты. Диаграммы бывают плоскостные и объемные (изображаются в виде секторов, полос (лент) или столбиков), линейные и фигурные.

Графики строятся для характеристики структуры; сравнения по промышленным предприятиям; оценки динамики и выполнения плана; характеристики вариации; оценки взаимосвязей.

В практике статистического анализа чаще применяются диаграммы – графики, показывающие соотношение сравниваемых величин.

Линейные диаграммы применяются для характеристики вариации в рядах распределения, а также взаимосвязи между явлениями, оценки изменений явлений во времени (динамики). Они строятся в линейной системе координат. Даты или периоды времени откладывают по оси абсцисс, уровни ряда динамики или темпы их изменения - по оси ординат. Полученные точки соединяют ломаной линией.

График дискретного ряда может быть представлен полигоном распределения (рис. 1.2).

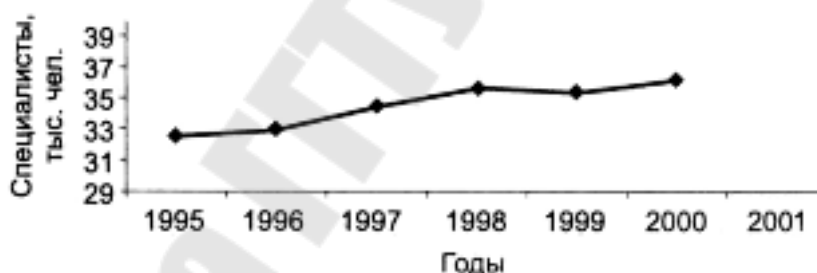


Рис. 1.2

Графическое изображение интервального вариационного ряда имеет вид *гистограммы* (столбиковой диаграммы).

Гистограмма распределения предприятий по численности промышленно-производственного персонала на рис. 1.3 построена по данным табл. 1.6.

Таблица 1.6

Показатель	Численность промышленно-производственного персонала, чел				
	100-200	200-300	300-400	400-500	500-600
Число предприятий	10	2	16	4	20
Накопленная численность предприятий	10	12	28	32	52

Для получения гистограммы по оси абсцисс откладывают интервалы ряда, высота которых равна частотам, отложенным по оси ординат.

Если вариационный ряд имеет неравные интервалы, по оси ординат наносят показатели плотности распределения (количество единиц каждой группы, приходящееся на единицу величины интервала).

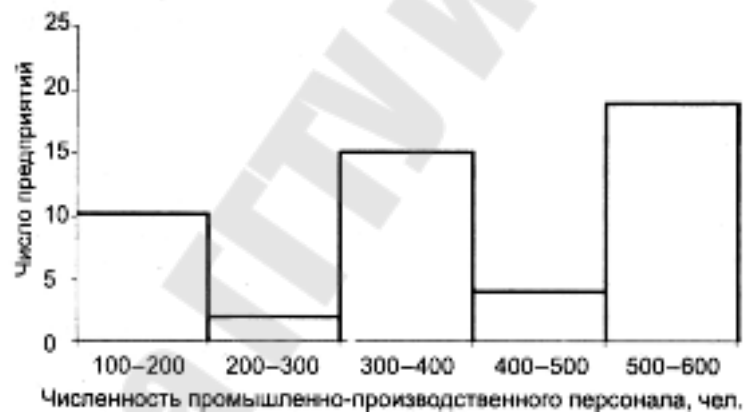


Рис. 1.3

При использовании на оси ординат данных накопленного (аккумулятивного) ряда (табл. 1.6) график принимает вид *кумуляты* (рис. 1.4).

Нижней границе первого интервала при построении кумуляты соответствует частота, равная 0, а верхней границе - частота данного интервала.

Верхняя граница второго интервала находится на уровне суммы частот двух первых интервалов и т.п.

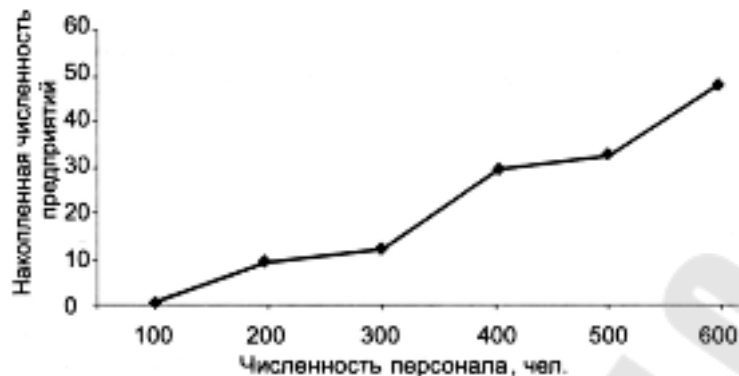


Рис. 1.4

Использование кумулянт удобно для сравнения вариационных рядов, анализа концентрации производства.

Полосовые (ленточные) диаграммы, в отличие от гистограмм, имеют горизонтально расположенные полосы (ленты).

Если на столбиковых диаграммах хотят показать структуру явлений, столбики (полосы) делят на части, пропорциональные изучаемым величинам.

Структуру и структурные сдвиги процессов и явлений хорошо отображает *секторная диаграмма*. Она имеет вид круга, разделенного радиусами на секторы. Секторы пропорциональны удельному весу частей изучаемой совокупности.

Например, электробаланс промышленного предприятия по целевому назначению за 2011 год выглядит следующим образом: электрические аппараты – 28,2%; электродвигатели – 70,5%; потери в сетях – 1,3% (рис. 1.5).

График должен иметь:

- графический образ (точки, столбики, секторы и т.д.);
- поле (место, где располагается графический образ);
- пространственные ориентиры (оси абсцисс, ординат, круговые поля и т.д.);
- масштаб (условную меру перевода числовой величины в графическую);
- название графика, подписи на осях координат с названием показателей и единиц их измерения, пояснения условных обозначений к графику.

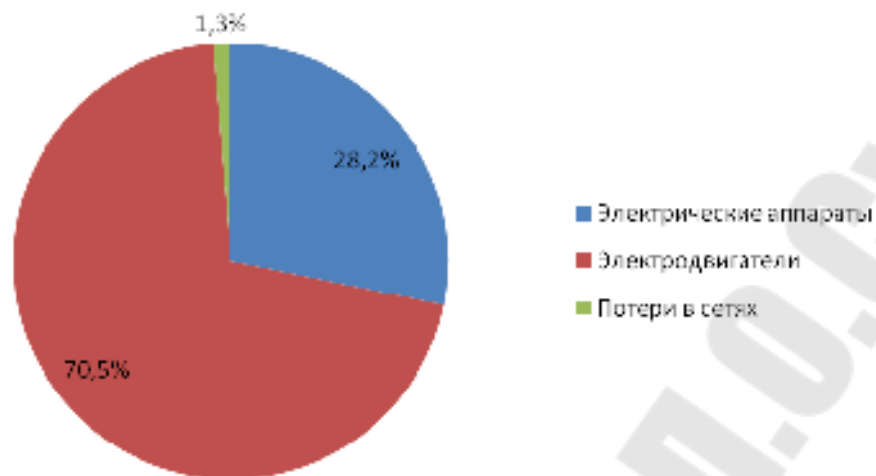


Рис. 1.5

По данным разработочных таблиц можно построить графики, показывающие взаимосвязь между двумя признаками: факторным и результативным.

Так, на основании данных табл. 1.7 строится график зависимости между производством продукции (факторный группировочный признак) и среднесписочной численностью рабочих (результативный признак).

Таблица 1.7

Группы предприятий по выпуску продукции, млн.руб.	Число предприятий	Выпуск продукции в среднем на одно предприятие, млн.руб.	Среднесписочная численность рабочих в среднем на одно предприятие, чел.
186,6 – 331,8	11	268,55	233
331,8 – 477,0	6	384,98	271
477,0 – 622,0	3	525,40	363
622,0 – 767,4	2	711,55	406
767,4 – 912,6	2	864,65	323

По оси абсцисс откладываем среднее значение факторного признака, оси ординат - результативного. По полученной ломаной кривой делаем выводы о характере зависимости между двумя показателями.

Далее проводится аналогичная работа, меняются местами признаки, т.е. группируются предприятия по среднесписочной численности рабочих.

1.1.4 Экспериментальное определение точечных и интервальных характеристик случайных величин

Полученные данные располагают в определенной последовательности и разбивают на несколько групп. Для сравнения групп интервалы делают одинаковыми. Дискретные и непрерывные величины разделяют на группы по разному. Если случайная переменная величина дискретна и имеет небольшой диапазон колебаний между крайними значениями, то интервал группирования принимается равным единице исследуемой величины.

Например, (в результате опыта получены следующие значения измеряемой величины X (приведены в порядке проведения испытаний)

17; 18; 18; 21; 18; 17; 19; 19; 17; 18;
 17; 21; 20; 19; 19; 16; 19; 19; 19; 20;
 18; 18; 18; 19; 16; 17; 18; 20; 16; 19.

Разность между максимальным и минимальным значениями измеряемой величины невелика.

$X_{\max} - X_{\min} = 5$. Поэтому интервал группирования принимаем равным единице, а группы составляем из отдельных значений измеряемой величины. После этого приступаем к распределению отдельных величин по группам.

Таблица 1.8

Значение измеряемой величины	Частота наблюдаемого значения m_i	Относительная частота m_i / n
16	3	0,10
17	5	0,17
18	8	0,27
19	9	0,30
20	3	0,10
21	2	0,06
	30	1,0

Если случайная переменная прерывна и имеет большой диапазон колебаний между отдельными значениями, то несколько таких значений следует объединить в один интервал.

Если изучаемая величина является непрерывной, то полученные значения объединяются в интервалы. Выбор интервала зависит от объема измерений, диапазона колебаний между крайними значениями, и цели исследования. Из статистических соображений число интервалов рекомендуется вычислять из соотношения

$$K = \sqrt{n};$$

где K – число интервалов ≥ 5 ;

n – объем измерений.

Например, получены следующие измерения X

4,26; 4,17; 4,25; 4,25; 4,22; 4,25; 4,25; 4,40; 4,05; 4,15;
4,40; 4,17; 4,27; 4,10; 4,25; 4,30; 4,25; 4,20; 4,20; 3,90.

$X_{\max} = 4,40$ $X_{\min} = 3,90$

Разделив разность между этими значениями на количество имеющихся групп $K = \sqrt{n} = \sqrt{20} \approx 5$, получим приблизительное значение ширины интервала группирования

$$\Delta X = \frac{4.40 - 3.90}{5} = 0.1$$

Ширина интервала не должна быть меньше цены деления измерительного прибора.

Возможны затруднения при включении значений, совпадающих с границами группы. На практике обычно спорную величину, совпадающую с границей группы, записывают в высшую группу.

Результаты группирования представим в таблице 1.9

Таблица 1.9

Границы группы	Середина интервала	Частота наблюдаемого значения	Относительная частота m_i / n
3,90 – 4,00	3,95	1	0,05
4,00 – 4,10	4,05	2	0,10
4,10 – 4,20	4,15	5	0,25
4,20 – 4,30	4,25	10	0,50
4,30 – 4,40	4,35	2	0,1
		20	1,0

Законы распределения случайной величины представим графически (рис. 1.6)

По таблицам 1.8 и 1.9 можно найти соответствующие статистические характеристики среднее арифметическое \bar{x} , дисперсию $S(x)$ и среднеквадратическое отклонение $\sigma_{ст}(x)$

$$\bar{x}^* = \frac{\sum X_i m_i}{\sum m_i}; S^*(x) = \frac{1}{n} \sum (X_i - \bar{X})^2;$$

$$\sigma_{ст}^* = \sqrt{\frac{1}{n} \sum_1^n (X_i - \bar{X})^2}.$$

Для практических задач для оценки случайной величины прибегают к ограниченному числу опытов с целью получения соответствующих характеристик происходящих случайных процессов. Таким образом, мы вынуждены прибегать к выборке.

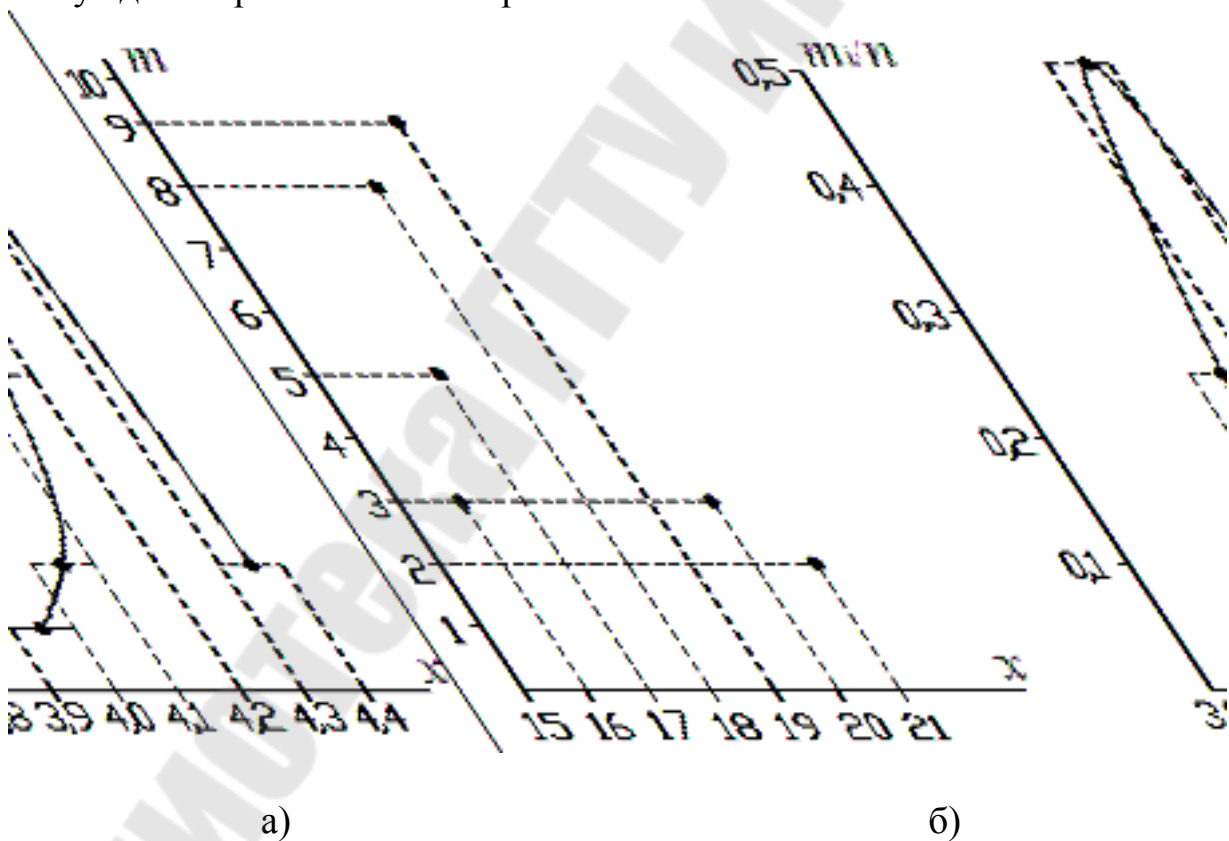


рис 1.6

Замена большого числа опытов, называемая генеральной совокупностью, ограниченным количеством испытаний (выборкой) оправдано

лишь тогда, когда числовые характеристики, полученные по выборке, будут достаточно близкими по своим значениям и характеристикам генеральной совокупности.

Такое условие неизбежно ведет к необходимости теоретического обоснования объема и способов выборки.

1.2 Выравнивание статистических рядов

Во всяком статистическом распределении присутствуют элементы случайности, связанные с количеством наблюдений, качеством опытов, полученными результатами. Только при очень большом числе наблюдений эти элементы случайности сглаживаются, и случайное явление обнаруживает в полной мере присущую ему закономерность. На практике число наблюдений ограничено и мы вынуждены считаться с тем, что любому статистическому распределению свойственны в большей или меньшей мере черты случайности. Поэтому при обработке статистического материала приходится решать вопрос о том, как подобрать для данного статистического ряда теоретическую кривую распределения, выражающую лишь существенные черты статистического материала, а не случайности, связанные с недостаточным объемом экспериментальных данных. Такая задача называется задачей **выравнивания (сглаживания) статистических рядов**.

Задача выравнивания заключается в подборе теоретической плавной кривой распределения, с той или иной точки зрения наилучшим образом описывающую данное статистическое распределение (рис. 1.7).

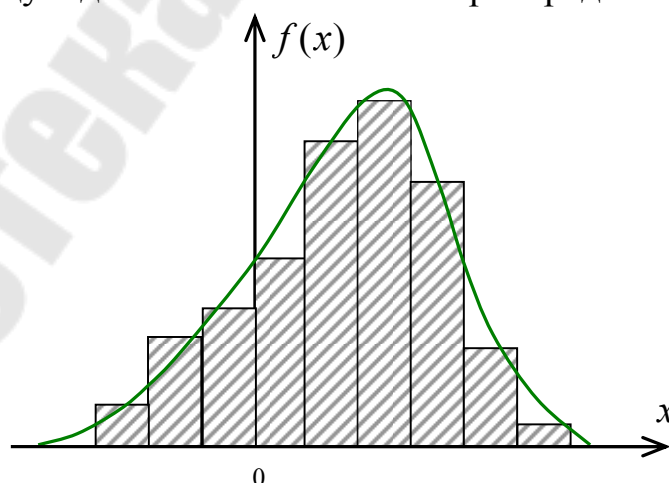


Рис. 1.7

Задача о наилучшем выравнивании статистических рядов есть задача в значительной мере неопределенная и решение ее зависит от того, что условиться считать «наилучшим». При сглаживании эмпирических зависимостей часто исходят из так называемого принципа или метода наименьших квадратов, считая, что наилучшим приближением к эмпирической зависимости в данном классе функций является такое, при котором сумма квадратов отклонений обращается в минимум. При этом вопрос о том, в каком именно классе функций следует искать наилучшее приближение, решается из соображений, связанных с физикой решаемой задачи, с учетом характера полученной эмпирической кривой и степени точности произведенных наблюдений. Часто принципиальный характер функции, выражающей исследуемую зависимость, известен заранее из теоретических соображений, из опыта же требуется получить лишь некоторые численные параметры, входящие в выражение функции; именно эти параметры подбираются с помощью метода наименьших квадратов.

В задачах выравнивания статистических рядов, как правило, принципиальный вид теоретической кривой выбирается заранее из соображений, связанных с существом задачи, а в некоторых случаях просто с внешним видом статистического распределения. Аналитическое выражение выбранной кривой распределения зависит от некоторых параметров; задача выравнивания статистического ряда переходит в задачу рационального выбора тех значений параметров, при которых соответствие между статистическим и теоретическим распределениями оказываются наилучшим.

Предположим, что исследуемая величина X есть ошибка измерения, возникающая в результате суммирования множества независимых элементарных ошибок; тогда из теоретических соображений можно считать, что величина X подчиняется нормальному закону:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2 \cdot \sigma^2}},$$

и задача выравнивания переходит в задачу о рациональном выборе параметров m и σ в данном выражении.

Когда заранее известно, что величина X распределяется статистически приблизительно равномерно на некотором интервале, тогда мож-

но поставить задачу о рациональном выборе параметров того закона равномерной плотности

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{при } \alpha < x < \beta, \\ 0 & \text{при } x < \alpha; x > \beta, \end{cases}$$

которым можно наилучшим образом заменить (выровнять) заданное статистическое распределение.

Следует иметь в виду, что любая аналитическая функция $f(x)$, с помощью которой выравнивается статистическое распределение, должна обладать основными свойствами плотности распределения:

$$f(x) \geq 0;$$

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

Предположим, что, исходя из тех или иных соображений, нами выбрана функция $f(x)$, удовлетворяющая последним условиям, с помощью которой мы хотим выровнять данное статистическое распределение; в выражение этой функции входят несколько параметров a, b, \dots ; требуется подобрать эти параметры так, чтобы функция $f(x)$ наилучшим образом описывала данный статистический материал. Один из методов, применяемых для решения этой задачи, – это так называемый **метод моментов**.

Согласно методу моментов, параметры a, b, \dots выбираются с таким расчетом, чтобы несколько важнейших числовых характеристик (моментов) теоретического распределения были равны соответствующим статистическим характеристикам. Например, если теоретическая кривая $f(x)$ зависит только от двух параметров a и b , эти параметры выбираются так, чтобы математическое ожидание m_x и дисперсия D_x теоретического распределения совпадали с соответствующими статистическими характеристиками m_x^* и D_x^* . Если кривая $f(x)$ зависит от трех параметров, можно подобрать их так, чтобы совпали первые три момента, и т.д. При выравнивании статистических рядов может оказаться полезной специально разработанная система *кривых Пирсона*, каждая из которых зависит в общем случае от четырех параметров. При вырав-

нивании эти параметры выбираются так, чтобы сохранить первые четыре момента статистического распределения (математическое ожидание, дисперсию, третий и четвертый моменты).

Пример 1.1

В таблице 1.10 приведено статистическое распределение величины X . Требуется выровнять это распределение с помощью нормального закона.

Таблица 1.10

Границы x_i	-4; -3	-3; -2	-2; -1	-1; 0	0; 1	1; 2	2; 3	3; 4
x'_i	- 3,5	- 2,5	- 1,5	- 0,5	0,5	1,5	2,5	3,5
m_i	6	25	72	133	120	88	46	10
p_i^*	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

Здесь x_i – интервалы измерений; x'_i – середины интервалов; m_i – число наблюдений в данном интервале; $\sum_{i=1}^k m_i = n$, $p_i^* = \frac{m_i}{n}$ – соответствующие частоты.

Решение. Нормальный закон распределения вероятностей зависит от двух параметров m и σ , которые входят в формулу плотности распределения вероятностей

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2 \cdot \sigma^2}}.$$

Подберём эти параметры так, чтобы сохранить первые два момента статистического распределения – математическое ожидание и дисперсию.

Вычислим приближенно статистическое среднее параметра X по формуле

$$m_x^* \approx M^*[X] = \sum_{i=1}^k x'_i \cdot p_i^* = \sum_{i=1}^k x'_i \cdot \frac{m_i}{n},$$

где x'_i – середина каждого интервала.

$$m_x^* = (-3,5) \cdot 0,012 + (-2,5) \cdot 0,05 + (-1,5) \cdot 0,144 + (-0,5) \cdot 0,266 + 0,5 \cdot 0,240 + 1,5 \cdot 0,176 + 2,5 \cdot 0,092 + 3,5 \cdot 0,020 = 0,168.$$

Для определения статистической дисперсии D_x^* вычислим второй начальный момент α_s^* по формуле:

$$\alpha_s^* = \sum_{i=1}^k (x'_i)^s \cdot p_i^*.$$

При $s = 2$ и $k = 8$ последняя формула примет вид:

$$\alpha_2^* = \sum_{i=1}^{k=8} (x'_i)^2 \cdot p_i^*.$$

Выполним численные подстановки:

$$\alpha_2^* = (-3,5)^2 \cdot 0,012 + (-2,5)^2 \cdot 0,05 + (-1,5)^2 \cdot 0,144 + (-0,5)^2 \cdot 0,266 + 0,5^2 \cdot 0,240 + 1,5^2 \cdot 0,176 + 2,5^2 \cdot 0,092 + 3,5^2 \cdot 0,020 = 2,126.$$

$$\text{Тогда } D_x^* = \alpha_2^* - (m_x^*)^2 = 2,126 - 0,168^2 = 2,098.$$

Определение D_x^* через второй начальный момент можно рекомендовать только в случае, когда математическое ожидание m_x^* исследуемой случайной величины X сравнительно невелико.

Статистическая дисперсия D_x^* может быть определена и по выра-

жению: $D_x^* = \sum_{i=1}^k (x'_i - m_x^*)^2 \cdot p_x^*$. Выполним численные подстановки:

$$D_x^* = (-3,5 - 0,168)^2 \cdot 0,012 + (-2,5 - 0,168)^2 \cdot 0,05 + (-1,5 - 0,168)^2 \cdot 0,144 + (-0,5 - 0,168)^2 \cdot 0,266 + (0,5 - 0,168)^2 \cdot 0,240 + (1,5 - 0,168)^2 \cdot 0,176 + (2,5 - 0,168)^2 \cdot 0,092 + (3,5 - 0,168)^2 \cdot 0,020 = 2,098.$$

Выбираем параметры m и σ^2 нормального закона распределения так, чтобы выполнялись условия: $m = m_x^*$ и $\sigma^2 = D_x^*$, то есть примем: $m = 0,168$; а $\sigma = \sqrt{2,098} = 1,448$.

Напишем выражение нормального закона:

$$f(x) = \frac{1}{1,448 \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-0,168)^2}{2 \cdot 1,448^2}}.$$

Пользуясь справочными таблицами, вычислим значения $f(x)$ на границах разрядов

Таблица 1.11

x	-4	-3	-2	-1	0	1	2	3	4
$f(x)$	0,004	0,025	0,090	0,199	0,274	0,234	0,124	0,0041	0,008

Построим на одном графике гистограмму и выравнивающую её кривую распределения.

