

ОГЛАВЛЕНИЕ

	Лекция 1	
Введение.....		2
Глава 1. Цели и задачи математического моделирования процессов и систем		2
1.1. Понятие «математическая модель».....		2
1.2. Классификация математических моделей.....		4
Контрольные вопросы к лекции 1		5
	Лекция 2	
1.3. Геометрическое представление математических моделей		6
Глава 2. Теоретические математические модели аналитического типа		7
2.1. Построение математической модели сверления лазером		8
Контрольные вопросы к лекции 2		9
	Лекция 3	
2.2. Линейные математические модели		10
2.3. Исследование простейшей математической модели работы газотурбинного двигателя.....		12
2.4. Нелинейные детерминированные модели		14
2.4.1. Полиномиальные модели		14
2.4.2. Позиномные модели		15
Контрольные вопросы к лекции 3		15
	Лекция 4	
(2.4.3. Математическая модель кратчайшего пути		16
Контрольные вопросы к лекции 4		18
	Лекция 5	
2.5. Математическая модель в виде обыкновенных дифференциальных уравнений		19
2.6. Модели, заданные в виде уравнений в частных производных		20
Контрольные вопросы к лекции 5		22
	Лекция 6	
2.7. Стохастические модели.....		23
Контрольные вопросы к лекции 6		25
	1 рубежный контроль	
	Лекция 7	
Глава 3. Эмпирические математические модели.....		26
3.1 Идентификация эмпирических математических моделей		26
3.2. Использование метода наименьших квадратов		28
Контрольные вопросы к лекции 7		29
	Лекция 8	
3.3. Статистические методы проверки адекватности математических моделей.....		30
Контрольные вопросы к лекции 8		32
	Лекция 9	
3.4. Идентификация параметров математической модели силы резания токарной операции		33
Контрольные вопросы к лекции 9		35
	Лекция 10	
3.5. Выбор оптимальной эмпирической модели		36
3.6. Использование критерия Фишера для проверки значимости высших степеней математической модели		37
Контрольные вопросы к лекции 10		38
	Лекция 11	
Глава 4. Математические модели теории принятия решений		39
4.1. Общие сведения о теории принятия решений.....		39
4.2. Общая математическая модель формирования оптимальных решений		40
4.3. Построение и решение оптимизационной задачи принятия решения (Задача о баке)		40
Контрольные вопросы к лекции 11		42
	Лекция 12	
4.4. Многокритериальные задачи принятия решений		43
4.5. Построение решений, оптимальных по Парето (Двухкритериальная задача о баке)		44
Контрольные вопросы к лекции 12		45

ВВЕДЕНИЕ

Вторая половина XX века связана с появлением и широким распространением новой методологии исследования сложных объектов и систем. В ее основе лежит метод математического моделирования и реализованные на его основе вычислительные эксперименты. Математические модели использовались и раньше. Они позволяли уже тогда анализировать недоступные или несуществующие объекты и процессы. Например:

- 1) Планета Уран была открыта путем анализа возмущений орбит трех планет (Левеверье).
- 2) К.Э. Циолковский показал, что для преодоления земного притяжения требуется первая космическая скорость, а не скорость света.

Однако считалось, что методы математического моделирования не пригодны для исследования сложных технических, экономических, биологических и социальных систем. В области техники отсутствие объективных математических методов привело, с одной стороны, к созданию многочисленных частных, так называемых инженерных методик расчета, носивших рецептурный характер, а с другой – к полному безраздельному господству эмпирики (натурных экспериментов).

Недостаточно полная проработка вариантов приводила к субъективным решениям.

Положение начало меняться во второй половине XX в. при развитии средств вычислительной техники, в частности современных ЭВМ, которое дало в руки исследователей новое эффективное средство моделирования сложных систем. В настоящее время не существует объектов, при изучении которых не применялись бы методы математического моделирования. Разработаны и активно используются математические модели технических устройств, модели разнообразных технологических процессов, экономические модели предприятий, регионов и целых государств, экологические модели, модели геологических и геофизических процессов, модели социальных систем, биологические и медицинские модели.

Глава 1. ЦЕЛИ И ЗАДАЧИ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ПРОЦЕССОВ И СИСТЕМ

1.1. ПОНЯТИЕ «МАТЕМАТИЧЕСКАЯ МОДЕЛЬ»

Математическое моделирование позволяет до создания реальной системы (объекта) или возникновения реальной ситуации рассмотреть возможные режимы работы, выбрать оптимальные управляющие воздействия, составить объективный прогноз будущих состояний системы.

Вычислительные эксперименты, проводимые на основе математических моделей, помогают увидеть за частным общее, развить универсальные методы анализа объектов различной физической природы, познать свойства изучаемых процессов и систем.

Наконец, математическое моделирование является основой интенсивно разрабатываемых автоматизированных систем проектирования, управления и обработки данных.

|| *Основная задача математического моделирования – выделение законов в природе, обществе и технике и запись их на языке математики.*

Например:

- 1) Зависимость между массой тела m , действующей на него силой F и ускорением его движения a записывается в форме 2-го закона Ньютона: $F = m \cdot a$;
- 2) Зависимость между напряжением в электрической цепи U , ее сопротивлением R и силой тока I записывается в виде закона Ома: $I = U/R$.

Существует множество определений математической модели.

Приведем одно из них:

Математической моделью некоторого объекта, процесса или явления будем называть запись его свойств на формальном языке с целью получения нового знания (свойств) об изучаемом процессе путем применения формальных методов.

Альтернативой формальному (математическому) подходу является экспериментальный подход. К его недостаткам можно отнести:

- 1) высокая стоимость подготовки и проведения экспериментов;
- 2) получение частного знания (знания о конкретном объекте исследования, а не о классе объектов).

Например, пусть требуется определить воздействие x на некоторый процесс или объект, при котором его результирующая характеристика y имеет максимально возможное значение (Рис. 1.1).

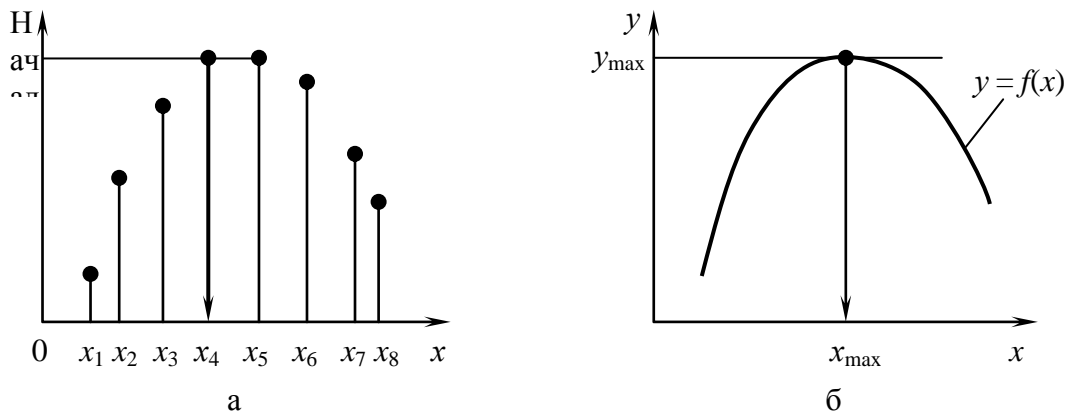


Рис. 1.1.

На рис. 1.1. а) показан эмпирический (экспериментальный) подход к решению поставленной задачи, который состоит в экспериментальном определении значения параметра y для нескольких значений входного воздействия x . Среди них найдено наибольшее, и оно принимается за максимум. Как видим из этого рисунка, возможно несколько значений воздействия x (x_4 и x_5), при которых y имеет наибольшее значение, но ни одно из них не является настоящим максимумом, который, возможно, лежит между ними.

Математический подход (рис. 1.1. б) предполагает наличие математической модели процесса типа $y = f(x)$. Взяв производную $\frac{df}{dx}$ и приравняв ее к нулю, получим уравнение, решением которого является точное значение x_{\max} , доставляющее максимум функции y .

Схема применения математической модели при решении реальных задач имеет вид, показанный на рис. 1.2.

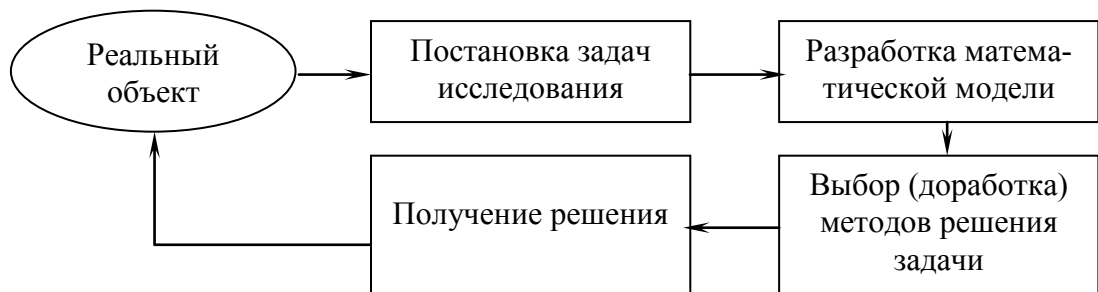


Рис. 1.2

Модель сложного объекта (процесса, системы) не может быть простой. Из чего следует, что процесс использования математических моделей реальных си-

|| стем является итерационным процессом, когда последовательно уточняется (дорабатывается) математическая модель и методы решения стоящих задач.

Важнейшей характеристикой моделей является их точность, адекватность действительности. При этом важно иметь в виду, что все модели представляют собой приближенное описание реальных объектов (процессов) и поэтому принципиально неточны. Интегральная оценка модели может быть получена путем сравнения результатов моделирования и экспериментальных данных для конкретных объектов или режимов.

Для оценки значимости совпадения или несовпадения модельных и экспериментальных результатов широко используются методы математической статистики. Вместе с тем не следует переоценивать результаты такой проверки. Хорошее совпадение модельных и экспериментальных данных, вообще говоря, не доказывает точности модели, а лишь подтверждают ее функциональную пригодность для моделирования. Всегда может быть предложена модель, обеспечивающая лучшее совпадение с экспериментом, но не лучшее описание моделируемого объекта или процесса.

1.2. КЛАССИФИКАЦИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Существует несколько схем классификации математических моделей. Все они достаточно условны. Одна из таких схем приведена на рис. 1.3.



Рис. 1.3

Все математические модели по использованному формальному языку можно разбить на аналитические и имитационные.

Аналитические – модели, в которых используется стандартный математический язык. Имитационные – модели, в которых использован специальный язык моделирования или универсальный язык программирования.

Аналитические модели могут быть записаны в виде формул или уравнений. Если какой-либо процесс не может быть описан в виде аналитической модели, его описывают с помощью специального алгоритма или программы. Такая модель является имитационной.

Аналитические модели в свою очередь разбиваются на теоретические и эмпирические модели. Теоретические модели отражают реальные структуры и процессы в исследуемых объектах, то есть, опираются на теорию их работы. Эмпирические модели строятся на основе изучения реакций объекта на изменение условий окружающей среды. При этом теория работы объекта не рассматривается, сам объект представляет собой так называемый «черный ящик», а модель – некоторую интерполяционную зависимость. Эмпирические модели могут быть построены на основе экспериментальных данных. Эти данные получают непосредственно на исследуемых объектах или с помощью их физических моделей.

По форме описания аналитические модели подразделяются на линейные и нелинейные.

Если все входящие в модель величины не зависят от времени, то имеем статическую модель объекта или процесса, в противном случае получаем динамическую модель.

В детерминированных моделях все взаимосвязи, переменные и константы заданы точно, что приводит к однозначному определению результирующей функции. Если часть или все параметры, входящие в модель по своей природе являются случайными величинами или случайными функциями, то модель относят к классу стохастических моделей. В стохастических моделях задаются законы распределения случайных величин, что приводит к вероятностной оценке результирующей функции.

Если аналитическое исследование может быть доведено до конца, модели называются аналитически разрешимыми. В противном случае говорят о численно разрешимых аналитических моделях.

КОНТРОЛЬНЫЕ ВОПРОСЫ К ЛЕКЦИИ 1

1. Что позволяет осуществить математическое моделирование до создания реальной системы, объекта?
2. Что позволяют увидеть вычислительные эксперименты?
3. Сформулируйте основную задачу математического моделирования.
4. Дайте определение математической модели.
5. Какой подход решения научных задач является альтернативным математическому моделированию?
6. Перечислите основные недостатки экспериментального подхода.
7. Что является важнейшей характеристикой математической модели?
8. На какие два вида делятся математические модели?
9. Перечислите виды аналитических математических моделей.
10. Дайте краткую характеристику видов моделей.

1.3. Геометрическое представление математических моделей

Геометрически математическая модель может быть представлена как некоторая поверхность отклика, соответствующая расположению точек $W = W(x)$ в k -мерном факторном пространстве X .

Наглядно можно представить себе только одномерную и двумерную поверхности отклика, причем в последнем случае удобно пользоваться топографическим способом изображения рельефа поверхности с помощью линий уровня (изолиний), построенных в двумерном факторном пространстве X . (Рис. 1.4).

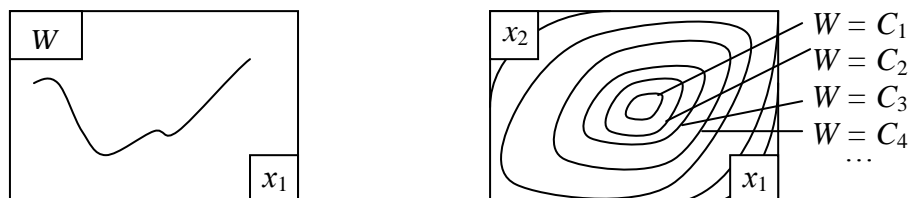


Рис. 1.4

Область, в которой определена поверхность отклика, называется областью определения X^* .

Эта область составляет, как правило, лишь часть полного факторного пространства X ($X^* \subset X$) и выделяется с помощью ограничений, наложенных на управляющие переменные x_i , записанных в виде равенств

$$\begin{aligned} x_i &= C_i, \quad i = 1, \dots, m; \\ f_j(x) &= C_j, \quad j = 1, \dots, l \end{aligned}$$

или неравенств

$$\begin{aligned} x_{i \min} &\leq x_i \leq x_{i \max}, \quad i = 1, \dots, k; \\ f_j(x) &\leq C_j, \quad j = 1, \dots, n, \end{aligned}$$

При этом функции $f_j(x)$ могут зависеть как одновременно от всех переменных, так и от некоторой их части.

Ограничения типа неравенств характеризуют или физические ограничения на процессы в изучаемом объекте (например, ограничения температуры), или технические ограничения, связанные с условиями работы объекта (например, предельная скорость резания).

Возможности исследования моделей существенно зависят от свойств (рельефа) поверхности отклика, в частности, от количества имеющихся на ней «вершин» и ее контрастности.

Количество вершин (впадин) определяет модальность поверхности отклика.

Если в области определения на поверхности отклика имеется одна вершина (впадина), модель называется унимодальной.

Характер изменения функции при этом может быть различным (Рис. 1.5).

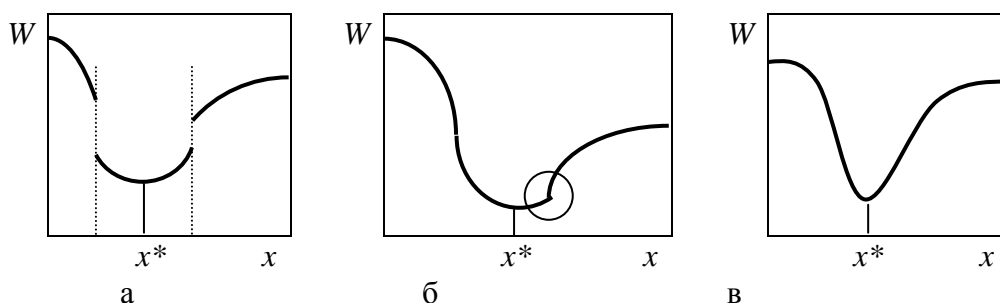


Рис. 1.5

Модель может иметь разрывы первого рода (см. рис. 1.5. а). Непрерывная унимодальная модель может иметь точки разрыва производной – разрывы второго рода (см. рис. 1.5. б). На рис. 1.5 в показана непрерывно-дифференцируемая унимодальная модель.

Для всех трех случаев, представленных на рис. 1.5, выполняется общее требование унимодальности:

Если $W(x^*) = \text{extr } W$, то из условия $x_1 < x_2 < x^*$ ($x_1 > x_2 > x^*$) следует $W(x_1) < W(x_2) < W(x^*)$, если extr – максимум, или $W(x_1) > W(x_2) > W(x^*)$, если extr – минимум, то есть, по мере удаления от экстремальной точки значение функции $W(x)$ непрерывно падает (растет).

Наряду с унимодальными бывают полимодальные модели (Рис. 1.6).

