

Федеральное агентство по образованию
Томский государственный университет
систем управления и радиоэлектроники

Г.Н. Решетникова

МОДЕЛИРОВАНИЕ СИСТЕМ

Учебное пособие

Рекомендовано Сибирским региональным отделением
учебно-методического объединения по образованию
в области радиотехники, электроники, медицинской техники
и автоматизации в качестве учебного пособия
для студентов специальностей 210100 (220201)
«Управление и информатика в технических системах»
и 210200 (220301) «Автоматизация технологических
процессов и производств в приборостроении»

Томск
2005

УДК 681.51.015.26(075.8)

ББК 22.181я73

Р 47

Рецензенты:

кафедра оптимизации систем управления

Томск. политехн. ун-та, зав. каф., профессор, д-р техн. наук **В.А. Силич**;

профессор кафедры прикладной математики Томск. гос. ун-та

д-р техн. наук **В.И. Смагин**.

Решетникова Г.Н.

Р 47 Моделирование систем : учеб. пособие / Г.Н. Решетникова ;
Федеральное агентство по образованию, Томск. гос. ун-т систем
упр. и радиоэлектроники. — Томск : Томск. гос. ун-т систем упр.
и радиоэлектроники, 2005. — 261 с.

ISBN 5-86889-255-0

Даны основные понятия моделирования систем, выделены отдельные виды моделирования: математическое, физическое, статистическое, имитационное, агрегативное.

Рассмотрены численные методы и методы проектирования следящих систем адаптивного управления для многомерных стохастических моделей объектов, функционирующих в условиях неполной информации о состоянии объекта и его параметров. Приведены примеры построения математических моделей и краткое описание программных средств, используемых для моделирования.

Для студентов вузов, обучающихся по специальности «Информатика и управление в технических системах», а также специалистов, желающих познакомиться с методами компьютерного моделирования.

УДК 681.51.015.26(075.8)

ББК 22.181я73

ISBN 5-86889-255-0

© Решетникова Г.Н., 2005

© Томск. гос. ун-т систем упр.
и радиоэлектроники, 2005

Содержание

Предисловие	7
1. ОСНОВНЫЕ ПОНЯТИЯ МОДЕЛИРОВАНИЯ СИСТЕМ	9
1.1. Модель	10
1.1.1. Свойства моделей	10
1.1.2. Классификации моделей	12
1.1.3. Основное определение модели	14
1.2. Система	14
1.2.1. Историческое развитие понятия «система»	15
1.2.2. Модель «черного ящика»	16
1.2.3. Модель состава системы	17
1.2.4. Модель структуры системы	18
1.2.5. Структурная схема системы	22
1.2.6. Классификации систем	22
1.2.7. Основные понятия и определения	26
1.2.8. Анализ и синтез систем	28
1.2.9. Основные системные принципы	30
1.3. Моделирование	31
1.3.1. Классификация методов представления систем	31
1.3.2. Методы формализованного представления систем	32
1.3.3. Методы неформализованного представления систем ...	33
1.3.4. Классификация видов моделирования	35
2. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ	38
2.1. Цели и задачи математического моделирования	38
2.2. Этапы построения математических моделей	39
2.3. Основные характеристики математических моделей	40
2.4. Классификация математических моделей	44
2.5. Формы представления математических моделей	44
2.5.1. Непрерывно-детерминированные модели	44
2.5.2. Дискретно-детерминированные модели	46
2.5.3. Дискретно-вероятностные модели	49
2.5.4. Непрерывно-вероятностные модели	50
2.6. Методы упрощения математических моделей	51
3. ЭЛЕМЕНТЫ ТЕОРИИ ПОДОБИЯ	54
3.1. Использование теории подобия при физическом моделировании	54
3.1.1. Анализ размерностей	55
3.1.2. Физическое моделирование	57
3.2. Математическое подобие	59
3.3. Элементы статистической теории подобия	61
4. СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ	65
4.1. Методы имитации случайных элементов на ЭВМ	65
4.1.1. Принципы моделирования случайных элементов	66

4.1.2. Типы датчиков базовых случайных величин	68
4.1.3. Методы построения программных датчиков	69
4.2. Моделирование случайных величин	71
4.2.1. Моделирование дискретных случайных величин	71
4.2.2. Моделирование непрерывных случайных величин	73
4.2.3. Моделирование случайных процессов	76
4.3. Методы Монте-Карло	77
4.3.1. Простейший метод Монте-Карло вычисления интегралов	78
4.3.2. Геометрический метод Монте-Карло вычисления интегралов	82
5. ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ	86
5.1. Принципы построения имитационных моделей	87
5.2. Организация процесса моделирования	88
5.3. Способы имитации	91
5.4. Этапы имитационного моделирования	92
5.5. Планирование имитационных экспериментов	95
5.5.1. Стратегическое планирование имитационного эксперимента	96
5.5.2. Тактическое планирование имитационного эксперимента	98
5.6. Обработка и анализ результатов моделирования	99
5.6.1. Оценка качества имитационной модели	99
5.6.2. Подбор параметров распределений	103
5.6.3. Оценка влияния и взаимосвязи факторов	106
5.7. Достоинства и недостатки имитационного моделирования	110
6. АГРЕГАТИВНОЕ МОДЕЛИРОВАНИЕ	111
6.1. Основные характеристики агрегативной модели	111
6.2. Математическое описание процесса функционирования агрегата	113
6.3. Пример агрегативного моделирования	116
7. ЧИСЛЕННЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ	118
7.1. Элементы теории погрешности	119
7.2. Приближение данных	122
7.2.1. Интерполирование	122
7.2.2. Сплайн-функции	131
7.2.3. Аппроксимация данных методом наименьших квадратов	136
7.3. Численное дифференцирование	143
7.4. Численное интегрирование	144
7.4.1. Интерполяционные квадратурные формулы	145
7.4.2. Квадратурные формулы наивысшей алгебраической степени точности. Квадратурные формулы Гаусса	150

7.5. Решение нелинейных уравнений и систем	152
7.5.1. Решение нелинейных уравнений	153
7.5.2. Метод Лобачевского при решении алгебраических уравнений	161
7.5.3. Решение систем нелинейных уравнений	166
7.6. Решение задач матричной алгебры	167
7.6.1. Некоторые понятия матричной алгебры	167
7.6.2. Обусловленность систем и матриц	171
7.6.3. Метод Гаусса	174
7.6.4. Метод квадратного корня	176
7.6.5. Итерационные методы решения систем линейных алгебраических уравнений	178
7.6.6. Метод Данилевского при решении полной проблемы собственных значений. Построение канонической формы Фробениуса	183
7.6.7. Метод Фаддеева	186
7.7. Решение обыкновенных дифференциальных уравнений и систем	187
7.7.1. Методы Эйлера и Рунге — Кутта при решении задачи Коши для дифференциального уравнения первого порядка	187
7.7.2. Правило Рунге	190
7.7.3. Методы Эйлера и Рунге — Кутта при решении задачи Коши для систем дифференциальных уравнений	191
7.7.4. Решение задачи Коши для систем линейных дифференциальных уравнений	192
7.7.5. Методы Адамса при решении задачи Коши для обыкновенных дифференциальных уравнений	193
7.7.6. Метод сеток решения линейных краевых задач	195
7.7.7. Метод коллокации решения краевых задач	197
8. МОДЕЛИРОВАНИЕ СИСТЕМ УПРАВЛЕНИЯ	200
8.1. Описание систем в пространстве состояний	200
8.2. Аналитическое конструирование оптимальных регуляторов	203
8.2.1. Оптимальное управление при минимизации классического квадратичного функционала	205
8.2.2. Оптимальное управление при минимизации функционала обобщенной работы	207
8.2.3. Оптимальное управление при минимизации локального квадратичного критерия	210
8.2.4. Синтез следящей системы управления	211
8.3. Моделирование систем оптимального управления	212
8.3.1. Основные понятия цифровых систем управления	213
8.3.2. Моделирование поведения управляемого объекта	214
8.3.3. Синтез оптимального управления	217

8.4. Моделирование систем управления при случайных внешних воздействиях	222
8.4.1. Моделирование поведения объекта при наличии внешних возмущений	222
8.4.2. Описание математической модели измерительного комплекса	224
8.4.3. Оценивание состояния модели объекта	225
8.4.4. Синтез управляющих воздействий по оценкам состояния	228
8.5. Синтез адаптивной следящей системы	230
8.5.1. Основные понятия адаптивных систем управления ...	231
8.5.2. Одновременное оценивание состояния и параметров модели объекта	232
8.6. Учет ограничений и запаздываний по управлению	236
8.7. Общая схема синтеза адаптивных систем управления	236
9. ПРИМЕРЫ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ	239
9.1. Построение математических моделей прямолинейного и вращательного движения	239
9.1.1. Математическая модель прямолинейного движения	239
9.1.2. Математическая модель вращательного движения	241
9.2. Макроэкономическая модель динамики фондов производственного накопления и потребления	241
9.3. Построение математических моделей производства, хранения и сбыта товара повседневного спроса	243
9.4. Построение математической модели вертикального движения ракеты	246
10. ПРОГРАММНЫЕ СРЕДСТВА МОДЕЛИРОВАНИЯ	248
10.1. Языки программирования	248
10.1.1. Язык Fortran	248
10.1.2. Язык C	249
10.1.3. Язык C++	250
10.1.4. Языки Pascal и Object Pascal	252
10.2. Системы разработки программного обеспечения	252
10.2.1. Microsoft Visual C++ и MFC	252
10.2.2. Borland Delphi	253
10.2.3. Borland C++Builder	253
10.3. Специализированные математические пакеты	253
10.3.1. Система Maple	254
10.3.2. Система Mathematica	255
10.3.3. Система MatLAB	256
10.3.4. Система MathCAD	257
10.3.5. Система STATISTICA	258
10.3.6. Система EXCEL	258
ЛИТЕРАТУРА	259

Предисловие

Данное учебное пособие предназначено для студентов вузов, обучающихся по специальности «Информатика и управление в технических системах» и полностью соответствует содержанию Государственного образовательного стандарта по дисциплине «Моделирование систем». Оно может быть полезным для студентов, обучающихся по инженерно-техническим и экономическим специальностям, а также для специалистов, желающих познакомиться с методами компьютерного моделирования.

Понятие системы настолько широко, что может быть применено практически к любой сфере человеческой деятельности. Соответственно, и моделирование систем как инструмент анализа и синтеза реальных объектов, процессов, явлений не ограничивается рамками отдельной специальности. Поэтому в первых главах учебного пособия основное внимание уделено общеметодологическим аспектам моделирования: определениям, классификации, описанию отдельных типов и видов моделей и их соотносению с теми или иными классами реальных объектов и процессов.

Развитие науки, техники и общества в целом настоятельно требуют использования все более сложных математических моделей изучаемых явлений и процессов. При решении задач, связанных с повышением эффективности производства, управлением различными объектами, обнаружением радио- и гидролокационных сигналов на фоне помех и так далее, все чаще приходится иметь дело с объектами, математические модели которых описываются динамическими стохастическими системами, в которых возмущения или помехи, действующие на объект, рассматриваются как случайные величины. В связи с этим более подробно рассматриваются вопросы статистического моделирования как универсального средства представления большинства сложных систем.

Переживаемый в настоящее время этап научно-технического прогресса является этапом информатизации, который характеризуется расширением внедрения информационно-вычислительных систем, созданием и развитием методов автоматического управления в технике, экономике, медицине, биологии и других видах деятельности. В современной теории автоматического управления доминирует подход, предусматривающий концепцию совмещенного синтеза систем управления, как самый приемлемый для реализации на управляющих компьютерах. Наиболее перспективными методами, позволяющими решать задачи совмещенного синтеза, являются методы, основанные на оптимизации квадратичных критериев, а использование прогнозирующих моделей и скользящего интервала времени позволяет существенно расширить возможности синтеза адаптивного управления. В настоящем учебном пособии рассматриваются методы синтеза систем

управления для решения задач слежения для многомерных стохастических моделей объектов, функционирующих в условиях неполной информации о состоянии объекта и его параметрах. Заметим, что задача слежения является одной из основных в теории управления. Такие задачи возникают при переводе технологического процесса из одного режима в другой, при управлении подвижными объектами для маневрирования. Типичным примером задачи слежения является задача прецизионного слежения антенны радиолокатора за летящим объектом при воздействии на антенну ветровых возмущений и при непредвиденных маневрах летящей цели. Единственным способом решения таких задач является имитационное моделирование, при котором возникает необходимость в численных методах. В связи с этим в пособии рассмотрены численные методы решения некоторых задач, при этом более подробно рассмотрены методы приближения данных и решения обыкновенных дифференциальных уравнений и систем, которые используются при моделировании систем управления.

Несмотря на наличие большого количества научной и учебной литературы, посвященной, с одной стороны, проблемам общей теории систем, и, с другой стороны, достаточно частным прикладным вопросам моделирования, существует дефицит учебно-методического обеспечения, интегрирующего общие и частные аспекты моделирования систем независимо от их природы и происхождения. В учебном пособии сделана попытка такой интеграции с определенным акцентом в сторону моделирования систем управления, продиктованного как потребностями конкретной специальности, так и научными интересами автора. Кроме того, все рассмотренные методы доведены до непосредственно реализуемых алгоритмов, что вызвано необходимостью организации компьютерного практикума.

В последней главе учебного пособия приводится краткое описание программных средств, как инструментальных, так и специализированных математических пакетов, наиболее часто используемых для моделирования. Эта глава написана в соавторстве с Борисовым Сергеем Ивановичем, преподавателем кафедры компьютерных систем в управлении и проектировании Томского государственного университета систем управления и радиоэлектроники.

В основе пособия лежат учебные курсы «Моделирование систем управления», «Численные методы», «Пакеты прикладных программ», читаемые автором студентам Томского государственного университета систем управления и радиоэлектроники (ТУСУРа) и факультета прикладной математики и кибернетики Томского государственного университета (ТГУ).

1. ОСНОВНЫЕ ПОНЯТИЯ МОДЕЛИРОВАНИЯ СИСТЕМ

Подход к объектам исследования как к системам выражает одну из главных особенностей современной науки. Идея построения теории, применимой к системам любой природы, была выдвинута австрийским биологом Л. фон Берталанфи. Один из путей реализации этой идеи Берталанфи видел в определении структурных сходств законов, установленных в различных дисциплинах, их обобщении и выведении общесистемных закономерностей.

Вначале теорию систем активно развивали философы. Однако философская терминология не всегда легко применима к практической деятельности, поэтому потребности практики одновременно со становлением теории систем привели к возникновению направления, названного *исследованием операций*. При постановке и исследовании сложных проблем проектирования и управления широкое распространение получил термин *системотехника*. В связи с неточным переводом этот термин стал использоваться в основном в приложениях системных методов к техническим направлениям, а для других направлений был предложен термин *системология*. В настоящее время в дисциплинах, связанных с исследованием и проектированием сложных систем, используются термины *системные исследования* и, значительно реже, *системный подход*. Наиболее конструктивным из направлений системных исследований считается *системный анализ*, появившийся в 1948 году в связи с задачами военного управления. В отличие от других направлений системных исследований, в системном анализе предлагается *методика* проведения системных исследований и осуществляется работа с целями, т.е. их формулирование, исследование и структуризация. Таким образом, *системный анализ* — это методология исследования целенаправленных систем. При этом разработка методики, выбор методов и приемов выполнения базируются на использовании понятий и закономерностей теории систем.

Необходимость использования системного анализа объясняется следующими основными причинами:

- большинство традиционных научных дисциплин (биология, химия, физика и т.д.) в качестве предмета исследования рассматривают систему — совокупность элементов, находящихся в отношениях и связях друг с другом и представляющих собой единое целое;

- внедрение вычислительной техники и средств автоматизации привело к тому, что главными объектами научного исследования, технического проектирования и конструирования оказались сложные динамические системы.

1.1. Модель

Первоначально *моделью* называли некоторое вспомогательное средство, объект, который в определенной ситуации заменял другой. При этом далеко не сразу была понята универсальность законов природы, всеобщность моделирования, т.е. не просто возможность, но и необходимость представления любых знаний в виде моделей.

В XX веке понятие модели становится более общим, охватывающим как реальные, так и идеальные модели. При этом понятие абстрактной модели вышло за пределы математических моделей и стало относиться к любым знаниям и представлениям о мире. Споры вокруг такого широкого толкования понятия модели продолжаются и поныне. Таким образом, сначала в сфере научных дисциплин информационного, кибернетического, системного направления, а затем и в других областях науки модель стала осознаваться как нечто универсальное, хотя и реализуемое различными способами, т.е. *модель есть способ существования знаний*.

1.1.1. Свойства моделей

Для того чтобы модель отвечала своему назначению, недостаточно взять готовую или создать новую. Необходимо существование условий, обеспечивающих функционирование модели, отсутствие таких условий лишает ее модельных свойств. Например, программа для ЭВМ — это машинная модель алгоритма, программисты знают, что малейшее несогласование в программе с языком машины полностью обесценивает ее. Таким образом, для реализации модельных свойств необходимо, чтобы модель была согласована со средой, в которой ей предстоит функционировать, входила в эту среду не как чуждый элемент, а как естественная часть. Такое свойство согласованности с культурной средой обозначают термином *ингерентность* (от англ. *inherent* — внутренний, собственный, накрепко связанный, существующий как неотъемлемая часть).

Между моделью и отображаемой ею реальностью существуют как соответствия, так и различия.

Мир, частью которого мы являемся, бесконечен, как бесконечен и любой объект не только в пространстве и времени, но и в связях с другими объектами. Но ограничены как наши собственные ресурсы (число нервных клеток, время, которое мы можем затратить на реше-

ние какой-то задачи и т.д.), так и внешние ресурсы, которые мы можем вовлечь в конкретный процесс практической или теоретической деятельности. Таким образом, необходимо познавать бесконечный мир конечными средствами, и способ преодоления этого противоречия состоит в построении моделей. При этом необходимо обязательно учитывать, чем отличаются модели и реальные объекты и в каком смысле и до какой степени можно отождествлять модель с оригиналом.

Главные различия между моделью и действительностью — это конечность, упрощенность и приближенность модели.

Модель *конечна*, так как:

- она отображает оригинал лишь в конечном числе отношений;
- ресурсы моделирования конечны.

Модель всегда *упрощенно* отображает оригинал. Упрощенность характеризует качественные различия модели и оригинала. Причины упрощенности состоят в следующем:

- сама конечность моделей делает их упрощенность неизбежной;
- модель отображает только главные, наиболее существенные свойства в исследуемом явлении (идеальный газ, математический маятник, абсолютно черное тело и т.д.);
- ограничены средства оперирования с моделью (например, ресурсы ЭВМ).

Модель всегда *приближенно* (приблизительно) отображает действительность. Приближенность модели характеризует количественные различия между моделью и оригиналом (например, карта страны — приближенная модель страны).

Таким образом, различие между моделью и оригиналом вызвано тем, что мы можем отображать реальность лишь в конечном числе отношений, конечными средствами. В результате упрощение и приближенность модели необходимы и неизбежны. Различие само по себе не может быть ни большим, ни маленьким: оно либо есть, либо его нет. Величину, меру, степень приемлемости различия можно ввести, только соотнеся их с целью моделирования.

Соответствие модели оригиналу связано с понятием *адекватности* модели. Модель, с помощью которой успешно достигается поставленная цель, называется *адекватной этой цели*. В ряде случаев удастся ввести некоторую меру адекватности модели, т.е. указать способ сравнения двух моделей по степени успешности достижения цели. Если такой способ приводит к количественно выражаемой мере адекватности, то можно ставить задачи об *идентификации* модели (нахождении в заданном классе моделей наиболее адекватной), об исследовании *чувствительности* и *устойчивости* моделей (зависимости меры адекватности модели от ее точности), об *адаптации* моделей (подстройке параметров модели с целью повышения адекватности) и т.д.

Поскольку различия между моделью и реальностью неизбежны и неустранимы, существует предел истинности, правильности знаний, сконцентрированных в моделях. Об истинности, правильности или ложности модели самой по себе говорить бессмысленно, так как только в практическом соотношении модели с отображаемой ею реальностью выявляется степень истинности модели. Изменение условий, в которых ведется сравнение, весьма существенно влияет на результат: именно из-за этого возможно существование двух противоречивых, но «одинаково» истинных моделей одного объекта (волновая и корпускулярная модели света или электрона различны, противоположны и истинны каждая в своих условиях).

1.1.2. Классификации моделей

Существуют различные классификации моделей. Так как всякая деятельность направлена на достижение определенной цели, то и модель отображает не сам по себе объект-оригинал, а то, что соответствует поставленной цели. Это позволяет все разнообразное множество моделей разделить на основные типы — по типам целей. Поэтому модели можно разделить на познавательные и прагматические, что соответствует делению целей на теоретические и практические.

Познавательные модели являются формой организации и представления знаний, средством соединения новых и существующих знаний. Поэтому при обнаружении расхождения между моделью и реальностью решается задача устранения этого расхождения с помощью изменения модели. Таким образом, познавательная деятельность ориентирована в основном на приближение модели к отображаемой реальности.

Прагматические модели являются средством управления, средством организации практических действий, способом представления образцово правильных действий или их результата, т.е. являются рабочим представлением целей. Поэтому использование прагматических моделей состоит в изменении реальности для приближения ее к модели при обнаружении расхождения между ними. Таким образом, прагматические модели носят нормативный характер, играют роль стандарта, образца, под которые «подгоняются» и деятельность, и ее результат. Примерами прагматических моделей могут служить планы и программы действий, алгоритмы и т.д.

Основное различие между познавательными и прагматическими моделями состоит в том, что познавательные модели отражают существующее, а прагматические — не существующее, но желаемое и возможно осуществимое.

Другим принципом классификации целей моделирования может служить деление моделей на статические и динамические. *Статическими* называются модели, описывающие взаимосвязи между параметра-

ми, характеризующими объект в одно и то же время. *Динамическими* называются модели, отображающие процесс изменения состояния объекта во времени.

При построении моделей, сознательно создаваемых человеком, используются средства сознания и окружающего материального мира. В соответствии с этим модели делятся на абстрактные (идеальные) и материальные (реальные).

Абстрактные модели являются идеальными конструкциями, построенными средствами мышления, сознания. Очевидно, что к абстрактным моделям относятся все языковые конструкции. Но так как неоднозначность естественного языка мешает в некоторых видах практики, то были созданы более точные (профессиональные) языки, которые завершаются формализованным языком математики.

Модели, создаваемые из реальных объектов и процессов, называются *реальными* или *физическими*. Чтобы некоторая материальная конструкция могла быть отображением, т.е. замещала в каком-то смысле оригинал, между оригиналом и моделью должно быть установлено *отношение похожего подобия*. Существуют разные способы установления такого подобия.

Прямое подобие устанавливается в результате цепочки физических взаимодействий в процессе создания модели. Примерами таких отображений являются фотографии, масштабированные модели самолетов, кораблей, макеты зданий, куклы, выкройки и т.д.

Косвенное подобие между оригиналом и моделью объективно существует в природе и обнаруживается в виде совпадения или достаточной близости их абстрактных моделей. Например, подопытные животные у медиков — аналоги человеческого организма, автопилот — аналог летчика, электрический ток может моделировать транспортные потоки информации в сетях связи. Наиболее известным примером является электромеханическая аналогия, так как некоторые закономерности электрических и механических процессов описываются одинаковыми уравнениями, различие состоит лишь в разной физической интерпретации переменных, входящих в эти уравнения. Роль моделей, обладающих косвенным подобием оригиналу, очень велика.

Условное подобие устанавливается в результате соглашения. Примерами условного подобия служат деньги (модель стоимости), удостоверения личности (официальная модель владельца), географические карты (модели местности) и т.д. С моделями условного подобия приходится иметь дело очень часто, так как они являются способом материального воплощения абстрактных моделей, вещественной формой, в которой абстрактные модели могут передаваться от одного человека к другому, храниться до момента их использования. Это достигается с помощью соглашения о том, какое состояние реального объекта ставится в соответствие данному элементу абстрактной модели. Такое соглашение

принимает вид совокупности правил построения моделей условного подобия и пользования ими.

В некотором смысле промежуточное, посредническое положение занимают реальные модели, имеющие абстрактное содержание — *знаковые* модели. Это специфические модели условного подобия. Например, теория связи, теория информации, радиотехника, теория управления и ряд других наук имеют дело с такими моделями, которые получили название *сигналов*. В этих науках правила построения и способы использования сигналов, названные кодом, кодированием и декодированием, сами стали предметом исследований. В науках, изучающих создание и использование этих моделей самим человеком, у моделей условного подобия имеется своя специфика, специальные методы исследования и название — *знаки*. Возникшая в связи с этим область знаний получила название *семиотики*. Семиотика изучает знаки не в отдельности, а как входящие в знаковые системы. Языкознание, картография, криптография, графология, нумизматика, информатика занимаются исследованием моделей условного подобия.

Существуют и другие классификации моделей, например по характеру используемого материала, по цвету и т.д.