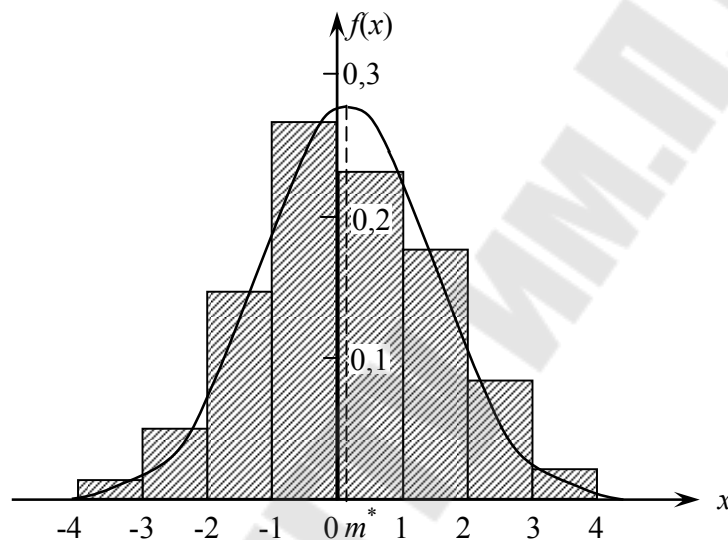


Рис. 1.8

Из графика видно, что теоретическая кривая распределения $f(x)$ сохраняет существенные особенности статистического распределения.

Пример 1.2

Пусть произведено 400 измерений параметра X . Результаты измерений представлены в виде статистического ряда:

Таблица 1.12

Границы	20; 30	30; 40	40; 50	50; 60	60; 70	70; 80	80; 90	90; 100
x_i								
m_j	21	72	66	38	51	56	64	32
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

Выровнять статистический ряд с помощью закона равномерной плотности.

Решение. Закон равномерной плотности выражается формулой

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{при } \alpha < x < \beta, \\ 0 & \text{при } x < \alpha; x > \beta, \end{cases}$$

и зависит от двух параметров α и β . Эти параметры следует выбрать так, чтобы сохранить первые два момента статистического распределения – математическое ожидание m_x^* и дисперсию D_x^* . Выражения математического ожидания и дисперсии для закона равномерной плотности:

$$m_x = \frac{\alpha + \beta}{2}; \quad D_x = \frac{(\beta - \alpha)^2}{12}.$$

Для того чтобы упростить вычисления, связанные с определением статистических моментов, перенесем начало отсчета в точку $x_0 = 60$ и примем за x'_i каждого разряда его середину. Ряд распределения примет вид:

Таблица 1.13

x'_i	- 35	- 25	- 15	- 5	5	15	25	35
p_i^*	0,052	0,180	0,165	0,095	0,128	0,140	0,160	0,080

Приближенное значение статистического среднего X' равно

$$m_{x'}^* = \sum_{i=1}^{k=8} x'_i \cdot p_i^* = (-35) \cdot 0,052 + (-25) \cdot 0,180 + \dots + 35 \cdot 0,080 = 0,26.$$

Второй статистический момент величины X' равен:

$$\alpha_2^* = \sum_{i=1}^{k=8} (x'_i)^2 \cdot p_i^* = (-35)^2 \cdot 0,052 + (-25)^2 \cdot 0,180 + \dots + (35)^2 \cdot 0,080 = 447,8,$$

откуда статистическая дисперсия: $D_x^* = \alpha_2^* - (m_x^*)^2 = 447,7$.

Переходя к прежнему началу отсчета, получим новое статистическое среднее: $m_x^* = m_{x'}^* + 60 = 60,26$ и ту же статистическую дисперсию:

$D_x^* = D_x^* = 447,7$. Параметры закона равномерной плотности определяются уравнениями: $\frac{\alpha + \beta}{2} = 60,26$; $\frac{(\beta - \alpha)^2}{12} = 447,7$.

Решая эти уравнения относительно α и β , имеем: $\alpha \approx 23,6$; $\beta \approx 96,9$; откуда $f(x) = \frac{1}{\beta - \alpha} = \frac{1}{73,3} \approx 0,0136$. На рис.1.9 показаны гистограмма и выравнивающий ее закон равномерной плотности $f(x)$.

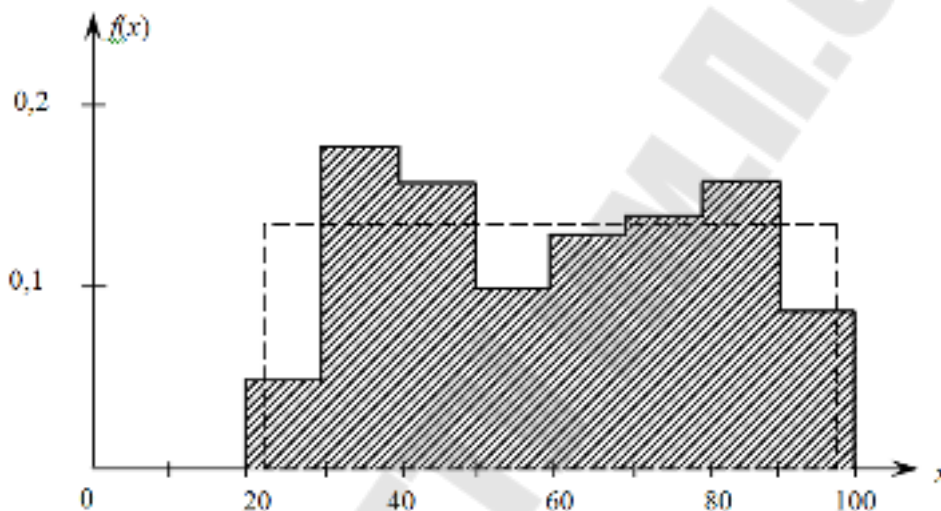


Рис.1.9

1.3 Применение критериев согласия при проверке гипотез

Рассмотрим вопросы, связанные с проверкой правдоподобия гипотез, о согласованности теоретического и статистического распределения.

Пусть данное статистическое распределение выравнено с помощью некоторой теоретической кривой $f(x)$ (рис.1.10). Как бы хорошо ни была подобрана теоретическая кривая, между нею и статистическим распределением неизбежны некоторые расхождения. Возникает вопрос: эти расхождения объясняются только случайными обстоятельствами, связанными с ограниченным числом наблюдений, или они являются существенными и связаны с тем, что подобранная кривая плохо выравнивает данное статистическое распределение. Для ответа на такой вопрос служат «критерии согласия».

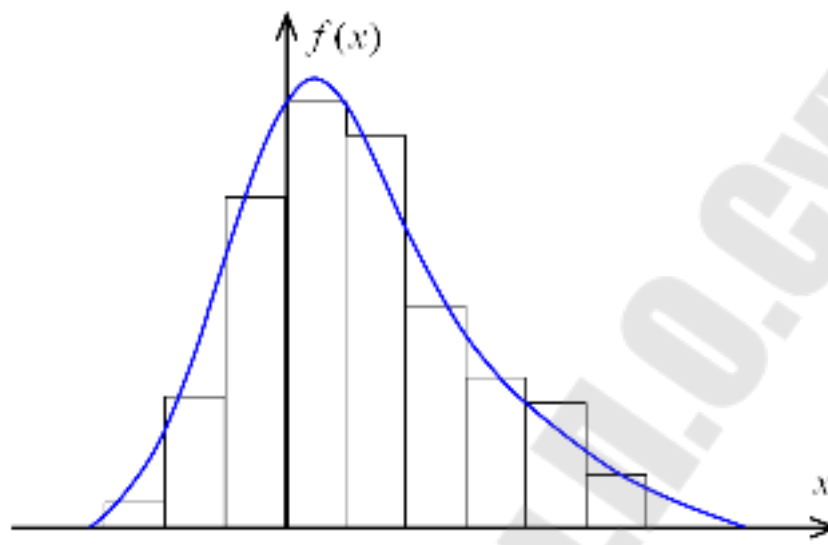


Рис. 1.10

Идея применения критериев согласия заключается в следующем. На основании данного статистического материала необходимо проверить гипотезу H , состоящую в том, что случайная величина X подчиняется некоторому закону распределения. Закон может быть задан в той или иной форме: например, в виде функции распределения $F(x)$ или в виде плотности распределения $f(x)$, или же в виде совокупности вероятностей p_i , где p_i – вероятность того, что величина X попадает в пределы i – го разряда.

Поскольку функция распределения $F(x)$ является наиболее общей и определяет собой любую другую, будем формулировать гипотезу H , как состоящую в том, что величина X имеет функцию распределения $F(x)$.

Чтобы принять или опровергнуть гипотезу H , рассмотрим некоторую величину U , характеризующую степень расхождения теоретического и статистического распределений. Величина U может быть выбрана различными способами; например, в качестве U можно взять сумму квадратов отклонений теоретических вероятностей p_i от соответствующих частот p_i^* или же максимальное отклонение статистической функции распределения $F^*(x)$ от теоретической $F(x)$ и т.д. Если величина U выбрана тем или иным способом, значит это некоторая **случайная величина**. Закон распределения этой случайной величины зависит от за-

кона распределения случайной величины X , над которой проводились опыты, и от числа опытов n . Если гипотеза H верна, то закон распределения величины U определяется законом распределения величины X (функцией $F(x)$) и числом n .

Пусть закон распределения нам известен. В результате данной серии опытов обнаружено, что выбранная мера расхождения U приняла некоторое значение u . Спрашивается, можно ли объяснить это случайными причинами или это расхождение слишком велико и указывает на наличие существенной разницы между теоретическим и статистическим распределениями и, следовательно, на непригодность гипотезы H ? Для ответа на этот вопрос предположим, что гипотеза H верна, и вычислим в этом предположении вероятность того, что за счет случайных причин, связанных с недостаточным объемом опытных данных, мера расхождения U окажется не меньше, чем наблюдаемое нами в опыте значение u , т.е. вычислим вероятность события:

$$U \geq u.$$

Если эта вероятность весьма мала, то гипотезу H следует отвергнуть как мало правдоподобную; если же эта вероятность значительна, следует признать, что экспериментальные данные не противоречат гипотезе H .

Каким же способом следует выбирать меру расхождения U ? Оказывается, что при некоторых способах ее выбора закон распределения величины U обладает простыми свойствами и при достаточно большом n практически не зависит от функции $F(x)$. Такими мерами расхождения и пользуются в статистике в качестве критериев согласия.

Один из наиболее применяемых критериев согласия – так называемый «критерий χ^2 » Пирсона.

Пусть произведено n независимых опытов, в каждом из которых случайная величина X приняла определенное значение. Результаты опытов сведены в k разрядов и представлены статистическим рядом:

Таблица 1.14

Границы X_i	$x_1; x_2$	$x_2; x_3$	$x_3; x_4$	$x_k; x_{k+1}$
p_i^*	p_1^*	p_2^*	p_3^*	p_k^*

Требуется проверить, согласуются ли экспериментальные данные с гипотезой о том, что случайная величина X имеет данный закон распределения (заданный функцией распределения $F(x)$ или плотностью $f(x)$). Назовем этот закон распределения «теоретическим».

Зная теоретический закон распределения, можно определить теоретические вероятности попадания случайной величины в каждый разряд:

$$p_1; p_2; p_3; \dots; p_k.$$

Проверяя согласованность теоретического и статистического распределений, будем исходить из расхождений между теоретическими вероятностями p_i и наблюдаемыми частотами p_i^* . Выбираем в качестве меры расхождения между теоретическим и статистическим распределениями сумму квадратов отклонений $(p_i^* - p_i)$, взятых с некоторыми коэффициентами c_i :

$$U = \sum_{i=1}^k c_i \cdot (p_i^* - p_i)^2.$$

К.Пирсон показал, что если принять $c_i = \frac{n}{p_i}$, то при больших n закон распределения величины U практически не зависит от функции распределения $F(x)$ и от числа опытов n , а зависит только от числа разрядов k , а именно, этот закон при увеличении n приближается к так называемому «распределению χ^2 ».

При таком выборе коэффициентов c_i мера расхождения обозначается χ^2 :

$$\chi^2 = n \sum_{i=1}^k \frac{(p_i^* - p_i)^2}{p_i}.$$

Для удобства вычислений можно ввести n под знак суммы и, учитывая, что $p_i^* = \frac{m_i}{n}$, где m_i – число значений в i -м разряде, привести формулу к виду:

$$U = \chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}. \quad (1.3)$$

Распределение χ^2 зависит от параметра r , называемого числом «степеней свободы» распределения. Число «степеней свободы» r равно

числу разрядов k минус число независимых условий («связей»), наложенных на частоты p_i^* . Примерами таких условий могут быть:

$$\sum_{i=1}^k p_i^* = 1, \quad (1) \quad \sum_{i=1}^k x_i' p_i^* = m_x, \quad (2) \quad \sum_{i=1}^k (x_i' - m_x^*)^2 p_i^* = D_x,$$

если требуется, чтобы сумма частот была равна единице (требование накладывается во всех случаях) – (1); если подбираем условие совпадения теоретического и статистического средних значений – (2) и, если требуется совпадения теоретической и статистической дисперсий (3) и т.д.

Для распределения χ^2 составлены специальные таблицы, которые дают возможность оценить степень согласованности теоретического и статистического распределений. Если величина X действительно распределена по закону $F(x)$, то вероятность p , определенная по таблице, есть вероятность того, что за счет чисто случайных причин мера расхождения теоретического и статистического распределений (4.1) будет не меньше, чем фактически наблюдаемое в данных опытах значение χ^2 . Если эта вероятность p весьма мала, то результаты опыта следует считать *противоречащим* гипотезе H о том, что закон распределения величины X есть $F(x)$. Если вероятность p сравнительно велика, можно признать расхождения между теоретическим и статистическим распределениями несущественными и отнести их за счет случайных причин. Гипотезу H о том, что величина X распределена по закону $F(x)$, можно считать правдоподобной.

Последовательность применения критерия χ^2 к оценке согласованности теоретического и статистического распределений сводится к следующему:

- 1) Определяется мера расхождения χ^2 по формуле (1.3).
- 2) Определяется число степеней свободы r как число разрядов k минус число наложенных связей s : $r = k - s$.
- 3) По r и χ^2 с помощью справочных таблиц определяется вероятность того, что величина, имеющая распределение χ^2 с r степенями свободы, превзойдет данное значение χ^2 . Если эта вероятность весьма мала (0,1 и менее), гипотеза отбрасывается как неправдоподобная. Если эта вероятность относительно велика, гипотезу можно признать не противоречащей опытным данным.

Следует заметить, что при пользовании критерием χ^2 достаточно большим должно быть общее количество опытов n и числа наблюдений m_i в отдельных разрядах.

Пример 1.3

Проверить согласованность теоретического и статистического распределения для выровненного статистического ряда по нормальному закону (см. пример 1.1).

Решение. Пользуясь теоретическим нормальным законом распределения с параметрами $m = 0,168$, $\sigma = 1,448$, находим вероятности попадания в разряды по формуле

$$p_i = \Phi^* \left(\frac{x_{i+1} - m}{\sigma} \right) - \Phi^* \left(\frac{x_i - m}{\sigma} \right),$$

где x_i , x_{i+1} – границы i -го разряда.

$$p_{-4;-3} = \Phi^* \left(\frac{-3 - 0,168}{1,448} \right) - \Phi^* \left(\frac{-4 - 0,168}{1,448} \right) = 0,0124 \text{ и т.д.}$$

Составляем сравнительную таблицу чисел попаданий в разряды m_i и соответствующих значений np_i ($n = 500$).

Таблица 1.15

X_i	-4;-3	-3;-2	-2;-1	-1; 0	0;1	1;2	2;3	3;4
m_i	6	25	72	133	120	88	46	10
np_i	6,2	26,2	71,2	122,2	131,8	90,5	38,2	10,5

Определим значение меры расхождения по выражению

$$\chi^2 = \sum_{i=1}^8 \frac{(m_i - np_i)^2}{np_i}$$

$$\begin{aligned} \chi^2 &= \frac{(6 - 6,2)^2}{6,2} + \frac{(25 - 26,2)^2}{26,2} + \frac{(72 - 71,2)^2}{71,2} + \frac{(133 - 122,2)^2}{122,2} + \\ &+ \frac{(120 - 131,8)^2}{131,8} + \frac{(88 - 90,5)^2}{90,5} + \frac{(46 - 38,2)^2}{38,2} + \frac{(10 - 10,5)^2}{10,5} = 3,94. \end{aligned}$$

Определяем число степеней свободы как число разрядов минус число наложенных связей s (в данном случае $s = 3$):

$$r = 8 - 3 = 5.$$

По таблицам χ^2 (справочных приложений) находим для $r = 5$:

$$\text{при } \chi^2 = 3,00 \quad p = 0,70;$$

$$\text{при } \chi^2 = 4,35 \quad p = 0,50.$$

Следовательно, искомая вероятность p при $\chi^2 = 3,94$ приближенно равна 0,56. Эта вероятность малой не является; поэтому гипотезу о том, что величина X распределена по нормальному закону, можно считать правдоподобной.

Пример 1.4

Проверить согласованность теоретического и статистического распределения для выровненного статистического ряда по закону равномерной плотности (см. пример 1.2).

Решение. Значения p_i вычисляем как вероятности попадания на участки (20; 30), (30; 40), и т.д. для случайной величины, распределенной по закону равномерной плотности на отрезке (23,6; 96,9), по выражению:

$$P(a < X < b) = \frac{b - a}{\beta - a}.$$

Значения вероятности p_i на отрезке (23,6; 96,9) равна 0,136 (вычислено по статистическому ряду).

Составляем сравнительную таблицу чисел попаданий в разряды m_i и соответствующих значений np_i ($n = 400$).

Таблица 1.16

Интервал	20;30	30;40	40;50	50;60	60;70	70;80	80;90	90;100
m_i	21	72	66	38	51	56	64	32
$n \cdot p_i$	34,9	54,57	54,57	54,57	54,57	54,57	54,57	37,7

Определим значение меры расхождения χ^2 по выражению

$$\chi^2 = \sum_{i=1}^8 \frac{(m_i - np_i)^2}{np_i}$$

$$\chi^2 = \frac{(21 - 34,9)^2}{34,9} + \frac{(72 - 54,57)^2}{54,57} + \frac{(66 - 54,57)^2}{54,57} + \frac{(38 - 54,57)^2}{54,57} + \frac{(51 - 54,57)^2}{54,57} + \frac{(56 - 54,57)^2}{54,57} + \frac{(64 - 54,57)^2}{54,57} + \frac{(32 - 37,7)^2}{37,7} = 21,282.$$

Определяем число степеней свободы как число разрядов минус число наложенных связей s (в данном случае $s = 3$):

$$r = 8 - 3 = 5$$

По таблицам χ^2 (справочных приложений) находим для $r = 5$:
при $\chi^2 = 20,50$ $p = 0,001$.

Следовательно, наблюдаемое нами расхождение между теоретическим и статистическим распределениями могло бы за счет чисто случайных причин появиться лишь с вероятностью $p \approx 0,001$. Так как эта вероятность очень мала, следует признать экспериментальные данные *противоречащими* гипотезе о том, что величина X распределена по закону равномерной плотности.

Кроме критерия χ^2 , для оценки степени согласованности теоретического и статистического распределений на практике применяется еще ряд других критериев. Из них мы вкратце остановимся на критерии А. Н. Колмогорова.

В качестве меры расхождения между теоретическим и статистическим распределениями А. Н. Колмогоров рассматривает максимальное значение модуля разности между статистической функцией распределения $F^*(x)$ и соответствующей теоретической функцией распределения.

$$D = \max |F^*(x) - F(x)|. \quad (1.4)$$

Основанием для выбора в качестве меры расхождения величины D является простота ее вычисления. Вместе с тем она имеет достаточно простой закон распределения. А. Н. Колмогоров доказал, что, какова бы ни была функция распределения $F(x)$ непрерывной случайной величины X , при неограниченном возрастании числа независимых наблюдений n вероятность неравенства

$$D\sqrt{n} \geq \lambda \quad (1.5)$$

стремится к пределу

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda^2} \quad (1.6)$$

Значения вероятности $P(\lambda)$, подсчитанные по формуле (1.6) приведены в табл. 1.17

Таблица 1.17

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0,0	1,000	0,7	0,711	1,4	0,040
0,1	1,000	0,8	0,544	1,5	0,022
0,2	1,000	0,9	0,393	1,6	0,012
0,3	1,000	1,0	0,270	1,7	0,006
0,4	0,997	1,1	0,178	1,8	0,003
0,5	0,964	1,2	0,112	1,9	0,002
0,6	0,864	1,3	0,068	2,0	0,001

Схема применения критерия А. Н. Колмогорова следующая: строятся статистическая функция распределения $F^*(x)$ и предполагаемая теоретическая функция распределения $F(x)$, и определяется максимум D модуля разности между ними (рис. 1.11).

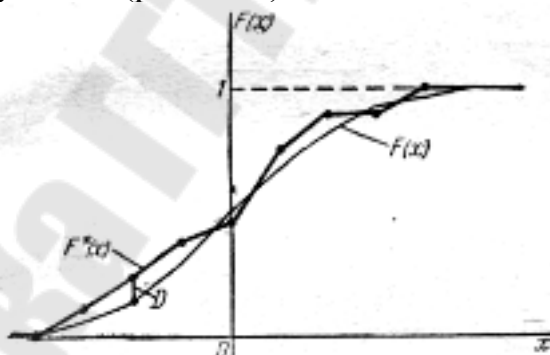


Рис. 1.11

Далее определяется величина

$$l = D\sqrt{n} \quad (1.7)$$

и по табл. 1.17 находится вероятность $P(X)$. Это есть вероятность того, что (если величина X действительно распределена по закону $F(x)$) за счет чисто случайных причин максимальное расхождение между $F^*(x)$ и $F(x)$ будет не меньше, чем фактически наблюдаемое. Если вероятность $P(\lambda)$ весьма мала, гипотезу следует отвергнуть как неправдоподобную;

при сравнительно больших $P(\lambda)$ ее можно считать совместимой с опытными данными.

Критерий А. Н. Колмогорова своей простотой выгодно отличается от критерия χ^2 , поэтому его весьма охотно применяют на практике. Следует, однако, оговорить, что этот критерий можно применять только в случае, когда гипотетическое распределение $F(x)$ полностью известно заранее из каких-либо теоретических соображений, т. е. когда известен не только вид функции распределения $F(x)$, но и все входящие в нее параметры. Такой случай сравнительно редко встречается на практике. Обычно из теоретических соображений известен только общий вид функции $F(x)$, а входящие в нее числовые параметры определяются по данному статистическому материалу. При применении критерия χ^2 это обстоятельство учитывается соответствующим уменьшением числа степеней свободы распределения χ^2 . Критерий А. Н. Колмогорова такого согласования не предусматривает. Если все же применять этот критерий в тех случаях, когда параметры теоретического распределения выбираются по статистическим данным, критерий дает заведомо завышенные значения вероятности $P(\lambda)$; поэтому мы в ряде случаев рискуем принять как правдоподобную гипотезу, в действительности плохо согласующуюся с опытными данными.

Справочные таблицы по критериям согласия приведены в приложении [2].

2 КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

2.1 Связи функциональные и статистические

Связи между различными явлениями в природе сложны и многообразны, однако их можно определенным образом классифицировать. В технике и естествознании часто речь идет о функциональной зависимости между переменными x и y , когда каждому возможному значению x поставлено в однозначное соответствие определенное значение y . Это может быть, например, зависимость между давлением и объемом газа (закон Бойля – Мариотта).

В реальном мире многие явления природы происходят в обстановке действия многочисленных факторов, влияние-каждого из которых ничтожно, а число их велико. В этом случае связь теряет свою одно-

значность и изучаемая физическая система переходит не в определенное состояние, а в одно из возможных для нее состояний. Здесь речь может идти лишь о так называемой статистической связи. Статистическая связь состоит в том, что одна случайная переменная реагирует на изменение другой изменением своего закона распределения. Следовательно, для изучения статистической зависимости нужно знать аналитический вид двумерного распределения. Однако нахождение аналитического вида двумерного распределения по выборке ограниченного объема, во-первых, громоздко, во-вторых, может привести к значительным ошибкам. Поэтому на практике при исследовании зависимостей между случайными переменными X и Y обычно ограничиваются изучением зависимости между одной из них и условным математическим ожиданием другой, т.е. $M(Y|X=x)=\varphi(x)$ или $M(X|Y=y) = \varphi(y)$,

где $M(Y|X=x)$ – математическое ожидание случайной величины Y при условии, что случайная величина X приняла значение x . Аналогично для $M(X|Y = y)$.

Вопрос о том, что принять за зависимую переменную, а что – за независимую, следует решать применительно к каждому конкретному случаю.

Знание статистической зависимости между случайными переменными имеет большое практическое значение: с ее помощью можно прогнозировать значение зависимой случайной переменной в предположении, что независимая переменная примет определенное значение. Однако, поскольку понятие статистической зависимости относится к осредненным условиям, прогнозы не могут быть безошибочными. Применяя некоторые вероятностные методы, как будет показано далее, можно вычислить вероятность того, что ошибка прогноза не выйдет за определенные границы.