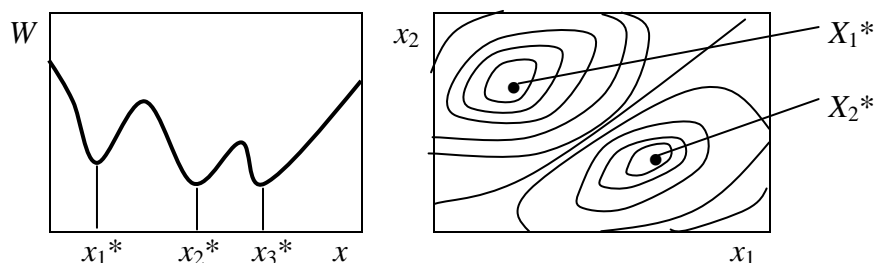


Рис. 1.6

Другим важным свойством поверхности отклика является ее контрастность, показывающая чувствительность результирующей функции к изменению факторов. Контрастность характеризуется величинами производных. Продемонстрируем характеристики контрастности на примере двумерной поверхности отклика (Рис. 1.7). Точка a расположена на «склоне», характеризующем равную контрастность по всем переменным x_i ($i=1,2$); точка b расположена в «овраге», в котором различная контрастность по различным переменным (имеем плохую обусловленность функции); точка c расположена на «плато», на котором низкая контрастность по всем переменным x_i говорит о близости экстремума.

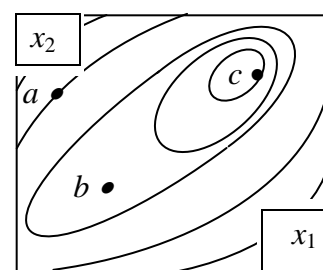


Рис. 1.7

Глава 2. ТЕОРЕТИЧЕСКИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ АНАЛИТИЧЕСКОГО ТИПА

Простейшие аналитические модели могут быть заданы явно в виде функции одной или нескольких переменных.

Обычно в виде функций задаются общие законы природы или общие закономерности, полученные в результате интегрирования дифференциальных уравнений. Примером такой модели может служить знаменитая формула К.Э. Циолковского:

$$\Delta v_{\text{ла}} = v \ln \frac{M_0}{M_k},$$

определяющая приращение скорости ракеты $\Delta v_{\text{ла}}$ при импульсном сжигании топлива через скорость истечения рабочего тела v и отношение начальной M_0 и конечной M_k масс ракеты.

Модель, заданная в явном виде, дает исчерпывающее описание исследуемого объекта. Она позволяет построить зависимость его характеристик от управляющих факторов, взять производные и найти экстремумы модели, определить характеристики модели в окрестности экстремумов и т.д.

Очень удобна графическая интерпретация таких моделей. Однако модели в виде формул могут быть разработаны только для очень простых объектов.

2.1. Построение математической модели сверления лазером

Примером аналитической теоретической модели может служить модель, описывающая глубину отверстия при лазерном сверлении.

Резание и сверление металлов весьма важно для многих областей техники. Значительный интерес представляет создание новых устройств, предназначенных для специальных материалов, а также для тех случаев, когда желательно обеспечить некоторую степень автоматизации указанных процессов. В последнее время для этого были предприняты попытки использования мощных лазеров.

Основная идея состоит в том, чтобы сфокусировать значительную мощность на малой площади поверхности материала, создавая таким образом интенсивный нагрев и испарение с последующим образованием отверстия. При сверлении необходимо постараться обеспечить такие условия процесса, чтобы проделанное отверстие прямо проходило сквозь материал, и избежать, таким образом, затекания расплавленного металла обратно в отверстие и застывания его там.

Построим математическую модель, главная применимость которой – глубокое сверление. При помощи модели попытаемся ответить на вопрос, как быстро можно проделать отверстие, используя пучок излучения высокой мощности, и на какую глубину.

Рассмотрим высокоэнергетический пучок лазерного излучения, сфокусированный на малом участке поверхности металла (Рис. 2.1). Определенная доля энергии поглощается, а остальная часть отражается. Поглощение энергии происходит внутри слоя, толщина которого много меньше миллиметра, вызывает поверхностный нагрев материала и рост температуры поверхности. Температура растет не безгранично. Существует два процесса, ограничивающие рост температуры:

- перенос тепла в глубь материала от нагретых к холодным участкам, обусловленный теплопроводностью;
- испарение. Когда температура материала достигает точки кипения, скрытое тепло поглощается без дальнейшего увеличения температуры в процессе испарения материала.

При удалении пара от поверхности материала в металле образуется выемка.

Задача количественного описания этого процесса и вызывает необходимость математического моделирования.

Будем рассматривать модель, описывающую процесс разрушения материала, при котором вся энергия лазерного излучения используется только для испарения материала.

Этот предельный режим испарения может возникать двумя путями:

- когда энергия поступает на поверхность слишком быстро, так что тепло не успевает распространиться в глубь металла;
- плотность мощности пучка постоянна, а распределение температуры впереди границы области испарения приближается к стационарному.

Предположим, что мощность W распределена по некоторой площади A поверхности; излучение приложено по нормали к поверхности (см. рис. 2.1). За интервал времени Δt поступает энергия $W \cdot \Delta t$. Пусть глубина возникающей выемки равна ΔS , тогда объем испарившегося материала равен $A \cdot \Delta S$. Используя закон сохранения энергии, получим

$$h \cdot \rho \cdot A \cdot \Delta S = W \cdot \Delta t,$$

где h – количество тепла, требуемое для испарения единицы массы материала; ρ – плотность материала.

Преобразуем это выражение и положим $\Delta t \rightarrow 0$, получим скорость роста глубины выемки:

$$\frac{dS}{dt} = \frac{W}{A} \cdot \frac{1}{h\rho}.$$

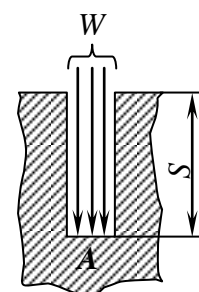


Рис. 2.1

Это уравнение показывает, что для любого материала предельная скорость пропорциональна плотности энерговыведения W/A . Интегрируя это уравнение и полагая $S = 0$ при $t = 0$, найдем глубину выемки в произвольный момент времени t :

$$S(t) = \frac{1}{h\rho A_0} \int_0^t W dt \quad (2.1)$$

или

$$S(t) = \frac{E(t)}{h\rho A},$$

где $E(t)$ – полная энергия, выделенная источником за промежуток времени $(0, t)$.

Таким образом, в предельном режиме испарения глубина выемки зависит только от полной энергии, поступившей на поверхность. Формула (2.1) представляет собой теоретическую аналитически-разрешимую динамическую детерминированную модель.

На практике всегда существует перенос некоторого количества тепла в материал за счет теплопроводности. Общая задача движения границы раздела фаз с учетом теплопроводности известна как *задача Стефана*. Ее решение представляет определенные математические трудности.

Контрольные вопросы к лекции 2

11. В виде чего может быть представлена математическая модель геометрически?
12. Что такое область определения математической модели?
13. Какая модель называется унимодальной?
14. Как задаются математические модели аналитического типа?
15. Приведите пример математической модели аналитического типа.
16. Какие задачи позволяет решить модель, заданная в явном виде?
17. Какой предельный режим рассматривается при построении математической модели сверления лазером?
18. Какой закон используется при построении математической модели сверления лазером?
19. Назовите процессы, препятствующие росту температуры при лазерном сверлении.
20. На какие вопросы можно ответить, используя математическую модель сверления лазером?
21. К какому типу принадлежит модель зависимости глубины выемки от длительности импульса?

2.2. Линейные математические модели

Наиболее простыми являются так называемые линейные детерминированные модели. Они задаются в виде линейной формы управляющих переменных (x):

$$W = a_0 + a_1x_1 + \dots + a_kx_k$$

при линейных ограничениях вида

$$\begin{aligned} b_{1j}x_1 + b_{2j}x_2 + \dots + b_{kj}x_k &\geq b_j, \quad j = 1, \dots, q_1; \\ c_{1j}x_1 + c_{2j}x_2 + \dots + c_{kj}x_k &= c_j, \quad j = 1, \dots, q_2; \\ d_{1j}x_1 + d_{2j}x_2 + \dots + d_{kj}x_k &\leq d_j, \quad j = 1, \dots, q_3. \end{aligned}$$

Общее число ограничений $m = q_1 + q_2 + q_3$ может превосходить число переменных ($m > k$). Кроме того, обычно вводится условие положительности переменных ($x_i \geq 0$).

Поверхность отклика для линейной модели представляет собой гиперплоскость. Например, рассмотрим линейную модель двух переменных следующего вида:

$$W = -2x_1 - 3x_2 \tag{2.2}$$

при следующих ограничениях

$$\begin{aligned} 2x_1 + 3x_2 &\leq 18; \\ x_1 - 4x_2 &\leq 4; \\ -2x_1 + x_2 &\leq 2; \\ x_1 &\geq 0; \quad x_2 &\geq 0. \end{aligned} \tag{2.3}$$

Область допустимых значений (область определения) $OABCD$ для модели (2.2) образована ограничениями (2.3) (Рис. 2.2). Поверхность отклика представляет собой плоский многоугольник $OA'B'C'D'$ (рис. 2.2, б).

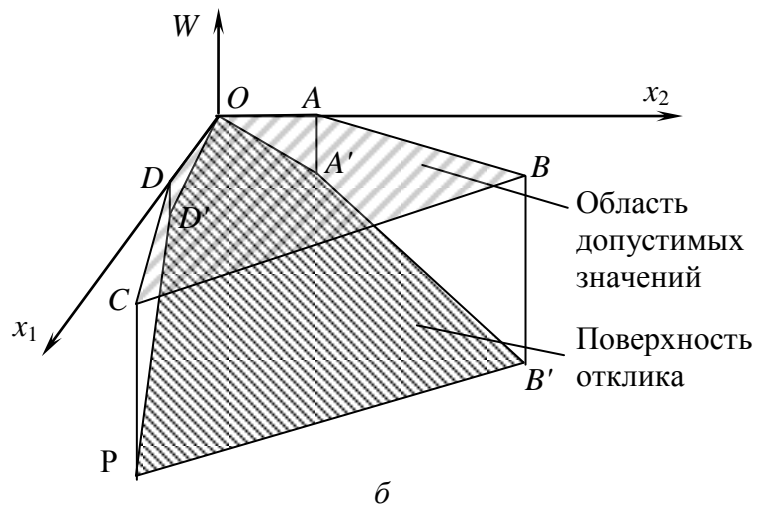
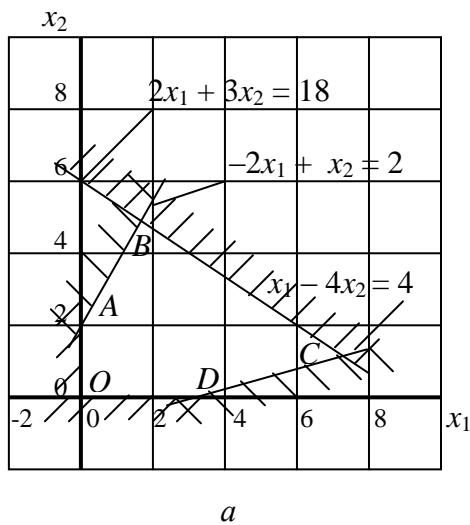


Рис. 2.2

При определенном соотношении ограничений множество допустимых решений может отсутствовать (пусто). Пример такого множества показан на рис. 2.3. Прямые AC и BC ограничивают область допустимых значений сверху. Третье ограничение отсекает область допустимых значений снизу от прямой AB . Таким образом, общей области, удовлетворяющей всем трем ограничениям, нет.

Линейные модели достаточно просты и поэтому, с одной стороны, предполагают существенное упрощение задачи, а с другой – допускают разработку простых и эффективных методов решения.

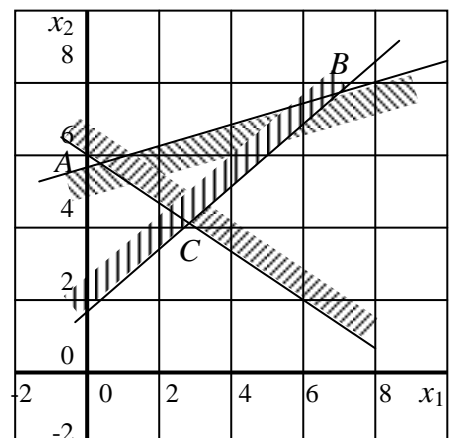


Рис. 2.3

При исследовании ДЛА линейные модели используются редко и почти исключительно при приближенном описании задач.

Линейные модели могут использоваться при поэтапной аппроксимации нелинейных моделей (линеаризация задачи). Особенно эффективен этот прием при изучении небольших областей исследуемого пространства. Представление отдельных участков нелинейной поверхности отклика линейной моделью лежит в основе большой группы методов оптимизации, так называемых методов с линейной тактикой.

Исследование линейных моделей не представляет труда. В частности влияние каждой из переменных на характеристики модели вида

$$W = a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k$$

задается ее коэффициентами:

$$\frac{\partial W}{\partial x_i} = a_i, \quad i = 1, \dots, k.$$

Для нахождения оптимума линейной модели $W_{\text{опт}}$ разработан эффективный симплекс-метод.

К линейным иногда сводятся простейшие модели стоимости, рассматриваемые как совокупность производимых затрат.

Примером такой модели является классическая модель стоимости перевозок (транспортная задача) (Рис. 2.4).

Имеется k пунктов производства ($i = 1, \dots, k$) и m пунктов потребления ($j = 1, \dots, m$) некоторого продукта. Количество продукта, произведенного в каждом из k пунктов производства, равно a_i ; количество продукта, необходимого в каждом из m пунктов потребления, равно b_j .

Предполагается равенство общего производства и потребления:

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

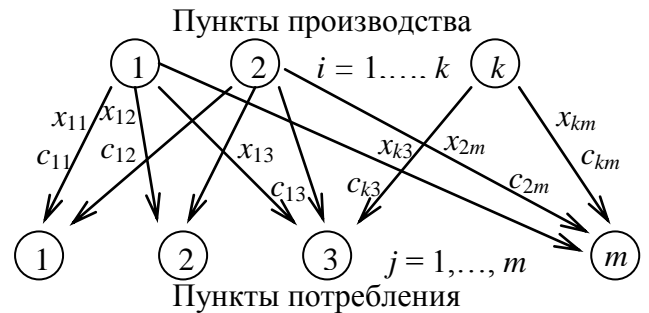


Рис. 2.4

Количество продукта, перевозимого из i -го пункта производства в j -й пункт потребления, равно x_{ij} ; стоимость перевозки единицы этого продукта – c_{ij} .

Суммарная стоимость перевозок C_{Σ} задается линейной моделью:

$$C_{\Sigma} = \sum_{i=1}^k \sum_{j=1}^m c_{ij}x_{ij}$$

при следующих ограничениях

$$\sum_{j=1}^m x_{ij} = a_i; \quad \sum_{i=1}^k x_{ij} = b_j; \quad x_{ij} \geq 0.$$

К линейным также относятся модели в виде линейных дифференциальных уравнений (обыкновенных или в частных производных).

Линейное обыкновенное дифференциальное уравнение n -го порядка имеет вид

$$a_n \frac{d^n x(t)}{dt^n} + a_{n-1} \frac{d^{n-1} x(t)}{dt^{n-1}} + \dots + a_1 \frac{dx(t)}{dt} + a_0 x(t) = f(t). \quad (2.10)$$

Начальные условия записываются как

$$x(0) = C_0, \quad x'(0) = C_1, \quad x''(0) = C_2, \dots, \quad x^{(n-1)}(0) = C_{n-1}.$$

Линейное дифференциальное уравнение в частных производных имеет вид

$$a_0 \frac{\partial \Phi(t)}{\partial t} + a_1 \frac{\partial \Phi(t)}{\partial x_1} + a_2 \frac{\partial \Phi(t)}{\partial x_2} + \dots + a_k \frac{\partial \Phi(t)}{\partial x_k} = f(x_1, x_2, \dots, x_k, t).$$

Модель, заданная в виде дифференциального уравнения в частных производных, включает начальные и граничные условия (условия на границе области определения функции $\Phi(t)$).

2.3. Исследование простейшей математической модели работы газотурбинного двигателя

Газотурбинный двигатель (ГТД) является основной силовой установкой современных самолетов.

Схема ГТД имеет вид, показанный на рис. 2.5.

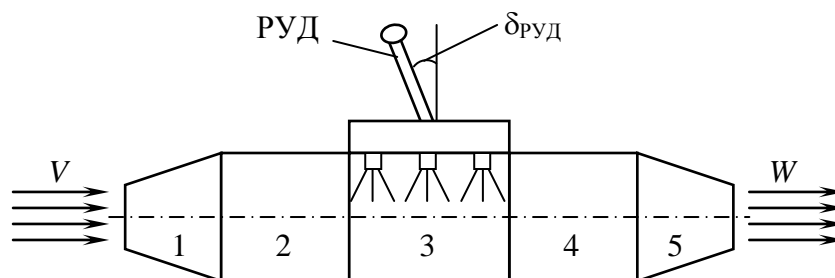


Рис. 2.5

Здесь 1 – входной диффузор; 2 – компрессор; 3 – камера сгорания; 4 – турбина; 5 – выходное сопло.

