

В.Б. Пономарев

А.Б. Лошкарев

**МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ
ТЕХНОЛОГИЧЕСКИХ
ПРОЦЕССОВ**

Федеральное агентство по образованию

ГОУ ВПО «Уральский государственный технический университет–УПИ»

В.Б. Пономарев

А.Б. Лошкарев

**МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ**

Курс лекций

Научный редактор – проф., канд. техн. наук В.Я. Дзюзер

Екатеринбург

2006

УДК 666.9.001.575 (042.4)

ББК 35.41в6

П 56

Рецензенты:

Уральский региональный центр информатизации Уральского государственного университета им. А.М. Горького (А.И. Петров – главный инженер);

В.Т. Стефаненко – зав. лабораторией, кандидат технических наук, старший научный сотрудник (ФГУП ВУХИН).

Пономарев В.Б.

П56 Математическое моделирование технологических процессов: курс лекций / В.Б. Пономарев, А.Б. Лошкарев. Екатеринбург : ГОУ ВПО УГТУ–УПИ, 2006. 129 с.

ISBN

Курс лекций содержит основные определения моделирования технологических процессов, методики и примеры оптимального планирования и обработки экспериментальных данных, линейного программирования технологических задач.

Курс лекций разработан в соответствии с учебным планом специальности 270101 – Механическое оборудование и технологические комплексы предприятий строительных материалов, изделий и конструкций для студентов очной и заочной форм обучения, а также рекомендован для специальностей:

240304 – Химическая технология тугоплавких неметаллических и силикатных материалов;

270106 – Производство строительных материалов, изделий и конструкций.

Библиогр.: 23 назв. Рис. 37. Табл. 20.

УДК 666.9.001.575 (042.4)

ББК 35.41в6

ISBN

© ГОУ ВПО «Уральский государственный
технический университет – УПИ», 2006

© В.Б. Пономарев,

А.Б.

Лошкарев,

2006

ОГЛАВЛЕНИЕ

1. ОБЩИЕ ПОЛОЖЕНИЯ.....	4
2. МОДЕЛИРОВАНИЕ МАТЕРИАЛЬНОЕ.....	7
3. МОДЕЛИРОВАНИЕ АНАЛОГОВОЕ.....	12
4. АНАЛИЗ РАЗМЕРНОСТЕЙ.....	13
5. МОДЕЛИРОВАНИЕ МАТЕМАТИЧЕСКОЕ.....	17
5.1. Классификация математических моделей.....	17
5.2. Вычислительный эксперимент.....	19
5.3. Основы информационных технологий на базе ПЭВМ.....	23
5.4. Примеры построения математических моделей.....	30
5.5. Вычислительные алгоритмы.....	36
6. СПОСОБЫ ПРОВЕДЕНИЯ ЭКСПЕРИМЕНТА.....	43
7. ЛИНЕЙНЫЕ РЕГРЕССИОННЫЕ МОДЕЛИ.....	48
8. ЗАДАЧИ ОПТИМИЗАЦИИ.....	58
8.1. Параметры и факторы оптимизации.....	58
8.2. Методы нахождения оптимума.....	62
8.3. Воспроизводимость и рандомизация опытов.....	65
8.4. Экспериментально–статистические модели.....	69
8.5. Полный факторный эксперимент.....	70
8.6. Дробный факторный эксперимент	86
9. МЕТОДЫ ОПТИМИЗАЦИИ.....	88
9.1. Метод крутого восхождения.....	88
9.2. Симплексный метод.....	93
9.3. Контурно–графический анализ.....	101
10. ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ.....	104
10.1. Основные понятия линейного программирования.....	104
10.2. Симплекс–метод линейного программирования.....	117
10.3. Моделирование планирования выпуска продукции.....	122
БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....	127

1. ОБЩИЕ ПОЛОЖЕНИЯ

Что такое модель?

МОДЕЛЬ (франц. *modèle*, итал. *modello*, от лат. *modulus* — мера, мерило, образец, норма):

- образец, служащий эталоном (стандартом) для серийного или массового воспроизведения (модель автомобиля, модель одежды и т. п.), а также тип, марка какого-либо изделия, конструкции;
- изделие (изготовленное из дерева, глины, воска, гипса и др.), с которого снимается форма для воспроизведения в другом материале (металле, гипсе и др.);
- человек, позирующий художнику (натурщик), и вообще изображаемые объекты («натура»);
- устройство, воспроизводящее, имитирующее (обычно в уменьшенном масштабе) строение и действие какого-либо другого устройства в научных, практических (например, в производственных испытаниях) или спортивных целях.

Перед тем как запустить в производство новый самолет, его обкатывают в аэродинамической трубе – это модель. Для того чтобы продемонстрировать систему кровообращения, лектор обращается к нарисованному плакату – это модель. На стене висит картина Айвазовского «Девятый вал» – это модель.

Другими словами, модель – это такой материальный или мысленно представляемый объект, который в процессе познания замещает объект – оригинал, сохраняя некоторые важные его черты.

Модель – это упрощенная система, отражающая отдельные стороны явления изучаемого объекта. Каждый изучаемый процесс можно описать различными моделями, при этом ни одна модель не может сделать это абсолютно полно и всесторонне. Однако использование упрощенной модели, отражающей отдельные черты исследуемого объекта, позволяет яснее увидеть взаимосвязь причин и следствий, входов и выходов, быстрее сделать необходимые выводы, принять правильные решения.

Таким образом, модель нужна для того, чтобы:

- 1) понять, как устроен объект, его структуру, свойства, законы взаимодействия с окружающим миром;
- 2) научиться управлять объектом или процессом, определить наилучшие способы управления;
- 3) прогнозировать последствия воздействий на объект.

Процесс построения модели называется *моделированием*. Условно можно выделить *материальное* и *идеальное* моделирование [3].

Материальным (физическим) моделированием принято называть моделирование, при котором реальному объекту противопоставляется увеличенная или уменьшенная копия, изученные свойства которой переносятся на объект при помощи теории подобия.

При материальном моделировании исследование объекта происходит при его воспроизведении в ином масштабе. Здесь возможен количественный перенос результатов эксперимента с модели на оригинал. Однако для анализа сложных объектов и процессов, каковыми являются большинство электронных схем, конструкций и технологических процессов производства радиоэлектронной техники, приборостроения, машиностроения и других промышленных отраслей, применение материального моделирования затруднительно, поскольку приходится использовать большое число критериев и ограничений, которые могут быть несовместимы, а зачастую и невыполнимы.

Примерами материального моделирования являются макеты, механические модели.

Идеальным моделированием называется моделирование, при котором реальному объекту противопоставляется описание его в форме речи, графики, таблиц, математических выражений.

К идеальным моделям относятся [23,24]:

- *модели словесные*. Речь является уникальной системой кодирования информации. С помощью речи можно описать любые предметы и процессы, однако это можно сделать только при помощи человека, то есть эти модели являются

субъективными. Построить по словесному описанию действующую модель практически невозможно;

- *модели графические.* Рисунки, чертежи и блок–схемы содержат большой объем информации, но и они являются статическими моделями, оживающими только через восприятие их человеком;

- *функциональные модели* описывают функции, выполняемые основными составными частями предприятия. Они разрабатываются для того, чтобы получить общее представление о процессе. Для примера рассмотрим общий план части типового цементного завода (рис. 1.1). Назначением этой части завода является получение однородного материала определенного химического состава с соответствующими размерами зерен для подачи его в сушильную печь. Сырье подается из хранилища в сырьевую мельницу, смешивается в гомогенизаторе и отправляется в сушильную печь. Таким образом, рисунок и описание процесса составляют вместе функциональную модель;

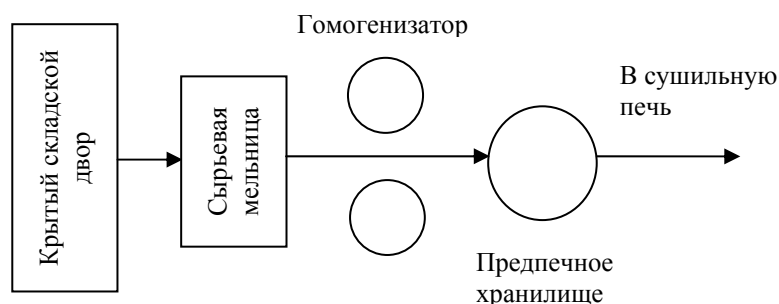


Рис. 1.1. Общий план участка

- *модели математические.* Математическое моделирование является методом качественного или количественного описания объектов или процессов, при этом реальный объект, процесс или явление упрощается, схематизируется и описывается определенным уравнением. В большинстве случаев математическая модель представляет собой уравнение регрессии, то есть геометрическое место точек математических ожиданий условных распределений целевой функции.

2. МОДЕЛИРОВАНИЕ МАТЕРИАЛЬНОЕ

Материальным (физическим) является вид моделирования, который состоит в замене изучения некоторого объекта или явления экспериментальным исследованием его модели, имеющей ту же физическую природу.

В науке любой эксперимент, производимый для выявления тех или иных закономерностей изучаемого явления или для проверки правильности и границ применимости найденных теоретическим путем результатов, по существу представляет собою моделирование, т. к. объектом эксперимента является конкретная модель, обладающая необходимыми физическими свойствами, а в ходе эксперимента должны выполняться основные требования, предъявляемые к физическому моделированию.

В технике физическое моделирование используется при проектировании и сооружении различных объектов для определения на соответствующих моделях тех или иных свойств (характеристик) как объекта в целом, так и отдельных его частей.

К физическому моделированию прибегают не только по экономическим соображениям, но и потому, что натурные испытания очень трудно или вообще невозможно осуществить, когда слишком велики (малы) размеры натурального объекта или значения других его характеристик (давления, температуры, скорости протекания процесса и т. п.).

В основе физического моделирования лежат теория подобия и анализ размерностей [9, 20]. Необходимыми условиями физического моделирования являются геометрическое подобие (подобие формы) и физическое подобие модели и природы: в сходственные моменты времени и в сходственных точках пространства значения переменных величин, характеризующих явления для природы, должны быть пропорциональны значениям тех же величин для модели. Наличие такой пропорциональности позволяет производить пересчет экспериментальных результатов, получаемых для модели, на природу путем умножения каждой из определяемых величин на постоянный для всех величин данной размерности множитель — коэффициент подобия.

Геометрическим называется подобие, при котором соблюдается равенство отношений всех сходственных линейных размеров модели и натуры. Безразмерные множители называют *константами* подобия. Например, если размеры одного треугольника a, b, c , а подобного ему a', b', c' , то:

$$a/a' = b/b' = c/c' = k_1 = \text{const.}$$

Таким образом, *константой подобия* называется безразмерное отношение сходственных размеров подобных фигур.

Кроме константы подобия различают еще *инварианту подобия*, которой является безразмерное отношение каких-либо двух размеров одной из фигур, равное отношению сходственных размеров другой подобной фигуры:

$$a/b = a'/b' = i_1 = \text{inv.}$$

Физическое подобие предполагает постоянное отношение физических свойств натуры и модели, например плотность, вязкость и др.

Поскольку физические величины связаны определенными соотношениями, вытекающими из законов и уравнений физики, то, выбрав некоторые из них за основные, можно коэффициенты подобия для всех других производных величин выразить через коэффициенты подобия величин, принятых за основные.

Например, в механике основными величинами считают обычно длину l , время t и массу m . Тогда, поскольку скорость $v = l/t$, коэффициент подобия скоростей $k_v = v_n/v_m$ (индекс «н» у величин для натуры, «м» — для модели) можно выразить через коэффициенты подобия длин $k_l = l_n/l_m$ и времен $k_t = t_n/t_m$ в виде:

$$k_v = \frac{l_n / t_n}{l_m / t_m} = \frac{l_n t_m}{l_m t_n} = \frac{k_l}{k_t}.$$

Поскольку на основании второго закона Ньютона сила F связана с ускорением w соотношением $F = mw$, то $k_F = k_m \cdot k_w$ (где, в свою очередь, $k_w = k_v/k_t$) и т. д. Из наличия таких связей вытекает, что для данного физического явления некоторые безразмерные комбинации величин, характеризующих это явление, долж-

ны иметь для модели и природы одно и то же значение. Эти безразмерные комбинации физических величин называются критериями подобия. Равенство всех критериев подобия для модели и природы является необходимым условием физического моделирования. Однако добиться этого равенства можно не всегда, т.к. не всегда удастся одновременно удовлетворить всем критериям подобия.

Чаще всего к физическому моделированию прибегают при исследовании различных механических (включая гидроаэромеханику и механику деформируемого твердого тела), тепловых и электродинамических явлений. При этом число и вид критериев подобия для каждого моделируемого явления зависит от его природы и особенностей.

Приведем несколько примеров критериев подобия.

Рейнольдса число (названо по имени О. Рейнольдса), для течений вязких жидкостей и газов, характеризующее соотношение между инерционными силами и силами вязкости:

$$Re = \rho v l / \mu,$$

где ρ – плотность, кг/м³;

μ – динамический коэффициент вязкости жидкости или газа, Н/(с·м²);

v – характерная скорость потока, м/с;

l – характерный линейный размер, м.

Так, при течении в круглых цилиндрических трубах обычно принимают $l = d$, где d – диаметр трубы, а $v = v_{cp}$, где v_{cp} – средняя скорость течения; при обтекании тел l – длина или поперечный размер тела, а $v = v_{\infty}$, где v_{∞} – скорость невозмущенного потока, набегающего на тело.

От числа Рейнольдса зависит также режим течения жидкости, характеризуемый $Re_{кр}$. При $R < Re_{кр}$ возможно лишь ламинарное течение жидкости, а при $Re > Re_{кр}$ течение может стать турбулентным. Значение $Re_{кр}$ зависит от вида течения. Например, для течения вязкой жидкости в круглой цилиндрической трубке $Re_{кр} = 2300$.

Фруда число, применяемое в случаях, когда существенно воздействие силы тяжести (в гидроаэромеханике, например при движении твердых тел в воде, в динамической метеорологии). Число Фруда характеризует соотношение между инерционной силой и силой тяжести, действующими на элементарный объем жидкости или газа:

$$Fr = v^2/gl,$$

где v – скорость течения (или скорость движущегося тела), м/с;

g – ускорение силы тяжести, м/с²;

l – характерный размер потока или тела, м.

Введено в 1870 г. английским ученым У. Фрудом. Условие подобия — равенство числа Фруда для модели и для натуральных объектов — применяют при моделировании движения кораблей, течений воды в открытых руслах, испытаниях моделей гидротехнических сооружений и др.

Нуссельта число (по имени немецкого физика В. Нуссельта), безразмерный параметр, характеризующий интенсивность конвективного теплообмена между поверхностью тела и потоком газа (жидкости):

$$Nu = \alpha l/\lambda,$$

где α – коэффициент теплообмена, Вт/(м²К), $\alpha = Q/(\Delta T \cdot S)$;

Q – количество тепла, отдаваемого (или получаемого) поверхностью тела в единицу времени, Вт;

ΔT – разница между температурой поверхности тела и температурой газа (жидкости) вне пограничного слоя, К, $\Delta T > 0$;

S – площадь поверхности, м²;

l – характерный размер, м;

λ – коэффициент теплопроводности газа, Вт/(м·К).

Фурье число (название по имени Ж. Фурье), для нестационарных тепловых процессов. Характеризует соотношение между скоростью изменения тепловых условий в окружающей среде и скоростью перестройки поля температуры внут-

ри рассматриваемой системы (тела), зависит от размеров тела и коэффициента его температуропроводности. Число Фурье обозначают F_0 и определяют формулой

$$F_0 = at_0/l^2,$$

где a – коэффициент температуропроводности, m^2/c , $a = \lambda/(\rho \cdot c)$;

ρ – плотность, kg/m^3 ;

c – удельная теплоемкость, $Dж/(kg \cdot K)$;

l – характерный линейный размер тела, m ;

t_0 – характерное время изменения внешних условий, c .

Когда при физическом моделировании необходимо обеспечить равенство нескольких критериев, возникают значительные трудности, часто непреодолимые, если только не делать модель тождественной натуре, что фактически означает переход от физического моделирования к натурным испытаниям. Поэтому на практике нередко прибегают к приближенному моделированию, при котором часть процессов, играющих второстепенную роль, или совсем не моделируется, или моделируется приближенно.

3. МОДЕЛИРОВАНИЕ АНАЛОГОВОЕ

Это один из важнейших видов моделирования, основанный на аналогии явлений, имеющих различную физическую природу, но описываемых одинаковыми математическими (дифференциальными, алгебраическими или какими-либо другими) уравнениями [4].

Простой пример — две системы, первая из которых, имеющая механическую природу, состоит из оси, передающей вращение через пружину и маховик, погруженный частично в вязкую тормозящую жидкость, валу, жестко связанному с маховиком. Вторая система — электрическая — состоит из источника электродвижущей силы, соединенного через катушку индуктивности, конденсатор и активное сопротивление со счетчиком электрической энергии. Если подобрать значения индуктивности, емкости и сопротивления так, чтобы они определенным образом соответствовали упругости пружины, инерции маховика и трению

жидкости, то эти системы обнаружат структурное и функциональное сходство (даже тождество), выражаемое, в частности, в том, что они будут описываться одним и тем же дифференциальным уравнением с постоянными коэффициентами вида

$$a \frac{d^2 z}{dt^2} + b \frac{dz}{dt} + c \cdot z = \omega .$$

Это уравнение может служить «теоретической моделью» обеих систем, любая же из них — «экспериментальной моделью» этого уравнения и «аналоговой моделью» друг друга. Эта аналогия лежит в основе электрического моделирования механических систем: электрические модели гораздо более удобны для экспериментального исследования, нежели моделируемые механические.

В настоящее время значение аналогового моделирования значительно уменьшилось, поскольку моделирование на ЭВМ имеет большие преимущества перед ним в отношении точности моделирования и универсальности. Но в достаточно фиксированных и специальных задачах аналоговое моделирование имеет свои преимущества (простота, а тем самым и дешевизна технического выполнения).

4. АНАЛИЗ РАЗМЕРНОСТЕЙ

Планирование эксперимента заключается в том, чтобы получить максимальный объем информации при наименьших затратах на эксперимент. Самым известным способом добиться компактности плана эксперимента является анализ размерностей – метод установления связи между физическими величинами, существенными для изучаемого явления, основанный на рассмотрении размерностей этих величин.

Теорема Бэкингема

Для правильного применения анализа размерностей необходимо знать количество фундаментальных переменных.

Фундаментальной переменной называется любая величина, оказывающая влияние на эксперимент и способная изменяться независимо от других переменных.

Первая часть теоремы Бэкингема гласит: «Если какое-либо уравнение однородно относительно размерностей, то его можно преобразовать к соотношению, содержащему набор безразмерных комбинаций величин».

Однородным относительно размерностей является уравнение, форма которого не зависит от выбора основных единиц. Например, в уравнение

$$F = \frac{\pi \cdot d^3}{6} \cdot \rho \cdot g$$
 можно подставлять значения в [м], и [мм], и [см], тогда как в

уравнение $q = A \cdot \exp(T)$ необходимо подставлять температуру только в градусах Кельвина, то есть оно неоднородно относительно размерностей.

Вторая часть теоремы Бэкингема (ПИ – теорема): «Если существует однозначное соотношение $f(A_1, A_2, \dots, A_n) = 0$ между n – физическими величинами, для описания которых используется k основных единиц, то существует также соотношение $f(\Pi_1, \Pi_2, \dots, \Pi_{n-k}) = 0$ между $(n - k)$ безразмерными комбинациями, составленными из этих физических величин». Многие размерные системы могут иметь несколько решений, и хотя все решения правильны, ценность их неодинакова, определяется физическим смыслом полученных комбинаций и во многом зависит от субъективного опыта экспериментатора.

Решение теоремы Бэкингема осуществляется методом последовательного исключения размерностей, или его еще называют «методом Ипсена» [20]. Сущность метода заключается в следующем:

– выбрать независимые переменные, оказывающие влияние на систему. Необходимо рассматривать также размерные коэффициенты и физические константы, если они играют важную роль (например, ускорение свободного падения). Это наиболее ответственный этап метода Ипсена;

– выбрать систему основных размерностей;

– записать размерности выбранных независимых переменных и составить безразмерные комбинации. Решение будет правильным, если каждая комбинация является *безразмерной*, число комбинаций не *меньше предсказанного Пи-теоремой*, каждая переменная *встречается в комбинациях хотя бы один раз*;

– изучить полученные комбинации с точки зрения физического смысла.

Рассмотрим на примере получение комбинаций безразмерных величин. Если, например, согласно опытным данным, некоторая величина x зависит от параметров y, z, s, w , то общий вид зависимости между данными величинами:

$$f(x, y, z, s, w) = 0 \quad \text{или} \quad x = f(y, z, s, w).$$

Определим, например, скорость v , с которой упадет на землю свободно падающее с высоты h тело массы m . Поскольку искомая величина может зависеть от ускорения свободного падения g , высоты h и массы m , то выражение для v можно представить в виде

$$v = C \cdot h^x g^y m^z, \quad (4.1)$$

где C – некоторая безразмерная постоянная;

x, y, z – числа, подлежащие определению.

Приравниваем размерности левой и правой частей (4.1):

$$LT^{-1} = L^x (LT^{-2})^y M^z.$$

Показатели степеней L, M, T в левой и правой частях должны быть равны, поэтому:

$$L: \quad 1 = x + y;$$

$$T: \quad -1 = -2y;$$

$$M: \quad 0 = z.$$

Отсюда $z = 0, y = 1/2, x = 1/2$, и формула (4.1) принимает вид:

$$v = C \sqrt{gh}.$$

Истинное значение скорости $v = \sqrt{2gh}$, то есть анализ размерностей дал возможность определить характер зависимости v от g, h, m с точностью до числового множителя C .

В следующем примере определим время опоздания студента на первую пару. Предположим, что время опоздания студента O [Т] зависит от следующих факторов:

S – продолжительность сна [Т], равная разности между временем подъема и временем отхода ко сну. (Чем меньше продолжительность сна, тем больше вероятность «проспать» – опоздать на первую пару);

Q – количество съеденного вечером калорийного ужина [М] (чем больше это количество, тем крепче сон);

P – атмосферное давление [М/(Л·Т)], определяющее сонливость при низком давлении;

t_1 – комнатная температура внутри помещения [Θ];

t_2 – температура на улице [Θ]. (Разность температур определяет «нежелание» выходить из теплого помещения на холод);

V – средняя скорость транспорта [Л/Т].

Общая зависимость выразится в виде:

$$O = S^a Q^b P^c t_1^d t_2^e V^f.$$

Подставим вместо переменных их размерности:

$$T^1 = T^a M^b (M/L \cdot T)^c \Theta^d \Theta^e (L/T)^f.$$

Напишем уравнения для каждой размерной величины, причем делитель пойдет со знаком минус:

$$T: \quad 1 = a - c - f;$$

$$M: \quad 0 = b + c;$$

$$L: \quad 0 = -c + f;$$

$$\Theta: \quad 0 = d + e.$$

Преобразуем равенства:

$$d = -e; \quad c = f; \quad b = -c = -f; \quad a = 1 + c + f = 1 + 2 \cdot f.$$

Подставим в основное уравнение:

$$O = S^{(1+2f)} Q^{-f} P^f t_1^d t_2^{-d} V^f = S^{(1)} S^{2f} Q^{-f} P^f (t_1/t_2)^d V^f$$

или

$$O = S \left(\frac{S^2 PV}{Q} \right)^f \cdot \left(\frac{t_1}{t_2} \right)^d.$$

Изучим полученное уравнение с точки зрения физического смысла. Поскольку время опоздания студента должно быть тем больше, чем меньше скорость движения транспорта, степень f должна иметь $\text{sign} = -1$. Сигнатура степени d соответствует физическому смыслу. Окончательно:

$$O = S \left(\frac{S^2 PV}{Q} \right)^{-f} \cdot \left(\frac{t_1}{t_2} \right)^d.$$

Таким образом, время опоздания студента на первую пару пропорционально количеству съеденного калорийного ужина, обратно пропорционально продолжительности сна, обратно пропорционально атмосферному давлению, обратно пропорционально средней скорости движения общественного транспорта и прямо пропорционально симплексу температур.

5. МОДЕЛИРОВАНИЕ МАТЕМАТИЧЕСКОЕ

5.1. Классификация математических моделей

Рассмотрим классификацию математических моделей.

1. Модели прогноза, или расчетные модели без управления

Основное назначение этих моделей – дать прогноз о поведении системы во времени и в пространстве, зная начальное состояние и информацию о поведении ее на границе.

Примерами могут служить модели распределения тепла, электрического поля, химической кинетики, гидродинамики.

2. Оптимизационные модели:

- Стационарные модели используются на уровне проектирования различных технологических систем.

- Динамические – как на уровне проектирования, так и, главным образом, для оптимального управления различными процессами – технологическими, экономическими и др.

В задачах оптимизации имеется два направления.

К первому относятся *детерминированные* задачи. Вся входная информация в них является полностью определяемой.

Второе направление относится к *стохастическим* процессам. В этих задачах некоторые параметры носят случайный характер или содержат элемент неопределенности.

Методы отыскания экстремума функции многих переменных с различными ограничениями часто называются методами математического программирования.

Этапы математизации знаний

Современная методология науки выделяет три этапа математизации знаний:

- математическая обработка эмпирических (экспериментальных) данных;
- моделирование;
- относительно полные математические теории.

Первый этап – это математическая, чаще всего именно количественная обработка эмпирических (экспериментальных) данных. Это этап выявления и выделения чисто функциональных феноменологических взаимосвязей (корреляций) между входными сигналами (входами) X_i и выходными реакциями (откликами) Y_i на уровне целостного объекта (явления, процесса), которые наблюдают в экспериментах с объектами–оригиналами. Данный этап математизации имеет место во всякой науке и может быть определен как этап первичной обработки ее эмпирического материала.

Второй этап математизации знаний определим как модельный. На этом этапе математизации, т.е. этапе математического моделирования, осуществляется попытка теоретического воспроизведения некоторого интересующего нас объекта–оригинала в форме другого объекта – математической модели.

Третий этап – это этап относительно полной математической теории данного уровня организации материи в данной или рассматриваемой предметной области. Третий этап предполагает существование логически полной системы понятий и аксиом. Математическая теория дает методологию и язык, пригодные для описания явлений, процессов и систем различного назначения.