2.2 Основные положения корреляционного анализа

Статистические связи между переменными можно изучать методом корреляционного и регрессионного анализа. С помощью этих методов решают разные задачи; требования, предъявляемые к исследуемым переменным, в каждом методе различны.

Основная задача корреляционного анализа – выявление связи между случайными переменными путем точечной и интервальной оценки

парных коэффициентов корреляции, вычисления и проверки значимости множественных коэффициентов корреляции и детерминации, оценки частных коэффициентов корреляции. Корреляционный анализ позволяет также оценить функцию регрессии одной случайной переменной на другую. Предпосылки корреляционного анализа следующие: 1) переменные величины должны быть случайными; 2) случайные величины должны иметь совместное нормальное распределение.

Рассмотрим простейший случай корреляционного анализа – двумерную модель. Введем основные понятия и опишем принцип проведения корреляционного анализа. Результаты этого параграфа легко обобщить на случай многомерной модели. Пусть X и Y – случайные переменные, имеющие совместное нормальное распределение. В этом случае, связь между X и Y можно описать коэффициентом корреляции ρ . Этот коэффициент определяется как ковариация между X и Y , отнесенная к их средним квадратическим отклонениям:

$$\rho = \frac{K_{XY}}{\sigma_x \sigma_y}, \text{ или } \rho = M \left\{ \frac{(X - M(x))(Y - M(y))}{\sigma_x \sigma_y} \right\} \quad (2.1)$$

Оценкой коэффициента корреляции является выборочный коэффициент корреляции r . Для его нахождения необходимо знать оценки следующих параметров: $M(x)$, $M(y)$, σ_x , σ_y . Наилучшей оценкой математического ожидания является среднее арифметическое, т.е.

$M(x) = \bar{x} = \sum_{i=1}^n x_i / n$. Оценкой дисперсии служит выборочная дисперсия, т. е.

$$\sigma_x^2 = s_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / n \quad (2.2)$$

Тогда выборочный коэффициент корреляции

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n s_x s_y}. \quad (2.3)$$

Коэффициент ρ называют также парным коэффициентом корреляции, а r – выборочным парным коэффициентом корреляции.

При совместном нормальном законе распределения случайных величин X и Y , используя рассмотренные выше параметры распределения и коэффициент корреляции, можно получить выражение для условного математического ожидания, т. е. записать выражение для функции регрессии одной случайной величины на другую. Так, функция регрессии Y на X имеет вид

$$M(Y \setminus X = x) = M(y) + \rho \frac{\sigma_y}{\sigma_x} [X - M(x)]; \quad (2.4)$$

функция регрессии X на Y – следующий вид:

$$M(X \setminus Y = y) = M(x) + \rho \frac{\sigma_x}{\sigma_y} [Y - M(y)]; \quad (2.5)$$

Выражения $\rho \frac{\sigma_y}{\sigma_x}$ и $\rho \frac{\sigma_x}{\sigma_y}$ называют коэффициентами регрессии

Подставив в (2.4) соответствующие оценки параметров, получим уравнения регрессии, график которых – прямая линия.

Запишем уравнение регрессии y на x и x на y :

$$\bar{y}(x) = \bar{y} + r \frac{s_y}{s_x} (x - \bar{x}), \quad \bar{x}(y) = \bar{x} + r \frac{s_x}{s_y} (y - \bar{y}). \quad (2.6)$$

Таким образом, в корреляционном анализе на основе оценок параметров двумерной нормальной совокупности получаем оценки тесноты связи между случайными переменными. Особенностью корреляционного анализа является строго линейная зависимость между переменными. Это обуславливается исходными предпосылками. На практике корреляционный анализ можно применять для обработки наблюдений, сделанных на предприятиях при нормальных условиях работы, если случайные изменения того или другого фактора вызывают случайные изменения свойств продукции.

2.3 Свойства коэффициента корреляции

Коэффициент корреляции является одним из самых распространенных способов измерения связи между случайными переменными. Рассмотрим некоторые свойства этого коэффициента.

Коэффициент корреляции принимает значения на интервале $(-1, +1)$.

Если коэффициент корреляции положителен, то связь между переменными также положительна и значения переменных увеличиваются или уменьшаются одновременно. Если коэффициент корреляции имеет отрицательное значение, то при увеличении одной переменной уменьшается другая.

Коэффициент корреляции не зависит от выбора начала отсчета и единицы измерения, т. е. от любых постоянных a_1 и a_2 , b_1 и b_2 таких, что $a_1 > 0$ и $a_2 > 0$, т. е.

$$\rho(a_1X + b_1; a_2Y + b_2) = \rho_{xy}. \quad (2.7)$$

Таким образом, переменные X и Y можно уменьшать или увеличивать в a раз, а также вычитать или прибавлять к значениям X и Y одно и то же число b . В результате величина коэффициента корреляции не изменится.

Если коэффициент корреляции $\rho_{xy} = 0$, то случайные переменные некоррелированы. Понятие некоррелированности не следует смешивать с понятием независимости, независимые величины всегда некоррелированы. Однако обратное утверждение неверно: некоррелированные величины могут быть зависимы и даже функционально, однако эта связь не линейная.

Выборочный коэффициент корреляции вычисляют по формуле (2.3). Имеется несколько модификаций этой формулы, которые удобно использовать при той или иной форме представления исходной информации. Так, при малом числе наблюдений выборочный коэффициент корреляции удобно вычислять по формуле

$$r = \frac{n \sum_x xy - \sum_x x \sum_y y}{\sqrt{n \sum_x x^2 - (\sum_x x)^2} \sqrt{n \sum_y y^2 - (\sum_y y)^2}}. \quad (2.8)$$

Если информация имеет вид корреляционной таблицы, то удобно пользоваться формулой

$$r = \frac{\sum_{x,y} xym_{xy} - \sum_x xm_x \sum_y ym_y / \sum_{x,y} m_{xy}}{\sum_{x,y} m_{xy} s_x s_y}, \quad (2.9)$$

где $\sum m_x = \sum m_y = m_{xy} = n$; m_x – суммарная частота наблюдаемого значения признака x при всех значениях y ; m_y – суммарная частота наблюдаемого значения признака y при всех значениях x ; m_{xy} – частота появления пары признаков (x, y) .

Из формулы (2.3) очевидно, что $r_{xy} = r_{yx}$, т. е. величина выборочного коэффициента корреляции не зависит от порядка следования переменных, поэтому обычно пишут просто r .

Эмпирическое корреляционное отношение является одним из показателей, характеризующих тесноту корреляционной зависимости, то есть степень ее приближения к функциональной связи.

Корреляционное отношение определяется по формуле:

$$\eta = \sqrt{\frac{\delta^2}{\sigma^2}}, \quad (2.10)$$

где δ^2 – групповая дисперсия;
 σ^2 – общая дисперсия.

Корреляционное отношение изменяется от 0 до 1.

Для качественной оценки тесноты связи на основе показателя эмпирического корреляционного отношения можно пользоваться табл. 2.1

Таблица 2.1

Величина η	0,1-0,3	0,3-0,5	0,5-0,7	0,7-0,9	0,9-0,99
Сила связи	слабая	умеренная	заметная	тесная	весьма тесная

2.4 Поле корреляции. Вычисление оценок параметров двумерной модели

На практике для вычисления оценок параметров двумерной модели удобно использовать *корреляционную таблицу и поле корреляции*. Пусть, например, изучается зависимость между объемом выполненных работ (y) и накладными расходами (x). Имеем выборку из генеральной совокупности, состоящую из 150 пар переменных (x_i, y_i) . Считаем, что предпосылки корреляционного анализа выполнены.

Пару случайных чисел (x_i, y_i) можно изобразить графически в виде точки с координатами (x_i, y_i) . Аналогично можно изобразить весь набор пар случайных чисел (всю выборку). Однако при большом объеме выборки это затруднительно. Задача упрощается, если выборку упорядочить, т.е. переменные сгруппировать. Сгруппированные ряды могут быть как дискретными, так и интервальными.

По осям координат откладывают или дискретные значения переменных, или интервалы их изменения. Для интервального ряда наносят координатную сетку. Каждую пару переменных из данной выборки изображают в виде точки с соответствующими координатами для дискретного ряда или в виде точки в соответствующей клетке для интервального ряда.

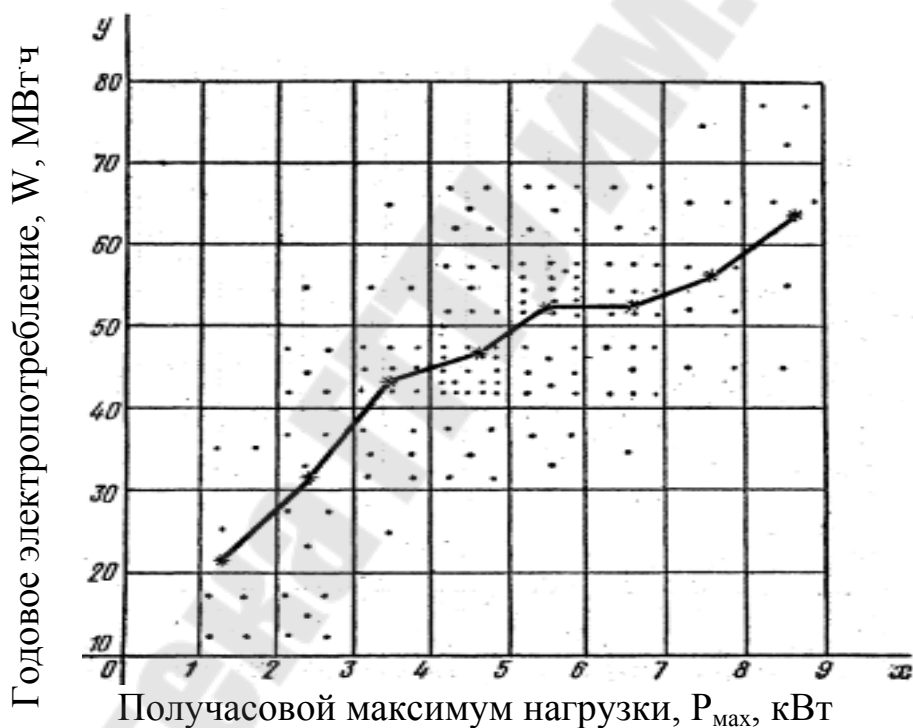


Рис. 2.1

Такое изображение корреляционной зависимости называют *полем корреляции*. На рисунке изображено поле корреляции для выборки, состоящей из 150 пар переменных (ряд интервальный). Если вычислить средние значения y в каждом интервале изменения x [обозначим их $y_i(x)$], нанести эти точки на рис.2.1 и соединить между собой, то получим ломаную линию, по виду которой можно судить, как в среднем меняются y в зависимости от изменения x . По виду этой линии можно

также сделать предположение о форме связи между переменными. В данном случае ломаную линию можно аппроксимировать прямой линией, так как она достаточно хорошо приближается к ней.

Таблица 2.1

Годовое электропотребление, y , МВт	Получасовой максимум нагрузки, x , кВт								
	1-2 1,5	2-3 2,5	3-4 3,5	4-5 4,5	5-6 5,5	6-7 6,5	7-8 7,5	8-9 8,5	m_y
10-20 15	4	5							9
20-30 25	1	3	1						5
30-40 35	2	3	6	5	3	1			20
40-50 45		5	9	19	8	7	2	1	51
50-60 55		1	2	7	16	9	4	2	41
60-70 65			1	5	6	4	2	2	20
70-80 75							1	3	4
m_x	7	17	19	36	33	21	9	8	150

По выборочным данным можно построить также корреляционную таблицу (табл. 2.1). Корреляционную таблицу, как и поле корреляции, строят по сгруппированному ряду (дискретному или интервальному). Таблица построена на основе интервального ряда. В первой строке и первом столбце таблицы помещают интервалы изменения x и y и значения середин интервалов. Так, например, 1,5 – середина интервала изменения $x = 1 + 2,15$ – середина интервала изменения $y = 10 + 20$. В ячейки, образованные пересечением строк и столбцов, заносят частоты попадания пар значений (x, y) в соответствующие интервалы по x и y . Например, частота 4 означает, что в интервал изменения y от 10 до 20 попало 4 пары наблюдавшихся значений. Эти частоты обозначают m_{xy} .

В 9-й строке и 10-м столбце находятся значения m_x и $m_y - m_{xy}$ по соответствующим столбцу и строке.

Как будет показано в дальнейшем, корреляционной таблицей удобно пользоваться при вычислении коэффициентов корреляции и параметров уравнений регрессии.

Корреляционная таблица построена на основе интервального ряда, поэтому для оценок параметров воспользуемся формулами для вычисления средней арифметической и дисперсии. Имеем:

$$\bar{x} = \frac{\sum_x x m_x}{\sum_x m_x}; \bar{y} = \frac{\sum_y y m_y}{\sum_y m_y}; \quad (2.11)$$

$$s_x^2 = \frac{\sum_x x^2 m_x}{\sum_x m_x} - \left(\frac{\sum_x x m_x}{\sum_x m_x} \right)^2; s_y^2 = \frac{\sum_y y^2 m_y}{\sum_y m_y} - \left(\frac{\sum_y y m_y}{\sum_y m_y} \right)^2.$$

Пример 2.1

По данным, приведенным в таблице 2.1, вычислить оценки параметров двумерной модели.

Решение. Вычислим все необходимые для расчета суммы по формулам (2.6)-(2.8).

Найдем оценки параметров \bar{x} , \bar{y} , s_x , s_y . Получаем:

$$\bar{x} = \frac{735}{150} = 4,9; \bar{y} = \frac{7110}{150} = 47,4;$$

$$s_x = \sqrt{\frac{4053,5}{150} - \left(\frac{735}{150} \right)^2} = 1,74;$$

$$s_x = \sqrt{\frac{363950}{150} - \left(\frac{7110}{150}\right)^2} = 13,4;$$

Вычислим выборочный коэффициент корреляции:

$$r = \frac{37175,0 + \frac{735 \cdot 7110}{150}}{150 \cdot 1,74 \cdot 13,4} = 0,67$$

2.5 Корреляционное отношение

На практике часто предпосылки корреляционного анализа нарушаются: один из признаков оказывается величиной не случайной, или признаки не имеют совместного нормального распределения. Однако статистическая зависимость между ними существует. Для изучения связи между признаками в этом случае существует общий показатель зависимости признаков, основанный на показателе изменчивости – общей (или полной) дисперсии.

Полной называется дисперсия признака относительно его математического ожидания. Так, для признака Y это $y_y^2 = M[Y - M(Y)]^2$. Дисперсию y_y^2 можно разложить на две составляющие, одна из которых характеризует влияние фактора X на Y , другая – влияние прочих факторов. Очевидно, чем меньше влияние прочих факторов, тем теснее связь, тем более приближается она к функциональной. Представим y_y^2 в следующем виде:

$$y_y^2 = M[M(Y|X = x) - M(Y)]^2 + M[Y - M(Y|x)]^2. \quad (2.12)$$

Первое слагаемое обозначим $\delta_{y|x}^2$. Это дисперсия функции регрессии относительно математического ожидания признака (в данном случае признака Y); она измеряет влияние признака X на Y . Второе слагаемое обозначим $\sigma_{y|x}^2$. Это дисперсия признака Y относительно функции рег-

рессии. Ее называют также *средней* из условия дисперсий или *остаточной* дисперсией; $\sigma_{y|x}^2$ измеряет влияние на Y прочих факторов.

Тесноту связи удобно оценивать в единицах общей дисперсии σ_Y^2 , т.е. рассматривать отношение $\frac{\{M[\bar{y}(x) - M(Y)]\}^2}{\sigma_Y^2}$. Эту величину обозначают $\eta_{TY|x}^2$ и называют *теоретическим корреляционным отношением*. Таким образом,

$$\eta_{TY|x}^2 = \frac{M[\bar{y}(x) - M(Y)]^2}{\sigma_Y^2} = \frac{\delta_{Y|X}^2}{\sigma_Y^2}. \quad (2.13)$$

Разделив обе части равенства (2.12) на σ_Y^2 , получим

$$1 = \frac{\sigma_{Y|X}^2 + \delta_{Y|X}^2}{\sigma_Y^2}, \text{ или } 1 = \frac{\sigma_{Y|X}^2}{\sigma_Y^2} + \eta_{TY|x}^2.$$

Из последней формулы имеем

$$\eta_{TY|x}^2 = 1 - \frac{\sigma_{Y|X}^2}{\sigma_Y^2}. \quad (2.14)$$

При вычислении η^2 по выборочным данным получаем *выборочное корреляционное отношение*. Обозначим его $\bar{\eta}^2$. Вместо дисперсий в этом случае используются их оценки. Тогда формула (2.14) примет вид

$$\bar{\eta}^2 = \frac{s_{Y|X}^2}{s_Y^2}. \quad (2.15)$$

Пример 2.2

По данным, приведенным в табл. 3.3, вычислить выборочное корреляционное отношение.

Решение. Вычислим $\overline{\eta^2}$ по формуле, приведенной выше. Представим s_y^2 и $s_{y|x}^2$ в следующем виде:

$$s_y^2 = \frac{\sum_y y^2 m_y}{\sum_y m_y} - \left(\frac{\sum_y y m_y}{\sum_y m_y} \right)^2.$$

$$s_{y|x}^2 = \frac{1}{\sum_{x,y} m} \sum_{x,y} [\overline{y(x)} - \overline{y}]^2 m_x$$

Суммы, необходимые для вычислений, приведены в табл.3.3 и 2.2.

Имеем $s_y^2 = 0,6746$

Найдем \overline{y}

$$\overline{y} = \frac{\sum_y y m_y}{\sum_y m_y} = 4,824.$$

Имеем $s_{y|x}^2 = 0,5528$.

Таким образом, $\overline{\eta^2} = 0,8194$.

Таблица 2.2

$\overline{y(x)}$	$\overline{y(x)} - \overline{y}$	m_x	$[\overline{y(x)} - \overline{y}]^2 m_x$
6,06	1,23	8	12,1032
5,29	0,47	9	1,9881
5,20	0,38	6	0,8664
4,50	-0,32	8	0,8192
4,53	-0,29	6	0,5046
4,16	-0,66	5	2,1780
3,80	-1,02	4	4,1616
3,70	-1,12	4	5,0176
Σ		50	27,6387

3 РЕГРЕССИОННЫЙ АНАЛИЗ

3.1 Определение формы связи. Понятие регрессии

Определить форму связи – значит выявить механизм получения зависимой случайной переменной. При изучении статистических зависимостей форму связи можно характеризовать функцией регрессии (линейной, квадратной, показательной и т. д.).

Условное математическое ожидание $M(Y|X=x)$ случайной переменной Y , рассматриваемое как функция x , т. е. $M(Y|X=x) = \varphi(x)$, называется функцией регрессии случайной переменной Y относительно X (или функцией регрессии Y по X). Точно так же условное математическое ожидание $M(X|Y=y)$ случайной переменной X , т. е. $M(X|Y=y) = \varphi(y)$, называется функцией регрессии случайной переменной X относительно Y (или функцией регрессии X по Y).

На примере дискретного распределения найдем функцию регрессии.

Пример 3.1

Задана двухмерная случайная переменная с плотностью вероятности

$$f(x, y) = \begin{cases} x + 2y & \text{при } 0 \leq x \leq 1; 0 \leq y \leq 1, \\ 0 & \text{для всех остальных значений} \end{cases}$$

Найти функцию регрессии случайной переменной Y для случая, когда X принимает значение x .

Решение: Для определения функции регрессии нужно знать условную плотность вероятности переменной Y при $X = x$. Имеем

$$f(Y|X=x) = \frac{f(x, y)}{\int_0^1 f(x, y) dy} = \frac{x + 2y}{\int_0^1 (x + 2y) dy} = \frac{x + 2y}{x + 1}$$

Тогда функция регрессии

$$M(Y|X=x) = \int_0^1 y f(Y|X=x) dy = \int_0^1 y \frac{x + 2y}{x + 1} dy = \frac{0.5x + 2/3}{x + 1}$$

Функция регрессии имеет важное значение при статистическом анализе зависимостей между переменными и может быть использована для прогнозирования одной из случайных переменных, если известно

значение другой случайной переменной. Точность такого прогноза определяется дисперсией условного распределения.

Несмотря на важность понятия функции регрессии, возможности ее практического применения весьма ограничены. Как видно из приведенного примера, для оценки функции регрессии необходимо знать аналитический вид двумерного распределения (X, Y) . Только в этом случае можно точно определить вид функции регрессии, а затем оценить параметры двумерного распределения. Однако, для подобной оценки мы чаще всего располагаем лишь выборкой ограниченного объема, по которой нужно найти вид двумерного распределения (X, Y) , а затем вид функции регрессии. Это может привести к значительным ошибкам, так как одну и ту же совокупность точек (x_i, y_i) на плоскости можно одинаково успешно описать с помощью различных функций. Именно поэтому возможности практического применения функции регрессии ограничены. Для характеристики формы связи при изучении зависимости используют понятие кривой регрессии.

Кривой регрессии Y по X (или Y на X) называют условное среднее значение случайной переменной Y , рассматриваемое как функция определенного класса, параметры которой находят методом наименьших квадратов по наблюдаемым значениям двумерной случайной величины (x, y) , т. е.

$$\bar{y}(x) = \varphi(x, b_1, b_2, \dots, b_m). \quad (3.1)$$

Аналогично определяется кривая регрессии X по Y (X на Y):

$$\bar{x}(y) = \varphi(y, c_1, c_2, \dots, c_m). \quad (3.2)$$

Кривую регрессии называют также эмпирическим уравнением регрессии или просто уравнением регрессии. Уравнение регрессии является оценкой соответствующей функции регрессии.

Возникает вопрос: почему для определения кривой регрессии используют именно условное среднее $\bar{y}(x)$. Функция $y(x)$ обладает одним замечательным свойством: она дает наименьшую среднюю погрешность оценки прогноза. Предположим, что кривая регрессии – произвольная функция. Средняя погрешность прогноза по кривой регрессии определяется математическим ожиданием квадрата разности между измеренной величиной и вычисленной по формуле кривой регрессии, т. е. $M[Y - \varphi(x)]^2$. Естественно потребовать вычисления такой кривой регрес-

сии, средняя погрешность прогноза по которой была бы наименьшей. Таковой является $\varphi(x) = \bar{y}(x)$. Это следует из свойств минимальности рассеивания около центра распределения $\bar{y}(x)$. Если рассеивание вычисляется относительно $\varphi(x)$ не равно $\bar{y}(x)$, то средний квадрат отклонения увеличивается. Поэтому можно сказать, что кривая регрессии, выражаемая как $\bar{y}(x)$, минимизирует среднюю квадратическую погрешность прогноза величины Y по X .

3.2 Основные положения регрессионного анализа

Под регрессией понимается функция, предназначенная для описания зависимости изменения результативных признаков под влиянием колебаний признаков-факторов. Понятие регрессии введено в статистическую науку по предложению английского ученого Ф. Гальтона.

В корреляционно-регрессионном методе парной корреляционной взаимосвязи соответствует однофакторная регрессионная модель, множественной взаимосвязи – множественная регрессия. Поэтому наличие корреляционной связи между параметрическими признаками позволяет приближенно предоставить значения результативного признака в виде некоторой функции от величины одного или нескольких факторных признаков.

Функцию, показывающую корреляционную зависимость между признаками, принято называть *уравнением регрессии*. Если оно связывает лишь два признака, то представляет собой *уравнение парной регрессии*; если отражает зависимость результативного признака от двух, трех и более факторных признаков – это *уравнение множественной регрессии*.

Нами показано, что при выявлении корреляционной формы, связывающей результативный признак с одним факторным, помогает графическое изображение корреляционной связи в виде поля корреляции. Обычно считают, что увеличение результативного и факторного признаков в арифметической прогрессии при прямой связи требует применения прямолинейной, а при обратной – гиперболической регрессии.

Прямая связь, при которой результативный признак увеличивается в арифметической прогрессии, а факторный повышается быстрее признака-результата, требует применения параболической или показательной регрессии. Уравнение множественной регрессии обычно выражается либо прямой, представляющей собой функцию многих переменных, либо степенной функцией.

Составление уравнения регрессии означает прежде всего определение его параметров, используя для этого, где возможно, способ наименьших квадратов, согласно которому сумма квадратов отклонений фактических значений результативного признака от теоретических значений, рассчитанных по уравнению регрессии, должна быть наименьшей, т. е.

$$\sum (y - y_x) \rightarrow \min, \quad (3.3)$$

где y – фактические варианты признака-результата; y_x – теоретические значения признака-результата.

Это условие приводит к системе нормальных уравнений, решение которых позволяет определить параметры уравнения регрессии. Заметим при этом, что число нормальных уравнений на одно больше числа входящих в уравнение регрессии факторов. Если известны параметры уравнения, то, подставляя в него принятые значения факторных признаков, можно рассчитать теоретическое значение результативного признака, что делает удобным применение корреляционных уравнений при прогнозировании результативных признаков.