1.1.3. Основное определение модели

Поскольку модель является целевым отображением, отсюда следует множественность моделей одного и того же объекта. Для разных целей обычно требуются разные модели, поэтому различия в определениях модели неизбежны. На настоящий момент наиболее полное определение модели дали Ф.И. Перегудов и Ф.П. Тарасенко. *Модель есть отображение абстрактное или реальное; статическое или динамическое; инженерное; конечное, упрощенное, приближенное; имеющее, наряду с безусловно-истинным, условно-ложное и ложное содержание; проявляющееся и развивающееся в процессе его создания и практического использования. Это громоздкое определение можно заменить его кратким эквивалентом: модель есть системное отображение оригинала.*

1.2. Система

Термин *система* используют в тех случаях, когда хотят охарактеризовать исследуемый или проектируемый объект или процесс как нечто целое, сложное, о котором невозможно сразу дать представление, изобразив графически или описав математическим выражением. Описание системы является моделью данной системы, поэтому границы между понятиями «модель» и «система» являются условными.

1.2.1. Историческое развитие понятия «система»

Существует несколько десятков определений системы, которые изменялись по мере развития теории систем не только по форме, но и по содержанию. В первых определениях в той или иной форме говорилось о том, что система — это *элементы* (части, компоненты) и *связи* (отношения) между ними. Так, Л. фон Бенталанфи определил систему как совокупность элементов, находящихся в определенных отношениях друг с другом и со средой. Позднее для уточнения элементов и связей в определении включаются *свойства*. Затем в определении системы появляется понятие *цель*; вначале — в неявном виде, потом — в виде конечного результата, системообразующего критерия, функции, а позднее и с явным упоминанием о цели. В некоторых определениях уточняются условия целеобразования — *среда, интервал времени*, т.е. период, в рамках которого будут существовать система и ее цели. Так, в определении В.Н. Сагатовского система определяется как конечное множество функциональных элементов и отношений между ними, выделенное из среды в соответствии с определенной целью в рамках определенного временного интервала. Далее в определение системы начинают включать *наблюдателя* — лицо, представляющее объект или процесс в виде системы при исследовании или принятии решения. Первое определение системы, в которое в явном виде включен наблюдатель, дал Ю.И. Черняк: система есть отражение в сознании субъекта (исследователя, наблюдателя) свойств объекта и их отношений в решении задачи исследования, познания.

Любая деятельность человека носит целенаправленный характер. Цели, которые ставит перед собой человек, редко достижимы только за счет его собственных возможностей или внешних сил, имеющих у него в данный момент. Появляется *проблемная ситуация*. Проблемность существующего положения осознается поэтапно: от смутного ощущения, что «что-то не так», к осознанию потребности, затем к выявлению проблемы и, наконец, к формулировке цели. *Цель* — это субъективный образ (абстрактная модель) несуществующего, но желаемого состояния среды, которое решило бы возникшую проблему. Вся последующая деятельность, способствующая решению этой проблемы, направлена на достижение поставленной цели. Это действия, направленные на отбор из окружающей среды объектов, свойства которых можно использовать для достижения цели, и на объединение этих объектов надлежащим образом, т.е. это работа по созданию того, что называется системой. Таким образом, *система есть средство достижения цели*.

Многие авторы развивали определение системы до различной степени формализации. Определение — это языковая модель системы, и, следовательно, разная языковая среда, в силу ингерентности модели, также обуславливает видоизменение определений (уже только поэтому

отличаются, например, философское и математическое определения системы).

Таким образом, при проведении системного анализа необходимо прежде всего отобразить ситуацию с помощью как можно более полного определения системы, а затем, выделив наиболее существенные компоненты, сформулировать «рабочее» определение, которое может уточняться, расширяться или сужаться в зависимости от хода анализа.

1.2.2. Модель «черного ящика»

Приведенное выше определение системы ничего не говорит о ее внутреннем устройстве. Поэтому систему можно изобразить в виде непрозрачного «ящика», выделенного из окружающей среды, где *среда* — совокупность всех объектов, изменение свойств которых влияет на систему, и тех, чьи свойства меняются при взаимодействии с системой. Таким образом, уже эта максимально простая модель отражает важные свойства системы: *целостность* и *обособленность от среды*. Но хотя система обособлена, выделена из среды, она не является полностью изолированной, так как достигнутая цель — это запланированные заранее изменения в окружающей среде. Таким образом система связана со средой и с помощью этих связей воздействует на среду, эти связи называются *выходами* системы. Кроме того, так как система является средством достижения цели, то должны существовать и связи, обеспечивающие ее использование и воздействие на нее. Эти связи называются *входами* системы. Такая модель системы получила название «*черного ящика*», так как полностью отсутствуют сведения о внутреннем содержании системы, а перечисляются только входные и выходные связи системы со средой (рис. 1.1).

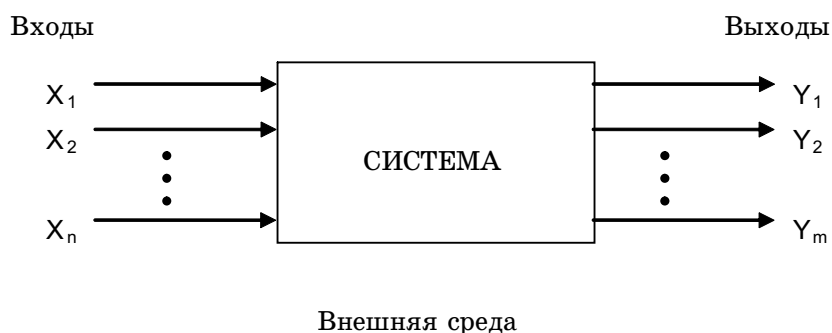


Рис. 1.1

Построение модели «черного ящика» не является тривиальной задачей, так как ответ на вопрос о том, сколько и какие именно входы

и выходы следует включать в модель, не прост и не всегда однозначен. Главной причиной множественности входов и выходов в модели «черного ящика» является то, что всякая реальная система, как и любой объект, взаимодействует с объектами окружающей среды неограниченным числом способов. Строя модель системы, нужно из бесчисленного множества связей отобрать конечное число для включения в список входов и выходов. Критерием отбора при этом является целевое назначение системы, существенность той или иной связи по отношению к поставленной цели. Однако тот факт, что исключаемые из рассмотрения связи признаны несущественными, не лишает их реальности, так как они все равно действуют на систему. И нередко оказывается, что казавшееся несущественным на самом деле является важным и должно быть обязательно учтено. Заметим, что модель «черного ящика» является не только полезной, но в ряде случаев единственно возможной при изучении систем (например, при исследовании психики человека).

1.2.3. Модель состава системы

При рассмотрении любой системы ее целостность и обособленность выступают как внешние свойства. Внутреннее содержание системы является неоднородным, состоящим из отдельных частей. Те части системы, которые представляются как неделимые, называются *элементами*, а те части, которые состоят из двух и более элементов, — *подсистемами*. Можно ввести обозначения, указывающие на иерархию частей системы (подподсистемы, подсистемы определенного уровня и т.д.). В результате получается *модель состава системы*, описывающая, из каких элементов и подсистем состоит система.

Построение модели состава системы является неоднозначным, так как:

- понятие элементарности можно определить по-разному, вследствие чего получаются разные модели состава;
- модель состава является целевой, поэтому для разных целей получаются разные модели состава одной и той же системы;
- всякое деление системы на элементы, подсистемы и т.д. является относительным и в определенной степени условным (например, тормозную систему автомобиля можно отнести либо к ходовой части, либо к подсистеме управления).

Таким образом, границы между подсистемами и между системой и окружающей средой являются условными, относительными. Модель состава системы ограничивается снизу тем, что считается элементом, а сверху — границей системы. При этом границы определяются целями построения модели системы.

1.2.4. Модель структуры системы

Так как система имеет целевую направленность, то необходимо установить между элементами системы определенные связи (отношения). Понятие *связь* входит в любое определение системы. Связь обеспечивает возникновение и сохранение в системе целостных свойств. Это понятие одновременно характеризует и строение (статику), и функционирование (динамику) системы.

Связь определяют как ограничение степени свободы элементов. Действительно, элементы, вступая во взаимодействие (связь) друг с другом, утрачивают часть свойств, которыми они потенциально обладали в свободном состоянии. Связь можно охарактеризовать:

- направлением — *направленные и ненаправленные*;
- силой — *сильные и слабые*;
- характером или видом — *связи подчинения, связи порождения или генетические, равноправные или безразличные, связи управления.*

Связи в конкретных системах могут быть одновременно охарактеризованы несколькими из названных признаков. Простейшими связями являются последовательное и параллельное соединения элементов и обратная связь, играющая важную роль в моделировании (рис. 1.2).

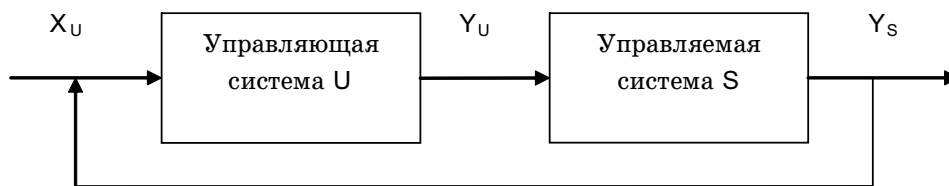


Рис. 1.2

Обратная связь может быть *положительной* — сохраняющей тенденции происходящих в системе изменений того или иного выходного параметра, и *отрицательной* — противодействующей тенденциям изменения выходного параметра, т.е. направленной на сохранение, стабилизацию требуемого значения параметра.

Обратная связь является основой саморегулирования, развития систем, приспособления их к изменяющимся условиям существования.

Совокупность необходимых и достаточных для достижения цели связей (отношений) между элементами системы называется структурой системы. Структура (от латинского structure — строение, расположение, порядок) отражает определенные взаимосвязи, взаиморасположение составных частей системы, ее устройство. При этом в сложных системах структура включает не все элементы и связи между ними, а лишь наиболее существенные, мало изменяющиеся и обеспечивающие существование системы и ее основных свойств.

Существуют различные виды и формы представления структур.

Сетевая структура, или *сеть*, представляет декомпозицию системы во времени (рис. 1.3).

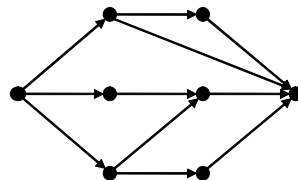


Рис. 1.3

Такие структуры могут отображать порядок действия технической системы (телефонная, электрическая сеть), этапы деятельности человека — сетевой график, сетевой план. При применении сетевых моделей пользуются определенной терминологией: вершина, ребро, путь, критический путь и т.д. Элементы сети могут быть расположены последовательно и параллельно. Сети бывают: однонаправленные, с обратными связями, с циклами. Для анализа сложных сетей используется математический аппарат теории графов, прикладная теория сетевого планирования и проектирования.

Иерархические структуры представляют собой декомпозицию системы в пространстве. Все компоненты (вершины, узлы) и связи (дуги, соединения узлов) существуют в этих структурах одновременно (не разнесены во времени). Такие структуры могут иметь достаточно большое число уровней декомпозиции. Структуры, в которых каждый элемент нижнего уровня подчинен одному узлу (вершине) верхнего уровня, называются *древовидными* структурами или *иерархическими* структурами с «сильными» связями (рис. 1.4). Структуры, в которых каждый элемент нижнего уровня подчинен нескольким узлам верхнего уровня, называются иерархическими структурами со «слабыми» связями (рис. 1.5).

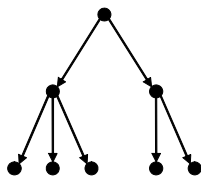


Рис. 1.4

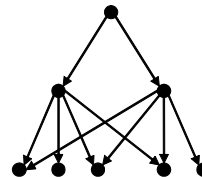


Рис. 1.5

Наибольшее распространение имеют древовидные иерархические структуры, с помощью которых представляются, например, конструкции сложных технических изделий и комплексов (рис. 1.6).

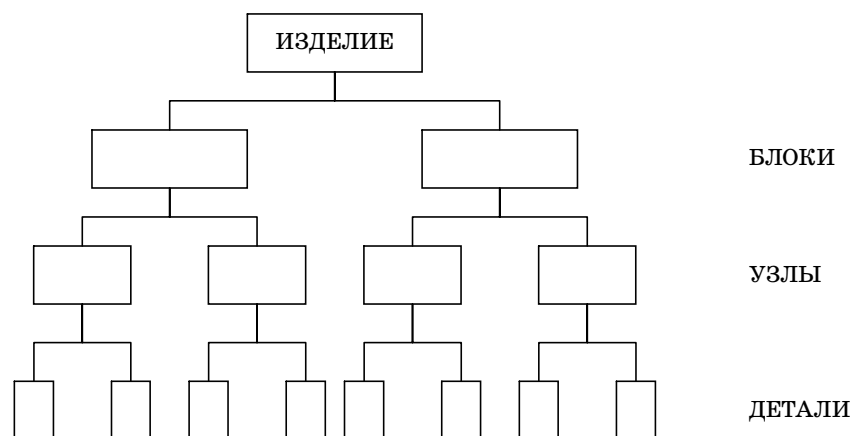


Рис. 1.6

Иерархии со «слабыми» связями применяют, например, для представления некоторых видов организационных структур.

В общем случае термин *иерархия* означает соподчиненность. Термин возник как наименование «служебной лестницы» в религии, широко применялся для характеристики взаимоотношений в аппарате управления государством, армией и т.д. Затем концепция иерархии была распространена на любой согласованный по подчиненности порядок объектов. Поэтому в иерархических структурах важно лишь выделение уровней соподчиненности, а между уровнями и компонентами в пределах уровня, в принципе, могут быть любые взаимоотношения. Существуют структуры, которые используют иерархический принцип, но имеют специфические особенности, так называемые *многоуровневые иерархические структуры*. Месаровичем были предложены особые классы иерархических структур типа *страт*, *слоев*, *эшелонов*, отличающиеся различными принципами взаимоотношений элементов в пределах уровня и различным правом вмешательства элементов высшего уровня в организацию взаимоотношений между элементами нижнего уровня.

При отображении сложных систем основная проблема состоит в том, чтобы найти компромисс между простотой описания и детализацией описания, позволяющей отразить многочисленные особенности конкретного объекта. Один из путей решения этой проблемы — задание системы семейством моделей, каждая из которых описывает поведение системы с точки зрения соответствующего уровня абстрагирования. Для каждого уровня существуют характерные особенности, законы и принципы, с помощью которых описывается поведение системы на этом уровне. Такое представление названо *стратифицированным*, а уровни абстрагирования — *стратами*.

П р и м е р

Отображение ЭВМ в виде двух страт (рис. 1.7): нижняя — физические операции; верхняя — математические и логические операции (программирование и реализация программ, информационные потоки).

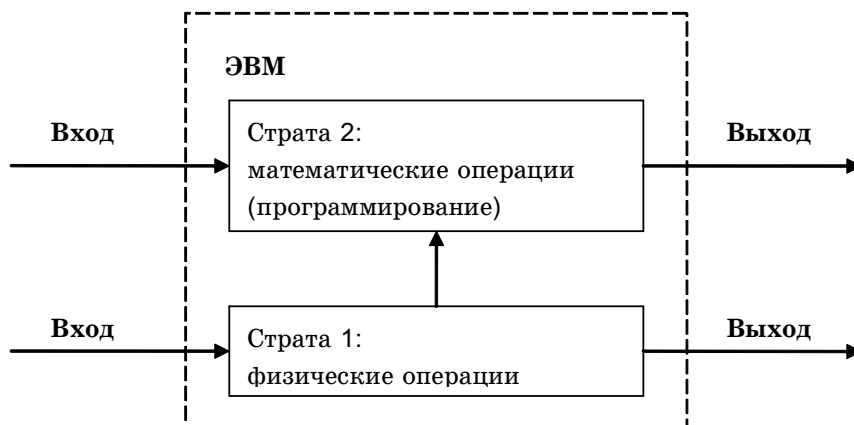


Рис. 1.7

Страты могут выделяться по разным принципам. Например, при представлении системы управления предприятием страты могут соответствовать сложившимся уровням управления: управление технологическими процессами и организационное управление. Если предприятие входит в объединение, то добавляется уровень управления объединением.

Второй вид многоуровневой структуризации — это *слои*. Для уменьшения неопределенности ситуации выделяются уровни сложности принимаемого решения — слои, т.е. определяется совокупность последовательно решаемых проблем. При этом их выделение осуществляется таким образом, чтобы решение проблем верхнего уровня определяло ограничения (допустимую степень упрощения) при моделировании на нижнем уровне, т.е. снижало неопределенность на нижнем уровне без утраты замысла решения общей проблемы.

Многоэшелонная иерархическая структура определяется следующим образом: система представляется в виде относительно независимых, взаимодействующих между собой подсистем, при этом некоторые или все подсистемы имеют права принятия решений, а иерархическое расположение подсистем (многоэшелонная структура) определяется тем, что некоторые из них находятся под влиянием или управляются вышестоящими. Уровень такой иерархии называют *эшелоном*.

Матричные структуры могут представлять взаимоотношения между уровнями иерархии. Например, матричное представление может быть

использовано при оформлении планов, поскольку помимо иерархической соподчиненности тематической основы плана в нем должны быть указаны исполнители, сроки выполнения, формы отчетности и т.д. Разновидность такого вида матричных представлений иерархических взаимоотношений используется в толковых словарях, в информационно-поисковых языках дескрипторного типа, в автоматизированных диалоговых процедурах анализа целей и функций, потому что при использовании таких диалоговых процедур первоначально неизвестно количество ветвей на каждом уровне.

Различные виды структур имеют специфические особенности и могут рассматриваться как самостоятельные понятия теории систем и системного анализа.

1.2.5. Структурная схема системы

При всем многообразии реальных систем принципиально различных типов моделей систем немного: модель типа «черный ящик», модель состава, модель структуры, а также их разумные сочетания и, прежде всего, объединение всех трех моделей — *структурная схема системы* («белый ящик», «прозрачный ящик»). В структурной схеме указываются все элементы системы, связи между элементами внутри системы и связи определенных элементов с окружающей средой (входы и выходы системы). Структурная схема является наиболее подробной и полной моделью любой системы, поэтому *систему можно определить как совокупность взаимосвязанных элементов, которая обособлена от среды и взаимодействует с ней как целое. Это второе определение системы.*

Все структурные схемы имеют много общего. Если абстрагироваться от содержательной стороны структурных схем, оставив только элементы и связи между ними, то такая схема называется *графом*. Кроме этого, структурная схема может быть представлена в матричной форме, в форме теоретико-множественных описаний и т.д.

1.2.6. Классификации систем

Существует много классификаций систем, позволяющих сравнивать и различать системы. Цель любой классификации — ограничить выбор подходов к отображению системы, сопоставить выделенным классам приемы и методы системного анализа и дать рекомендации по выбору методов моделирования для соответствующего класса систем. Система, в принципе, может быть охарактеризована несколькими признаками, т.е. ей может быть найдено место в разных классификациях, каждая из которых может оказаться полезной при выборе методов моделирования.

Наиболее известными являются следующие классификации систем: *по происхождению* (искусственные, естественные, смешанные); *по виду отображаемого объекта* (технические, экономические, биологические и т.д.); *по описанию переменных* (качественные, количественные, смешанное описание переменных); *по способу управления* (управляемые извне, самоуправляемые, с комбинированным управлением). Системы делят на *статические*, которые отображают поведение системы в фиксированный момент времени, и *динамические*, которые отображают поведение системы во времени, *открытые*, способные обмениваться со средой массой, энергией и информацией, и *закрытые*, или *замкнутые*, изолированные от среды. По степени ресурсной обеспеченности управления системы делятся на *обычные*, *малые*, *энергокритичные*, *простые*, *большие* и *сложные*.

Любая классификация, отражая объективные различия, носит целевой характер и является относительной, условной. Поэтому возможны различные классификации одного и того же множества систем.

Оценивая классификации с точки зрения их использования при выборе методов моделирования систем, следует отметить, что рекомендации, вплоть до выбора математических методов моделирования, имеются в них только для классов относительно низкой сложности, а для более сложных систем оговаривается, что дать такие рекомендации затруднительно.

Поэтому была предложена классификация, в которой делается попытка связать выбор методов моделирования со всеми классами систем. Основанием для этой классификации является *степень организованности*. В соответствии с этой классификацией выделяются три класса систем: *хорошо организованные*, *плохо организованные*, или *диффузные*, системы и *самоорганизующиеся* системы, которые включают классы саморегулирующихся, самообучающихся, самонастраивающихся и тому подобных систем.

Приведем краткую характеристику этих классов систем.

1. *Класс хорошо организованных систем*. Представление объекта или процесса в виде хорошо организованной системы возможно в случаях, когда удастся определить все элементы системы и их взаимосвязи между собой и с целями системы в виде детерминированных зависимостей. На представлении этим классом систем основано большинство моделей физических процессов и технических систем. Однако для сложных объектов формирование таких моделей является достаточно трудным делом. Для отображения сложного объекта в виде хорошо организованной системы приходится выделять существенные и не учитывать относительно несущественные для конкретной цели компоненты. При необходимости более детального описания нужно уточнить цель, указав с какой степенью глубины необходимо знание исследуемого объек-

та, и построить новую систему с учетом уточненной цели.

П р и м е р

При исследовании сложного радиоэлектронного устройства после предварительного его отображения с помощью обобщенной блок-схемы разрабатывают принципиальную схему, проводят соответствующие расчеты для определения номиналов элементов, реализующих необходимый режим функционирования, и т.д.

При представлении объекта в виде хорошо организованной системы задачи выбора целей и определения средств их достижения (элементов, связей) не разделяются. Проблемная ситуация может быть описана в виде выражений, связывающих цель со средствами (критериев функционирования, показателей эффективности, целевых функций и т.п.), которые могут быть представлены сложными математическими выражениями. При этом говорят, что цель представлена в виде критерия функционирования или эффективности.

Представление объекта в виде хорошо организованной системы применяется в тех случаях, когда может быть предложено детерминированное описание и экспериментально показана правомерность его применения, т.е. экспериментально доказана адекватность модели реальному объекту или процессу.

2. *Класс плохо организованных, или диффузных, систем.* При представлении объектов в виде систем этого класса не ставится задача определения всех компонентов и их связи с целями системы. Система характеризуется некоторым набором макропараметров и закономерностями, которые выявляются на основе исследования не всего объекта или класса явлений, а путем изучения выборки компонентов, характеризующих исследуемый объект или процесс. На основе этого, выборочного, исследования получают характеристики или закономерности (статистические, экономические и т.п.) и распространяют их на поведение системы в целом. При этом делаются соответствующие оговорки. Например, при получении статистических закономерностей их распространяют на поведение системы с определенной вероятностью, которая оценивается с помощью специальных приемов, изучаемых математической статистикой.

П р и м е р

При использовании газа для прикладных целей его свойства не определяют путем точного описания поведения каждой молекулы, а характеризуют макропараметрами — давлением, относительной проницаемостью, постоянной Больцмана и т.д. Основываясь на этих параметрах, разрабатывают приборы и устройства, использующие свойства газа.

3. *Класс самоорганизующихся систем.* Этот класс характеризуется рядом признаков и особенностей, приближающих рассматриваемые системы к реальным развивающимся системам. Эти особенности, как правило, обусловлены наличием в системе активных элементов и носят двойственный характер: они являются новыми полезными свойствами для существования системы, приспособляемости ее к изменяющимся условиям среды, но в то же время вызывают неопределенность и затрудняют управление системой.

К таким особенностям относятся:

- нестационарность (изменчивость, нестабильность) отдельных параметров и стохастичность поведения;
- непредсказуемость поведения системы в конкретных условиях, но в то же время наличие предельных возможностей, определяемых имеющимися ресурсами (элементами, их свойствами) и характерными для определенного типа систем структурными связями;
- способность адаптироваться к изменяющимся условиям среды и как внешним, так и внутренним помехам, что, казалось бы, является полезным свойством, но адаптивность может проявляться и по отношению к управляющим воздействиям, что затрудняет управление системой;
- способность противостоять энтропийным (разрушающим систему) тенденциям, обусловленная наличием активных элементов, стимулирующих обмен материальными, энергетическими и информационными продуктами со средой и проявляющих собственные «инициативы», благодаря чему в таких системах не выполняется закон возрастания энтропии и даже наблюдается противоположная тенденция, т.е. собственно самоорганизация, развитие;
- способность вырабатывать варианты поведения и (при необходимости) изменять свою структуру, сохраняя при этом целостность и основные свойства;
- способность и стремление к целеобразованию: в отличие от закрытых технических систем, которым цели задаются извне, в системах с активными элементами цели формируются внутри системы;
- неоднозначность использования понятий (например, таких как цель, средство, система, подсистема и т.п.).

Большинство из приведенных особенностей являются специфическими, отличающими этот класс систем от других, при этом существует принципиальная ограниченность формализованного описания развивающихся, самоорганизующихся систем. Поэтому для моделирования таких систем необходимо сочетание количественных и качественных методов.

Выделенные классы систем удобно использовать как подходы на начальном этапе решения любой задачи. Этим классам поставлены в

соответствие методы формализованного и неформализованного представления систем.