Цикл работы ГТД включает следующие этапы:

- 1) Набегающий со скоростью V поток воздуха через диффузор поступает в компрессор.
- 2) Компрессор, вращаясь на одном валу с турбиной, сжимает воздух, который поступает в камеру сгорания.
- 3) В камеру сгорания постоянно впрыскивается топливо (керосин), которое смешивается со сжатым воздухом.
- 4) Газ, образующийся от сгорания, поступает на турбину, которая разгоняет его до скорости W .
- 5) С этой скоростью газ через сопло выбрасывается в атмосферу.

За счет того, что $W > V$, образуется сила тяги P , которая позволяет самолету осуществлять полет в атмосфере.

Изменение силы тяги осуществляется путем изменения скорости впрыска топлива в камеру сгорания с помощью перемещения ручки управления двигателем (РУД). Перемещение РУД на определенный угол $\delta_{\text{РУД}}$ осуществляется либо вручную летчиком, либо с помощью исполнительного устройства по сигналам от САУ полетом. Увеличение значения $\delta_{\text{РУД}}$ вызывает возрастание силы P , а уменьшение – убывание этой силы.

ГТД является сложной технической системой, в которой протекает значительное число физических и химических процессов. Двигатель оснащен всевозможными устройствами автоматики, системами поворота и охлаждения турбинных лопаток и т.д. Естественно, математическое описание функционирования ГТД также будет достаточно громоздким, включающим в себя системы дифференциальных уравнений в частных производных, обыкновенных дифференциальных уравнений, трансцендентных функций, алгоритмы цифровой системы управления двигателем. Такие модели используются в процессе проектирования ГТД.

Для решения задач управления полетом используется более простая модель ГТД, представляющая собой зависимость силы тяги P от угла $\delta_{\text{РУД}}$ отклонения РУД. Процесс изменения силы тяги описывается обыкновенным дифференциальным уравнением вида:

$$\tau \frac{dP}{dt} + P = k\delta_{\text{РУД}}, \quad (2.11)$$

где $\tau > 0$ – постоянная времени двигателя, зависящая кроме конструктивных характеристик также от температуры окружающего воздуха, его влажности и других внешних факторов; k [кг/град] – коэффициент пропорциональности.

Начальное условие для уравнения (2.11) записывается как

$$P(0) = P_0. \quad (2.12)$$

Таким образом, уравнение (2.11) совместно с начальным условием (2.12) представляет собой простейшую математическую модель работы ГТД, записанную в виде обыкновенного дифференциального уравнения 1-го порядка.

Для определения коэффициента пропорциональности k используются градуировочные графики зависимости тяги от угла поворота РУД, построенные на основе экспериментальных данных. Тангенс угла наклона графика равен искомому коэффициенту.

Интегрирование уравнения (2.11) с начальным условием (2.12) позволяет выяснить, как изменяется сила тяги во времени (рис. 2.6).

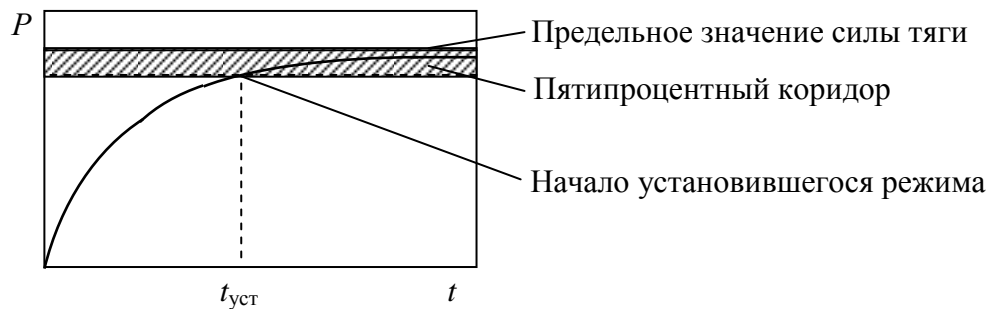


Рис. 2.6

При отклонении РУД тяга P нарастает и затем стабилизируется на определенном предельном значении, т.е. ГТД является инерционным объектом.

Предельное значение силы тяги получаем из (2.11), когда скорость ее изменения равна нулю:

$$\frac{dP}{dt} = 0 \Rightarrow P_{пред} = k\delta_{РУД}. \quad (2.13)$$

Длительность нарастания зависит от значения постоянной времени двигателя τ . Процесс считается установившимся при $t = t_{уст}$, когда тяга входит в так называемый пятипроцентный коридор от предельного значения силы тяги (рис. 2.6). Чем больше τ , тем инерционнее двигатель и, следовательно, больше $t_{уст}$.

На рис. 2.7 показано поведение силы тяги в зависимости от угла отклонения РУД при $\tau = 0,5$.

Сила тяги при взлете, когда РУД отклонена на 10° , выходит на установившийся режим на третьей секунде и достигает величины 3390 кг. Через десять секунд после взлета, когда РУД отклонена на 20° , сила тяги устанавливается на величине 6780 кг, и еще через десять секунд, когда РУД отклонена на 30° , сила тяги устанавливается на величине 10170 кг. Предельное значение силы тяги равно 14270 кг.

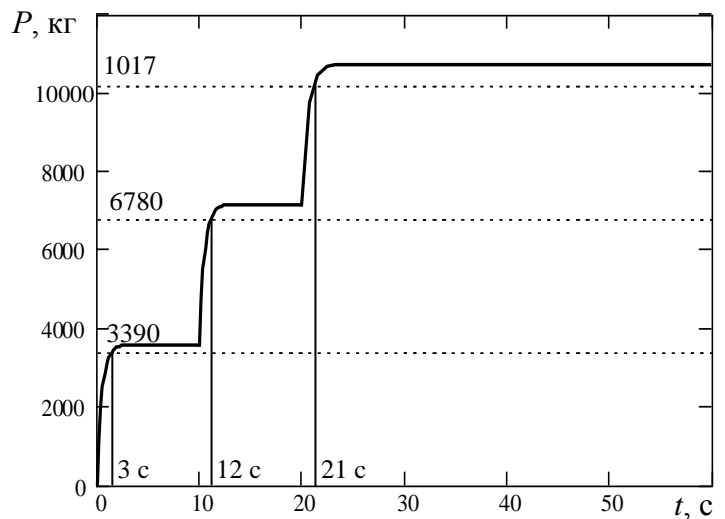


Рис. 2.7

2.4. Нелинейные детерминированные модели

Нелинейные детерминированные модели обладают бóльшей точностью и гибкостью. Они могут быть заданы в виде нелинейной функции одной или нескольких переменных или в виде дифференциальных уравнений (обыкновенных или в частных производных). Наиболее распространенными среди нелинейных моделей при описании ДУ и ДЛА являются:

- полиномиальные функции;
- позиномные функции;
- тригонометрические функции;
- экспоненциальные функции;
- обыкновенные дифференциальные уравнения;
- дифференциальные уравнения в частных производных др.

Нелинейные модели могут быть записаны в виде функционала, зависящего от управляющих переменных x и некоторых функций $f(x)$ всех или части этих переменных: $W = W(x, f(x))$. При этом функции $f(x)$ могут представлять собой функционалы, зависящие от промежуточных функций $f^*(x)$ и т.д. На класс функций $f(x)$, $f^*(x)$ не накладывается никаких ограничений, однако предполагается возможность однозначного перехода от вектора управляющих параметров x к общей характеристике модели W .

Область определения модели может быть ограничена с помощью равенств или неравенств:

$$\begin{aligned} x_i &= c_i, & i &= 1, \dots, m; \\ f(x) &= c_j, & j &= 1, \dots, l; \\ x_{i \min} &\leq x_i \leq x_{i \max}, & i &= 1, \dots, k; \\ f_j(x) &\leq c_j, & j &= 1, \dots, n. \end{aligned}$$

По существу под определение нелинейной модели подпадает любое математическое описание ДУ и ДЛА, не укладывающееся в рамки более простых моделей.

2.4.1. Полиномиальные модели

Полиномиальные модели основаны на идее приближенного представления модели конечным числом членов ряда Тейлора:

$$W(x) = W(x_0) + \sum_{i=1}^k \frac{\partial W(x_0)}{\partial x_i} (x_i - x_{i0}) + \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \frac{\partial^2 W(x_0)}{\partial x_i \partial x_j} (x_i - x_{i0})(x_j - x_{j0}) + \dots$$

Наиболее простой из моделей этого класса является квадратичная модель:

$$W(x) = a_0 + \sum_{i=1}^k a_i x_i + \sum_{\substack{i=1 \\ j \geq i}}^k a_{ij} x_i x_j$$

при ограничениях

$$\sum_{i=1}^k b_{ij} x_i \geq b_j, \quad j = 1, \dots, q_1; \quad \sum_{i=1}^k c_{ij} x_i = c_j, \quad j = 1, \dots, q_2; \quad \sum_{i=1}^k d_{ij} x_i \leq d_j, \quad j = 1, \dots, q_3.$$

Квадратичные модели широко используются для представления экспериментальных данных при идентификации ДЛА и их элементов.

Квадратичные модели используются для аппроксимации отдельных участков поверхности отклика, когда линейное приближение оказывается недостаточным, например, в окрестности экстремума, и лежит в основе нелинейных методов оптимизации. Если квадратичная модель также оказывается недостаточно точной, то используются полиномиальные модели более высоких порядков.

Исследование полиномиальных моделей частично можно осуществить аналитическими методами. Например, аналитически можно определить степень влияния отдельных переменных на характеристики модели.

2.4.2. Позиномные модели

Позиномные модели основаны на представлении модели в виде суммы произведений степенных функций:

$$W(x) = \sum_{j=1}^m c_j x_1^{\alpha_{1j}} x_2^{\alpha_{2j}} \dots x_k^{\alpha_{kj}} = \sum_{j=1}^m c_j \prod_{i=1}^k x_i^{\alpha_{ij}}, \quad (2.14)$$

где x_i – управляющие переменные, α_{ij} – произвольные положительные числа, $c_j \geq 0$ – обеспечивает выпуклость модели.

Величины α_{ij} , c_j рассчитываются на основе статистических данных, отражающих опыт производства соответствующих узлов и систем.

Позиномные модели можно использовать для описания стоимости сложных систем.

К позиномным моделям сводится задача выбора геометрических характеристик ряда технических устройств, в том числе элементов ДЛА, например, электромагнитов, силовых ферм и т.д.

Исследование позиномных моделей сложнее, чем моделей полиномиального типа, и осуществляется в основном численными методами. Однако, при $m = 1$ и $x_1 > 0$, $x_2 > 0, \dots$, $x_k > 0$ в формуле (2.4) существует способ приведения позинома к линейному виду.

В этом частном случае модель (2.4) будет выглядеть в следующем виде:

$$W(x_1, x_2, \dots, x_k) = c \cdot x_1^{\alpha_1} x_2^{\alpha_2} \dots x_k^{\alpha_k}.$$

Прологарифмируем обе части этого равенства, получим

$$\ln W = \ln c + \alpha_1 \ln x_1 + \alpha_2 \ln x_2 + \dots + \alpha_k \ln x_k. \quad (2.15)$$

Введем обозначения логарифмов переменных W , x_1 , x_2, \dots, x_k и константы c :

$$Y = \ln W; \quad C = \ln c; \quad X_i = \ln x_i; \quad i = 1, \dots, k.$$

Выражение (2.5) примет линейный вид

$$Y(X_1, X_2, \dots, X_k) = C + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k.$$

Для поиска оптимальных решений на основе позиномных моделей разработан специальный аппарат – так называемое геометрическое программирование.

Контрольные вопросы к лекции 3

1. С какими значениями величин оперируют детерминированные модели?
2. Как выглядит линейная детерминированная модель в общем виде?
3. Что представляет собой поверхность отклика для линейной модели?
4. Приведите модель стоимости перевозок.
5. Где используются линейные детерминированные модели?
6. Приведите простейшую математическую модель изменения силы тяги ГТД.
7. К какому типу она относится?
8. Где она может быть использована?
9. Приведите модель установившегося процесса горизонтального полета самолета.
10. Что и как можно определить с ее помощью?
11. Какие виды нелинейных математических моделей Вы знаете?
12. Приведите общий вид квадратичного полинома.
13. Приведите формулу позинома.
14. Как привести позином к линейному виду (при каком условии)?

2.4.3. Математическая модель кратчайшего пути

В качестве примера применения нелинейных статических моделей рассмотрим задачу описания двумерного движения точки по ограниченной области (рис. 2.8). Такая задача может возникнуть при определении координат опорных точек движения инструмента на станке с ЧПУ.

Найдем кратчайший путь от точки A с координатами (x_A, y_A) до точки B с координатами (x_B, y_B) на плоскости, из которой исключена область D , определенная неравенством $x^2 + y^2 \leq R^2$.

Кратчайшим расстоянием между двумя точками на плоскости является соединяющий их отрезок прямой.

Пусть расстояние между точками A и B равно p и центр окружности, ограничивающей область D , лежит посередине между точками A и B . Тогда

$$x_A = -\frac{p}{2}; y_A = 0; x_B = \frac{p}{2}; y_B = 0.$$

Рассмотрим путь ACB , где точка C имеет координаты $(0, y_C)$, а y_C – достаточно велико, чтобы отрезки AC и CB не пересекались с областью D . Тогда по теореме Пифагора

$$ACB = 2\sqrt{\left(\frac{p}{2}\right)^2 + y_C^2}.$$

Отсюда видно, что при убывании y_C путь сокращается. Будем уменьшать y_C до тех пор, пока AC не коснется окружности в точке E_1 ($C \rightarrow C_1$). Этот путь является наилучшим среди путей, составленных из двух отрезков прямых линий.

Обозначим через α угол $\angle E_1OC_1$, тогда $E_1C_1 + C_1F_1 = 2R \operatorname{tg} \alpha$;

Длина дуги E_1F_1 определяется по формуле $\operatorname{arc} E_1F_1 = 2R\alpha$.

Но $\operatorname{tg} \alpha > \alpha$ для всех $\alpha \in \left(0; \frac{\pi}{2}\right)$.

Следовательно, путь, состоящий из отрезка AE_1 , дуги E_1F_1 и отрезка F_1B , является более коротким, чем AC_1B .

На этой стадии решения задачи мы выяснили, что кратчайший путь состоит из двух отрезков прямых линий и дуги окружности.

Для окончательного решения задачи рассмотрим путь AE_2 , дуга E_2F_2 , F_2B , где $\angle AOE_2 = \angle BOF_2 = \beta$.

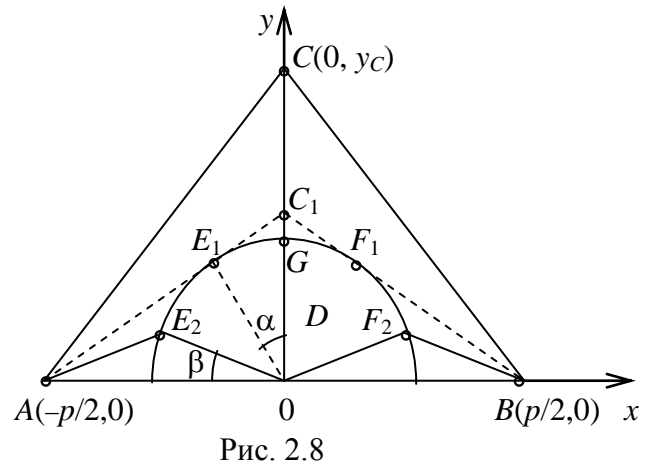
Длину этого пути обозначим через S . Получим математическую модель пути:

$$S = 2(AE_2 + \operatorname{arc} E_2G) = 2 \left[\sqrt{\left(\frac{p}{2} - R \cos \beta\right)^2 + R^2 \sin^2 \beta} + R \left(\frac{\pi}{2} - \beta\right) \right]. \quad (2.16)$$

$$0 \leq \beta \leq \beta_{\text{ор}} = \frac{\pi}{2} - \alpha, \quad (2.16')$$

где α – угол между прямой E_1O и осью OY .

Ограничение (2.16') вводится потому, что при $\beta > \frac{\pi}{2} - \alpha$ прямая AE_1 пересечет область D , а этого не должно быть.



Задача заключается в определении угла β_0 , при котором путь S будет минимальным. Необходимым условием минимума функции $S(\beta)$ является равенство нулю производной:

$$\frac{dS}{d\beta} = 0. \quad (2.17)$$

Рассмотрим частный случай:

$$P = 4; R = 1.$$

$$\text{Тогда} \quad \alpha = \frac{\pi}{2} - \arccos\left(\frac{R}{p/2}\right) = \arcsin\left(\frac{R}{p/2}\right) = \arcsin\left(\frac{1}{2}\right) = \frac{\pi}{6}.$$

Подставив значения p и R в математическую модель (2.6), получим

$$S = 2 \left[\sqrt{(2 - \cos \beta)^2 + \sin^2 \beta} + \frac{\pi}{2} - \beta \right].$$

Произведя некоторые преобразования, получим

$$S = 2 \left(\sqrt{5 - 4 \cos \beta} + \frac{\pi}{2} - \beta \right).$$

Возьмем производную по β от этого выражения и приравняем ее к нулю.

$$\frac{4 \sin \beta}{2\sqrt{5 - 4 \cos \beta}} - 1 = 0.$$

Получили уравнение, решив которое относительно β , найдем значение угла β_0 , при котором S минимально. Опустив промежуточные преобразования, получим $\cos \beta = 1/2$. То есть $\beta = \pi/3$.