Уравнение регрессии может показать связь между признаками более точно, если оно построено на основании достаточно большой статистической совокупности. Но поскольку оно все-таки выражает приближенную меру связи, то уравнение регрессии нередко называют *моделью* связи между признаками.

Основная задача регрессионного анализа – изучение зависимости между результативным признаком Y и наблюдавшимся признаком X , оценка функции регрессии,

Предпосылки регрессионного анализа:

1) Y – независимые случайные величины, имеющие постоянную дисперсию;

2) X – величины наблюдаемого признака (величины не случайные);

3) условное математическое ожидание $M(Y|X=x)$ можно представить в виде

$$M(Y|X = x) = \varphi(x) = \epsilon_0 + \epsilon_1 x. \quad (3.4)$$

Выражение (3.4), называется функцией регрессии (или модельным уравнением регрессии) Y на X . Оценке в этом выражении подлежат параметры β_0 и β_1 , называемые коэффициентами регрессии, а также $\sigma_{\text{ост}}^2$ – остаточная дисперсия.

Остаточной дисперсией называется та часть рассеивания результативного признака, которую нельзя объяснить действием наблюдаемого признака. Остаточная дисперсия может служить для оценки точности подбора вида функции регрессии (модельного уравнения регрессии), полноты набора признаков, включенных в анализ. Оценки параметров функции регрессии находят, используя метод наименьших квадратов.

Рассмотрим линейный регрессионный анализ. Линейным он называется потому, что изучает лишь те виды зависимостей $\varphi(x)$, которые линейны по оцениваемым параметрам, хотя могут быть нелинейны по переменным X .

Например, зависимости

$$M(Y|X = x) = \varphi_1(x) = \epsilon_0 + \epsilon_1 x, M(Y \setminus X = x) = \varphi_2(x) = \beta_0 + \beta_1 / x,$$

$M(Y \setminus X = x) = \varphi_3(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ линейны относительно параметров $\beta_0, \beta_1, \beta_2$, хотя вторая и третья зависимости нелинейны относительно переменных x . Вид зависимости $\varphi(x)$ выбирают, исходя из визуальной оценки характера расположения точек на поле корреляции; опыта предыдущих исследований; соображений профессионального характера, основанных на знании физической сущности процесса.

Важное место в линейном регрессионном анализе занимает так называемая «нормальная регрессия». Она имеет место, если сделать предположения относительно закона распределения случайной величины Y . Предпосылки «нормальной регрессии»:

1) Y – независимые случайные величины, имеющие постоянную дисперсию и распределенные по нормальному закону;

2) X – величины наблюдаемого признака (величины не случайные);

3) условное математическое ожидание $M(Y|X=x)$ можно представить в виде (3.4).

В этом случае оценки коэффициентов регрессии – несмещенные с минимальной дисперсией и нормальным законом распределения. Из этого положения следует, что при «нормальной регрессии» имеется возможность оценить значимость оценок коэффициентов регрессии, а также построить доверительный интервал для коэффициентов регрессии и условного математического ожидания $M(Y|X=x)$.

3.3 Линейная регрессия

Рассмотрим простейший случай регрессионного анализа – модель вида (3.4), когда зависимость $\varphi(x)$ линейна и по оцениваемым параметрам и по переменным. Оценки параметров модели (3.4) β_0 и β_1 обозначим b_0 и b_1 . Оценку остаточной дисперсии $\sigma_{ост}^2$ обозначим $s_{ост}^2$. Подставив в формулу вместо параметров их оценки, получим уравнение регрессии $\bar{y}(x) = b_0 + b_1 \cdot x_1$, коэффициенты которого b_0 и b_1 находят из условия минимума суммы квадратов отклонений измеренных значений результативного признака y_i от вычисленных по уравнению регрессии $y(x_i)$:

$$Q = \sum_{i=1}^n [y_i - y(x_i)]^2 = \min, \text{ или}$$

$$Q = \sum_{i=1}^n [y_i - b_0 - b_1 x_i]^2 = \min.$$

Составим систему нормальных уравнений: первое уравнение

$$\frac{\partial Q}{\partial b_0} = 2 \sum_{i=1}^n [y_i - b_0 - b_1 x_i] = 0$$

$$\sum_{i=1}^n y_i - \sum_{i=1}^n a_0 - \sum_{i=1}^n b_1 x_i = 0$$

$$\text{откуда } nb_0 + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

второе уравнение

$$\frac{\partial Q}{\partial b_1} = 2 \sum_{i=1}^n x_i [y_i - b_0 - b_1 x_i] = 0$$

$$\text{откуда } b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$$

Итак,

$$\begin{cases} nb_0 + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases} \quad (3.5)$$

Оценки, полученные по способу наименьших квадратов, обладают дисперсией в классе линейных оценок. Решая систему (3.5) относительно b_0 и b_1 , найдем оценки параметров b_0 и b_1 :

$$b_1 = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2 / n}, \quad (3.6)$$

$$b_0 = \left(\sum_{i=1}^n y_i \right) / n - b_1 \left(\sum_{i=1}^n x_i \right) / n = \bar{y} - b_1 \bar{x}, \quad (3.7)$$

Остается получить оценку параметра σ_{ocm}^2 . Имеем

$$s_{ocm}^2 = \frac{1}{n-2} \sum_{i=1}^n [y_i - y(x_i)]^2, \quad (3.8)$$

где n – количество наблюдений.

Если n велико, то для упрощения расчетов наблюдавшиеся данные принято группировать, т.е. строить корреляционную таблицу. Формулы для нахождения коэффициентов регрессии по сгруппированным данным те же, что и для расчета по несгруппированным данным, но суммы

$$\sum_{i=1}^n x_i, \sum_{i=1}^n y_i, \sum_{i=1}^n x_i y_i, \sum_{i=1}^n x_i^2, \sum_{i=1}^n y_i^2, \text{ заменяют на}$$

$$\sum_x x m_x, \sum_y y m_y, \sum_{xy} x y m_{xy}, \sum_x x^2 m_x, \sum_y y^2 m_y,$$

где m_x , m_y и m_{xy} – частоты повторений соответствующих значений переменных. В дальнейшем часто используется этот наглядный прием вычислений.

Пример 3.2

По данным, приведенным в табл. 2.1, оценить параметры уравнения регрессии. Предполагается, что связь выражается формулой (3.4).

Решение. При сгруппированных данных удобно использовать следующие формулы:

$$b_1 = \frac{\sum_{x,y} x y m_{xy} - \left(\sum_x x m_x \sum_y y m_y \right) / \sum_{xy} m_{xy}}{\sum_x x^2 m_x - \left(\sum_x x m_x \right)^2 / \sum_{xy} m_{xy}}$$

$$b_0 = \sum_y y m_y / \sum_y m_y - b_1 \sum_x x m_x / \sum_x m_x$$

Используя суммы, вычисленные в табл. 2.2, получаем:
 $b_1 = 5,167; b_0 = 22,082$.

Уравнение регрессии имеет вид $\bar{y}(x) = 22,082 + 5,167x$.

Промежуточные результаты вычислений оценки остаточной дисперсии приведены в табл. 3.2. По данным табл. 3.1 находим:
 $y = 7110/150 = 47,4$, $\bar{x} = 735/150 = 4,9$.

Для сгруппированного ряда остаточную дисперсию оценим по формуле

$$s_{ocm}^2 = \left\{ \sum_x [y_i - y(x_i)]^2 m_x \right\} / (n - 2),$$

откуда $s_{ocm}^2 = 1585,073/148 = 10,711$, $s_{ocm} = 3,38$.

3.4 Нелинейная регрессия

Рассмотрим случай, когда зависимость нелинейна по переменным x , например модель вида

$$M(Y|X = x) = \varphi(x) = \beta_0 + \beta_1/x. \quad (3.9)$$

На рис. 3.1 изображено поле корреляции. Очевидно, что зависимость между Y и X нелинейная и ее графическим изображением является не прямая, а кривая. Оценкой выражения (3.7) является уравнение регрессии

$$\hat{y}(x) = b_0 - b_1 / x,$$

где b_0 и b_1 – оценки коэффициентов регрессии β_0 и β_1 .

Принцип нахождения коэффициентов тот же – метод наименьших квадратов, т.е.

$$Q = \sum_{i=1}^n [y_i - y(x_i)]^2 = Q \min, \text{ или}$$

$$Q = \sum_{i=1}^n [y_i - b_0 - b_1 / x_i]^2 = Q \min$$

Таблица 3.1

Годовое электропо- требление, у, МВт·ч	Получасовой максимум нагрузки, х, кВт											
	1 – 2 1,5	2 – 3 2,5	3 – 4 3,5	4 – 5 4,5	5 – 6 5,5	6 – 7 6,5	7 – 8 7,5	8 – 9 8,5	m_y	ym_y	y^2m_y	$\sum yxm_{xy}$
10 – 20 15	4	5							9	135	2025	277,5
20 – 30 25	1	3	1						5	125	3125	312,5
30 – 40 35	2	3	6	5	3	1			20	700	24500	2695,0
40 – 50 45		5	9	19	8	7	2	1	51	2295	103275	10912,5
50 – 60 55		1	2	7	16	9	4	2	41	2255	124025	12897,5
60 – 70 65			1	5	6	4	2	2	20	1300	84500	7605,0
70 – 80 75							1	3	4	300	22500	2475,0
m_x	7	17	19	36	33	21	9	8	150	7110	363950	37175,0
xm_x	10,5	42,5	66,5	162,0	181,5	136,5	67,5	68,0	735,0			
x^2m_x	15,75	106,25	232,75	729,0	998,25	887,25	506,25	578,0	4053,5			

Таблица 3.2

y_i	$y(x_i)$	m_x	$y_i - y(x_i)$	$[y_i - y(x_i)]^2 m_x$	$y_i - \bar{y}$	$(y_i - \bar{y})^2 m_x$	$(x_i - \bar{x})^2 m_x$
22,14	29,832	7	-7,69	414,952	-25,26	4466,473	80,92
31,47	35	17	-3,53	211,835	-15,93	4314,003	97,92
42,89	40,166	19	2,72	140,982	-4,51	386,462	37,24
48,33	45,336	36	3	322,704	0,93	31,136	5,78
52,57	50,5	33	2,07	141,402	5,17	882,054	11,88
52,62	55,668	21	-3,05	195,196	5,12	550,502	53,76
57,22	60,834	9	-3,61	117,548	9,82	867,892	60,84
63,75	66,002	8	-2,25	40,554	15,35	1884,980	103,68
Σ		150		1585,073		13383,502	452,02

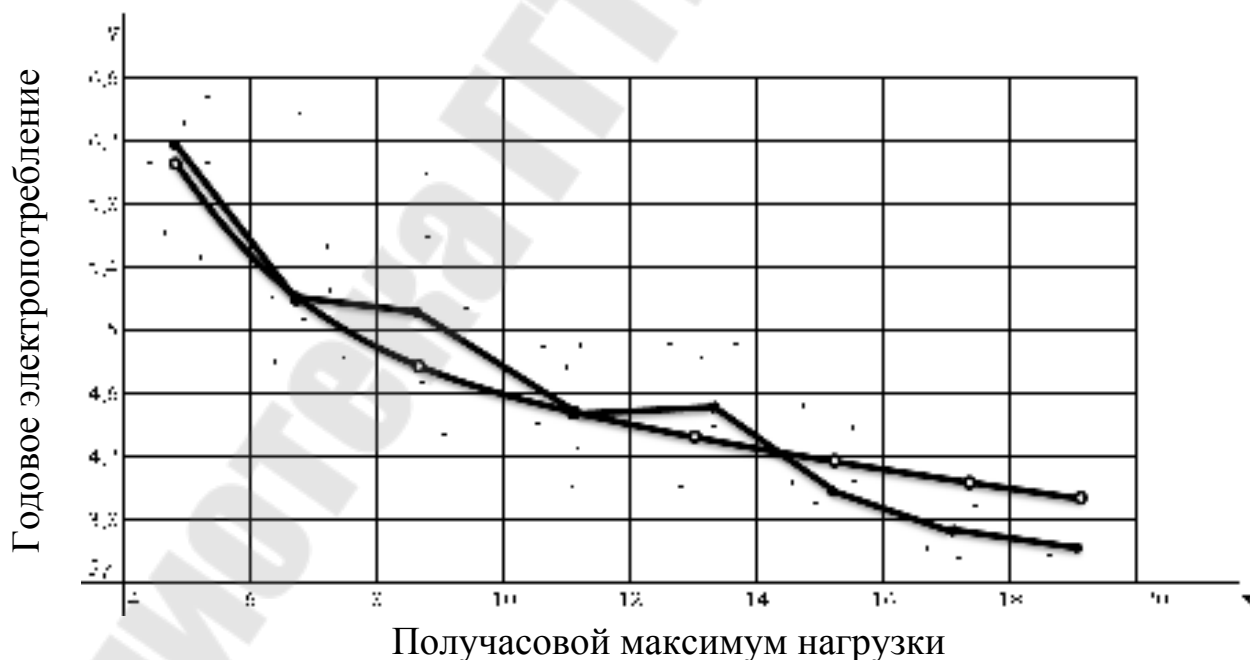


Рис.3.1

Дифференцируя последнее равенство по b_0 и b_1 и приравнивая правые части нулю, получаем так называемую систему нормальных уравнений:

$$\begin{cases} \sum_{i=1}^n b_0 + \sum_{i=1}^n b_1 / x_i = \sum_{i=1}^n y_i \\ \sum_{i=1}^n b_0 / x_i + \sum_{i=1}^n b_1 / x_i^2 = \sum_{i=1}^n y_i / x_i \end{cases} \quad (3.10)$$

Пример 3.3.

По данным, приведенным в таблице 3.3, вычислить коэффициенты уравнения регрессии. Связь между переменными имеет вид (3.9).

Решение.

В табл. 3.3 приведены все необходимые суммы для составления нормальных уравнений, а также средние значения $\bar{y}_i(x)$ для каждого интервала изменения x . Значения $\bar{y}_i(x)$ приведены в последней строке таблицы и нанесены на поле корреляции (рис. 3.1). Решив уравнения (3.10), можно вычислить значения $y_i(x)$ и построить линию регрессии. Запишем систему уравнений (3.10) в виде

$$\begin{cases} nb_0 + b_1 \sum_x m_x / x = \sum_y y m_y \\ b_0 \sum_x m_x / x + b_1 \sum_x m_x / x^2 = \sum_{xy} y m_{xy} / x \end{cases}$$

или

$$\begin{cases} 50b_0 + 5,5203b_1 = 241,2 \\ 5,5203b_0 + 0,7265b_1 = 28,3714 \end{cases}$$

откуда $b_1 = 14,881$, $b_0 = 3,181$. Уравнение регрессии (3.9) теперь принимает вид $\bar{y}(x) = 3,181 + 14,881/x$.

Используя формулу для оценки дисперсии по данным сгруппированного ряда, вычислим $s_{\text{ост}}^2$ (табл. 3.4).

Среднее значение $\bar{y} = \sum y m_y / n = 241,2 / 50 = 4,82$

Таблица 3.3

Годовое электропо- требление, у, МВт·ч	Получасовой максимум нагрузки, х, кВт											
	4 – 6 5	6 – 8 7	8 – 10 9	10 – 12 11	12 – 14 13	14 – 16 15	16 – 18 17	18 – 20 19	m_y	ym_y	$y^2 m_y$	ym_{xy} / x
3,4 – 3,8 3,6							2	3	5	18	64,8	0,9918
3,8 – 4,2 4,0				1	1	3	2	1	8	32	128	2,1524
4,2 – 4,6 4,4			1	4	2	2			9	39,6	174,24	3,3523
4,6 – 5,0 4,8		3	1	3	3				10	48	230,4	5,0074
5,0 – 5,4 5,2		3	2						5	26	135,2	3,3842
5,4 – 5,8 5,6	2	2	2						5	28	156,8	4,4621
5,8 – 6,2 6,0	3		1						4	24	144	4,2666
6,2 – 6,6 6,4	3	1							4	25,6	163,84	4,7546

Продолжение таблицы 3.3

m_x	8	9	6	8	6	5	4	4	50	241,2	1197,28	28,3714
$\frac{1}{x} m_x$	1,600	1,2857	0,6667	0,7273	0,4615	0,3333	0,2353	0,2105	5,5203			
$\frac{1}{x^2} m_x$	0,3200	0,1837	0,0741	0,0661	0,0355	0,0222	0,0138	0,0111	0,7265			
$\bar{y}_{i,x} = \frac{\sum y m_{xy}}{m_x}$	6,05	5,29	5,20	4,50	4,53	4,16	3,80	3,70				

Таблица 3.4

y_i	$y(x_i)$	m	x	$y - y(x_i)$	$[y_i - y(x_i)]^2 m$	$(y_i - \bar{y})^2 m$
6,05	6,16	8	5	-0,11	0,0968	12,1032
5,29	5,31	9	7	-0,02	0,0036	1,9881
5,20	4,83	6	9	+0,37	0,8214	0,8664
4,50	4,53	8	11	-0,03	0,0072	0,8192
4,53	4,32	6	13	+0,21	0,2646	0,5046
4,16	4,17	5	15	-0,01	0,0005	2,1780
3,80	4,05	4	17	-0,25	0,2500	4,1616
3,70	3,96	4	19	-0,26	0,2704	5,0176
Σ					1,7145	27,6387

Остаточная дисперсия $s_{\text{ост}}^2 = 1,7145/48 = 0,358$.

Полученная линия регрессии нанесена на рис 3.1. Как следует из рисунка, форма связи (гипербола) выбрана правильно. Полученная линия регрессии является хорошей аппроксимацией эмпирических данных.

В общем случае нелинейной зависимости между переменными Y и X связь может выражаться многочленом k -й степени от x :

$$M(Y|X = x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_k x^k$$

Коэффициенты регрессии определяют по принципу наименьших квадратов. Система нормальных уравнений имеет вид

$$\left\{ \begin{array}{l} b_0 n + b_1 \sum_{i=1}^n x_i + b_2 \sum_{i=1}^n x_i^2 + \dots + b_k \sum_{i=1}^n x_i^k = \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 + b_2 \sum_{i=1}^n x_i^3 + \dots + b_k \sum_{i=1}^n x_i^{k+1} = \sum_{i=1}^n y_i x_i \\ \dots \\ b_0 \sum_{i=1}^n x_i^k + b_1 \sum_{i=1}^n x_i^{k+1} + b_2 \sum_{i=1}^n x_i^{k+2} + \dots + b_k \sum_{i=1}^n x_i^{2k} = \sum_{i=1}^n y_i x_i^k \end{array} \right.$$

Вычислив коэффициенты системы, ее можно решить любым известным способом.

Пример 3.4.

По данным, приведенным в таблице 3.5, оценить параметры уравнения регрессии. Переменные связаны между собой зависимостью $M(Y|X = x) = \beta_0 + \beta_1 x + \beta_2 x^2$ (уравнение параболы).

Решение. Система нормальных уравнений имеет вид

$$\begin{aligned} b_0 + b_1 \sum_1^{50} x_i + b_2 \sum_1^{50} x_i^2 &= \sum_1^{50} y_i, \\ b_0 \sum_1^{50} x_i + b_1 \sum_1^{50} x_i^2 + b_2 \sum_1^{50} x_i^3 &= \sum_1^{50} y_i x_i, \\ b_0 \sum_1^{50} x_i^2 + b_1 \sum_1^{50} x_i^3 + b_2 \sum_1^{50} x_i^4 &= \sum_1^{50} y_i x_i^2. \end{aligned}$$

Используя данные табл. 3.5, получаем следующую систему нормальных уравнений:

$$\begin{cases} 50b_0 + 125b_1 + 365b_2 = 151 \\ 125b_0 + 365b_1 + 1175b_2 = 405 \\ 365b_0 + 1175b_1 + 4025b_2 = 1237 \end{cases} \quad \text{или} \quad \begin{cases} b_0 + 2,5b_1 + 7,3b_2 = 3,02 \\ b_0 + 2,92b_1 + 9,4b_2 = 3,24 \\ b_0 + 3,22b_1 + 11,03b_2 = 3,39 \end{cases}$$

Решая данную систему уравнений, найдем значения коэффициента регрессии: $b_0 = 1,424$, $b_1 = 0,799$, $b_2 = -0,055$.

Уравнение регрессии имеет вид

$$\bar{y} = 1,424 + 0,799x - 0,055x^2.$$

Оценим остаточную дисперсию. Все вычисления сведены в табл. 3.6.

Таблица 3.5

y	x				m_y	ym_y	yxm_{xy}	yx^2m_{xy}
	1	2	3	4				
5				1	1	5	20	80
4		3	6	6	15	60	192	648
3	3	8	6	3	20	60	147	411
2	5	4	3		12	24	44	96
1	2				2	3	2	2
m_x	10	15	15	10	50	151	405	1237
xm_x	10	30	45	40	125			
x^2m_x	10	60	135	160	365			
x^3m_x	10	120	405	640	1175			
x^4m_x	10	240	1215	2560	4025			

Таблица 3.6

y_i	$y(x_i)$	m	x	$[y_i - y(x_i)]^2 m$	$y_i - \bar{y}$	$(y_i - \bar{y})^2 m$
2,100	2,168	10	1	0,046	-0,92	8,464
2,933	2,802	15	2	0,256	-0,087	0,1125
3,200	3,326	15	3	0,237	+0,180	0,486
3,800	3,740	10	4	0,036	+0,780	6,084
Σ				0,575		15,1465

Остаточная дисперсия

$$S_{\text{ост}}^2 = \frac{\sum [(y_i - y(x_i))]^2 m_x}{n - m - 1} = \frac{0,575}{47} = 0,0122.$$

4 ОПТИМИЗАЦИОННЫЕ ЗАДАЧИ ЭЛЕКТРОСНАБЖЕНИЯ

4.1 Задачи электроснабжения, требующие поиска оптимальных решений

Задачи развития и функционирования электроэнергетики часто ставятся и разрешаются как математические экстремальные задачи при определенных ограничениях, т.е. как задачи исследования операций. Это обусловлено стремлением получать не просто любые решения реальных энергетических задач, а оптимальные. Возможность получения таких решений связана с тем, что в задачах обычно существует множество допустимых решений, называемых альтернативами. Иногда число альтернатив настолько велико, что без применения математических моделей и методов не представляется возможным найти оптимальный вариант или даже близкий к нему.

Приведем перечень энергетических задач, решения которых осуществляются в большей или меньшей степени с помощью математических методов.

1. Планирование развития генерирующих мощностей для покрытия графика электрических нагрузок экономического района.
2. Планирование развития электрических сетей в развивающейся электроэнергетической системе.
3. Оптимизация уровней токов короткого замыкания в электроэнергетической системе.
4. Оптимальное управление качеством электроэнергии в системах.

Кратко рассмотрим частную задачу планирования развития электрической сети (статическую задачу оптимизации сети) с тем, чтобы уяснить в первом приближении ее сущность, увидеть множество альтернатив и определить необходимость применения математических методов. Статическая задача оптимизации сети состоит в синтезе наилучшей в смысле приведенных затрат конфигурации сети, обеспечивающей необходимую надежность питания узлов нагрузки и требуемые параметры электрического режима.

Известно, что проектируемая сеть может иметь различное напряжение. Однако конкурентноспособные напряжения сети ограничиваются обычно двумя, реже тремя, уровнями, в зависимости от расстояний и величин передаваемых нагрузок. При этом число возможных линий планируемой сети, обеспечивающих связь узлов нагрузки

с центром питания и между собой, весьма велико. Линии могут иметь различные сечения, в узлах нагрузки могут устанавливаться компенсирующие устройства, главные понизительные подстанции узлов нагрузок могут иметь силовые трансформаторы, отличающиеся параметрами и конструкцией. Все это создает такое множество альтернатив, что даже при относительно небольшом числе узлов нагрузок простой их перебор и экономическая оценка оказываются иногда невыполнимыми. Следовательно, необходимы математические методы с их ограниченным и целенаправленным анализом вариантов, дающими возможность выбрать наилучший.

Существует ряд электроэнергетических задач меньшего масштаба, связанных с развитием, функционированием или техническим проектированием отдельных подсистем, в которых успешно применяются математические методы поиска оптимальных решений.

Сеть электроснабжения промышленных предприятий представляет собой сложную систему, состоящую из подсистем внешнего, внутризаводского и внутрицехового электроснабжения, взаимодействующую с питающей энергосистемой и технологической системой основного производства. Функционирование системы электроснабжения осуществляется на основе значительных потоков информации между этой системой и внешними системами, а также циркулирующих внутри нее. Поэтому синтез оптимальной системы электроснабжения промышленного предприятия представляет собой сложную задачу, строгого математического решения которой в настоящее время нет.

Между тем современные системы электроснабжения концентрируют значительные капитальные вложения, а доля электрозатрат в себестоимости основной продукции электроемких производств нередко превышает 20%. Таким образом, необходимо стремиться к выявлению наиболее эффективного варианта проектируемой системы электроснабжения, достичь которого можно на основе применения математических методов в рамках автоматизированной системы проектирования электроснабжения.