1.2.7. Основные понятия и определения

Пусть s — элемент системы; l — связь между элементами; S — система, т.е. совокупность элементов со связями и целью функционирования. При этом выделяют понятия *большая система* — это система, состоящая из большого числа однотипных элементов с однотипными связями, и *сложная система* — это система, состоящая из разнотипных элементов с разнотипными связями. Заметим, что термины «большая система» и «сложная система» еще окончательно не установились. Поэтому *системы, моделирование которых затруднительно вследствие их размерности, будем называть большими системами. Сложными системами будем называть системы, в моделях которых недостаточно информации для эффективного управления.*

Сложные системы встречаются в различных областях материального производства, научных исследований и повседневной жизни. К основным отличительным характеристикам сложных систем относятся следующие:

- наличие большого числа взаимосвязанных и взаимодействующих между собой элементов;
- наличие целевой функции, на оптимизацию которой направлено функционирование системы и которая не совпадает с целевыми функциями элементов, составляющих систему;
- наличие управления и разветвленной информационной сети, осуществляющей многочисленные связи системы как внутри ее элементарной базы, так и с внешней средой.

П р и м е р ы

Телефонная сеть города; информационная система; АСУ технологическим процессом; компьютерная сеть; система управления городским транспортом; вычислительная система; крупный морской или авиационный порт; робототехнический комплекс; отраслевая экономика и т.д.

В соответствии с введенными обозначениями под сложной системой S будем понимать тройку множеств

$$S = \{ \{s\}, \{l\}, F \},$$

где $\{s\}$ — множество элементов системы; $\{l\}$ — комплекс как взаимных связей между элементами, так и связей между элементами и внешней средой; F — функция, описывающая цель функционирования системы.

Процессы, происходящие в сложных системах, как правило, сразу не удастся представить в виде математических соотношений или алго-

ритмов. Поэтому для того чтобы охарактеризовать стабильность ситуации или ее изменения, используют специальные термины, заимствованные теорией систем из теории автоматического регулирования, биологии, философии.

Состояние — множество характеристик элементов системы. Его определяют либо через входные воздействия и выходные сигналы (результаты), либо через макропараметры, макросвойства системы (давление, скорость, ускорение).

Поведение — способность системы переходить из одного состояния в другое. Этим понятием пользуются, когда неизвестны закономерности (правила, алгоритмы, функции) перехода из одного состояния в другое.

Равновесие — способность системы в отсутствии внешних возмущающих воздействий или при постоянных воздействиях сохранять свое состояние сколь угодно долго. Это состояние называют состоянием равновесия.

Устойчивость — способность системы возвращаться в состояние равновесия, после того как она была из этого состояния выведена под влиянием внешних возмущающих воздействий. Состояние равновесия, в которое система способна возвращаться, называют устойчивым состоянием равновесия.

Процесс (динамика) — множество значений состояний системы, изменяющихся во времени.

Функционирование — процесс, который происходит в системе и окружающей ее среде, стабильно реализующий поставленную цель.

Цель функционирования системы — достижение желаемого состояния системы.

Заметим, что определение цели влечет:

- необходимость постановки локальных целей для элементов системы;
- целенаправленное вмешательство в процесс функционирования системы, называемое *управлением*.

Развитие — необратимое, направленное, закономерное изменение, в результате которого возникает новое качественное состояние объекта, его состава или структуры. Развитием называют то, что происходит с системой при изменении ее целей. Характерной чертой развития является тот факт, что существующая структура перестает соответствовать новой цели и для обеспечения новой функции системы приходится изменять структуру, а иногда и состав системы, т.е. перестраивать всю систему.

Исследование процесса развития, соотношения развития и устойчивости, исследование механизмов, лежащих в их основе, — наиболее

сложные задачи теории систем. При этом целесообразно выделить особый класс систем — *развивающиеся системы*, обладающие особыми свойствами и требующие использования специальных подходов к их моделированию.

1.2.8. Анализ и синтез систем

Среди множества задач исследования сложных систем можно выделить два больших класса:

- 1) задачи *анализа*, направленные на изучение свойств функционирования системы в зависимости от ее структуры;
- 2) задачи *синтеза*, связанные с выбором структуры и значений параметров по заданным свойствам системы.

Основной операцией анализа является представление целого в виде частей. При решении задач системных исследований объектами анализа являются как сами системы, так и цели, для достижения которых проводятся эти исследования. В результате анализа решаемые системой задачи разбиваются на подзадачи, системы — на подсистемы, цели — на подцели. Процесс разбиения продолжается до тех пор, пока объект анализа не будет представлен в виде совокупности элементарных компонентов. Необходимость таких действий обусловлена тем, что для отдельных подсистем объекта существенно проще предложить математическое описание, чем для всего объекта. Операция разложения целого на части называется *декомпозицией*. Цель декомпозиции — представить систему в виде иерархической структуры. Обычно декомпозицию проводят высококвалифицированные эксперты, имеющие богатый опыт работы в данной области.

Для каждой конкретной системы, каждой задачи и цели исследования существует некоторая разумная степень декомпозиции, переступать которую нецелесообразно. Это вызвано тем, что при слишком подробном представлении системы ее математическая модель будет содержать большое количество математических операторов, отражающих модели элементов, и при построении структуры исследуемой системы необходимо будет учесть большое количество взаимосвязей для этих элементов. Такая система может оказаться трудно реализуемой. Поэтому процесс декомпозиции целесообразно проводить путем постепенной детализации используемых моделей, т.е. *итеративно*.

Операцией, противоположной декомпозиции, является *агрегирование* — объединение частей в целое. Цель агрегирования — составление модели системы из моделей составляющих компонентов. Если декомпозиция осуществляется сверху вниз, то агрегирование идет снизу вверх.

Как и в случае с декомпозицией, техника агрегирования основана на использовании определенных моделей исследуемой системы. Разные постановки задач приводят к разным целям агрегирования и, сле-

довательно, к необходимости использования разных моделей. Таким образом, тип *агрегативной модели* определяется постановкой задачи и общей целью проводимого исследования, где под агрегативной моделью понимается результат выполнения операции агрегирования. В общем случае *агрегирование определяют как установление отношений на заданном множестве агрегатов*, где под агрегатом понимается некоторое абстрактное математическое описание моделей составляющих компонентов исследуемой системы.

Как правило, объектом системных исследований являются большие или сложные системы широкой прикладной направленности. Системный анализ применяется для решения задач исследования технических, социальных, природных систем, т.е. объектом анализа может быть и технический процесс, и экологическая ситуация обширной территории, и технико-экономическое развитие промышленного объекта, и социально-психологические исследования внутри коллектива. Естественно, что приходится наблюдать и описывать разнообразные процессы и структуры в ходе проведения исследований. Количество таких моделей многообразно и требует для своего описания применения разнообразных моделей. Однако, несмотря на многообразие описываемых процессов и структур, они могут быть классифицированы и представлены в виде ограниченного набора классов-агрегатов. Наиболее часто используются следующие виды агрегатов.

1. *Агрегаты-структуры*. Как и любой вид агрегата, структура является моделью системы и определяется совокупностью: объект, цель и средства моделирования. В результате получается многообразие типов структур: сетевые, древовидные, матричные и т.д. При синтезе создается структура будущей системы.

2. *Агрегаты-операторы*. Тип агрегата-оператора имеет место тогда, когда агрегируемые признаки фиксируются в числовых шкалах. В этом случае задается отношение на множестве признаков в виде числовой функции многих переменных, которая и является агрегатом. Основное применение агрегаты-операторы находят при описании динамических свойств системы. Представление зависимости выходных показателей системы в виде функционала от входных переменных есть пример агрегата-оператора.

3. *Агрегаты-статистики*. Процессы функционирования реальных сложных систем во многих случаях носят случайный характер. Выходные характеристики таких систем принимают случайные значения из множества величин, описываемых некоторой функцией распределения $F(\theta, t)$, где θ — вектор параметров закона распределения; t — некоторый момент времени. Если элементы вектора параметров функции распределения выражаются через достаточные статистики, тогда нет необходимости хранить всю информацию о реализованных характеристиках системы. Эту информацию можно заменить оценками парамет-

ров, которые получены по реализованным результатам наблюдений. Достаточные статистики — это агрегаты, которые извлекают всю полезную информацию об интересующем параметре из совокупности наблюдений. Примерами достаточных статистик являются параметры нормального закона распределения — математическое ожидание и дисперсия, параметр экспоненциального закона распределения — λ -характеристика. Использовать достаточные статистики необходимо с большой осторожностью. Их применение оправдано только в том случае, когда обоснован вид закона распределения, который описывает совокупность выходных величин. Дело в том, что агрегирование в данном случае является необратимым преобразованием, которое может привести к потере информации. Например, по сумме нельзя восстановить совокупность случайных величин слагаемых.

4. *Агрегат как случайный процесс.* Если процесс функционирования реальной сложной системы по своему существу носит характер случайного процесса, для агрегата как математической модели системы используются основные понятия теории случайных процессов. Например, случайный процесс, протекающий в любой физической системе, представляет собой случайные переходы системы из состояния в состояние. Состояние системы может быть охарактеризовано с помощью численных переменных: в простейшем случае — одной, а в более сложных — несколькими.

Таким образом, исследование сложных систем возможно только при использовании методов анализа и синтеза в совокупности. Они дополняют, а не заменяют друг друга.

1.2.9. Основные системные принципы

Понятие системы является основополагающим для всех направлений научной и практической деятельности. Оно базируется на следующих основных системных принципах:

- *целостность* — невозможность свести свойства системы к сумме свойств составляющих ее элементов и вывести из элементов свойства системы, а также зависимость каждого элемента системы от его места, функций и так далее внутри системы;
- *структурность* — возможность описания системы через установление структуры (связей и отношений) системы и обусловленность поведения системы не столько свойствами отдельных элементов, сколько свойствами ее структуры;
- *взаимозависимость системы и среды* — формирование и проявление свойств системы в процессе взаимодействия со средой;
- *многоуровневость и иерархичность* — представление самой системы

в виде элемента более общей системы и каждого ее элемента, в свою очередь, в виде системы;

- *множественность описания системы* — построение множества различных моделей исследуемой системы, каждая из которых соответствует определенной цели.

1.3. Моделирование

Так как любая деятельность человека направлена на достижение определенной цели, то важнейшим организующим элементом такой деятельности является цель — образ желаемого будущего, т.е. модель состояния, на реализацию которого и направлена деятельность. Системность деятельности проявляется в том, что она осуществляется по определенному плану или алгоритму. Следовательно, *алгоритм* — образ будущей деятельности, ее модель.

Как правило, деятельность редко осуществляется по жесткой программе, с учетом только начальных (априорных) данных и без учета происходящего на промежуточных этапах. Чаще приходится оценивать текущий результат предыдущих действий (апостериорные данные) и выбирать следующий шаг из числа возможных. Это означает, что приходится сравнивать последствия всех возможных шагов, не выполняя их реально, т.е. целесообразная деятельность невозможна без моделирования.

Моделирование — это исследование явлений, процессов или систем путем построения и изучения их моделей; использование моделей для определения или уточнения характеристик и рационализации способов построения вновь конструируемых объектов. Моделирование — одна из основных категорий теории познания. На идее моделирования, по существу, базируется любой метод научного исследования — как теоретический (при котором используются различного рода знаковые, абстрактные модели), так и экспериментальный (использующий физические модели).

1.3.1. Классификация методов представления систем

Постановка любой задачи заключается в том, чтобы перевести ее словесное, вербальное описание в формальное, т.е. получить математическую зависимость между величинами в виде формул. Если полученная формальная модель опирается на фундаментальный закон или подтверждается экспериментом, то этим доказывается ее адекватность отображаемой ситуации, и модель рекомендуется для решения задач соответствующего класса. По мере усложнения задач получение модели и доказательство ее адекватности усложняется. Постепенно эксперимент

становится дорогим и опасным, а то и практически нереализуемым, и постановка задачи, формирование модели становится составной частью процесса *принятия решения*.

П р и м е р

Пусть требуется доставить груз из пункта *A* в пункт *B*. В терминах проблемы принятия решений эту задачу можно представить следующим образом: задана цель — достичь пункта *B*; имеются средства — путь длиной *L* и транспорт,двигающийся со скоростью *v*. Требуется обеспечить реализацию цели.

Для решения задачи нужно определить взаимосвязи цели со средствами ее достижения. В данном случае в качестве выражения, связывающего цель со средствами, можно использовать закон равномерного прямолинейного движения, который имеет вид $t = L / v$ или в общем случае $t = f(L, v)$. Если не ставится никаких дополнительных условий, то задача считается решенной, так как получено выражение, связывающее цель со средствами ее достижения.

Если в задачу вводятся *ограничения*, например требование доставки груза к определенному времени и ограничение затрат на транспортировку, то для решения задачи формируется комплекс соотношений, включающий, наряду с выражением, которое связывает цель со средствами, неравенства или равенства, отражающие ограничения. Такая задача решается методами математического программирования.

Таким образом, если удастся получить выражения, связывающие цель со средствами ее достижения, то задача практически всегда решается, так как полученное формализованное представление задачи позволяет применять и формализованные методы ее решения.

В то же время при постановке задачи в числе ограничений могут быть и принципиально не формализуемые. Например, даже в рассматриваемом простейшем примере могут быть выдвинуты ограничения по безопасности транспортировки груза, по выбору транспорта в соответствии с видом груза и т.д. В этих случаях полностью формализованная постановка задачи оказывается невозможной.

Поэтому методы моделирования систем можно разделить на два больших класса: *методы формализованного представления систем* и *методы, направленные на активизацию использования интуиции и опыта специалистов*. При этом строгого разделения на формальные и неформальные методы не существует, и можно говорить только о большей или меньшей степени формализованности либо большей или меньшей опоре на интуицию. Такое разделение находится в соответствии с основной идеей системного анализа, которая состоит в сочетании формальных и неформальных представлений в моделях и методиках. В свою очередь, существуют классификации методов внутри этих классов.

1.3.2. Методы формализованного представления систем

Первоначально исследователи предлагали различные классификации формализованного представления систем и старались поставить им в соответствие определенные методы моделирования, позволяющие наилучшим образом отразить особенности того или иного класса. В большинстве первоначально применявшихся классификаций выделялись детерминированные и вероятностные методы, затем появились классификации, в которых выделялись теоретико-множественные представления, графы, математическая логика и некоторые новые разделы математики. Например, в классификации современного математического аппарата инженера В.П. Сигорский выделил: множества, матрицы, графы, логику, вероятности. В одной из первых классификаций, предложенных специально для целей системных исследований А.И. Кухтенко, выделяются следующие формализованные методы изучения сложных систем: общеалгебраический, теоретико-множественный, логико-лингвистический, информационный и эвристический.

Ф.Е. Темников предложил классификацию, в которой выделяются следующие обобщенные группы методов: аналитические; статистические; теоретико-множественные представления; логические представления; лингвистические представления; семиотические представления; графические.

Кроме того, в математике постоянно возникают новые направления на пересечении методов, отнесенных к обобщенным группам. В частности, на пересечении аналитических и теоретико-множественных представлений возникла и развивается *алгебра групп*; в рамках алгебры групп и теории множеств начала развиваться *комбинаторика*; теоретико-множественные и графические представления стали основой возникновения *топологии*; статистические и теоретико-множественные методы инициировали возникновение *теории размытых множеств*, которая, в свою очередь, явилась началом развития нового направления — *нечетких формализаций* и т.д.

Практически невозможно создать единую классификацию, которая включала бы все разделы современной математики, использующиеся для моделирования систем.

1.3.3. Методы неформализованного представления систем

Методы, направленные на активизацию использования интуиции и опыта специалистов, возникли и развивались как самостоятельные и для обобщения в теории систем. Вначале их называли *качественными* или *экспертными*, поскольку они представляют собой подходы, в той или иной форме активизирующие выявление и обобщение мнений экс-

пертов. Возникновение этих методов, как правило, связано с конкретными условиями проведения исследований. Самыми известными являются следующие методы:

- *методы типа мозговой атаки* или *коллективной генерации идей* основаны на гипотезе, что среди большого числа идей имеются по меньшей мере несколько хороших, которые нужно выявить;
- *методы типа сценариев* основаны на подготовке любых документов, содержащих анализ рассматриваемой проблемы и предложения по ее решению или по развитию системы независимо от формы представления;
- *методы структуризации* основаны на структурных представлениях разного рода, позволяющих разделить сложную проблему на более мелкие, лучше поддающиеся исследованию;
- *методы типа дерева целей* основаны на использовании иерархической структуры, получаемой путем расчленения общей цели на подцели нижележащих уровней, т.е. на построении структуры целей;
- *методы экспертных оценок* базируются на том, что неизвестная характеристика исследуемого явления трактуется как случайная величина, отражением закона распределения которой является индивидуальная оценка эксперта о достоверности и значимости того или иного события, при этом предполагается, что истинное значение исследуемой характеристики лежит внутри диапазона экспертных оценок и обобщенное коллективное мнение является достоверным;
- *методы типа Дельфи* основаны на итеративной процедуре повышения объективности экспертных опросов за счет использования количественных методов и обратной связи, т.е. осуществляется ознакомление экспертов с предыдущими результатами и учет этих результатов при оценке значимости мнений экспертов;
- *морфологические методы (метод Цвикки)* основаны на систематическом нахождении всех возможных вариантов решения поставленной проблемы или реализации системы путем комбинирования основных (выделенных исследователем) структурных элементов системы или их признаков;
- *методы организации сложных экспертиз* основаны на повышении объективности экспертных оценок путем расчленения большой первоначальной неопределенности проблемы на более мелкие, лучше поддающиеся осмыслению: способ усложненной экспертной процедуры, при котором выделяются группы критериев оценки и вводятся весовые коэффициенты, повышающие объективность результирующих оценок; метод решающих матриц, при котором осуществляется стратифицированное представление проблемы, в результате чего оценка относительной важности сложной альтернативы сводится к последовательности более частных альтернатив, и это обеспечивает при прочих равных условиях большую достоверность оценок.

Основную конструктивную идею моделирования при неформализованном представлении систем можно сформулировать следующим образом. Разрабатывается знаковая система, с помощью которой фиксируются известные на данный момент компоненты и связи. Затем, путем преобразования полученного отображения с помощью установленных правил (структуризации или декомпозиции, композиции и поиска мер близости) получают новые, неизвестные ранее компоненты, взаимоотношения, зависимости, которые могут либо послужить основой для принятия решений, либо подсказать последующие шаги на пути подготовки решения. Таким образом можно накапливать информацию об объекте, фиксируя при этом все новые компоненты и связи, и, применяя их, получать отображения последовательных состояний системы, постепенно создавая все более адекватную модель реального, изучаемого или создаваемого объекта. При этом информация может поступать от специалистов различных областей знаний и накапливаться по мере поступления. Адекватность модели также доказывается последовательно, путем оценки правильности отражения в каждой последующей модели компонентов и связей, необходимых для достижения поставленных целей. Такое моделирование становится своеобразным механизмом развития системы. Практическая реализация такого механизма связана с необходимостью разработки языка моделирования процесса принятия решения. В основу такого языка (языковой системы) может быть положен один из методов моделирования систем, например теоретико-множественные представления, математическая логика, математическая лингвистика, имитационное моделирование и т.д.

1.3.4. Классификация видов моделирования

Все классификации носят целевой характер. Они позволяют выделить и структурировать интересующие свойства и характеристики систем. При этом любая классификация условна, она является лишь средством, помогающим ориентироваться в огромном числе разнообразных методов моделирования. Например, классификация систем по виду использованной модели и способу решения поставленной задачи выглядит следующим образом (рис. 1.8).

Физическое моделирование предполагает, что в качестве модели используется либо сама исследуемая система, либо другая система с той же или подобной физической природой. Примером физического моделирования является исследование свойств летательных аппаратов в аэродинамической трубе.

Под *математическим моделированием* понимается процесс установления соответствия данной реальной системе S некоторой математической

модели и исследование этой модели, позволяющее получить характеристики реальной системы.

Для *аналитического моделирования* характерно то, что процессы функционирования элементов системы записываются в виде некоторых математических соотношений или логических условий. Аналитическая модель может быть исследована следующими методами:

- аналитическим, когда стремятся получить в общем виде явные зависимости для искомых характеристик;
- численным;
- качественным, когда, не имея решений в явном виде, можно найти некоторые свойства решения.

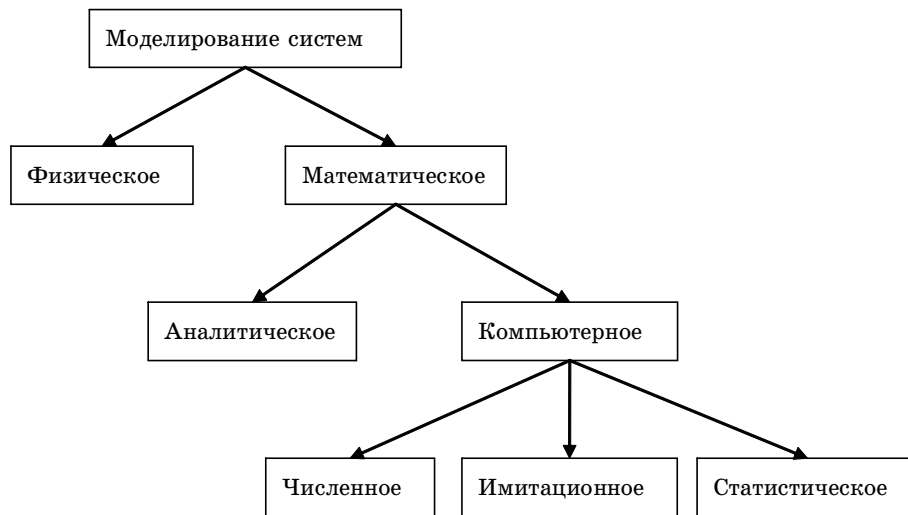


Рис. 1.8

Для *компьютерного моделирования* характерно то, что математическая модель системы представлена в виде программы для ЭВМ — компьютерной модели, позволяющей проводить с ней вычислительные эксперименты. В зависимости от математического аппарата, используемого при построении модели, и способа организации вычислительных экспериментов можно выделить три взаимосвязанных вида компьютерного моделирования:

- численное,
- статистическое;
- имитационное.

При *численном моделировании* для построения компьютерной модели

используются методы вычислительной математики, а вычислительный эксперимент заключается в численном решении некоторых математических уравнений.

Статистическое моделирование — вид компьютерного моделирования, позволяющий получать статистические данные о процессах в моделируемой системе S .

Имитационное моделирование — вид компьютерного моделирования, для которого характерно воспроизведение на ЭВМ процесса функционирования исследуемой сложной системы. При этом имитируются элементарные явления, составляющие процесс, с сохранением их логической структуры, последовательности протекания во времени, что позволяет получить информацию о состоянии системы S в заданные моменты времени.

2. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Основной целью рациональной деятельности человека является создание высокоэффективных систем в экономике, технике, экологии, производстве, обществе. Высокая эффективность предполагает достижение экстремума некоторых числовых характеристик — показателей эффективности. В связи с этим возникает необходимость в использовании математических методов, но к реальному явлению, объекту эти методы непосредственно не применимы. Поэтому необходимо, прежде всего, построить *математическую модель*, т.е. приближенно описать реальное явление или объект с помощью математических соотношений. Математическое моделирование представляет собой метод исследования реальных явлений или процессов путем построения их математических моделей и исследования этих моделей.

Математическое моделирование существует фактически с тех времен, когда математический аппарат начинал применяться для решения практических задач. Однако эти попытки часто заходили в тупик как из-за сложности самих моделей (нелинейные уравнения; большое число переменных, параметров, уравнений; недостаток априорной информации и др.), так и невозможности их исследования традиционными аналитическими методами. С появлением и совершенствованием ЭВМ ситуация коренным образом изменилась, появились новые понятия и подходы. В настоящее время *сущность математического моделирования состоит в замене исходной системы ее математической моделью и дальнейшем исследовании этой модели на ЭВМ.*

2.1. Цели и задачи математического моделирования

Метод математического моделирования, сводящий исследование реальных объектов и процессов к решению математических задач, занимает ведущее место среди других методов исследования, особенно с появлением ЭВМ. Он позволяет проектировать новые технические средства, работающие в оптимальных режимах, используется для решения сложных задач науки и техники, является необходимым при прогнозировании экономических ситуаций. Математические модели применяются в самых различных областях знания, они являются важным элементом автоматизированных систем управления.