Чтобы убедиться, что найденное значение является точкой минимума, необходимо исследовать вторую производную от (2.16). Если она больше нуля при $\beta = \beta_0$, то $S(\beta)$ действительно минимальна в этой точке.

Вторая производная от $S(\beta)$ имеет вид

$$\frac{d^2 S}{d\beta^2} = -4 \frac{2 \cos^2(\beta) - 5 \cos(\beta) + 2}{[5 - 4 \cos(\beta)]^{3/2}}.$$

Подставив в нее найденное значение $\beta_0 = \pi/3$, получим

$$\left. \frac{d^2 S}{d\beta^2} \right|_{\beta_0 = \pi/3} = 0.$$

Равенство нулю второй производной требует дополнительного исследования критической точки. Необходимо найти первую, не обращающуюся в нуль, производную. Если она нечетного порядка, функция не имеет в исследуемой точке ни максимума, ни минимума. Если она четного порядка и больше нуля, исследуемая точка является минимумом. Проверим третью производную от $S(\beta)$ по β :

$$\left. \frac{d^3 S}{d\beta^3} \right|_{\beta = \frac{\pi}{3}} = - \frac{[4(4 \cos^2(\beta) - 10 \cos(\beta) + 13)] \sin(\beta)}{\sqrt{(5 - 4 \cos(\beta))^5}} \Bigg|_{\beta = \frac{\pi}{3}} = -2.$$

Отсюда имеем, что при $\beta = \pi/3$ функция $S(\beta)$ не имеет ни максимума, ни минимума. Действительно, из графика функции $S(\beta)$ (рис. 2.9) видно, что на отрезке (2.16') функция (2.16) монотонно убывает. В точке $\beta_0 = \pi/3$, совпадающей с $\beta_{\text{огр}}$, кривая имеет точку перегиба. Наименьшее в области определения значение находится на границе этой области. Следовательно, путь AE_1GF_1B действительно кратчайший и его длина равна $S(\pi/3) = 4,511$.

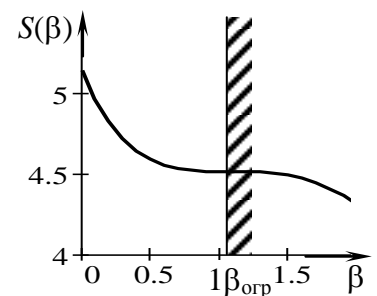


Рис. 2.9

Покажем, что математическая модель (2.6) для любых p и R монотонно убывает на отрезке $0 \leq \beta \leq \beta_{\text{оп}} = \pi/2 - \alpha$ и, следовательно, имеет наименьшее значение при $\beta = \pi/2 - \alpha$. Для этого необходимо показать, что вторая производная от $S(\beta)$ на интересующем нас отрезке не превышает нуля.

Вторая производная от функции (2.6) имеет вид

$$\frac{d^2 S}{d\beta^2} = 2R \frac{p \sin(\beta) - \sqrt{p^2 - 4pR \cos(\beta) + 4R^2}}{\sqrt{p^2 - 4pR \cos(\beta) + 4R^2}}.$$

Покажем, что она не превышает нуля:

$$2R \frac{p \sin(\beta) - \sqrt{p^2 - 4pR \cos(\beta) + 4R^2}}{\sqrt{p^2 - 4pR \cos(\beta) + 4R^2}} \leq 0.$$

Разделив обе части неравенства на $2R$ и умножив на корень квадратный (это можно сделать, не нарушив неравенства, так как $R > 0$, а корень квадратный представляет собой длину отрезка, т. е. тоже больше нуля), получим

$$p \sin(\beta) \leq \sqrt{p^2 - 4pR \cos(\beta) + 4R^2}.$$

Возведя обе части в квадрат (на рассматриваемом отрезке $\sin(\beta) > 0$) и произведя некоторые преобразования, получим

$$p^2 \cos^2(\beta) \geq 4pR \cos(\beta) - 4R^2.$$

В левой части неравенства $\cos^2(\beta)$ можно заменить его минимальным значением, т.е. нулем, а в правой части – максимальным значением, т.е. единицей. Тогда получим

$$0 \geq 4pR - 4R^2 \quad \text{или} \quad p \geq R.$$

Но p действительно больше R (см. рис. 2.5).

Таким образом, аналитическую модель пути (формула (2.6)) мы использовали для доказательства того, что при $\beta = \pi/2 - \alpha$ путь является кратчайшим. Зная это, можно определить координаты опорных точек движения инструмента на станке с ЧПУ при любых значениях величин p и R :

$$A(-p/2, 0); E_1(-R \sin(\alpha), R \cos(\alpha)); F_1(R \sin(\alpha), R \cos(\alpha)); B(p/2, 0),$$

где $\alpha = \arcsin\left(\frac{R}{p/2}\right)$.

Контрольные вопросы к лекции 4

1. К какому типу можно отнести модель кратчайшего расстояния между двумя точками?
2. Является ли найденное значение угла β точкой минимума пути?
3. Является ли путь S при найденном значении угла β кратчайшим?

2.5. Математическая модель в виде обыкновенных дифференциальных уравнений

Математическая модель в виде одного или нескольких обыкновенных дифференциальных уравнений (ОДУ) широко используются при изучении переходных процессов в системах автоматического регулирования (САР), при описании баллистики летательных аппаратов, а также при описании процессов движения (поток, частицы, механические элементы).

В простейшем случае модель может иметь вид линейного дифференциального уравнения n -го порядка:

$$a_n \frac{d^n x(t)}{dt^n} + a_{n-1} \frac{d^{n-1} x(t)}{dt^{n-1}} + \dots + a_1 \frac{dx(t)}{dt} + a_0 x(t) = f(t)$$

или системы дифференциальных уравнений 1-го порядка

$$\frac{dx_1}{dt} = f_1(t, x_1, \dots, x_n);$$

$$\frac{dx_2}{dt} = f_2(t, x_1, \dots, x_n);$$

.....

$$\frac{dx_n}{dt} = f_n(t, x_1, \dots, x_n).$$

Часто встречаются смешанные задачи, а также нелинейные ОДУ.

Модель, заданная в виде дифференциальных уравнений, должна включать в себя необходимый набор начальных условий:

$$x(0) = C_0, \quad x'(0) = C_1, \quad x''(0) = C_2, \dots, \quad x^{(n-1)}(0) = C_{n-1}$$

или $x_1(0) = C_1, \quad x_2(0) = C_2, \dots, \quad x_n(0) = C_n.$

Исследование моделей, заданных в виде обыкновенных дифференциальных уравнений, осуществляется аналитическими и численными методами. Наиболее полными являются аналитические решения, обеспечивающие всесторонний анализ полученных результатов. Но такие решения получены лишь для ограниченного числа дифференциальных уравнений. Численные методы решения позволяют найти лишь конкретные значения изучаемой функции при заданной комбинации исходных данных. Для анализа модели можно использовать некоторую совокупность решений. Однако, очевидно, что результаты анализа в этом случае могут зависеть от выбора этой совокупности.

В качестве простейшего примера математической модели механической системы может быть рассмотрена модель движения груза массой m , закрепленного на вертикальной стенке с помощью пружины жесткостью C и совершающего колебательное движение вдоль оси x в среде с вязкостью ν (Рис. 2.10).

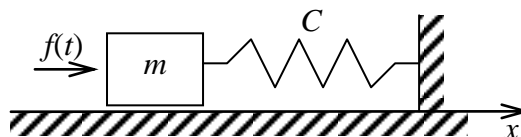


Рис. 2.10

Возмущающая сила, вызывающая колебания, зависит от времени $f(t)$. Наряду с возмущающей силой $f(t)$ на груз действует сила инерции $m \frac{d^2 x(t)}{dt^2}$, сила вязкого трения $\nu \frac{dx(t)}{dt}$, усилие пружины $\frac{1}{C} x(t)$. Все эти силы тормозят движение груза.

Согласно принципу Даламбера сумма всех сил, действующих на груз должна равняться нулю:

$$m \frac{d^2 x(t)}{dt^2} + v \frac{dx(t)}{dt} + \frac{1}{C} x(t) - f(t) = 0. \quad (2.18)$$

Начальные условия характеризуют начальное положение и начальную скорость груза:

$$x(0) = x_0; \quad x'(0) = 0. \quad (2.19)$$

Уравнение (2.18) совместно с начальными условиями (2.19) представляет собой математическую модель рассматриваемой механической системы.

2.6. Модели, заданные в виде уравнений в частных производных

Ряд задач, связанных с использованием физических полей, приводит к моделям в виде дифференциальных уравнений в частных производных.

Особенностью таких задач является то, что изучаемые параметры изменяются не только во времени, но и зависят от координат x, y, z рассматриваемого пространства. Такие модели называются нестационарными. Модели, в которых параметры не зависят от времени, называются стационарными.

К таким моделям сводятся описания полей температур в элементах конструкции двигателя и полей скоростей при течении жидкости (газа). Уравнениями в частных производных описываются колебания элементов конструкции и поля напряжений, возникающих при работе этих элементов.

Линейное дифференциальное уравнение в частных производных имеет вид

$$a_0 \frac{\partial \Phi(t)}{\partial t} + a_1 \frac{\partial \Phi(t)}{\partial x_1} + a_2 \frac{\partial \Phi(t)}{\partial x_2} + \dots + a_k \frac{\partial \Phi(t)}{\partial x_k} = f(x_1, x_2, \dots, x_k, t).$$

Математическая модель, описанная дифференциальными уравнениями в частных производных, должна включать в себя необходимые для решения задачи краевые условия:

1. Должна быть задана область D , ограниченная поверхностью (на плоскости – кривой) Γ , в которой определяется решение.
2. Должны быть заданы условия на границе Γ этой области.

В случае нестационарного поля эти граничные условия, так же как и сама область могут меняться во времени.

Граничные условия могут быть 1-го, 2-го и 3-го рода:

а) Граничные условия 1-го рода предусматривают задание на границе величины искомой функции:

$$\Phi|_{\Gamma} = f_1(\Gamma) \quad \text{– для стационарного поля;}$$

$$\Phi(t)|_{\Gamma} = f_1(\Gamma, t) \quad \text{– для нестационарного поля.}$$

б) Граничные условия 2-го рода – предусматривают задание производной искомой функции:

$$\left. \frac{\partial \Phi}{\partial x_i} \right|_{\Gamma} = f_2(\Gamma) \quad \text{– для стационарного поля;}$$

$$\left. \frac{\partial \Phi}{\partial x_i} (t) \right|_{\Gamma} = f_2(\Gamma, t) \quad \text{– для нестационарного поля.}$$

в) Граничные условия 3-го рода – предусматривают комбинации функции и ее производной:

$$\left(a\Phi + b \frac{\partial \Phi}{\partial x_i} \right) \Big|_{\Gamma} = f_3(\Gamma) \quad \text{– для стационарного поля;}$$

$$\left(a(t)\Phi(t) + b(t) \frac{\partial \Phi}{\partial x_i}(t) \right) \Big|_{\Gamma} = f_3(\Gamma, t) \text{ — для нестационарного поля.}$$

3. Для нестационарных полей должны быть заданы одно или два начальных условия, характеризующих состояние поля в начальный момент времени:

$$\begin{aligned} \Phi(x_i) \Big|_{t=0} &= f_4(x_i); \\ \frac{\partial \Phi}{\partial t}(x_i) \Big|_{t=0} &= f_5(x_i), \quad (i = 1, 2, 3). \end{aligned}$$

Здесь x_i — координаты пространства.

Совокупность уравнений и краевых (и начальных) условий полностью определяет модель и позволяет провести ее исследование.

Решение часто задается в виде семейств изолиний $\Phi = \text{const}$ (Рис. 2.11).

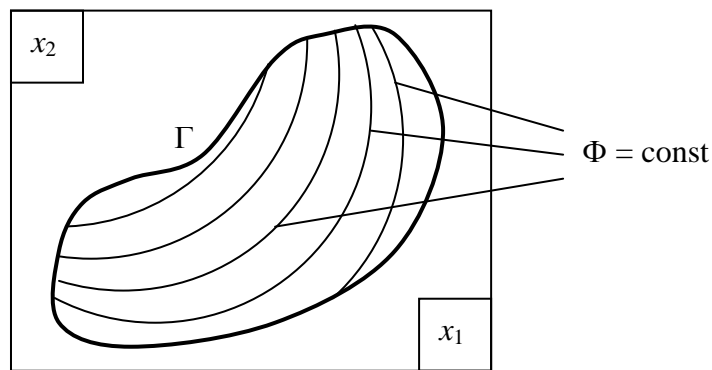


Рис. 2.11

В качестве примера рассмотрим хорошо изолированный металлический прут, нагреваемый с одной стороны. С другой стороны помещен измеритель температуры (Рис. 2.12). Величина подогрева $x(t)$ в момент времени t является входным сигналом, а измеряемая на другом конце температура $y(t)$ — выходным сигналом.

Обозначим через ξ расстояние от измерителя до точки прутка. Температура в этой точке z будет описываться функцией вида

$$z = z(t, \xi).$$

Уравнение теплопроводности для одномерного случая для определения функции z будет иметь вид:

$$\frac{\partial z(t, \xi)}{\partial t} = K \frac{\partial^2 z(t, \xi)}{\partial \xi^2},$$

где K — коэффициент теплопроводности.

Начальным условием в данном случае является начальное распределение температуры (при $t = 0$) по прутку: $z(0, \xi) = \varphi(\xi)$.

Граничные условия определяются тремя условиями:

а) Нагрев прутка на правом конце

$$\left. \frac{\partial z(t, \xi)}{\partial \xi} \right|_{\xi=L} = Kx(t).$$

б) На левом конце подвод тепла отсутствует

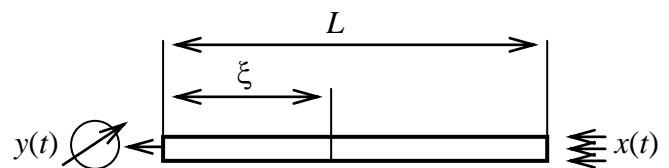


Рис. 2.12

$$\left. \frac{\partial z(t, \xi)}{\partial \xi} \right|_{\xi=0} = 0.$$

в) Показания на измерителе температур ($\xi = 0$) в момент времени t определяется следующим выражением

$$z(t, 0) = y(t).$$

Таким образом, для вычисления температуры на расстоянии L от измерителя по формуле для $y(t)$ необходимо проинтегрировать дифференциальное уравнение с учетом начальных и граничных условий, т.е. получить функцию $z(t, \xi)$. Затем следует проградуировать измеритель температуры, т.е. определить соответствие между $x(t)$ и $y(t)$, задавая различные значения $x(t)$ и вычисляя $z(t, L)$.

Контрольные вопросы к лекции 5

1. Где используются математические модели в виде обыкновенных дифференциальных уравнений?
2. Что должна включать в себя математическая модель в виде обыкновенных дифференциальных уравнений?
3. Какими методами осуществляется исследование моделей, заданных в виде обыкновенных дифференциальных уравнений?
4. Запишите математическую модель движения груза массой m , закрепленного на вертикальной стенке с помощью пружины жесткостью C и совершающего колебательное движение вдоль оси x в среде с вязкостью ν .
5. Какой принцип используется при построении этой модели?
6. К какому типу относится эта модель?
7. Где используются математические модели в виде дифференциальных уравнений в частных производных?
8. Что является особенностью математических моделей в виде дифференциальных уравнений в частных производных?
9. Что должна включать в себя математическая модель в виде дифференциальных уравнений в частных производных?
10. Какого типа бывают граничные условия?
11. Приведите математическую модель распределения температурного поля в металлическом прутке, нагреваемом с одной стороны.

2.7. Стохастические модели

Точные величины и зависимости, используемые в детерминированных моделях, представляют собой лишь некоторые средние значения (математические ожидания) реальных случайных величин (зависимостей). Так, физические константы, характеризующие материалы и рабочие тела (предел прочности материала σ , теплопроводность λ , плотность ρ и т.д.) меняются в зависимости от партии материала и условий окружающей среды. Всегда имеется определенный разброс размеров деталей l , расходов топлива в системах подачи. Все это приводит к тому, что и результирующие функции, характеризующие процесс, также носят случайный характер. Результаты, полученные с помощью детерминированной модели, представляют собой математические ожидания этих характеристик. При этом конкретные данные для конкретной системы могут существенно отличаться от этих математических ожиданий. Например, ресурс конкретного двигателя может существенно отличаться от среднего ресурса двигателей данного типа. Для учета таких отличий вводятся всевозможные «запасы прочности», призванные гарантировать работоспособность реальных объектов при неблагоприятном стечении обстоятельств.