4.2 Понятие об управлении. Принципы исследования операций и основные понятия

Термин «исследование операций» появился в период второй мировой войны, когда для решения военно-технических задач создава-

лись смешанные группы ученых разных специальностей, и в настоящее время прочно вошел в научный лексикон. Под «операцией» следует понимать любую организованную, или точнее, организуемую деятельность. Именно эта деятельность в широком смысле слова и является предметом «исследованием операций». Конечная цель такого исследования – практические рекомендации, позволяющие руководителю более обоснованно принимать решения, т.е. рационально управлять. «Исследование операций» представляет собой использование научного метода для обеспечения исполнительных органов количественным обоснованием решений, принимаемых ими в процессе управления. Таким образом, целью исследования мыслится управление. А отсюда непосредственно вытекает, что исследование операций как научная дисциплина является не чем иным, как составной частью науки об управлении – кибернетики.

Исследование операций представляет собой *поиск сходных структур* в различных явлениях и применение к ним общих методов анализа. Недавно исследование операций было определено еще более четко как «научная подготовка решения». Дадим несколько определений «исследование операций», приведенных в различных работах:

- это методология научного поиска и вместе с тем инструмент для решения важных практических задач;
- это научный подход к решению задач организационного управления.

Операциями называют действия, в которых ответственный руководитель использует наличные людские и материальные ресурсы для достижения определенной цели в условиях положительного или отрицательного влияния случайных внешних факторов.

Исследование операций – это прикладное направление кибернетики, изучающее способы совершенствования и повышения эффективности организации, планирования и управления в различных системах на основе количественных методов.

Цель исследования операций – определить, как необходимо организовать работу системы (объекта, процесса), чтобы ее (его) функционирование обеспечивало максимальный эффект, выработать соответствующие научно обоснованные решения и рекомендации по оптимальному функционированию системы на основе использования практически всех существующих методов анализа и синтеза (оптимизации) систем и ЭВМ.

Система - это совокупность взаимосвязанных, взаимодействующих элементов (людей, машин, ...), выполняющих определенную задачу.

В настоящее время уже выделился определенный класс задач, эффективно решаемых в рамках исследования операций: распределительные; управление запасами; замены оборудования; упорядочения и согласования; выбора оптимальных режимов движения; массового обслуживания; состязательные; поиска и др.

Решение задач исследования операций предполагает выполнение следующих основных этапов:

- 1 Постановки задачи и выбора критериев оптимизации;
- 2 Выявление основных особенностей, взаимосвязей и количественных закономерностей исследуемой системы;
- 3 Построение математической модели системы (процесса);
- 4 Исследования математической модели с помощью специализированных алгоритмов и программ, поиска оптимальных параметров системы (процесса, операции).

4.3 Модели, применяемые для решения оптимизационных задач

Моделирование означает осуществление каким-либо способом отображения и воспроизведения действительности для изучения имеющихся в ней объективных закономерностей. Это – метод познания, при котором изучаемый объект – оригинал – находится в некотором соответствии с другим объектом – моделью.

Модель – это абстрактное описание того или иного объекта (явления) реального мира. Модель используется для получения такой информации, которую невозможно получить путем непосредственного исследования оригинала. Для этого модель должна быть не только сходной с оригиналом, но и отличаться от него, т.е. модель не тождественна оригиналу. Оригиналу может быть очень многообразен. Модель же должна отражать определенные свойства оригинала.

Существует два основных типа моделирования: физическое и математическое. При физическом моделировании изучаемого объекта выделяется какой-либо процесс и воспроизводится в другом масштабе. Но при этом возникают отличия от реального процесса из-за изменения масштаба и обрыва связей с другими процессами. Математиче-

ские модели – (аналоговые), основаны на формальном совпадении математических уравнений, описывающих процесс в оригинале и в модели. Содержанием математических моделей являются системы уравнений, описывающих процессы в объектах. Для построения математической модели используют формальные математические категории: уравнения, неравенства, области определения переменных, вероятности и другие. Одни и те же уравнения могут описывать различные процессы.

Математические модели создаются для исследования явлений реального мира с тем, чтобы проникнуть в их сущность, изучить влияние различных факторов на их протекание и в конечном итоге управлять ими. Это положение в равной мере относится к математическим моделям, отражающим энергетические операции или операции, протекающие в любых технических, экономических, биологических или социальных системах.

Целевая направленность математических моделей отражается в их структуре. Остановимся на таких компонентах структуры модели, как целевая функция (критериальная функция, функция цели) и ограничения. Обычно достижение цели связывается с различными планами действий (альтернативами). Выделение лучшей из альтернатив осуществляется с помощью некоторого критерия, способного сопоставить различные планы действий по степени их предпочтения. Критерий, выраженный через управляемые переменные модели, и представляют собой критериальную целевую функцию.

Однако поиск наилучшего варианта действий на основе принятого критерия обычно лимитирован рядом обстоятельств, называемых ограничениями. Они касаются ресурсов в широком смысле этого слова и базируются на концепции ограниченности средств достижения цели, которая сводится к признанию дефицитности трудовых, материально-технических, природных, финансовых, информационных и других ресурсов. Ограничения реализуются в моделях в форме уравнений и неравенств, выраженных через управляемые переменные.

4.4 Классификация методов оптимизации

Классификация моделей математического программирования и соответствующих методов их анализа может быть произведена по различным признакам. Укажем на некоторые из них:

по временному признаку (статические модели и модели динамические);

по порядку математических соотношений, соответствующих целевой функции и ограничениям (линейные и нелинейные модели);

по признаку дифференцируемости целевой функции и функциональных ограничений;

по типу управляемых переменных (модели с целочисленными, дискретными, булевыми переменными и переменными непрерывными).

Общую задачу математического программирования можно сформулировать, например, так: определить вектор $X = (x_1; x_2; \dots, x_n)^t$, который является решением задачи

$$\begin{aligned} Z(X) &\rightarrow \min (\max); \\ g_i(X) &\leq (\geq) 0, \quad i = 1, 2, \dots, m; \\ h_j(X) &= 0, \quad j = 1, 2, \dots, n, \end{aligned} \quad (4.1)$$

где t – индекс транспонирования.

В модели (4.1) первое соотношение представляет собой критериальную функцию, минимизируемую (максимизируемую) в данном случае, а следующие соотношения – ограничения, которые формируют область допустимых решений. Эта область представляет собой множество в n – мерном евклидовом пространстве E^n .

Если функции $Z(X)$, $g_i(X)$, $h_j(X)$ линейны, то модель (4.1) называется линейной, а задача, которую она отражает, - задачей линейного программирования. Линейные целевые функции являются наиболее “простыми” функциями, поэтому анализ моделей линейного программирования в известной мере проще, чем другие модели математического программирования, а методы линейного программирования достаточно глубоко разработаны.

Если хотя бы одна из этих функций не линейна, то модель (4.1) называется нелинейной, а задача, которая представлена в ней, - задачей нелинейного программирования. Методы анализа таких задач называют методами нелинейного программирования.

Задачи нелинейного программирования обычно принято подразделять на задачи выпуклого программирования и невыпуклого, или многоэкстремальные. Эти же определения можно распространить и на соответствующие модели.

Задача (4.1) будет называться задачей выпуклого программирования, если выпукла целевая функция и выпукло множество, задаваемое ограничениями. Для анализа моделей выпуклого программирования

ния разработан мощный аппарат выпуклого программирования. К задачам выпуклого программирования принадлежат и задачи квадратичного программирования, т.е. такие, у которых целевая функция представляет сумму линейной и квадратичной форм относительно управляемых переменных

$$F(x) = a + \sum_{i=1}^n c_i x_i + \sum_{i=1, t=1}^n d_{i,t} x_i x_t, \quad (4.2)$$

а ограничения – линейны.

Многоэкстремальные задачи порождаются в случаях, если целевая функция либо гладкая, но невыпуклая (содержащая локальные оптимумы), либо при гладкой и выпуклой критериальной функции оптимизация рассматривается на невыпуклом множестве.

Наконец, задачи линейного и нелинейного программирования, а также многоэкстремальные задачи могут быть дискретными или непрерывными в зависимости от того, накладываются или нет условия целочисленности на все или некоторые управляемые переменные.

Таким образом, выраженные через управляемые переменные целевая функция и ограничения представляют собой математическую модель явления или системы. Математические модели, если они адекватны реальной действительности, позволяют производить численные эксперименты и на их основе строить суждения о реакции моделируемой системы на различные воздействия.

В задачах математического программирования различают оптимумы локальные и глобальные.

Локальным минимумом (максимумом) называется точка, в окрестности которой нет других точек, удовлетворяющих ограничениям задачи и доставляющих целевой функции меньшие (большие) значения. Если непрерывная функция на замкнутом ограниченном множестве в евклидовом пространстве имеет несколько локальных минимумов (максимумов), то наименьший (наибольший) из них называется глобальным минимумом (максимумом).

При определенных условиях локальные оптимумы могут совпадать с глобальными. Это отмечается для выпуклых и вогнутых функций, рассматриваемых на выпуклых множествах в евклидовом пространстве.

Метод динамического программирования предназначен для анализа моделей, содержащих целевую функцию специальной структуры. Речь идет о сепарабельной целевой функции, которая может быть представлена в виде суммы или произведения n функций одной переменной, т.е.

$$F(x) = F(x_1, x_2, \dots, x_n) = \sum_{i=1}^n f_i(x_i) \quad (4.3)$$

или

$$F(x) = F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_i(x_i). \quad (4.4)$$

В случае (4.3) функция $F(x)$ называется аддитивной, а в случае (4.4) – мультипликативной.

Метод динамического программирования применим для анализа многошаговых процессов и является средством анализа многоэкстремальных задач.

Метод прямого перебора. Если известна функциональная связь целевой функции Y и искомой переменной X , то можно последовательно вычислить значения целевой функции для некоторых значений искомой переменной. Вычисления повторяются до тех пор, пока не будет найден \min (\max) значения целевой функции:

$$Y = f(x_1, \dots, x_i, \dots, x_n, u_1, \dots, u_j, \dots, u_m), \\ x_i = x_{0i} + \Delta x_i k \quad (k = 0, 1, 2, \dots, l).$$

Этот метод может быть использован для решения задач исследования операций, если имеется одна искомая переменная или несколько с небольшим диапазоном изменения искомым переменных.

Классический метод дифференциального исчисления. Если известна функциональная связь целевой функции с искомыми переменными $Y = f(x_1, \dots, x_i, \dots, x_n, u_1, \dots, u_j, \dots, u_m)$, которая обладает непрерывными первыми частными производными, то, определив частные производные по своим искомым переменным, приравняв частные производные от Y по x_i к нулю и решив систему уравнений

$$\frac{\partial Y}{\partial x_1} = 0,$$

.....

$$\frac{\partial Y}{\partial x_n} = 0,$$

найдем значения x_{ic} , дающие стационарные значения целевой функции, среди которых находятся оптимальные.

Метод неопределенных множителей Лагранжа. С помощью этого метода можно определить экстремальные точки функции многих переменных при наличии дополнительных связей между оптимизируемыми параметрами.

Пусть имеется целевая функция $Y = f(x_1, \dots, x_i, \dots, x_n, u_1, \dots, u_j, \dots, u_m)$, экстремум которой требуется определить, причем существуют дополнительные условия

$$\varphi_k(x_1, \dots, x_i, \dots, x_n, u_1, \dots, u_j, \dots, u_m) = 0 \quad (k = 1, 2, \dots, p).$$

Введя p дополнительных множителей $\lambda_1, \lambda_2, \dots, \lambda_k, \dots, \lambda_p$, построим новую функцию

$$L(x_1, \dots, x_i, \dots, x_n, u_1, \dots, u_j, \dots, u_m, \lambda_1, \lambda_2, \dots, \lambda_k, \dots, \lambda_p) = \\ = f(x_1, \dots, x_i, \dots, x_n, u_1, \dots, u_j, \dots, u_m) - \sum_{k=1}^p \lambda_k * \varphi_k(x_1, \dots, x_n, u_1, \dots, u_m),$$

где λ_k - множитель Лагранжа.

Необходимые условия экстремума состоят в равенстве нулю всех первых частных производных от L по $x_1, \dots, x_i, \dots, x_n, \lambda_1, \dots, \lambda_k, \dots, \lambda_p$:

$$\frac{\partial L}{\partial x_i} = 0 \quad (i = 1, 2, \dots, n);$$

$$\frac{\partial L}{\partial \lambda_k} = 0 \quad (k = 1, 2, \dots, p).$$

В результате получается $n+p$ уравнений с $n+p$ неизвестными $((x_1, \dots, x_i, \dots, x_n), (\lambda_1, \dots, \lambda_k, \dots, \lambda_p))$. Решение этих уравнений относительно переменных x и λ дает возможность определить положение стационарной точки. Использование вспомогательной функции $L(x, \lambda)$ позволяет заменить задачу с дополнительными условиями задачей без них.

5 ЗАДАЧИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Понятие линейного программирования. Линейное программирование – раздел математического программирования, применяемый при разработке методов отыскания экстремума линейных функций нескольких переменных при линейных дополнительных ограничениях, налагаемых на переменные.

По типу решаемых задач его методы разделяются на универсальные и специальные. С помощью универсальных методов могут решаться любые задачи линейного программирования (ЗЛП). Специальные методы учитывают особенности модели задачи, её целевой функции и системы ограничений.

Особенностью задач линейного программирования является то, что экстремума целевая функция достигает на границе области допустимых решений. Классические же методы дифференциального исчисления связаны с нахождением экстремумов функции во внутренней точке области допустимых значений.

Методами ЗЛП могут решаться задачи: о наилучшем использовании ресурсов; о выборе оптимальных технологий; о размещении ремонтов специализированного оборудования; транспортная задача

(задача развития электрических сетей); о размещении заказа и т.д. В перечисленных задачах может определяться или максимум или минимум целевой функции.

Задача развития электрических сетей (транспортная задача). Модель транспортной задачи – транспортировка некоторого однородного продукта (электроэнергии) от производителей к потребителям, при этом должен иметься баланс между суммарным спросом потребителей и возможностями поставщиков по их удовлетворению. Причем потребителям безразлично, из каких пунктов производства будет поступать электроэнергия, лишь бы их заявки были полностью удовлетворены. Так как от схемы прикрепления потребителей к поставщикам существенно зависит объем передаваемой энергии, возникает задача о наиболее рациональной схеме электрических сетей (рациональное напряжение, количестве и сечений линий, трансформаций и др.), при котором потребности полностью удовлетворяются, все производство электроэнергии технологически обеспечивается, а затраты на проектирование сетей и передачу энергии потребителям минимальны.

Задача формулируется так. Имеется m пунктов производства, в каждом из которых может быть произведено b_i ($i = 1, m$) электроэнергии и n пунктов потребления электроэнергии, где потребность составляет a_j ($j = 1, n$) единиц. Известны величины c_{ij} – затраты на транспортировку на единицу электроэнергии из i – го пункта в j – й пункт потребления. Обозначим через x_{ij} количество электроэнергии, передаваемого из i – го пункта производства в j – й пункт потребления. Матрица $[c_{ij}]$ – называется матрицей тарифов, $X = [x_{ij}]$ – матрицей транспортировки. С целью удобства построения математической модели матрицы тарифов и транспортировки совмещают в одну, именуемую макетом транспортной задачи (развития сетей) (табл.5.1).

Таблица 5.1

b_i	a_j							
	a_1	a_2	a_3	...	a_j	...	a_{n-1}	a_n
b_1	C_{11} x_{11}	C_{12} x_{12}	C_{13} x_{13}	...	C_{1j} x_{1j}	...	$C_{1(n-1)}$ $x_{1(n-1)}$	C_{1n} x_{1n}
b_2	C_{21} x_{21}	C_{22} x_{22}	C_{23} x_{23}	...	C_{2j} x_{2j}	...	$C_{2(n-1)}$ $x_{2(n-1)}$	C_{2n} x_{2n}
...
b_i	C_{i1} x_{i1}	C_{i2} x_{i2}	C_{i3} x_{i3}	...	C_{ij} x_{ij}	...	$C_{i(n-1)}$ $x_{i(n-1)}$	C_{in} x_{in}
...
b_m	C_{m1} x_{m1}	C_{m2} x_{m2}	C_{m3} x_{m3}	...	C_{mj} x_{mj}	...	$C_{m(n-1)}$ $x_{m(n-1)}$	C_{mn} x_{mn}

Математическая модель задачи развития электрической сети: целевая функция, описывающая транспортные затраты

$$Z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} \cdot x_{ij} \rightarrow \min \quad (5.1)$$

при ограничениях: на возможности производителя – весь продукт из пунктов производства может быть транспортирован

$$\sum_{i=1}^m x_{ij} \leq b_i \quad (i=1, \dots, m) \quad (5.2)$$

на спрос потребителей, который должен быть полностью удовлетворен:

$$\sum_{j=1}^n x_{ij} = a_i \quad (j=1, \dots, n) \quad (5.3)$$

при условии не отрицательности переменных, исключающем обратные перетоки

$$x_{ij} \geq 0 \quad (i=1, \dots, m; j=1, \dots, n) \quad (5.4)$$

Основные виды записи ЗЛП.

Общей задачей линейного программирования (ОЗЛП) называют задачу, которую математически представляют в виде следующих линейных соотношений:

– целевая функция – $W = \sum_{j=1}^n c_j x_j \rightarrow \max(\min) \quad (5.5)$

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq b_i & (i=1, \dots, m_1) \\ \sum_{j=1}^n a_{ij} x_j = b_i & (i=m_1+1, \dots, m_2) \\ \sum_{j=1}^n a_{ij} x_j \geq b_i & (i=m_2+1, \dots, m) \end{cases} \quad (5.6)$$

где $-C_j, a_{ij}, b_i$ – заданные действительные числа; $x_j \geq 0$ ($j = 1, \dots, n$) – искомые неотрицательные переменные параметры, удовлетворяющие условию задачи и ограничениям; $X = (X_1; \dots; X_n)^t$ – план задачи.

Канонической формой записи ЗЛП называют задачу

$$\min W = \sum_{j=1}^n c_j \cdot x_j \quad (5.7)$$

$$\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = (1, 2, \dots, m) \quad (5.8)$$

$$x_j \geq 0 \quad (j = 1, \dots, n). \quad (5.9)$$

Каноническая форма записи ЗЛП может быть представлена в матричной форме и в векторной форме. Введем обозначения:

$$C = [c_1, c_2, \dots, c_n],$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad A_0 = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix},$$

где C – матрица-строка; A – матрица системы уравнений; X – матрица – столбец переменных; A_0 – матрица-столбец свободных членов.

Каноническая форма задачи с использованием матричной записи примет вид

$$\min W = [c_1, c_2, \dots, c_n] \cdot [x_1 \quad x_2 \quad \dots \quad x_n]^T$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}, \quad X \geq 0$$

или $\min W = CX, \quad AX = A_0, \quad X \geq 0.$

Векторная форма записи ЗЛП.

Для столбцов матрицы А введем обозначения:

$$A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \quad A_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ \dots \\ a_{m3} \end{pmatrix}, \dots, A_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix}, \dots, A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}.$$

Тогда задача (5.7) – (5.9) в векторной форме записи примет вид:

$$\min W = cx$$

$$A_1x_1 + A_2x_2 + \dots + A_jx_j + \dots + A_nx_n = A_0, \quad x \geq 0,$$

где cx – скалярное произведение векторов $c=(c_1;c_2;\dots;c_n)$ и $x=(x_1; x_2;\dots; x_n)$.

Геометрическая интерпретация задачи линейного планирования

Геометрическая интерпретация задачи позволяет наглядно определить структуру, выявить особенности и открыть пути исследования более сложных свойств. Задачи линейного планирования с двумя переменными можно решить графически. В трехмерном пространстве графическое решение усложняется, а в n – мерном пространстве графическое решение невозможно.

Случай двух переменных не имеет особого практического значения, однако его рассмотрение проясняет свойства общей задачи линейного планирования, приводит к идее ее решения, делает геометрически наглядными способы решения и пути их практической реализации.

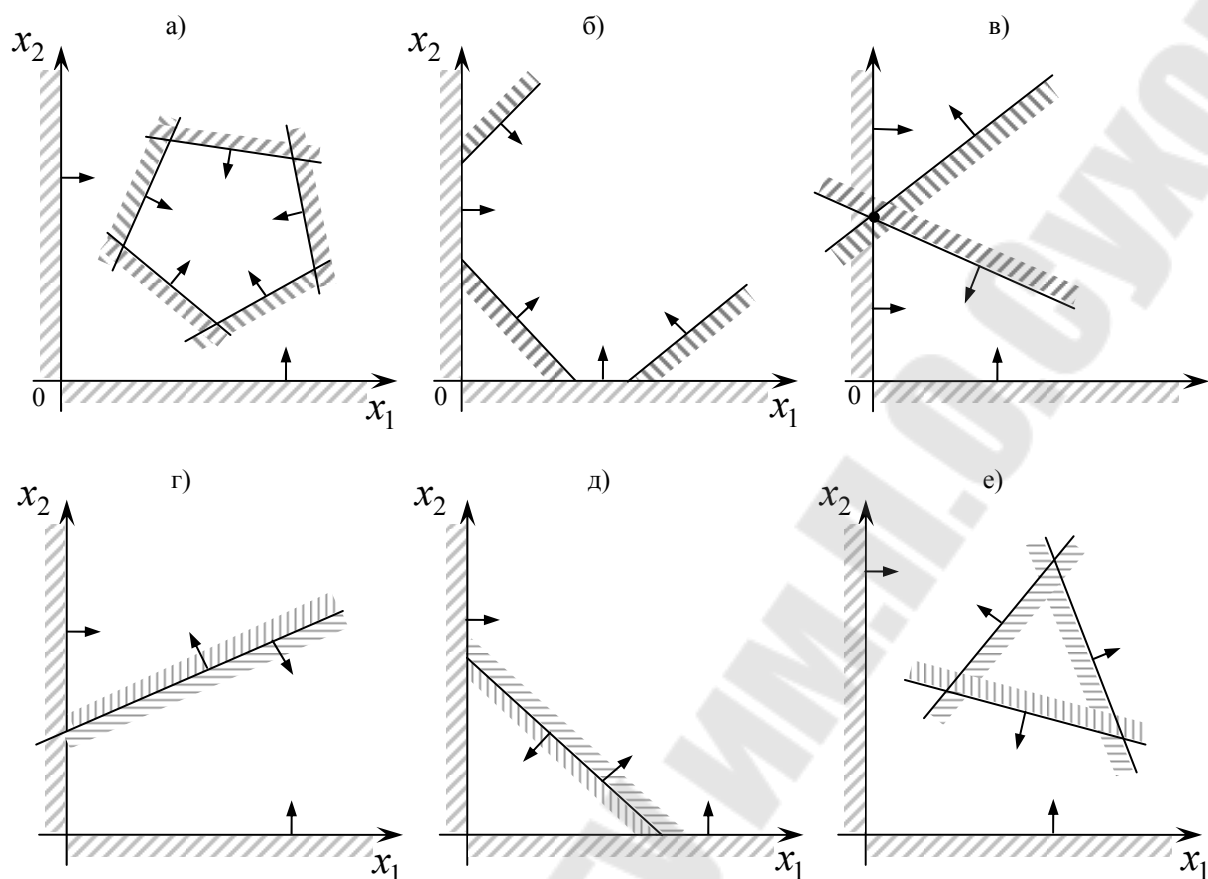


Рис.5.1

Перейдем к геометрической интерпретации целевой функции. Пусть область допустимых решений ЗЛП – непустое множество, например многоугольник ABCDE (рис.5.2). Выберем произвольное значение целевой функции $Z = Z_0$. Получим $c_0 + c_1x_1 + c_2x_2 = Z_0$. Это уравнение прямой линии. В точках прямой MN целевая функция сохраняет одно и то же постоянное значение Z_0 . Считая в равенстве (5.10) Z параметром, получим уравнение семейства параллельных прямых, называемых *линиями уровня целевой функции (линиями постоянного значения)*.