5.2. Вычислительный эксперимент

Ни одно техническое достижение не повлияло так на интеллектуальную деятельность человека, как электронно–вычислительные машины. Увеличив в десятки и сотни миллионов раз скорость выполнения арифметических и логических операций, колоссально повысив тем самым производительность интеллектуального труда человека, ЭВМ вызвали коренные изменения в области обработки информации. В начале XXI века компьютеры стали настолько совершенными, что появилась реальная возможность использовать их в научных исследованиях не только в качестве больших арифмометров, но и обратиться с их помощью к изучению таких разделов математики, которые ранее были практически не доступны для исследований. Это было осознано еще при решении на несовершенных ЭВМ сложных математических задач ядерной физики, баллистики, прикладной небесной механики.

Основой *вычислительного эксперимента* [15, 17, 21] является математическое моделирование, теоретической базой – прикладная математика, а технической – мощные электронно–вычислительные машины.

Научное исследование реального процесса можно проводить теоретически или экспериментально независимо друг от друга. Такой путь познания истины носит односторонний характер. В современных условиях развития науки и техники стараются проводить комплексное исследование объекта.

Вычислительный эксперимент – это эксперимент над математической моделью объекта на ЭВМ, который состоит в том, что по одним параметрам модели вычисляются другие ее параметры и на этой основе делаются выводы о свойствах явления, описываемого математической моделью.

Основные этапы вычислительного эксперимента:

- проведение натурального эксперимента;
- построение математической модели;
- выбор и применение численного метода для нахождения решения;
- обработка результатов вычислений;
- сравнение с результатами натурального эксперимента;
- принятие решения о продолжении натуральных экспериментов;
- продолжение натурального эксперимента для получения данных, необходимых для уточнения модели;
- накопление экспериментальных данных;
- построение математической модели;
- автоматическое построение программной реализации математической модели;
- автоматизированное нахождение численного решения;
- автоматизированное преобразование вычислительных результатов в форму, удобную для анализа;
- принятие решения о продолжении натуральных экспериментов.

Видоизмененная цепочка, реализованная в виде единого программного комплекса, и составляет «технология» вычислительного эксперимента.

Сферы применения вычислительного эксперимента и математического моделирования

В современной науке и технике появляется все больше областей, задачи в которых можно и нужно решать методом вычислительного эксперимента, с помощью математического моделирования. Обратим внимание на некоторые из них.

- *Энергетическая проблема.* Прогнозирование атомных и термоядерных реакторов на основе детального математического моделирования происходящих в них физических процессов. В этой области работа ведется очень успешно. Вычислительный эксперимент тесно сопрягается с натурным экспериментом и по-

могает, заменяет и удешевляет весь исследовательский цикл, существенно его ускоряя.

- *Космическая техника.* Расчет траекторий летательных аппаратов, задачи обтекания, системы автоматического проектирования. Обработка данных натурального эксперимента, например радиолокационных данных, изображений со спутников, диагностика плазмы. Здесь очень важной оказывается проблема повышения качества приборов, и в частности измерительной аппаратуры. Между тем в настоящее время показано, что, используя измерительный прибор среднего качества и присоединив к нему ЭВМ, можно на основе специальных алгоритмов получить результаты, которые дал бы измерительный прибор очень высокого качества. Таким образом, сочетание измерительного прибора с компьютером открывает новые возможности.

- *Технологические процессы.* Получение кристаллов и пленок, которые, кстати, нужны для создания вычислительной техники, для решения проблем в области элементарной базы (что невозможно без математического моделирования); моделирование теплового режима конструктивных узлов перспективных ЭВМ, процессов лазерной плазмы, технологии создания материалов с заданными свойствами (это одна из основных задач любой технологии).

- *Экологические проблемы.* Вопросы прогнозирования и управления экологическими системами могут решаться лишь на основе математического моделирования, поскольку эти системы существуют в «единственном экземпляре».

- *Гео- и астрофизические явления.* Моделирование климата, долгосрочный прогноз погоды, землетрясений и цунами, моделирование развития звезд и солнечной активности, фундаментальные проблемы происхождения и развития Вселенной.

- *Химия.* Расчет химических реакций, определение их констант, исследование химических процессов на макро- и микроуровне для интенсификации химической технологии.

- *Биология.* Особо следует отметить интерес к математическому моделированию в связи с изучением фундаментальных проблем этой науки (генетики, морфогенеза) и разработкой новых методов биотехнологии.

К основным преимуществам вычислительного эксперимента можно отнести следующие:

- возможность исследования объекта без модификации установки или аппарата;
- возможность исследования каждого фактора в отдельности, в то время как в реальности они действуют одновременно;
- возможность исследования нереализуемых на практике процессов.

Вычислительный эксперимент включает в себя следующие этапы (рис. 5.1):

- 1) физическое описание процесса, то есть уяснение закономерности протекаемых явлений;
- 2) разработка математической модели;
- 3) алгоритм или метод решения уравнений;
- 4) разработка программ;
- 5) проведение расчетов, анализ результатов и оптимизация.

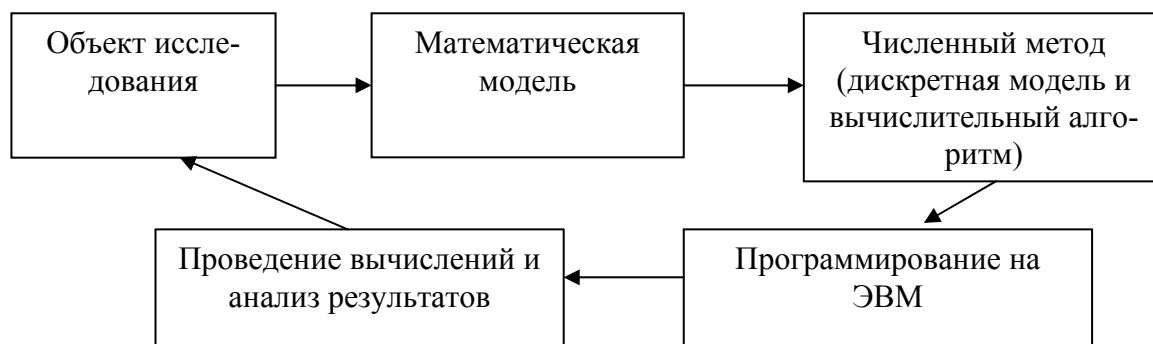


Рис. 5.1. Схема вычислительного эксперимента

Тем самым основу вычислительного эксперимента составляет триада: *модель – алгоритм – программа*. Опыт решения крупных задач показывает, что метод математического моделирования и вычислительный эксперимент соединяют в себе преимущества традиционных теоретических и экспериментальных методов исследования.

5.3. Основы информационных технологий на базе ПЭВМ

Любая задача с применением ПЭВМ решается следующим образом

(рис. 5.2):

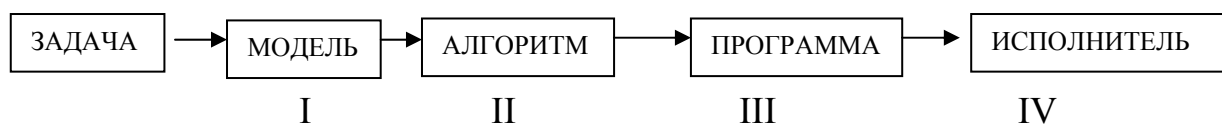


Рис. 5.2. Последовательность решения задачи

На основе реального объекта строится модель, которая ограничивается критериями подобия.

При решении задачи по классической технологии необходимо довести модель до математической модели. При этом, если исследователь – «инженер–технолог», то для решения задачи на I и II этапах требуется профессиональный математик, а на III и IV – программист–системотехник. Такой стиль работы называют процедурным.

НИТ (новая информационная технология) предусматривает другой стиль работы – средоориентированный, когда используется пакет прикладных программ [19, 22].

Пакеты прикладных программ

Пакет прикладных программ (ППП) состоит из функционального наполнения и системной части. Функциональное наполнение представляет собой, грубо говоря, набор отдельных программ, решающих конкретные задачи. Эти задачи объединены одной направленностью, или, как говорят, предметной областью. Дело в том, что ППП не является универсальным, он проблемно–ориентирован, т.е. предназначен для решения определенного класса задач.

Если это задачи механики сплошной среды, то в функциональное наполнение могут входить, например, программы для расчета уравнений газовой ди-

намики, уравнений теплопроводности, уравнений для электромагнитного поля, уравнений для излучения, фазовых переходов и т.д.

Содержание каждой такой индивидуальной программы, или «модуля», специфично, однако требования к оформлению входной и выходной информации унифицированы. Эти модули представляют собой своеобразные «черные ящики», которые можно соединять в цепочки, ветви так, чтобы в конце концов получить заданную программу.

Системная часть выполняет функции сервисного характера. Основные задачи здесь состоят в следующем. Прежде всего необходимо организовать хранение функционального наполнения. Но хранить, в данном случае, – не значит ограничиться записью информации на каких-либо носителях. В этом архиве должен быть порядок: по первому требованию указанный модуль должен быть направлен «в работу».

Главное назначение системной части ППП – обеспечивать возможность сборки из отдельных модулей полной программы, способной решать заданную задачу. Для этого вычислитель, создающий программу, должен общаться с пакетом – давать приказы, воспринимать ответную информацию.

Конечно, это очень упрощенная схема работы с пакетом, но она отражает характерные этапы такой деятельности.

Кроме того, для того чтобы пользоваться пакетом и, значит, грамотно вести расчеты, совсем не обязательно самому обладать высокой квалификацией программиста или математика-вычислителя (ведь именно они должны создавать эти пакеты). Поэтому пакеты программ должны быть такими, чтобы к их помощи могли прибегнуть не только математики, но и специалисты других сфер научной деятельности, прошедшие сравнительно небольшой курс математического обучения.

Кратко ППП можно определить как набор программ, которые адекватно покрывают определенную предметную область. Общение с машиной при этом сводится к двум этапам:

- 1) объективное построение среды (подбор ППП под определенную задачу);

2) программная настройка среды (чаще всего происходит автоматически).

СРЕДА должна при этом:

- обеспечивать хранение информации;
- иметь определенный внешний вид (пользователь должен сразу догадаться, какую задачу можно решать в данной среде);
- обладать определенным поведением (пользователь должен сразу понять, как себя вести в данной среде).

Любой ППП обладает высокоразвитым языком, и когда нас не устраивает внешний вид среды и ее поведение, мы можем настраивать среду, то есть менять ее вид и поведение.

Типы сред:

- операционные системы;
- операционные оболочки;
- инструментальные средства;
- текстовые процессоры;
- электронные таблицы;
- графические пакеты;
- математические пакеты;
- системы управления базами данных;
- системы программирования;
- системы управления базами знаний;
- экспертные системы;
- игры
- и т.д.

Операционная система – это определенный набор программ, которые выполняют самые простые функции для обслуживания решения задач, задают способ хранения информации [12] (например, MS DOS (дискетная операционная система Майкрософт). Система работает с файлами, представляющими собой набор информации, занимающими определенную длину (размер в байтах) и имеющими

уникальное имя и расширение. Файловая система – система распределения памяти в вычислительной машине. Один лист машинописного текста, набранного в текстовом редакторе DOS, занимает примерно 2 килобайта памяти). Каждый файл содержит информацию о времени и дате своего рождения. Файлы можно копировать, перемещать, переименовывать, удалять и редактировать. Сама по себе оперативная система имеет модульный характер, то есть состоит из нескольких программ. Ядро системы – та часть, которая постоянно находится в оперативной памяти, транзиты – перемещаемые модули ОС. Оперативную систему можно представить в виде следующей схемы (рис. 5.3):

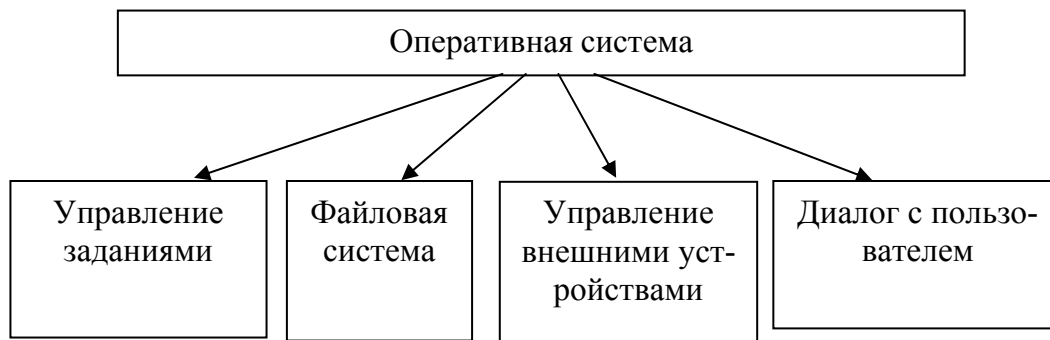


Рис. 5.3. Схема оперативной системы

Файловая система требует управления. Программа управления файловой системой называется MS.SYS. Управление внешними устройствами складывается из физического и логического управления. Физическое управление осуществляется аппаратно (контроллерами устройств). Логическое – специальными программами – драйверами устройств. Для этого предназначена программа IO.SYS (система ввода – вывода input–output sys). В файле IO.SYS находится базовая система ввода вывода BIOS DOS, которая включает в себя набор стандартных драйверов ввода–вывода. В этом файле содержатся драйверы:

- накопитель на гибких магнитных дисках;
- накопитель на жестких магнитных дисках (винчестерах);
- дисплей и клавиатура;
- принтер и принтерные порты;

- последовательные порты;
- пустое устройство (NUL);
- системные часы
- и др.

С помощью команды DEVICE в файле конфигурации DOS можно добавить свои драйверы:

- расширенный драйвер дисплея ANSI.SYS;
- драйвер псевдодиска VDISK.SYS – устройство, эквивалентное жесткому диску, но находящееся в оперативной памяти);
- драйвер мышки и т.д.

После включения компьютера и загрузки команд IO.SYS и MSDOS.SYS компьютер все еще не в состоянии воспринимать команды от пользователя. Для возможности войти в контакт с компьютером на «естественном» языке предназначен файл COMMAND.COM.

Для удобства работы несколько файлов, относящихся к одному ППП, могут иметь одинаковое имя, но различаться расширением, которое отделяется от имени точкой и может содержать три символа (например, файлы с расширением .exe, .com, .bat – выполняемые, относятся к программам, остальные означают отношение к файлам помощи, текстовым файлам, файлам библиотек и баз данных и т.д.).

Операционные оболочки – это среды, которые позволяют выполнять действия на уровне ОС, но имеют более наглядный и удобный интерфейс (связь с пользователем).

До девяностых годов прошлого века самой распространенной оболочкой была Norton Commander, представляющая два окна, в которых отражаются все имена файлов с их размером и датой рождения. Под окнами располагается набор команд (меню), позволяющих выполнять все основные действия над файлами по команде одной из функциональных клавиш. С 1996 года начался повальный пе-

переход на windows – оболочки, которые, при наличии мультимедиа, обеспечивают наиболее комфортное общение с машиной.

Инструментальные средства – это пакеты, которые позволяют решить проблемы технического характера на уровне машины. К ним относятся антивирусные программы, программы проверки и лечения жесткого диска, программы восстановления потерянных файлов и папок, программы дефрагментации диска и т.п.

Текстовые процессоры – пакеты программ для подготовки и оформления документов (Блокнот, WordPad, Microsoft Word и др.).

Электронные таблицы. Удобны, когда на модельном уровне наша задача сводится к таблицам, причем, изменяя какое-либо одно значение, можно пересчитать всю таблицу. Самый распространенный процессор – Microsoft Excel.

Графические пакеты используются для построения графики (рисунки, чертежи, эскизы, дизайн, мультипликация, рекламные ролики и т.п.). Работа со сканером, плоттером, графопостроителем.

Математические пакеты – группа программ, которые удобно применять, когда наша модель сводится к математической модели.

Системы управления базами данных – упорядочивание большого объема информации, фильтрация, выборка информации.

Системы программирования позволяют программировать на языке машины. Преобразование мнемочкодов (языка буквенных кодов) в машинные коды осуществляет специальная программа – ассемблер, являющаяся языком низкого уровня программирования. Языки высокого уровня позволяют писать программы понятными словами удобно и просто, но работают гораздо медленнее, чем ассемблер, состоящий из набора «шаманских» символов, которые трудно выучить. Первым языком высокого уровня был ФОРТРАН (formula translation) для решения бухгалтерских программ. Очень распространены языки турбопascal, предназначенный для написания не очень больших программ, и СИ⁺⁺ – для профессиональных программистов.

В общем виде все вышесказанное можно представить в виде следующей схемы (рис. 5.4):



Рис. 5.4. Структура ПЭВМ

5.4. Примеры построения математических моделей

Исследование любого реального объекта (явления природы, производственного процесса, экономического планирования и т.д.) начинается с формализации объекта, с построения математической модели: выделяются наиболее существенные черты и свойства и описываются с помощью математических соотношений.

Например, необходимо определить площадь поверхности письменного стола. Определяем ширину и длину стола и перемножаем, то есть реальный объект – поверхность стола – заменяется абстрактной математической моделью – прямоугольником. Площадь такого прямоугольника приближенно принимается за исследуемую площадь. Однако человеческий глаз как измерительный инструмент не отличается высокой точностью, поэтому модель нужно проверить, например, следующим образом: измерить все противоположные стороны и диагонали, и если они попарно равны, можно принять модель прямоугольника, в про-

тивном случае получится модель четырехугольника. При более высокой точности можно учесть закругления стола.

Рассмотрим задачу по механике (рис. 5.5). Камню на Земле сообщили начальную скорость V_0 под углом α к ее поверхности. Необходимо найти траекторию камня и вычислить расстояние до его падения.

Допустим, что камень вылетает при помощи катапульты, что позволяет уточнить характерные размеры камня, его начальную скорость, массу. Зададимся граничными условиями:

- Земля – инерционная система отсчета;
- ускорение свободного падения g постоянно;
- кривизной Земли пренебрегаем;
- действием воздуха на движущийся камень пренебрегаем.

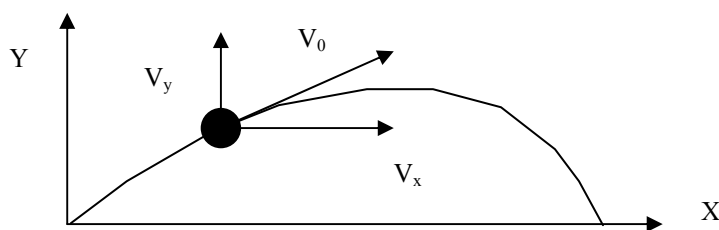


Рис. 5.5. Проекция скорости движения камня

При сделанных предположениях проекция камня на ось X будет двигаться равномерно со скоростью $V_x = V_0 \cdot \cos\alpha$, проекция на ось Y будет двигаться равноускоренно с ускорением $a_y = -g$ и начальной скоростью $V_y = V_0 \cdot \sin\alpha$, тогда:

$$\left\{ \begin{array}{l} X = t \cdot V_x = t \cdot V_0 \cdot \cos\alpha \\ Y = t \cdot V_y - \frac{gt^2}{2} = t \cdot V_0 \cdot \sin\alpha - \frac{gt^2}{2} \end{array} \right\}.$$

Выразим время t через координату $X(t)$:

$$t = \frac{X}{V_0 \cos\alpha}$$

и подставим в выражение $Y(t)$:

$$Y = \frac{X \cdot V_0 \sin \alpha}{V_0 \cos \alpha} - \frac{g}{2} \left(\frac{X}{V_0 \cos \alpha} \right)^2 = X \cdot \operatorname{tg} \alpha - \frac{g}{2} \left(\frac{X}{V_0 \cos \alpha} \right)^2.$$

Эта парабола пересекает ось Y в следующих случаях: при $X = 0$

$$\text{и при } X = \frac{2 \cdot \operatorname{tg} \alpha}{g} (V_0 \cos \alpha)^2 = \frac{V_0^2}{g} 2 \cdot \sin \alpha \cdot \cos \alpha = \frac{V_0^2}{g} \cdot \sin 2\alpha.$$

Таким образом, длина падения камня $L = \frac{V_0^2}{g} \cdot \sin 2\alpha$.

Соответствие математической модели изучаемому объекту

Математическая модель никогда не бывает тождественна рассматриваемому объекту, не передает всех его свойств и особенностей. Она является приближенным описанием объекта и носит всегда приближенный характер. Точность соответствия определяется степенью соответствия, адекватности модели и объекта.

При построении математической модели приходится выдвигать дополнительные предположения – гипотезы. Модель поэтому еще называют *гипотетической*. Основным критерием применимости модели является эксперимент. Критерий практики позволяет сравнивать гипотетические модели и выбирать из них наиболее подходящую.

Например, нужно установить пределы применимости нашей модели с камнем. Рассмотрим силы воздействия на камень (рис. 5.6).

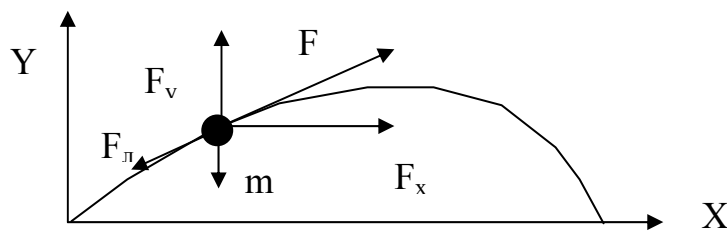


Рис. 5.6. Проекции сил, действующих на камень

Лобовое сопротивление действует в сторону, противоположную движению:

$$F_{\text{л}} = \xi S \rho V^2,$$

где ξ – коэффициент лобового (аэродинамического) сопротивления;

S – площадь миделевого сечения, м²;

ρ – плотность воздуха, кг/м³;

V – скорость движения камня, м/с.

Причем ξ зависит от формы тела и характеристики процесса обтекания (числа Re): $\xi = f(Re)$,

$$Re = \frac{V d \rho}{\mu},$$

где μ – динамическая вязкость воздуха, Па·с.

Катапультной камень можно забросить на расстояние до 100 м, с высотой траектории 20÷30 м и скоростью до 30 м/с. Диаметр камня может достигать 200 мм. Для нормальных условий ($t = 15$ °С, $\rho = 1,3$ кг/м³, $\mu = 1,7 \cdot 10^{-5}$ Па·с) оценим величину Re :

$$Re = \frac{30 \cdot 0,2 \cdot 1,3}{1,7 \cdot 10^{-5}} = 4,6 \cdot 10^5.$$

На основании теоретических и экспериментальных исследований для шара при $3 \cdot 10^5 < Re \leq 7 \cdot 10^6$ коэффициент лобового сопротивления $\xi = 0,15$.

Полагая $S = \frac{\pi d^2}{4}$, силу можно рассчитать по уравнению $F_{\text{л}} = \xi \frac{\pi}{4} d^2 \rho V^2$

(квадратичная зависимость от скорости).

Сравним эту силу с основной силой в рассматриваемой задаче – с силой тяжести:

$$P = mg = \frac{\pi d^3}{6} \rho_{\text{ч}} g,$$

где $\rho_{\text{ч}}$ – плотность материала камня ($\rho_{\text{ч}} = 2300$ кг/м³).

Составим отношение:

$$\frac{F_{л}}{P} = \frac{\xi \frac{\pi}{4} d^2 \rho V^2}{\frac{\pi d^3}{6} \rho_{ч} g} = \frac{3}{2} \xi \frac{\rho}{\rho_{ч}} \frac{V^2}{dg} \quad (5.1)$$

Поскольку ранее мы определили $l = \frac{V_0^2}{g} \text{Sin}2\alpha$ (при $\alpha = 45^\circ$, $\text{Sin}2\alpha = 1$),

тогда:

$$l = \frac{V_0^2}{g} \quad \text{или} \quad V_0^2 = lg.$$

Подставим полученные соотношения в формулу 5.1:

$$\frac{F_{л}}{P} = \frac{3}{2} \xi \frac{\rho}{\rho_{ч}} \frac{V^2}{dg} = \frac{3}{2} \xi \frac{\rho}{\rho_{ч}} \frac{l}{d}.$$

При $l = 100$ м, $d = 0,2$ м, $\xi = 0,15$ получим $\frac{F}{l} = 0,03$.

Обозначим через Δl абсолютную погрешность дальности броска, тогда при заданных условиях и дальности 100 м $\Delta l = 2 \div 3$ м.

Развитие и уточнение математической модели

Рассмотрим применимость нашей модели с камнем к пушечному ядру ($l = 1000$ м, $d = 0,14$ м, $\rho_{ч} = 7000$ кг/м³ (плотность чугуна), $\xi = 0,15$).

Отношение $\frac{F}{l} = 0,15$ (то есть ошибка в определении дальности порядка 15 %, или 150 м).

Для того чтобы применить эту модель к пушечному ядру, нужно вернуться к граничным условиям – последний пункт сформулировать таким образом: воздух воздействует на ядро с силой $F_{л}$, противоположной скорости.

Решение задачи при этом осуществляется численными методами.

Переход от гладкоствольного к нарезному оружию позволил увеличить дальность стрельбы при скоростях снаряда, превышающих скорость звука.

При этом для сверхзвуковых скоростей зависимость $\xi = f(V)$, причем она резко возрастает с увеличением скорости снаряда. Это еще более усложняет задачу расчетов дальности стрельбы.

Дальнейшее увеличение скорости тела привело к необходимости ввести еще одно уточнение – учет вращения Земли вокруг своей оси. В северном полушарии все движущиеся тела отклоняются вправо, в Южном – влево (правый берег реки обычно подмывается сильнее левого, а на ж/д колеях быстрее снашивается правая сторона рельса по ходу движения поезда).

5.5. Вычислительные алгоритмы

Алгоритм – это конечный набор правил, позволяющих чисто механически решать любую конкретную задачу из некоторого класса однотипных задач [8]. При этом подразумевается:

- исходные данные могут изменяться в определенных пределах (*массовость алгоритма*);
- процесс применения правил к исходным данным (путь решения задачи) определен однозначно (*детерминированность алгоритма*);
- на каждом шаге процесса применения правил известно, что следует считать результатом этого процесса (*результативность алгоритма*).