Значительно более полные и объективные результаты можно получить при переходе от детерминированных к стохастическим моделям, то есть при переходе от точно заданных величин к соответствующим случайным величинам.

При этом константы (σ , λ , ρ , l, \dots) заменяются случайными величинами ξ_σ , ξ_λ , ξ_ρ , ξ_l, \dots , подчиненными определенным законам распределения.

Однократное исследование стохастической модели приведет к некоторой случайной величине функции отклика ξ_w , представляющей собой, вообще говоря, ограниченную ценность. Для получения значимых результатов необходимо провести многократное исследование модели и получить распределение результирующей характеристики в интересующем исследователя диапазоне. Поверхность отклика в этом случае представляет собой некий размытый слой переменной плотности.

Такой метод исследования стохастической модели получил название метода статистических испытаний или метода Монте-Карло.

Трудоёмкость исследования стохастических моделей существенно выше, чем моделей детерминированных:

1. Значительно возрастает объем исходной информации: замена констант случайными величинами, введение законов распределения этих величин усложняют модель.
2. Для получения распределения результирующей функции необходимо многократное исследование модели.

С другой стороны, полученное при статистическом моделировании распределение характеристик системы дает в руки исследователя чрезвычайно ценную информацию: Такое распределение позволяет оценить не только среднее значение изучаемой величины, но и разброс этих значений, вероятности появления тех или иных значений при конкретном испытании (например, вероятность выхода из строя ДЛА через тот или иной промежуток времени) и их зависимость от различных факторов.

Очень часто используют нормальный или гауссовский закон распределения, для которого плотность вероятности $f(x)$ и функция распределения $P(x)$ задаются следующими соотношениями:

Вероятность того, что случайная величина попадет в интервал $(x, x+dx)$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}};$$

Вероятность того, что случайная величина попадет в интервал $(-\infty, x)$:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

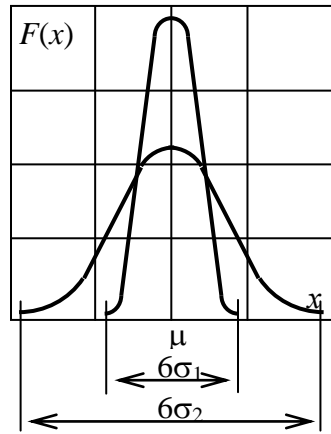


Рис. 2.13

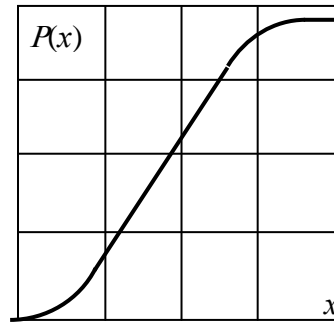


Рис. 2.14

Для случайной величины ξ , распределенной по нормальному закону, $\mu = M(\xi)$, $\sigma = \sigma(\xi)$ (Рис. 2.13, 2.14). Случайная величина распределена в интервале $\mu \pm 3\sigma$. По нормальному закону распределены обычно характеристики материалов, размеры деталей, ресурсы элементов ДЛА.

Наряду с нормальным используются и другие законы распределения случайных величин. Например, равномерное распределение – задает равновероятностные на отрезке $[a, b]$ случайные величины. (Рис. 2.15, 2.16). Плотность вероятности и функция распределения при равномерном распределении определяются по формулам:

$$f(x) = \begin{cases} 0, & -\infty < x < a; \\ \frac{1}{b-a}, & a \leq x \leq b; \\ 0, & b < x < \infty. \end{cases} \quad P(x) = \begin{cases} 0, & -\infty < x < a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b < x < \infty. \end{cases}$$

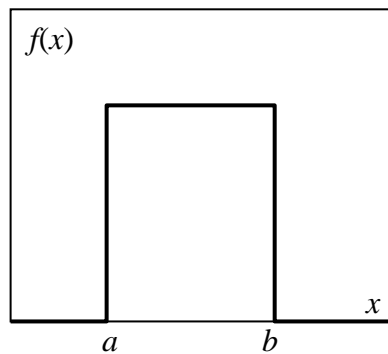


Рис. 2.15

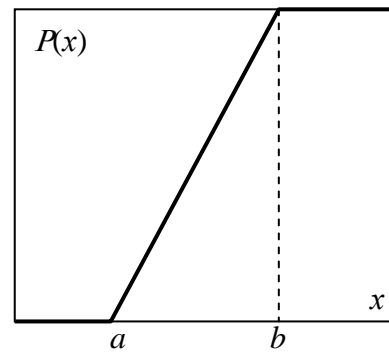


Рис. 2.16

Выбор закона распределения для конкретной случайной величины, входящей в стохастическую модель, может быть обоснован экспериментально или теоретически.

Конкретные параметры распределения (μ , σ , ...) всегда определяются на основе экспериментальных данных. Оценка параметров нормального распределения на основе выборки $\{x_i\}$ из n случайных значений величины x дается соотношениями:

$$\mu = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}}.$$

При использовании метода статистических испытаний характеристики изучаемой системы оцениваются на основе некоторой ограниченной выборки реализаций. Поэтому важно определить достоверность этой оценки.

Вероятность p пребывания системы в некотором состоянии (например, вероятность того, что время работы элемента ДЛА до первого отказа составит не менее t часов), определяется частотой этого события при моделировании:

$$p \approx \frac{n_+}{n},$$

где n_+ – число реализаций, при которых наблюдалось изучаемое состояние системы (время работы ДЛА до первого отказа превысило t); n – общее число реализаций.

Эта оценка является приближенной, так как определяется на основе ограниченной выборки. Отношение $\frac{n_+}{n}$ называется выборочной статистикой.

Ошибка моделирования определяется отклонением выборочной статистики от вероятности

$$\delta = \left| \frac{n_+}{n} - p \right|.$$

Можно показать, что эта ошибка удовлетворяет неравенству

$$\delta \leq \sqrt{\frac{p(1-p)}{\alpha n}}, \quad (2.20)$$

Здесь p – вероятность рассматриваемого состояния; α – вероятность невыполнения оценки (2.20) (уровень риска). Доверительная вероятность выполнения этой оценки равна $1 - \alpha$.

Из (2.20) следует, что погрешность стохастического моделирования обратно пропорциональна \sqrt{n} . То есть увеличение точности при стохастическом моделировании требует значительного увеличения числа реализаций. Для уменьшения погрешности в 10 раз необходимо увеличить число реализаций (а значит и время счета) в 100 раз. Поэтому метод статистических испытаний не может дать решения с очень высокой степенью точности. Считается, что допустимая ошибка может составлять 1-5% максимальной величины, полученной при моделировании.

Величина ошибки зависит также от вероятности p оцениваемого состояния и допустимого уровня риска α . Обычно α задают на одном из фиксированных уровней ($\alpha = 0,005; 0,01; 0,025; 0,05; 0,1 \dots$).

Контрольные вопросы к лекции 6

1. Что представляют собой величины, входящие в стохастическую модель?
2. Что представляет собой поверхность отклика моделей, исследуемых методом статистических испытаний?
3. В чем заключается метод Монте-Карло?
4. Какие трудности возникают при исследовании стохастических моделей?
5. Какую информацию дает в руки исследователя полученное при статистическом исследовании распределение характеристик системы?
6. Какие законы распределения случайной величины Вы знаете?
7. Как выглядит плотность распределения для нормального закона?
8. Как выглядит плотность распределения для закона равной вероятности?
9. Как определяются оценки математического ожидания и дисперсии случайной величины?
10. Что такое выборочная статистика?
11. Почему она называется «выборочная»?
12. От чего зависит погрешность стохастического моделирования?

Глава 3. ЭМПИРИЧЕСКИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

3.1 Идентификация эмпирических математических моделей

Переход к эмпирическим моделям предполагает заведомый отказ от аналитических методов исследования. Поэтому эмпирические модели более разнообразны и включают в себя различные по форме математические зависимости.

При разработке эмпирической математической модели предполагается использование экспериментальных данных, полученных при испытаниях объектов. Результаты таких испытаний всегда представляют собой наборы величин, характеризующих работу объекта или системы при различных сочетаниях управляющих параметров.

Наиболее эффективным средством представления результатов экспериментов в системах математического моделирования являются эмпирические модели.

При построении эмпирической модели обычно предполагается, что физическая теория работы объекта отсутствует или по тем или иным причинам не может быть использована.

Объект идентификации представляет собой так называемый «черный ящик» с некоторым числом регулируемых (или, по крайней мере, измеряемых) входов x и одним или несколькими наблюдаемыми (измеряемыми) выходами (Рис. 3.1).

Здесь x_i – управляющие переменные; ω_i – неопределенности (шумы); q_i – ограничения; W – характеристическая функция.

Задачей идентификации является построение модели объекта по результатам наблюдений его реакции на возмущения внешней среды.

При этом необходимо учитывать ошибки, возникающие при измерении характеристик объекта.

Требуется построить зависимость (модель)

$$W = f(x),$$

которая описывает характеристики изучаемой системы.

Это уравнение называется уравнением регрессии и описывает поверхность (гиперповерхность) отклика, характеризующую эмпирическую модель.

Обычно предполагается, что имеющиеся экспериментальные данные дают достаточно информации для воссоздания математического описания объекта.

На рис. 3.2 показано решение задачи идентификации для некоторого набора данных, полученное с помощью линейной регрессионной зависимости: $W = a + bx$.

Идентификацию модели начинают с выбора формы модели, т.е. вида функции $f(x)$. При этом на практике может встретиться два случая:

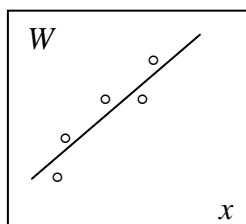


Рис. 3.2

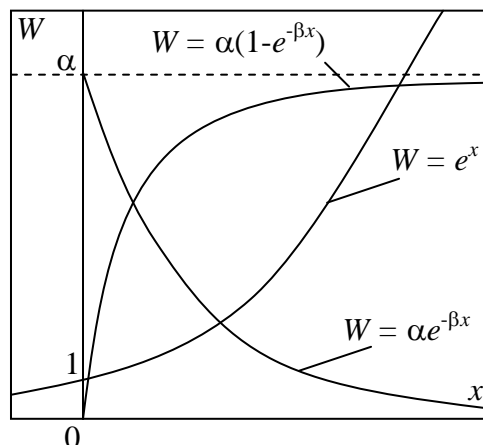


Рис. 3.3

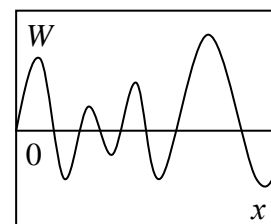


Рис. 3.4

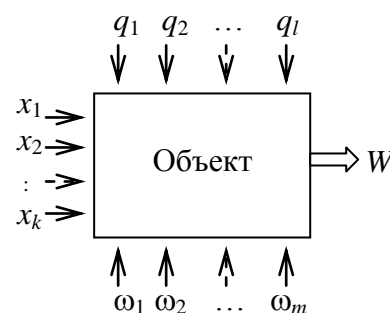


Рис. 3.1

1) Форма математической модели известна заранее, а задача идентификации сводится к определению коэффициентов этой модели. Так, описание ряда затухающих или развивающихся процессов дается зависимостями экспоненциального типа (Рис. 3.3). Задача исследования является определением коэффициентов α , β .

2) Форма математической модели заранее неизвестна. В этом случае для идентификации модели используются отрезки бесконечных рядов, а задача заключается в определении числа членов ряда и коэффициентов при этих членах. Модель может быть представлена в виде

$$W = \sum_{i=1}^k \beta_{0i} f_0(x_i) + \sum_{i=1}^k \beta_{1i} f_1(x_i) + \dots + \sum_{i=1}^k \beta_{li} f_l(x_i),$$

где $f_q(x_i)$ – некоторые заданные функции; β_{qi} – коэффициенты регрессии; $q = 0, 1, \dots, l$.

В одномерном случае ($k = 1$) уравнение принимает вид

$$W = \beta_0 f_0(x) + \beta_1 f_1(x) + \dots + \beta_l f_l(x).$$

Конкретный вид модели зависит от выбора функций $f_q(x)$, по которым производится разложение W . Например, при описании колебательных процессов удобно использовать ряд Фурье $W = \alpha_0 + \sum_{k=1}^{\infty} (\alpha_k \cos kx + \beta_k \sin kx)$ (Рис. 3.4).

Часто в качестве функций $f_0(x), f_1(x), f_2(x), \dots, f_l(x)$ выступают степенные функции $x^0, x^1, x^2, \dots, x^l$. Если ограничиться первыми членами разложения, то уравнения сведутся к линейным, квадратичным и другим полиномиальным моделям. Однако пока остается неясным, сколько членов ряда обеспечивает наилучшее описание изучаемого процесса.

Обычно берут количество экспериментальных точек значительно больше, чем количество коэффициентов регрессии. В этом случае нельзя построить поверхность отклика, проходящую через все экспериментальные точки. Да этого и не требуется. При этом, однако, можно построить приближенную модель, обеспечивающую в некотором смысле наилучшее совпадение с экспериментальными данными.

Например, прямая a построена по 10-ти экспериментальным точкам методом наименьших квадратов (Рис. 3.5); кривая b – квадратичная модель; c – полиномиальная модель 3-го порядка достаточно хорошо соответствует исходному экспериментальному материалу, хотя проходит не через все экспериментальные точки.

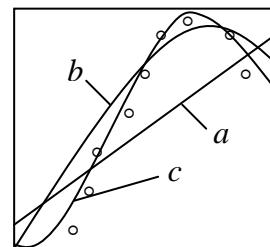


Рис. 3.5

Таким образом, для любой экспериментальной выборки могут быть предложены различные модели идентификации. Конкретная форма модели зависит от выбора функций $f_q(x)$ и количества членов ряда.

Сама постановка задачи идентификации включает в себя элемент неопределенности, возможность множественности решений. Важно выбрать лучшее или, по крайней мере, достаточно хорошее из этих решений.

Для оценки точности модели естественно использовать величины отклонений, полученных в эксперименте величин W_j и их оценок Wm_j , предсказанных моделью

$$\varepsilon_j = W_j - Wm_j. \quad (3.1)$$

Исключительное распространение получил метод наименьших квадратов отклонений реальных значений оцениваемой величины от значений, предсказанных моделью.

Специальные методы планирования эксперимента позволяют существенно повысить объем получаемой информации, улучшают характеристики эмпирических моделей, а также упрощают процедуру обработки экспериментальных данных. Однако на практике очень часто приходится иметь дело с неорганизованным (пассивным) экспериментом. Связано это, по крайней мере, с тремя причинами:

1) Исследователь может только наблюдать входы системы, но не может их регулировать, что полностью исключает возможность планирования эксперимента (типичная ситуация: астроном – галактика).

2) Неизвестны диапазоны возможного изменения переменных (входов), что затрудняет планирование эксперимента и исключает возможность использования ряда эффективных методов планирования.

3) Приходится строить модели идентификации на основе уже полученных ранее беспорядочных экспериментальных данных.

3.2. Использование метода наименьших квадратов

В качестве простого примера построения модели методом наименьших квадратов рассмотрим задачу восстановления математического описания некоторого процесса по результатам эксперимента.

Предполагается, что процесс описывается одномерным уравнением 2-го порядка

$$W = a_0 + a_1x + a_2x^2, \quad 0 \leq x \leq 6.$$

Считаем, что величина x измеряется точно, а W – с ошибкой ε , имеющей нормальное распределение с нулевым математическим ожиданием и единичной дисперсией $M(\varepsilon) = 0, \sigma^2(\varepsilon) = 1$.

Выборка десяти случайных пар (x, \tilde{W}) представлена в табл. 3.1 в графах 2 и 3.

Таблица 3.1

№	x	\tilde{W}	Wm	ε
1	2	3	4	5
1	4,8608	9,28	8,848	0,432
2	4,2396	9,40	8,821	0,579
3	2,7792	7,88	7,460	0,420
4	0,5988	1,86	2,039	-0,179
5	3,2136	7,77	8,056	-0,286
6	4,5156	8,73	8,874	-0,144
7	5,9340	8,33	8,118	0,212
8	1,5852	5,16	4,994	0,166
9	4,4880	7,28	8,872	-1,592
10	4,0932	9,22	8,767	0,453

Метод наименьших квадратов заключается в том, что неизвестные (искомые) коэффициенты a_0, a_1, a_2 должны минимизировать функцию, представляющую собой сумму квадратов невязок ε_j :

$$G = \sum_{j=1}^{10} \varepsilon_j = \sum_{j=1}^{10} (Wm_j - \tilde{W}_j)^2.$$

Минимум некоторой функции, как известно, находится в точке (a_0^*, a_1^*, a_2^*) , где все частные производные этой функции по переменным a_0, a_1, a_2 равны нулю.