Возникает вопрос: как установить направление возрастания (убывания) целевой функции? Найдём частные производные целевой функции по x_1 и x_2 :

$$\frac{\partial Z}{\partial x_1} = c_1; \quad \frac{\partial Z}{\partial x_2} = c_2. \quad (5.12)$$

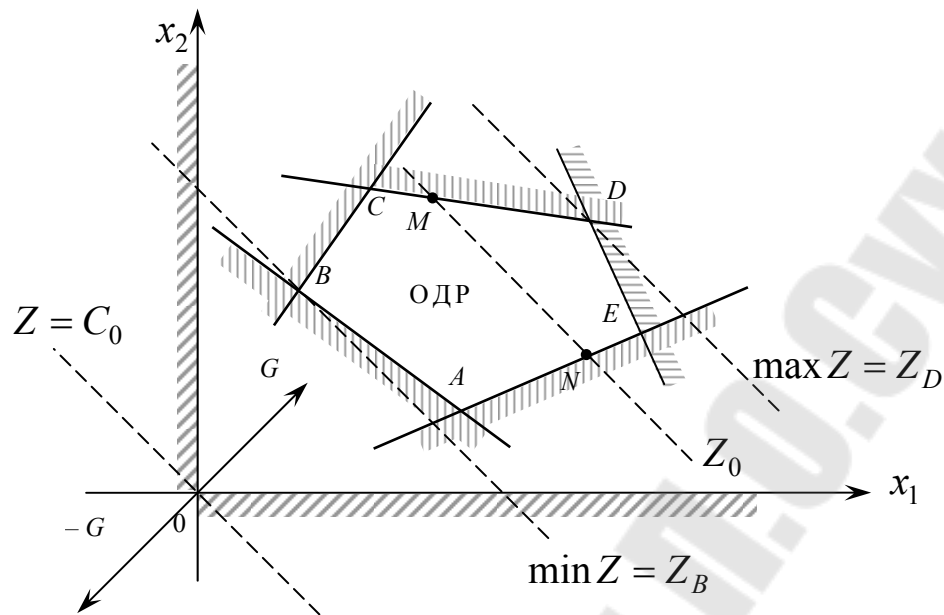


Рис 5.2

Частные производные (5.12) функции показывают скорость ее возрастания вдоль осей. Следовательно, c_1 и c_2 – скорость возрастания Z соответственно вдоль осей Ox_1 и Ox_2 . Вектор $G = (c_1; c_2)$ называется *градиентом функции*. Он показывает направление наискорейшего возрастания целевой функции:

$$G = (\partial Z / \partial x_1, \partial Z / \partial x_2).$$

Вектор $-G$ показывает направление наискорейшего убывания целевой функции. Его называют *антиградиентом*.

Вектор $G = (c_1, c_2)$ перпендикулярен к прямым $Z = \text{const}$ семейства $c_0 + c_1x_1 + c_2x_2 = Z$.

Из геометрической интерпретации элементов ЗЛП следует порядок её графического решения.

1. С учетом системы ограничений строим область допустимых решений.

2. Строим вектор $G = (c_1, c_2)$ наискорейшего возрастания целевой функции – вектор градиентного направления.

3. Проводим произвольную линию уровня $Z = Z_0$ (проще всего провести линию $Z = c_0$, перпендикулярную к вектору G).

4. При решении задачи на максимум перемещаем линию уровня $Z = Z_0$ в направлении вектора G так, чтобы она касалась ОДР в её крайнем положении (крайней точке) (на рис.5.2 – до точки Д). В слу-

чае решения задачи на минимум линию уровня $Z = Z_0$ перемещаем в направлении антиградиента (на рис. 5.2 – до точки В).

5. Определяем оптимальное решение, т.е. значения $X^* = (x_1^*; x_2^*)$ и экстремальное значение целевой функции $Z^* = z(x^*)$.

Из геометрической интерпретации условий ограничений и целевой функции возможно выявить особенности решения ЗЛП, а именно (см. рис.5.3):

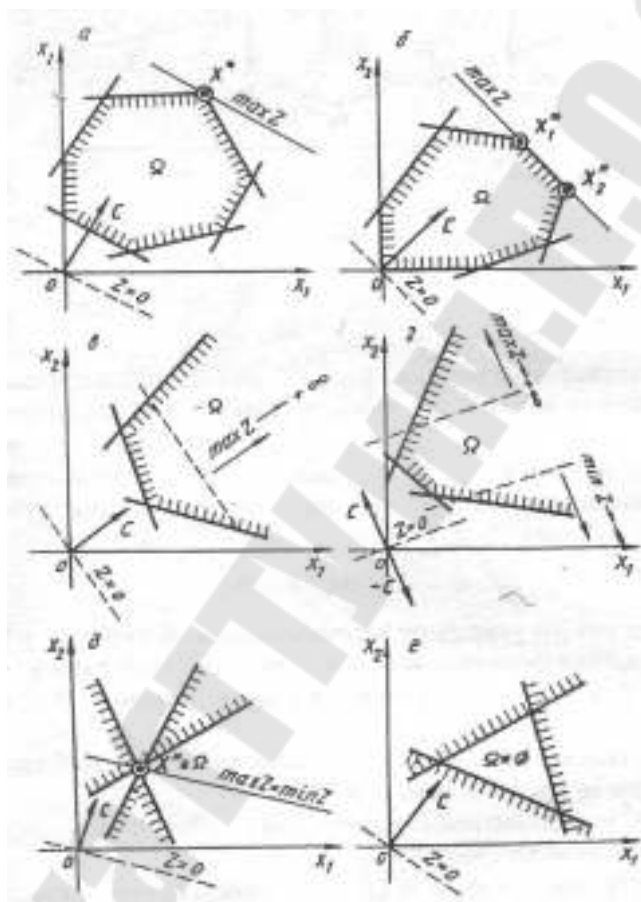


Рис.5.3

1) – оптимальное решение единственное: линия уровня и область допустимых решений в соответствующем положении имеют одну общую точку (рис.5.3, а);

2) – оптимальных решений бесконечное множество: в соответствующем положении линия уровня проходит через сторону области допустимых решений (рис.5.3, б);

3) – целевая функция не ограничена: линия уровня, сколько бы ее ни перемещали, не может занять соответствующего положения (рис.5.3, в,г);

4) – область допустимых решений состоит из единственной точки, где целевая функция достигает одновременно и максимального, и минимального значений (рис. 5.3, д);

5) – задача не имеет решения: область допустимых решений – пустое множество, т.е. система ограничений несовместна (рис. 5.3 е).

При количестве переменных более трех ЗЛП теряет геометрическую наглядность, но идея получения решения (графического) сохраняет смысл и для случая многомерного пространства.

Переход к канонической форме. Как следует из общей задачи линейного планирования, в которой для целевой функции может определяться либо максимум, либо минимум, в них большинство ограничений задаются неравенствами. Наиболее же широко используемые методы решения задач линейного планирования применяются лишь к задачам, записанным в канонической форме. Поэтому приходится переходить от любой формы ЗЛП к ее каноническому виду, причем нужно быть уверенным, что эти формы эквивалентны.

Пусть исходная ЗЛП имеет вид

$$\max Z = C_0 + \sum_1^n C_j x_j \quad (5.13)$$

$$\sum_{i,j=1}^{m_1,n} a_{ij} x_j \leq b_i \quad (i = 1, m_1) \quad (5.14)$$

$$\sum_{i,j=1}^{m,n} a_{ij} x_j \geq b_i \quad (i = m_1 + 1, m) \quad (5.15)$$

$$x_j \geq 0 \quad (j = 1, n) \quad (5.16)$$

Канонической формой записи ЗЛП называют задачу

$$Z = c_0 - \left(\sum_1^n c_j x_j \right) \rightarrow \min \quad (5.17)$$

$$\sum_{i,j=1}^{m,n} a_{ij} x_j = b_i \quad (5.18)$$

$$x_j \geq 0 \quad (j = 1, \dots, n). \quad (5.19)$$

При необходимости задачу максимизации можно заменить задачей минимизации, и наоборот. Для функции одной переменной это утверждение очевидно. В самом деле, если x^* – точка максимума функции $y = f(x)$, то для функции $y = -f(x)$ она является точкой минимума, так как графики функции $f(x)$ и $-f(x)$ – симметричны относительно оси абсцисс (рис. 5.4).

Итак, $\max f(x^*) = -\min(-f(x^*))$.

То же самое имеет место и в случае функции многих переменных (n):

$$\max f(x_1^*, \dots, x_n^*) = -\min(-f(x_1^*, \dots, x_n^*)).$$

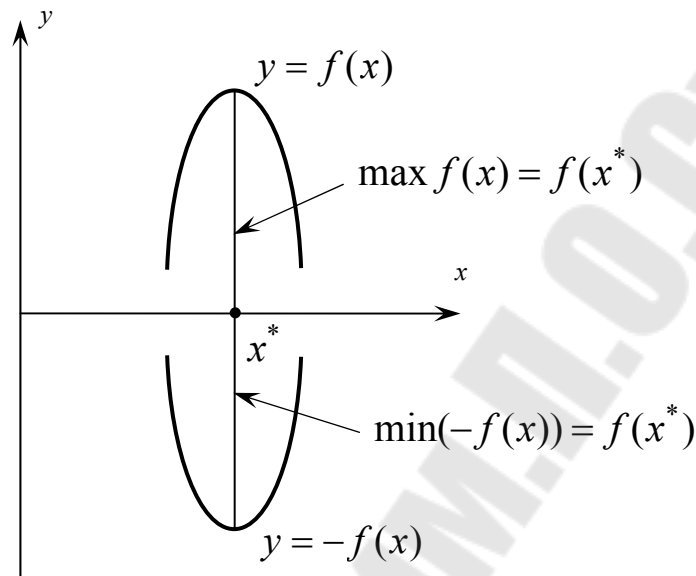


Рис. 5.4

Преобразуем систему ограничений к каноническому виду. Введем m дополнительных неотрицательных переменных $x_{n+i} \geq 0$ ($i = 1, m$). Для того чтобы неравенства типа \leq (5.14) преобразовались в равенства, к их левым частям прибавим дополнительные переменные $x_{n+i} \geq 0$ ($i = 1, m_1$), после чего система неравенств (5.14) примет вид:

$$\sum_{i,j}^{m_1;n} a_{ij}x_n + x_{n+i} = b_i. \quad (5.20)$$

Для того чтобы неравенства типа \geq (5.15) преобразовать в равенства, из их левых частей вычтем дополнительные переменные $x_{n+i} \geq 0$ ($i = m_1 + 1, m$). Получим

$$\sum_{m_1+1,1}^{m,n} a_{ij}x_n - x_{n+i} = b_i \quad (5.21)$$

Дополнительные переменные x_{n+i} ($i = 1, m$) в целевую функцию вводятся с коэффициентами, равными нулю. Целевую функцию умножаем на минус единицу. Получим задачу:

$$\min(-Z) = -\left(c_0 + \sum_1^n c_j x_j + \sum_n^1 0 \cdot x_{n+i} \right) \quad (5.22)$$

$$\sum_{i,j=1}^{m_1;n} a_{ij}x_n + x_{n+i} = b_i \quad (5.23)$$

$$\sum_{i=m_1+1, j=1}^{m,n} a_{ij}x_n - x_{n+i} = b_i \quad (5.24)$$

$$x_j \geq 0 \quad (j = 1, n) \quad x_{n+i} \geq 0 \quad (i = 1, m) \quad (5.25)$$

Задача (5.22) – (5.25) имеет каноническую форму. Задачи (5.13) – (5.16) и (5.22) – (5.25) тесно связаны между собой.

Симплексный метод

Для решения ЗЛП в аналитической форме применяют один из методов решения – симплекс метод (метод последовательного улучшения решения), который предполагает: 1) умение находить начальный опорный план; 2) наличие признака оптимальности плана; 3) умение переходить к не худшему опорному плану.

Для построение начального опорного плана исходные данные задачи должны быть представлены в канонической форме. Пусть ЗЛП представлена системой ограничений в каноническом виде:

$$\sum_{j=1}^n a_{ij} \cdot x_j = b_i, \quad b_i \geq 0 \quad (i = 1, m).$$

Говорят, что ограничение ЗЛП имеет предпочтительный вид, если при не отрицательности правой части ($b_i \geq 0$) левая часть ограничения содержит переменную, входящую с коэффициентом, равным единице, а в остальные ограничения – равенства – с коэффициентом, равным нулю. Например, в системе ограничений

$$\begin{aligned} \underline{x_1} + 2x_2 - 4x_4 &= 5, \\ 2x_2 + \underline{x_3} + 2x_4 &= 8, \\ x_2 - 3x_4 &= 3 \end{aligned}$$

первое и второе ограничения имеют предпочтительный вид, третье – нет.

Если каждое ограничение-равенство ЗЛП в каноническом виде содержит переменную, входящую в левую часть с коэффициентом, равным единице, а во все остальные с коэффициентом, равным нулю (при не отрицательности правых частей), то говорят, что система ограничений представлена в предпочтительном виде. В этом случае легко найти ее опорное решение (базисное с не отрицательными координатами): все свободные переменные нужно приравнять нулю, тогда базисные переменные (БП) будут равны свободным членам. Если

полученный план будет иметь не более m отличных от нуля координат, то он будет опорным.

Приравнивание предпочтительных переменных к правым частям дает базисное решение, т.е. крайнюю точку многоугольника. Поэтому предпочтительные – базисные. Переменные, приравниваемые нулю, – свободные.

Пусть система ограничений имеет вид

$$\sum_{j=1}^n a_{ij} \cdot x_j \leq b_i, \quad b_i \geq 0 \quad (i = 1, m).$$

Приведем систему к каноническому виду

$$\sum_{j=1}^n a_{ij} \cdot x_j + x_{n+i} = b_i, \quad b_i \geq 0 \quad (i = 1, m),$$

которая имеет предпочтительный вид.

Следовательно, начальный опорный план примет вид: $X_0 = (0; 0; \dots; 0; b_1; b_2; \dots; b_m)$. В целевую функцию дополнительные переменные вводятся с коэффициентами, равными нулю.

Пусть система ограничений имеет вид

$$\sum_{j=1}^n a_{ij} \cdot x_j \geq b_i, \quad b_i \geq 0 \quad (i = 1, \dots, m).$$

Приведем систему к каноническому виду

$$\sum_{j=1}^n a_{ij} \cdot x_j - x_{n+i} = b_i, \quad b_i \geq 0 \quad (i = 1, \dots, m),$$

которая не имеет предпочтительный вид, так как дополнительные переменные x_{n+i} входят в левую часть (при $b_i \geq 0$) с коэффициентом -1 . Поэтому, базисный план $X_0 = (0; 0; \dots; 0; -b_1; -b_2; \dots; -b_m)$ является недопустимым. В практических задачах всегда имеются условия ограничений, которые в каноническом виде преобразуются из соотношений (5.18), поэтому можно перейти к улучшенному решению.

Приведем последовательность шагов при решении задачи оптимизации: шаг первый: Приводим задачу к канонической форме

$$Z = c_0 - \left(\sum_{j=1}^n c_j \cdot x_j \right) \rightarrow \min;$$

$$x_i = b_i - \sum a_{ij} \cdot x_j - \text{БП};$$

$$x_j \geq 0 - \text{свободные переменные.}$$

шаг второй: Находим начальный опорный план, приравняв $x_j = 0$, тогда $Z = c_0$, $x_i = [b_i]$.

шаг третий: Проверяем, оптимально ли найденное решение. Если решение оптимально, то вычисления окончены, в противном случае необходим переход к следующему опорному плану.

Выполнение этого шага требует анализа коэффициентов в функции цели при свободных переменных и проверки не отрицательности значений БП. Если значения БП удовлетворяют условию не отрицательности, то решение одно из допустимых (т.е. это одна из вершин ОДР). На оптимальность решения будут указывать отрицательные значения коэффициентов c_j в целевой функции. В случае наличия положительных коэффициентов c_j решение может быть улучшено.

шаг четвертый: Переход к новой вершине ОДР, в которой значение целевой функции меньше начального. Наибольшее из положительных коэффициентов c_j будет указывать на ту свободную переменную x_j , которая может изменяться от нуля до значений, не нарушающих условия ограничений.

Приравнивают БП к нулю для всех ограничений на данном шаге. Находят отношения $\left\{ \frac{b_i}{a_{ij}} \right\}$ для j – ой свободной переменной при условии равенства нулю других свободных переменных ($a_{ij} > 0$).

Наименьшее из положительных $\left\{ \frac{b_i}{a_{ij}} \right\}$ укажет ту БП, которая станет

свободной, т.е. укажет уравнение ограничений, из которого определится новая БП, выраженная через новые свободные переменные. Все остальные БП и функцию цели нужно пересчитать через новые свободные переменные. Т. о. будет получен очередной опорный план. Далее повторение шагов 3 и 4. Выполнение шагов один – четыре предполагает последовательные преобразования условий ограничений и функции цели. Поэтому, в дальнейшем будем говорить о решениях, полученных методом симплекс – преобразований.

Все преобразования симплекс – метода можно выполнить в табличной форме. **Табличная форма** симплекс – метода предполагает представление исходной задачи в канонической форме (см. выше). Для внесения исходных данных в таблицу запишем условие задачи следующим образом:

$$Z = c_0 - \left(\sum_{j=1}^n c_j \cdot x_j \right) \rightarrow \min;$$

$$x_i = b_i - \left(\sum_{j=1}^n a_{ij} \cdot x_j \right) - \text{БП};$$

$x_j \geq 0$ – свободные переменные.

Коэффициенты при свободных переменных в уравнениях БП и функции цели внесем в таблицу, причем разместим их в верхние части клеток со знаками, указанными в скобках соответствующих уравнений.

Таблица 5.2

БП	Сво- бодный член b_i	Свободные переменные					
		x_1	x_2	...	x_j	...	x_n
x_{n+1}	b_1 $b_i \cdot (-\lambda a_{1j})$	a_{11} $a_{i1} \cdot (-\lambda a_{1j})$	a_{12} $a_{i2} \cdot (-\lambda a_{1j})$...	a_{1j} $-\lambda \cdot a_{1j}$...	a_{1n} $-a_{in} \lambda a_{1j}$
x_{n+2}	b_2 $b_i \cdot (-\lambda a_{2j})$	a_{21} $a_{i1} \cdot (-\lambda a_{2j})$	a_{22} $a_{i2} \cdot (-\lambda a_{2j})$...	a_{2j} $-\lambda \cdot a_{2j}$...	a_{2n} $-a_{in} \lambda a_{2j}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_i	b_i $\lambda \cdot b_i$	a_{i1} $\lambda \cdot a_{i1}$	a_{i2} $\lambda \cdot a_{i2}$...	a_{ij} $\lambda = \frac{1}{a_{ij}}$...	a_{in} $\lambda \cdot a_{in}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
x_m	b_m $b_i \cdot (-\lambda a_{mj})$	a_{m1} $a_{i1} \cdot (-\lambda a_{mj})$	a_{m2} $-a_{i2} \lambda a_{mj}$	⋮	a_{mj} $-\lambda \cdot a_{mj}$...	a_{mn} $-a_{in} \lambda a_{mj}$
Z_1	c_0 $b_i \cdot (-\lambda c_j)$	c_1 $a_{i1} \cdot (-\lambda c_j)$	c_2 $-a_{i2} \lambda c_j$...	c_j $-\lambda \cdot c_j$...	c_n $-a_{in} \lambda c_j$

Столбец «свободные члены» определяет первое начальное решение при равенстве нулю свободных переменных:

$$Z_1(x_j, x_i) = c_0 \quad X \begin{bmatrix} 0 & 0 & \dots & 0 & b_1 & b_2 & \dots & b_m \end{bmatrix}.$$

Далее выполняем шаги 3 и 4 симплекс - преобразований. Наибольший коэффициент при свободных переменных в функции цели определяет разрешающий столбец (пусть это будет x_j). Находим наименьшее

положительное отношение $\left\{ \frac{b_i}{a_{ij}} \right\}$, которое определяет разрешающую

строку. Элемент, стоящий на пересечении разрешающего столбца и разрешающей строки, называют генеральным (обведем его кружком). Разрешающая строка показывает: какая базисная переменная поменяется со свободной переменной $x_i \leftrightarrow x_j$.

Для пересчета коэффициентов базисных переменных и функции цели через новые свободные переменные выполним следующее:

- 1) находим $\lambda = 1/a_{ij}$; a_{ij} – генеральный элемент;
- 2) все коэффициенты разрешающей строки умножим на λ (кроме генерального), а коэффициенты разрешающего столбца - на "- λ " и запишем в нижней части клеток;
- 3) выделим старые значения коэффициентов разрешающей строки (\downarrow) и новые значения коэффициентов разрешающего столбца (\uparrow);
- 4) числа, вводимые в нижнюю часть клетки на пересечении строки l и столбца S находим перемножением старого значения коэффициентов разрешающей строки и нового значения коэффициентов разрешающего столбца.

После заполнения, всех клеток таблицы осуществляют ее преобразование в новую таблицу:

Во все верхние отделения клеток разрешающей строки и разрешающего столбца заполняются значения из нижних отделений предыдущей таблицы; в остальные клетки помещают алгебраическую сумму значений данной клетки.

Анализируя полученную таблицу, видим, что решение «улучшено», т.е. $Z_1 > Z_2$. Далее выполняем действия, аналогичные вышеописанным.

Таблица 5.3

Базисные переменные	Свободный член b_i	Свободные переменные					
		x_1	x_2	...	x_j	...	x_n
x_{n+1}	b'_1	a'_{11}	a'_{12}	...	$-\lambda a_{1j}$...	a'_{1n}
x_{n+2}	b'_2	a'_{21}	a'_{22}	...	$-\lambda a_{2j}$...	a'_{2n}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_j	λb_i	λa_{i1}	λa_{i2}	...	λ	...	λa_{in}
\vdots	\vdots	\vdots	\vdots	...	\vdots	...	\vdots
x_m	b'_m	a'_{m1}	a'_{m2}	...	$-\lambda a_{mj}$...	a'_{mn}
Z_2	C'_0	C'_1	C'_2	...	$-\lambda C_j$...	C'_n

6 ЗАДАЧИ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Задачи математического программирования

$$\max (\min) Z = f(x); \quad (6.1)$$

$$\begin{cases} \varphi_i(x) \{ \leq, =, \geq \} b_i \quad (i=1, \dots, m) \\ x \geq 0 \end{cases}, \quad x = (x_1, \dots, x_n), \quad (6.2)$$

в которых либо целевая функция (6.1), либо ограничения (6.2), либо и то и другое нелинейны, называются нелинейными.

Сложность решения задач нелинейного планирования заключается кроме нелинейности условий задачи еще и в том, что некоторые переменные могут изменяться не непрерывно, а принимать ряд заданных фиксированных значений. Кроме того, сложность решений и в том, что целевые функции могут иметь не один, а несколько максимумов (минимумов), и нужно найти глобальный экстремум.

Геометрическая интерпретация задач нелинейного планирования аналогична задачам линейного планирования.

Общая постановка задачи нелинейного планирования может быть сформулирована следующим образом: найти параметры X^* (x_1, x_2, \dots, x_n), обращающих целевую функцию $W = f(x_1, x_2, \dots, x_n)$ в $\max(\min)$ при условии наложенных ограничений

$$\left. \begin{array}{l}
 \varphi_1(x_1, x_2, \dots, x_n) \leq b_1; \\
 \varphi_2(x_1, x_2, \dots, x_n) = b_2; \\
 \dots\dots\dots \\
 \varphi_i(x_1, x_2, \dots, x_n) \geq b_i; \\
 \dots\dots\dots \\
 \varphi_m(x_1, x_2, \dots, x_n) \leq b_m; \\
 x_j \geq 0 \quad j = 1, 2, \dots, n;
 \end{array} \right\}$$

число ограничений меньше числа переменных; и (или) целевая функция, и (или) ограничения представляются нелинейными зависимостями.

Нелинейные задачи составляют широкий класс настолько сложных задач, что до сих пор невозможно разработать общие методы, подобные симплекс-методу в линейном программировании, которые позволяли бы решать любые нелинейные задачи. Но, несмотря на отсутствие универсальных методов, разработаны способы решения специальных классов задач, и прежде всего задач с выпуклыми (вогнутыми) функциями $f(x)$ и $\varphi_i(x)$.

Особенности решения задач нелинейного программирования.

При решении задач нелинейного программирования очень важно знать: 1) выпукло или не выпукло множество решений? 2) является ли критериальная функция $W = f(x)$ выпуклой или вогнутой, или не относится ни к тому, ни к другому классу?

Множество выпукло, если оно содержит точки А и В, а так же все точки прямой АВ. (Примеры множеств, обладающих свойством выпуклости, приведены на рис.6.1 а, б. Множества, не обладающих свойством выпуклости, приведены на рис.6.1.в, г, д).