Математическое моделирование позволяет достичь следующих целей:

- 1) получить новые или уточнить существующие знания о реальных процессах и объектах;
- 2) соединить достоинства теоретических и экспериментальных методов, т.е. математическую модель и эксперимент на ЭВМ;
- 3) гарантировать высокую эффективность применения вычислительной техники и основанных на ней средств информатики и автоматизации: САПР, АРМ, АСУ, ГАП, экспертных систем;
- 4) исследовать такие системы, натурное и физическое моделирование которых экономически не оправдано или трудноосуществимо (требуется уникальное дорогостоящее оборудование; эксперименты с такими системами, как АСУ предприятием, отраслью и т.п.);
- 5) исследовать системы, натурное и физическое моделирование которых связано с опасными для здоровья человека условиями (высокой радиацией, токсичностью и т.п.);
- 6) исследовать не существующие еще в реальности перспективные системы на стадии их проектирования;
- 7) исследовать труднодоступные объекты — удаленные в пространстве и во времени (моделирование атмосферы планет, процесса зарождения Вселенной и т.п.);
- 8) исследовать объекты, не наблюдаемые вследствие их размера и длительности существования (модели макро- и микромира);
- 9) исследовать экономические, социальные, экологические и биологические системы.

2.2. Этапы построения математических моделей

Математическая модель — это приближенное описание какого-либо класса явлений внешнего мира, выраженное с помощью математической символики. Анализ математических моделей позволяет проникнуть в суть изучаемого явления, осуществить прогнозирование и управление. Процесс математического моделирования, т.е. изучения рассматриваемого явления с помощью математических моделей, можно разделить на четыре этапа.

Первый этап — формулирование законов, связывающих основные объекты модели. Этот этап требует знания фактов, относящихся к изучаемым явлениям, и глубокого проникновения в их взаимосвязи. Этап завершается записью в математической форме сформулированных качественных представлений о связях между объектами модели.

Второй этап — исследование математических задач, к которым приводят математические модели рассматриваемого явления. Основным вопросом здесь является решение *прямой задачи*, т.е. получение в результате анализа модели выходных данных для дальнейшего их

сопоставления с результатами наблюдений изучаемых явлений. На этом этапе важную роль играют математический аппарат, необходимый для анализа математических моделей, и вычислительная техника, которая является средством для получения количественной выходной информации при решении сложных математических задач.

Третий этап — выяснение, является ли построенная модель адекватной рассматриваемому явлению, т.е. согласуются ли результаты решения математических задач с результатами наблюдений рассматриваемого явления в пределах принятой точности наблюдений. Если все параметры модели определены и отклонения математических решений от наблюдений рассматриваемого явления находятся в пределах точности наблюдений, то математическая модель является адекватной, т.е. принимается для описания рассматриваемого явления. Если отклонения математических решений от наблюдений выходят за пределы точности наблюдений, то математическая модель является неадекватной и не может быть принята для описания рассматриваемого явления. Часто при построении модели некоторые ее характеристики остаются неопределенными. Задачи, в которых определяются характеристики модели (параметрические, функциональные) таким образом, чтобы результаты математического решения совпадали с наблюдениями рассматриваемого явления в пределах точности наблюдений, называются *обратными задачами*. Если математическая модель такова, что ни при каком выборе ее характеристик эти условия не удовлетворяются, то модель непригодна для исследования рассматриваемых явлений.

Четвертый этап — анализ модели в связи с накоплением данных об изучаемых явлениях и модификация модели. В процессе развития науки и техники данные об изучаемых явлениях все более и более уточняются, и наступает момент, когда выводы, получаемые на основании принятой математической модели, не соответствуют знаниям о рассматриваемом явлении. Таким образом, возникает необходимость в построении новой или модификации существующей математической модели. Может возникнуть ситуация, когда априорные знания (гипотезы) об изучаемых явлениях оказались неверными, и тогда требуется пересмотреть существующие знания и выдвинуть новые гипотезы.

2.3. Основные характеристики математических моделей

Пусть $[t_0, T]$ — временной интервал моделирования системы S (интервал модельного времени), где t_0 — время начала моделирования (достаточно часто полагают $t_0 = 0$); T — время окончания моделирования; $t \in [t_0, T]$ — текущее значение модельного времени.

Построение математической модели системы S начинается с определения параметров и переменных, характеризующих процесс функционирования системы.

Параметры системы $\theta_1, \theta_2, \dots, \theta_m$ — это характеристики системы, остающиеся постоянными на всем интервале моделирования $[t_0, T]$.

Если значения $\{\theta_i\}$ определены на некотором множестве Θ , т.е.

$$\theta = \begin{pmatrix} \theta_1 \\ \dots \\ \theta_m \end{pmatrix} \in \Theta \subset R^m,$$

то говорят, что имеется параметрическое семейство систем.

Множество переменных разбивается на два подмножества — зависимых и независимых переменных.

К *независимым* переменным относятся следующие характеристики.

1. *Входные воздействия (сигналы)* u_1, u_2, \dots, u_{n_1} . Входные воздействия в момент $t \in [t_0, T]$ характеризуются вектором

$$u = u(t) = \begin{pmatrix} u_1(t) \\ \dots \\ u_{n_1}(t) \end{pmatrix} \in U \subset R^{n_1}.$$

Среди $\{u_i\}$ могут быть *управляющие* воздействия, например $u_1, u_2, \dots, u_{n'_1}$, ($n'_1 \leq n_1$), а остальные $(n_1 - n'_1)$ воздействия — *неуправляющие*.

2. *Воздействия внешней среды* v_1, v_2, \dots, v_{n_2} . Среди них могут быть контролируемые и неконтролируемые, наблюдаемые и ненаблюдаемые, детерминированные и случайные воздействия. В момент $t \in [t_0, T]$ они характеризуются вектором

$$v = v(t) = \begin{pmatrix} v_1(t) \\ \dots \\ v_{n_2}(t) \end{pmatrix} \in V \subset R^{n_2}.$$

3. *Переменные, характеризующие состояние системы* x_1, x_2, \dots, x_{n_3} . В отличие от $\{\theta_i\}$, состояния $\{x_i\}$ характеризуют свойства системы, изменяющиеся во времени. Состояние системы в момент $t \in [t_0, T]$ описывается вектором

$$x = x(t) = \begin{pmatrix} x_1(t) \\ \dots \\ x_{n_3}(t) \end{pmatrix} \in X \subset R^{n_3},$$

где X — пространство состояний, или фазовое пространство системы (множество возможных значений вектора x). Если $t_1 < t_2 < \dots$ — моменты изменения состояния системы, то последовательность $x(t_1), x(t_2), \dots$ называется *фазовой траекторией* системы.

К *зависимым* переменным относятся следующие характеристики.

1. Выходные характеристики (сигналы) системы y_1, y_2, \dots, y_{n_4} , определяемые в момент $t \in [t_0, T]$ вектором

$$y = y(t) = \begin{pmatrix} y_1(t) \\ \dots \\ y_{n_4}(t) \end{pmatrix} \in Y \subset R^{n_4}.$$

Связи между зависимыми и независимыми переменными изображены на рис. 2.1.

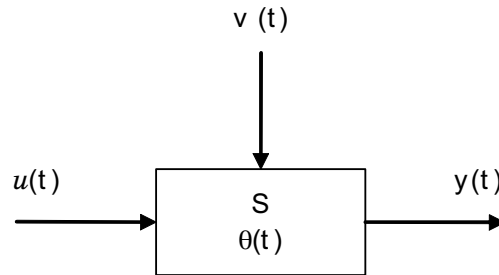


Рис. 2.1

2. Показатели эффективности функционирования системы $\omega_1, \omega_2, \dots, \omega_k$, которые характеризуют степень достижения системой ее цели (т.е. характеризуют качество функционирования системы) и в момент $t \in [t_0, T]$ образуют вектор

$$\omega = \omega(t) = \begin{pmatrix} \omega_1(t) \\ \dots \\ \omega_k(t) \end{pmatrix} \in W \subset R^k.$$

Показатели эффективности представляются в виде некоторых функционалов

$$\omega_i = \omega_i(u(t), v(t), x(t), y(t), \theta, t), \quad t \in [t_0, T], \quad i = \overline{1, k}.$$

При наличии в системе случайных факторов, например случайных воздействий внешней среды, значения $\{\omega_i\}$ являются также случайными и поэтому не могут служить показателями эффективности. В этих случаях в качестве показателей эффективности используют средние значения $\{W_i\}$, определяемые соотношениями

$$W_i = M \{ \omega_i \}, \quad i = \overline{1, k},$$

где $M \{ \}$ — символ математического ожидания.

Процесс функционирования системы во времени описывается операторными соотношениями, заданными аналитически или алгоритмически, для состояний, выходных характеристик и показателей эффективности системы

$$\left. \begin{aligned} x(t) &= F_1(u^{(t)}, v^{(t)}, \theta, t); \\ y(t) &= F_2(u^{(t)}, v^{(t)}, x^{(t)}, \theta, t); \\ \omega(t) &= F_3(u^{(t)}, v^{(t)}, x^{(t)}, y^{(t)}, \theta, t), \end{aligned} \right\} \quad (2.1)$$

где $u^{(t)}$ обозначает реализацию процесса $u(t)$ на отрезке $[0, t]$, аналогично обозначим $x^{(t)}, y^{(t)}$. Через $F_1(\cdot)$, $F_2(\cdot)$, $F_3(\cdot)$ обозначены соответствующие операторы, описывающие динамику зависимых и независимых переменных и показателей эффективности.

Зависимости (2.1) называются законами функционирования системы S ; зависимость $y = y(t)$, $t \in [t_0, T]$, называется выходной траекторией системы, а зависимость $x = x(t)$, $t \in [t_0, T]$, — фазовой траекторией.

В выборе переменных $x(t)$, характеризующих состояние системы в момент времени $t \in [t_0, T]$, обычно имеется произвол, который используется так, чтобы упростить закон функционирования (2.1) и привести его к виду

$$\left. \begin{aligned} x(t) &= F(x^{(0)}, u^{(t)}, v^{(t)}, \theta, t), \\ y(t) &= G_1(x^{(t)}, t), \\ \omega(t) &= G_2(x^{(t)}, t), \end{aligned} \right\} \quad (2.2)$$

где $F(\cdot)$, $G_1(\cdot)$, $G_2(\cdot)$ — некоторые операторы; $x^{(0)}$ — начальное состояние системы; $t \in [t_0, T]$.

Закон (2.2) отличается от (2.1) следующими особенностями:

- состояние системы S в момент времени t зависит от начального состояния системы $x^{(0)}$;
- выходные характеристики и показатели эффективности системы в момент времени t зависят от состояний $x^{(t)}$ и текущего времени.

Таким образом, математической моделью системы является множество переменных $u, v, \theta, x, y, \omega$ вместе с законом функционирования в виде (2.1) или (2.2).

2.4. Классификация математических моделей

Операторные соотношения в (2.1), (2.2) могут быть заданы аналитически, т.е. с помощью функциональных соотношений или логических условий, или алгоритмически. В зависимости от способа задания закона функционирования математические модели делятся на *аналитические* и *алгоритмические*.

Заметим, что время t может рассматриваться и как непрерывная переменная $t \in [t_0, T]$, и как дискретная

$$t = j\Delta, \quad j = 0, 1, \dots, N, \quad N = [T / \Delta],$$

где Δ — шаг дискретизации. В соответствии с этим имеем *непрерывные* и *дискретные* математические модели. Если математическая модель не содержит случайных элементов, то имеем *детерминированную* модель, в противном случае — *вероятностную* модель.

Таким образом, по признакам непрерывности и стохастичности можно выделить четыре обширных класса математических моделей: *непрерывно-детерминированные* модели; *дискретно-детерминированные*; *непрерывно-вероятностные* и *дискретно-вероятностные*.

2.5. Формы представления математических моделей

2.5.1. Непрерывно-детерминированные модели

В этих моделях время $t \in [0, T]$ полагается непрерывной переменной, а случайными факторами в системе пренебрегают. Основной математический аппарат, используемый при построении и исследовании непрерывно-детерминированных моделей, — это теория дифференциальных и интегральных уравнений. Рассмотрим особенности непрерывно-детерминированного моделирования на примерах использования дифференциальных уравнений.

П р и м е р

Пусть система S описывает электрический колебательный контур с двумя параметрами: емкостью θ_1 и индуктивностью $1/\theta_2$. Состояние этой системы в момент t будем характеризовать величиной заряда конденсатора $x(t)$, которая измеряется прибором с некоторой погрешностью $v(t) \in [-\varepsilon, \varepsilon]$. Тогда, как известно из физики, математическая модель такой системы будет иметь вид

$$\theta_2 \frac{\partial^2 x(t)}{\partial t^2} + \frac{1}{\theta_1} x(t) = u(t), \quad y(t) = x(t) + v(t), \quad (2.3)$$

где $u(t)$ — внешнее детерминированное электрическое воздействие на контур. Если $u(t) = 0$, $x(0) = A \cos(\phi)$, то решение однородного уравнения (2.3) имеет вид

$$x(t) = A \cos\left(\frac{1}{\sqrt{\theta_1 \theta_2}} t + \phi\right),$$

где A — амплитуда, ϕ — фаза колебаний.

При построении непрерывно-детерминированных моделей систем автоматического управления приходится использовать дифференциальные уравнения высших порядков.

П р и м е р

Пусть S — одноканальная система автоматического управления, структура которой представлена на рис. 2.2.

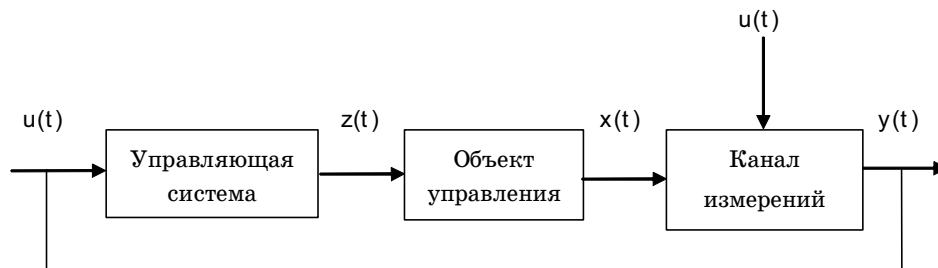


Рис. 2.2

Объект управления описывается дифференциальным уравнением k -го порядка

$$\Psi\left(\frac{\partial^k x(t)}{\partial t^k}, \dots, \frac{\partial x(t)}{\partial t}, x(t), \frac{\partial^l z(t)}{\partial t^l}, \dots, \frac{\partial z(t)}{\partial t}, z(t), \theta, t\right) = 0, \quad (2.4)$$

где $\Psi(\cdot)$ — некоторая функция от $(k+l+m+3)$ переменных; $x(t)$ — состояние объекта управления, измеряемое с некоторой ε — погрешностью в канале измерения:

$$y(t) = x(t) + v(t), \quad |v(t)| \leq \varepsilon, \quad (2.5)$$

$z(t)$ — воздействующий на объект управления сигнал, формируемый управляющей системой по заданному алгоритму

$$z(t) = \Phi(u(t), y(t)). \quad (2.6)$$

Соотношения (2.4)–(2.6) определяют непрерывно-детерминированную модель.

Нередко используется частный случай (2.4) — линейное дифференциальное уравнение, коэффициенты которого являются параметрами θ модели:

$$\theta_1 \frac{\partial^k x(t)}{\partial t^k} + \dots + \theta_k \frac{\partial x(t)}{\partial t} + \theta_{k+1} x(t) + \theta_{k+2} \frac{\partial^l z(t)}{\partial t^l} + \dots + \theta_{k+l+2} z(t) = 0. \quad (2.7)$$

В прикладных задачах переноса излучения (например, при расчете ядерных реакторов) при построении непрерывно-детерминированной модели приходится использовать дифференциальные уравнения в частных производных и интегральные уравнения.

2.5.2. Дискретно-детерминированные модели

В отличие от непрерывно-детерминированных, в этих моделях время t является дискретной переменной: $t = j\Delta$, где Δ — шаг дискретизации; $j = 0, 1, 2, \dots$ — дискретное время.

Основной математический аппарат, используемый при построении дискретно-детерминированных моделей, — это теория разностных уравнений и аппарат дискретной математики, в частности теория конечных автоматов.

Разностное уравнение — это уравнение, содержащее конечные разности искомой функции

$$\Phi(x_j, x_{j+1}, \dots, x_{j+n}, u_j, u_{j+1}, \dots, u_{j+m}, \theta, j) = 0, \quad (2.8)$$

где $x_j = x(j\Delta)$, $u_j = u(j\Delta)$ — соответственно состояние системы и внешнее воздействие в дискретный момент времени j .

Отметим, что дискретно-детерминированные модели в виде (2.8) часто возникают как промежуточные при исследовании непрерывно-детерминированных моделей на ЭВМ, когда аналитическое решение дифференциального уравнения получить не удастся и приходится применять разностные схемы.

П р и м е р

Рассмотрим $x(t)$ — величину заряда конденсатора из (2.3), $i(t)$ — величину силы тока в дискретные моменты времени $t = j\Delta$. Для разностного приближения величины $i(t)$ воспользуемся формулой

$$i_j = \frac{x_{j+1} - x_j}{\Delta}.$$

Дискретную аппроксимацию уравнения $\frac{\partial i(t)}{\partial t} = \frac{\partial^2 x(t)}{\partial t^2}$ осуществим

двумя следующими способами:

$$\frac{\partial^2 x(t)}{\partial t^2} \approx \frac{i_j - i_{j-1}}{\Delta} = \frac{x_{j+1} - 2x_j + x_{j-1}}{\Delta^2}; \quad (2.9)$$

$$\frac{\partial^2 x(t)}{dt^2} \approx \frac{i_{j+1} - i_j}{\Delta} = \frac{x_{j+2} - 2x_{j+1} + x_j}{\Delta^2}. \quad (2.10)$$

Тогда однородную часть математической модели (2.3) можно описать двумя разностными уравнениями в соответствии с (2.9) и (2.10). Тогда

$$x_{j+1} - \left(2 - \frac{\Delta^2}{\theta_1 \theta_2}\right) x_j + x_{j-1} = 0; \quad (2.11)$$

$$x_{j+2} - 2x_{j+1} + \left(1 + \frac{\Delta^2}{\theta_1 \theta_2}\right) x_j = 0. \quad (2.12)$$

Решая разностные уравнения при заданных начальных значениях x_0, x_1 и x_{-1} , получим фазовые траектории $\{x_j\}, j = 1, 2, \dots$. Анализ устойчивости разностных уравнений позволяет заключить, что амплитуда колебаний фазовой траектории при использовании уравнения (2.12) возрастает со скоростью, сравнимой с ростом функции

$$\left(1 + \frac{\Delta^2}{\theta_1 \theta_2}\right)^{j/2},$$

что не соответствует предположению об ограниченности амплитуды непрерывной модели (2.3). Поэтому можно сделать вывод о неадекватности разностного уравнения (2.12) аппроксимируемой модели (2.3). Разностное уравнение (2.11) является адекватным модели (2.3). Следовательно, исследование устойчивости разностных схем необходимо при аппроксимации непрерывных моделей дискретными.

Другой математический аппарат для построения дискретно-детерминированных моделей представляет теория конечных автоматов. *Конечный автомат* — это математическая модель дискретной системы S , которая под действием входных сигналов $u \in U$ вырабатывает выходные сигналы $y \in Y$ и может иметь некоторые изменяемые внутренние состояния $x \in X$; при этом U, Y, X — конечные множества.

Конечный автомат S характеризуется шестью элементами:

- входным алфавитом U ;
- выходным алфавитом Y ;
- внутренним алфавитом состояний X ;
- начальным состоянием $x_0 \in X$;
- функцией переходов $x' = \Phi(x, u): XU \rightarrow X$;
- функцией выходов $y = \psi(x, u): XU \rightarrow Y$.

Процесс функционирования конечного автомата S состоит в следующем. На j -м такте ($j = 0, 1, \dots$) на вход автомата, находящегося в состоянии $x_j \in X$, поступает входной сигнал $u_j \in U$, на который автомат

реагирует переходом на $(j+1)$ -м такте в состояние $x_{j+1} \in X$ и выдачей выходного сигнала $y_j \in Y$.

Например, *конечный автомат Мили* описывается следующими рекуррентными соотношениями:

$$\left. \begin{aligned} x_{j+1} &= \Phi(x_j, u_j), \quad x(0) = x_0, \\ y_j &= \psi(x_j, u_j), \quad j = 0, 1, \dots \end{aligned} \right\} \quad (2.13)$$

На рис. 2.3 изображен граф, описывающий булевский автомат Мили с $x_0 = 0$ и алфавитами $U = Y = \{0, 1\}$, $X = \{0, 1, 2\}$. Вершины графа соответствуют внутренним состояниям автомата, а дуги — соответствующим переходам. Каждая дуга из x в x' помечается парой (u, y) , где u, y — входной и выходной сигналы, при которых осуществляется переход из x в x' .

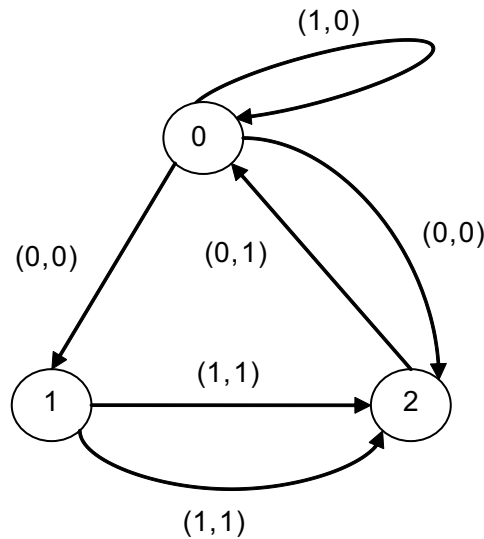


Рис. 2.3

Отметим, что конечные автоматы являются адекватной математической моделью элементов и узлов ЭВМ, устройств контроля, регулирования, управления систем коммутации в технике передачи данных.

Современное развитие дискретно-детерминированных моделей нашло отражение в теории *сетей Петри*. Математический формализм этого аппарата моделирования основывается на *теории комплектов*, обобщающей теорию множеств. Комплект, подобно множеству, есть набор элементов из некоторой области X , отличие лишь в том, что допускается присутствие в X одного и того же элемента несколько раз.

2.5.3. Дискретно-вероятностные модели

Главное отличие дискретно-вероятностных моделей от дискретно-детерминированных заключается в учете случайных элементов. Основной математический аппарат, используемый при построении и исследовании дискретно-вероятностных моделей, — это теория разностных стохастических уравнений и теория вероятностных автоматов.

Разностное стохастическое уравнение — это уравнение, которое содержит случайные параметры θ или (и) случайные входные воздействия $\{u_j\}$.

Пусть на вероятностном пространстве (Ω, F, P) определены случайный вектор параметров

$$\theta = \begin{pmatrix} \theta_0 \\ \dots \\ \theta_{m-1} \end{pmatrix} \in \mathbb{R}^m$$

и случайная последовательность входных воздействий $u_0, u_1, u_2, \dots \in U$.

Нелинейное разностное стохастическое уравнение имеет вид

$$\Phi(x_j, x_{j-1}, \dots, x_{j-n}, u_j, u_{j-1}, \dots, u_{j-l}, \theta, j) = 0, \quad (2.14)$$

где $j = n, n+1, \dots$; x_0, x_1, \dots, x_{n+1} — заданные начальные состояния системы; $\Phi(\cdot)$ — заданная функция $n + l + m + 3$ переменных.

Решением этого уравнения является определенная на (Ω, F, P) случайная последовательность состояний моделируемой системы $x_{n+2}, x_{n+3}, \dots \in X$.

Если $\Phi(\cdot)$ линейна по $\{x_j, u_j\}$, то (2.14) примет вид

$$\sum_{i=0}^n \theta_i x_{j-i} + \sum_{k=0}^l \theta_{k+n+1} u_{j-k} = u_j, \quad j = n+1, n+2, \dots, \quad (2.15)$$

где $\theta = (\theta_0, \dots, \theta_{n+l+1})^T$ — вектор $m = n + l + 2$ параметров. Уравнение (2.15) в математической статистике известно как уравнение авторегрессии и скользящего среднего порядка (n, l) . Частными случаями (2.15) являются модель авторегрессии $AR(n)$ порядка n :

$$\sum_{i=0}^n \theta_i x_{j-i} + u_j = 0,$$

и модель $СС(l)$ скользящего среднего порядка l :

$$x_j = \sum_{i=0}^l \theta_i u_{j-i}.$$

П р и м е р

Дискретно-вероятностная модель электрического колебательного контура, находящегося под влиянием случайных внешних воздействий $\{u_j\}$, имеет вид

$$x_j - 2x_{j-1} + \left(1 + \frac{\Delta^2}{\theta_1\theta_2}\right)x_{j-1} - \frac{\Delta}{\theta_2}u_{j-2} = 0,$$

$$y_j = x_j + u_j, \quad j = 2, 3, \dots,$$

где ошибка измерения $\{u_j\}$ представляется в виде стационарной случайной последовательности с нулевым математическим ожиданием и дисперсией σ^2 .

Другой математический аппарат построения дискретно-вероятностных моделей сложных систем представляет теория вероятностных автоматов. *Вероятностный автомат*, определенный на (Ω, F, P) , есть конечный автомат, в котором функция переходов $x' = \Phi(x, u, \omega)$, $\omega \in \Omega$, и функция выхода $y = \psi(x, u, \omega)$ являются случайными функциями, имеющими некоторые вероятностные распределения.