В большинстве задач математическая модель выражается в виде формулы. Формула определяет последовательность математических операций, которые нужно выполнить для вычисления исходной величины. Однако известны задачи, для которых ответ легко может быть найден, хотя он не описывается в виде формулы (например, правило поразрядного сложения чисел столбиком).

Если модель описывает зависимость между исходными данными и искомыми величинами, то *алгоритм представляет собой последовательность действий, которые надо выполнить, чтобы от исходных данных перейти к искомым величинам.*

Удобной формой записи алгоритма является блок – схема. Она не только достаточно наглядно описывает алгоритм, но и представляет собой основу для составления программы. Каждый класс математических моделей имеет свой метод решения, который реализуется в алгоритме. Поэтому очень важной является классификация задач по виду математической модели. При таком подходе задачи, различные по содержанию, можно решать с помощью одного и того же алгоритма. Алгоритм записывают с помощью обычных математических символов. Для того чтобы он мог быть прочитан ЭВМ, необходимо составить программу. Программа – это описание алгоритма решения задачи, заданное на языке ЭВМ. Алгоритмы и программы объединяются понятием «математическое обеспечение».

Алгоритмы задач принятия решений, как правило, настолько сложны, что без применения ЭВМ реализовать их практически невозможно.

Например, число π определялось при помощи вычисления периметров вписанного и описанного многоугольников диаметром $d = 1$. С ростом числа сторон периметры вписанных многоугольников растут, а описанных – убывают.

$$\lim_{n \rightarrow \infty} (p_{\text{вп}}) = \lim_{n \rightarrow \infty} (p_{\text{оп}}) = \pi .$$

Архимед дошел в вычислениях до 96 – угольника, получив следующую

оценку: $3 \frac{10}{71} \leq \pi \leq 3 \frac{1}{7} .$

В первой половине XV века придворный астроном хана Улукбека Аль-Каши вычислил π с 17 знаками после запятой [8]. Он дошел до $6 \cdot 2^{27}$ –угольника. К концу XIX века английский математик В. Шенкс вычислил 707 знаков числа π , затратив на это более 20 лет. Однако в 1945 г. было обнаружено, что В. Шенкс допустил ошибку на 520–м знаке, и все его последующие вычисления пошли насмарку. В настоящее время вычисление π с точностью до 500000 знаков с применением ЭВМ осуществляется за несколько минут. Данный пример показывает решение задачи бесконечного сходящегося процесса. Алгоритмы решения таких задач называют *вычислительными алгоритмами*, а основанные на них методы решения математических задач – *численными*

на них методы решения математических задач – *численными* методами. До появления ЭВМ численные методы использовались крайне редко в силу чрезвычайной трудоемкости вычислений.

Нахождение корня непрерывной функции численными методами

Для примера рассмотрим алгоритмы решения уравнений [7]. Методы решения линейных и квадратных уравнений были известны еще древним грекам. Однако уравнения пятой и более степеней неразрешимы через коэффициенты с помощью арифметических действий и извлечения корней.

Пусть дана непрерывная функция $f(x)$ и требуется найти корень уравнения $f(x) = 0$. Предположим, что найден отрезок $[a, b]$, такой, что $f(a) \cdot f(b) < 0$. Тогда, согласно теореме Больцано–Коши¹, внутри отрезка $[a, b]$ существует точка «с», в которой значение функции равно нулю.

Метод «вилки», или **метод бисекций**, заключается в построении последовательно вложенных друг в друга отрезков, на концах которых функция принимает значения разных знаков. Каждый последующий отрезок получают делением пополам предыдущего. Процесс построения последовательности отрезков позволяет найти нуль функции $f(x)$ с любой заданной точностью. Опишем один шаг итераций.

Пусть на $(n - 1)$ шаге найден отрезок:

$$[a_{n-1}, b_{n-1}] \subset [a, b], \text{ такой, что } f(a_{n-1}) \cdot f(b_{n-1}) < 0.$$

Делим его пополам точкой $c = (a_{n-1} + b_{n-1})/2$ и вычисляем $f(c)$. Если $f(c) = 0$, то «с» – корень уравнения. Если $f(c) \neq 0$, то из двух половин отрезка выберем ту, на концах которой функция имеет противоположные знаки, так как один из корней лежит на этой половине. Таким образом,

$$\begin{aligned} a_n = a_{n-1}, \quad b_n = c, & \quad \text{если } f(c) \cdot f(a_{n-1}) < 0, \\ a_n = c, \quad b_n = b_{n-1}, & \quad \text{если } f(c) \cdot f(a_{n-1}) > 0. \end{aligned}$$

¹ Теорема о существовании корня непрерывной функции. Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и принимает на его концах значения разных знаков, то на этом отрезке существует, по крайней мере, один корень уравнения.

Если требуется найти корень с точностью ε , то деление продолжается до тех пор, пока длина отрезка не станет меньше 2ε . Тогда координата середины отрезка и есть значение корня с требуемой точностью. Метод бисекций («вилки») – простой и надежный метод для поиска простого корня уравнения $f(x) = 0$. Он сходится для любых непрерывных функций, в том числе недифференцируемых. Скорость сходимости невелика. Для достижения точности ε необходимо совер-

шить N итераций, где $N \cong \log_2 \frac{b-a}{\varepsilon}$. Если на

отрезке $[a,b]$ находится несколько корней уравнения, то процесс нахождения корня сходится к одному из них.

Метод простых итераций

Метод простых итераций (последовательных приближений) решения уравнения $f(x) = 0$ состоит в замене исходного уравнения эквивалентным ему уравнением $x = \varphi(x)$ и построении последовательности $x_{n+1} = \varphi(x_n)$, сходящейся при $n \rightarrow \infty$ к точному решению. Сформулируем достаточные условия сходимости метода простых итераций.

Теорема. Пусть функция $\varphi(x)$ определена и дифференцируема на $[a,b]$, причем все ее значения $\varphi(x) \in [a,b]$. Тогда, если существует число q , такое, что $|\varphi'(x)| \leq q < 1$ на отрезке $[a,b]$, то последовательность $x_{n+1} = \varphi(x_n)$, $n = 0, 1, 2, \dots$, сходится к единственному на $[a,b]$ решению уравнения $x = \varphi(x)$ при любом начальном значении $x_0 \in [a,b]$, т.е.:

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \varphi(x_n) = c, \quad f(c) = 0, \quad c \in [a,b].$$

При этом, если на отрезке $[a,b]$ производная $\varphi'(x)$ положительна, то:

$$|c - x_n| \leq \frac{q}{1-q} |x_n - x_{n-1}|,$$

если $\varphi'(x)$ отрицательна, то:

$$|c - x_n| \leq |x_n - x_{n-1}|.$$

Опишем один шаг итераций. Исходя из найденного на предыдущем шаге значения x_{n-1} , вычисляем $y = \varphi(x_{n-1})$. Если $|y - x_{n-1}| > \varepsilon$, полагаем $x_n = y$ и выполняем очередную итерацию.

Если же значение $|y - x_{n-1}| < \varepsilon$, то вычисления заканчивают и за приближенное значение корня принимают величину $x_n = y$. Погрешность полученного результата зависит от знака производной $\varphi'(x)$. При $\varphi'(x) > 0$ корень найден с погрешностью $\frac{q\varepsilon}{1-q}$; если $\varphi'(x) < 0$, то погрешность не превышает ε .

Метод допускает прямую геометрическую интерпретацию. Построим графики функций $y = x$, $y = \varphi(x)$. Корнем уравнения $y = \varphi(x)$ является абсцисса точки пересечения кривой $y = \varphi(x)$ с прямой $y = x$ (рис. 5.7).

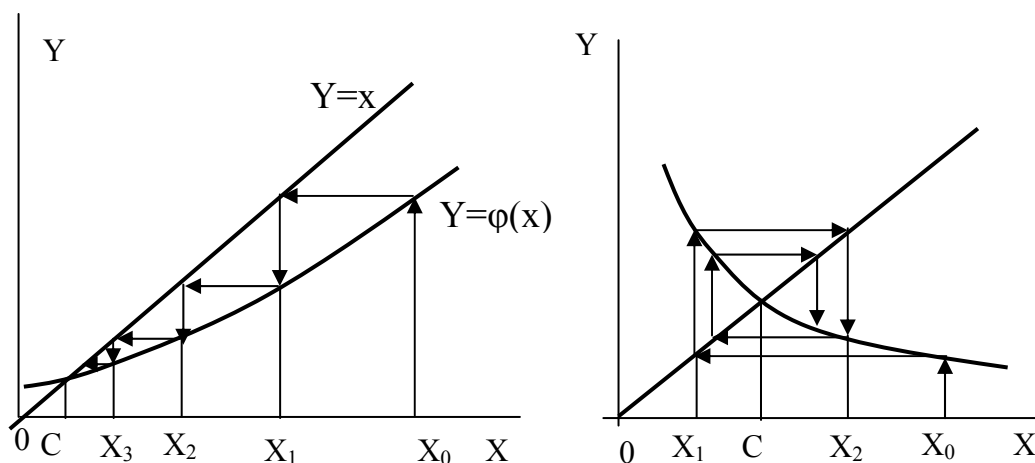


Рис. 5.7. Метод простых итераций

Взяв в качестве начальной произвольную точку $x_0 \in [a, b]$, строим ломаную линию. Абсциссы вершин этой ломаной представляют собой последовательные приближения корня «С». Скорость сходимости к корню тем выше, чем меньше число q .

Метод Ньютона (метод касательных)

Метод состоит в построении итерационной последовательности:

$$x_{n+1} = x_n - f(x_n)/f'(x_n),$$

сходящейся к корню уравнения $f(x) = 0$.

Сформулируем достаточные условия сходимости метода.

Теорема. Пусть $f(x)$ определена и дважды дифференцируема на $[a,b]$, причем $f(a) \cdot f(b) < 0$, а производные $f'(x)$, $f''(x)$ сохраняют знак на отрезке $[a,b]$. Тогда, исходя из начального приближения $x_0 \in [a,b]$, удовлетворяющего неравенству $f(x_0) \cdot f''(x_0) > 0$, можно построить последовательность

$x_{n+1} = x_n - f(x_n)/f'(x_n)$, $n=0,1,2,\dots$, сходящуюся к единственному на $[a,b]$ решению уравнения $f(x)$.

Метод Ньютона допускает простую геометрическую интерпретацию. Если через точку с координатами $(x_n; f(x_n))$ (рис. 5.8) провести касательную, то абсцисса точки пересечения этой касательной с осью «Ох» и есть очередное приближение x_{n+1} корня уравнения $f(x) = 0$.

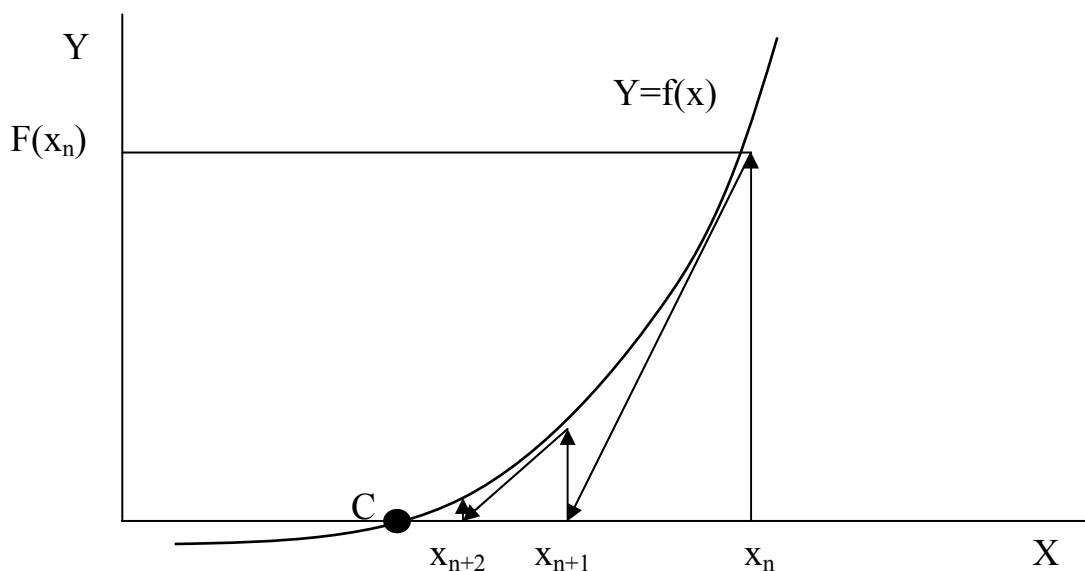


Рис. 5.8. Метод касательных

Если необходимо найти корень с точностью ε , то итерационный процесс можно прекращать, когда:

$$|x_n - x_{n-1}| \leq \varepsilon_0 = \sqrt{\frac{2m_1\varepsilon}{M_2}},$$

где m_1 – наименьшее значение модуля первой производной $|f'(x)|$ на отрезке $[a,b]$,

M_2 – наибольшее значение модуля второй производной $|f''(x)|$ на отрезке $[a,b]$.

Опишем один шаг итераций. Если на $(n - 1)$ шаге очередное приближение x_{n-1} не удовлетворяет условию окончания процесса, то вычисляем величины $f(x_{n-1})$, $f'(x_{n-1})$ и следующее приближение корня:

$$x_n = x_{n-1} - f(x_{n-1})/f'(x_{n-1}).$$

При выполнении условия $|x_n - x_{n-1}| \leq \sqrt{\frac{2m_1 \varepsilon}{M_2}}$ величину x_n принимаем

за приближенное значение корня C , вычисленное с точностью ε .

6. СПОСОБЫ ПРОВЕДЕНИЯ ЭКСПЕРИМЕНТА

Активный эксперимент состоит в целенаправленном изменении входных параметров технологического процесса. В основе этого метода лежит планирование эксперимента. Активный эксперимент позволяет за счет целенаправленного изменения входных параметров получать необходимый объем информации при существенно меньшем числе опытов, чем при пассивном эксперименте. Различают следующие активные методы эксперимента.

Эмпирический метод

При использовании эмпирических методов математическое описание составляется следующим образом:

- 1) проводятся эксперименты методом «черного ящика», т.е. изучается реакция объекта на различные возмущения;
- 2) осуществляется статистическая обработка результатов и поиск наилучшей формы аппроксимации полученных данных;
- 3) строится математическое описание.

Единственным критерием применимости полученного математического описания является наибольшая простота уравнений при хорошей аппроксимации экспериментальных данных.

Достоинства:

- простота описания;

- доступность получения моделей;
- возможность построения модели при отсутствии теории процесса.

Недостатки:

- невозможность применения модели для режимов, в которых не проводились измерения;
- невозможность применения модели при переходе к другим установкам;
- невозможность экстраполяции результатов.

Рассмотрим пример (рис. 6.1). Модель строилась для значений в интервале $[a, b]$. Получена квадратичная зависимость «2». Видно, что в интервале $[a, b]$ модель хорошо описывает процесс, протекающий в оригинале, экспериментальная зависимость $Y = f(X)$ отображается кривой «1». При выходе величины значения X за пределы отрезка $[a, b]$ модель (кривая «2») дает значительные погрешности.

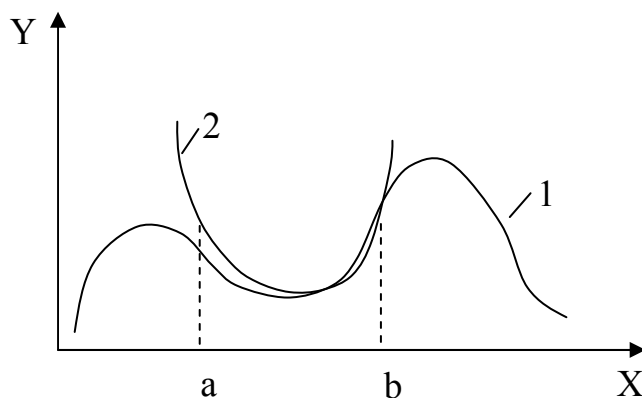


Рис. 6.1. Аппроксимация экспериментальных данных

Эмпирические методы полезны и применимы для изучения сложных систем, если их структура не изменяется во времени, теория процесса неизвестна и (или) когда необходимо быстро получить модель без исследования процесса.

Экспериментально – аналитический метод

При использовании этого метода исследователь пытается определить физическую сущность явлений, протекающих в объекте. Используется декомпозиция сложного явления, т.е. на основе анализа определяются более простые, элементарные процессы, которые можно исследовать более доступными способами. После анализа влияния элементарных процессов на процесс в целом несущест-

венные факторы отбрасываются и выбирается тот элементарный процесс, который оказывает наиболее существенное влияние. Затем составляется математическое описание, причем не в форме полинома, а в виде зависимости, которая характерна для данного элементарного процесса. Влияние остальных элементарных процессов учитывается посредством изменения коэффициентов, входящих в эту зависимость.

В качестве примера рассмотрим построение модели для описания процесса переноса тепла в неподвижном зернистом слое в вертикальном направлении (рис. 6.2).

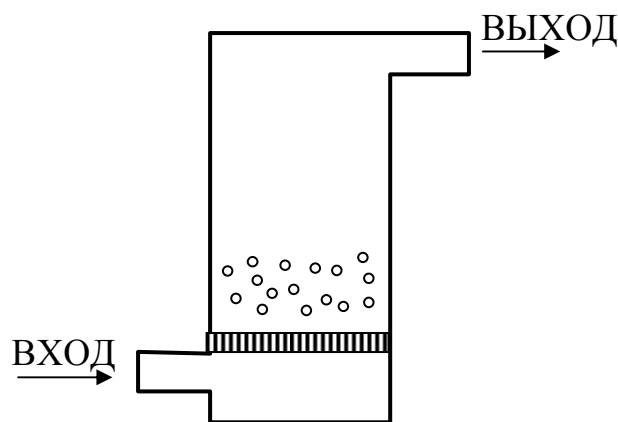


Рис. 6.2. Установка «кипящего» слоя

Процесс переноса тепла осуществляется за счет следующих процессов: теплопроводности, теплопередачи и излучения.

При температурах менее 800 К и малых линейных скоростях потока газа перенос тепла в основном определяется теплопроводностью. Этот процесс описывается уравнением Фурье $q_t = \lambda \frac{dT}{dX}$.

$$q_t = \lambda \frac{dT}{dX}$$

Однако пользоваться этим уравнением еще нельзя, т.к. в нем не учтены теплопередача и излучение (остальные элементарные составляющие процесса переноса тепла). Для их учета вместо истинного значения вводится некоторое «эффективное» значение, которое определяется экспериментально, причем тогда

$$q_t = \lambda_{\text{эф}} \frac{dT}{dX}$$

Уравнение является экспериментально–аналитической моделью процесса переноса тепла в неподвижном зернистом слое.

Совершенно очевидно, что $\lambda_{\text{эф}}$ не является физической константой, а зависит от условий экспериментов, при которых она была получена, и от масштабов установки.

Достоинства: лучше описывает нелинейные свойства объекта моделирования, т.к. позволяет более надежно выбирать вид уравнения.

Недостатки: эффективные коэффициенты изменяются в зависимости от условия проведения опытов, поэтому экспериментально – аналитическая модель справедлива лишь в том интервале, в котором производился эксперимент.

Сопоставим эмпирический и экспериментально–аналитический методы построения математических моделей.

Экспериментально–аналитический метод имеет преимущество перед чисто экспериментальным в том, что он отражает теорию процесса. Для учета влияния явлений, не учтенных при составлении модели, вводятся эффективные коэффициенты. В первом методе эксперимент необходим для получения модели, во втором – для определения коэффициентов модели.

Теоретический метод

Этот метод предполагает составление математического описания на основе детального изучения и глубокого понимания физических и химических закономерностей процессов, протекающих в нем. Составленное на основе этого метода математическое описание дает возможность с большей точностью предсказывать результаты протекания процесса в заданных нами условиях.

Теоретический метод – наиболее надежный способ составления математического описания.

В математическое описание объекта входят представленные ниже составляющие (рис. 6.3).

Материальные и энергетические балансы составляются на основе закона сохранения энергии и массы: «приход» – «расход» = «накопление».

Ограничения могут быть обусловлены технологическими, техническими или экономическими причинами.



Рис. 6.3. Математическое описание объекта

Достоинства: возможность широкой экстраполяции, разделение сложного процесса на отдельные составляющие и исследование процесса по частям облегчает составление модели процесса в целом, возможность изучения процесса на разных уровнях.

Недостатки: трудность создания надежной теории сложных процессов, невозможность использования при неизвестном механизме процесса, большие затраты времени.

Выбор того или иного метода зависит от важности и степени сложности процесса. Для крупных многотоннажных производств необходимы хорошие модели, здесь применяют теоретический метод. Этим же методом пользуются при создании принципиально новых технологических процессов.

Для мелких производств со сложным характером процесса используют экспериментальный метод. На практике, как правило, используется разумное сочетание всех методов.

7. ЛИНЕЙНЫЕ РЕГРЕССИОННЫЕ МОДЕЛИ

Регрессионная модель для одной переменной управления

Разработку моделей установившихся процессов для действующих предприятий и проверку теоретических моделей, построенных на основе использования физических законов, можно осуществлять экспериментально.

Регрессионный анализ – это метод построения модели, наиболее соответствующий набору экспериментальных данных.

Под наилучшим соответствием понимается, что функция ошибки, являющаяся показателем разности между моделью и данными, должна быть минимизирована. Такой функцией ошибки обычно служит сумма квадратов ошибок (разностей между измеренным значением в данной точке и величиной, предсказанной в модели). Это называется подбором экспериментальных формул по методу наименьших квадратов.

На рис. 7.1 показаны n выборки экспериментальных данных (X_1, Y_1) , (X_2, Y_2) , ..., (X_n, Y_n) .

Допустим, что модель представляет собой прямую линию:

$$Y^p = A_0 + A_1 X,$$

где Y^p – величина, предсказываемая регрессионной моделью.

Требуется получить такие значения коэффициентов A_0 , A_1 , при которых сумма квадратов ошибок является минимальной. Ошибка E для каждой точки определяется как расстояние по вертикали от этой точки до прямой линии (модели).

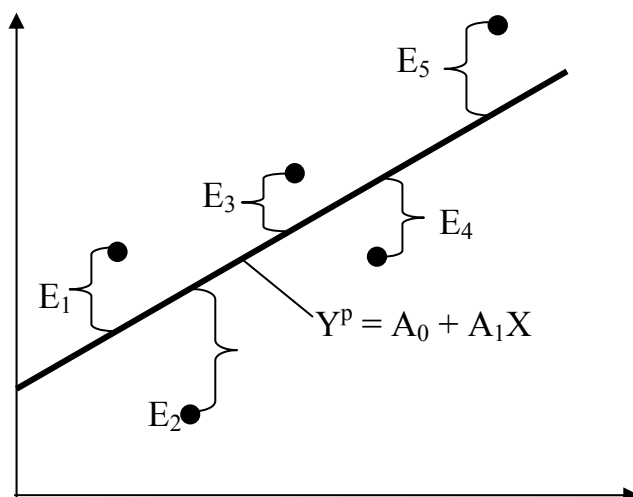


Рис. 7.1. Линия регрессии

Обозначим:

$$Y_1^p = A_0 + A_1 X_1,$$

$$Y_2^p = A_0 + A_1 X_2,$$

.....

$$Y_n^p = A_0 + A_1 X_n,$$

тогда ошибки будут выражаться в виде

$$E_1 = Y_1^p - Y_1 = A_0 + A_1 X_1 - Y_1,$$

$$E_2 = Y_2^p - Y_2 = A_0 + A_1 X_2 - Y_2,$$

.....

$$E_n = Y_n^p - Y_n = A_0 + A_1 X_n - Y_n.$$

Функция ошибки F определяется выражением $F = E_1^2 + E_2^2 + \dots + E_n^2$, или

$$F = \sum_{i=1}^n (A_0 + A_1 X_i - Y_i)^2.$$

Для получения таких значений A_0 и A_1 , при которых функция F является минимальной, применяются обычные методы математического анализа. Услови-

ям минимума являются $\frac{\partial F}{\partial A_0} = 0$ и $\frac{\partial F}{\partial A_1} = 0$.

Дифференцируя F, получаем

$$\begin{aligned} \frac{\partial F}{\partial A_0} &= \frac{\partial}{\partial A_0} \sum_{i=1}^n (A_0 + A_1 X_i - Y_i)^2 = \sum_{i=1}^n 2(A_0 + A_1 X_i - Y_i) = \\ &= 2(nA_0 + A_1 \sum X_i - \sum Y_i) = 0, \end{aligned}$$

откуда
$$nA_0 + \left(\sum_{i=1}^n X_i \right) A_1 + \sum_{i=1}^n Y_i = 0. \quad (7.1)$$

Аналогично

$$\begin{aligned} \frac{\partial F}{\partial A_1} &= \frac{\partial}{\partial A_1} \sum_{i=1}^n (A_0 + A_1 X_i - Y_i)^2 = \sum_{i=1}^n 2X_i (A_0 + A_1 X_i - Y_i) = \\ &= 2A_0 \sum_{i=1}^n X_i + 2A_1 \sum_{i=1}^n X_i^2 - 2 \sum_{i=1}^n Y_i X_i = 0, \end{aligned}$$

$$\text{откуда } \left(\sum_{i=1}^n X_i \right) A_0 + \left(\sum_{i=1}^n X_i^2 \right) A_1 = \sum_{i=1}^n X_i Y_i. \quad (7.2)$$

Решая систему двух линейных алгебраических уравнений (7.1) и (7.2), можно получить значения A_0 и A_1 . В матричном представлении эти уравнения имеют вид

$$\begin{pmatrix} \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i^2 \end{pmatrix} \cdot \begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix}. \quad (7.3)$$

Решая уравнение (7.3), получаем

$$A_0 = \frac{\sum_{i=1}^n Y_i \sum_{i=1}^n X_i^2 - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i X_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2},$$

$$A_1 = \frac{n \sum_{i=1}^n Y_i X_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2},$$

где n – число выборок экспериментальных данных.