Для определения частных производных, распишем функцию G через ее предполагаемый вид:

$$G = \sum_{j=1}^{10} (a_0 + a_1x_j + a_2x_j^2 - \tilde{W}_j)^2.$$

Возьмем от функции G производные по a_0, a_1, a_2 :

$$\frac{\partial G}{\partial a_0} = \sum_{j=1}^{10} [2(a_0 + a_1x_j + a_2x_j^2 - \tilde{W}_j) \cdot 1];$$

$$\frac{\partial G}{\partial a_0} = \sum_{j=1}^{10} [2(a_0 + a_1 x_j + a_2 x_j^2 - \tilde{W}_j) \cdot x_j];$$

$$\frac{\partial G}{\partial a_1} = \sum_{j=1}^{10} [2(a_0 + a_1 x_j + a_2 x_j^2 - \tilde{W}_j) \cdot x_j^2].$$

Приравняв эти выражения к нулю и произведя некоторые преобразования, получим систему линейных алгебраических уравнений третьего порядка с тремя неизвестными, коэффициенты которой вычисляются по известным данным из табл. 3.1:

$$\begin{cases} a_0 \cdot 10 + a_1 \sum_{j=1}^{10} x_j + a_2 \sum_{j=1}^{10} x_j^2 = \sum_{j=1}^{10} \tilde{W}_j; \\ a_0 \cdot \sum_{j=1}^{10} x_j + a_1 \sum_{j=1}^{10} x_j^2 + a_2 \sum_{j=1}^{10} x_j^3 = \sum_{j=1}^{10} \tilde{W}_j x_j; \\ a_0 \cdot \sum_{j=1}^{10} x_j^2 + a_1 \sum_{j=1}^{10} x_j^3 + a_2 \sum_{j=1}^{10} x_j^4 = \sum_{j=1}^{10} \tilde{W}_j x_j^2. \end{cases}$$

Решая полученную систему, получим $a_0 = -0,161$; $a_1 = 3,929$; $a_2 = -0,427$.

Таким образом, математическая модель будет иметь вид

$$Wm = -0,161 + 3,929x - 0,427x^2. \quad (3.2)$$

Проверим адекватность модели методом Фишера. Для этого заполним четвертый и пятый столбцы таблицы 3.1, подставляя в математическую модель (3.2) и затем в формулу (3.1) значения x_j из первого столбца.

Определим число степеней свободы системы по формуле

$$f_s = n - m - 1,$$

где $n = 10$ – количество экспериментальных точек; $m = 3$ – количество неизвестных коэффициентов. То есть $f_s = 6$.

Выборочная дисперсия вычисляется по формуле

$$s^2(\varepsilon) = \frac{1}{f_s} \sum_{j=1}^{10} (\tilde{W}_j - W_{mj})^2 = \frac{1}{6} \sum_{j=1}^{10} (\tilde{W}_j - W_{mj})^2 = 0,607.$$

Критерий Фишера вычисляется по формуле

$$F = \frac{s^2(\varepsilon)}{\sigma^2(W)} = 0,607.$$

По статистическим таблицам при 5%-м уровне риска ($\alpha = 0,05$) находим пороговое значение критерия Фишера

$$F_{f,\alpha} = F_{6,0,05} = 2,01.$$

Так как полученное значение F меньше критического (порогового), гипотеза об адекватности модели реальному процессу принимается.

Контрольные вопросы к лекции 7

1. Что является исходным материалом при построении эмпирической модели?
2. Как используется физическая теория работы объекта при построении эмпирической модели?
3. Что при этом представляет собой объект идентификации?
4. Сформулируйте задачу идентификации.
5. Что такое уравнение регрессии?
6. С чего начинается процесс идентификации?
7. От чего зависит конкретная форма модели?
8. Перечислите причины проведения unplanned эксперимента.
9. В чем заключается метод наименьших квадратов?

3.3. Статистические методы проверки адекватности математических моделей

Если имеются или могут быть получены необходимые и достоверные экспериментальные данные, для проверки адекватности моделей можно использовать методы математической статистики.

Математически задача проверки адекватности модели формулируется как задача проверки предположения о том, что значение отклика модели W_m отличается от реального отклика системы W не более чем на заданную величину ε^ :*

$$|W(x) - W_m| = |\varepsilon| \leq \varepsilon^*, \quad x \in X^*. \quad (3.3)$$

Однако, истинное значение отклика системы никогда неизвестно. Полученный в результате эксперимента отклик \tilde{W} в силу неконтролируемого дрейфа системы, разброса характеристик ее элементов и, наконец, просто ошибок измерения представляет собой случайную величину, отличающуюся от W . Поэтому при сравнении результатов математического и физического экспериментов (W_{mi}, \tilde{W}_i) будет получена совокупность случайных величин $\{\varepsilon_i\}$: $\tilde{W}_i(x) - W_{mi}(x) = \varepsilon_i, \quad i = 1, \dots, n$, среди которых могут оказаться как величины, удовлетворяющие условию (3.3), так и не удовлетворяющие ему.

Можно ли считать, что полученные отклонения ($\varepsilon_i > \varepsilon^*$) объясняются случайными причинами или их наличие должно быть признано существенным, что приводит к отказу от проверяемой модели. Для решения этого вопроса на основе выборки случайных величин $\{\varepsilon_i\}$ строят статистические критерии, по которым оценивают адекватность модели.

Гипотеза об адекватности модели действительности (гипотеза H_0) может быть сформулирована как предположение о том, что полученная совокупность $\{\varepsilon_i\}$ не дает оснований отказаться от рассматриваемой модели. Иными словами, модель удовлетворяет заданной точности ε^ .*

Альтернативная гипотеза H_1 состоит в том, что модель не отвечает заданным требованиям (3.3) и, следовательно, должна быть отвергнута.

Так как выборка $\{\varepsilon_i\}$ случайна, решение о выборе одной из гипотез H_0 или H_1 носит вероятностный характер. При этом может быть допущена ошибка первого рода, состоящая в отказе от правильной модели (принимается H_1 , когда верна H_0), или ошибка второго рода, состоящая в принятии ошибочной модели (принимается H_0 , когда верна H_1). Вероятность ошибки первого рода обозначают через α , второго рода – β . Принято называть α риском разработчика, β – риском потребителя. Разумеется, желательно минимизировать как α , так и β . Однако, при заданном объеме экспериментальной выборки уменьшение α влечет за собой увеличение β .

На практике α задается на определенном уровне ($\alpha = 0,05; 0,01; 0,005; 0,001$), при этом в 100 α % случаев правильная модель отвергается.

Величина $1 - \beta$ характеризует вероятность отказа от ошибочной модели, называется мощностью критерия и является мерой его эффективности.

Выбор вероятностей ошибок α и β при проверке конкретной модели зависит от ответственности решений, принимаемых на основе моделирования.

Например, если модель предназначена для управления двигателем летательного аппарата, необходимо в первую очередь минимизировать β , так как в данном случае принятие неверной модели, а значит, возможность ошибочных решений при управлении представляет больший вред, чем отказ от правильной модели.

Для оценки гипотезы об адекватности модели существует несколько критериев:

- 1) Критерий согласия χ^2 Пирсона.
- 2) Критерий Смирнова-Колмогорова.
- 3) Критерий Фишера и др.

При использовании критерия χ^2 проверке подлежит гипотеза о том, что рассматриваемая модель адекватна исследуемой системе с вероятностью p (например, $p = 0,95$). Это значит, что при n независимых испытаниях np значений ε_i должно удовлетворять условию (3.3) и лишь в $(1-p)n$ случаях это условие может быть нарушено.

В результате случайного эксперимента для этих событий будут получены частоты v_1 и v_2 : $v_1 \approx pn$; $v_2 \approx (1-p)n$; ($v_1 + v_2 = n$).

Частоты v_1 и v_2 отличаются от точных вероятностных оценок или из-за несоответствия модели действительности (заданная вероятность p не соблюдается), или из-за случайных отклонений.

Для оценки предположения о том, что отклонения v_1 и v_2 от соответствующих вероятностей случайны, строится функция

$$U^* = \frac{(v_1 - pn)^2}{pn} + \frac{[v_2 - (1-p)n]^2}{(1-p)n},$$

представляющая собой сумму квадратов отклонений, нормированных на соответствующие вероятности.

Полученное значение U^* сравнивается с табличным значением при заданном уровне риска α . Если U^* превышает пороговое значение $\chi_{1,\alpha}^2$, модель должна быть отвергнута, и принимается гипотеза H_1 . Если $U^* \leq \chi_{1,\alpha}^2$, экспериментальные данные не противоречат гипотезе об адекватности модели, и принимается гипотеза H_0 .

Необходимым условием использования критерия χ^2 является многочисленность экспериментальных данных (не меньше 20).

Критерий Смирнова-Колмогорова основан на максимальном значении отклонений

$$S = \sup \{ \varepsilon_i \} = \sup \{ \tilde{W}_i - W_{mi} \}.$$

Для заданной экспериментальной выборки строится вспомогательная функция

$$\lambda_n^* = S\sqrt{n},$$

которая сравнивается с пороговым значением $\lambda_{n,\alpha}$, определенным по таблицам распределения функции Смирнова-Колмогорова.

При $\lambda_n^* > \lambda_{n,\alpha}$ модель должна быть отвергнута, а при $\lambda_n^* \leq \lambda_{n,\alpha}$ экспериментальные данные не противоречат гипотезе об адекватности модели.

Критерий Смирнова-Колмогорова целесообразно использовать при относительно малых выборках, когда критерий χ^2 оказывается неэффективным.

Критерий Фишера осуществляется путем анализа дисперсий. Если дисперсия, характеризующая ошибку эксперимента $\sigma^2(W)$, известна, вычисляется выборочная дисперсия $S^2(\varepsilon)$ и составляется F -отношение:

$$F_{f_s,\infty} = \frac{S^2(\varepsilon)}{\sigma^2(W)}.$$

Полученную величину F -отношения сравнивают с пороговым значением критерия Фишера $F_{f_s,\infty,\alpha}$ при заданном уровне риска α .

При $F_{f_s,\infty} \leq F_{f_s,\infty,\alpha}$ полученная величина $S^2(\varepsilon)$ может быть объяснена случайным разбросом экспериментальных данных и, следовательно, нет оснований для отказа от проверяемой модели.

Если $F_{f_s,\infty} > F_{f_s,\infty,\alpha}$, полученное расхождение результатов моделирования и экспериментальных данных значимо и, следовательно, модель должна быть отвергнута как недостаточно точная.

Контрольные вопросы к лекции 8

1. Сформулируйте задачу проверки адекватности модели.
2. Что означает понятие «адекватность математической модели»?
3. В чем заключается ошибка первого рода?
4. В чем заключается ошибка второго рода?
5. Какие критерии проверки адекватности математической модели Вы знаете?
6. Охарактеризуйте каждый из этих критериев.

3.4. Идентификация параметров математической модели силы резания токарной операции

Построим математическую модель силы резания при обработке круглой детали на токарном станке (Рис. 3.6).

Сила резания P описывается математической моделью в виде позинома

$$P = C S^\alpha V^\beta t^\gamma, \quad (3.4)$$

где S – продольная подача; V – скорость резания; t – глубина резания; C , α , β , γ – неизвестные параметры.

Формула (3.4) является справочной. Для определения неизвестных параметров воспользуемся методом наименьших квадратов.

Пусть проведено n экспериментов, результаты которых сведены в таблицу 3.2.

Таблица 3.2

№	S (мм/об)	V (мм/с)	t (мм)	P (КГ)
1	S_1	V_1	t_1	P_1
2	S_2	V_2	t_2	P_2
⋮	⋮	⋮	⋮	⋮
n	S_n	V_n	t_n	P_n

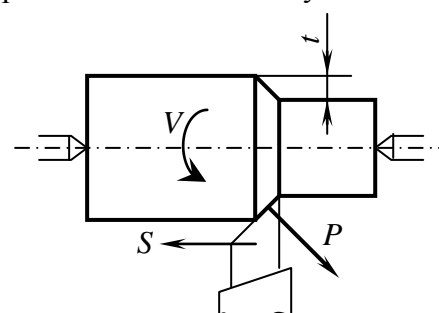


Рис. 3.6

Для упрощения решения поставленной задачи прологарифмируем выражение (3.4):

$$\ln P = \ln C + \alpha \ln S + \beta \ln V + \gamma \ln t.$$

Введем обозначения $\bar{P} = \ln P$; $\bar{C} = \ln C$; $\bar{S} = \ln S$; $\bar{t} = \ln t$.

Тогда формула (3.4) преобразуется к линейному виду:

$$\bar{P} = \bar{C} + \alpha \bar{S} + \beta \bar{V} + \gamma \bar{t}. \quad (3.5)$$

Метод наименьших квадратов сведется к минимизации функции

$$G = \sum_{i=1}^n (\bar{P}_i - \bar{P}_{mi})^2 = \sum_{i=1}^n (\bar{P}_i - \bar{C} - \alpha \bar{S}_i - \beta \bar{V}_i - \gamma \bar{t}_i)^2 \rightarrow \min,$$

где \bar{P}_i – логарифмы экспериментальных значений силы резания, взятых из табл. 3.2; \bar{P}_{mi} – логарифмы силы резания, предсказанные с помощью математической модели (3.5), $\bar{S}_i, \bar{V}_i, \bar{t}_i$ – логарифмы экспериментальных значений подачи, скорости и глубины резания, взятых из той же табл. (3.2), \bar{C} – логарифм неизвестного параметра C .

Возьмем производные от функции G по $\bar{C}, \alpha, \beta, \gamma$ и приравняем их к нулю:

$$\frac{\partial G}{\partial C} = -2 \sum_{i=1}^n (\bar{P}_i - \bar{C} - \alpha \bar{S}_i - \beta \bar{V}_i - \gamma \bar{t}_i) = 0;$$

$$\frac{\partial G}{\partial \alpha} = -2 \sum_{i=1}^n (\bar{P}_i - \bar{C} - \alpha \bar{S}_i - \beta \bar{V}_i - \gamma \bar{t}_i) \bar{S}_i = 0;$$

$$\frac{\partial G}{\partial \beta} = -2 \sum_{i=1}^n (\bar{P}_i - \bar{C} - \alpha \bar{S}_i - \beta \bar{V}_i - \gamma \bar{t}_i) \bar{V}_i = 0;$$

$$\frac{\partial G}{\partial \gamma} = -2 \sum_{i=1}^n (\bar{P}_i - \bar{C} - \alpha \bar{S}_i - \beta \bar{V}_i - \gamma \bar{t}_i) \bar{t}_i = 0.$$

Разделим обе части уравнений на -2 ; вынесем $\bar{C}, \alpha, \beta, \gamma$ за знак суммы; перенесем члены, не зависящие от $\bar{C}, \alpha, \beta, \gamma$, в правую часть:

$$\begin{aligned}
n\bar{C} + \alpha \sum_{i=1}^n \bar{S}_i + \beta \sum_{i=1}^n \bar{V}_i + \gamma \sum_{i=1}^n \bar{t}_i &= \sum_{i=1}^n \bar{P}_i; \\
\bar{C} \sum_{i=1}^n \bar{S}_i + \alpha \sum_{i=1}^n \bar{S}_i^2 + \beta \sum_{i=1}^n \bar{V}_i \bar{S}_i + \gamma \sum_{i=1}^n \bar{t}_i \bar{S}_i &= \sum_{i=1}^n \bar{P}_i \bar{S}_i; \\
\bar{C} \sum_{i=1}^n \bar{V}_i + \alpha \sum_{i=1}^n \bar{S}_i \bar{V}_i + \beta \sum_{i=1}^n \bar{V}_i^2 + \gamma \sum_{i=1}^n \bar{t}_i \bar{V}_i &= \sum_{i=1}^n \bar{P}_i \bar{V}_i; \\
\bar{C} \sum_{i=1}^n \bar{t}_i + \alpha \sum_{i=1}^n \bar{S}_i \bar{t}_i + \beta \sum_{i=1}^n \bar{V}_i \bar{t}_i + \gamma \sum_{i=1}^n \bar{t}_i^2 &= \sum_{i=1}^n \bar{P}_i \bar{t}_i.
\end{aligned} \tag{3.6}$$

Получили систему линейных алгебраических уравнений четвертого порядка, коэффициентами которой являются суммы произведений логарифмов экспериментальных данных. Решив полученную систему, найдем искомые значения коэффициентов \bar{C} , α , β , γ линейной модели (3.5).

Для определения параметров исходной модели (3.4) необходимо для коэффициента C (только для него) проделать операцию, обратную логарифмированию – потенцирование: $C = e^{\bar{C}}$. Коэффициенты α , β , γ получаются непосредственно из решения системы (3.6).

Если в распоряжении исследователя имеются экспериментальные данные, для проверки адекватности математической модели действительности можно использовать методы математической статистики. Рассматриваемый ниже метод пригоден при изучении любых математических моделей. Однако конкретный анализ проводится на примере построенной модели силы резания при точении с помощью критерия согласия χ^2 , предложенного Пирсоном.