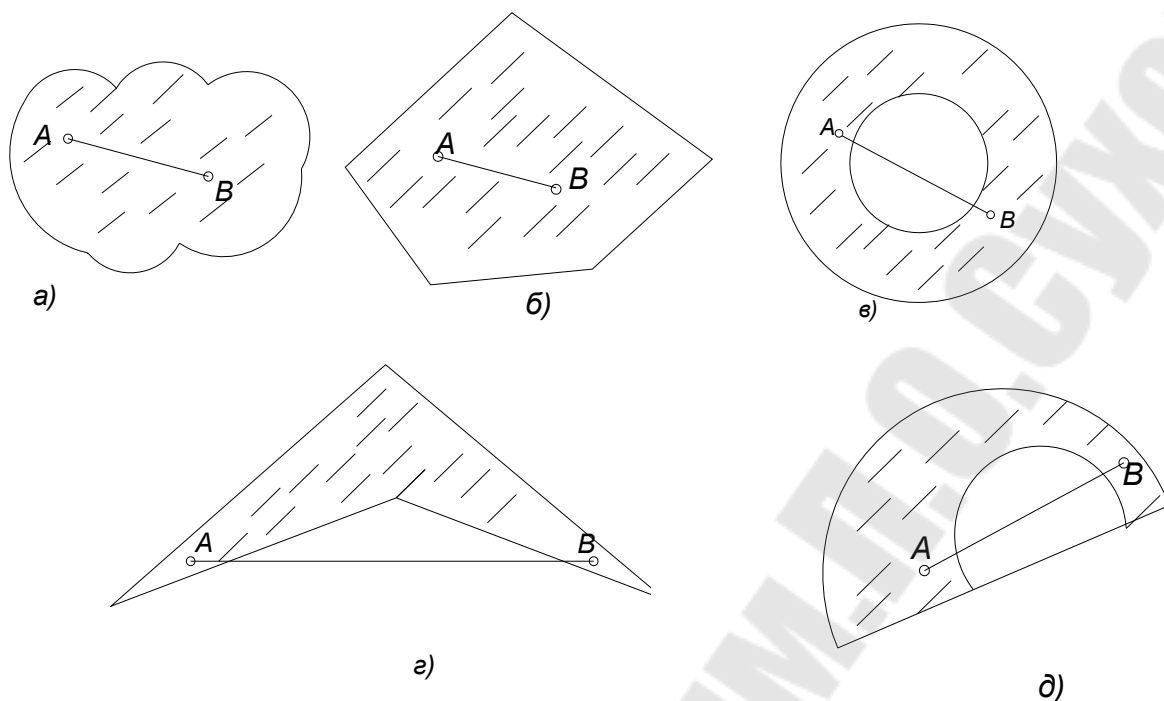


Рис 6.1 (а, б, в, г д)

Функция $y = f(x)$, определенная на выпуклом множестве X , называется выпуклой, если отрезок, соединяющий любые две его точки, принадлежит графику или располагается выше его (рис. 6.2 а). Функция $y = f(x)$, определенная на выпуклом множестве X , называется вогнутой, если отрезок, соединяющий любые две его точки, принадлежит графику или располагается ниже его (рис. 6.2. б).

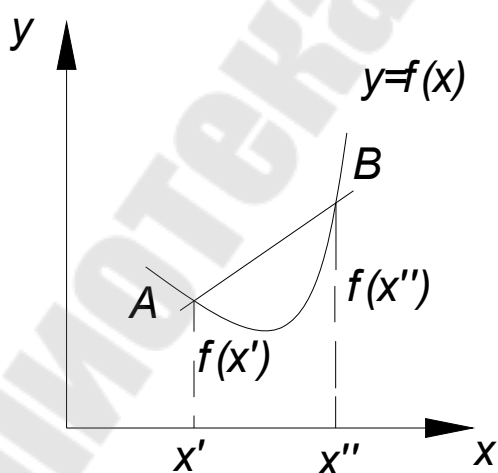


Рис. 6.2.а

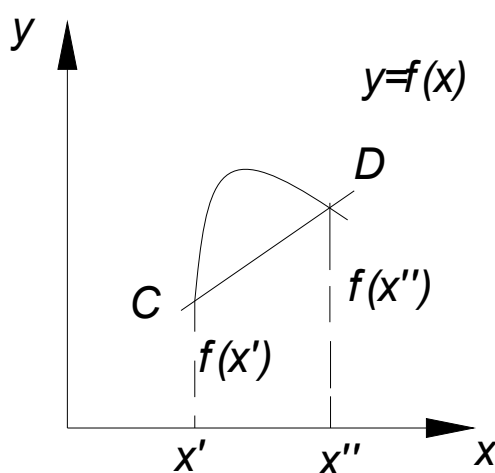


Рис.6.2. б

Аналогичные понятия можно привести и для функций многих переменных.

В математике доказывается ряд теорем, которые позволяют определять глобальные экстремумы. Так, для задач, в которых множество допустимых решений выпукло: 1) если $f(x)$ – выпуклая функция, то локальный минимум, определенный на выпуклом множестве X , совпадает с ее глобальным минимумом на этом множестве; 2) если $f(x)$ – вогнутая функция на заданном выпуклом множестве X , то локальный максимум $f(x)$ является глобальным.

Методы решения задач нелинейного программирования

Метод неопределенных множителей Лагранжа является классическим методом решения задач математического программирования. При практическом применении метода могут встретиться значительные вычислительные трудности, сужающие область его использования. Метод Лагранжа является аппаратом, используемым для обоснования различных современных численных методов.

Рассмотрим задачу оптимизации

$$\max (\min) Z = f(x) \quad (6.3)$$

$$\varphi_i(x) = b_i \quad (i=1, \dots, m), \quad x = (x_1, x_2, \dots, x_n). \quad (6.4)$$

Эта задача выделяется из задачи (6.1), (6.2) тем, что среди ограничений (6.4) нет неравенств, нет условий не отрицательности переменных, их дискретности, $m < n$ и функции $f(x)$ и $\varphi_i(x)$ непрерывны и имеют частные производные, по крайней мере второго порядка.

Последовательность решения данной задачи методом неопределенных множителей Лагранжа:

1) составить функцию Лагранжа

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{i=1}^m \lambda_i (b_i - \varphi_i(x_1, \dots, x_n)) \quad (6.5)$$

$\lambda_1, \dots, \lambda_m$ – множители Лагранжа;

2) найти частные производные функции Лагранжа по всем переменным $x_1, x_2, \dots, x_n, \lambda_1, \dots, \lambda_m$

$$\begin{cases} \frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial \varphi_i}{\partial x_j} = 0 & (j = 1, 2, \dots, n) \\ \frac{\partial L}{\partial \lambda_i} = b_i - \varphi_i = 0 & (i = 1, 2, \dots, m) \end{cases} \quad (6.6)$$

Получили систему уравнений(6.6), состоящую из $n + m$ уравнений. Решить полученную систему (если это окажется возможным!) и найти таким образом все стационарные точки функции Лагранжа;

3) из стационарных точек, взятых без координат $\lambda_1, \dots, \lambda_m$, выбрать точки, в которых функция $f(x)$ имеет условные локальные экстремумы при наличии ограничений (6.4).

Признаком существования минимума функции $f(x)$ в стационарной точке x^* является выполнение в ней достаточных условий минимума – выпуклости функции в окрестности этой точки, что может быть представлено следующим образом [5]:

$$\frac{\partial^2 f(x)}{\partial x_1^2} \Big|_{x = x^*} > 0;$$

$$\begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \Big|_{x = x^*} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_2^2} \Big|_{x = x^*} \end{bmatrix} > 0$$

$$\begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_3} \Big|_{x = x^*} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_2^2} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_3} \Big|_{x = x^*} \\ \frac{\partial^2 f(x)}{\partial x_3 \partial x_1} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_3 \partial x_2} \Big|_{x = x^*} & \frac{\partial^2 f(x)}{\partial x_3^2} \Big|_{x = x^*} \end{bmatrix} > 0$$

т.е. все определители должны быть положительны.

Этот метод можно обобщить и на случай, когда переменные не отрицательны и некоторые ограничения заданы в форме неравенств.

Методы спуска. Численные (поисковые) методы играют существенную роль при решении многих прикладных задач, в том числе электроэнергетики. Это обусловлено рядом причин, среди которых главное место занимает разнообразие целевых функций и ограничений, а также форм их задания.

Допустим, что рассматривается задача безусловной минимизации целевой функции $F(x)$. Сущность всех методов решения этой задачи, о которых далее пойдет речь, состоит в построении последова-

тельности точек $x^{(0)}, x^{(1)}, \dots, x^{(p)}, \dots$, монотонно уменьшающих значение целевой функции, т.е.

$$F(x^{(0)}) \geq F(x^{(1)}) \geq F(x^{(2)}) \geq \dots \geq F(x^{(p)}) \geq \dots$$

Такие методы (алгоритмы) называют методами спуска (или подъема при максимизации целевой функции). Их важнейшей характеристикой является сходимость, которая состоит в том, что при беспредельном увеличении последовательность $x^{(0)}, x^{(1)}, \dots, x^{(p)}, \dots$ сходится в точке глобального (локального) минимума.

Использование всех методов спуска сводится к следующей схеме. Пусть на p -й итерации (p -м шаге) имеется точка $x^{(p)}$. Тогда выбирают направление спуска – единичный вектор $l^{(p)}$, определяющий это направление, а также длину шага $\alpha_p > 0$ вдоль этого направления. Очередную точку вычисляют по следующей формуле:

$$x^{(p+1)} = x^{(p)} + \alpha_p l^{(p)}, \quad p = 0, 1, \dots$$

Различные методы спуска отличаются друг от друга подходом к выбору длины шага α_p (при $\alpha_p = \alpha$ спуск осуществляется с постоянным шагом) и вектора $l^{(p)}$. Все методы спуска можно разделить на три группы в зависимости от того, какие характеристики целевой функции $F(x)$ используются для определения α_p и $l^{(p)}$. К первой группе относят методы, требующие только вычисления значений целевой функции. Эти методы называются методами нулевого порядка. Методы первого порядка требуют вычисления значений функции и ее первых производных. Наконец, методы второго порядка предполагают исследование и вторых производных.

С помощью методов нулевого порядка можно решать задачи более широкого класса, чем с помощью методов первого и второго порядка. В частности, на их основе могут минимизироваться не дифференцируемые функции, задаваемые таблично или алгоритмически.

Широкое распространение в энергетических расчетах нашел метод нулевого порядка – метод покоординатного спуска. В соответствии с этим методом направление спуска выбирают параллельно координатным осям. Сначала проводят спуск вдоль оси Ox_1 , затем – вдоль оси Ox_2 и т.д. вплоть до оси Ox_n .

Градиентные методы. Градиентным методом можно решать, вообще говоря, любую нелинейную задачу. Однако при этом находится лишь локальный экстремум. Поэтому применять этот метод рационально при решении задач выпуклого программирования, в которых любой локальный экстремум является одновременно и глобальным.

Задачи без ограничений. Рассмотрим задачу максимизации нелинейной дифференцируемой функции $f(X)$. Суть градиентного поиска точки максимума X^* очень проста: надо взять произвольную точку X_0 и с помощью градиента $\nabla f(X_0)$, вычисленного в этой точке, определить направление, в котором возрастает с наибольшей скоростью, а затем, сделав небольшой шаг в найденном направлении, перейти в новую точку X_1 . Потом снова определить наилучшее направление $\nabla f(X_1)$ для перехода в очередную точку X_2 и т.д. Таким образом, надо построить последовательность точек $X_0, X_1, X_2, \dots, X_k, \dots$ так, чтобы она сходилась к точке X^* , т.е. для точек последовательности выполнялись условия

$$f(X_0) < f(X_1) < f(X_2) < \dots < f(X_k) < \dots$$

Графическое решение данной задачи можно представить в виде рис. 6.3

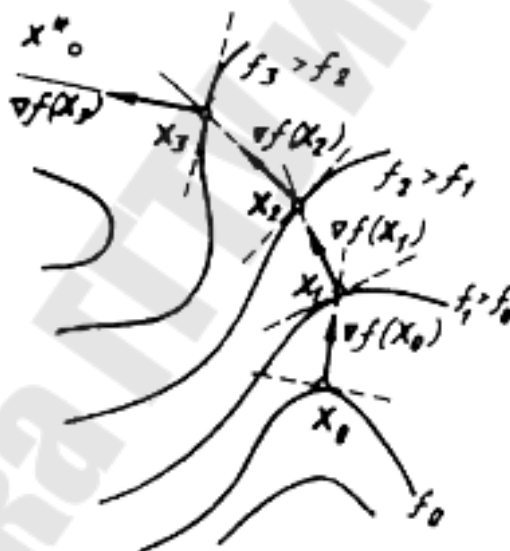


Рис. 6.3

Градиентные методы позволяют получать точное решение за бесконечное число шагов и только в некоторых случаях – за конечное. В связи с этим градиентные методы относят к приближенным методам решения.

Движение из точки X_k в новую точку X_{k+1} осуществляется по прямой, проходящей через точку X_k и имеющей уравнение

$$X_{k+1} = X_k + \alpha_k \nabla f(X_k) \quad , \quad (6.7)$$

где α_k – числовой параметр, от которого зависит длина шага.

Градиентные методы отличаются друг от друга способом выбора величины шага – значения параметра α_k . Можно, например, двигаться из точки в точку с постоянным шагом $\alpha_k = \alpha$, т.е. при любом k

$$X_{k+1} = X_k + \alpha \nabla f(X_k) \quad .$$

Если при этом окажется, что $f(X_{k+1}) < f(X_k)$, то следует возвратиться в точку X_k и уменьшить значение параметра α , например до $\alpha/2$.

Если ищется приближенное решение, то поиск можно прекратить, основываясь на следующих соображениях. После каждой серии из определенного числа шагов сравнивают достигнутые значения целевой функции $f(X)$. Если после очередной серии изменение $f(X)$ не превышает некоторого наперед заданного малого числа ε , поиск прекращают и достигнутое значение $f(X)$ рассматривают как искомым приближенный максимум, а соответствующее ему X принимают за X^* .

Если целевая функция $f(X)$ вогнутая (выпуклая), то необходимым и достаточным условием оптимальности точки X^* является равенство нулю градиента функции в этой точке.

Распространенным является вариант градиентного поиска, называемый *методом наискорейшего подъема* (наискорейшего спуска – если решается задача минимизации). Суть его в следующем. После определения $\nabla f(X_k)$ в точке X_k движение вдоль прямой $X = X_k + \alpha_k \nabla f(X_k)$ производится до точки X_{k+1} , в которой достигается максимальное значение функции $f(X)$ в направлении градиента $\nabla f(X_k)$; затем в этой точке вновь определяется градиент, и движение совершается по прямой $X = X_{k+1} + \alpha_{k+1} \nabla f(X_{k+1})$ в направлении нового градиента $\nabla f(X_{k+1})$ до точки X_{k+2} , в которой достигается максимальное в этом направлении значение $f(X)$. Движение продолжается до тех пор, пока не будет достигнута точка X^* , соответствующая наибольшему значению целевой функции $f(X)$. На рис.6.4 приведена схема движения к оптимальной точке X^* методом наискорейшего подъема. В данном случае направление градиента $\nabla f(X_k)$ в точке X_k является касательным к линии уровня поверхности $f(X)$ в точке X_{k+1} , следовательно, градиент $\nabla f(X_{k+1})$ в точке X_{k+1} ортогонален градиенту $\nabla f(X_k)$.

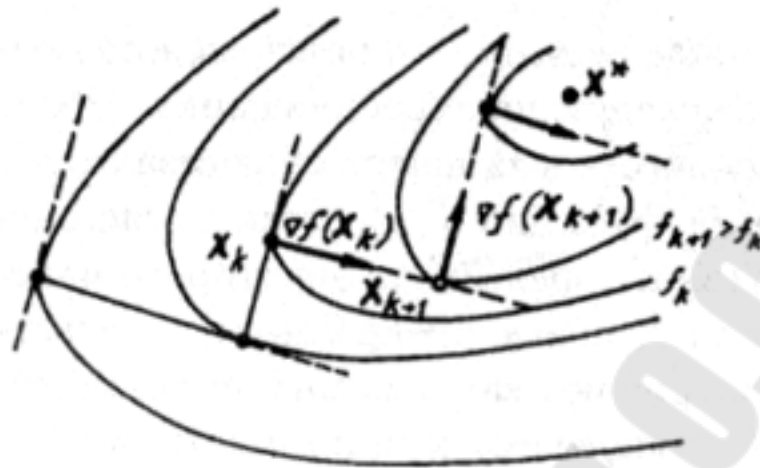


Рис.6.4

Перемещение из точки X_k в точку $X_{k+1} = X_k + \alpha \nabla f(X_k)$ сопровождается возрастанием функции $f(X)$ на величину

$$\begin{aligned} \Delta f(X) &= f(X_{k+1}) - f(X_k) = f(X_{k+1,1}; \dots; X_{k+1,n}) - f(X_{k,1}; \dots; X_{k,n}) = \\ &= f \left\{ x_{k,1} + \alpha_k \frac{\partial f(x_k)}{\partial x_1}; \dots; x_{k,n} + \alpha_k \frac{\partial f(x_k)}{\partial x_n} \right\} - f(x_{k,1}; \dots; x_{k,n}) \end{aligned} \quad (6.8)$$

Из выражения (6.8) видно, что Δf является функцией переменной α_k , т.е. $\Delta f = \Delta f(\alpha_k)$. При прохождении максимума функции $f(x)$ в направлении градиента $\nabla f(X_k)$ необходимо выбирать шаг перемещения (множитель α_k), обеспечивающий наибольшее возрастание приращению функции, именно функции $\Delta f(\alpha_k)$. Величина α_k , при которой достигается наибольшее значение $\Delta f(\alpha_k)$, может быть определена из необходимого условия экстремума функции $\Delta f(\alpha_k)$:

$$d(\Delta f(\alpha_k))/d\alpha_k = 0 \quad (6.9)$$

Найдем выражение для производной, дифференцируя равенство (6.8) по α_k как сложную функцию:

$$\frac{d(\Delta f(\alpha_k))}{d\alpha_k} = \frac{\partial f(x_{k+1})}{\partial x_1} \cdot \frac{\partial f(x_k)}{\partial x_1} + \dots + \frac{\partial f(x_{k+1})}{\partial x_n} \cdot \frac{\partial f(x_k)}{\partial x_n} = \nabla f(X_{k+1}) \nabla f(X_k)$$

Подставляя этот результат в равенство (6.9), получаем

$$d(\Delta f(\alpha_k))/d\alpha_k = \nabla f(X_{k+1})\nabla f(X_k);$$

Это равенство имеет простое геометрическое истолкование: градиент в очередной точке X_{k+1} ортогонален градиенту в предыдущей точке X_k .

Задачи с линейными ограничениями. Рассмотрим задачу выпуклого программирования с линейными ограничениями:

$$\max Z = f(x) \quad (6.10)$$

$$ax \leq a_{i0} \quad (i=1, 2, \dots, m) \quad (6.11)$$

$$x \geq 0, \quad (6.12)$$

где $x = (x_1; x_2; \dots; x_n)$; $a_i = (a_{i1}; a_{i2}; \dots; a_{in})$.

Предполагается, что $f(x)$ является вогнутой функцией и имеются непрерывные частные производные в каждой точке допустимой области.

Начнем с геометрической иллюстрации процесса решения задачи (рис.6.5). Пусть начальная точка X_0 расположена внутри допустимой области. Из точки X_0 можно двигаться в направлении градиента $\nabla f(X_0)$, пока $f(X)$ не достигнет максимума. В нашем случае $f(X)$ все время возрастает, поэтому остановиться надо в точке X_1 на граничной прямой.

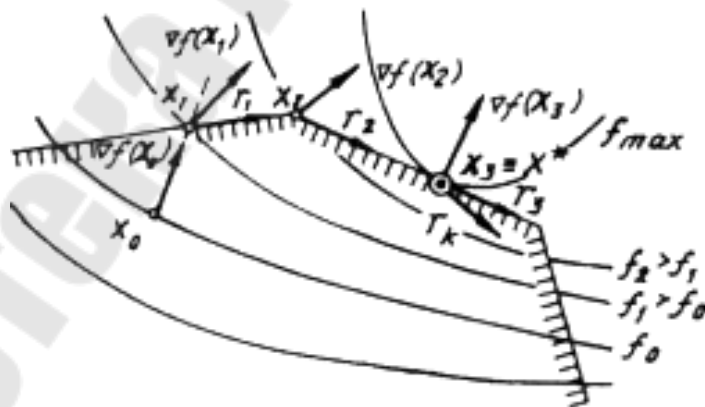


Рис.6.5

Как видно из рисунка, дальше двигаться в направлении градиента $\nabla f(X_1)$ нельзя, т.к. выйдем из области допустимых решений. Поэтому надо найти другое направление перемещения, которое, с одной стороны, не выходит из допустимой области, а с другой – обеспечивает наибольшее возрастание $f(X)$. Такое направление определяет век-

тор r_1 , составляющий с вектором $\nabla f(X_1)$ наименьший острый угол по сравнению с любым другим вектором, выходящим из точки X_1 и лежащим в допустимой области. Аналитически такой вектор найдется из условия максимизации скалярного произведения $\nabla f(X_1) r_1 > 0$. В данном случае вектор r_1 , указывающий наивыгоднейшее направление, совпадает с граничной прямой.

Таким образом, на следующем шаге двигаться надо по граничной прямой до тех пор, пока возрастает $f(X)$; в нашем случае – до точки X_2 . Из рисунка видно, что далее следует перемещаться в направлении вектора r_2 , который находится из условия максимизации скалярного произведения $\nabla f(X_2) r_2 > 0$, т.е. по граничной прямой. Движение заканчивается в точке X_3 , поскольку в этой точке завершается оптимизационный поиск, т.к. в ней функция $f(X)$ имеет локальный максимум. Ввиду вогнутости оптимизационной функции, в этой точке $f(X)$ достигает глобального максимума в допустимой области. Градиент в точке максимума $X_3 = X^*$ составляет тупой угол с любым вектором r_k из допустимой области, проходящей через X_3 , поэтому скалярное произведение $\nabla f(X_3) r_k$ будет отрицательным для любого допустимого r_k , кроме r_3 , направленного по граничной прямой. Для него скалярное произведение $\nabla f(X_3) r_3 = 0$, т.к. $\nabla f(X_3)$ и r_3 взаимно перпендикулярны (граничная прямая касается линии уровня поверхности $f(X)$, проходящей через точку максимума X^*). Это равенство и служит аналитическим признаком того, что в точке X_3 функция $f(X)$ достигла максимума.

Аналитическое решение задачи (6.10) – (6.12) требует, чтобы при выборе α_k определение очередной точки $X_{k+1} = X_k + \alpha_k \nabla f(X_k)$ осталось в допустимой области. Существуют различные методы оптимизации длины шага (α_k) и проверки принадлежности очередной точки X_k допустимой области.

Задачи с нелинейными ограничениями. Если в задачах с линейными ограничениями движение по граничным прямым оказывается возможным и даже целесообразным, то при нелинейных ограничениях, определяющих выпуклую область, любое сколь угодно малое перемещение из граничной точки может сразу вывести за пределы области допустимых решений, и возникает необходимость в возвращении в допустимую область (рис. 6.6). Подобная ситуация характерна

для задач, в которых экстремум функции $f(X)$ достигается на границе области. В связи с этим

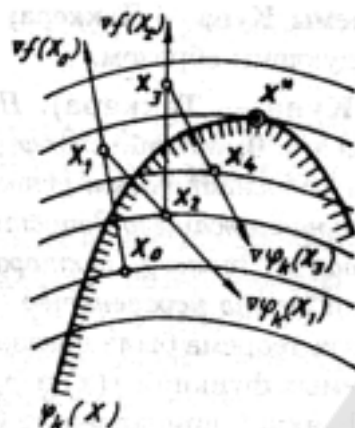


Рис 6.6

применяются различные способы перемещения, обеспечивающие построение последовательности точек, расположенных вблизи границы и внутри допустимой области, или зигзагообразное движение вдоль границы с пересечением последней. Как видно из рисунка, возврат из точки X_1 в допустимую область следует осуществлять вдоль градиента той граничной функции $\varphi_k(X)$, которая оказалась нарушенной. Это обеспечит отклонение очередной точки X_2 в сторону точки экстремума X^* . Признаком экстремума в подобном случае будет коллинеарность векторов ∇f и $\nabla \varphi_k$.

7 ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

Динамическое программирование (динамическое планирование)

– это метод нахождения оптимальных решений в задачах с многошаговой (многоэтапной) структурой. Это процессы планирования и управления, развивающиеся во времени. Однако, метод динамического программирования может использоваться при решении задач, где время вообще не фигурирует; разделение на шаги в таких задачах вводится искусственно. Поэтому «динамика» задач динамического программирования заключается в методе решения.

В практике встречаются задачи, которые связаны с развитием производства во времени (распределительные сети на перспективу), оптимальным управлением качества электрической энергии и др. Их решают либо путем составления статических моделей для каждого периода, либо путем составления единой динамической задачи опти-

мального программирования с применением многошаговой процедуры принятия решений.

Особенности задач динамического программирования.

1. Рассматривается система, состояние которой на каждом шаге определяется вектором x_t . Дальнейшее изменение ее состояния зависит только от данного состояния x_t и не зависит от того, каким путем система пришла в него.

2. На каждом шаге выбирается одно решение u_t , под воздействием которого система переходит из предыдущего состояния x_{t-1} в новое x_t . Это новое состояние является функцией состояния на начало интервала x_{t-1} и принятого в начале интервала решения u_t , т.е.

$$x_t = x_t(x_{t-1}, u_t)$$

3. Действие на каждом шаге связано с определенным выигрышем или потерей, которые зависят от состояния на начало шага (этапа) и принятого решения.

4. На векторы состояния и управления могут быть наложены ограничения, объединение которых составляет область допустимых решений $u \in \Omega$.

5. Требуется найти такое допустимое управление u_t для каждого шага t , чтобы получить экстремальное значение функции цели за все T шагов.

Любую допустимую последовательность действий для каждого шага, переводящую систему из начального состояния в конечное, называют *стратегией управления*. Допустимая стратегия управления, обеспечивающая функции цели экстремальное значение, называется *оптимальной*.