Модели множественной линейной регрессии

Модель множественной линейной регрессии представлена уравнением

$$Y^p = A_0 + \sum_{j=1}^n A_j X_j.$$

Задача состоит в том, чтобы получить такие значения коэффициентов A_0, A_1, \dots, A_k , при которых сумма квадратов ошибок (разностей между данными, предсказываемыми регрессионной моделью, и выборкой из n экспериментальных данных) является минимальной. Функция ошибки при этом

$$F = (A_0 + A_1 X_{11} + A_2 X_{21} + \dots + A_k X_{k1} - Y_1)^2 + \\ + (A_0 + A_1 X_{12} + A_2 X_{22} + \dots + A_k X_{k2} - Y_2)^2 + \dots + \\ + (A_0 + A_1 X_{1n} + A_2 X_{2n} + \dots + A_k X_{kn} - Y_n)^2.$$

Минимизируя функцию F, положим

$$\frac{\partial F}{\partial A_0} = \frac{\partial F}{\partial A_1} = \dots = \frac{\partial F}{\partial A_k} = 0.$$

В матричном виде система линейных уравнений для определения коэффициентов модели имеет вид

$$\begin{pmatrix} n & \sum_{i=1}^n X_{1i} & \sum_{i=1}^n X_{2i} & \dots & \sum_{i=1}^n X_{ki} \\ \sum_{i=1}^n X_{1i} & \sum_{i=1}^n X_{1i}^2 & \sum_{i=1}^n X_{1i} X_{2i} & \dots & \sum_{i=1}^n X_{1i} X_{ki} \\ \sum_{i=1}^n X_{2i} & \sum_{i=1}^n X_{2i} X_{1i} & \sum_{i=1}^n X_{2i}^2 & \dots & \sum_{i=1}^n X_{2i} X_{ki} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{i=1}^n X_{ki} & \sum_{i=1}^n X_{ki} X_{1i} & \sum_{i=1}^n X_{ki} X_{2i} & \dots & \sum_{i=1}^n X_{ki}^2 \end{pmatrix} \cdot \begin{pmatrix} A_0 \\ A_1 \\ A_2 \\ \dots \\ A_k \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_{1i} Y_i \\ \sum_{i=1}^n X_{2i} Y_i \\ \dots \\ \sum_{i=1}^n X_{ki} Y_i \end{pmatrix},$$

где n – число экспериментальных точек;

i – номер точки.

Ошибки эксперимента

При измерении физических величин существует три основных источника ошибок:

- 1) основной чувствительный элемент неправильно отражает измеряемую величину (например, в спиртовой манометр залита питьевая вода);
- 2) неспособность индикатора правильно отражать реакцию чувствительного элемента (неправильная калибровка прибора);
- 3) неспособность наблюдателя правильно регистрировать показания прибора (если он заменил содержимое манометра, то может вообще снять показания с другого прибора).

Эти источники ошибок приводят к двум классам ошибок:

- случайным,
- систематическим.

Случайная ошибка – когда при последовательных измерениях постоянной величины каждый раз получаются разные числовые значения.

Систематическая ошибка – когда среднее значение последовательных отсчетов отклоняется от заранее известного точного значения на какую–либо постоянную величину.

Систематическая ошибка устраняется путем калибровки или ремонта прибора.

Для описания случайных ошибок применяют теории вероятностей.

Для характеристики частоты появления различных значений случайной величины X в теории вероятности применяют различные законы распределения случайной величины, причем во всех случаях кривая плотности распределения

вероятностей определяется соотношением
$$\int_{-\infty}^{+\infty} p(x) dx = 1.$$

Наиболее часто используемым в теории вероятностей является «нормальный» закон распределения (распределение Гаусса), плотность вероятности которого описывается выражением

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right].$$

Координата центра распределения может быть определена несколькими способами:

- *медиана* – такая точка на оси абсцисс, слева и справа от которой вероятности появления различных значений случайной величины равны друг другу и составляют 50 %;
- *мода* – только для симметричных распределений точка на оси абсцисс, имеющая максимальную плотность распределения;

- *математическое ожидание* – центр тяжести распределения, то есть такая точка на оси абсцисс, относительно которой опрокидывающий момент равен нулю или

$$\bar{X} = \int_{-\infty}^{+\infty} x \cdot p(x) dx.$$

Для дискретного распределения, например для описания разброса координат частиц, сыпавшихся с наклонной плоскости (рис. 7.2):

$$\bar{X} = \frac{1}{n} \sum_i x_i \cdot p(x),$$

где n – общее количество частиц;

$p(x)$ – количество частиц, попавших в i -ю ячейку;

x_i – расстояние до середины i -й ячейки.

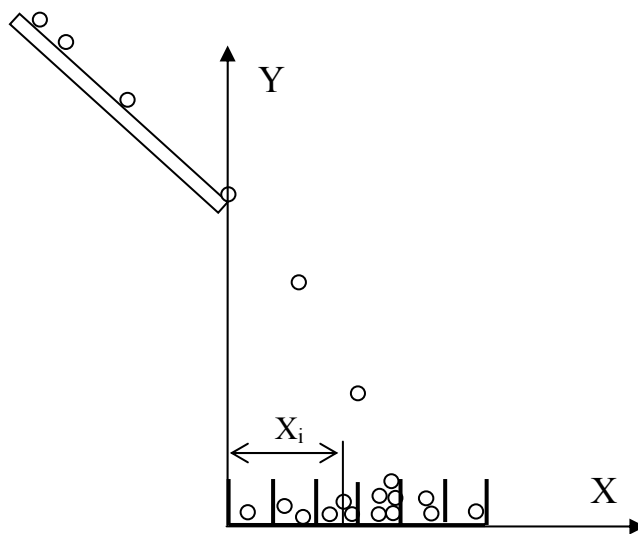


Рис. 7.2. Схема разброса координат частиц, сыпавшихся с наклонной плоскости

Если из всех наблюдавшихся значений погрешности вычесть систематическую составляющую, то есть перенести начало координат в центр распределения, то такое распределение называется *центрированным*.

Для описания различных свойств распределений используют такие параметры, как моменты, причем первый центральный момент называется *математическим ожиданием*.

Центральный момент k -го порядка для дискретной случайной величины выражается в виде

$$\mu_k = \frac{1}{n} \sum_i (x_i - \bar{X})^k \cdot p(x).$$

Второй центральный момент называется дисперсией случайной величины

$$D = \frac{1}{n} \sum_i (x_i - \bar{X})^2 \cdot p(x).$$

Для более наглядной характеристики самого рассеяния пользуются корнем квадратным из дисперсии – среднеквадратичным отклонением (СКО)

$$\sigma = \sqrt{D}.$$

Третий центральный момент характеризует асимметрию, то есть скошенность распределения, когда один спад – крутой, а другой – пологий. Для относительной характеристики асимметрии используют безразмерный коэффициент

асимметрии $S = \frac{\mu_3}{\sigma^3}.$

Четвертый центральный момент характеризует протяженность распределения, его относительное значение $\varepsilon = \frac{\mu_4}{\sigma^4}$ называют эксцессом. Не нужно путать его с коэффициентом эксцесса $\gamma = \varepsilon - 3.$

Для классификации распределений по форме удобнее использовать другую функцию – контрэксцесс $\chi = \frac{1}{\sqrt{\varepsilon}}.$

Одним из условий правомерности статистической обработки выборки является требование ее однородности. Отсчеты, резко отклоняющиеся по своим значениям от большинства других отсчетов, принято называть промахами и исключать из выборки. Наиболее предпочтительной является методика исключения промахов, предложенная Г.А. Агебяном [1], который рекомендует оценки

\bar{X} и σ определять без использования отсчетов, предполагаемых промахами, а границу цензурирования назначать в зависимости от объема выборки n при

$$\begin{aligned} 6 < n \leq 100 & & |X_{гр}| = 4\sigma; \\ 100 < n \leq 1000 & & |X_{гр}| = 4,5\sigma; \\ 1000 < n \leq 10000 & & |X_{гр}| = 5\sigma. \end{aligned}$$

Если промахи попадают в этот интервал, их включают в расчет оценки \bar{X} и σ и все заново пересчитывают.

Для определения формы распределения, медианы и других характеристик выборка должна быть представлена в виде гистограммы, состоящей из m столбцов с определенной протяженностью d соответствующих им интервалов. При большом m гистограмма будет отличаться от плавной кривой распределения вследствие изрезанности многими всплесками и провалами, при слишком малом m гистограмма будет отличаться от действительной кривой распределения вследствие слишком крупной ступенчатости, из-за чего характерные особенности будут просто потеряны.

При объеме выборки, равном n , и величине контрэксцесса χ наиболее предпочтительной для определения количества интервалов является формула И.У.Алексеевой [2]

$$m = \frac{1}{3} \sqrt[5]{\frac{n^2}{\chi^8}}.$$

Значение m принято выбирать нечетным, для того чтобы не искажалась середина кривой распределения.

Обычно мерой ошибки регрессионной модели служит стандартное (среднеквадратичное) отклонение s , определяемое по формуле

$$s = \sqrt{\frac{\sum_{i=1}^n E_i^2}{n-2}} = \sqrt{\frac{\sum_{i=1}^n (A_0 + A_1 X_i - Y_i)^2}{n-2}}.$$

Для нормально распределенных процессов приблизительно 66 % точек находится в пределах одного стандартного отклонения от модели и 95 % точек – в

пределах двух стандартных отклонений. Стандартное отклонение – важный показатель для решения вопроса о достоверности модели. Большая ошибка может означать, что модель не соответствует процессу, который послужил источником экспериментальных данных, однако она может быть вызвана и другой причиной – значительным разбросом данных измерений. В этом случае, возможно, потребуется взять большее количество выборок.

8. ЗАДАЧИ ОПТИМИЗАЦИИ

8.1. Параметры и факторы оптимизации

Оптимизация технологического процесса производства любой продукции содержит важный этап – определение (отыскание) математической модели или уравнения связи выходного показателя качества изделия (целевой функции, параметра оптимизации) с параметрами этого изделия или технологического процесса (входными факторами).

Поиск оптимальных условий является одной из наиболее распространенных научно–технических задач. Процесс решения этих задач называется процессом оптимизации или просто оптимизацией. Примером оптимизации является поиск оптимального состава многокомпонентных смесей или сплавов, повышение производительности или эффективности работы действующих установок, повышение качества продукции, снижение затрат на производство изделий и т.п.

Для описания объекта исследования используют схему «черного» ящика (рис. 8.1).

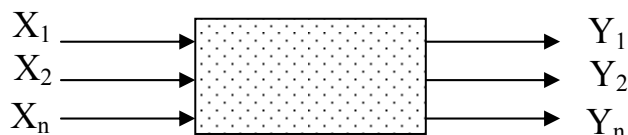


Рис. 8.1. Схема «черного» ящика

Стрелки справа изображают численные характеристики целей исследования, называемые еще *критериями оптимизации, целевыми функциями, функциями отклика*.

Стрелки слева называют *факторами*, то есть способами воздействия на объект.

Любое экспериментальное исследование содержит три этапа:

- 1) этап постановки задачи;
- 2) этап планирования и проведения эксперимента;
- 3) анализ и интерпретация результатов.

Главной трудностью на этапе постановки задачи является переход с языка специальности на язык планирования эксперимента, на язык математики. Построение математической модели технологического процесса в зависимости от поставленной задачи может преследовать следующие цели:

- минимизировать расход материала на единицу выпускаемой продукции при сохранении качества;
- произвести замену дорогостоящих материалов на более дешевые или дефицитных – на распространенные;
- сократить время обработки в целом или на отдельных операциях, перевести отдельные режимы в некритические зоны;
- снизить трудовые затраты на единицу продукции;
- улучшить частные показатели и общее количество готовой продукции;
- повысить однородность продукции;
- улучшить показатели надежности;
- увеличить надежность и быстродействие управления;
- увеличить эффективность контроля качества;
- создать условия для автоматизации процесса управления и т.п.

Параметры оптимизации

Прежде всего необходимо выбрать зависимую переменную Y , которую впредь будем называть целевой функцией или параметром оптимизации, за ко-

торый принимают один из показателей качества продукции либо по каждой технологической операции отдельно, либо по всему технологическому процессу сразу.

В зависимости от объекта и цели исследования параметры оптимизации могут быть самыми разными. Некоторая квалификация их представлена на следующей схеме (рис. 8.2).

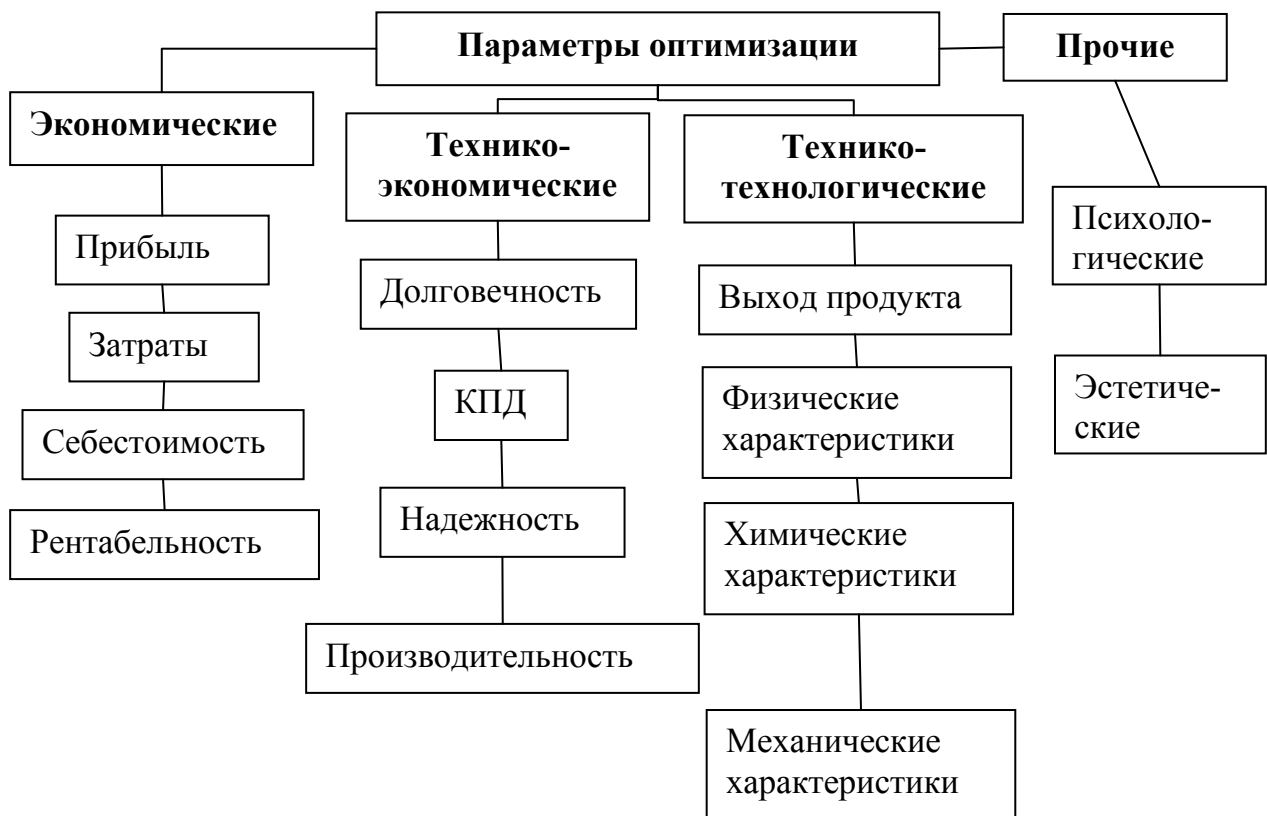


Рис. 8.2. Параметры оптимизации

Параметр оптимизации должен соответствовать следующим требованиям:

- параметр должен измеряться при любом изменении (комбинации) режимов технологического процесса;
- параметр должен быть статистически эффективным, то есть измеряться с наибольшей точностью;
- параметр должен быть информационным, то есть всесторонне характеризовать технологический процесс (операцию);

- параметр должен иметь физический смысл, то есть должна присутствовать возможность достижения полезных результатов при соответствующих условиях процесса;

- параметр должен быть однозначным, то есть должна минимизироваться или максимизироваться только одна целевая функция.

В тех случаях, когда оптимизироваться должны две целевые функции (P и W), их можно объединить в один параметр оптимизации посредством линейной комбинации. В этом случае целевая функция имеет вид

$$Y = \Psi_1 P + \Psi_2 W,$$

где Ψ_1 и Ψ_2 – весовые коэффициенты.

Факторы оптимизации

За фактор оптимизации принимают контролируемую величину объекта (изделия, процесса, операции), то есть величину, характеризующую то или иное свойство объекта или режим технологического оборудования. Эта величина, числовое значение которой измеряется в пределах (границах) изменения, должна влиять на параметр оптимизации.

Различают *качественные* факторы, к которым относятся различные вещества, технологические способы, аппараты и т.д., и *количественные*, которые можно оценить количественно.

Требования к факторам:

- должны быть управляемыми;
- иметь как можно большую точность замера;
- должны быть совместимы (все их комбинации осуществимы и безопасны);
- должны быть однозначны;
- должны быть независимы друг от друга.

8.2. Методы нахождения оптимума

Как ставить эксперимент, чтобы найти оптимум при минимуме затрат?

Существует несколько вариантов:

- перебор всех значений факторов – очень трудоемкая операция;
- случайный выбор некоторых состояний и определение откликов в надежде на оптимальный вариант. Таким способом можно попасть в оптимум быстро, но существует также вероятность перебирать их слишком долго;
- построение математической модели и предсказание по ней значений откликов, которые не изучались экспериментально.

Метод покоординатного спуска

Пусть нужно найти наименьшее значение целевой функции. Выберем какую-нибудь начальную точку M_0 (рис. 8.3) и зафиксируем значения всех факторов, кроме первого.

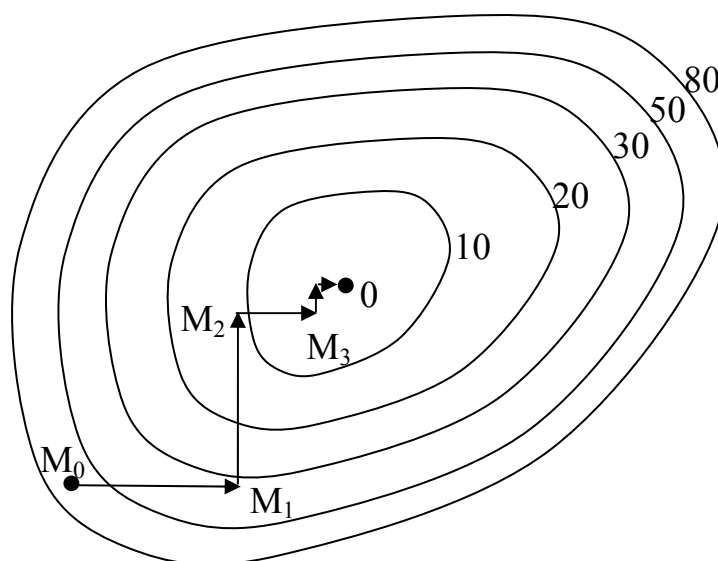


Рис. 8.3. Поиск наименьшего значения функции методом покоординатного спуска

Изменяя значения этого фактора, будем двигаться в сторону убывания функции, пока не дойдем до ее минимума. Обозначим эту точку M_1 . Фиксируем теперь все факторы, кроме второго, и движемся к минимуму функции (точка M_3). Фиксируем следующий фактор и так далее. Дойдя до последнего фактора, снова вернемся к первому фактору и продолжим поиск наименьшего значения функции.

Метод градиентного спуска

Рассмотрим функцию $f(x, y, z)$. Вычислим ее частные производные и образуем с их помощью вектор, который называют градиентом функции:

$$\text{grad}(f(x, y, z)) = \frac{\partial f}{\partial x}(x, y, z)\bar{i} + \frac{\partial f}{\partial y}(x, y, z)\bar{j} + \frac{\partial f}{\partial z}(x, y, z)\bar{k},$$

где \bar{i} , \bar{j} , \bar{k} – единичные векторы, параллельные координатным осям.

Направление градиента является направлением наиболее быстрого возрастания функции в данной точке. Противоположное направление называют антиградиентом. Метод градиентного спуска реализуется следующим образом.

Выберем начальную точку M_0 , вычисляем в этой точке градиент функции и делаем небольшой шаг в обратном антиградиентном направлении. В новой точке повторяем процедуру, пока не дойдем до точки минимума функции (рис. 8.4). Метод градиентного спуска требует вычисления градиента целевой функции на каждом шаге. Если функция задана аналитически, то это, как правило, не проблема: для частных производных, определяющих градиент, можно получить явные формулы. В противном случае частные производные приходится вычислять приближенно, заменяя их соответствующими разностными отноше-

ниями $\frac{\partial f}{\partial x_i} \approx \frac{f(x_1, \dots, x_i + \Delta x_i, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{\Delta x_i}$.

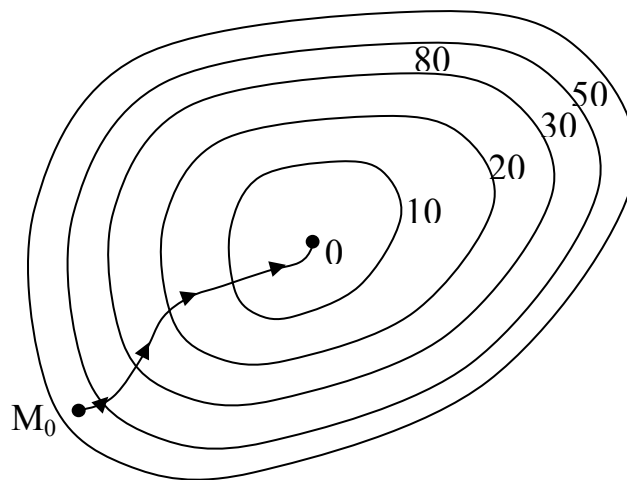


Рис. 8.4. Поиск наименьшего значения функции методом градиентного спуска

Проблема многоэкстремальности

На рис. 8.5 приведены линии уровня функции с двумя локальными минимумами в точках O_1 и O_2 .

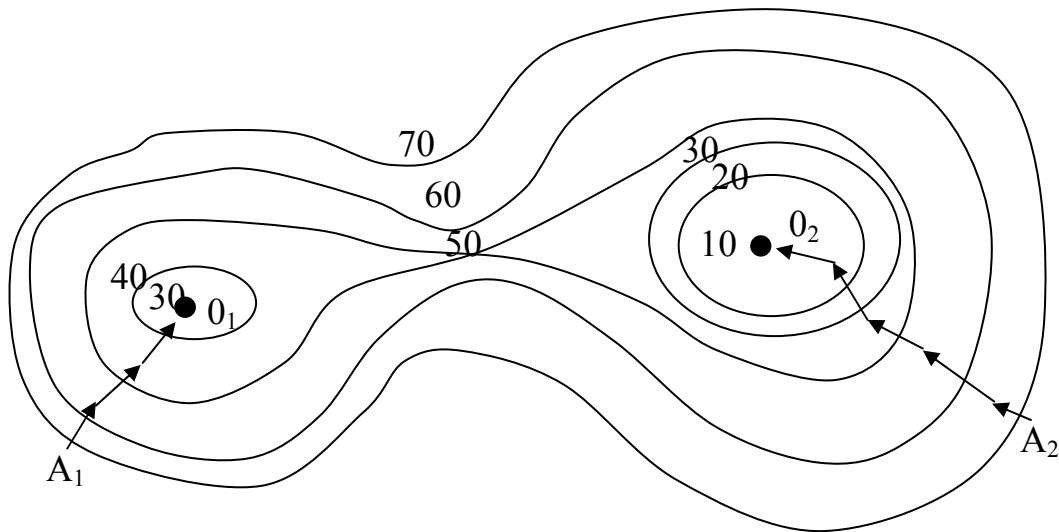


Рис. 8.5. Пример функции с двумя локальными минимумами

Такие функции принято называть многоэкстремальными. В нашем случае наименьшее значение функция принимает в точке O_2 . Начиная поиск минимума, из точки A_1 мы найдем только локальный минимум O_1 , так и не приходя к настоящему наименьшему значению O_2 . Универсального способа решения задачи многоэкстремальности не существует. Самый простой способ состоит в том, что проводят поиск несколько раз, начиная его с разных точек. Если получаются разные результаты, выбирают наименьший из них.

Во всех случаях необходимо иметь математическую модель функции оптимума.

8.3. Воспроизводимость и рандомизация опытов

Прежде чем приступить к планированию эксперимента, необходимо убедиться в воспроизводимости опытов. Для этого производят несколько серий параллельных опытов в рассматриваемой области изменения влияющих факторов [18]. Для удобства обработки данные заносят в табл. 8.1.

Результаты опытов

Номер серии опытов	Результаты параллельных опытов				Средне- арифметическое	Дисперсия
1	y_{11}	y_{12}	...	y_{1k}	\hat{y}_1	S^2_1
2	y_{21}	y_{22}	...	y_{2k}	\hat{y}_2	S^2_2
...						
N	y_{N1}	y_{N2}	...	y_{Nk}	\hat{y}_N	S^2_N

Для каждой серии параллельных опытов вычисляют среднее арифметическое значение функции отклика:

$$\hat{y}_j = \frac{1}{k} \sum_{i=1}^k y_{ji} \quad (j=1,2,\dots,N),$$

где k – число параллельных опытов, проведенных при одинаковых условиях.

Обычно N и k принимают от 2 до 4.

Вычисляют оценку дисперсии для каждой серии параллельных опытов:

$$S_j^2 = \frac{1}{k-1} \sum_{i=1}^k (y_{ji} - \hat{y}_j)^2.$$

Для проверки воспроизводимости опытов находят отношение наибольшей из оценок дисперсий к сумме всех оценок дисперсий:

$$G_p = \frac{\max S_j^2}{\sum_{j=1}^N S_j^2}.$$

Эта величина называется расчетным значением критерия Кохрена.

Некоторые значения критерия Кохрена приведены в табл. 8.2, они соответствуют доверительной вероятности $P = 0,95$, с которой принимается гипотеза воспроизводимости опытов.

Значение критерия Кохрена при $P=0,95$

N	F=k-1							
	1	2	3	4	5	6	7	8
2	0,999	0,975	0,939	0,906	0,877	0,853	0,833	0,816
3	0,967	0,871	0,798	0,746	0,707	0,677	0,653	0,633
4	0,907	0,768	0,684	0,629	0,590	0,560	0,637	0,518
5	0,841	0,684	0,598	0,544	0,507	0,478	0,456	0,439
6	0,781	0,616	0,532	0,480	0,445	0,418	0,398	0,382
7	0,727	0,561	0,480	0,431	0,397	0,373	0,354	0,338
8	0,680	0,516	0,438	0,391	0,360	0,336	0,319	0,304
9	0,639	0,478	0,403	0,358	0,329	0,307	0,290	0,277

Примечание

N – общее количество оценок дисперсий;

F = (k – 1) – число степеней свободы.