2.5.4. Непрерывно-вероятностные модели

При построении и исследовании непрерывно-вероятностных моделей используются теория стохастических дифференциальных уравнений и теория массового обслуживания.

Стохастическое дифференциальное уравнение (в форме Ито) имеет вид

$$dx(t) = a(t, x(t))dt + b(t, x(t))d\omega(t), \quad (2.16)$$

где $x(t)$ — случайный процесс, определяющий состояние системы в момент времени t ; $\omega(t)$ — стандартный винеровский случайный процесс; $a(\cdot)$, $b(\cdot)$ — коэффициенты переноса и диффузии соответственно. При некоторых условиях гладкости на коэффициенты $a(\cdot)$, $b(\cdot)$ решение этого стохастического дифференциального уравнения является марковским диффузионным процессом. Распределение вероятностей процесса удовлетворяет уравнениям Колмогорова. Такая непрерывно-вероятностная модель часто используется при моделировании стохастических систем управления, процессов тепло- и массообмена.

Теория массового обслуживания применяется для построения математических моделей таких сложных систем, для которых характерно:

- наличие потока многих заявок на выполнение определенных операций (заявок на обслуживание);
- наличие многократно повторяемых операций (выходной поток).

Теория массового обслуживания разрабатывает и исследует математические модели различных по своей природе процессов функционирования экономических, производственных, технических и других систем, например: поставок сырья и комплектующих изделий предприятию; поступления заданий на ЭВМ от удаленных терминалов; вызовов на телефонных станциях и т.д. При этом для функционирования

таких систем характерна стохастичность: случайность моментов времени появления заявок на обслуживание и т.д.

Сложная система S , описываемая как система массового обслуживания, состоит из $L \geq 1$ взаимосвязанных и взаимодействующих элементов-приборов обслуживания Π_1, \dots, Π_L . Прибор обслуживания Π_i состоит из накопителя заявок H_i , в котором могут одновременно находиться l_i заявок ($0 \leq l_i \leq m_i$), и канала K_i обслуживания заявок; m_i ($0 \leq m_i \leq \infty$) — емкость накопителя H_i , т.е. число мест в очереди для ожидания начала обслуживания в канале K_i , ($i = \overline{1, L}$). На каждый элемент прибора Π_i поступают потоки событий; в накопитель H_i — поток заявок $\{v_i\}$, на канал K_i — поток обслуживания $\{u_i\}$. Поток заявок $\{v_i\}$ представляет последовательность времени между моментами появления заявок на входе и образует подмножество неуправляемых переменных. Поток $\{u_i\}$ представляет собой последовательность интервалов времени между моментами начала и окончания обслуживания заявок и образует подмножество управляемых переменных. Заявки, обслуженные каналом K_i , и заявки, покинувшие прибор Π_i по различным причинам не обслуженными (например, из-за конечности m_i), образуют выходной поток $\{y_i\}$ — последовательность интервалов времени между моментами выхода заявок; это и есть выходной сигнал системы массового обслуживания.

Процесс функционирования прибора Π_i можно представить как процесс изменения состояний его элементов во времени $x_i = x_i(t) \in X \subset R^2$, где $x_{i1}(t)$ — число заявок, находящихся в накопителе в момент времени t ; $x_{i2}(t)$ — состояние канала обслуживания K_i ($x_{i2}(t) = 0$, если канал свободен, $x_{i2}(t) = 1$, если канал занят).

Приборы обслуживания Π_1, \dots, Π_L связаны и взаимодействуют между собой. Если каналы $\{K_i\}$ различных приборов обслуживания соединены параллельно, то имеет место многоканальное обслуживание. Если же приборы $\{\Pi_i\}$ или их параллельные блоки соединены последовательно, то имеет место многофазное обслуживание.

2.6. Методы упрощения математических моделей

Почти все реальные объекты и процессы чрезвычайно сложны, и их описание включает в себя бесконечное число элементов, переменных, параметров, соотношений, ограничений и т.д. Поэтому при построении моделей объектов и процессов приходится отбрасывать большую часть реальных характеристик и оставлять только те, которые воссоздают идеализированный вариант исследуемого объекта. Все модели являются либо упрощенными представлениями реального мира, либо абстракциями. Сходство модели с объектом, который она

отображает, называется степенью изоморфизма. Для того чтобы быть *изоморфной* (т.е. идентичной или сходной по форме), модель должна удовлетворять двум условиям:

- 1) должно существовать взаимно однозначное соответствие между элементами модели и элементами представляемого объекта;
- 2) должны существовать точные соотношения или взаимодействия между элементами.

Степень изоморфизма модели относительна, и большинство моделей скорее *гомоморфны*, т.е. существует сходство по форме при различии основных структур, причем имеет место лишь поверхностное подобие между различными группами элементов модели и объекта. Гомоморфные модели являются результатом *упрощения* и *абстракции*.

Для разработки идеализированной гомоморфной модели обычно разбивают систему на подсистемы или элементы для их анализа и интерпретации. Такой способ действий зависит от наличия подсистем или элементов, которые в первом приближении не зависят друг от друга или достаточно просто взаимодействуют между собой. Например, для анализа работы автомашины можно проанализировать последовательно работу двигателя, коробки передач, привод, систему подвески и т.д. С такого рода анализом при построении модели близко связан процесс упрощения реальной системы, где под упрощением понимается пренебрежение несущественными деталями или принятие предположений о более простых соотношениях. Например, часто предполагается, что между переменными имеет место линейная зависимость вместо существующей нелинейной, газы считаются идеальными, давление — адиабатическим, проводимость — однородной и т.д.

При построении моделей систем управления также прибегают к упрощению. Так, переменные считаются либо детерминированными при чрезвычайно упрощенной трактовке реальности, либо стохастическими, которые описываются известными вероятностными функциями распределения. Достаточно часто предполагается, что зависимость между переменными носит линейный характер, а параметры модели объекта являются точно заданными, несмотря на то, что эти допущения не совсем правомерны. Такие предположения часто используются для построения математических моделей систем управления.

Другим аспектом анализа является *абстракция* (от латинского *abstractio* — отвлечение) — *форма познания, основанная на мысленном выделении существенных свойств и связей предмета и отвлечении от других частных его свойств и связей. Абстракция является результатом процесса абстрагирования.*

Основными типами абстракции являются:

- *изолирующая абстракция* — вычленение исследуемого явления из некоторой целостности;

- *обобщающая абстракция* — описание обобщенной картины явлений;
- *идеализация* — замещение реального явления идеализированной схемой.

Большинство моделей — это абстракции в том смысле, что они стремятся представить качества и поведение моделируемого объекта в форме или способом, отличающимся от их действительной реализации.

После анализа и моделирования частей и элементов системы приступают к их объединению в единое целое, т.е. путем синтеза конструируют некоторое приближение к реальной ситуации. При этом необходимо, чтобы используемые для синтеза части были выбраны корректно и корректно установлено их взаимодействие.

Если эти условия соблюдены, то процессы анализа, абстракции, упрощения и синтеза в итоге приведут к созданию модели, которая *аппроксимирует* поведение исследуемой реальной системы. При этом необходимо учитывать, что модель является только приближением (аппроксимацией), поэтому она будет себя вести не как реальный объект.

3. ЭЛЕМЕНТЫ ТЕОРИИ ПОДОБИЯ

Исследование систем с помощью моделей может быть обосновано только в том случае, когда модель адекватно описывает процессы и явления, происходящие в системах. В основе доказательства степени соответствия модели и объекта лежит *теория подобия*. Моделирование представляет собой процесс проведения исследований объекта, который базируется на подобии модели и объекта и включает в себя построение модели, ее изучение и перенос полученных результатов на объект исследования. При этом под моделью понимается объект, явление, процесс, система, знаковое образование, математическое выражение, которые находятся в отношении подобия к исследуемому объекту.

Математическое моделирование использует подобие между величинами, входящими в математические выражения, описывающие поведение изучаемого объекта. Физическое моделирование использует подобие между объектом и моделью, имеющей физическую природу. Основой физического или математического экспериментального исследования являются методы теории подобия, которые применяются как при постановке экспериментов, так и при обработке данных о результатах этих экспериментов.

3.1. Использование теории подобия при физическом моделировании

При физическом моделировании модель воспроизводит изучаемый процесс с сохранением его физической природы. Преимущества физического моделирования перед натурным экспериментом заключаются в том, что условия моделирования могут значительно отличаться от условий, свойственных исследуемому объекту или процессу, так как нет необходимости сохранять размеры сооружений, скорости течения жидкостей и газов, нагрузки на элементы конструкций и т.д.

Научной основой моделирования с использованием физических моделей является *теория подобия* — учение об исследовании физических явлений, основанное на понятии о физическом подобии. Два физических явления подобны, если по численным значениям характеристик одного можно получить численные значения характеристик другого явления простым пересчетом, аналогично переходу от одной системы единиц измерения к другой. Для всякой совокупности подобных явлений все соответствующие безразмерные

характеристики (безразмерные комбинации из размерных величин) имеют одинаковое численное значение. Верно и обратное — если все соответствующие безразмерные характеристики двух явлений одинаковы, то эти явления физически подобны.

3.1.1. Анализ размерностей

Метод установления связи между существенными для изучаемого явления физическими величинами, основанный на рассмотрении размерностей этих величин, называется анализом размерностей. В этом методе рассматриваются проблемы установления соответствия различных систем единиц измерения, вопросы о выборе первичных величин и соответствующих им опытных единиц измерения, образование вторичных величин измерения для величин, определяемых через первичные.

Существуют различные системы единиц измерения. Это система СГС, в которой первичны сантиметр, грамм-масса и секунда; система МКГСС, в которой первичны метр, килограмм-сила и секунда; система СИ, в которой первичны метр, килограмм-масса, секунда, ампер, кельвин, кандела и др. Выражение производной единицы измерения через основные называется *формулой размерности*, которая записывается через символы первичных единиц измерения и имеет вид степенных одночленов.

Например, в системе СГС формулы размерности содержат три аргумента: единицу длины L , единицу времени T и единицу массы M . На основании определения силы через массу и ускорение формула размерности силы K имеет вид

$$K = \frac{ML}{T^2} = MLT^{-2}.$$

В системе СГС формулы для любой величины N механической, тепловой или электромагнитной природы имеют вид

$$N = L^l T^t M^m, \quad (3.1)$$

где показатели степеней l , t , m — некоторые целые или дробные действительные числа, которые называются *показателями размерности* или *размерностью* величины N . При этом считают, что размерность первичной величины в отношении себя равна единице, а в отношении любой другой первичной величины — нулю.

Формулы размерности позволяют определить масштабные множители для пересчета соответствующих характеристик при изменении величин первичных единиц измерения. Если вместо заданных единиц измерения длины L , времени T и массы M перейти к новым единицам измерения, меньшим для длины в α раз, для времени в β раз и для массы в γ раз, то новая единица измерения для величины N

с размерностью по формуле (3.1) будет меньше первоначальной в $\alpha^l \beta^t \gamma^m$ раз. Если $l = t = m = 0$, то численное значение величины не зависит от выбора масштабов для первоначальных единиц и, следовательно, такая величина будет безразмерной.

Физические закономерности, вообще говоря, не зависят от выбора системы единиц измерения. Это обстоятельство обуславливает особую структуру функций и функционалов, выражающих с помощью размерных величин физические закономерности, не зависящие от системы единиц измерения. Эта специальная структура функциональных соотношений устанавливается π -теоремой, согласно которой всякое соотношение между размерными характеристиками, имеющими физический смысл, представляет собой соотношение между отвлеченными безразмерными комбинациями из размерных определяемых и определяющих величин. Причем среди этих величин должны быть и размерные физические постоянные, имеющие существенное значение в рассматриваемых явлениях.

Таким образом, необходимо ввести систему определяющих параметров постоянных или переменных, вытекающую из постановки выделяемого класса задач и характеризующую полностью для данной среды каждую отдельно взятую задачу. *Определяющие параметры — это величины, характеризующие каждый отдельный эксперимент, которые являются необходимыми и достаточными для повторения и сравнения различных экспериментов.* В число определяющих параметров необходимо включать независимые переменные типа координат точек пространства и времени, физические параметры типа коэффициента упругости, вязкости и т.п. В случае физического подобия критерии подобия могут быть получены без математического описания объектов, на основании анализа размерностей и π -теоремы.

Если рассматриваемая задача сформулирована как математическая, т.е.

$$a = f(a_1, a_2, \dots, a_n), \quad (3.2)$$

где a — искомая величина, a_1, a_2, \dots, a_n — определяющие параметры, то определяющие параметры — это все исходные данные, которые нужно знать для вычисления функции (3.2). В зависимости от постановки задачи число параметров n может быть конечным или бесконечным. Обычно для рассматриваемого класса задач можно ограничиваться случаями, когда n конечно и, вообще говоря, невелико. Параметры a_1, a_2, \dots, a_n состояются из величин, входящих в основные уравнения, граничные условия и начальные данные.

Выводы анализа размерностей получаются на основании π -теоремы о существовании соотношения вида (3.2), которое выполняется для определяемой величины a и определяющих величин a_1, a_2, \dots, a_n в лю-

бой системе единиц измерения.

Система аргументов в соотношении (3.2) должна в смысле теории размерностей обладать полнотой. Это значит, что если величина a отлична от нуля или бесконечности, то существуют числа $\alpha_1, \alpha_2, \dots, \alpha_n$ такие, что размерности величины a и комбинаций $a_1^{\alpha_1}, a_2^{\alpha_2}, \dots, a_n^{\alpha_n}$ одинаковы. На основании этого и π -теоремы соотношение (3.2) принимает вид

$$\pi = \frac{a}{a_1^{\alpha_1} a_2^{\alpha_2} \dots a_n^{\alpha_n}} = g(\pi_1, \pi_2, \dots, \pi_{n-k}),$$

где $\pi_1, \pi_2, \dots, \pi_{n-k}$ — степенные безразмерные одночлены (аналогичные $\pi = 3, 14, \dots$), образованные из a_1, a_2, \dots, a_n , причем $k \leq n$. Число k равно числу параметров с независимыми размерностями среди a_1, a_2, \dots, a_n .

Основная практическая польза применения анализа размерностей и π -теоремы состоит в установлении возможности перенесения результатов опыта, проведенного в одних условиях, в другие условия, в которых опыт не проводился. Например, по опытам с водой, движущейся по трубе с фиксированным диаметром, можно автоматически давать нужные ответы о движении воды в трубах той же формы, но других размеров, о движении в трубах нефти, воздуха и т.п. Однако для возможности перенесения результатов данного опыта на другие опыты необходимо обеспечивать одинаковые значения безразмерных параметров. Значения некоторых из этих величин легко обеспечиваются геометрическими условиями постановки опытов, значения других параметров связаны с физико-механическими условиями проведения опытов.

3.1.2. Физическое моделирование

Анализ размерностей и теория подобия тесно связаны между собой и положены в основу экспериментов с моделями. В таких экспериментах осуществляются замены изучения некоторого физического явления изучением аналогичного явления на модели меньшего или большего масштаба (обычно в специальных лабораторных условиях).

После установления системы параметров, определяющих выделенный класс явлений, устанавливаются условия подобия двух явлений. Пусть установлено n независимых параметров, некоторые из которых являются безразмерными. Пусть размерности определяющих параметров и физических постоянных выражены через размерности k из этих параметров с независимыми размерностями $k \leq n$. Тогда из n величин можно записать только $n-k$ независимых безразмерных комбинаций, составленных из определяющих параметров. Среди всех безразмерных величин всегда можно указать некоторую базу — систему безразмер-

ных величин, которые определяют собой все остальные.

Для подобию двух явлений необходимо и достаточно, чтобы численные значения безразмерных комбинаций, составленных из полного перечня определяющих параметров, образующих базу, были одинаковыми в этих двух явлениях. Условия постоянства базы отвлеченных параметров, составленных из заданных величин, определяющих рассматриваемое явление, называются *критериями подобию*.

Например, в гидродинамике важнейшими критериями подобию являются:

- число Рейнольдса, характеризующее соотношение между инерционными силами и силами вязкости: $R = \frac{\rho v l}{\mu}$;
- число Фруда, характеризующее соотношение между инерционными силами и силами тяжести: $Fr = \frac{v^2}{\sqrt{g l}}$;
- число Маха, учитывающее сжимаемость газа: $M = \frac{v}{a}$.

Здесь приняты следующие обозначения: ρ — плотность, v — скорость, l — линейный размер, μ — динамический коэффициент вязкости, g — ускорение силы тяжести, a — скорость звука.

П р и м е р

В задаче об установившемся обтекании тела несжимаемой вязкой жидкостью все безразмерные величины, характеризующие движение в целом, определяются тремя параметрами: углами α, β (направление поступательной скорости тела относительно его поверхности) и числом Рейнольдса R . Условия физического подобию (критерии подобию) представляются соотношениями

$$\alpha = \text{const}, \quad \beta = \text{const}, \quad R = \frac{Pvd}{\mu} = \text{const}.$$

Таким образом, при моделировании результаты опытов с моделью можно переносить на физические явления только при одинаковых значениях α, β и R . Первые два условия всегда легко осуществить на практике, третье — труднее, особенно в тех случаях, когда модель меньше обтекаемого тела. Поэтому для сохранения величины числа Рейнольдса необходимо либо увеличивать скорость обтекаемого потока, что обычно практически не осуществимо, либо существенно изменять плотность и вязкость жидкости. На практике эти обстоятельства приводят к большим затруднениям при изучении аэродинамического сопротивления, например, продувка самолетов в натуральную величину в аэродинамических трубах.

При моделировании движения тел в газе с большими скоростями

необходимо иметь одинаковые значения числа Маха на модели и в натуре, а при моделировании плавания кораблей по воде необходимо обеспечивать равенство для натуре и модели чисел Фруда и Рейнольдса.

С помощью моделирования можно изучать явления, протекающие в течение длительного или очень короткого периода, так как на моделях подобные явления могут продолжаться несколько часов или дней.

Физическое моделирование имеет ограниченную сферу применения. Более широкими возможностями обладает математическое моделирование.

3.2. Математическое подобие

Математическое подобие (аналогия) — это сходство по количественным признакам, имеющим математическое выражение в виде некоторых уравнений. *Математические аналоги* — это объекты, описываемые сходственными уравнениями, которые получаются приравниванием нулю сходственных функций, т.е. функций одинакового вида, отличающихся только аргументами и отличными от нуля постоянными коэффициентами, например $z_1 = a_{10}x_{11} \sin(a_{11}x_{12})$ и $z_2 = a_{20}x_{21} \sin(a_{21}x_{22})$. *Сходственные переменные* — переменные, входящие под знаки сходственных функций одинаковым образом (z_1 и z_2 , x_{11} и x_{21} , x_{12} и x_{22}). Два объекта подобны, если:

- они имеют сходственные математические описания в форме уравнений вида

$$F(z, x_i, a_j, t_s, D_s) = 0, \quad i = 1, 2, \dots, \quad j = 1, 2, \dots, \quad s = 1, 2, \dots, \quad (3.3)$$

где $D_s = \frac{d}{dt_s}$;

- сходственные переменные (z_1 и z_2 , x_{1i} и x_{2i} , t_{1s} и t_{2s}) связаны постоянными коэффициентами подобия — константами подобия:

$$C_z = \frac{z_1}{z_2}, \quad C_{x_i} = \frac{x_{1i}}{x_{2i}}, \quad C_{t_s} = \frac{t_{1s}}{t_{2s}} = \frac{D_{2s}}{D_{1s}}. \quad (3.4)$$

При условиях (3.4) соответствующие сходственные уравнения, функции и переменные называются *подобными*. Благодаря константам подобия результаты, полученные для одного объекта, могут быть трансформированы в соответствующие результаты для подобного объекта. Необходимое условие подобия — совместность уравнений (3.3) и (3.4).

Константы подобия (3.4) связаны определенными уравнениями констант. Для их вывода сходственные уравнения (3.3) приводятся к безразмерной форме

$$\Phi(z, x_i, a_j, t_s, D_s) \pm 1 = 0 \quad (3.5)$$

и произведения степеней z , x_i , a_j , t_s , D_s объединяются в безразмерные степенные комплексы вида

$$\pi_r = a_j z^{\alpha_r} x_i^{\beta_{i,r}} t_s^{\gamma_{s,r}} D_s^{\delta_{s,r}},$$

называемые *критериями подобия*. В результате безразмерные функции $\Phi(\cdot)$ представляются безразмерными *критериальными функциями*

$$\Psi(\pi_r) = \Phi(z, x_i, a_j, t_s, D_s),$$

а безразмерная форма уравнения (3.5) представляется *критериальным уравнением*

$$\Psi(\pi_r) \pm 1 = 0.$$

В случае подобия сходственные критерии равны:

$$\pi_{1r} = \pi_{2r},$$

что записывается символически в виде $\pi_r = \text{idem}$ (*idem* — неизменно).

Уравнения констант подобия имеют вид

$$\frac{\pi_{1r}}{\pi_{2r}} = \frac{a_{1j}}{a_{2j}} C_z^{\alpha_r} C_{x_i}^{\beta_{i,r}} C_{t_s}^{\delta_{s,r} - \gamma_{s,r}} = 1. \quad (3.6)$$

Уравнения системы (3.6) должны быть совместными и независимыми. Если они несовместны, то подобие невозможно ни при каких значениях констант. Зависимые уравнения из системы (3.6) необходимо исключить. Число независимых уравнений равно числу m независимых критериев подобия π_r , которое определяет π -теорема. Согласно π -теореме зависимость, связывающая $n = k + m$ переменных и постоянных размерных величин, среди которых k величин обладают независимыми размерностями, может быть преобразована в зависимость между $m = n - k$ независимыми безразмерными степенными комплексами n величин.

П р и м е р

Пусть модели объектов описываются уравнениями

$$D_1 z_1 + a_{11} z_1 - a_{12} x_1 = 0;$$

$$D_2 z_2 + a_{21} z_2 - a_{22} x_2 = 0,$$

где

$$D_1 = \frac{d}{dt_1}; \quad D_2 = \frac{d}{dt_2}.$$

Необходимо привести эти уравнения к безразмерной форме, например, вида

$$\frac{D_1 z_1}{a_{12} x_1} + \frac{a_{11} z_1}{a_{12} x_1} - 1 = 0;$$

$$\frac{D_2 z_2}{a_{22} x_2} + \frac{a_{21} z_2}{a_{22} x_2} - 1 = 0$$

и получить критериальные уравнения

$$\pi_{11} + \pi_{12} - 1 = 0; \quad \pi_{21} + \pi_{22} - 1 = 0,$$

где

$$\pi_{11} = \frac{D_1 z_1}{t_1 a_{12} x_1}; \quad \pi_{12} = \frac{a_{11} z_1}{t_1 a_{12} x_1}; \quad \pi_{21} = \frac{D_2 z_2}{t_2 a_{22} x_2}; \quad \pi_{22} = \frac{a_{21} z_2}{t_2 a_{22} x_2}.$$

Тогда уравнения констант будут иметь вид

$$\frac{\pi_{11}}{\pi_{21}} = \frac{a_{22} C_z}{a_{12} C_x C_t} = 1, \quad \frac{\pi_{12}}{\pi_{22}} = \frac{a_{11} a_{22} C_z}{a_{21} a_{12} C_x} = 1, \quad (3.7)$$

где

$$C_z = \frac{z_1}{z_2}; \quad C_x = \frac{x_1}{x_2}; \quad C_t = \frac{t_1}{t_2} = \frac{D_2}{D_1}.$$

Одна из констант может быть выбрана произвольно, две другие однозначно определяются из уравнений (3.7).

Частными случаями математического подобия являются:

- геометрическое — подобие геометрических образов;
- временное — подобие функций времени, при котором временная константа показывает, в каком отношении находятся такие параметры функций, как период, временная задержка и т.д.

3.3. Элементы статистической теории подобия

В общем случае параметры исследуемых систем, процессов или явлений могут представлять собой случайные величины. Поэтому необходимо применять основные положения теории подобия с учетом стохастического характера процессов и явлений, происходящих в объектах. Принципы подобия в стохастическом смысле основаны на том, что сравниваемые параметры являются случайными величинами, а критерии подобия — функциями этих случайных величин.

Пусть имеется соотношение вида

$$\rho = \rho_1 \wedge \rho_2 \wedge \dots \wedge \rho_m, \quad (3.8)$$

где $\rho_i = \frac{x_{1i}}{x_{2i}}$, $i = \overline{1, m}$, — статистики критериев, сформированные для каждой группы одноименных параметров; $\{x_{1i}\}$ — параметры системы-оригинала; $\{x_{2i}\}$ — параметры модели.

В стохастической постановке принято считать, что две системы подобны, когда функции распределения параметров, характеризующих эти системы,

равны, а статистика критерия подобия находится в пределах границ доверительного интервала.