Гипотеза H_0 формулируется как предположение о том, что отклонение ε экспериментальных данных от значений, предсказанных моделью (3.4), с вероятностью p (доверительная вероятность) укладываются в некоторый толерантный интервал $\pm \varepsilon^*$. Если это предположение правильно, то в толерантный интервал $(P \pm \varepsilon^*)$ должно укладываться np отклонений $\varepsilon_i = |P_i - P_{mi}|$. Вне толерантного интервала должно оказаться $(1-p)n$ отклонений. Для ограниченной случайной выборки из n наблюдений эти события будут наблюдаться с частотой v_1 и v_2 , лишь приближенно совпадающие с соответствующими вероятностями:

$$v_1 \approx np; \quad v_2 \approx (1-p)n; \quad v_1 + v_2 = n.$$

Необходимо установить, можно ли объяснить эти отклонения случайными причинами (в этом случае можно принять гипотезу H_0) или же они не случайны – статистически значимы (в этом случае нужно принять альтернативную гипотезу H_1).

Для этого вычисляется некоторая величина U , называемая статистикой:

$$U = \frac{(v_1 - np)^2}{np} + \frac{(v_2 - (1-p)n)^2}{(1-p)n}.$$

Эту величину нужно сравнить с пороговым значением χ^2 -критерия ($\chi^2_{1,\alpha}$) при принятом уровне риска α . Если $U \leq \chi^2_{1,\alpha}$, наблюдаемые отклонения частот от соответствующих вероятностей можно объяснить случайностью и нет оснований для отказа от нуль-гипотезы H_0 . Если $U > \chi^2_{1,\alpha}$, то или произошло маловероятное событие $(1-p)$, или наблюдаемые отклонения не случайны. В этом случае принимается гипотеза H_1 .

Вывод о правильности гипотезы H_1 , вообще говоря, не требует безоговорочного отказа от проверяемой модели:

1) Можно изменить исходные предположения с тем, чтобы увеличить толерантный интервал $\pm \varepsilon^*$ или уменьшить доверительную вероятность p . При этом уменьшатся отклонения v_1 и v_2 от соответствующих вероятностей, и проверка может привести к принятию

гипотезы H_0 . В этом случае моделью можно пользоваться, но нужно признать, что ее точность оказалась ниже, чем первоначально предполагалось.

2) Можно уменьшить уровень риска α (то есть вероятность отказа от правильной модели в результате неудачного эксперимента). Это приводит к увеличению порогового значения $\chi^2_{1,\alpha}$. Это, в свою очередь, может изменить оценку значения U . Однако нужно помнить, что при этом увеличивается риск признать правильной ошибочную модель.

3) Можно потребовать увеличения объема выборки, что, разумеется, приведет к увеличению точности оценки модели и уменьшению риска ошибок.

При проверке адекватности моделей действительности всегда рассматривается случай, когда за пределами толерантного интервала оказалось больше точек, чем ожидалось ($v_1 < pn$; $v_2 > (1-pn)n$). В противном случае опасений за точность модели не возникает, однако можно предположить, что величина толерантного интервала задана необоснованно большой. Если в результате проверки по критерию χ^2 в этом случае будет получена величина $U > \chi^2_{1,\alpha}$, то завышение толерантного интервала (или занижение доверительной вероятности p) статистически значимо, и необходимо уменьшить ε^* или увеличить p . В обоих случаях нужно признать, что модель оказалась точнее, чем ожидалось.

Контрольные вопросы к лекции 9

1. Приведите общий вид математической модели силы резания при точении.
2. Как привести модель, заданную в виде позинома, к линейному виду?
3. Каким методом найдены параметры линейной модели?
4. В чем заключается этот метод?
5. Как перейти от линейной модели к позиному?
6. Сформулируйте нуль-гипотезу проверки построенной модели на адекватность.
7. Что такое доверительная вероятность?
8. Перечислите меры, которые можно применить в случае неадекватности построенной математической модели.
9. В каком случае можно не проверять модель на адекватность?

3.5. Выбор оптимальной эмпирической модели

Принцип наименьших квадратов позволяет найти наилучшую модель идентификации для исследуемой экспериментальной выборки с заданным уравнением регрессии вида

$$W = \sum_{i=1}^k \beta_{0i} f_0(x_i) + \sum_{i=1}^k \beta_{1i} f_1(x_i) + \dots + \sum_{i=1}^k \beta_{li} f_l(x_i).$$

Если имеются достаточно веские основания для выбора формы этого уравнения, никаких проблем не возникает. Однако, в большинстве случаев конкретная форма модели заранее неизвестна и может, вообще говоря, быть различной.

На первый взгляд может показаться, что более сложная модель (увеличение степени полинома) всегда обеспечивает получение бóльшей точности. На самом деле это не так. При переходе к полиномам более высокой степени можно, конечно, получить лучшее согласие регрессионной кривой с экспериментальными данными. Для $m = n$ это согласие будет абсолютным, но при этом получится худшее согласие с истинным характером процесса $W(x)$. Дело в том, что экспериментальные данные представляют собой случайные величины и содержат лишь ограниченную информацию о характере $W(x)$. Увеличение степени полинома целесообразно лишь до тех пор, пока из экспериментальной выборки извлекается надежная информация. Таким образом, возникает проблема выбора формы модели.

Подход к решению этой проблемы основан на статистическом исследовании уравнений регрессии.

1) Метод всех возможных регрессий основан на последовательном изучении всех возможных моделей ($m < n$), из которых отбирается лучшая модель.

Метод представляется мало пригодным для анализа сложных систем, так как отличается высокой трудоемкостью.

2) Метод исключения предполагает исследование наиболее полной (в пределах разумного) модели и последовательную проверку на значимость всех ее членов. При этом для каждого из членов модели вычисляется величина критерия Фишера F . На основе полученного множества $\{F_i\}$ выбирается член уравнения регрессии, соответствующий минимальному значению критерия F_i . Если это минимальное значение меньше критического при выбранном уровне риска ($F_i < F_{кр \alpha}$), то соответствующий член исключается из регрессионного уравнения как несущественный, после чего все коэффициенты регрессии пересчитываются заново и вновь осуществляется проверка их значимости.

Если $F_i > F_{кр \alpha}$, то все члены модели существенны и уравнение регрессии остается в первоначальном виде. Однако, если это произошло уже на первом шаге исследования, стоит рассмотреть целесообразность усложнения первоначальной модели.

Трудоемкость этого метода меньше, чем метода всех возможных регрессий.

3) Метод включений по существу противоположен методу исключений и предусматривает последовательное включение в модель новых членов с проверкой их статистической значимости.

Трудоемкость этого метода существенно меньше трудоемкости рассмотренных выше методов.

Существуют и некоторые другие методы подбора оптимального уравнения регрессии.

Общим недостатком всех рассмотренных ранее методов является использование для оценки модели того же экспериментального материала, на основе которого эта модель построена.

4) Иной подход основан на использовании регуляризации. При этом подходе все экспериментальные данные разбиваются на две части: обучающую (n_1) и проверочную (n_2). Первая из них используется для определения коэффициентов регрессии модели, вторая – для оценки модели в целом.

Оптимальные по этому подходу модели мало чувствительны к небольшим изменениям исходных данных.

Число точек обучающей последовательности должно быть, по крайней мере, на единицу больше числа коэффициентов регрессии ($n_1 > m+1$). Для повышения достоверности результатов этот запас должен быть существенно увеличен ($n_1 \geq (2 \dots 3)m$). Проверочная последовательность должна включать в себя хотя бы одну точку.

В ряде случаев в качестве критерия регуляризации удобно использовать критерий несмещенности, обеспечивающий наименьшее изменение модели при изменении состава обучающей последовательности. При этом весь экспериментальный массив разбивается на две одинаковые по величине последовательности ($n_1 = n_2$), каждая из которых поочередно используется в качестве обучающей. В результате их использования определяются две независимые, одинаковые по форме модели $W_m^*(n_1)$ и $W_m^{**}(n_2)$. Оптимальная модель ищется по всем точкам выборки:

Критерий регуляризации всегда имеет четко выраженный минимум, что обеспечивает объективное выделение модели оптимальной сложности.

3.6. Использование критерия Фишера для проверки значимости высших степеней математической модели

Критерий Фишера может быть использован для сравнения точности двух (или нескольких) конкурирующих моделей.

Пусть рассматриваются две модели изучаемой системы (ω_1, ω_2), приводящие к двум различным множествам значений функции отклика: W_{m1}, W_{m2} .

Будем считать, что модель ω_2 более подробна и предположительно более точна, чем ω_1 . Для каждой из моделей может быть составлена остаточная сумма квадратов:

$$S_1 = \sum_{i=1}^n (W_i - W_{m1i})^2; f_{S_1} = n - m_1 - 1;$$

$$S_2 = \sum_{i=1}^n (W_i - W_{m2i})^2; f_{S_2} = n - m_2 - 1$$

и подсчитаны средние квадраты этих сумм (выборочные дисперсии):

$$s_1^2 = \frac{S_1}{f_{S_1}} = \frac{\sum_{i=1}^n (W_i - W_{m1i})^2}{n - m_1 - 1}; s_2^2 = \frac{S_2}{f_{S_2}} = \frac{\sum_{i=1}^n (W_i - W_{m2i})^2}{n - m_2 - 1}.$$

Для сравнения моделей подсчитывается так называемая дополнительная сумма квадратов SS , связанная с дополнительными данными, введенными в модель ω_2 , и характеризующаяся внесенными в нее уточнениями; а также число степеней свободы этой дополнительной суммы квадратов:

$$SS = S_1 - S_2; f_{SS} = f_{S_1} - f_{S_2} = m_2 - m_1.$$

Средний квадрат дополнительной суммы определяется соотношением

$$s_{SS}^2 = \frac{SS}{f_{SS}} = \frac{\sum_{i=1}^n (W_i - W_{m1i})^2 - \sum_{i=1}^n (W_i - W_{m2i})^2}{m_2 - m_1}.$$

Если известна дисперсия экспериментальных данных $\sigma^2(W)$, то роль дополнительной информации, содержащейся в модели ω_2 , оценивается путем сравнения F -отношения с пороговым (критическим) значением критерия Фишера:

$$F_1 = \frac{s_{SS}^2}{\sigma^2(W)}.$$

Если дисперсия экспериментальных данных $\sigma^2(W)$ неизвестна, сравнение проводится с оценкой дисперсии для упрощенной модели s_1^2 :

$$F_2 = \frac{s_{SS}^2}{s_1^2}.$$

Если полученное значение критерия Фишера значимо:

$$F_1 > F_{m_2 - m_1, \alpha} \quad \text{или} \quad F_2 > F_{m_2 - m_1, n - m_1 - 1, \alpha},$$

то дополнительная информация, заложенная в модели ω_2 , существенна, и модель ω_2 действительно отличается от модели ω_1 . В противном случае уточнения, вносимые моделью ω_2 , неразличимы на фоне шума; с точки зрения точности модели равноценны, и предпочтение должно быть отдано более простой модели ω_1 .

В частном случае полиномиальных моделей, представляющих собой конечные отрезки бесконечных рядов, этим методом можно проверить целесообразность включения в модель членов ряда с более высокими степенями.

Рассмотрим пример, приведенный в п. 3.2 (лекция 7) с построенной моделью в виде полинома второй степени. Проанализируем целесообразность использования для данной выборки кубической модели типа $W_m = a_0 + a_1x + a_2x^2 + \boxed{a_3x^3}$.

В результате расчетов методом наименьших квадратов можно найти коэффициенты такой модели:

$$W_m = -0,878 + 4,98x - 0,768x^2 + 0,032x^3.$$

Вычислим остаточную сумму квадратов:

$$S_3 = \sum_{i=1}^{10} (W_i - W_{mi})^2 = 4,537.$$

Напомним, что остаточная сумма квадратов для модели второго порядка имела значение $S_2 = 4,393$.

Дополнительная сумма квадратов

$$SS = S_2 - S_3 = 0,144; \quad f_{SS} = m_3 - m_2 = 4 - 3 = 1.$$

Средний квадрат дополнительной суммы

$$s_{SS}^2 = \frac{SS}{f_{SS}} = 0,144.$$

Критерий Фишера

$$F = \frac{s_{SS}^2}{s_2^2} = \frac{0,144}{0,607} = 0,2206.$$

По статистическим таблицам распределения определяем критическое (пороговое) значение критерия Фишера при количестве степеней свободы $f = n - m_2 - 1 = 10 - 3 - 1 = 6$ и уровне риска $\alpha = 0,05$

$$F_{f_{SS}, f, \alpha} = 5,59.$$

В связи с тем, что полученное из расчета значение критерия Фишера меньше критического, можно считать, что член третьего порядка не добавляет существенной информации и, следовательно, он является незначимым. Переход к модели третьего порядка нецелесообразен.

Контрольные вопросы к лекции 10

1. Обеспечивает ли получение бóльшей точности более сложная модель?
2. Перечислите методы выбора оптимальной модели.
3. На чем основан метод всех возможных регрессий?
4. На чем основан метод исключения?
5. На чем основан метод включений?
6. На чем основан подход регуляризации?
7. Опишите критерий проверки значимости высших степеней математической модели.

Глава 4. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ

4.1. Общие сведения о теории принятия решений

Принятие решений является основой любой деятельности человека.

Простейшая схема принятия решений включает в себя некоторую цель и совокупность способов ее достижения (Рис. 4.1).

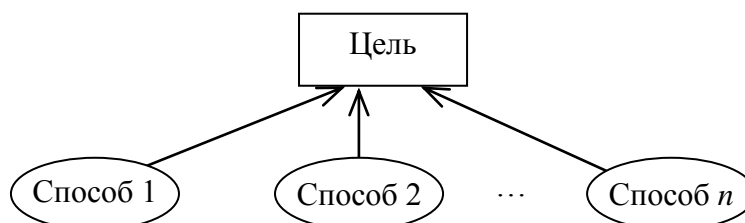


Рис. 4.1

Под целью будем понимать конкретный конечный результат, который необходимо получить путем выбора и реализации тех или иных способов деятельности. При этом в любом процессе принятия решений обязательно присутствует субъект принятия решений, который в общем случае представляется группой лиц, ответственных за целеполагание, формирование вариантов способов действий и, главное, за выбор конкретного решения.

При формировании оптимальных решений обязательным является наличие критериев оптимальности решений или по-другому целевых функций.

Критерием оптимальности называется математическое выражение, позволяющее количественно оценить степень достижения поставленной цели при выборе того или иного решения.

Задача ПР называется однокритериальной, если выбираемое решение служит достижению одной цели. Например, выбор управленческого решения по производственной программе предприятия, позволяющего получить максимум прибыли (цель) от реализации продукции.

Во многих ситуациях ПР объективно присутствует несколько целей.

Задачи ПР, удовлетворяющих нескольким целям, называются многокритериальными задачами. Например, при выборе проектных решений по новому пассажирскому самолету требуется обеспечить максимальное число пассажиров (цель 1) при минимальном расходе топлива (цель 2).

Отметим, что если в однокритериальных задачах возможно получение единственного оптимального решения (Рис. 4.2 а), то в многокритериальных ЗПР такая возможность отсутствует (Рис. 4.2 б).

В многокритериальных задачах возможно получение совокупности компромиссных вариантов (СКВ) решений на интервале $[x_{1 \text{ опт}}, x_{2 \text{ опт}}]$.

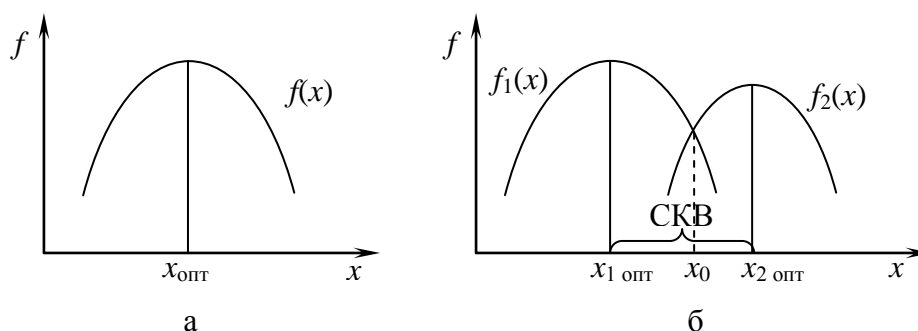


Рис. 4.2

4.2. Общая математическая модель формирования оптимальных решений

В математических моделях принятия решений в качестве нового знания выступает оптимальное решение, которое в наилучшем смысле соответствует достижению поставленной цели (целей).

Введем в рассмотрение n -мерный вектор $X = (x_1, x_2, \dots, x_n)$, определяющий количественные характеристики формируемого решения.

Обозначим через a, b, c вектора соответствующих размерностей, описывающие количественные характеристики неконтролируемых факторов.

Для оценки эффективности различных вариантов решений будем использовать специальным образом сформированную функцию:

$$W = f(c, X),$$

которая называется критерием оптимальности решений или целевой функцией задачи ПР.

Тогда выбор оптимального решения $X_{\text{опт}}$ будем осуществлять, исходя из требования $W(x) \rightarrow \max(\min)_{x \in X}$.

Множество X должно быть допустимым с точки зрения учета условий принятия решений (ограничений).

Пусть ЛПР обладает для достижения цели вектором ресурсов b . Представим в виде вектор-функции $\varphi(a, X)$ фактический расход ресурсов при использовании вектора решений X и вектора некоторых факторов a .