Если выполняется условие $G_p \leq G$, то опыты считаются воспроизводимыми, а оценки дисперсий – однородными.

Если проверка на воспроизводимость дала отрицательный результат, то остается признать либо невоспроизводимость эксперимента относительно управляемых переменных вследствие наличия флуктуаций неуправляемых и неконтролируемых переменных, создающих на выходе большой уровень «шума», либо наличие грубого промаха в строке, откуда взята дисперсия S_j^2 . В первом случае следует увеличить число параллельных опытов, во втором – найти грубый промах и заменить его результатом доброкачественного измерения при соответствующей комбинации факторов. Если это по каким-то причинам невозможно, то, чтобы не нарушать предпосылки использования критерия Кохрена, на место грубого промаха следует поместить среднюю арифметическую величину \hat{y}_j данной строки.

Пример. В эксперименте измерялся выход продукта реакции (y) в зависимости от температуры (x_1) и концентрации вещества (x_2) (Табл. 8.3).

Условия проведения опытов и результатов измерений

Номер серии опытов	Условия опытов		Результаты измерений		Средне-арифметическое \hat{y}_j	Дисперсия S_j^2
	$x_1, ^\circ\text{C}$	$x_2, \%$	$y_{j1}, \%$	$y_{j2}, \%$		
1	24	45	35,0	36,0	35,5	0,50
2	24	55	39,3	38,1	38,7	0,72
3	26	45	31,8	33,4	32,6	1,28

Расчетное значение критерия Кохрена находим по формуле:

$$G_p = \frac{1,28}{0,5 + 0,72 + 1,28} = 0,51.$$

Соответствующее значение критерия Кохрена $G = 0,967$ берем из таблицы. Оно найдено для следующих параметров: $P = 0,95$; $N = 3$; $f = k-1 = 2-1 = 1$.

Условие $G_p \leq G$ выполнено, следовательно, опыты можно считать воспроизводимыми.

Вычисление погрешности эксперимента

Оценки однородных дисперсий нескольких серий параллельных опытов можно усреднить и найти величину:

$$S_y^2 = \frac{1}{N} \sum_{j=1}^N S_j^2,$$

называемую оценкой дисперсии воспроизводимости. С ней связано число степеней свободы $f = N(k-1)$. В рассматриваемом примере

$$S_y^2 = 1/3 (0,5 + 0,72 + 1,28) = 0,83 \quad f = N(k-1) = 3(2-1) = 3.$$

Оценку дисперсии среднего значения рассчитывают по формуле:

$$S_{\bar{y}}^2 = \frac{S_y^2}{k}$$

с числом степеней свободы $f = N(k-1)$. $S_{\bar{y}}^2 = 0,83/2 \sim 0,42$.

В тех случаях, когда из-за недостатка времени, трудоемкости или высокой стоимости эксперимента опыты не дублируются, при обработке экспериментальных данных используют S^2_y .

Рандомизация

Для того чтобы компенсировать систематические погрешности эксперимента, используют прием, называемый рандомизацией. Он заключается в том, что опыты проводят в случайной последовательности, которая устанавливается при помощи таблицы случайных чисел.

Пусть требуется рандомизировать 6 опытов. Поставим им в соответствие любые 6 последовательных случайных чисел (одинаковые числа не допускаются), например:

№1	60
№2	12
№3	05
№4	15
№5	34
№6	30

Расположив случайные числа в порядке возрастания или убывания, получаем искомую последовательность опытов (№3, №2, №4, №6, №5, №1).

8.4. Экспериментально–статистические модели

Под математическим описанием процесса будем понимать систему уравнений, связывающих функции отклика с влияющими факторами. В простейшем случае это может быть одно уравнение. Часто математическое описание называют *математической моделью*.

Математической моделью в планировании эксперимента принимают уравнение, связывающее параметр оптимизации с факторами:

$$Y = (X_1, X_2, \dots X_n).$$

Каждый фактор может принимать в опыте одно из нескольких значений – *уровней*. Чтобы узнать число различных состояний, достаточно число уровней

факторов «n» (при условии, что оно одинаково для всех факторов) возвести в степень числа факторов «m»:

$$N = p^m,$$

где p – число уровней.

Так, для пяти факторов с пятью уровнями $N = 3125$, а для десяти факторов на четырех уровнях – уже свыше 1 000 000.

С помощью математических методов оптимального планирования эксперимента можно получить математическую модель процесса даже при отсутствии сведений о его механизме. Это в ряде случаев бывает очень полезно.

Ценность математического описания заключается в том, что оно:

- представляет информацию о влиянии факторов;
- позволяет количественно определить значения функций отклика при заданном режиме ведения процесса;
- может служить основой для оптимизации.

Следует отметить, что на основе методов планирования эксперимента можно количественно описать также свойства таких продуктов, как сплавы, пластмассы, резина, керамика, ситаллы, бетоны и т. п.

Математические модели, получаемые с помощью методов планирования эксперимента, принято называть экспериментально–статистическими.

8.5. Полный факторный эксперимент

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется *полным факторным экспериментом*.

Метод *полного факторного эксперимента* дает возможность получить математическое описание исследуемого процесса в некоторой *локальной* области факторного пространства, лежащей в окрестности выбранной точки с координатами $(x_{01}, x_{02}, \dots, x_{0n})$. Перенесем начало координат факторного пространства в эту точку (рис. 8.6).

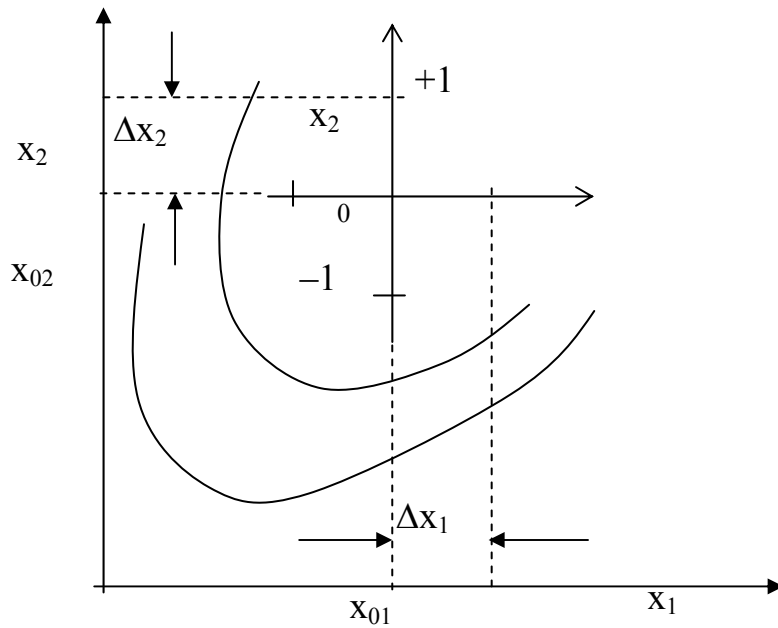


Рис. 8.6. Введение кодированных переменных

С этой целью введем новые переменные:

$$X_i = \frac{x_i - x_{0i}}{\Delta x_i} \quad (i = 1, 2, \dots, n),$$

где Δx_i — масштаб по оси x_i .

Иногда величину X_i называют *кодированной* переменной.

Функцию отклика в окрестности нового начала координат разложим в ряд Тейлора

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{12} x_1 x_2 + \dots + \beta_{11} x_1^2 + \dots,$$

где β_0 — значение функции отклика в начале координат;

$$\left. \begin{aligned} \beta_i &= \frac{\partial y}{\partial x_i}, \\ \beta_{ij} &= \frac{\partial^2 y}{\partial x_i \partial x_j}, \\ \beta_{ii} &= \frac{1}{2} \frac{\partial^2 y}{\partial x_i^2} \end{aligned} \right\}$$

и т. д.

Метод полного факторного эксперимента служит для получения математического описания процесса в виде отрезка ряда Тейлора. При этом обычно ограничиваются линейной частью разложения и членами, содержащими произведения факторов в первой степени. Таким образом, удастся находить уравнение локального участка поверхности отклика, если его кривизна не слишком велика.

Следует отметить, что коэффициенты искомого уравнения определяются на основе экспериментальных данных и, следовательно, несут на себе отпечаток погрешностей эксперимента. Чтобы подчеркнуть это обстоятельство, в уравнении вместо символов β , обозначающих истинные значения коэффициентов, пишут b , подразумевая под этим соответствующие выборочные оценки.

Итак, с помощью полного факторного эксперимента ищут математическое описание процесса в виде уравнения:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_{12}X_1X_2 + \dots + b_{11}X_1^2 + \dots$$

Его называют уравнением регрессии, а входящие в него коэффициенты — коэффициентами регрессии. Для удобства вычислений коэффициентов регрессии все факторы в ходе полного факторного эксперимента варьируют на двух уровнях, соответствующих значениям кодированных переменных +1 и -1.

В табл. 8.4 приведены условия опытов полного двухфакторного эксперимента. Часть таблицы, обведенная пунктиром, называется матрицей планирования.

Таблица 8.4

Полный двухфакторный эксперимент

Номер опыта	Факторы		Функция отклика
	X_1	X_2	
1	+1	-1	Y_1
2	+1	-1	Y_2
3	+1	+1	Y_3
4	+1	+1	Y_4

Как видно из рис. 8.7, опыты, приведенные в табл. 8.4, соответствуют на факторной плоскости вершинам квадрата с центром в начале координат.

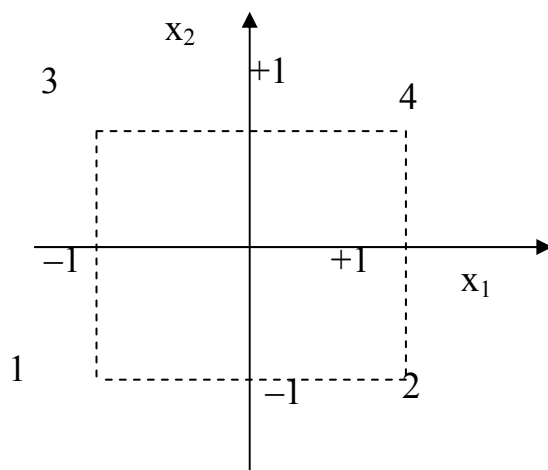


Рис. 8.7. Опыты полного двухфакторного эксперимента

В следующей табл. 8.5 приведены условия опытов полного трехфакторного эксперимента. Эти опыты соответствуют в факторном пространстве вершинам куба с центром в начале координат.

Таблица 8.5

Полный трехфакторный эксперимент

Номер опыта	Факторы			Функция отклика
	X_1	X_2	X_3	
1	-1	-1	-1	Y_1
2	+1	-1	-1	Y_2
3	-1	+1	-1	Y_3
4	+1	+1	-1	Y_4
5	-1	-1	+1	Y_5
6	+1	-1	+1	Y_6
7	-1	+1	+1	Y_7
8	+1	+1	+1	Y_8

Из табл. 8.4, 8.5 видны основные принципы построения матриц планирования полного факторного эксперимента:

- уровни варьирования первого фактора чередуются от опыта к опыту;
- частота смены уровней варьирования каждого последующего фактора вдвое меньше, чем у предыдущего.

Матрица планирования полного факторного эксперимента обладает следующими свойствами:

1) симметричность относительно центра эксперимента, то есть алгебраическая сумма элементов вектор – столбца каждого фактора равна нулю:

$$\sum_i^N X_{ji} = 0,$$

где j – номер фактора;

N – число опытов;

2) условие нормировки – сумма квадратов элементов каждого столбца равна числу опытов:

$$\sum_i^N X_{ji}^2 = N;$$

3) сумма почленных произведений любых двух вектор – столбцов равна нулю. Это свойство называют ортогональностью матрицы планирования:

$$\sum_i^N X_{ji} \cdot U_{ji} = 0;$$

4) рототабельность матрицы, то есть все точки в матрице подбираются так, чтобы точность предсказания значений параметра оптимизации была одинакова на равных расстояниях от центра эксперимента и не зависела от направления.

Интервал варьирования факторов

Масштаб по оси x_i (см. рис. 8.6) называют интервалом или шагом варьирования факторов (Δx_i).

Интервалом варьирования факторов называется число, добавление которого к основному (нулевому) уровню показывает верхний, а вычитание – нижний уровень факторов.

Шаг (интервал) варьирования по каждой переменной выбирается таким, чтобы приращение величины выходного параметра Y к базовому значению Y^* при реализации шага можно было выделить на фоне «шума» при небольшом числе параллельных опытов.

Если нет никаких указаний на величину шага Δx_i , то в первом приближении можно выбрать $\Delta x_i = 0,15x_{0i}$, т.е. принять за шаг 15 % – ное отклонение от

базового уровня X_{0i} . Такой шаг предоставляет достаточную гарантию того, что фактор X_i вызовет заметную реакцию Y , если связь между ними существует.

Размер интервала варьирования определяется многими факторами, но упрощенно можно ограничиться следующим (рис. 8.8):



Рис. 8.8. Рекомендации по выбору интервалов варьирования факторов

- если интервал составляет менее 10 % от области определения, он считается узким;
- если не более 30 % – средним;
- более 30 % – широким.

Точность фиксирования (определения) факторов определяется точностью приборов и стабильностью в ходе опыта. Упрощенно можно полагать, что если погрешность составляет:

- < 1 % – высокая точность,
- < 5 % – средняя точность,
- >10 % – низкая точность эксперимента.

Свойства полного факторного эксперимента 2^2

Математическая модель

Описание исследуемого объекта нельзя получить в виде точной формулы, справедливой во всем диапазоне существования аргументов. Оно может быть лишь приближенным и на небольшом участке в окрестностях выбранной базовой точки. Аппроксимация искомой математической зависимости представляет собой некоторый полином – отрезок ряда Тейлора.

Рассмотрим снова матрицу планирования 2^2 . Предположим, что для движения к оптимуму нужна линейная модель типа $Y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2$.

Наша задача – найти неизвестные коэффициенты, причем эксперимент, содержащий конечное число опытов, позволяет получить только выборочные оценки для коэффициентов уравнения. Их точность зависит от свойств выборки и нуждается в статистической проверке.

Оценки коэффициентов вычисляются по простой формуле:

$$b_j = \frac{\sum_i X_{ji} \cdot Y_i}{N}, \quad j = 0, 1, 2, \dots k. \quad (8.1)$$

Для нашего случая:

$$b_1 = \frac{(-1)Y_1 + (+1)Y_2 + (-1)Y_3 + (+1)Y_4}{4},$$
$$b_2 = \frac{(-1)Y_1 + (-1)Y_2 + (+1)Y_3 + (+1)Y_4}{4}.$$

Коэффициент $b_0 = \bar{Y}$ есть среднееарифметическое значение переменной Y

$$b_0 = \frac{Y_1 + Y_2 + Y_3 + Y_4}{4}. \quad (8.2)$$

Коэффициенты при переменных указывают на силу влияния факторов. Чем больше коэффициент по абсолютной величине, тем больше влияние на эксперимент оказывает данный фактор.

Знак плюс говорит о том, что параметр оптимизации увеличивается с увеличением фактора, минус – наоборот.

Планируя эксперимент на первом этапе, мы стремимся получить линейную модель. Однако у нас нет гарантии, что в выбранных интервалах варьирования модель линейна. Один из часто встречающихся видов нелинейности связан с тем, что эффект одного фактора зависит от уровня, на котором находится другой фактор, то есть просматривается эффект взаимодействия двух факторов.

Для оценки эффекта взаимодействия матрица планирования дополняется еще одним столбцом (см., например, табл. 8.6).

Таблица 8.6

Учет эффекта взаимодействия

Номер	X ₁	X ₂	X ₁ ·X ₂	Y
1	-1	-1	+1	Y ₁
2	+1	-1	-1	Y ₂
3	-1	+1	-1	Y ₃
4	+1	+1	+1	Y ₄

Дополнительный коэффициент можно представить в виде

$$b_{12} = \frac{(+1)Y_1 + (-1)Y_2 + (-1)Y_3 + (+1)Y_4}{4}.$$

Модель будет выглядеть следующим образом:

$$Y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_{12} \cdot X_1 \cdot X_2.$$

После определения оценок коэффициентов регрессии необходимо проверить гипотезу о значимости коэффициентов b_i . Лучше всего это сделать в виде *нуль-гипотезы*, т.е. гипотезы о равенстве $b_i = 0$.

Если она подтвердилась, то коэффициент b_i следует признать статистически незначимым и отбросить из искомой модели; если гипотеза не подтвердилась, то соответствующий коэффициент b_i следует признать значимым и включить в модель.

Проверка гипотезы проводится с помощью t – критерия Стьюдента, который при проверке нуль – гипотезы формируется в виде

$$t = \frac{|b_i|}{\sqrt{S_{b_i}^2}} = \frac{|b_i|}{\sqrt{S_{\bar{y}}^2}} \cdot \sqrt{N},$$

где $S_{b_i}^2$ – дисперсия ошибки определения коэффициента b_i .

Некоторые значения t – критерия представлены в табл. 8.7.

Таблица 8.7

Распределение Стьюдента

f	Доверительная вероятность						
	0,7	0,8	0,9	0,95	0,98	0,99	0,999
1	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	1,336	1,886	2,920	4,303	6,965	9,925	31,598
3	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	1,156	1,476	1,943	2,571	3,365	4,032	6,859
6	1,134	1,440	1,895	2,447	3,143	3,707	5,959
7	1,119	1,415	1,860	2,365	2,998	3,499	5,405
8	1,108	1,397	1,833	2,306	2,896	3,355	4,781

При полном и дробном факторном планировании для всех i :

$$S_{b_i}^2 = \frac{S_{\bar{y}}^2}{N}, \quad (8.3)$$

доверительные интервалы:

$$\Delta b_i = \pm t \cdot S_{b_i}.$$

Если величина коэффициента регрессии превышает Δb_i , найденное для

q % – ного уровня значимости и числа степеней свободы,

$$f = N(k - 1),$$

где N – число серий параллельных опытов,

k – число параллельных опытов.

При этом нуль–гипотеза отвергается, коэффициент считается значимым и его следует включить в искомую модель.

Статистическая незначимость коэффициента b_i может быть обусловлена следующими причинами:

- уровень базового режима близок к точке частного экстремума по переменной X_i или по произведению переменных;
- шаг варьирования Δx_i выбран малым;
- данная переменная (или произведение переменных) не имеет функциональной связи с выходным параметром Y ;
- велика ошибка эксперимента вследствие наличия неуправляемых и неконтролируемых переменных.

Поскольку ортогональное планирование позволяет определять доверительные границы для каждого из коэффициентов регрессии в отдельности, то если какой–либо из коэффициентов окажется незначимым, он может быть отброшен без пересчета всех остальных. После этого математическая модель объекта составляется в виде уравнения связи выходного параметра Y и переменных X_i , включающего только значимые коэффициенты.

Проверка адекватности модели

Первый вопрос, который нас интересует после вычисления коэффициентов регрессии, – это проверка ее пригодности или проверка адекватности модели.

Для этого определяют дисперсию адекватности (или остаточную дисперсию):

$$S_{ад}^2 = \frac{1}{f} \sum_i \Delta Y_i^2, \quad (8.4)$$

где f – число степеней свободы (разность между числом опытов и числом коэффициентов (констант), которые уже вычислены по результатам этих опытов независимо друг от друга),

$$f = N - (k + 1)$$

здесь N – число опытов;

k – число коэффициентов регрессии b_i .

Так для ПФЭ 2^3 $f = 8 - (3 + 1) = 4$, для ПФЭ 2^2 $f = 4 - (2 + 1) = 1$.

Вычислим разность $\Delta Y_i^2 = (Y_i^{\text{эксперим}} - Y_i^{\text{расчетн}})^2$.

После этого для проверки адекватности используем критерий Фишера:

$$F = \frac{S_{\text{ад}}^2}{S^2} \geq F_{\text{табл}} \quad (8.5)$$

Для определения $F_{\text{табл}}$ необходимо знать число степеней свободы для дисперсий воспроизводимости и адекватности. В табл. 8.8 представлены некоторые значения критерия Фишера с доверительным уровнем вероятности 95 %.

Таблица 8.8

Фишеровское распределение величин для $P=0,95$

	Число степеней свободы дисперсии адекватности									
	1	2	3	4	5	6	7	8	9	10
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9
2	18,51	19,00	19,16	19,75	19,30	19,33	19,35	19,37	19,38	19,40
3	10,13	9,55	9,28	9,12	9,01	8,44	8,89	8,85	8,81	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98

Если расчетное значение критерия Фишера не превышает табличного, то с соответствующей доверительной вероятностью модель можно считать адекватной.

Если гипотеза адекватности отвергается, то модель признается неадекватной экспериментальным данным. *Неадекватность модели не означает ее неправильности!* Неадекватность модели может означать, что не весь перечень влияющих факторов был принят во внимание, или что необходимо перейти к более сложной форме уравнения связи, или выбрать другой шаг варьирования по одному или нескольким факторам и т.п. Однако все достижения неадекватной модели (отсев незначимых факторов, оценка дисперсии эксперимента и другие) остаются в силе.

Пример. Рассмотрим химический процесс, в котором выход продукта реакции y (%) зависит от температуры реакционной смеси x_1 (°C) и концентрации реагента x_2 (%). Требуется с помощью полного факторного эксперимента найти математическое описание этого процесса в окрестности точки факторного пространства с координатами $x_{01} = 50$ °C и $x_{02} = 25$ %.

Решение. Математическое описание рассматриваемого процесса будем искать в виде уравнения регрессии $Y = b_0 + b_1X_1 + b_2X_2$, где кодированные переменные связаны с температурой и концентрацией следующими соотношениями:

$$X_1 = \frac{x_1 - x_{01}}{\Delta x_1}, \quad X_2 = \frac{x_2 - x_{02}}{\Delta x_2}.$$

При проведении полного факторного эксперимента зададимся условиями, приведенными в табл. 8.9.

Таблица 8.9

Основные характеристики экспериментов

Характеристика эксперимента	X_1 , °C	X_2 , %
Основной уровень	50	25
Интервал варьирования	5	1
Верхний уровень	55	26
Нижний уровень	45	24

Матрица планирования и результаты полного факторного эксперимента представлены в табл. 8.10.

Таблица 8.10

Полный двухфакторный эксперимент

Номер опыта	X ₁	X ₂	x ₁ , °C	x ₂ , %	Y, %
1	- 1	-1	45	24	35,5
2	+ 1	-1	55	24	38,7
3	- 1	+1	45	26	32,6
4	+1	+1	55	26	36,2

На основании результатов полного факторного эксперимента рассчитаем коэффициенты регрессии, пользуясь формулами (8.1) и (8.2):

$$b_0 = \frac{1}{4}(35,5 + 38,7 + 32,6 + 36,2) = 35,6,$$

$$b_1 = \frac{1}{4}(-35,5 + 38,7 - 32,6 + 36,2) = 1,95,$$

$$b_2 = \frac{1}{4}(-35,3 - 38,7 + 32,6 + 36,2) = -1,35.$$

Будем считать, что оценка дисперсии среднего значения S_y^2 равна 0,42. Примем также, что с этой величиной связаны три степени свободы:

$$(f = N(k - 1) = 3(2 - 1) = 3).$$

Ошибку в определении коэффициентов регрессии вычислим по формуле (8.3):

$$S_b = \sqrt{\frac{S_y^2}{N}} = \sqrt{\frac{0,42}{4}} = 0,32.$$

Пользуясь табл. 8.7, найдем, что для доверительной вероятности $P = 0,95$ и трех степеней свободы значение критерия Стьюдента $t = 3,18$. Тогда $S_b \cdot t = 0,32 \cdot 3,18 = 1,03$.

Для оценки значимости коэффициентов регрессии рассмотрим следующие соотношения:

$$|b_0| = 35,6 > S_b \cdot t, |b_1| = 1,95 > S_b \cdot t, |b_2| = 1,35 > S_b \cdot t.$$

Видно, что все коэффициенты регрессии значимы. Следовательно, иско-
мое уравнение имеет вид:

$$y = 35,6 + 1,95X_1 - 1,35X_2.$$

Для проверки адекватности уравнения регрессии найдем расчетные значения функции отклика:

$$Y^p_1 = 35,6 + 1,95(-1) - 1,35(-1) = 35,0,$$

$$Y^p_2 = 35,6 + 1,95(+1) - 1,35(-1) = 38,9,$$

$$Y^p_3 = 35,6 + 1,95(-1) - 1,35(+1) = 32,3,$$

$$Y^p_4 = 35,6 + 1,95(+1) - 1,35(+1) = 36,2.$$

По формуле (8.4) вычислим оценку дисперсии адекватности:

$$S^2_{ад} = \frac{1}{N - (k + 1)} \sum_{j=1}^N (y_j^э - y_j^p)^2 =$$

$$= \frac{1}{4 - 3} [(35,5 - 35,0)^2 + (38,7 - 38,9)^2 + (32,6 - 32,3)^2 + (36,2 - 36,2)^2] = 0,38.$$

С ней связано число степеней свободы $f = N - (k + 1) = 4 - 3 = 1$. Расчетное значение критерия Фишера находим по формуле (8.5):

$$F = \frac{S^2_{ад}}{S^2} = \frac{0,38}{0,42} = 0,905.$$

Оно не превосходит значения, приведенного в табл. 8.8, ($F^{таб} = 10,13$), следовательно, нельзя сказать, что уравнение регрессии неадекватно.

При оптимизации мы стремимся сделать эффекты взаимодействия как можно меньшими. В задачах интерполяции, напротив, их выявление часто важно и интересно. С ростом числа факторов число возможных взаимодействий быстро растет. Так, для матрицы 2^3 (табл. 8.11) таких взаимодействий уже четыре.