Геометрическая интерпретация задачи динамического программирования состоит в следующем. Пусть n – размерность пространства состояний. В каждый момент времени координаты системы имеют вполне определенные значения. С изменением времени могут изменяться значения координат вектора состояния. Назовем переход системы из одного состояния в другое траекторией движения в пространстве состояний. Наглядно задачу динамического программирования можно интерпретировать в случае, если пространство состояний двумерно. Область возможных состояний в этом случае изобразится некоторой фигурой Ω , начальное и конечное состояния системы – точками $x_0, x_T \in \Omega$ (рис.7.1).

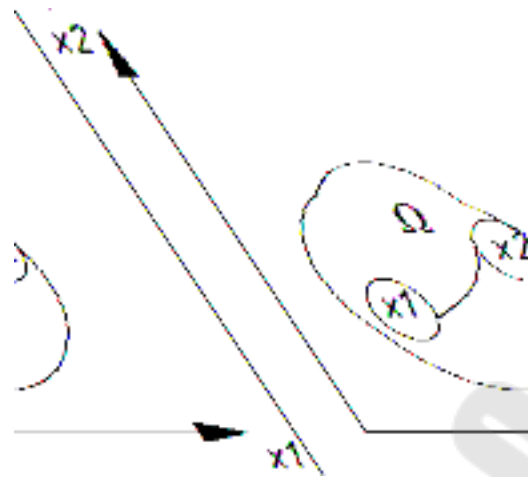


Рис.7.1

Управление – это воздействие, переводящее систему из начального состояния в конечное. Для многих задач известны область X_0 и X_T . Тогда допустимые управления переводят точки из области X_0 в X_T . Задача динамического программирования геометрически может быть сформулирована следующим образом: найти такую траекторию, начинающуюся в области X_0 и оканчивающуюся в области X_T , для которой функция цели достигает экстремального значения.

Принципы динамического программирования. Любую многошаговую задачу можно решать по-разному. Во-первых, можно считать неизвестными величинами u_t и находить экстремум целевой функции одним из существующих методов оптимизации, т.е. искать сразу все элементы решения на всех N шагах. Отметим, что этот путь не всегда приводит к цели, особенно когда функция заданна в виде таблиц или число переменных очень велико. Второй путь основан на идее проведения оптимизации поэтапно. Поэтапность отнюдь не предполагает изолированности в оптимизации этапов. Наоборот, управление на каждом шаге выбирается с учетом всех его последствий. Чаще второй способ оптимизации оказывается проще, чем первый, особенно при большом числе шагов. Идея постепенной, пошаговой оптимизации составляет суть метода динамического программирования. Оптимизация одного шага проще оптимизации всего процесса в целом.

Приведем пример. Пусть проектируется распределительная сеть между пунктами А и В (в условиях существующей городской застройки).

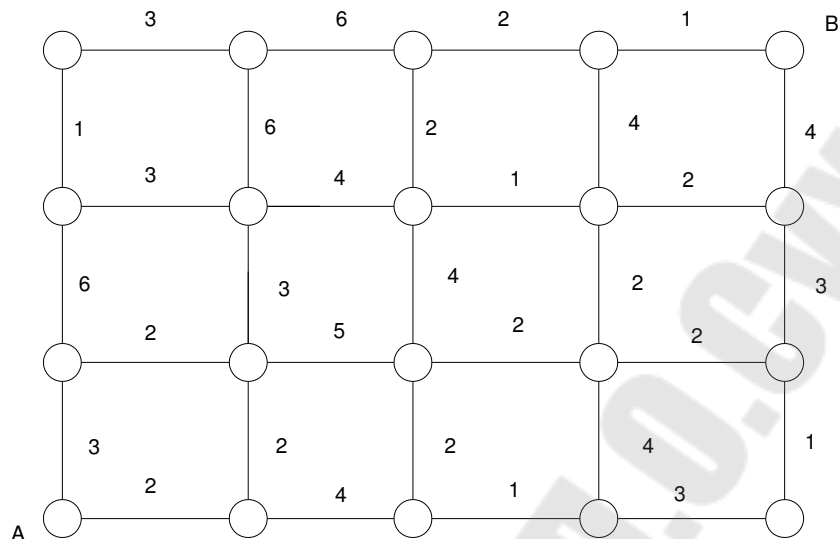


Рис.7.2

Различные варианты трассы требуют неодинаковых затрат в связи с особенностью уличных застроек, необходимостью электрообеспечения определенного объекта, естественных препятствий и т.д. Требуется так выполнить сеть, связывающих пункты А и В, чтобы суммарные затраты на сооружение были минимальны.

Отметим, что в задаче нет естественного деления на шаги. Это деление вводится искусственно, для чего схему между А и В разбивается на N частей и за шаг оптимизации принимается каждая такая часть.

Одним из условий применимости метода динамического программирования является возможность разбиения процесса оптимизации решения на ряд однотипных шагов (этапов), каждый из которых планируется отдельно, но с учетом состояния системы на начало этапа и последствий принятого решения. Однако из этого правила есть исключение. Среди всех шагов существует один, который может планироваться без учета последствий. Это последний шаг. Он может быть изучен и спланирован сам по себе наилучшим (в смысле выбранного критерия) образом, поскольку за ним нет больше этапов. Отсюда получаем одну из специфических особенностей динамического программирования: всю вычислительную процедуру программирования целесообразно разворачивать от конца к началу. Раньше всех планируется последний $N - й$ шаг, за ним $(N-1)$ -й и т.д. Но как найти оптимальное управление u_N на $N - м$ шаге, если оно определяется не только целью управления, но и состоянием системы на начало этого шага? Сделать это можно на основе предположений об ожидаемых

исходах предшествующего, но еще не исследованного этапа, т.е. о значениях x_{N-1} .

Для каждого возможного исхода x_{N-1} на (N-1)-м этапе находим оптимальное управление на N-м этапе. Такой набор оптимальных управлений, зависящих от возможных исходов предыдущего этапа, называется *условно-оптимальным решением* $u_N^*(x_{N-1})$. Завершив анализ конечного этапа, требуя, чтобы функция цели достигала экстремального значения на двух последних этапах вместе. Это дает условно-оптимальное решение на предпоследнем этапе $u_{N-1}^*(x_{N-2})$, т.е. делаются всевозможные предположения о том, чем кончился предыдущий (N-2)-й шаг, и для каждого из предположений находится такое управление на (N-1)-м шаге, при котором эффект за последние два шага (из них последний уже оптимизирован) будет максимален. Тем самым мы найдем для каждого исхода (N-2)-го шага условно-оптимальное управление на (N-1)-м и условно-оптимальное значение функции цели на последних двух шагах. Прделав такой поиск условно-оптимальных управлений для каждого шага от конца к началу, найдем последовательность условно-оптимальных управлений $u_1^*(x_0), u_2^*(x_1), \dots, u_{N-1}^*(x_{N-2}), u_N^*(x_{N-1})$.

Условно-оптимальные управления дают возможность найти не условное, а просто оптимальное управление на каждом шаге. В самом деле, пусть начальное состояние x_0 известно. Тогда, проделав процедуру движения от конца к началу, находим $u_1^*(x_0)$. Так как начальное состояние x_0 определяется однозначно, это оптимальное управление для первого шага. Вместе с тем находим экстремальное значение целевой функции относительно всего процесса. Зная оптимальное действие (с точки зрения всего процесса) для первого шага, выявим, к какому состоянию перейдет система в результате этого действия, т.е. найдем оптимальное состояний системы x_1^* на начало второго этапа. Но для всех возможных состояний на начало второго этапа выявлены оптимальные управления. Таким образом, зная x_1^* , установим оптимальное управление для второго этапа $u_2^*(x_1^*)$ и т.д. Прделав обратное движение от начала к концу, найдем оптимальные управления для всех этапов.

Таким образом, в процессе оптимизации управления методом динамического программирования многошаговый процесс проходит дважды. Первый раз от конца к началу, в результате чего находятся условно-оптимальные управления и условно-оптимальное значение

функции цели для каждого шага, в том числе оптимальное управление для первого шага и оптимальное значение функции цели для всего процесса. Второй раз – от начала к концу, в результате чего находятся уже оптимальные управления на каждом шаге с точки зрения всего процесса. Первый этап сложнее и длительнее второго, на втором остается лишь рекомендации, полученные на первом. Следует отметить, что понятия «конец» и «начало» можно поменять местами и разворачивать процесс оптимизации в другом направлении. С какого конца начать – диктуется удобством выбора этапов и возможных состояний на их начало.

Принцип оптимальности. Оптимальное управление на каждом шаге определяется состоянием системы на начало этого шага и целью управления. Оптимальная стратегия обладает таким свойством, что, каковы бы ни были начальное состояние и начальные решения, последующие решения должны приниматься исходя из оптимальной стратегии с учетом состояния, вытекающего из первого решения. Принцип оптимальности выражается в составлении определенных рекуррентных соотношений (функциональных уравнений Р. Беллмана).

Рекуррентные соотношения методов динамического программирования (функциональные уравнения Р. Беллмана). В основе динамического программирования лежит принцип оптимальности, указывающий на процедуру построения оптимального управления. Так как оптимальной стратегией может быть только та, которая одновременно оптимальна и для любого количества оставшихся шагов, ее можно строить по частям: сначала для последнего этапа, затем для двух последних, для трех и т.д., пока не приходим к первому шагу.

Приведем математическую формулировку принципа оптимальности для задач с аддитивным критерием оптимальности (сепарабельная функция цели). Для простоты будем считать, что начальное x_0 и конечное x_T состояния системы заданы. Обозначим через $z_1(x_0, u_1)$ значение функции цели на первом этапе при начальном состоянии системы x_0 и при управлении u_1 , через $z_2(x_1, u_2)$ - соответствующее значение функции цели только на втором этапе, ..., через $z_i(x_{i-1}, u_i)$ - на i - м этапе, ..., через $z_N(x_{N-1}, u_N)$ - на N -м этапе. Очевидно, что

$$Z = z(x_0, u) = \sum_{i=1}^N Z_i(x_{i-1}, u_i) \quad (7.1)$$

Надо найти оптимальное управление $u^* = (u_1^*; u_2^*; \dots; u_N^*)$ такое, что обеспечивает экстремум целевой функции (7.1) при ограничениях $u \in \Omega$.

Для решения этой задачи введем обозначения. Пусть $\Omega_N, \Omega_{N-1, N}, \dots, \Omega_{1, N} \equiv \Omega$ – соответственно области определения на последнем этапе, двух последних и т.д.; Ω – область определения исходной задачи. Обозначим через $F_1(x_{N-1}), F_2(x_{N-2}), \dots, F_k(x_{N-k}), \dots, F_N(x_0)$ – соответственно условно-оптимальные значения функции цели на последнем этапе, двух последних и т.д., на k последних и т.д., на всех N этапах.

Начинаем с последнего этапа. Пусть x_{N-1} – возможные состояния системы на начало N – го этапа. Находим:

$$F_1(x_{N-1}) = \max(\min)_{u_N \in \Omega} z_N(x_{N-1}, u_N) \quad (7.2)$$

Для двух последних этапов получаем

$$F_2(x_{N-2}) = \max(\min)_{u_{N-1} \in \Omega} [z_{N-1}(x_{N-2}, u_{N-1}) + F_1(x_{N-1})] \quad (7.3)$$

Аналогично:

$$F_3(x_{N-3}) = \max(\min)_{u_{N-2} \in \Omega} [z_{N-2}(x_{N-3}, u_{N-2}) + F_2(x_{N-2})] \quad (7.4)$$

$$\dots \dots \dots$$

$$F_k(x_{N-k}) = \max(\min)_{u_{N-k+1} \in \Omega} [z_{N-k+1}(x_{N-k}, u_{N-k+1}) + F_{k-1}(x_{N-k+1})] \quad (7.5)$$

$$\dots \dots \dots$$

$$F_N(x_0) = \max(\min)_{u_1 \in \Omega} [z_1(x_0, u_1) + F_{N-1}(x_1)] \quad (7.6).$$

Выражение (7.6) представляет собой математическую запись принципа оптимальности. Выражение (7.5) – общая форма записи условно-оптимального значения функции цели для k оставшихся этапов. Выражения (7.2) – (7.6) называются *рекуррентными соотношениями (функциональными уравнениями) Беллмана*. Отчетливо просматривается их рекуррентный (возвратный) характер, т.е. для нахождения оптимального управления на N шагах нужно знать условно-оптимальное управление на предшествующих $N-1$ этапах и т.д.

Примеры решения задач методом динамического программирования

1. Решим с помощью метода динамического программирования задачу определения оптимальной степени участия синхронных двигателей в компенсации реактивной мощности $Q_{\Sigma} = 1,8$ Мвар, если каждый из двигателей может выдать реактивную мощность соответственно: $Q_{д1} = 0,6$ Мвар; $Q_{д2} = 0,5$ Мвар и $Q_{д3} = 1,2$ Мвар. Потери мощности ΔP , возникающие в синхронном двигателе, зависят от реактивной мощности $Q_{д}$, генерируемой двигателем. ΔP_i представляются нелинейными зависимостями от $Q_{дi}$.

Величина потерь мощности, возникающих в первом двигателе составит - $\Delta P_1 = 0,00749 Q_1 + 0,01104 Q_1^2$,

во втором двигателе - $\Delta P_2 = 0,01294 Q_2 + 0,02244 Q_2^2$

и в третьем - $\Delta P_3 = 0,00973 Q_3 + 0,00797 Q_3^2$.

Задача состоит в том, чтобы отыскать такие мощности $Q_i, i = 1,2,3$, которые обеспечили бы минимум суммарных потерь $\Delta P_{\Sigma} = \Delta P_1 + \Delta P_2 + \Delta P_3$ при соблюдении ограничений по величинам $Q_{дi}, i = 1,2,3$, а также по суммарной реактивной мощности $Q_{\Sigma} = 1,8$ Мвар, которую должны скомпенсировать синхронные двигатели.

В математическом плане задача сводится к минимизации целевой функции $F(x) = \sum f_i(x_i) \rightarrow \min;$
(х.х7)

$$f_1(x_1) = 0,00749 x_1 + 0,01104 x_1^2,$$

$$f_2(x_2) = 0,01294 x_2 + 0,02244 x_2^2,$$

$$f_3(x_3) = 0,00973 x_3 + 0,00797 x_3^2.$$

при соблюдении ограничений: $x_1 + x_2 + x_3 = 1,8;$

$$x_1 \leq 0,6;$$

$$x_2 \leq 0,5;$$

$$x_3 \leq 1,2;$$

$$x_i \geq 0, i = 1,2,3.$$

При использовании метода динамического программирования производим дискретизацию переменных $\Delta x = 0,1$. т.е. строим сетку $X = 0,0; 1,8$ с шагом Δx .

Формирование рекуррентных соотношений начинается с $k = 1$. Поскольку здесь не требуется распределения ресурса между переменными, то рекуррентное соотношение примет вид $F_1(X) = f_1(X)$; поэтому

$$F_1(0) = f_1(0) = 0,00749 \cdot 0 + 0,01104 \cdot 0^2 = 0;$$

$$F_1(0,1) = f_1(0,1) = 0,00749 \cdot 0,1 + 0,01104 \cdot 0,1^2 = 0,00086;$$

.....

$$F_1(0,6) = f_1(0,6) = 0,00749 \cdot 0,6 + 0,01104 \cdot 0,6^2 = 0,00847.$$

Расчеты $F_1(X)$ прекращаются для $X = 0,6$, поскольку имеются ограничения $x_1 \leq 0,6$.

На втором шаге рекуррентное соотношение имеет вид

$$F_2(X) = \min [F_1(X-x_2) + f_2(x_2)], \text{ в силу чего будем иметь:}$$

$$\text{для } X = 0 \quad F_2(0) = 0 + 0,01294 \cdot 0 + 0,02244 \cdot 0^2 = 0;$$

$$\text{для } X = 0,1 \quad F_2(0,1) = 0,00086 + 0,01294 \cdot 0 + 0,02244 \cdot 0^2 = 0,00086 \quad (x_2=0)$$

$$F_2(0,1) = 0 + 0,01294 \cdot 0,1 + 0,02244 \cdot 0,1^2 = 0,00152 \quad (x_2 = 0,1)$$

т.е. минимальным $F_2(0,1) = 0,00086$ (отмечаем это значение);

$$\text{для } X = 0,2 \quad F_2(0,2) = 0,00194 + 0,01294 \cdot 0 + 0,02244 \cdot 0^2 = 0,00194 \quad (x_2=0,0);$$

$$F_2(0,2) = 0,00086 + 0,01294 \cdot 0,1 + 0,02244 \cdot 0,1^2 = 0,00238 \quad (x_2=0,1);$$

$$F_2(0,2) = 0 + 0,01294 \cdot 0,2 + 0,02244 \cdot 0,2^2 = 0,00349 \quad (x_2 = 0,2)$$

Подобные расчеты продолжаем до $X = 1,1$, чем обеспечивается одновременное соблюдение ограничений $x_1 \leq 0,6$ и $x_2 \leq 0,5$. При этом для $X = 1,1$ вычисляется $F_2(1,1)$ лишь для одного допустимого значения $x_2 = 0,5$:

$$F_2(1,1) = 0,00847 + 0,01294 \cdot 0,5 + 0,02244 \cdot 0,5^2 = 0,02055.$$

На третьем шаге по рекуррентному соотношению $F_3(X) = \min [F_2(X-x_3) + f_3(x_3)]$ вычисляется серия значений $F_3(1,8)$ для $x_3 = 0,7; 0,8; 0,9; 1,0; 1,1; 1,2$. Для $x_3 \leq 0,7$ расчеты не проводятся, поскольку в этом случае никакие сочетания x_1 и x_2 не обеспечат выполнения ограничения $x_1 + x_2 + x_3 = 1,8$.

Ход решения задачи методом динамического программирования при $\Delta x = 0,1$ Мвар представлен в табл. 7.1

Таблица 7.1

X	x_1	$F_1(X)$	x_2	$F_2(X)$	x_3	$F_3(X)$
0,0	0,0	0	0,0	0		
0,1	0,1	0,00086	0,0	0,00086		
0,2	0,2	0,00194	0,0	0,00194		
0,3	0,3	0,00324	0,0	0,00324		
0,4	0,4	0,00476	0,1	0,00476		
0,5	0,5	0,00651	0,1	0,00628		
0,6	0,6	0,00847	0,1	0,00803		
0,7			0,1	0,00998		
0,8			0,2	0,01196		
0,9			0,3	0,01437		
1,0			0,4	0,01724		
1,1			0,5	0,02055		
1,2						
1,3						
1,4						
1,5						
1,6						
1,7						
1,8					0,9	0,02958

Заполнением строки $X = 1,8$ завершается «прямой ход» метода динамического программирования. Из табл.7.1 видно, что минимальное значение целевой функции $F_3(1,8) = 0,02958$. Ему соответствует $x_3^* = 0,9$. Тогда $x_1 + x_2 = 1,8 - 0,9 = 0,9$. Из табл. 7.1 видно, что $F_2(0,9) = 0,01437$. Ему соответствует $x_2^* = 0,3$. Наконец, $x_1^* = 1,8 - x_3^* - x_2^* = 0,6$. Таким образом, получено решение $x_1^* = 0,6$; $x_2^* = 0,3$; $x_3^* = 0,9$, а $\Delta P_\Sigma = \Delta P_1 + \Delta P_2 + \Delta P_3 = F(x) = \sum f_i(x_i) \rightarrow \min = 0,02958$ МВт.

Решение задачи, приведенной на рис.7.2, начиная с «конца» приведено на рис.7.3, а начиная с «начала» - на рис. 7.4.

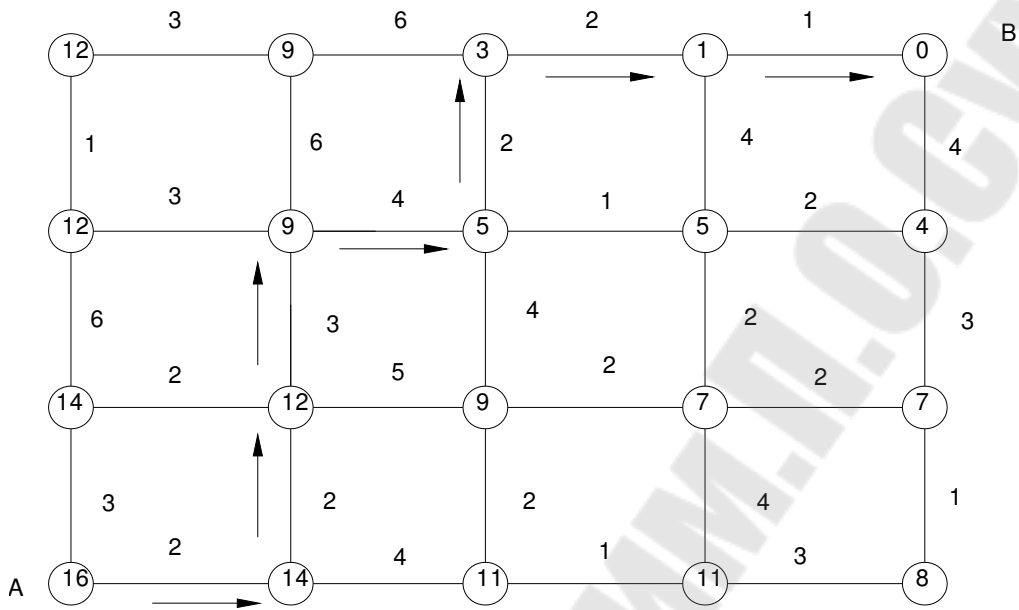


Рис.7.3

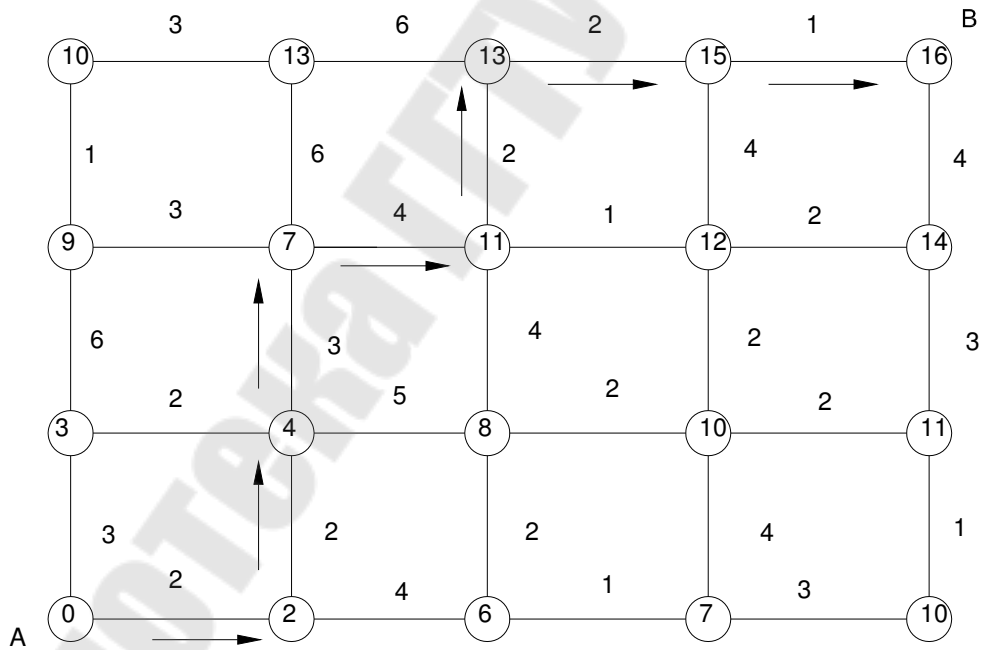


Рис.7.4

Полученные результаты идентичны.

ЛИТЕРАТУРА

1. Электрические системы. Математические задачи электроэнергетики / В. А. Веников [и др.]. – Москва : Высш. шк., 1981. – 288 с.
2. Математическая статистика/В.М.Иванова, В.Н. Калинина [и др.] – Москва : Высш. шк., 1981. – 371 с.
3. Вентцель, Е. С. Теория вероятностей / Е. С. Вентцель. – Москва: Наука, 1969. – 576 с.
4. Кузнецов А.В., Сакович В.А., Холод Н.И. Высшая математика. Математическое программирование – Минск: Вышэйшая школа, 2001. – 351 с.
5. Зайчинка Ю.П. Исследования операций. – Киев,: Вища школа, 1979. – 392 с.
6. Колесникова И.Н., Круглякова Д.В. Статистика – Минск: Вышэйшая школа, 2011. – 285с.
7. Шундалов Б.М. Статистика. Общая теория. – Минск: ИВЦ Минфина, 2006. – 288с.

**Алферова Тамара Викторовна
Попова Ольга Михайловна**

**МАТЕМАТИЧЕСКИЕ ЗАДАЧИ
ЭЛЕКТРОЭНЕРГЕТИКИ**

**Курс лекций
по одноименной дисциплине
для студентов специальности 1-43 01 03
«Электроснабжение (по отраслям)»
дневной и заочной форм обучения**

Подписано к размещению в электронную библиотеку
ГГТУ им. П. О. Сухого в качестве электронного
учебно-методического документа 16.01.13.

Рег. № 38Е.

E-mail: ic@gstu.by

<http://www.gstu.by>