Тогда $\varphi(a, X) \leq b$ есть ограничение.

Во многих задачах ПР учитывается условие $X \geq 0$.

Таким образом, общая математическая модель формирования оптимальных решений может быть представлена в следующем виде:

$$W = f(c, X) \rightarrow \max; \quad (4.1)$$

$$\varphi(a, X) \leq b; \quad (4.2)$$

$$X \geq 0. \quad (4.3)$$

Постановка задачи в этом случае выглядит следующим образом:

|| *Найти значение вектора X , доставляющего максимум (минимум) критерию оптимальности решений (4.1) и удовлетворяющего при этом условиям (4.2) и (4.3).*

Математическая модель ПР (4.1) – (4.3) является однокритериальной моделью.

Если ЛПР должен учитывать m целей, то, формализуя их в виде критериев оптимальности, получим:

$$W_1 = f_1(c_1, X) \rightarrow \max;$$

$$W_2 = f_2(c_2, X) \rightarrow \min;$$

.....

$$W_m = f_m(c_m, X) \rightarrow \text{extr};$$

(4.4)

где c_1, c_2, \dots, c_m – вектора неконтролируемых факторов.

Математическая модель (4.4), (4.2), (4.3) является многокритериальной моделью.

В реальных задачах ПР ограничения вида (4.2) могут включать в себя как неравенства вида « \leq », « \geq », « $=$ », так и их различные сочетания.

4.3. Построение и решение оптимизационной задачи принятия решения (Задача о баке)

Пусть требуется выбрать геометрические размеры цилиндрического бака объемом V из условия минимального расхода материала на его изготовление.

Для построения математической модели введем в рассмотрение вектор проектных решений $X = (r, h)$, где $2r, h$ – диаметр и высота бака (Рис. 4.3).

Если предположить, что бак изготавливается сваркой из трех деталей, то расход материала при произвольном векторе решений X будет равен площади поверхности бака:

$$S = 2\pi r^2 + 2\pi r h \rightarrow \min_{r,h}. \quad (4.5)$$

Согласно условиям задачи выражение (4.5) является целевой функцией (критерий оптимальности проектных решений).

Условие того, что бак должен иметь объем заданного значения V , представим в виде:

$$\pi r^2 h = V. \quad (4.6)$$

На компоненты вектора решений X необходимо наложить дополнительные условия:

$$R > 0, h > 0. \quad (4.7)$$

Выражения (4.5) – (4.7) описывают нелинейную однокритериальную модель формирования оптимальных решений, при $n = 2, m = 1$.

Пусть бак должен иметь минимальную трудоемкость его изготовления. Если считать трудоемкости изготовления крышки, дна и боковой стенки достаточно малыми величинами, то затраты времени на изготовление бака будут пропорциональны длине свариваемых швов:

$$T = c(4\pi r + h) \rightarrow \min_{r,h}, \quad (4.8)$$

где c – затраты времени на сварку единицы длины.

Выражения (4.5), (4.8), (4.6), (4.7) описывают двухкритериальную нелинейную модель формирования оптимальных решений.

При построении математической модели в этой задаче принятия решений были использованы известные геометрические закономерности.

Аналитическое решение задачи ПР возможно, если соответствующая математическая модель включает в себя ограничения типа равенств, то есть имеет вид:

$$W = f(c, X) \rightarrow \text{extr}_X;$$

$$\varphi(a, X) = b;$$

$$-\infty < X < \infty.$$

Такие задачи решаются обычно классическими методами условной оптимизации, которые предусматривают построение функции Лагранжа вида

$$L(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m) = f(c, x_1, \dots, x_n) + \sum_{j=1}^m \lambda_j [\varphi_j(a, x_1, \dots, x_n) - b_j], \quad (4.9)$$

где $\lambda_1, \lambda_2, \dots, \lambda_m$ – неопределенные множители Лагранжа.

Точки экстремума этой функции определяются из решения системы уравнений вида

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= 0, \quad i = 1, \dots, n; \\ \frac{\partial L}{\partial \lambda_j} &= 0, \quad j = 1, \dots, m. \end{aligned} \quad (4.10)$$

Решая эту систему, получим решение вида

$$\begin{aligned} x_i^{\text{опт}} &= \psi_i(a, b, c), \quad i = 1, \dots, n; \\ \lambda_j &= \lambda_j(a, b, c), \quad j = 1, \dots, m. \end{aligned} \quad (4.11)$$

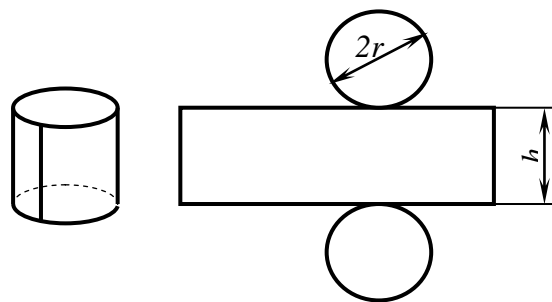


Рис. 4.3

Используем этот метод для решения однокритериальной задачи (4.8), (4.6) (без учета (4.5), (4.7)).

Функция Лагранжа имеет вид:

$$L(r, h, \lambda) = c(4\pi r + h) + \lambda(\pi r^2 h - V).$$

Система уравнений (4.17) относительно переменных r, h, λ :

$$\frac{\partial L}{\partial r} = 4\pi c + 2\lambda\pi r h = 0;$$

$$\frac{\partial L}{\partial h} = c + \lambda\pi r^2 = 0;$$

$$\frac{\partial L}{\partial \lambda} = \pi r^2 h - V = 0.$$

Имеем систему алгебраических уравнений, решая которую, получим значения неизвестных r, h (λ находить необязательно):

$$r = \sqrt[3]{\frac{V}{2\pi^2}}; \quad h = \sqrt[3]{4\pi V}; \quad \lambda = -c \sqrt[3]{\frac{4\pi}{V^2}}.$$

Таким образом, оптимальные размеры бака, найденные с помощью аналитического метода условной оптимизации, не зависят от затрат времени c на сварку единицы длины, но зависят от требуемого объема бака V . Требование (4.8) при этих значениях r и h выполняется, то есть трудоемкость будет минимальной.

Недостатками этого метода являются:

- 1) Не учитываются в явном виде условия неотрицательности (4.7).
- 2) Система уравнений (4.10) позволяет получить решение в форме (4.11) только для простых функций (4.5), (4.6).

Контрольные вопросы к лекции 11

1. Что включает в себя простейшая схема принятия решений?
2. Что такое цель?
3. Что такое критерий оптимальности?
4. Что такое однокритериальная ЗПР?
5. Что такое многокритериальная ЗПР?
6. Возможно ли получение единственного оптимального решения в многокритериальных задачах?
7. Напишите общий вид математической модели формирования оптимальных решений.
8. Сформулируйте задачу принятия решений.
9. Запишите критерий минимального расхода материала для задачи о баке.
10. Запишите критерий минимальной трудоемкости для задачи о баке.
11. Запишите общий вид функции Лагранжа.
12. Перечислите недостатки аналитического метода условной оптимизации.

4.4. Многокритериальные задачи принятия решений

Во многих задачах принятия решений имеется несколько целей, которые хочет достичь ЛПР. Такие задачи сводятся к многокритериальным задачам вида:

$$W_1 = f_1(X) \rightarrow \underset{X}{extr};$$

.....

$$W_k = f_k(X) \rightarrow \underset{X}{extr};$$

$$\varphi(X) \leq 0; \quad X \geq 0,$$

где $X(x_1, x_2, \dots, x_n)$ – вектор решений.

Наибольшее распространение на практике решения таких задач получил подход, связанный с работами итальянского математика-экономиста Викторио Парето. Он обеспечивает ЛПР возможность гибкого принятия решений. При оптимизации по Парето строится множество «неулучшаемых» решений, изменение каждого из которых ухудшает значение целевых функций $f_1(X), f_2(X), \dots, f_k(X)$.

Рассмотрим наиболее распространенную на практике двухкритериальную задачу оптимизации вида:

$$W_1 = f_1(x_1, x_2) \rightarrow \max; \tag{4.12}$$

$$W_2 = f_2(x_1, x_2) \rightarrow \max;$$

$$\varphi(x_1, x_2) \leq 0; \tag{4.13}$$

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

Условия (4.13) определяют множество допустимых решений и образуют на плоскости x_1Ox_2 некоторую область, каждой точке C которой соответствует точка C^* в пространстве значений критериев W_1OW_2 (Рис. 4.4). Ее координаты $W_1^{C^*}; W_2^{C^*}$ вычисляются по формулам (4.12) при $x_1 = x_1^C; x_2 = x_2^C$:

$$W_1^{C^*} = f_1(x_1^C, x_2^C); \quad W_2^{C^*} = f_2(x_1^C, x_2^C).$$

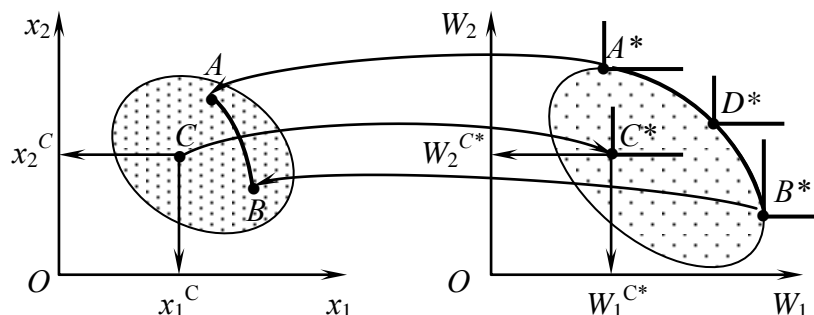


Рис. 4.4

Рассмотрим в множестве значений критериев четыре точки A^*, B^*, D^* и C^* (см. рис. 4.4). Точка A^* является оптимальной для критерия $W_2 = f_2(x_1, x_2)$, так как в этой точке критерий W_2 имеет максимальное значение. Аналогично точка B^* является оптимальной для критерия $W_1 = f_1(x_1, x_2)$. Точка C^* является «заведомо плохой» точкой, она не является оптимальной ни для одного критерия, так как в области значений критериев можно найти «более лучшую» точку D^* такую, что $W_1^{D^*} > W_1^{C^*}; W_2^{D^*} > W_2^{C^*}$. Для точек A^*, B^*, D^* более «лучших» точек в пространстве значений критериев не существует. Такие точки составляют множество решений, оптимальных по Парето в пространстве значений критериев. В нашем случае это точки кривой $A^*D^*B^*$. Для выделения «лучших» (неулучшаемых) точек используется понятие конуса K_i с вершиной в точке (W_1^i, W_2^i) (Рис. 4.5). Уравнения этого конуса имеют вид: $W_1 \geq W_1^i; W_2 \geq W_2^i$.

Правило выделения «лучших» точек:

Если в конусе K_i лежит хотя бы одна точка (W_1^j, W_2^j) , то она является более предпочтительной, чем точка (W_1^i, W_2^i) (см. рис. 4.5).

Тогда все точки множества значений критериев, для которых соответствующие конусы являются пустыми, являются парето-оптимальными решениями в пространстве значений критериев.

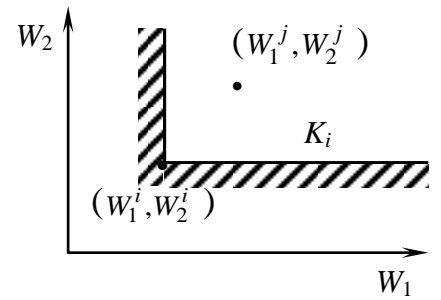


Рис. 4.5

Для нашего примера конусы, построенные во всех точках кривой $A^*D^*B^*$ (см. рис. 4.4), являются пустыми. Строя обратное отображение этих точек в пространство решений X , можно получить множество искомых решений (кривая AB на рис. 4.4), оптимальных по Парето. Такое множество называется множеством компромиссов, множеством эффективных точек или множеством Парето. Построив множество компромиссов, ЛПР выбирает в нем из неформальных соображений некоторую точку, которая является наилучшим компромиссом, по мнению ЛПР.

4.5. Построение решений, оптимальных по Парето (Двухкритериальная задача о баке)

Вернемся к рассмотренной в п. 4.5 задаче о баке, описанную формулами (4.5) – (4.8), и попробуем найти паретооптимальное ее решение. Для этого введем параметры $\alpha_1 > 0$ и $\alpha_2 > 0$, удовлетворяющие условию $\alpha_1 + \alpha_2 = 1$, и построим линейную свертку критериев (4.5) и (4.8):

$$F(X, \alpha_1, \alpha_2) = \alpha_1 S(X) + \alpha_2 T(X). \quad (4.14)$$

Положим $\alpha_1 = 1$, тогда α_2 обращается в ноль, свертка (4.14) принимает вид $F(X) = S(X)$ и задача превращается в однокритериальную. Решив эту задачу, найдем оптимальную точку X_S для обеспечения минимального расхода материала. Теперь положим $\alpha_1 = 0$, тогда $\alpha_2 = 1$, свертка (4.14) примет вид $F(X) = T(X)$. Решив эту однокритериальную задачу, найдем оптимальную точку X_T , обеспечивающую минимальную трудоемкость изготовления бака. Если провести аналогию с рис. 7.4, то X_S соответствует точке A , X_T соответствует точке B . Чтобы определить промежуточные точки (остальные компромиссные решения), введем обозначение $\alpha = \alpha_1$. Тогда $\alpha_2 = 1 - \alpha$. Формула (4.14) примет вид $F(X, \alpha) = \alpha S(X) + (1 - \alpha)T(X)$. Или:

$$F(r, h, \alpha) = \alpha S(r, h) + (1 - \alpha)T(r, h).$$

Функция Лагранжа запишется в виде

$$L(r, h, \lambda) = F(r, h, \alpha) + \lambda \varphi(r, h, \alpha).$$

Распишем функцию Лагранжа подробнее:

$$L(r, h, \lambda) = \alpha(2\pi r^2 + 2\pi r h) + (1 - \alpha)C(4\pi r + h) + \lambda(\pi r^2 h - V) \rightarrow \min_{r, h, \lambda}.$$

Чтобы найти минимум функции Лагранжа, нужно взять от нее производные по искомому переменным r , h , λ и приравнять их к нулю.

$$\begin{aligned} \frac{\partial L}{\partial r} &= 2\pi\alpha(2r + h) + 4\pi(1 - \alpha)C + 2\pi\lambda r h = 0; \\ \frac{\partial L}{\partial h} &= 2\alpha\pi r + (1 - \alpha)C + \pi\lambda r^2 = 0; \\ \frac{\partial L}{\partial \lambda} &= \pi r^2 h - V = 0. \end{aligned} \quad (4.15)$$

Получили систему трех алгебраических уравнений, решив которую найдем зависимость r , h , λ от α . Задавая α от 0 до 1, получим множество решений, оптимальных по

Парето. Поскольку аналитически решить систему (4.15) довольно сложно, можно воспользоваться любым численным методом, задавая предварительно значения α с любым приемлемым шагом.

Для примера эта задача была решена с шагом 0,1 в пакете *MathCad*. На рис. 4.6 показано полученное множество паретооптимальных решений.

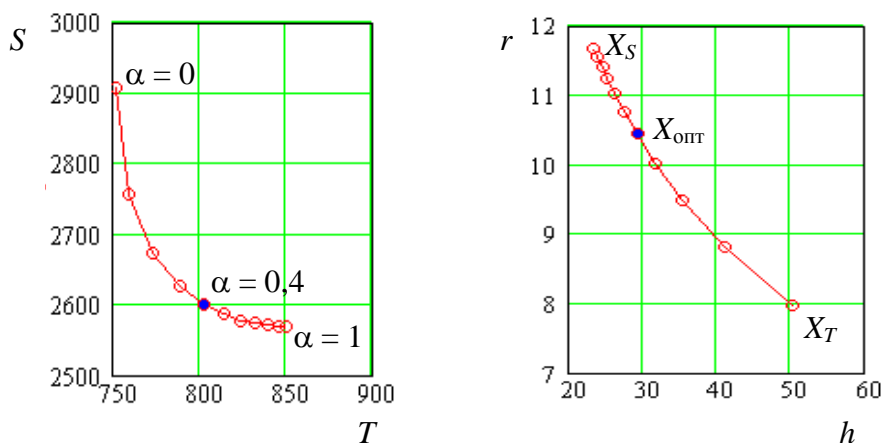


Рис. 4.6

ЛПР выбрал из этого множества точку при $\alpha = 0,4$, при котором $S = 2601$; $T = 802$; $r = 10,44$; $h = 29,2$. Эта точка устроила его потому, что при дальнейшем увеличении α S уменьшается уже незначительно и T имеет наименьшее значение из всех последующих.

Контрольные вопросы к лекции 12

1. Какие решения называются паретооптимальными?
2. Сформулируйте правило выделения лучших точек.
3. Что такое множество компромиссных решений?
4. Как получить множество компромиссных решений?
5. Запишите функцию Лагранжа для двухкритериальной задачи о баке.
6. Как найти минимум функции Лагранжа?