Таблица 8.11

Учет эффектов взаимодействия факторов ПФЭ 2^3

Номер опыта	X ₀	X ₁	X ₂	X ₃	X ₁ .X ₂	X ₁ .X ₃	X ₂ .X ₃	X ₁ .X ₂ .X ₃	Y
1	+1	-1	-1	-1	+1	+1	+1	-1	Y ₁
2	+1	+1	-1	-1	-1	-1	+1	+1	Y ₂
3	+1	-1	+1	-1	-1	+1	-1	+1	Y ₃
4	+1	+1	+1	-1	+1	-1	-1	-1	Y ₄
5	+1	-1	-1	+1	+1	-1	-1	+1	Y ₅
6	+1	+1	-1	+1	-1	+1	-1	-1	Y ₆
7	+1	-1	+1	+1	-1	-1	+1	-1	Y ₇
8	+1	+1	+1	+1	+1	+1	+1	+1	Y ₈

Уравнение регрессии:

$$Y = b_0X_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + b_{12} \cdot X_1X_2 + b_{23} \cdot X_2X_3 + b_{13} \cdot X_1X_3 + b_{123} \cdot X_1X_2 X_3 .$$

8.6. Дробный факторный эксперимент

Вернемся к матрице 2^2 с учетом взаимодействия факторов (см. табл. 8.6).

Уравнение регрессии $Y = b_0X_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_{12} \cdot X_1 \cdot X_2$.

Если есть основания считать, что процесс описывается линейной моделью, достаточно определить b_0, b_1, b_2 . При линейном приближении эффект взаимодействия стремится к нулю ($b_{12} \Rightarrow 0$) и этот вектор – столбец можно использовать для нового фактора X_3 . Матрица планирования запишется в виде табл. 8.12.

Таблица 8.12

Замена эффекта взаимодействия новым фактором

Номер опыта	X_0	X_1	X_2	X_3	Y
1	+1	-1	-1	+1	Y_1
2	+1	+1	-1	-1	Y_2
3	+1	-1	+1	-1	Y_3
4	+1	+1	+1	+1	Y_4

Уравнение регрессии $Y = b_0X_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3$.

Данное правило формулируется следующим образом.

Для того, чтобы сократить число опытов, нужно новому фактору присвоить вектор–столбец матрицы, принадлежащий взаимодействию, которым можно пренебречь.

Подводя итог, следует сформулировать основные правила проведения полного или дробного факторного эксперимента, которые должны включать:

- 1) цель исследования (от правильной формулировки целей и задач исследования во многом зависит успех эксперимента);
- 2) параметры, характеризующие процесс;
- 3) формулировку задачи оптимизации или интерполяции;
- 4) факторы, влияющие на процесс;
- 5) выбор варьируемых факторов;

- 6) технологию проведения эксперимента;
- 7) список необходимых замеров и анализов;
- 8) описание экспериментальной установки с основными геометрическими размерами и физическими параметрами;
- 9) выбор основного уровня и интервалов варьирования. Удобнее их записать в виде табл. 8.13.

Таблица 8.13

Основной уровень и интервалы варьирования факторов

Факторы	Уровни варьирования			Шаг варьирования
	-1	0	+1	
X_1				I_1
X_2				I_2
...				...

- 10) априорную информацию: опубликованные данные, сведения из отчетов;
- 11) матрицу планирования;
- 12) таблицу результатов эксперимента (таблицы и графики).

9. МЕТОДЫ ОПТИМИЗАЦИИ

9.1. Метод крутого восхождения

Мы рассмотрели методы построения экспериментально–статистических моделей в виде уравнений регрессии. Теперь рассмотрим вопрос о том, как использовать эти модели для оптимизации процессов или свойств многокомпонентных систем.

Следует иметь в виду, что качество процесса обычно характеризуется *несколькими* функциями отклика. Однако обычно невозможно найти такое сочетание значений влияющих факторов, при котором одновременно достигаются экстремумы всех интересующих экспериментатора функций отклика. Например, максимальная производительность оборудования и минимальная себестоимость продукции обычно достигаются при различных технологических режимах.

Важно отметить, что как влияющие факторы, так и функции отклика могут изменяться только в определенных пределах. Так, концентрации реагентов не могут быть отрицательными, температура и давление в аппарате не могут превышать безопасных пределов, себестоимость продукции должна быть не выше плановой и т.п. Следовательно, оптимизацию процессов, как правило, осуществляют в условиях *ограничений* на влияющие факторы и функции отклика.

Величина, характеризующая уровень оптимизации процесса, называется *критерием оптимальности*. В частном случае критерием оптимальности может быть одна из функций отклика, характеризующих процесс.

Оптимизация процесса представляет собой целенаправленный поиск значений влияющих факторов, при которых достигается экстремум критерия оптимальности (с учетом ограничений, наложенных на все влияющие факторы и функции отклика).

Известные ученые Д. Бокс и К. Уилсон предложили использовать для оптимизации результаты полного или дробного факторного эксперимента [18]. Сущность такой оптимизации состоит в следующем.

Пусть, например, критерием оптимальности служит функция отклика y , представленная в виде

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_{12} X_1 X_2 + \dots + b_{11} X_1^2 + \dots$$

Один из влияющих факторов принимают за *базовый*, и для него вычисляют произведение соответствующего коэффициента регрессии на шаг варьирования. Например, для первого фактора это произведение имеет вид $b_1 \Delta x_1$.

Затем для базового фактора выбирают шаг движения Δx_1^* , с которым будет осуществляться оптимизация. Обычно $\Delta x_1^* < \Delta x_1$.

После этого вычисляют отношение $\gamma = \frac{\Delta x_1^*}{b_1 \Delta x_1}$.

Для всех остальных факторов шаги движения к оптимальным значениям рассчитывают по формуле

$$\Delta x_i^* = \gamma b_i \Delta x_i. \quad (9.1)$$

Движение к оптимуму начинают *из центра* плана, который использовался для получения математического описания функции отклика (рис. 9.1).

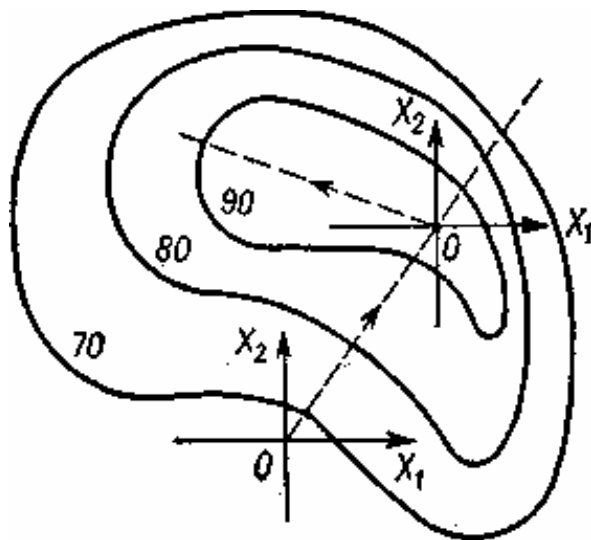


Рис. 9.1. Оптимизация по методу крутого восхождения

Значения факторов на каждом новом шаге находят путем прибавления Δx_i^* к соответствующим предыдущим значениям. Так осуществляется оптимизация по методу *крутого восхождения*.

Если же ищется минимум функции y , то новые значения факторов находят из предыдущих путем вычитания Δx_i^* . Такой способ оптимизации называют методом *наискорейшего спуска*.

Движение к оптимуму прекращают в следующих случаях:

- значения (одного или нескольких) факторов или функций отклика вышли на границы допустимых значений;
- достигнут экстремум критерия оптимальности y .

В первом случае на этом оптимизация заканчивается, а во втором – в области экстремума функции y ищут ее *новое* математическое описание, используя полный или дробный факторный эксперимент. Если удастся получить адекватное описание этой функции в виде полинома, то продолжают оптимизацию методом крутого восхождения (см. рис. 9.1). Очевидно, оптимум, найденный в результате первого крутого восхождения, был локальным.

Если же в области оптимума не удастся получить адекватного уравнения регрессии, то проводят анализ выбранных переменных и добавляют новые влияющие факторы либо увеличивают точность эксперимента.

Пример. Пусть в результате полного факторного эксперимента (см. стр. 85) получено адекватное уравнение регрессии

$$y_1 = 35,6 + 1,95X_1 - 1,35X_2.$$

Здесь, как и в предыдущем примере, y_1 – выход продукта реакции, X_1 – температура, X_2 – концентрация реагента.

Введем также в рассмотрение функцию отклика y_2 , характеризующую скорость химической реакции ($\text{кмоль} \cdot \text{м}^{-3} \cdot \text{ч}^{-1}$).

Требуется выполнение условия $y_2 \geq 2,5$.

Решение. Допустим, что ограничения на влияющие факторы имеют вид:

$$30^\circ \leq x_1 \leq 120^\circ,$$

$$10\% \leq x_2 \leq 70\%.$$

Будем оптимизировать выход продукта реакции методом крутого восхождения.

В качестве базового фактора возьмем температуру и примем шаг движения на крутом восхождении 4° , тогда $\gamma = \frac{\Delta x_1^*}{b_1 \Delta x_1} = \frac{4}{1,95 \cdot 5} = 0,41$.

Здесь Δx_1 взят по условиям полного факторного эксперимента (предыдущий пример).

Шаг по концентрации на крутом восхождении можно рассчитать по уравнению:

$$\Delta x_2^* = \gamma b_2 \Delta x_2 = 0,41(-1,35) \cdot 1 = -0,55\%.$$

Для удобства ведения эксперимента шаги движения, рассчитанные по формуле (9.1), можно несколько округлять. В данном случае удобно принять $\Delta x_2^* = -0,5\%$.

Результаты опытов, выполненных по методу крутого восхождения, приведены в табл. 9.1.

Результаты опытов по методу крутого восхождения

Характеристика опыта	x_1	x_2	y_1^3	y_2^3
Центр плана	50	25	35,1	2,9
Интервал варьирования	5	1	–	–
Шаг движения	4	–0,5	–	–
Крутое восхождение				
Номер опыта				
1	54	24,5	36,9	3,2
2	58	24,0	37,2	3,7
3	62	23,5	38,5	2,8
4	66	23,0	40,7	2,3
5	70	22,5	38,1	1,9
6	74	22,0	37,2	1,6

Примечание. y_1^3 – экспериментальные значения выхода продукта реакции; y_2^3 – экспериментально найденные скорости реакции.

Как видно из табл. 9.1, в опыте № 4 достигнут максимальный выход продукта реакции, однако скорость процесса в этом случае меньше допустимого значения. По–видимому, оптимальным режимом процесса следует считать условия опыта № 3. Ограничения на x_1 и x_2 в ходе оптимизации не нарушены.

9.2. Симплексный метод

Симплексом называется правильный многогранник, имеющий $n + 1$ вершину, где n – число факторов, влияющих на процесс. Так, если факторов два, то симплексом является правильный треугольник. Сущность симплексного метода оптимизации иллюстрирует рис. 9.2.

Начальная серия опытов соответствует вершинам исходного симплекса (точки 1, 2 и 3). Условия этих первых опытов берутся из области значений факторов, соответствующих наиболее благоприятным из известных режимов оптимизируемого процесса.

Сравнивая между собой результаты опытов в точках 1, 2 и 3, находят среди них самый «плохой» с точки зрения выбранного критерия оптимальности. Пусть, например, самым «неудачным» оказался опыт в точке 1.

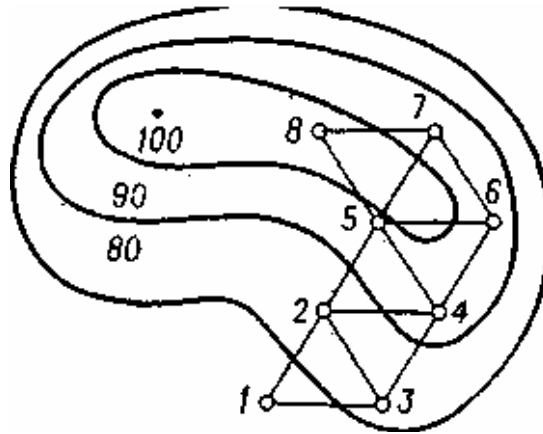


Рис. 9.2. Оптимизация по симплексному методу

Этот опыт исключают из рассмотрения, а вместо него в состав симплекса вводят опыт в точке 4, которая симметрична точке 1 относительно противоположной стороны треугольника, соединяющей точки 2 и 3.

Далее сравнивают между собой результаты опытов в вершинах нового симплекса, отбрасывают самый «неудачный» из них и переносят соответствующую вершину симплекса в точку 5. Затем рассмотренная процедура повторяется в течение всего процесса оптимизации.

Если достигнут экстремум критерия оптимальности, то дальнейшее движение симплекса прекращается. Это значит, что новый шаг возвращает исследователя в предыдущую точку факторного пространства.

Следует иметь в виду, что симплексный метод, так же как и метод крутого восхождения, является *локальным* методом поиска экстремума. Если существует несколько экстремумов критерия оптимальности, то этот метод позволяет найти тот из них, который расположен ближе к точкам исходного симплекса. Поэтому если есть подозрение о существовании нескольких экстремумов критерия оптимальности, то нужно осуществить их поиск, каждый раз начиная оптимизацию из новой области факторного пространства. Затем следует сравнить между собой найденные оптимальные условия и из всех вариантов выбрать наилучший.

При оптимизации необходимо принимать во внимание ограничения, наложенные на влияющие факторы и функции отклика.

Важно отметить, что при пользовании симплексным методом *не обязательно* дублировать опыты. Дело в том, что ошибка в отдельном опыте может только несколько замедлить оптимизацию. Если же последующие опыты выполняются безупречно, то движение к оптимуму продолжается.

Матрица опытов исходного симплекса в *кодированных* переменных приведена в табл. 9.2. Символом «О» обозначены координаты центра плана, т.е. основной уровень.

Величины, входящие в эту таблицу, рассчитываются по следующим формулам:

$$k_i = \sqrt{\frac{1}{2i(i+1)}}, \quad (9.2)$$

$$R_i = i \cdot k_i, \quad (9.3)$$

где i — номер фактора в матрице планирования.

Таблица 9.2

Матрица исходного симплекса

Номер опыта	X_1	X_2	...	X_{n-1}	X_n	Функция отклика
1	k_1	k_2	...	k_{n-1}	k_n	y_1
2	$-R_1$	k_2	...	k_{n-1}	k_n	y_2
3	0	$-R_2$...	k_{n-1}	k_n	y_3
...
$n-1$	0	0	...	k_{n-1}	k_n	y_{n-1}
n	0	0	...	$-R_{n-1}$	k_n	y_n
$n+1$	0	0	...	0	$-R_n$	y_{n+1}

Опыты, представленные в табл. 9.2, соответствуют вершинам симплекса, сторона которого равна единице, а центр совпадает с началом координат (в кодированных переменных).

Результаты расчетов для четырех факторов, выполненных на основании табл. 9.2 и формул (9.2) и (9.3), приведены в табл. 9.3.

Условия начальной серии опытов

Номер опыта	X ₁	X ₂	X ₃	X ₄
1	0,5	0,289	0,204	0,158
2	-0,5	0,289	0,204	0,158
3	0	-0,578	0,204	0,158
4	0	0	-0,612	0,158
5	0	0	0	-0,632

Аналогично можно рассчитать условия исходной серии опытов для большего количества факторов.

Очевидно, наибольшее количество опытов придется ставить *в начале* эксперимента. Затем на каждом шаге оптимизации выполняется *только один* опыт.

Приступая к оптимизации, необходимо с помощью табл. 9.2 или 9.3 рассчитать матрицу исходной серии опытов в *физических* переменных, преобразуя формулу

$$x_i = x_{i0} + \Delta x_i X_i, \quad (9.4)$$

где x_{i0} – основной (нулевой уровень);

X_i – кодированная переменная;

Δx_i – интервал варьирования.

В дальнейшем все операции производятся только с физическими переменными.

Условия каждого нового опыта рассчитываются по формуле

$$x_i = \frac{2}{n} \left(\sum_{j=1}^{n+1} x_{ji} - x_i^* \right) - x_i^*, \quad (9.5)$$

где n – число факторов в матрице планирования;

j – номер опыта;

i – номер фактора;

x_i^* – значение i -го фактора в самом «неудачном» опыте предыдущего симплекса.

Следует отметить, что на любом шаге оптимизации, осуществляемой симплексным методом, можно включить в программу исследований *новый фактор*,

который до тех пор не принимался во внимание, но оставался на постоянном уровне. При этом значения всех ранее рассматриваемых факторов рассчитываются по формуле

$$x_i = \frac{1}{n+1} \sum_{j=1}^{n+1} x_{ji}, \quad (9.6)$$

где $i = 1, 2, \dots, n$, т. е. является средним арифметическим значением соответствующих координат предыдущего симплекса.

Значение вновь вводимого фактора определяется по формуле

$$x_{n+1} = x_{0(n+1)} + \Delta x_{n+1}(R_{n+1} + k_{n+1}), \quad (9.7)$$

где $x_{0(n+1)}$ — основной уровень этого фактора;

Δx_{n+1} — выбранный шаг варьирования для данного фактора;

R_{n+1}, k_{n+1} — величины, рассчитываемые по формулам (9.2) и (9.3).

Отметим, что добавление нового фактора в состав *полного факторного* эксперимента сопровождается увеличением количества опытов *вдвое*. В этом смысле симплексный метод имеет очевидное *преимущество*.

В практику научных исследований симплексный метод был введен англичанином Ф. Химсвортом в 1962 году [18].

Пример. Пусть требуется с помощью симплексного метода оптимизировать выход целевого продукта y (%), который получается при взаимодействии двух реагентов с концентрациями x_1 и x_2 (кмоль·м⁻³) при температуре x_3 (°C).

Решение. Выберем основные уровни и шаги варьирования факторов и сведем их в табл. 9.4.

Таблица 9.4

Значения уровней факторов и шагов варьирования

Фактор	Основной уровень	Шаг варьирования
x_1 (кмоль·м ⁻³)	1,0	0,1
x_2 (кмоль·м ⁻³)	1,5	0,2
x_3 (°C)	60,0	5,0

Пользуясь формулой (9.4) и табл. 9.3, рассчитаем условия проведения первых четырех опытов:

$$\begin{aligned}
X_{11} &= 1 + 0,1 \cdot 0,5 = 1,05, & X_{21} &= 1 + 0,1 \cdot (-0,5) = 0,95, \\
X_{12} &= 1,50 + 0,2 \cdot 0,289 = 1,56, & X_{22} &= 1,50 + 0,2 \cdot 0,289 = 1,56, \\
X_{13} &= 60 + 5 \cdot 0,204 = 61, & X_{23} &= 60 + 5 \cdot 0,204 = 61, \\
X_{31} &= 1 + 0,1 \cdot 0 = 1, & X_{41} &= 1 + 0,1 \cdot 0 = 1, \\
X_{32} &= 1,50 + 0,2 \cdot (-0,578) = 1,38, & X_{42} &= 1,50 + 0,2 \cdot 0 = 1,5, \\
X_{33} &= 60 + 5 \cdot 0,204 = 61, & X_{43} &= 60 + 5 \cdot (-0,612) = 57.
\end{aligned}$$

Полученные результаты сведем в табл. 9.5. Здесь первый индекс обозначает номер опыта, а второй — номер фактора.

Сравнивая между собой результаты первых четырех опытов, видим, что самый низкий выход целевого продукта получился в третьем опыте. Этот опыт следует исключить из дальнейшего рассмотрения.

Заменим его опытом № 5, условия проведения которого рассчитаем по формуле (9.5):

$$\begin{aligned}
X_{51} &= 2/3(1,05 + 0,905 + 1 + 1 - 1) - 1 = 1; \\
X_{52} &= 2/3(1,56 + 1,56 + 1,38 + 1,5 - 1,38) - 1,38 = 1,7; \\
X_{53} &= 2/3(61 + 61 + 61 - 57 - 67) - 67 = 58.
\end{aligned}$$

Таблица 9.5

Условия и результаты планирования по симплексному методу

Номер опыта	X ₁	X ₂	X ₃	Функция отклика
1	1,05	1,56	61	72,3
2	0,905	1,56	61	70,1
3	1,00	1,38	61	65,4
4	1,00	1,50	57	68,2
5	1,00	1,70	58	73,9
6	1,00	1,72	63	76,5

В новом симплексе, образованном опытами № 1, 2, 4 и 5, самым «неудачным» является опыт № 4. Его заменим опытом № 6, условия которого найдем, пользуясь той же формулой (9.5).

Далее процедура оптимизации может быть продолжена аналогично.

Рассмотрим теперь вопрос о том, как включить в программу исследований еще один фактор, например скорость вращения мешалки. Пусть до этих пор она была постоянной и равной 500 об/мин. Теперь будем считать эту величину фактором x_4 и примем для нее шаг варьирования $\Delta x_4 = 100$ об/мин.

Предыдущий симплекс для трех факторов (табл. 9.5) состоит из опытов № 1, 2, 5 и 6. Для того чтобы из него получить новый симплекс для четырех факторов, введем опыт № 7 (табл. 9.6).

Условия проведения опыта № 7 найдем по формулам (9.6) и (9.7):

$$X_{71} = 1/4(1,05 + 0,95 + 2 \cdot 1,00) = 1,00,$$

$$X_{72} = 1/4(2 \cdot 1,56 + 1,70 + 1,72) = 1,64,$$

$$X_{73} = 1/4(2 \cdot 61 + 58 + 63) = 61,$$

$$X_{74} = 500 + 100(0,632 + 0,158) = 579 \approx 580.$$

Далее оптимизацию можно продолжить с учетом всех четырех факторов, пользуясь рассмотренной выше процедурой.

Таблица 9.6

Симплексный план эксперимента для четырех факторов

Номер опыта	X_1	X_2	X_3	X_4	Функция отклика
1	1,05	1,56	61	500	72,3
2	0,95	1,56	61	500	70,1
5	1,00	1,70	58	500	73,9
6	1,00	1,72	63	500	76,5
7	1,00	1,64	61	580	78,1

Таким образом, при симплекс – планировании:

1) удастся резко снизить число экспериментов по сравнению с методом полного факторного эксперимента, где, кроме того, добавление каждого нового фактора требует удвоения всего числа экспериментов, а при симплекс –

планировании – только одного нового опыта (если выбрано правильное направление) и еще одного (если выбрано неправильное направление);

2) получаемые результаты не зависят от формы поверхности отклика, так как из всех данных нас интересуют худшие результаты, и при отрицательных результатах экспериментатор возвращается назад и повторяет «кантование» симплекса;

3) не требуется проведения расчетов. Метод может быть также применен при изучении процессов, в которых функцию выхода нельзя измерить количественно, а можно только оценить полуколичественно или даже чисто качественно. При этом правила движения к оптимуму не теряют своей строгости.

Вместе с тем, используя метод симплекс – планирования:

- мы никогда не сможем оценить роль отдельных факторов;
- при исследовании сложных процессов не получим никакой информации о взаимодействии факторов.

К тому же экспрессность метода симплекс – планирования проявляется в полной мере лишь в тех случаях, когда затраты времени на проведение самого эксперимента незначительны и основное время экспериментатора уходит на расчеты (в случае постановки полного факторного эксперимента). В тех же случаях, когда эксперимент по своей природе является длительным (недели и месяцы), применение метода симплекс – планирования нерационально, так как последовательность получения точек может растянуться на неопределенно долгий срок, ибо построение следующего симплекса невозможно, прежде чем не будет реализован предыдущий. В этом случае целесообразно использование метода полного факторного эксперимента, позволяющего одновременно поставить хотя и большее число вариантов, но зато получить более полное представление о влиянии факторов и условиях движения к оптимуму.

9.3. Контурно–графический анализ

При изучении формы поверхности отклика исследователи в ряде случаев обходятся без составления математического описания функции отклика, используя приемы контурно–графического анализа [18]. Сущность его состоит в определенном расположении опытов в факторном пространстве, получении дополнительной информации путем линейной интерполяции экспериментальных данных и построения на факторной плоскости (или на двумерных сечениях) линий постоянного уровня функции отклика. Экспериментальные точки располагают таким образом, чтобы они охватывали всю область факторного пространства, представляющую интерес для исследователя. Схема, предложенная математиком П. Берчем, предполагает постановку шести опытов (рис. 9.3). Один из них проводится в центре исследуемой области, а остальные — в вершинах пятиугольника.

Схема математика В. Клеймана (рис. 9.4) требует постановки четырех опытов в вершинах прямоугольника и двух опытов – на его оси симметрии.

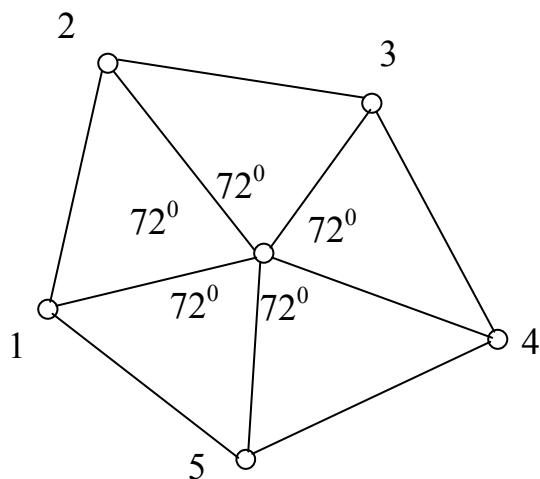


Рис. 9.3. Схема П. Берча

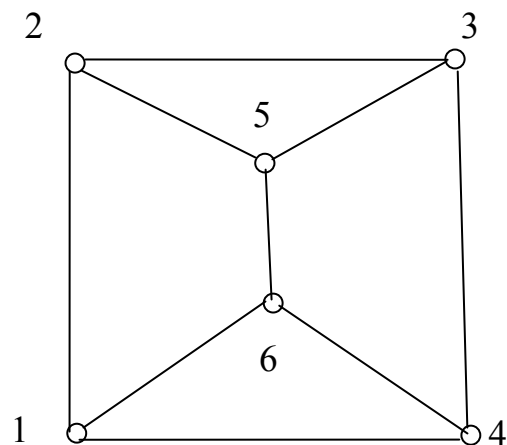


Рис. 9.4. Схема В. Клеймана

Разумеется, схемы П. Берча и В. Клеймана не исчерпывают всех возможных вариантов расположения экспериментальных точек в факторном пространстве. Однако следует иметь в виду, что расстояния между этими точками не должны быть слишком велики. В противном случае при большой

кривизне и сложной форме поверхности отклика погрешности линейной интерполяции наложат заметный отпечаток на результаты исследований.