Способ получения численного значения статистик критерия подобия состоит в следующем. Пусть случайные параметры x_{1i} имеют плотность распределения $f_{1i}(x_{1i})$, параметры x_{2i} имеют плотность распределения $f_{2i}(x_{2i})$, $i = \overline{1, m}$. Тогда в терминах задачи проверки статистических гипотез нулевая гипотеза H_0 будет состоять в том, что $f_{1i}(x_{1i}) = f_{2i}(x_{2i})$ для всех одноименных параметров. Альтернативная гипотеза заключается в том, что $f_{1i}(x_{1i}) \neq f_{2i}(x_{2i})$. Для критерия подобия можно записать, что сравниваемые системы подобны, если $\hat{\rho} \in [\rho_n, \rho_b]$. Здесь ρ_n, ρ_b — соответственно нижняя и верхняя границы доверительного интервала для критерия подобия, определяемые с некоторым уровнем значимости.

Для определения численного значения статистики критерия ρ при справедливости гипотезы H_0 необходимо по известным плотностям распределения случайных величин x_{1i} и x_{2i} определить плотности распределения $f_{1i}(x_{1i})$ и $f_{2i}(x_{2i})$, $i = \overline{1, m}$, от этих плотностей перейти к плотности распределения величины ρ , а затем вычислить выборочное значение статистики $\hat{\rho}$ и границы доверительного интервала $[\rho_n, \rho_b]$.

Так как в (3.8) $\rho_i = \frac{x_{1i}}{x_{2i}}$, $i = \overline{1, m}$, то сравнению подлежат параметры одного типа и множество величин ρ_i представляет собой набор независимых случайных величин. Тогда объект-оригинал и модель находятся в отношении подобия, если они подобны по каждому определяющему параметру. Теперь задача сводится к определению доверительного интервала по каждому критерию подобия ρ_i с одним и тем же уровнем значимости. Если такие интервалы найдены и для каждого критерия выполняются условия $\rho \in [\rho_n, \rho_b]$, то объекты можно считать подобными. Если хотя бы по одному критерию условие подобия не выполняется, то оснований считать объекты подобными нет. Заметим, что часть анализируемых параметров может иметь детерминированный характер, и подобие по этим параметрам определяется путем деления параметров объекта-оригинала на соответствующие параметры модели. В этом случае размерность задачи (3.8), требующей анализа с использованием аппарата случайных функций, существенно понижается.

Плотность распределения величины ρ определяется следующим образом. Имеем $\hat{\rho} = \frac{x_{1i}}{x_{2i}}$ — выборочное значение статистики критерия, определенное по результатам испытаний системы оригинала и модели,

$F_{\rho_i}(k)$ — функция распределения величины ρ_i . Не нарушая общности, опустим индекс в записи для функции распределения, так как аналогичные выкладки будут иметь место для любых параметров системы, имеющих характер случайных величин. Выражение для функции распределения имеет вид

$$F_{\rho}(k) = P\{\rho \leq k\} = P\left\{\frac{x_1}{x_2} \leq k\right\}.$$

Так как случайные величины x_1 и x_2 независимы, то их совместная плотность распределения равна произведению плотностей этих величин, а вероятность соотношения $\frac{x_1}{x_2} \leq k$ выражается интегралом от совместной плотности по области, определенной неравенствами $x_2 > 0$, $x_1 \leq x_2 k$:

$$F_{\rho}(k) = \iint_{\substack{x_2 > 0 \\ x_1 \leq x_2 k}} f_1(y)f_2(z)dydz. \quad (3.9)$$

Будем считать, что функция $F_{\rho}(k)$ дифференцируема, т.е. существует плотность $f_{\rho}(k)$.

Точные доверительные границы определяются из соотношений

$$\int_0^{\rho_n} f_{\rho}(k / H_0) dk = \alpha, \quad \int_{\rho_b}^{\infty} f_{\rho}(k / H_0) dk = \alpha, \quad (3.10)$$

где α — уровень значимости. В (3.10) неизвестными величинами являются значения границ доверительного интервала ρ_n, ρ_b , относительно которых необходимо решить эти интегральные уравнения. Определив функцию распределения статистики критерия (3.9), можно ограничиться вычислением приближенных границ доверительного интервала. Для этого необходимо получить плотность распределения статистики критерия ρ , вычислив математическое ожидание

$$m_{\rho} = \int_{\Omega_{\rho}} k f_{\rho}(k) dk$$

и среднее квадратическое отклонение статистики критерия

$$\sigma_{\rho}^2 = \int_{\Omega_{\rho}} k^2 f_{\rho}(k) dk - m_{\rho}^2,$$

где Ω_{ρ} — область определения критерия ρ . В общем случае область определения критерия $\Omega_{\rho} = [0, \infty]$. Определив математическое ожида-

ние и среднее квадратическое отклонение, можно вычислить приближенные границы доверительного интервала по формулам

$$\rho_n = m_p - t_\alpha \sigma_p, \quad \rho_b = m_p + t_\alpha \sigma_p,$$

где t_α — квантиль распределения Стьюдента, определенный для уровня значимости α .

Проверка гипотезы о подобии объекта-оригинала и модели принимается на основании сравнения оценок, полученных при функционировании объекта-оригинала и исследовании модели. При этом так как оценки определяются путем обработки выборки случайных чисел ограниченного объема, то они сами являются случайными величинами, имеющими плотность распределения $f_x(t)$.

Таким образом, проверяя по каждой группе одноименных параметров гипотезу о их подобии, можно сделать вывод о подобии системы оригинала и модели.

4. СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Под *статистическим моделированием* понимается воспроизведение с помощью ЭВМ функционирования вероятностной модели некоторого объекта. Цель моделирования такого рода состоит в оценивании с его помощью средних характеристик модели. Обычно это математическое ожидание величин, характеризующих систему, их дисперсии и ковариации.

Статистическое моделирование является неотъемлемой и органичной составляющей моделирования систем в силу, во-первых, случайного характера многих реальных явлений и процессов и, во-вторых, неизбежности рандомизации (искусственного внесения случайности) моделируемых систем. Рандомизация служит эффективным средством уменьшения количества существенных и зависимых и независимых факторов модели, подлежащих измерению, наблюдению и исследованию, за счет возможности учета и описания части таких факторов «в среднем», т.е. путем их статистического представления.

Методы статистического моделирования весьма многочисленны; здесь будут обсуждаться только те из них, которые представляют наибольший интерес с точки зрения практического применения к рассматриваемым вопросам.

4.1. Методы имитации случайных элементов на ЭВМ

Функционирование реальных объектов и процессов обычно осуществляется под влиянием случайных факторов. При моделировании случайные факторы рассматриваются либо как случайные входные (управляющие) сигналы, либо как неконтролируемые воздействия внешней среды.

Примерами таких случайных факторов являются: случайные интервалы времени между поступлениями заданий в вычислительную систему, деталей на производственный участок, вызовов в АТС; время обработки и передачи информации в системах передачи информации и т.д.

4.1.1. Принципы моделирования случайных элементов

Моделирование на ЭВМ случайного элемента подчиняется двум основным принципам:

1) сходство между случайным элементом-оригиналом Ξ^* и его моделью Ξ состоит в совпадении (близости) вероятностных законов распределения или числовых характеристик;

2) всякий случайный элемент Ξ определяется (конструируется) как некоторая функция от простейших случайных элементов, так называемых базовых случайных величин.

Простейшим для моделирования на ЭВМ случайным экспериментом является эксперимент, заключающийся в бросании точки наудачу в промежуток $[0, 1)$. Результатом этого эксперимента является координата точки. Математической моделью такого эксперимента является вероятностное пространство (Ω, F, P) , где $\Omega = [0, 1)$ — пространство элементарных событий (элементарное событие $\omega \in \Omega$ заключается в том, что координата брошенной точки равна ω); F — σ -алгебра, порожденная интервалами из Ω ; P — вероятностная мера, определенная для событий (подмножеств) $A \in F$ и совпадающая с мерой Лебега, так что для события $A = \{\omega : \omega \in [0, x]\}$

$$P(A) = P\{\omega \in [0, x]\} = x, \quad x \in [0, 1).$$

Базовой случайной величиной (БСВ) на (Ω, F, P) будем называть непрерывную случайную величину

$$\alpha = \alpha(\omega) = \omega,$$

равномерно распределенную на полуинтервале $[0, 1)$.

Функция распределения БСВ имеет вид

$$F_{\alpha}(x) = P\{\alpha < x\} = \begin{cases} 0, & x \leq 0, \\ x, & 0 < x < 1, \\ 1, & x \geq 1, \end{cases}$$

а плотность распределения определяется формулой

$$p_{\alpha}(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 0, & x \notin [0, 1). \end{cases}$$

Обозначим закон распределения α как $R[0, 1)$.

Математическое ожидание БСВ (первый начальный момент) определяется следующим образом:

$$\mu = M\{\alpha\} = \frac{1}{2};$$

а дисперсия (второй центральный момент):

$$\sigma^2 = D\{\alpha\} = M\{(\alpha - \mu)^2\} = \frac{1}{12}.$$

Наряду с простейшим экспериментом будем рассматривать составной случайный эксперимент, получающийся в результате r -кратного ($r \geq 1$) независимого повторения простейших экспериментов. Результатом составного случайного эксперимента является последовательность из r независимых БСВ $\alpha_1, \dots, \alpha_r$ таких, что

$$\alpha_i = \alpha_i(\omega) = \omega_i, \quad i = \overline{1, r}, \quad \omega = (\omega_1, \dots, \omega_r) \in \Omega^r,$$

где ω_i — координата точки, брошенной наудачу в $[0, 1)$ в i -м простейшем эксперименте.

Совместная плотность распределения вероятностей $\alpha_1, \dots, \alpha_r$

$$p_{\alpha_1, \dots, \alpha_r}(x_1, \dots, x_r) = \begin{cases} 1, & x_i \in [0, 1), \quad i = \overline{1, r}, \\ 0, & \text{в противном случае.} \end{cases}$$

Согласно второму принципу моделирования случайных элементов любой случайный элемент Ξ представляется для некоторого натурального r в виде функции $f(\cdot)$ от r независимых БСВ

$$\Xi = f(\alpha_1, \dots, \alpha_r).$$

Таким образом, задача моделирования произвольного случайного элемента Ξ^* разбивается на две подзадачи:

- 1) моделирование на ЭВМ независимых БСВ $\alpha_1, \dots, \alpha_r$ (блок Б1);
- 2) нахождение функции $f(\cdot)$ такой, чтобы случайный элемент Ξ обладал требуемыми вероятностным законом распределения и числовыми характеристиками (блок Б2).

Поэтому моделирующий алгоритм состоит из двух блоков (рис. 4.1).

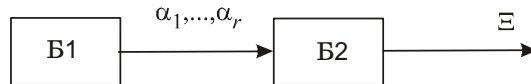


Рис. 4.1

Для определения одного и того же случайного элемента Ξ^* может быть предложено несколько вариантов функциональных преобразований. Обычно предпочтение отдается варианту, требующему меньших вычислительных затрат. Для этого используется понятие коэффициента использования БСВ.

Коэффициентом использования БСВ k называется величина, обратная числу r базовых случайных величин, используемых для моделирования одной реализации случайного элемента Ξ^* :

$$k = \frac{1}{r}, \quad 0 < k \leq 1.$$

Величина k является мерой вычислительных затрат на моделирование Ξ^* . Чем меньше k , тем больше затраты. Целесообразно выбирать такую функцию $f(\cdot)$, для которой k принимает наибольшее значение.

Очевидно, чтобы моделировать на ЭВМ случайные элементы с заданным вероятностным законом распределения, необходимо уметь моделировать БСВ. Так как БСВ α является абсолютно непрерывной случайной величиной, а на ЭВМ приходится иметь дело с дискретными случайными величинами, то моделирование БСВ основано на аппроксимации непрерывной случайной величины α дискретной случайной величиной α' .

4.1.2. Типы датчиков базовых случайных величин

Датчик базовой случайной величины — это устройство, позволяющее по запросу получать реализацию a или несколько независимых реализаций a_1, \dots, a_r базовой случайной величины α . Реализации a_1, \dots, a_r иногда называют случайными числами.

Существуют три типа датчиков базовой случайной величины:

- 1) табличный;
- 2) физический;
- 3) программный.

Табличный датчик — это таблица случайных чисел, представляющая собой экспериментально полученную выборку реализации равномерно распределенной в интервале $[0, 1)$ случайной величины. Применение табличных датчиков при моделировании на ЭВМ ограничено, так как достаточно часто используются миллионы случайных чисел, а это требует большого объема оперативной памяти для хранения таблицы.

Физический датчик — это специальное радиоэлектронное устройство, являющееся приставкой к ЭВМ, выходной сигнал которой имитирует значение базовой случайной величины. Это устройство состоит из источника флуктуационного шума (например, «флуктуационно шумящей» радиолампы), значение которого в произвольный момент времени является случайной величиной $\eta \geq 0$ с плотностью $p_\eta(y)$, и нелинейного преобразователя

$$\alpha = \frac{\{\eta\}_\Delta}{\Delta},$$

где $\{\eta\}_\Delta = \eta - \Delta \left[\eta / \Delta \right]$ — дробная часть числа η относительно $\Delta > 0$ ($[\cdot]$ — целая часть числа). Полагая Δ достаточно малой величиной, можно получить базовую случайную величину α .

Физический датчик обладает следующими недостатками:

- 1) невозможность повторения некоторой ранее полученной реализации a (поскольку $P\{\alpha = a\} = 0$);
- 2) наличие схемной нестабильности, приводящей к необходимости контроля работы датчика при очередном его использовании.

По этим причинам на современных компьютерах физические датчики используются крайне редко.

Программный датчик — это программа, служащая для определения на ЭВМ реализаций БСВ, т.е. для получения последовательности псевдослучайных чисел a_1, a_2, \dots . Термин «псевдослучайные числа» используется по следующим причинам:

- 1) по происхождению эти числа не случайные, так как они получаются согласно заданному детерминированному закону;
- 2) по вероятностным характеристикам эти числа похожи на реализации независимых БСВ.

4.1.3. Методы построения программных датчиков

Рассмотрим основные методы и алгоритмы, используемые при построении программных датчиков БСВ.

1. *Мультипликативный конгруэнтный метод (метод вычетов)*. Согласно этому методу псевдослучайная последовательность вычисляется по рекуррентным формулам

$$a_i = \frac{a_i^*}{M}, \quad a_i^* = (\beta a_{i-1}^*) \bmod M, \quad i = 1, 2, \dots, \quad (4.1)$$

где β, M, a_0^* — параметры программного датчика (натуральные числа): β — множитель, M — модуль, $a_0^* \in \{1, \dots, M-1\}$ — стартовое значение; операция $y = (z) \bmod M$ означает вычет числа z по модулю M , т.е.

$$y = z - M \left[\frac{z}{M} \right].$$

Из (4.1) следует:

- последовательность $\{a_i^*\}$, а значит и $\{a_i\}$, всегда «зацикливается», т.е. начиная с некоторого номера i_0 , образуется цикл, который повторяется бесконечное число раз;
- период последовательности $T \leq M-1$ (если $a_i^* = 0$, то $a_{i+j}^* = 0$ для любого $j \geq 1$).

Параметры β, M, a_i^* выбираются из условия максимума периода. Период T можно определить аналитически методами теории чисел или с помощью численных экспериментов на ЭВМ.

П р и м е р

Для типового программного датчика RANDU параметры β, M, a_i^* и T имеют следующие значения:

- для 16-разрядной ЭВМ
 $\beta = 259, \quad M = 32768, \quad a_0^* = 259, \quad T = M / 4 = 8192;$
- для 32-разрядной ЭВМ
 $\beta = 65539, \quad M = 2147483648, \quad a_0^* = 65539, \quad T = M / 4 = 536870912.$

2. *Метод, использующий линейные смешанные формулы.* В этом методе псевдослучайная последовательность вычисляется рекуррентно по так называемой линейной смешанной формуле, обобщающей (4.1):

$$a_i = \frac{a_i^*}{M}, \quad a_i^* = (\beta_1 a_{i-1}^* + \dots + \beta_p a_{i-p}^* + c) \bmod M, \quad i = 1, 2, \dots$$

Параметры датчика: p — порядок; $a_{-p+1}^*, a_{-p+2}^*, \dots, a_{-1}^*, a_0^*$ — стартовые значения; β_1, \dots, β_p — множители; c — приращение; M — модуль. Период датчика $T \leq M^p - 1$, т.е. растет с увеличением M и p .

3. *Метод, использующий нелинейные рекуррентные формулы.* Псевдослучайная последовательность генерируется с помощью квадратичного рекуррентного соотношения

$$a_i = \frac{a_i^*}{M}, \quad a_i^* = (\gamma (a_{i-1}^*)^2 + \beta a_{i-1}^* + c) \bmod M, \quad i = 1, 2, \dots \quad (4.2)$$

Параметры a_0^*, γ, β, M целесообразно выбирать с помощью следующего утверждения: если $M = 2^q$, $q \geq 2$, то наибольший период $T_{\max} = M$ последовательности (4.2) достигается тогда и только тогда, когда β, c — нечетные, γ — четное, причем

$$(\beta) \bmod 4 = (\gamma + 1) \bmod 4.$$

4. *Метод Макларена-Марсальи.* Метод основан на комбинировании двух простейших программных датчиков БСВ. Пусть $\{b_i\}, \{c_i\}$, $i = 0, 1, 2, \dots$, — псевдослучайные последовательности, порождаемые независимо работающими программными датчиками Д1 и Д2 соответственно; $V = \{V(0), \dots, V(K-1)\}$ — вспомогательная таблица K чисел. Вначале V -таблица заполняется K членами последовательности $\{b_i\}$:

$$V(j) = b_j, \quad j = \overline{0, K-1}.$$

Через $\{a_i\}$ будем обозначать комбинированную псевдослучайную последовательность. Член a_i является результатом последовательности операций

$$s = [c_i K], \quad a_i = V(s), \quad V(s) = b_{s+K}, \quad i = 0, 1, 2, \dots$$

Таким образом, Д2 осуществляет случайный выбор из V -таблицы и случайное заполнение числами, порожденными Д1. Этот метод позволяет ослабить зависимость между членами a_i и получить невероятно большие периоды, если периоды T_1, T_2 последовательностей $\{b_i\}, \{c_i\}$ — взаимно простые числа.

5. *Метод, основанный на свойстве воспроизводимости равномерного закона.* Пусть $\xi_1, \xi_2, \dots, \xi_N \in [0, 1)$ — независимые равномерно распределенные случайные величины; N — натуральное число. Если определить случайные величины

$$\eta_N = \xi_1 + \dots + \xi_N, \quad \zeta_N = (\eta_N) \bmod 1 = \eta_N - [\eta_N],$$

то случайная величина ζ_N имеет стандартное равномерное распределение $R[0, 1)$.

4.2. Моделирование случайных величин

4.2.1. Моделирование дискретных случайных величин

Моделированием случайной величины принято называть процесс получения на ЭВМ последовательности выборочных значений случайной величины с заданным законом распределения.

1. *Моделирование дискретной случайной величины.* Пусть ξ^* — дискретная случайная величина, принимающая $N < \infty$ возможных значений $c_0 < c_1 < \dots < c_{N-1}$ с заданными вероятностями p_0, p_1, \dots, p_{N-1} соответственно. Алгоритм моделирования реализации ξ случайной величины ξ^* состоит в следующем: вычисляется вспомогательный вектор

$$q = (q_0, q_1, \dots, q_{N-1}) = (p_0, p_0 + p_1, \dots, p_0 + p_1 + \dots + p_{N-2}, 1);$$

моделируется с помощью датчика БСВ реализация a базовой случайной величины α ; принимается решение о том, что реализацией ξ^* является x , определяемое по правилу

$$x = c_i, \text{ если } q_{i-1} \leq a \leq q_i \quad (q_{-1} \equiv 0), \quad i = \overline{0, N-1}.$$

Для того чтобы построить последовательность из n реализаций случайной величины ξ^* (случайную выборку объема n), достаточно повторить данный алгоритм n раз, перед каждым повторением алгоритма меняя стартовое число a_0^* в датчике БСВ. Коэффициент использования БСВ $k = 1$.

2. *Моделирование полной группы случайных событий.* Система подмножеств $Y_0, Y_1, \dots, Y_{N-1} \in F$ называется *полной группой случайных событий* на (Ω, F, P) , если:

- $\bigcup_{i=0}^{N-1} Y_i = \Omega$ — происходит хотя бы одно событие из группы;
- $Y_i \cap Y_j = \emptyset$ для $i \neq j$ — совместное наступление двух и более событий из группы невозможно;
- $P(Y_i) > 0, \quad i = \overline{0, N-1}$, — все события группы имеют ненулевые вероятности.

П р и м е р

В случайном эксперименте по бросанию монеты полную группу образуют два события: Y_0 — выпадение орла, Y_1 — выпадение решки.

Для моделирования полной группы случайных событий может быть применен алгоритм моделирования дискретной случайной величины. Действительно, обозначим: $p_i = P(Y_i)$, $i = \overline{0, N-1}$, — вероятность i -го события из группы; ξ^* — дискретная случайная величина, принимающая значения $0, 1, \dots, N-1$ с вероятностями p_0, p_1, \dots, p_{N-1} соответственно (ξ^* — номер события из полной группы, наступающего в случайном эксперименте). Тогда задача моделирования случайного события сводится к задаче моделирования случайной величины ξ^* .

3. *Моделирование однородной цепи Маркова.* Пусть $\xi_0^*, \xi_1^*, \dots, \xi_n^*$ — однородная цепь Маркова, т.е. последовательность дискретных случайных величин со значениями из множества $\{c_0, c_1, \dots, c_{N-1}\}$ ($c_0 < c_1 < \dots < c_{N-1}$) и распределением вероятностей, которое определяется:

- вектором вероятностей начального состояния $\pi = (\pi_0, \pi_1, \dots, \pi_{N-1})$,

где $\pi_i = P\{\xi_0^* = c_i\} \geq 0$, $\sum_{i=0}^{N-1} \pi_i = 1$;

- матрицей вероятностей одношаговых переходов $P = (p_{ij})$, $i, j = \overline{0, N-1}$, где $p_{ij} = P\{\xi_{t+1}^* = c_j / \xi_t^* = c_i\} \geq 0$, $t = \overline{0, n-1}$, $\sum_{j=0}^{N-1} p_{ij} = 1$.

Определим вспомогательные вектор и матрицу:

$$q = (q_j), \quad q_j = \sum_{k=0}^j \pi_k;$$

$$Q = (q_{ij}), \quad q_{ij} = \sum_{k=0}^j p_{ik}, \quad i, j = \overline{0, N-1}.$$

Так как $\xi_0^*, \xi_1^*, \dots, \xi_n^*$ — дискретные случайные величины, то для моделирования каждой из них применяется приведенный выше алгоритм моделирования дискретной случайной величины при использовании соответствующего вектора вероятностей. Пусть a_0, a_1, \dots, a_n — реализации БСВ, имитирующие случайную выборку из $R[0, 1)$. Моделирование последовательности $\xi_0^*, \xi_1^*, \dots, \xi_n^*$ осуществляется в два этапа:

1) моделирование ξ_0^* , где принимается решение, что реализацией ξ_0^* является $x_0 = c_i$, т.е. цепь Маркова в начальный момент $t = 0$ находится в состоянии номер $v_0 = i$, если

$$q_{i-1} \leq a_0 < q_i \quad (q_{-1} \equiv 0), \quad i \in \{0, 1, \dots, N-1\};$$

2) моделирование $\xi_1^*, \xi_2^*, \dots, \xi_n^*$, при этом для моделирования ξ_1^* используется v_0 -я вектор-строка $(q_{v_0, 0}, \dots, q_{v_0, N-1})$ матрицы Q и принимается решение, что реализацией ξ_1^* является $x_1 = c_i$, если $(v_1 = i)$

$$q_{v_0, i-1} \leq a_1 < q_{v_0, i} \quad (q_{v_0, -1} \equiv 0), \quad i = \overline{0, N-1}.$$

Аналогично, используя строку $(q_{v_k, 0}, \dots, q_{v_k, N-1})$ матрицы Q , получаем реализацию случайной величины ξ_k^* , $k = \overline{1, n}$.

4.2.2. Моделирование непрерывных случайных величин

При построении имитационных моделей сложных систем часто возникают задачи моделирования на ЭВМ непрерывной случайной величины ξ^* с заданной плотностью распределения вероятностей $f_0(x)$, $x \in R^1$. В некоторых случаях для решения данных задач удается применить метод *обратной функции*.

Определим функцию распределения ξ^* :

$$F_0(x) = \int_{-\infty}^x f_0(y) dy, \quad (4.3)$$

которую будем предполагать строго монотонно возрастающей. Через $x = F_0^{-1}(y)$ обозначим обратную функцию; она находится при решении уравнения

$$F(x) = y. \quad (4.4)$$

Теорема. Если α — БСВ, определенная на (Ω, F, P) , то случайная величина

$$\xi = F_0^{-1}(\alpha) \quad (4.5)$$

имеет заданную плотность распределения $f_0(\cdot)$.

Моделирующий алгоритм метода обратной функции состоит в следующем:

- 1) нахождение функции распределения $F_0(x)$ по заданной плотности распределения $f_0(x)$ согласно (4.3);
- 2) нахождение обратной функции F_0^{-1} путем решения относительно x уравнения (4.4);
- 3) моделирование реализации БСВ α и вычисление реализации случайной величины ξ по формуле (4.5).