На рис. 9.5 показано, как осуществляется контурно–графический анализ по схеме В. Клеймана. Сначала проводят намеченную серию опытов. Затем соединяют линейными отрезками соседние точки и находят методами линейной интерполяции значения функции отклика в серединах этих отрезков (как средне–арифметические значения результатов опытов в соединяемых экспериментальных точках). Наконец, проводят линии через точки с одинаковыми значениями функции отклика.

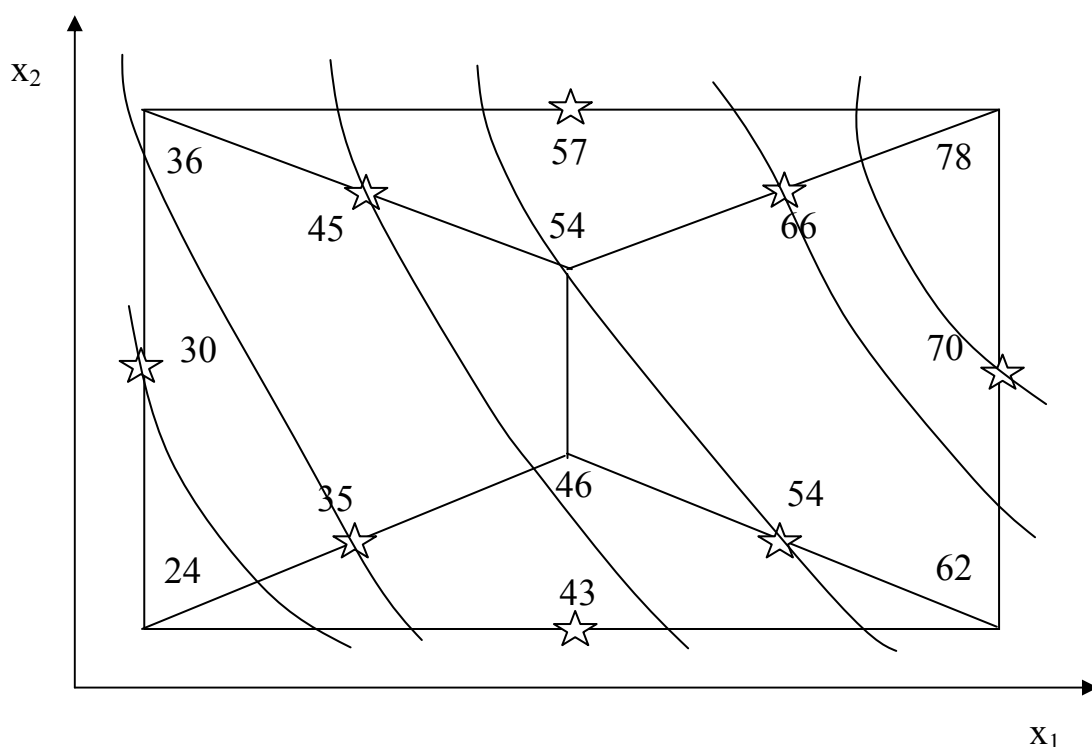


Рис. 9.5. Контурно–графический анализ по схеме Клеймана

Контурно–графический анализ по схеме П. Берча проводится аналогично.

Рассмотренной методикой можно пользоваться при построении двумерных сечений поверхности отклика.

Методом контурно–графического анализа можно строить на кальке контурные линии различных функций отклика, характеризующих процесс. Совмещая координатные оси этих графиков и просматривая кальки на свет, можно достаточно быстро выбирать оптимальные условия ведения процесса.

10. ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

10.1. Основные понятия линейного программирования

Линейным программированием называют задачи оптимизации, в которых целевая функция является линейной функцией своих аргументов, а условия, определяющие их допустимые значения, имеют вид линейных уравнений и неравенств [7].

Линейное программирование начало развиваться в первую очередь в связи с задачами экономики, с поиском способов оптимального распределения и использования ресурсов. Оно послужило основой широкого использования математических методов в экономике. Следует подчеркнуть, что в рамках реальных экономических задач число независимых переменных обычно бывает очень большим (тысячи, десятки тысяч аргументов). Поэтому практическая реализация алгоритмов решения таких задач принципиально невозможна без использования современной вычислительной техники.

Рассмотрим линейную целевую функцию с одной переменной управления:

$$F(Y, X) = A + BX + CY, \quad (10.1)$$

причем линейная модель физического процесса выражается как

$$Y = D + EX. \quad (10.2)$$

Подставив (10.1) в (10.2), получим G-форму целевой функции:

$$G(X) = A + BX + CD + CEX$$

или

$$G(X) = \psi_0 + \psi_1 X,$$

где $\psi_0 = A + CD$; $\psi_1 = B + CE$.

Видно, что при $\psi_1 > 0$ максимум достигается при $X = +\infty$, а минимум – при $X = -\infty$.

Таким образом, линейные целевые функции (как с одной переменной, так и с n -переменными) при отсутствии ограничений не имеют конечного оптимума, поэтому в задачах оптимизации целевой функции ограничения играют принципиальную роль. В дальнейшем будет показано, что совокупность любого числа линейных ограничений выделяет в пространстве

X_1, X_2, \dots, X_n некоторый выпуклый многогранник области возможных значений переменных управления. Экстремум целевой функции достигается в одной из его вершин.

При этом линиями равного уровня целевой функции являются линии, соединяющие точки, в которых значения целевой функции равны между собой.

Для линейной функции с двумя переменными управления:

$$G = \psi_0 + \psi_1 X_1 + \psi_2 X_2 ,$$

линии равного уровня, нанесенные на плоскость (X_1, X_2) , представляют собой прямые линии типа:

$$X_2 = \frac{(G - \psi_0)}{\psi_2} + \frac{\psi_1}{\psi_2} X_1 .$$

На рис. 10.1 показаны линии равного уровня целевой функции на плоскости (X_1, X_2) ; все линии параллельны между собой.

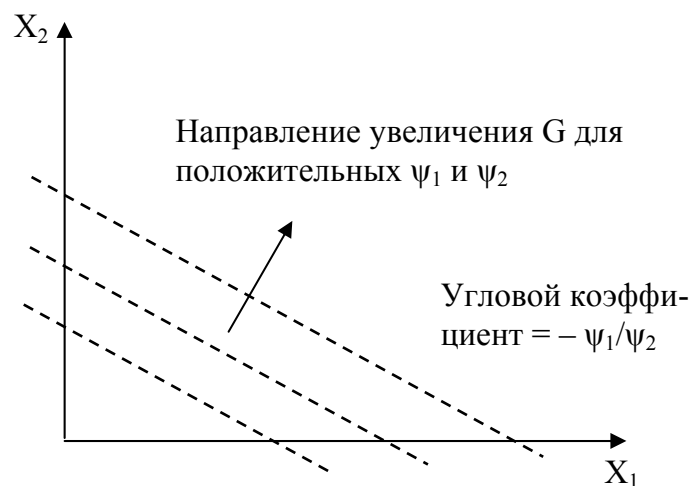


Рис. 10.1. Линии равного уровня целевой функции

Рассмотрим линейное программирование на примере максимизации функции $G = 25X_1 + 520X_2$ при условии, что ограничениями являются

$$3X_1 + X_2 \geq 8;$$

$$4X_1 + 3X_2 \geq 19;$$

$$X_1 + 3X_2 \geq 7;$$

$$0 \leq X_1 \leq 10;$$

$$0 \leq X_2 \leq 9.$$

На рис. 10.2 показаны область допустимого решения и линии равного уровня целевой функции. Координаты точек пересечения ограничивающих линий могут быть найдены алгебраическим либо графическим способом. В результате получим:

$$A(0,8); B(1,5); C(4,1); D(7,0); E(10,0); F(10,9); G(0,9).$$

Минимум находится в точке С, а максимум – в точке F, причем

$$G_{\min} = 150, \text{ а } G_{\max} = 700.$$

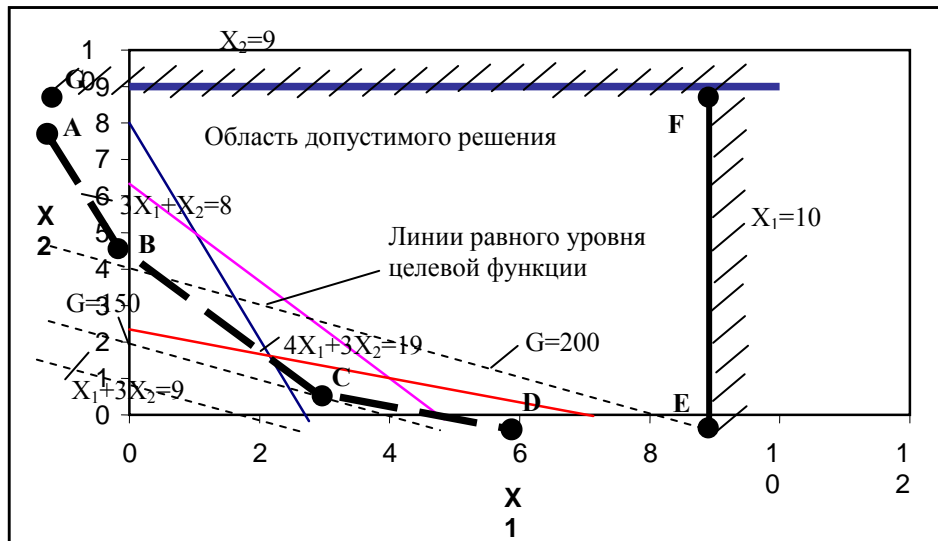


Рис. 10.2. Многогранник области возможных значений переменных управления

Заданная в стандартной форме основная задача линейного программирования состоит в следующем – найти минимальное (или максимальное) значение линейного выражения:

$$G = \psi_0 + \psi_1 X_1 + \psi_2 X_2 + \dots + \psi_m X_m$$

при наличии линейных ограничений, наложенных на неизвестные X_1, X_2, \dots, X_m , то есть при наличии ограничений вида:

$$\begin{cases} K_{11}X_1 + K_{12}X_2 + \dots + K_{1m}X_m = Q_1, \\ K_{21}X_1 + K_{22}X_2 + \dots + K_{2m}X_m = Q_2, \\ \dots \\ K_{p1}X_1 + K_{p2}X_2 + \dots + K_{pm}X_m = Q_p \end{cases}$$

и при условии неотрицательности всех переменных ($X_i > 0$).

Рассмотрим в качестве примера две типичные задачи линейного программирования: транспортную задачу и задачу о использовании ресурсов.

Транспортная задача

В городе имеется два склада цемента и два завода ЖБИ, потребляющих этот цемент. Ежедневно с первого склада вывозится 50 т цемента, со второго — 70 т. Этот цемент доставляется на заводы, причем первый завод получает 10 т, второй — 80 т. Допустим, что перевозка одной тонны цемента с первого склада на первый завод стоит 120 руб., с первого склада на второй завод — 160 руб., со второго склада на первый завод — 80 руб., и со второго склада на второй завод — 100 руб. Как нужно спланировать перевозки, чтобы их стоимость была минимальной?

Для того чтобы ответить на этот вопрос, дадим математическую постановку задачи. Обозначим через x_1 и x_2 количество цемента, который следует перевезти с первого склада соответственно на первый и второй заводы, а через x_3 и x_4 — количество цемента, который нужно перевезти со второго склада на первый и второй заводы. Эти условия приводят к системе уравнений:

$$\begin{cases} x_1 + x_2 = 50, \\ x_3 + x_4 = 70, \\ x_1 + x_3 = 40, \\ x_2 + x_4 = 80. \end{cases} \quad (10.3)$$

$$x_i \geq 0, \quad i = 1, 2, 3, 4. \quad (10.4)$$

Первые два уравнения системы определяют, сколько цемента нужно вывезти с каждого склада, два других уравнения показывают, сколько цемента нужно привезти на каждый завод. Неравенство (10.4) означает, что в обратном направлении с заводов на склады цемент не возят. Общая стоимость всех перевозок определяется формулой:

$$f = 120x_1 + 160x_2 + 80x_3 + 100x_4. \quad (10.5)$$

С математической точки зрения задача заключается в том, чтобы найти числа x_i , удовлетворяющие условиям (10.3), (10.4) и минимизировать стоимость перевозок (10.5).

Рассмотрим систему (10.3). Это система четырех уравнений с четырьмя неизвестными. Однако независимыми в ней являются только первые три уравнения, четвертое – их следствие (если сложить 1–е и 2–е уравнения и вычесть 3–е, получится 4–е). Таким образом, фактически нужно рассмотреть следующую систему, эквивалентную (10.3):

$$\begin{cases} x_1 + x_2 = 50, \\ x_3 + x_4 = 70, \\ x_1 + x_3 = 40. \end{cases} \quad (10.6)$$

В ней число уравнений на единицу меньше числа неизвестных, так что мы можем выбрать какое–нибудь неизвестное, например x_1 , и выразить через него с помощью уравнений (10.6) три остальные.

Соответствующие формулы имеют вид:

$$\begin{cases} x_2 = 50 - x_1, \\ x_3 = 40 - x_1, \\ x_4 = 70 - x_3 = 70 - 40 + x_1 = 30 + x_1. \end{cases} \quad (10.7)$$

Учитывая (10.4), получаем систему:

$$\begin{cases} x_1 \geq 0, \\ 50 - x_1 \geq 0, \\ 40 - x_1 \geq 0, \\ 30 + x_1 \geq 0, \end{cases} \quad (10.8)$$

из которой:

$$0 \leq x_1 \leq 40. \quad (10.9)$$

Таким образом, задавая любое x_1 , удовлетворяющее (10.9), и вычисляя x_2 , x_3 , x_4 по формулам (10.7), мы получим один из возможных планов перевозки. При реализации этого плана с каждого склада будет вывезено и на каждый завод доставлено нужное количество цемента.

Вычислим стоимость перевозок. Для этого подставим (10.9) в формулу (10.5). В результате получим $G = 14200 - 20x_1$.

Эта формула определяет величину G как функцию одной переменной x_1 , которую можно выбирать произвольно в пределах условий (10.9). Стоимость окажется минимальной, если мы придадим величине x_1 наибольшее возможное значение: $x_1 = 40$. Значения остальных величин x_i находятся по формулам (10.7).

Итак, оптимальный по стоимости план перевозок имеет вид

$$\begin{cases} x_1 = 40, \\ x_2 = 10, \\ x_3 = 0, \\ x_4 = 70. \end{cases}$$

Стоимость перевозок в этом случае составляет 13400 руб. При любом другом допустимом плане перевозок она окажется выше: $G > G_{\min}$.

Задача о использовании ресурсов

Мебельная фабрика выпускает стулья двух типов. На изготовление одного стула первого типа, стоимостью 800 руб., расходуется 2 п.м досок стандартного сечения, $0,5 \text{ м}^2$ обивочной ткани и 2 чел.–ч рабочего времени. Для стульев второго типа аналогичные данные составляют: 1200 руб., 4 п.м, $0,25 \text{ м}^2$ и 2,5 чел.–ч.

Допустим, что в распоряжении фабрики имеется 440 п.м досок, 65 м^2 обивочной ткани, 320 чел.–ч рабочего времени. Какое количество стульев каждого типа надо изготовить, чтобы в рамках этих ресурсов стоимость произведенной продукции была максимальной?

Для ответа на этот вопрос постараемся опять сформулировать задачу как математическую. Обозначим через x_1 и x_2 запланированное к производству число стульев соответственно первого и второго типов. Ограниченный запас сырья и трудовых ресурсов означает, что x_1 и x_2 должны удовлетворять неравенствам

$$\begin{cases} 2x_1 + 4x_2 \leq 440, \\ 1/2x_1 + 1/4x_2 \leq 65, \\ 2x_1 + 5/2x_2 \leq 320. \end{cases} \quad (10.10)$$

Кроме того, по смыслу задачи они должны быть неотрицательными:

$$\begin{cases} x_1 \geq 0, \\ x_2 \geq 0. \end{cases} \quad (10.11)$$

Стоимость запланированной к производству продукции определяется формулой:

$$G(x_1, x_2) = 800x_1 + 1200x_2. \quad (10.12)$$

Итак, с математической точки зрения задача составления оптимального по стоимости выпущенной продукции плана фабрики сводится к определению пары целых чисел x_1 , x_2 , удовлетворяющих линейным неравенствам (10.10), (10.11) и дающих наибольшее значение линейной функции (10.12). Мы опять получили типичную задачу линейного программирования. По своей постановке она несколько отличается от транспортной задачи, однако это различие не существенно.

Для анализа сформулированной задачи рассмотрим плоскость и введем на ней декартову систему координат x_1 , x_2 . Найдем на этой плоскости множество точек, координаты которых удовлетворяют (10.10), (10.11). Неравенства (10.11) означают, что это множество лежит в первой четверти. Выясним смысл ограничений, которые задаются неравенствами (10.10). Проведем на плоскости прямую, определяемую уравнением (рис. 10.3)

$$2x_1 + 4x_2 = 440. \quad (10.13)$$

Она делит плоскость на две полуплоскости. На одной из них, расположенной ниже прямой (10.13), функция $f_1(x_1, x_2) = 2x_1 + 4x_2 - 440$ принимает отрицательные значения; на другой, расположенной выше прямой (10.13), — положительные. Таким образом, первое из неравенств (10.10) выполняется на множестве точек, которое включает в себя прямую (10.13) и полуплоскость, расположенную ниже этой прямой. На рисунке соответствующая часть плоскости заштрихована.

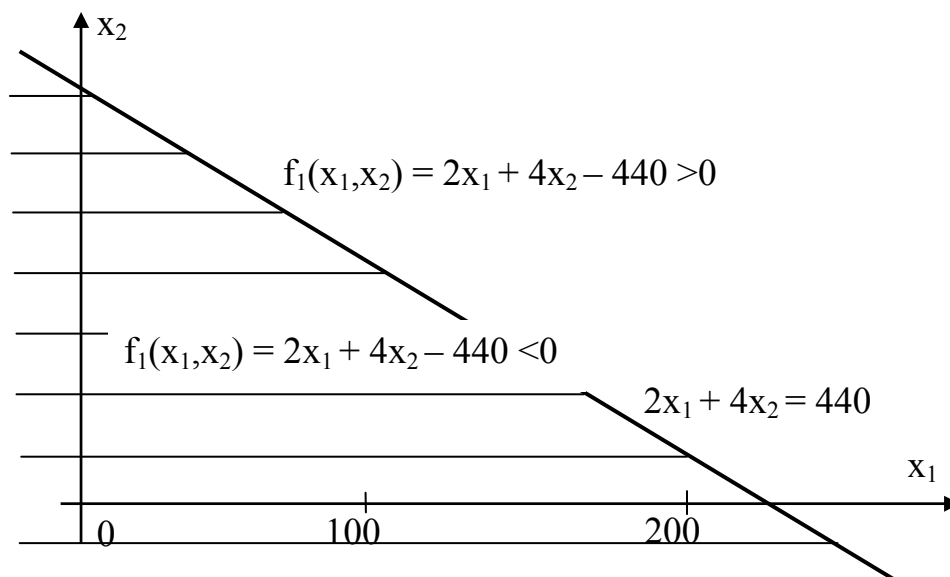


Рис. 10.3. Решение неравенства $2x_1 + 4x_2 \leq 440$

Совершенно аналогично можно найти множества точек, удовлетворяющих второму и третьему неравенствам из системы (10.10). Они показаны на рис. 10.4, 10.5 штриховкой.

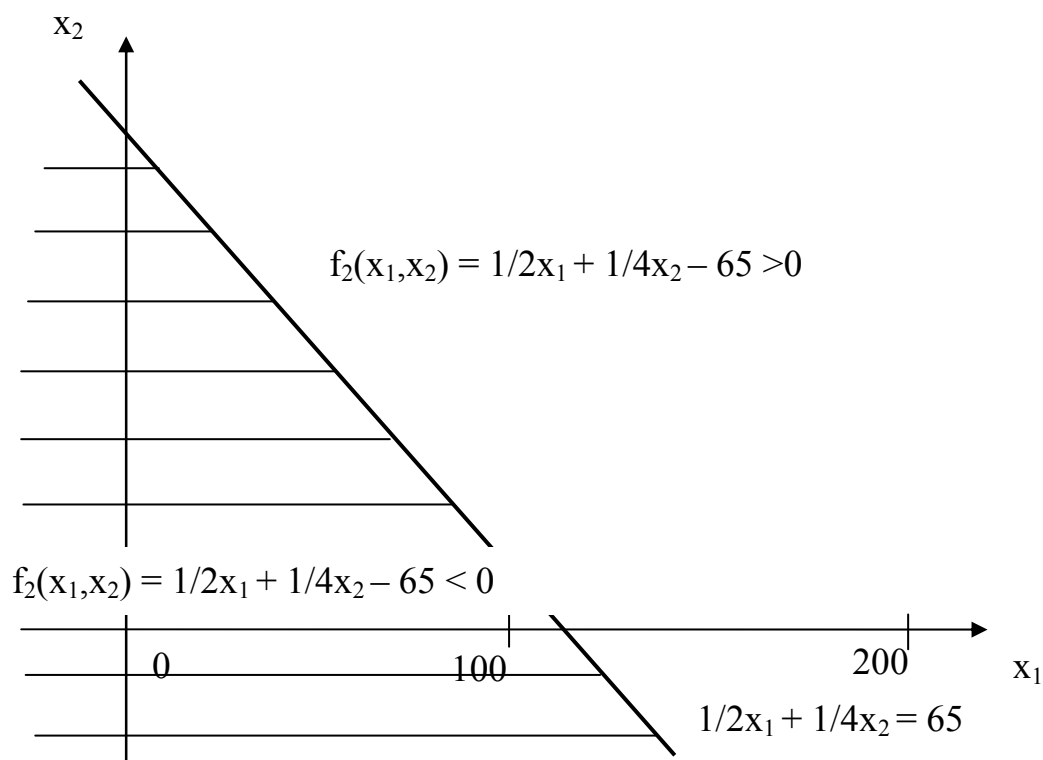


Рис. 10.4. Решение неравенства $1/2x_1 + 1/4x_2 \leq 65$

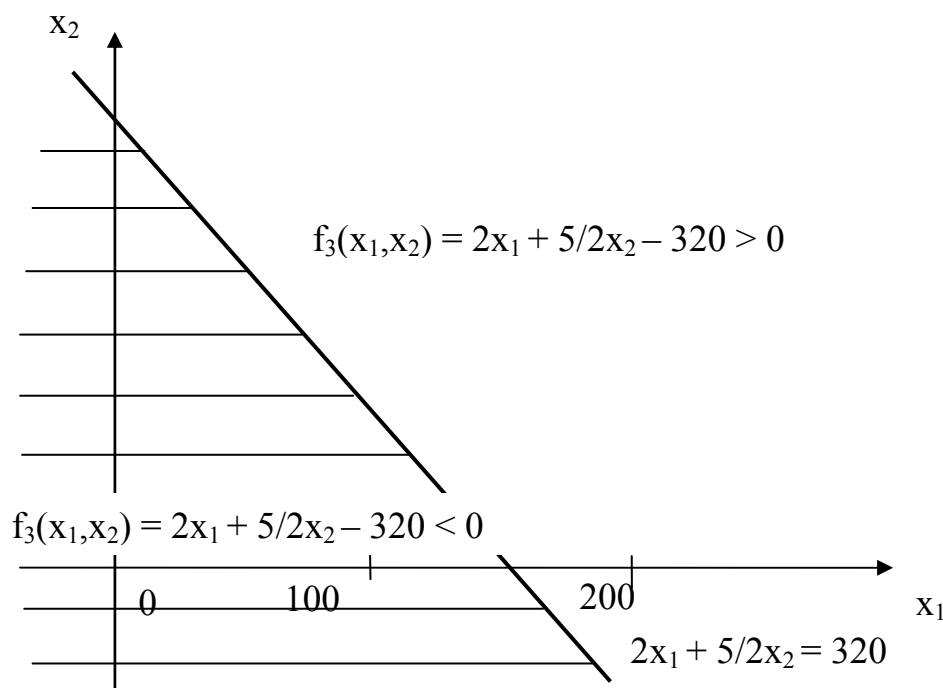


Рис. 10.5. Решение неравенства $2x_1 + 5/2x_2 \leq 320$

Возьмем пересечение трех найденных множеств и выделим его часть, расположенную в первой четверти. В результате получим множество точек, удовлетворяющих всей совокупности ограничений (10.10), (10.11). Данное множество имеет вид пятиугольника, показанного на рис. 10.6. Его вершинами являются точки пересечения прямых, на которых неравенства (10.10), (10.11) переходят в точные равенства. Координаты вершин пятиугольника указаны на рисунке.

Любой точке P с целочисленными координатами (x_1, x_2) , принадлежащей данному пятиугольнику, соответствует план выпуска стульев, который может быть выполнен при имеющихся запасах сырья и трудовых ресурсах (реализуемый план). Наоборот, если точка P не принадлежит пятиугольнику, то соответствующий план не может быть выполнен (нереализуемый план).

Рассмотрим на плоскости x_1, x_2 линии уровня целевой функции (10.12):

$$800x_1 + 1200x_2 = C. \quad (10.14)$$

Это уравнение описывает семейство прямых, параллельных прямой

$$800x_1 + 1200x_2 = 0.$$

При параллельном переносе этой прямой вправо параметр C возрастает, влево — убывает.

Свойства функции (10.12) тесно связаны с прямыми (10.14). Вдоль каждой из них она сохраняет постоянное значение, равное C , а при переходе с одной прямой на другую ее значение меняется.

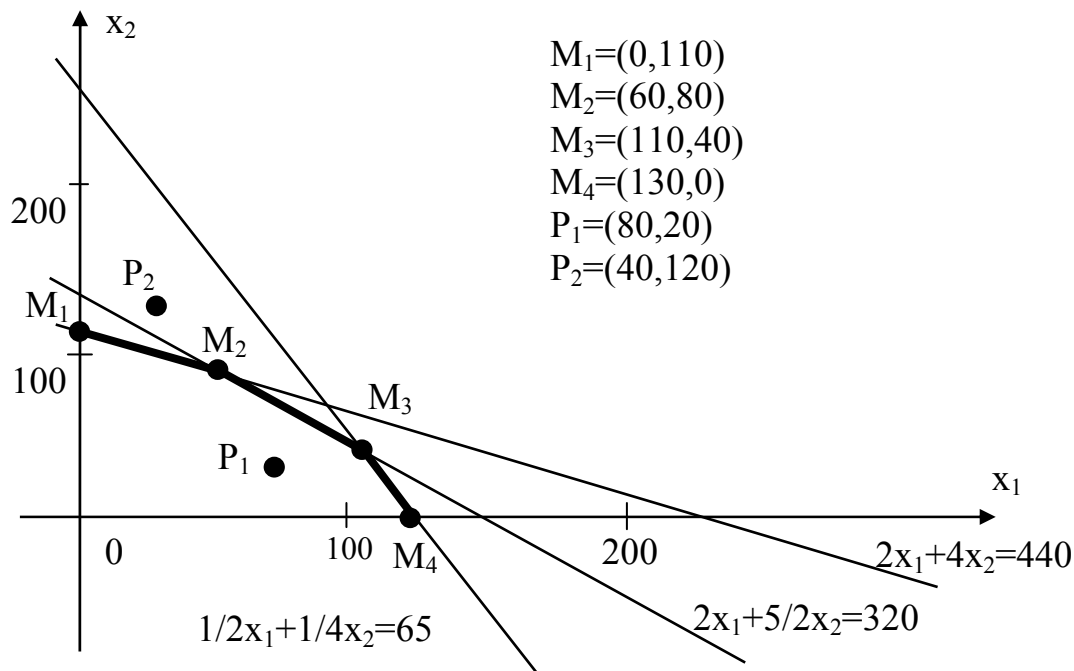


Рис. 10.6. Пятиугольник $OM_1M_2M_3M_4$, точки которого удовлетворяют системе неравенств (10.10), (10.11)

Будем рассматривать только первую четверть. Предположим, что мы перешли из точки P_1 , расположенной на одной прямой, в точку P_2 , расположенную на другой прямой (рис. 10.7). Если вторая прямая расположена дальше от начала координат, чем первая, то функция G при этом переходе возрастет. Отсюда следует важный вывод: оптимальный план должен располагаться на прямой семейства (10.14), наиболее удаленной от начала координат.

Этот вывод позволяет закончить решение задачи. Рассмотрим рис. 10.8. На нем воспроизведен пятиугольник реализуемых планов и нарисована прямая семейства (10.14), проходящая через точку M_2 с координатами $(60, 80)$. Она является предельной прямой семейства, имеющей общую точку с пятиугольником. Если мы попытаемся с помощью параллельного переноса отодвинуть ее дальше от начала координат, то получим прямую, не имеющую общих точек с пятиугольником, т. е. соответствующие планы нереализуемы.

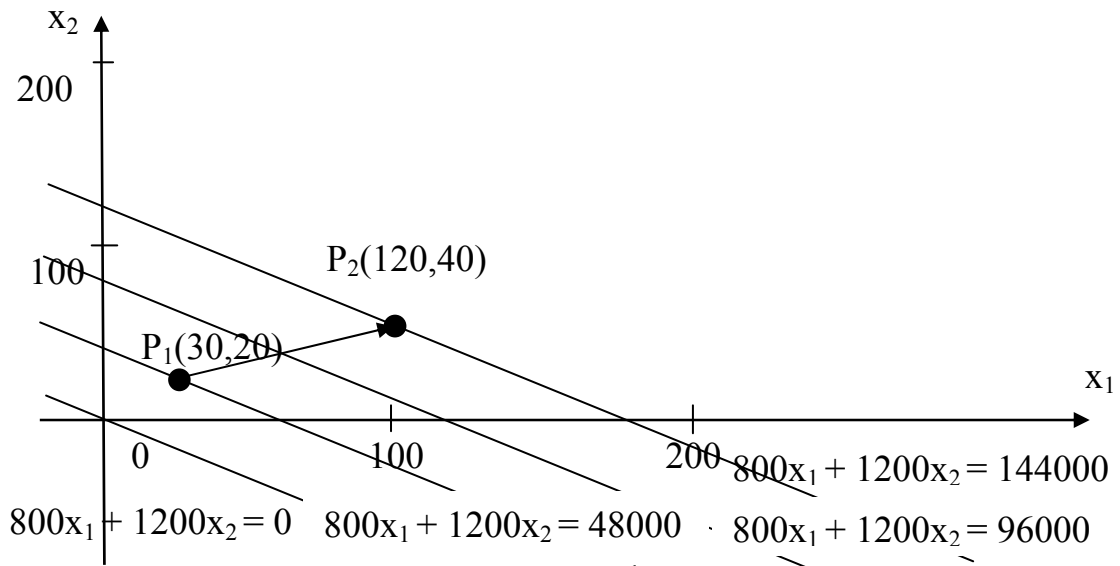


Рис. 10.7. Линии равного уровня функции $G(x_1, x_2) = 800x_1 + 1200x_2$

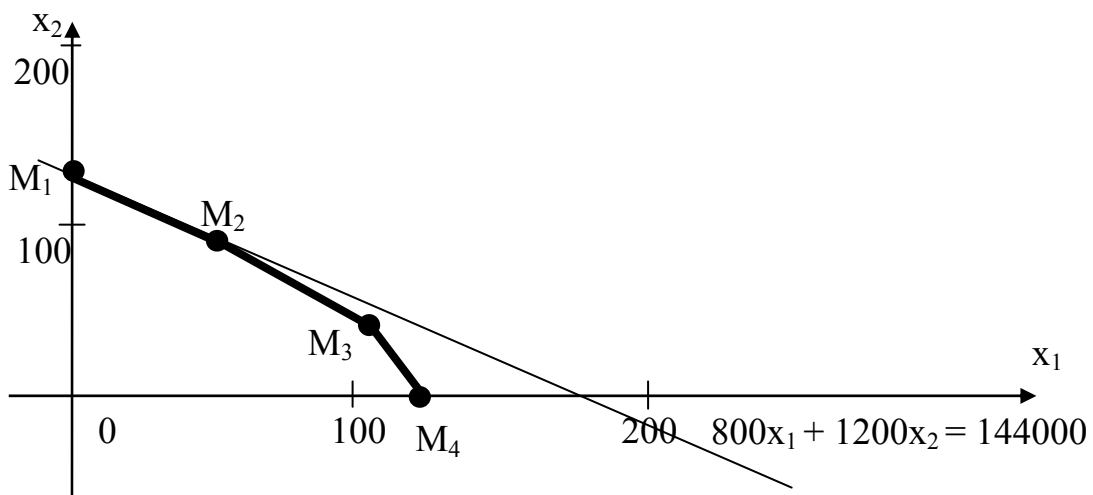


Рис. 10.8. Определение оптимального плана производства стульев

Итак, оптимальный план найден, – он предписывает производство 60 стульев первого типа и 80 стульев второго типа. Стоимость этой продукции 144000 руб. На выполнение плана нужно затратить: 440 п.м досок, 50 м² обивочной ткани, 320 чел. – ч рабочего времени.

Видно, что оптимальный план требует полного использования запаса досок и трудовых ресурсов, в то время как обивочная ткань будет израсходована не полностью – останется 15 м².

Этот результат ясен из рис. 10.8. Точка M_2 , определяющая оптимальный план, является вершиной пятиугольника.

Она лежит на пересечении прямых $2x_1 + 4x_2 = 440$ и $2x_1 + 5/2x_2 = 320$.

Уравнения этих прямых получаются из первого и третьего условий системы (10.10) при замене их на строгие равенства. Это означает полный расход досок и трудовых ресурсов. Однако точка M_2 не принадлежит прямой:

$$1/2x_1 + 1/4x_2 = 65,$$

так что второе условие (10.10) связано с ограниченным запасом обивочной ткани, имеет форму неравенства $50 < 65$.

Проведенный анализ показывает, что дальнейшее увеличение стоимости продукции регламентируется запасом досок и трудовыми ресурсами.

10.2. Симплекс–метод линейного программирования

Если количество переменных управления больше двух или трех, применять графический способ очень трудно, в этом случае удобнее применять аналитические методы, в частности симплекс–метод. Применительно к линейному программированию решение задачи разбивается на два этапа. На первом этапе находится какое–нибудь (даже не самое удачное) решение, удовлетворяющее поставленной задаче и ограничениям (начальный план). На втором этапе производится последовательное улучшение плана по определенным правилам, пока дальнейшее улучшение станет невозможным. Условимся называть переменные, которые входят в план, основными (базисными) переменными, остальные – неосновными.

Вычисления следует вести по следующей схеме:

а) отыскивается начальный план, где основные переменные этого плана выражаются через неосновные;

б) выражается значение минимизируемой функции G через неосновные переменные;

в) выбирается та из неосновных переменных, введение которой в план способно улучшить G ;

г) определяется, какая из основных переменных должна быть при этом исключена из плана и сделана неосновной;

д) вновь вводимая в план переменная выражается через переменную, выводимую из плана, и другие неосновные переменные;

е) все остальные основные переменные и значение минимизируемой функции G выражаются через новые неосновные переменные;

ж) повторение предыдущих операций.

Рассмотрим решение задачи на минимизацию стоимости топлива.

Пусть имеется несколько сортов топлива с различной зольностью, теплотворной способностью и стоимостью (табл. 10.1).

Таблица 10.1

Исходные данные

Сорт топлива	Зольность, кг	Теплотворная способность, ккал	Стоимость, \$
	в расчете на 1 кг		
Первый	0,02	10	0,10
Второй	0,04	8	0,06
Третий	0,08	4	0,02

Требуется определить стоимость закупки, а также сколько и какого топлива следует приобрести, если оговорено, что общая теплотворность топлива должна быть 4000 ккал, а общее количество золы – не превышать 56 кг.

Обозначим X_1, X_2, X_3 – неизвестное количество килограммов каждого топлива в наиболее целесообразном варианте заказа. Тогда общая стоимость топлива будет $G = 0,1X_1 + 0,06X_2 + 0,02X_3$, при этом должны выполняться следующие ограничения:

$$10X_1 + 8X_2 + 4X_3 = 4000, \quad (10.15)$$

$$0,02X_1 + 0,04X_2 + 0,08X_3 \leq 56.$$

Для того чтобы последнее неравенство превратить в равенство, введем дополнительную величину X_4 (количество фиктивного вещества, содержащего 1 %

золы, не дающего калорий, но и не имеющего стоимости) и перепишем последнее условие в виде $0,02X_1 + 0,04X_2 + 0,08X_3 + 0,01X_4 = 56$ или, после умножения на 100,

$$2X_1 + 4X_2 + 8X_3 + X_4 = 5600. \quad (10.16)$$

Определяем начальный план, из (10.15) получаем:

$$X_1 = 400 - 0,8X_2 - 0,4X_3. \quad (10.17)$$

Из выражения (10.16) находим:

$$X_4 = 5600 - 2X_1 - 4X_2 - 8X_3. \quad (10.18)$$

Подставив (10.17) в (10.18), получим:

$$X_4 = 5600 - 2(400 - 0,8X_2 - 0,4X_3) - 4X_2 - 8X_3,$$

или:

$$X_4 = 4800 - 2,4X_2 - 7,2X_3. \quad (10.19)$$

Примем в качестве основных переменных X_1 и X_4 , в качестве неосновных – X_2 и X_3 . Выразим через неосновные переменные функцию G .

$$G = 0,1(400 - 0,8X_2 - 0,4X_3) + 0,06X_2 + 0,02X_3,$$

или:

$$G = 40 - 0,02X_2 - 0,02X_3. \quad (10.20)$$

Теперь цель достигнута.

В качестве начального плана можно принять $X_2 = X_3 = 0$ (при этом G минимальна и равняется 40 \$); $X_1 = 400$; $X_4 = 4800$.

Постараемся уменьшить G .

Из выражения (10.20) видно, что введение X_2 в план способно уменьшить стоимость закупки (с увеличением X_2 значение G уменьшается). Посмотрим, в каких пределах можно допустить приобретение этого топлива. Из (10.18) при условии $X_1 \geq 0$

$$0 = 400 - 0,8X_2,$$

тогда $X_2 = 500$ кг. Аналогично из (10.19) при условии $X_4 \geq 0$

$$0 = 4800 - 2,4X_2.$$

Соответственно $X_2 = 2000$ кг.

Остановимся на первом значении. Меняем основные и неосновные переменные X_1 и X_2 .

$$\text{Из (10.18) } X_2 = \frac{1}{0,8}(400 - X_1 - 0,4X_3) = 500 - 1,25X_1 - 0,5X_3. \quad (10.21)$$

Подставив (10.21) в (10.19) и (10.20), имеем

$$\begin{aligned} X_4 &= 4800 - 2,4(500 - 1,25X_1 - 0,5X_3) - 7,2X_3, \\ X_4 &= 3600 + 3X_1 - 6X_3, \end{aligned} \quad (10.22)$$

$$G = 40 - 0,02(500 - 1,25X_1 - 0,5X_3) - 0,02X_3,$$

$$G = 30 + 0,025X_1 - 0,01X_3.$$

Вновь полученный план выглядит следующим образом (при G_{\min}):
введем X_3 вместо X_4 , используя выражение (10.22), получаем

$$X_3 = 600 + 0,5X_1 - \frac{1}{6}X_4,$$

$$X_2 = 500 - 1,25X_1 - \frac{600 + 0,5X_1 - \frac{1}{6}X_4}{2} = 200 - 1,5X_1 + \frac{1}{12}X_4,$$

$$G = 30 + 0,025X_1 - 0,01(X_3 = 600 + 0,5X_1 - \frac{1}{6}X_4),$$

$$G = 24 + 0,02X_1 + 0,001667X_4.$$

Полученный план: $X_1 = X_4 = 0$; $X_2 = 200$; $X_3 = 600$ и $G = 20$ \$.

То обстоятельство, что введение X_1 или X_4 в план не уменьшит стоимость закупки, указывает, что полученный план является наилучшим.

10.3. Моделирование планирования выпуска продукции

Рассмотрим некоторые определения [5]. *Организация* – сообщество людей, совместно реализующих некоторую программу или цель и действующих на основе определенных процедур и правил.

Предположим следующую структуру организации: *Предприятие* – *Совет директоров*.

Первый элемент – предприятие, производящее продукцию.

Обозначим количество продукции за период времени – Y , а затраты труда, сырья, энергии на выпуск продукции – Z .

Пара (Y, Z) – состояние организации. Очевидно, что существует функциональная зависимость Z_{\min} от Y . Обозначим ее $\varphi(Y)$ – функцией производственных издержек. Реальные затраты Z могут быть выше по причине недостаточной организации производства или недостаточной заинтересованности рабочих в максимальной эффективности производства. Поэтому возможным состоянием предприятия может быть любая пара (Y, Z) (рис. 10.9), так что $Z > \varphi(Y)$.

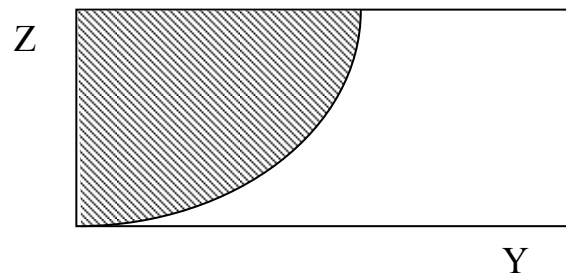


Рис. 10.9. Зависимость затрат от выпуска продукции

Второй элемент – Совет директоров. Он сам ничего не производит, а является органом управления предприятием. Нужны рычаги управления. В плановой системе – это план. Утрированно Совет директоров планирует предприятию только объем выпуска X (рис. 10.10).

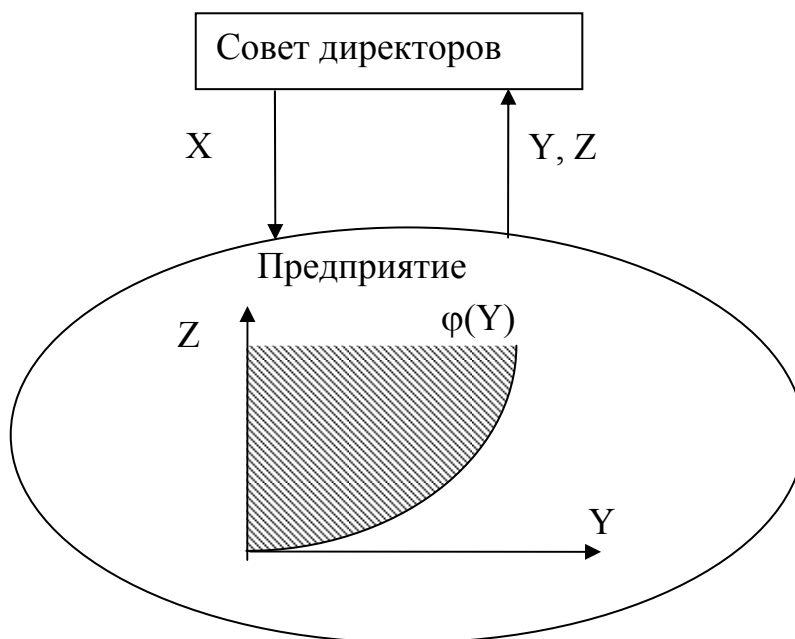


Рис. 10.10. Схема структуры Совет директоров – Предприятие

Механизм функционирования

План должен быть реально выполним и достаточно высок. Для этого *Совет директоров* (центр) должен знать возможности предприятия. Как правило, центр хуже знает возможности элементов системы, чем сами элементы.

Для эффективного руководства центр должен получить информацию от элементов о их возможностях. Таким образом, способ получения данных является одной из составных частей организационного механизма. Как правило, функция $\varphi(Y)$ описывается в виде параболы

$$\varphi(Y) = \frac{1}{2r} Y^2,$$

где r характеризует эффективность предприятия (чем она выше, тем с меньшими затратами осуществляется производство).

Простейший способ получить информацию – это спросить. Такой способ называется *встречным* (информация – снизу вверх, планы – сверху вниз).

Предприятию не всегда выгодно полностью раскрывать все свои резервы, то есть сообщать Z_{\min} . Поэтому в *Совет директоров* сообщается оценка S эффективности, не равная истинной, то есть $S \neq r$. Это называется свойством активности элементов организации, то есть их способность исказить информацию.

Процедура установления плана предприятию является процедурой планирования. Математически – это зависимость плана от информации, то есть $X = \pi(S)$.

Выполнение плана зависит от активности элементов организации, их способности работать с разной эффективностью, которая определяется в свою очередь заинтересованностью.

Предприятие, выпуская продукцию, получает прибыль (разность объема реализации и затрат). Если цена продукции λ , то объем реализации $\lambda \cdot Y$, а прибыль $\Pi = \lambda \cdot Y - Z$.

Желая увеличить прибыль, предприятие будет выпускать продукцию с минимальными затратами:

$$Z = \frac{1}{2r} Y^2,$$

$$\Pi = \lambda \cdot Y - \frac{1}{2r} Y^2.$$

Какой выпуск продукции обеспечит максимальную прибыль?

Преобразуем:

$$\begin{aligned} \Pi &= \lambda \cdot Y - \frac{1}{2r} Y^2 + \frac{1}{2} \lambda^2 r - \frac{1}{2} \lambda^2 r = \lambda \cdot Y \frac{2r}{2r} - \frac{1}{2r} Y^2 + \frac{1}{2} \lambda^2 r - \frac{1}{2} \lambda^2 \frac{r^2}{r} = \\ &= \frac{1}{2} \lambda^2 r - \frac{1}{2r} (Y^2 - 2\lambda r Y + \lambda r^2), \end{aligned}$$

$$\Pi = \frac{1}{2} \lambda^2 r - \frac{1}{2r} (Y - \lambda r)^2.$$

Легко заметить, что Π будет максимальна при $Y_0 = \lambda \cdot r$, то же самое можно определить при помощи производной $\Pi'(Y) = \lambda - Y/r = 0$, $Y = \lambda \cdot r$.

Если предприятие не несет никакой ответственности за невыполнение плана, то план в эту формулу не входит. Необходима система штрафов и поощрений за выполнение плана, установленного центром.

Пусть в случае невыполнения плана продукция оплачивается предприятию по цене λ_1 , меньшей, чем λ , то есть $\lambda_1 = \varepsilon \cdot \lambda$, $0 < \varepsilon < 1$. Фактически предприятие штрафует на величину $\lambda \cdot Y - \lambda_1 \cdot Y = (1 - \varepsilon) \lambda \cdot Y$.

С учетом штрафов экономический интерес предприятия можно описать зависимостью:

$$f(\lambda, X, Y) = \begin{cases} \varepsilon \cdot \lambda \cdot Y - \frac{1}{2r} Y^2, & \text{если } (Y < X) \\ \lambda \cdot Y - \frac{1}{2r} Y^2, & \text{если } (Y > X). \end{cases}$$

Эта функция называется целевой функцией предприятия, или системой стимулирования. Очевидно, что если план $X \leq \lambda \cdot r$, то он всегда будет выполнен или перевыполнен и центру никогда не придется штрафовать предприятие.

Если $X > \lambda \cdot r$, ($X = k \cdot Y$, $k > 0$), то есть прибыль при таком плане меньше максимальной, то это предприятию не интересно, так как если оно выполнит план, прибыль будет равна:

$$\lambda \cdot X - \frac{1}{2r} X^2 = \lambda \cdot k \cdot Y - \frac{1}{2r} k^2 \cdot Y^2 = \lambda \cdot Y - \frac{1}{2r} k \cdot Y^2,$$

а если не выполнит –

$$\varepsilon \cdot \lambda \cdot Y - \frac{1}{2r} Y^2.$$

Прибыль максимальна при выпуске $Y = \varepsilon \cdot \lambda \cdot r$:

$$\Pi_{\max} = \varepsilon^2 \lambda^2 r - \frac{1}{2r} \varepsilon^2 \lambda^2 r^2 = \frac{1}{2} \varepsilon^2 \lambda^2 r.$$

Если $\lambda \cdot X - \frac{1}{2r} X^2 > \Pi_{\max}$, то предприятию экономически выгодно выполнить план, если $< \Pi_{\max}$, – то невыгодно.

Решим квадратное уравнение $\lambda \cdot X - \frac{1}{2r} X^2 = \Pi_{\max}$,

его решение:

$$X_{\text{н}} = \lambda \cdot r \left(1 + \sqrt{1 - \varepsilon^2} \right) = \lambda \cdot r (1 + \eta), \quad \text{где} \quad \eta = \sqrt{1 - \varepsilon^2}.$$

Если план $\lambda \cdot r < X < X_{\text{н}}$, то прибыль при выполнении плана больше, чем при невыполнении, поэтому план будет выполнен.

Если $X > X_{\text{н}}$, то, наоборот, экономически выгоднее план не выполнять, так как он сверхнапряженный.

Таким образом, с помощью линейного моделирования можно решать как математические, так и экономические задачи.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Агекян Т.А. Основы теории ошибок для астрономов и физиков / Т.А. Агекян. М. : Наука, 1972.
2. Алексеева И.У. Теоретическое и экспериментальное исследование законов распределения погрешностей, их классификация и методы оценки их параметров : автореф. дис. ... канд. техн. наук / И.У. Алексеева. Л., 1975.
3. Амосов Н.М. Моделирование сложных систем / Н.М. Амосов. Киев: Наукова думка, 1968.
4. Бир С. Кибернетика и управление производством: пер. с англ. / С. Бир. М. : Наука, 1963.
5. Бурков В.Н. Человек. Управление. Математика / В.Н. Бурков. М. : Просвещение, 1989.
6. Емельянов А.С. Эконометрия и прогнозирование / А.С. Емельянов. М. : Экономика, 1985.
7. Вводные лекции по прикладной математике / А.Н. Тихонов [и др.]. М. : Наука, 1984.
8. Вирт Н. Алгоритмы и структуры данных / Н. Вирт. М. : Мир, 1989.
9. Гухман А.А. Введение в теорию подобия / А.А. Гухман. М. : Наука, 1963.
10. Докучаев А.А. Принципы обработки информации / А.А. Докучаев. СПб. : ТЭИ, 1995.
11. Докучаев А.А. Введение в табличный процессор MS–Excel for Windows / А.А. Докучаев, С.А. Мошенский. СПб. : ТЭИ, 1996.
12. Информатика : Энциклопедический словарь для начинающих / под ред. Д.А. Поспелова. М. : Наука, 1994.
13. Клини С. К. Введение в метаматематику : пер. с англ. / С. К. Клини. М. : Наука, 1957.
14. Ларичев О.И. Наука и искусство принятия решений / О.И. Ларичев. М. : Наука, 1979.

15. Ли Т.Г. Управление процессами с помощью вычислительных машин. Моделирование и автоматизация / Т.Г. Ли, Г.Э. Адамс, У.М. Гейнз / пер. с англ. под общ. ред. В.И. Мудрова. М : Советское радио, 1972.
16. Моисеев Н.Н. Математические методы системного анализа / Н.Н. Моисеев. М. : Наука, 1984.
17. Самарский А.А. Компьютеры, модели, вычислительный эксперимент / А.А. Самарский. М : Наука, 1988.
18. Саутин С.Н. Планирование эксперимента в химии и химической технологии / С.Н. Саутин. Л. : Химия, 1975.
19. Свириденко С.С. Современные информационные технологии / С.С. Свириденко. М. : Радио и связь, 1989.
20. Седов Л.И. Методы подобия и размерности в механике / Л.И. Седов. М. : Наука, 1972.
21. Сениченков Ю. Три урока по теме «Математическое моделирование и вычислительный эксперимент» с помощью Model Vision / Ю. Сениченков. М. : Наука, 1994.
22. Советов Б.Я. Информационная технология / Б.Я. Советов. М. : Высшая школа, 1994.
23. Хомский Н. Язык и мышление : пер. с англ. / Н. Хомский. М. : Наука, 1972.

Учебное издание

Владимир Борисович Пономарев

Александр Борисович Лошкарев

Математическое моделирование технологических процессов

Редактор Л.Ю. Козяйчева

ИД № 06263 от 12.11.2001 г.

Подписано в печать	Формат	60x84	1/16
Бумага писчая	Офсетная печать	Усл. печ. л. 7,44	Уч.
изд. л. 5,7	Тираж	Заказ	Цена «С»

Редакционно-издательский отдел ГОУ ВПО УГТУ-УПИ

620002, Екатеринбург, ул. Мира, 19