Коэффициент использования БСВ $k = 1$.

П р и м е р ы

1. Пусть $f_0(x)$ — плотность *равномерного распределения* $R[a, b]$, т.е.

$$f_0(x) = \begin{cases} \frac{1}{b-a}, & \text{если } x \in [a, b], \\ 0, & \text{если } x \notin [a, b]. \end{cases}$$

Согласно (4.3), (4.4)

$$F_0(x) = \begin{cases} 0, & \text{если } x < a, \\ \frac{x-a}{b-a}, & \text{если } x \in [a, b], \\ 1, & \text{если } x > b. \end{cases}$$

$$x = (b-a)y + a, \quad y \in R[0, 1).$$

Моделирование ξ^* осуществляется по формуле

$$\xi = (b-a)\alpha + a,$$

где α — БСВ.

2. Пусть $f_0(x)$ — плотность экспоненциального распределения, т.е.

$$f_0(x) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

По формулам (4.3), (4.4) получаем

$$F_0(x) = 1 - e^{-\lambda x}, \quad x \geq 0,$$

$$x = F_0^{-1}(y) = -(1/\lambda) \ln(1-\alpha).$$

Моделирование ξ^* с плотностью распределения $f_0(x)$ осуществляется по формуле

$$\xi = -\lambda^{-1} \ln(1-\alpha),$$

где α — БСВ. Так как α и $1-\alpha$ одинаково распределены, то эквивалентным моделирующим алгоритмом является следующий:

$$\xi = -\mu \ln(\alpha), \quad \mu = \lambda^{-1} = M\{\xi^*\}.$$

Так как $F_0^{-1}(x)$ аналитически не вычисляется, то метод обратной функции непосредственно применить нельзя.

3. Пусть $f_0(x)$ — плотность стандартного нормального распределения $N(0, 1)$. Функция распределения (функция Лапласа) описывается интегралом вероятностей

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

Известно, что $F_0(-x) = 1 - F_0(x)$. Если α — БСВ, то стандартная нормально распределенная величина η равна

$$\eta = \begin{cases} F_0^{-1}(\alpha), & \text{если } \frac{1}{2} \leq \alpha < 1, \\ -F_0^{-1}(1-\alpha), & \text{если } 0 < \alpha < \frac{1}{2}. \end{cases}$$

Можно аппроксимировать $F_0^{-1}(x)$, например если $\frac{1}{2} \leq \alpha < 1$, то

$$F_0^{-1}(\alpha) = \frac{2,30753 + 0,27061\theta}{1 + 0,99229\theta + 0,04481\theta^2} - \theta,$$

где $\theta = \sqrt{-2\ln(\alpha)}$. Погрешность аппроксимации не превосходит $3 \cdot 10^{-3}$.

Для моделирования стандартной нормально распределенной случайной величины $N\{0, 1\}$ достаточно часто используется *метод суммирования*. Согласно *центральной предельной теореме* если $\xi_1, \xi_2, \dots, \xi_n$ — последовательность взаимно независимых одинаково распределенных случайных величин ($M\{\xi_i\} = a$, $D\{\xi_i\} = \sigma^2$, $i = \overline{1, n}$), то функция распределения нормированной суммы

$$\eta_n = \frac{\sum_{i=1}^n \xi_i - na}{\sigma\sqrt{n}}$$

стремится к функции стандартного нормального распределения $N(0, 1)$. При этом вид $p_{\xi_i}(x)$, $i = \overline{1, n}$, значения не имеет. Тогда если $\xi_1, \xi_2, \dots, \xi_n$ —

реализации БСВ, т.е. $M\{\xi_i\} = \frac{1}{2}$, $D\{\xi_i\} = \frac{1}{12}$, $i = \overline{1, n}$, то

$$\eta_n = \sqrt{\frac{12}{n}} \sum_{i=1}^n \xi_i - \frac{n}{2}.$$

Для моделирования часто используется $n = 12$. Тогда величину

$$\eta_{12} = \sum_{i=1}^{12} \xi_i - 6$$

можно использовать в качестве стандартной нормально распределенной случайной величины. Нормально распределенная случайная величина τ , для которой $M\{\tau\} = \mu$, $D\{\tau\} = \sigma^2$, моделируется следующим образом:

$$\tau = \mu + \sigma\eta,$$

где τ — стандартная нормально распределенная случайная величина.

4. Пусть $x = (x_1, x_2, \dots, x_n)^T$ — случайный вектор. Рассмотрим моделирующий алгоритм для нормальных случайных векторов с характеристиками

$$M\{\xi_i\} = \mu_i, \quad M\{(\xi_i - \mu_i)^T (\xi_j - \mu_j)\} = R \delta_{ij}, \quad i, j = \overline{1, n},$$

где R — ковариационная матрица n -го порядка, δ_{ij} — символ Кронекера,

$$\delta_{ij} = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Моделирующий алгоритм состоит в следующем:

1) моделирование случайного вектора $\eta = (\eta_1, \eta_2, \dots, \eta_n)^T$, где η_i , $i = \overline{1, n}$, — независимые стандартные нормальные случайные величины;

2) построение компонентов вектора ξ по правилу

$$\xi_i = \mu_i + \sqrt{R_{ii}} \eta_i, \quad i = \overline{1, n}.$$

4.2.3. Моделирование случайных процессов

В задачах имитационного и статистического моделирования динамических сложных систем часто возникает необходимость реализации на ЭВМ случайного процесса $\xi^* = \xi_t^* = \xi^*(\omega, t) \in R^1$, $\omega \in \Omega$, $t \in [0, T]$, определенного на вероятностном пространстве (Ω, F, P) и временном промежутке $[0, T]$. Пусть $\Delta > 0$ — заданный шаг дискретизации времени; $T = M\Delta$; $t = m\Delta$, $m = \overline{0, M}$, где t — дискретное время; η_t — дискретный «белый шум», т.е. последовательность независимых в совокупности гауссовских случайных величин, распределенных по стандартному нормальному закону $N\{0, 1\}$. При этом последовательность $\xi_0^*, \xi_\Delta^*, \xi_{2\Delta}^*, \dots, \xi_{M\Delta}^*$ принято называть *временным рядом*.

Простейшей моделью случайного процесса ξ_t^* , часто встречающейся на практике, является *гауссовский временной ряд с трендом*

$$\xi_t^* = f(t) + \sigma \eta_t, \quad (4.6)$$

где $f(t)$ — детерминированная действительная функция времени, называемая *трендом*; $\sigma \geq 0$ — среднеквадратическое отклонение случайной ошибки. Если $f(t) \neq \text{const}$, временной ряд (4.6) является нестационарным. Имитация временного ряда ξ_t^* легко осуществляется по формуле (4.6); при этом последовательность независимых случайных величин реализуется рассмотренными ранее методами.

Обобщением модели (4.6) временного ряда ξ_t^* является модель

$$\xi_t^* = f(t) + \zeta_t^*,$$

где ζ_t^* — гауссовский стационарный временной ряд с нулевым математическим ожиданием $M \{\zeta_t^*\} = 0$ и ковариационной функцией $M \{\zeta_t^* \zeta_{t+\tau}^*\} = R_{\zeta^*}(\tau)$.

4.3. Методы Монте-Карло

Понятие «статистическое моделирование» тесно связано с понятием «метод Монте-Карло» — это метод решения задач, использующий случайные величины. Точнее под методом Монте-Карло понимается совокупность приемов, позволяющих получать решения математических или физических задач при помощи случайных многократных испытаний. На практике случайные испытания заменяются результатами некоторых вычислений, производимых над случайными числами.

Название «метод Монте-Карло» для методов, систематически использующих случайные величины, появилось в 1949 г. Создателями этого метода считают американских математиков Дж. Неймана и С. Улама. Само название «Монте-Карло» происходит от названия города Монте-Карло Княжества Монако, знаменитого своим игорным домом. Дело в том, что одним из простейших механических приборов для получения случайных величин является рулетка.

Методом Монте-Карло можно решать следующие задачи:

- 1) моделировать любой процесс, на протекание которого влияют случайные факторы;
- 2) для многих математических задач, не связанных с какими-либо случайностями, можно искусственно придумать вероятностную модель, позволяющую решать эти задачи.

Таким образом, метод Монте-Карло является универсальным методом решения математических задач. Метод Монте-Карло используется для решения систем линейных и нелинейных уравнений, решения дифференциальных и интегральных уравнений, вычисления интегралов высокой кратности, решений уравнений в частных производных и т.д.

Эффективное применение метода Монте-Карло стало возможным после появления ЭВМ, так как для получения достаточно точной оценки искомой величины требуется произвести вычисления для большого объема числовых данных. Это объясняется тем, что метод Монте-Карло использует различные предельные соотношения теории вероятностей — законы больших чисел и предельные теоремы.

Рассмотрим задачу вычисления кратных интегралов методами Монте-Карло. Для вычисления таких интегралов существуют *кубатурные формулы*, но они являются очень громоздкими, а для интегралов кратности выше третьей вообще неприменимыми. Рассмотрим два простейших варианта метода Монте-Карло, которые легко реализуются на ЭВМ.

Пусть требуется вычислить интеграл

$$I = \int_V f(x_1, \dots, x_n) dx_1 \dots dx_n, \quad (4.7)$$

где область V определяется неравенствами

$$\left. \begin{aligned} a_1 &\leq x_1 \leq b_1, \\ a_i &\leq \varphi_i(x_1, \dots, x_{i-1}) \leq x_i \leq \psi_i(x_1, \dots, x_{i-1}) \leq b_i, \quad i = \overline{2, n}. \end{aligned} \right\} \quad (4.8)$$

Суть методов Монте-Карло заключается в том, что для вычисления интеграла (4.7) строится последовательность $l_1, l_2, \dots, l_n, \dots$, для которой справедливо

$$\lim_{n \rightarrow \infty} P \{ |l_n - I| \leq \varepsilon \} = 1,$$

где $P \{ \cdot \}$ — вероятность события, l_n — приближенное значение интеграла, вычисленное на n -м шаге, $\varepsilon > 0$ — заданная точность.

4.3.1. Простейший метод Монте-Карло вычисления интегралов

Сначала рассмотрим применение методов Монте-Карло к вычислению однократного интеграла

$$I = \int_a^b f(x) dx, \quad (4.9)$$

где $f(x)$ задана на $[a, b]$. Выберем произвольную плотность распределения $p_\xi(x)$ случайной величины ξ . Плотность $p_\xi(x)$ определена на интервале $[a, b]$, причем $p_\xi > 0$ и $\int_a^b p_\xi(x) dx = 1$.

Введем новую случайную величину η , связанную со случайной величиной ξ формулой

$$\eta = \frac{f(\xi)}{p_\xi(\xi)}.$$

Математическое ожидание случайной величины η равно

$$M \{ \eta \} = \int_a^b \left[\frac{f(x)}{p_\xi(x)} \right] p_\xi(x) dx = \int_a^b f(x) dx = I.$$

Таким образом, можно вычислить значение интеграла (4.9), вычислив математическое ожидание случайной величины η . Для вычисления $M \{ \eta \}$ воспользуемся методами математической статистики. Пусть $\xi_1, \xi_2, \dots, \xi_N$ — N реализаций случайной величины ξ тогда при достаточно большом значении N имеем

$$I = M \{ \eta \} \approx \frac{1}{N} \sum_{j=1}^N \frac{f(\xi_j)}{p_\xi(\xi_j)}.$$

На основании правила трех сигм справедлива следующая формула:

$$P \left\{ \left| \frac{1}{N} \sum_{j=1}^N \frac{f(\xi_j)}{p_\xi(\xi_j)} - I \right| \leq 3 \sqrt{\frac{D(\eta)}{N}} \right\} = 0,997. \quad (4.10)$$

$D(\eta)$ — дисперсия случайной величины η , которая может быть вычислена по формуле

$$D(\eta) \approx \frac{1}{N} \sum_{j=1}^N (\eta_j^2) - \left[\frac{1}{N} \sum_{j=1}^N \eta_j \right]^2,$$

где

$$\eta_j = \frac{f(\xi_j)}{p_\xi(\xi_j)}.$$

Формула (4.10) означает, что с вероятностью, близкой к 1, абсолютная погрешность вычисления интеграла (4.9) не превосходит величины $3\sqrt{D(\eta)/N}$, т.е. погрешность метода вычисления интеграла приблизительно равна

$$\Delta_M \approx 3 \sqrt{\frac{D(\eta)}{N}}.$$

При реализации метода Монте-Карло обычно в качестве $p_\xi(x)$ используют равномерное распределение, т.е.

$$p_\xi(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b]; \\ 0, & x \notin [a, b]. \end{cases} \quad (4.11)$$

Для того чтобы получить реализации равномерно распределенной на $[a, b]$ случайной величины ξ , достаточно иметь реализации случайной величины τ , равномерно распределенной на $[0, 1]$, тогда

$$\xi = a + \tau(b - a).$$

Учитывая (4.11), получим

$$I \approx \frac{b-a}{N} \sum_{j=1}^N f(\xi_j),$$

$$D(\xi) \approx \frac{(b-a)^2}{N} \sum_{j=1}^N f^2(\xi_j) - I^2.$$

При вычислении кратного интеграла (4.7) область (4.8) с помощью линейной замены переменных

$$x_i = a_i + (b_i - a_i)t_i, \quad i = \overline{1, n}, \quad (4.12)$$

заключается в n -мерный единичный куб, т.е. интеграл в (4.7) приводится к виду

$$I = (b_1 - a_1) \dots (b_n - a_n) \int_{V^*} \dots \int f^*(t_1, \dots, t_n) dt_1 \dots dt_n, \quad (4.13)$$

где $(b_1 - a_1) \dots (b_n - a_n)$ — Якобиан преобразования,

$$f^*(t_1, \dots, t_n) = f(a_1 + (b_1 - a_1)t_1, \dots, a_n + (b_n - a_n)t_n),$$

область V^* определяется неравенствами

$$\left. \begin{aligned} 0 \leq t_1 \leq 1; \\ 0 \leq \varphi_i^*(t_1, \dots, t_{i-1}) \leq t_i \leq \psi_i^*(t_1, \dots, t_{i-1}) \leq 1, \quad i = \overline{2, n}. \end{aligned} \right\}$$

Здесь

$$\varphi_i^*(t_1, \dots, t_{i-1}) = \frac{1}{b_i - a_i} [\varphi_i(a_1 + (b_1 - a_1)t_1, \dots, a_{i-1} + (b_{i-1} - a_{i-1})t_{i-1}) - a_i],$$

$$\psi_i^*(t_1, \dots, t_{i-1}) = \frac{1}{b_i - a_i} [\psi_i(a_1 + (b_1 - a_1)t_1, \dots, a_{i-1} + (b_{i-1} - a_{i-1})t_{i-1}) - a_i],$$

$$i = \overline{2, n}$$

Тогда формулу (4.13) можно записать в виде

$$I = (b_1 - a_1) \dots (b_n - a_n) I^*, \quad (4.14)$$

где

$$I^* = \int_{V^*} \dots \int f^*(t_1, \dots, t_n) dt_1 \dots dt_n. \quad (4.15)$$

Метод Монте-Карло вычисления интеграла (4.15) заключается в следующем. Пусть N — совокупность точек $M_k(\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_n^{(k)})$, $k = \overline{1, N}$, координаты, которых являются независимыми случайными величинами, равномерно распределенными на интервале $[0, 1]$. Тогда полагаем

$$f^*(M_k) = \begin{cases} f^*(M_k), & \text{если } M_k \in V^*; \\ 0, & \text{если } M_k \notin V^*, \end{cases}$$

и

$$I^* \approx \frac{1}{N} \sum_{k=1}^N f^*(M_k), \quad (4.16)$$

при этом справедлива оценка

$$P \left\{ \left| I^* - \frac{1}{N} \sum_{k=1}^N f^*(M_k) \right| \leq 3 \sqrt{\frac{D^*(M)}{N}} \right\} = 0,997,$$

где

$$D^*(M) \approx \frac{1}{N} \sum_{k=1}^N (f^*(M_k))^2 - (I^*)^2. \quad (4.17)$$

При использовании этого метода необходимо задавать число N случайных точек такое, чтобы для требуемой точности вычисления интеграла $\varepsilon > 0$ было справедливо неравенство

$$3 \sqrt{\frac{D^*(M)}{N}} \leq \varepsilon.$$

Для определения N можно использовать итерационные алгоритмы. Например, можно сначала задать число точек, равное N_0 , а затем увеличивать его по формуле $N_k = N_{k-1} + H_N$, $k = 1, 2, \dots$, где $H_N > 1$ — некоторое заданное целое число. При этом в (4.16) и (4.17) значения функций в суммах необходимо добавлять, не пересчитывая их заново.

Другой алгоритм определения N состоит в следующем. Известно, что если заданы $\varepsilon > 0$ и $0 < \alpha < 1$, то с вероятностью α имеет место неравенство

$$N \geq \frac{t_\alpha^2}{\varepsilon^2} \left(\int_V \dots \int [f^*(t_1, \dots, t_n) - I^*]^2 dt_1 \dots dt_n \right),$$

если

$$\left| I^* - \frac{1}{N} \sum_{k=1}^N f^*(M_k) \right| < \varepsilon,$$

где t_α определяется по α из таблицы значений интеграла вероятностей. Для правила трех сигм $\alpha = 0,997$; $t_\alpha = 3$.

Тогда, задавая N_0 , вычисляют сначала значения

$$I_{N_0}^* = \frac{1}{N_0} \sum_{j=1}^{N_0} f^*(M_j),$$

$$D_{N_0}^* = \frac{1}{N_0} \sum_{j=1}^{N_0} [f^*(M_j)]^2 - (I_{N_0}^*)^2,$$

и очередное значение N_1 вычисляется по формуле

$$N_1 = \frac{\left(1 + 4 \sqrt{\frac{2}{N_0}} \right) D_{N_0}^*}{\varepsilon^2} t_\alpha^2.$$

Если $N_1 > N_0$, то увеличивают число случайных точек до N_1 и вычисляют $I_{N_1}^*$ и $D_{N_1}^*$ и т.д. Для прекращения вычислений необходимо выполнение условия

$$N_{k+1} \leq N_k.$$

Для уменьшения объема вычислений можно воспользоваться следующими рекуррентными формулами:

$$I_{N_{k+1}}^* = \frac{1}{N_{k+1}} \left[N_k I_{N_k}^* + \sum_{j=1}^{N_{k+1}-N_k} f^*(M_j) \right];$$

$$D_{N_{k+1}}^* = \frac{1}{N_{k+1}} \left[N_k D_{N_k}^* + N_{N_k}^2 I_{N_k}^2 + \sum_{j=1}^{N_{k+1}-N_k} (f^*(M_j))^2 \right] - I_{N_{k+1}}^2.$$

4.3.2. Геометрический метод Монте-Карло вычисления интегралов

Пусть при вычислении однократного интеграла (4.9) для подынтегральной функции $f(x)$ выполняется условие

$$c \leq f(x) \leq d.$$

Сначала проведем следующие преобразования: введем новую функцию

$$\phi(x) = \frac{f(x) - c}{d - c},$$

значения которой лежат в интервале $[0, 1]$. Тогда

$$\int_a^b f(x) dx = (d - c) \int_a^b \phi(x) dx + c(b - a);$$

затем, выполнив замену переменной $x = a + (b - a)t$, получим

$$I = \int_a^b f(x) dx = (d - c)(b - a) \int_0^1 \phi(a + (b - a)t) dt + c(b - a). \quad (4.18)$$

Таким образом, необходимо вычислить интеграл

$$I^* = \int_0^1 \phi(a + (b - a)t) dt,$$

а затем получить значение I согласно (4.18).

Для реализации метода генерируется N точек $M_k(\xi^{(k)}, \eta^{(k)})$, $k = \overline{1, N}$, где $\xi^{(k)}, \eta^{(k)}$ — независимые случайные величины, равномерно распределенные на интервале $[0, 1]$. Если для совокупности двух случайных величин $\xi^{(k)}$ и $\eta^{(k)}$ окажется, что

$$\phi(a + (b-a)\xi^{(k)}) < \eta^{(k)},$$

то событие считается неблагоприятным; если

$$\phi(a + (b-a)\xi^{(k)}) \geq \eta^{(k)},$$

то событие считается благоприятным, так как в этом случае точка с координатами $(\xi^{(k)}, \eta^{(k)})$ попадает в заштрихованную область (рис. 4.2), площадь которой равна значению интеграла I^* .

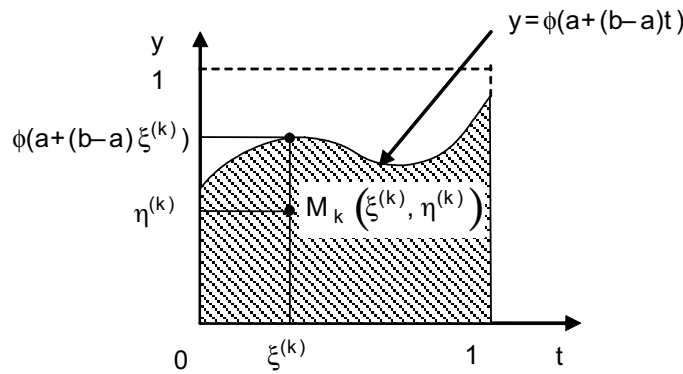


Рис. 4.2

Пусть из N точек v попали в заштрихованную область. Частота благоприятных событий $\frac{v}{N}$ будет приблизительно равна площади заштрихованной области, т.е.

$$I^* = \int_0^1 \phi(a + (b-a)t) dt \approx \frac{v}{N}.$$

Погрешность метода можно определить, воспользовавшись неравенством Чебышева

$$P \left\{ \left| \frac{v}{N} - I^* \right| \leq \frac{1}{2\sqrt{\alpha N}} \right\} = 1 - \alpha, \quad (4.19)$$

где $\alpha > 0$ — малая величина, обычно $\alpha = 0,01$. Формула (4.19) означает, что с вероятностью $(1 - \alpha)$ погрешность метода при вычислении интеграла I^* с помощью геометрического метода Монте-Карло приблизительно равна

$$\Delta_M \approx \frac{1}{2\sqrt{\alpha N}}.$$

При этом значение N можно определить из неравенства

$$\frac{1}{2\sqrt{\alpha N}} \leq \varepsilon. \quad (4.20)$$

При вычислении кратного интеграла (4.7) с областью интегрирования V , определяемой неравенствами (4.8), требуется, чтобы для подынтегральной функции $f(x_1, \dots, x_n)$ в области V выполнялось условие

$$c \leq f(x_1, \dots, x_n) \leq d.$$

Сделав замену переменных (4.12), получим для I выражение (4.13) или (4.14), (4.15). Введем новые функции

$$f^{**}(t_1, \dots, t_n) = \frac{f^*(t_1, \dots, t_n) - c}{d - c}$$

и

$$\bar{f}(t_1, \dots, t_n) = \begin{cases} 1, & \text{если } M(t_1, \dots, t_n) \in V^*; \\ 0, & \text{если } M(t_1, \dots, t_n) \notin V^*. \end{cases}$$

Тогда

$$I = (b_1 - a_1) \dots (b_n - a_n) \left\{ \int_{V^*} \dots \int \bar{f}(t_1, \dots, t_n) dt_1 \dots dt_n + (d - c) \int_{V^*} \dots \int f^{**}(t_1, \dots, t_n) dt_1 \dots dt_n \right\}$$

Для вычисления полученных интегралов генерируется N случайных точек $M_k(\xi_1^{(k)}, \dots, \xi_n^{(k)}, \eta^{(k)})$, $k = \overline{1, N}$, где $\xi_1^{(k)}, \dots, \xi_n^{(k)}, \eta^{(k)}$ — независимые случайные величины, равномерно распределенные на интервале $[0, 1]$. Для первого интеграла v_1 соответствует числу благоприятных событий, если выполняются неравенства

$$\left. \begin{aligned} 0 \leq \xi_1^{(k)} \leq 1; \\ \Phi_i^*(\xi_1^{(k)}, \dots, \xi_{i-1}^{(k)}) \leq \xi_i^{(k)} \leq \Psi_i^*(\xi_1^{(k)}, \dots, \xi_{i-1}^{(k)}), \quad i = \overline{1, n}, \end{aligned} \right\} \quad (4.21)$$

а v_2 соответствует числу благоприятных событий для второго интеграла, если выполняются неравенства (4.21) и

$$\eta^{(k)} \leq f^{**}(\xi_1^{(k)}, \dots, \xi_n^{(k)}).$$

Тогда значение I будет равно

$$I \approx (b_1 - a_1) \dots (b_n - a_n) \left(c \frac{v_1}{N} + (d - c) \frac{v_2}{N} \right).$$

Для определения необходимого числа N можно воспользоваться неравенством (4.20) или утверждением, что для заданных $\varepsilon > 0$ и $0 < \alpha < 1$ с вероятностью α имеет место неравенство

$$\left| I - \frac{v}{N} \right| < \varepsilon,$$

если

$$N \geq \frac{I(1-I)}{\varepsilon^2} t_\alpha^2,$$

где, как и ранее, t_α определяется по α из таблицы значений интеграла вероятностей. Тогда для определения N используется следующий алгоритм: задается некоторое значение N_0 , вычисляется интеграл I_{N_0} и полагается

$$N_1 = \frac{I_{N_0}(1-I_{N_0})}{\varepsilon^2} t_\alpha^2.$$

Если $N_1 > N_0$, то добавляют $(N_1 - N_0)$ точек, вычисляется значение интеграла I_{N_1} , и так до тех пор, пока для некоторого k не выполнится неравенство

$$N_{k+1} \leq N_k, \quad k = 0, 1, \dots$$

Здесь в качестве v можно использовать v_2 и определять N по точности вычисления интеграла для функции $f^{**}(t_1, \dots, t_n)$.

5. ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ

Управление в современном мире становится все более трудным делом. Это объясняется усложнением самих объектов и, следовательно, усложнением их моделей. Система, объединяющая все объекты и связи между ними, является сложной системой. Изменение одной из характеристик такой системы может привести к изменениям или создать потребность в изменениях в других частях системы. В связи с этим появилась одна из методологий системного анализа, которая призвана помочь осмыслить последствия таких изменений. Это *имитационное моделирование* (*имитировать* — значит вообразить, постичь суть явления, не прибегая к экспериментам на реальном объекте). Согласно Р. Шеннону, *имитационное моделирование есть процесс конструирования модели реальной системы и постановка экспериментов на этой модели с целью либо понять поведение системы, либо оценить (в рамках ограничений, накладываемых некоторым критерием или совокупностью критериев) различные стратегии, обеспечивающие функционирование данной системы.*

Имитационное моделирование — наиболее мощный и универсальный метод исследования и оценки эффективности систем, поведение которых зависит от воздействия случайных факторов. К таким системам можно отнести и летательный аппарат, и популяцию животных, и предприятие, работающее в условиях слаборегулируемых рыночных отношений. В основе имитационного моделирования лежит статистический эксперимент (метод Монте-Карло), реализация которого невозможна без применения средств вычислительной техники. Поэтому любая имитационная модель представляет собой, в конечном счете, более или менее сложный программный продукт.

Имитационное моделирование является экспериментальной и прикладной методологией, имеющей следующие цели:

- описание поведения систем;
- построение теорий и гипотез, которые могут объяснить наблюдаемое поведение;
- прогнозирование поведения систем, т.е. тех воздействий, которые могут быть вызваны изменениями в самой системе или изменениями способов ее функционирования.

В отличие от большинства технических методов, которые могут быть классифицированы в соответствии с научными дисциплинами, в которые они уходят своими корнями (например, с физикой или химией), имитационное моделирование применимо к любой отрасли науки.

5.1. Принципы построения имитационных моделей

Имитационной моделью сложной системы S называются программы или алгоритмы, позволяющие имитировать на ЭВМ поведение отдельных элементов системы S и связей между ними в течение заданного времени моделирования. Эксперименты с имитационной моделью на ЭВМ называются *имитационными (вычислительными) экспериментами*.

Все имитационные модели функционируют как модели типа черного ящика. Это означает, что они обеспечивают выдачу выходного сигнала системы, если на ее взаимодействующие подсистемы поступает входной сигнал. Поэтому для получения необходимой информации или результатов необходимо осуществлять *прогон* имитационных моделей. Имитационные модели не способны формировать свое собственное решение в том виде, в каком это имеет место в аналитических моделях, а могут лишь служить в качестве средства для анализа поведения системы в условиях, определенных экспериментатором. Следовательно, имитационное моделирование не теория, а методология решения проблем.

Прежде чем начать разработку имитационной модели, необходимо понять, что собой представляют структурные элементы, из которых она состоит. Несмотря на то что математическая или физическая структура модели чаще всего является очень сложной, основы ее построения достаточно просты. В общем виде структуру модели можно представить математически в виде

$$Z = F(X, Y),$$

где $Z = (z_1, \dots, z_{n_z})^T$ — результат действия модели; $X = (x_1, \dots, x_{n_x})^T$ — параметры и переменные, которыми можно управлять; $Y = (y_1, \dots, y_{n_y})^T$ — параметры и переменные, которыми нельзя управлять; F — функциональная зависимость X и Y .

Почти каждая модель представляет собой некоторую комбинацию следующих составляющих:

- элементов (подсистем, компонентов), т.е. тех составных частей, которые при объединении образуют систему;
- переменных двух типов — входных и выходных;
- параметров;
- функциональных зависимостей, описывающих поведение переменных и параметров;
- ограничений, устанавливающих пределы изменения значений переменных;
- целевой функции или функции критерия.

Существуют два типа целей имитационного моделирования: приобретение и сохранение. *Цели сохранения* связаны с сохранением или поддержанием каких-либо ресурсов (временных, энергетических и т.д.)

или состояний (уровня производства, скорости, направления и т.д.). Цели приобретения связаны с приобретением новых ресурсов (прибыли, заказчиков и т.д.) или достижением определенных состояний, к которым стремится организация или ее руководитель (захват части рынка и т.д.). Выражение для целевой функции должно быть однозначным определением целей и задач, с которыми должны соизмеряться принимаемые решения. Согласно словарю Вебстера *критерий* — это мерило оценки, правило или вид проверки, при помощи которых составляется правильное суждение о чем-либо. Функция критерия обычно является составной частью модели, и весь процесс работы с моделью направлен на оптимизацию или удовлетворение заданного критерия.

Процесс создания модели исследуемой системы может быть лучше всего определен как интуитивное искусство. Основой успешной методики моделирования является тщательная отработка модели — постепенное продвижение от простой модели ко все более сложной. *Искусство моделирования состоит в способности анализировать проблему, выделять из нее путем абстракции существенные черты, выбирать и должным образом модифицировать основные предположения, характеризующие систему, а затем отрабатывать и совершенствовать модель до тех пор, пока она не станет давать полезные для практики результаты.*

5.2. Организация процесса моделирования

При разработке имитационных моделей сложных систем приходится учитывать две особенности функционирования ЭВМ:

1) сложная система S , как правило, состоит из многих элементов, и все элементы системы функционируют одновременно. Однако в большинстве современных ЭВМ параллельное выполнение нескольких программ, имитирующих поведение отдельных элементов системы, невозможно;

2) поскольку имитационные модели — это программы для ЭВМ, то они должны оперировать с конечным множеством данных и, следовательно, имитировать поведение системы S не во все моменты времени $t \in [0, T]$ а лишь в некоторые моменты времени, составляющие конечное множество.

Чтобы обеспечить имитацию параллельных (одновременных) событий системы S на конечном множестве моментов времени, в имитационной модели используется специальная переменная t , называемая *системным модельным временем* или просто *модельным временем*.

Модельное время t следует отличать от других типов времени, используемых при имитационном моделировании систем, таких как t_p — реальное время функционирования системы S ; t_s — машинное время имитации, отражающее затраты ресурса времени ЭВМ на организацию имитационного моделирования.

Существуют два способа формирования конечного множества моментов времени моделирования, известных как принципы организации изменения модельного времени « Δt » и « Δx ».

Принцип « Δt » заключается в изменении модельного времени с фиксированным шагом Δt . В соответствии с принципом « Δt » изменение модельного времени t происходит через промежутки времени, равные Δt , т.е. t в течение времени моделирования T принимает конечное множество значений $\{0, \Delta t, 2\Delta t, \dots, \gamma\Delta t = T\}$. При этом событиям, попадающим в интервал постоянства модельного времени $\delta_r = ((r-1)\Delta t, r\Delta t)$, $r = \overline{0, \gamma}$, в имитационной модели присваивается один и тот же момент наступления $t = r\Delta t$. Выбор величины Δt существенно влияет как на быстродействие имитационного моделирования, так и на точность аппроксимации сложной системы S имитационной моделью, при этом:

- если Δt мало, то выполняется много лишних вычислений состояний системы в моменты, когда состояние не меняется, за счет этого возрастает t_{Σ} выполнения имитационного моделирования;
- даже при сравнительно малом значении Δt моменты наступления событий в системе, а следовательно, и моменты изменения состояния системы не совпадают с моментами наступления событий в имитационной модели, поэтому фазовая траектория, построенная с помощью имитационной модели, не совпадает с фазовой траекторией системы S .

Принцип « Δx » заключается в изменении модельного времени при скачкообразном изменении вектора состояния $x(t)$ системы S на некоторую величину Δx ($\Delta x \neq 0$). Для моментов времени t^* из множества $\mathfrak{X} \subset [0, T]$, сформированного по принципу « Δx », справедливо

$$x(t^* + 0) = x(t^*) + \Delta x, \quad t^* \in \mathfrak{X}.$$

Для моментов времени $t \in [0, T]$, не входящих во множество \mathfrak{X} , вектор состояний изменяется непрерывно (либо остается неизменным). Заметим, что скачкообразные изменения состояния системы S происходят при наступлении особых событий, таких как поступление управляющих сигналов и внешних воздействий, выдача выходных сигналов и т.п.

Пусть сложная система S состоит из N элементов $A^{(1)}, \dots, A^{(N)}$, поведение которых предполагается моделировать, т.е.

$$S = \{A^{(1)}, \dots, A^{(N)}\}.$$

Для каждого элемента $A^{(i)} \in S$, $i = \overline{1, N}$, существует локальное модельное время $t^{(i)} \in [0, T]$. Поведение элемента $A^{(i)} \in S$ в течение интервала моделирования определяется некоторой последовательностью действий

$$g_1^{(i)}, g_2^{(i)}, \dots, g_{M_i}^{(i)}, \quad g_j^{(i)} \in G, \quad j = \overline{1, M_i},$$

где G — множество всевозможных действий для элементов S . На множестве G выделим подмножество действий $D \subset G$, для выполнения которых в имитационной модели требуется некоторое ненулевое модельное время. Обозначим такие действия

$$d_1^{(i)}, \dots, d_{m_i}^{(i)}, \quad d_j^{(i)} \in D \subset G, \quad j = \overline{1, m_i}, \quad m_i \leq M_i, \quad i = \overline{1, N},$$

а интервалы модельного времени, затрачиваемые на выполнение этих действий, соответственно $\tau_1^{(i)}, \dots, \tau_{m_i}^{(i)}$. Последовательность $\{\tau_j^{(i)}\}$, $j = \overline{1, m_i}$, достаточно часто является последовательностью случайных величин с заданными законами распределения $L\{\tau_j^{(i)}\}$, $i = \overline{1, N}$.

Действия $\{d_j^{(i)}\} \in D$ приводят к наступлению в системе S особых событий $\{A_j^{(i)}\}$. События $\{C_j^{(i)}\}$, к которым приводят действия $\{g_j^{(i)}\}$, не входящие в D , т.е. не требующие затрат модельного времени, считаются *неособыми*.

Момент локального модельного времени наступления события $A_j^{(i)}$ для $A_j^{(i)} \in S$ определяется по формуле

$$t_j^{(i)} = t^* + \tau_j^{(i)}, \quad j = \overline{1, M_i}, \quad i = \overline{1, N},$$

где $\tau_j^{(i)}$ чаще всего имитируется на ЭВМ в соответствии с заданным законом распределения $L\{\tau_j^{(i)}\}$; t^* — текущее значение модельного времени.

Состояние системы S в момент времени $t \in [0, T]$ определяется вектором состояния $x(t) \in X \subset R^n$. Состояния системы в моменты $t_j^{(i)} \in \mathcal{X}$ наступления особых событий называются *особыми состояниями*, а состояние $x(0)$ — *начальным состоянием* системы.

В соответствии с принципом « Δx » изменение модельного времени происходит в моменты наступления событий или, что то же самое, в моменты особых состояний, а фазовая траектория, построенная с помощью имитационного моделирования, будет совпадать на множестве $\mathcal{X} \subset [0, T]$ с фазовой траекторией системы S .

П р и м е р

Применением принципов « Δt » и « Δx » в системах автоматического управления является квантование сигналов по времени и по уровню.

На практике отдается предпочтение принципу « Δx ». Принцип « Δt » используется лишь в случаях:

1) когда события $\{A_j^{(i)}\}$ таковы, что $t_j^{(i)} - t_{j-1}^{(i)} \cong \text{const}$ на всем интервале моделирования и, следовательно, можно подобрать значение Δt изменения модельного времени, обеспечивающее минимальную погрешность аппроксимации (например для разностных уравнений);

2) событий очень много, и они появляются группами. В этом случае за счет групповой обработки событий $\{A_j^{(i)}\}$, попавших внутрь очередного шага изменения Δt , удастся уменьшить затраты машинного времени.

В большинстве случаев события $\{A_j^{(i)}\}$ наступают через случайные интервалы времени $\{\tau_j^{(i)}\}$. Поэтому способ задания шага до следующего события экономичнее (в смысле затрат машинного времени) и точнее (в смысле точности аппроксимации фазовой траектории) способа фиксированного изменения модельного времени.

5.3. Способы имитации

Под *способом имитации системы* $S = \{A^{(1)}, A^{(2)}, \dots, A^{(N)}\}$ понимается способ формирования фазовой траектории системы, т.е. способ изменения вектора состояния $x(t)$ системы S . Возможны три способа изменения $x(t)$:

- 1) в моменты наступления событий $\{A_j^{(i)}\}$;
- 2) в результате действий $\{d_j^{(i)}\}$, на выполнение которых требуются затраты модельного времени $\{\tau_j^{(i)}\}$, при этом пара $\{d_j^{(i)}, \tau_j^{(i)}\}$ называется (i, j) -й *активностью* системы S ;
- 3) в результате выполнения хронологической последовательности событий и действий, называемой *процессом*.

В зависимости от того, какой из способов формирования фазовой траектории используется, различают следующие способы имитации: событийный, основанный на просмотре активностей, процессный, транзактный и агрегатный.

Сформулируем условия применения способов имитации для моделирования системы $S = \{A^{(1)}, A^{(2)}, \dots, A^{(N)}\}$.

Событийный способ заключается в следующем:

- 1) множество особых событий разбивается на небольшое число L типов

$$\{A^{(1)}, \dots, A^{(L)}\} \subset \{A_j^{(i)}\} \quad j = \overline{1, m_i}, \quad i = \overline{1, N}, \quad L \ll \sum_{i=1}^N m_i;$$

- 2) для каждого типа событий определяется последовательность действий, приводящая к изменению состояния системы S ;
- 3) определяются условия перехода от одного события к другому для всех типов событий;
- 4) интервалы времени между последовательными наступлениями событий являются случайными величинами с известными законами распределения вероятностей.

Способ, основанный на просмотре активностей, заключается в следующем:

- 1) все действия для элемента $A^{(i)}$ системы S являются различными и приводят к наступлению различных событий;
- 2) каждое действие $d_j^{(i)}$ характеризуется набором условий его выполнения;
- 3) времена выполнения действий $\{\tau_j^{(i)}\}$ являются случайными величинами с известным законом распределения вероятностей.

Моделирующий алгоритм, основанный на просмотре активностей, реализует просмотр всех наборов условий, а также обрабатывает активности $\{d_j^{(i)}, \tau_j^{(i)}\}$, условия для которых выполняются, т.е. моделирует время выполнения соответствующего действия и реализует само действие. Данный способ требует значительных затрат машинного времени для проверки многих условий. Его целесообразно применять в тех случаях, когда время выполнения действий зависит от состояния системы.

Процессный способ сочетает в себе черты событийного способа и способа, основанного на просмотре активностей. Он применяется, когда поведение элементов $\{A^{(i)}\}$ системы S может быть описано фиксированными для некоторого класса систем последовательностями событий и действий, так называемыми *процессами*. Для его реализации используются специализированные системы моделирования — языки моделирования или пакеты прикладных программ, обладающие специальными средствами, например операторами языка моделирования, удобными для описания данных процессов. Однако, поскольку набор подобных стандартных средств, как правило, ограничен, этот способ является менее гибким, чем событийный.

Транзактный способ имитации сформировался в результате развития процессного способа для моделирования систем массового обслуживания.

Агрегатный способ основывается на использовании агрегативных моделей.

5.4. Этапы имитационного моделирования

Процесс имитационного моделирования сложной системы S можно условно разделить на три последовательно выполняемых этапа:

- 1) построение математической (концептуальной) модели S' ;
- 2) разработка моделирующего алгоритма и построение имитационной модели S'' ;
- 3) исследование системы S с помощью модели S'' , т.е. проведение имитационных экспериментов, обработка и интерпретация результатов.

Процесс имитационного моделирования не является строго поступательным. Между этапами существует обратная связь, обеспечивающая уточнение, корректировку и учет дополнительной информации при разработке и использовании имитационной модели.

Дадим краткую характеристику каждому из этапов.

Этап 1. На основании изучения содержательного описания системы S , составленного в терминах предметной области, осуществляется переход к математической модели S' , которая представляет собой упрощенное математическое или алгоритмическое описание системы S .

Построение модели S' включает, в свою очередь, пять взаимосвязанных этапов:

- 1) постановка задачи и формулировка целей исследования;
- 2) анализ системы S , т.е. разбиение (декомпозиция) системы на элементы, допускающие удобное математическое или алгоритмическое описание и определение связей между элементами;
- 3) определение параметров, переменных и пространства состояний системы S (характеристик θ, v, u, y, x) и установление областей Θ, V, U, Y, X изменения для каждой характеристики;
- 4) выбор показателей эффективности функционирования модели, т.е. вектора ω ;
- 5) построение концептуальной модели S' системы S и проверка ее адекватности, т.е. использование математических моделей непрерывно-детерминированного, непрерывно-вероятностного, дискретно-детерминированного, дискретно-вероятностного типов и проверка гипотез, предположений и математических соотношений.

Этап 2. На данном этапе осуществляется переход от концептуальной модели S' к моделирующему алгоритму и имитационной модели S'' . Этот переход осуществляется также в пять основных этапов:

- 1) выбор способа имитации, а также вычислительных и программных средств реализации имитационной модели;
- 2) построение логической схемы моделирующего алгоритма;
- 3) алгоритмизация математических моделей, описывающих поведение элементов системы, и связей между ними в рамках выбранного способа имитации;
- 4) разработка имитационной модели — программирование моделирующего алгоритма;
- 5) отладка, тестирование и проверка адекватности имитационной модели.

Этап 3. Использование имитационной модели осуществляется в три этапа:

- 1) планирование имитационных экспериментов;
- 2) проведение имитационных экспериментов;
- 3) обработка, анализ и интерпретация результатов моделирования.

Пусть моделируемая на интервале $[0, T]$ система S характеризуется вектором параметров $\theta \in \Theta \subset \mathbb{R}^m$, причем значениями этого вектора в имитационной модели можно управлять. Кроме того, система S характеризуется вектором показателей эффективности функционирования на интервале $[0, t] \subset [0, T]$

$$\omega = \omega(0, t) = \begin{pmatrix} \omega_1(0, t) \\ \dots \\ \omega_k(0, t) \end{pmatrix} \in \mathbb{R}^k.$$

В результате имитационного моделирования системы S при заданных значениях вектора $\theta' \in \Theta$ и T получают:

1) фазовую траекторию системы S

$$\{x(t) \in X, t \in \mathfrak{X}\},$$

где $\mathfrak{X} \subset [0, T]$ — конечное множество моментов изменения модельного времени, получаемое по принципам « Δx » и « Δt », причем $0, T \in \mathfrak{X}$;

2) значения показателей эффективности $\omega(\theta', t')$ для каждого интервала модельного времени $(0, t')$, $t' \in \mathfrak{X}$, и для всего интервала моделирования в целом:

$$\omega(\theta') = \omega(\theta', T).$$

Имитационные модели могут использоваться для решения различных задач исследования сложных систем.

Задача 1

Оценивание значений показателей эффективности $\omega(\theta')$, $\theta' \in \Theta$.

Задача 2

Оценивание функциональной зависимости $\omega(\cdot)$ от $\theta \in \Theta$.

Задача 3

Сравнение эффективности функционирования системы S для различных значений параметров $\theta', \theta'' \in \Theta$, $\theta' \neq \theta''$.

Задача 4

Оптимизация системы S на множестве Θ , т.е. решение следующей задачи: $R(\theta) \rightarrow \max$ (или \min) по $\theta \in \Theta$, где $R(\cdot)$ — некоторый критерий оптимальности (обобщенный скалярный показатель эффективности) для системы S .

Значения параметров $\theta_1, \dots, \theta_m$ в ходе имитационного эксперимента могут задаваться неоднозначно. Выбор тех или иных значений определяется целью имитационного эксперимента. Существуют различные способы организации имитационных экспериментов, позволяющие получать выборки значений показателей эффективности, которые обладают заданными свойствами.

Процесс задания значений $\theta \in \Theta$, способа формирования выборки значений показателей эффективности $W(\theta)$, обеспечивающих эффективное решение задач исследования сложной системы, называется *планированием имитационных экспериментов*.

Таким образом, для решения задач 1–4 необходим этап планирования имитационного эксперимента, т.е. этап 3.

Имитационный эксперимент с моделью системы S'' , в результате которого для заданных значений $\theta' \in \Theta$ и T имитируется фазовая траектория и вычисляются показатели эффективности, называется *прогоном* имитационной модели.

В результате одного прогона имитационной модели S'' получаем одну реализацию показателя эффективности $\omega^{(1)}(\theta')$, $\theta' \in \Theta$. По одной случайной реализации показателя эффективности нельзя судить об эффективности функционирования системы S . Поэтому на третьем этапе для заданных значений $\theta' \in \Theta$ и T осуществляется $n > 1$ прогонов имитационной модели, в результате которых получается случайная выборка значений показателей эффективности

$$\omega^{(1)}(\theta'), \dots, \omega^{(n)}(\theta'), \quad \theta' \in \Theta,$$

где $\omega^{(i)}(\theta')$ — значение показателя эффективности функционирования системы S для i -го, $i = \overline{1, n}$, прогона имитационной модели.

Для оценивания эффективности функционирования системы по случайной выборке $\{\omega^{(i)}(\theta')\}$, $i = \overline{1, n}$, необходим этап обработки, анализа и интерпретации результатов имитационного моделирования, т.е. третий этап. На данном этапе используются методы статистического анализа данных.

5.5. Планирование имитационных экспериментов

Планирование имитационных экспериментов преследует две основные цели:

- 1) сокращение общего числа экспериментов при соблюдении требований к достоверности и точности их результатов;
- 2) повышение информативности каждого эксперимента.

Поиск плана эксперимента производится в так называемом факторном пространстве. *Факторное пространство* — это множество внешних и внутренних параметров модели, значения которых можно контролировать в процессе подготовки и проведения имитационных экспериментов. Поскольку факторы могут носить как количественный, так и качественный характер, то значения факторов обычно называют *уровнями*. Если при проведении эксперимента уровни факторов можно изменять, то эксперимент называется *активным*, в противном случае — *пассивным*.

Каждый из факторов имеет верхний и нижний уровни, расположенные симметрично относительно некоторого нулевого уровня. Точка в факторном пространстве, соответствующая нулевым уровням всех факторов, называется *центром плана*. *Интервалом варьирования* фактора называется некоторое число j , прибавление которого к нулевому уровню дает верхний уровень, а вычитание — нижний. Как правило, план эксперимента строится относительно одного (основного) выходного скалярного параметра Y , который называется *наблюдаемой переменной*.

Если моделирование используется как инструмент принятия решений, то в роли наблюдаемой переменной выступает показатель эффективности. При этом предполагается, что значение наблюдаемой переменной, полученное в ходе эксперимента, складывается из двух составляющих:

$$y = f(x) + e(x),$$

где $f(x)$ — функция *отклика* (неслучайная функция факторов); $e(x)$ — ошибка эксперимента (случайная величина); x — точка в факторном пространстве (определенное сочетание уровней факторов). Очевидно, что y является случайной переменной, так как зависит от случайной величины $e(x)$.

Дисперсия D_y наблюдаемой переменной характеризует точность измерений и называется *дисперсией воспроизводимости эксперимента*. Она равна дисперсии ошибки эксперимента: $D_y = D_e$. Эксперимент называется *идеальным*, если $D_y = 0$.

Существует два основных варианта постановки задачи планирования имитационного эксперимента.

1. Из всех допустимых необходимо выбрать такой план, который позволил бы получить наиболее достоверное значение функции отклика $f(x)$ при фиксированном числе экспериментов.

2. Выбрать такой допустимый план, при котором статистическая оценка функции отклика может быть получена с заданной точностью при минимальном числе экспериментов.

Решение задачи планирования в первой постановке называется *стратегическим планированием эксперимента*, во второй — *тактическим планированием*.

5.5.1. Стратегическое планирование имитационного эксперимента

Целью методов стратегического планирования является получение максимального объема информации об исследуемой системе в каждом эксперименте или, что то же самое, определение сочетаний уровней внешних и внутренних факторов, при которых может быть получена наиболее полная и достоверная информация о поведении исследуемой системы.

При стратегическом планировании эксперимента должны быть решены две основные задачи:

- 1) идентификация факторов;
- 2) выбор уровней факторов.

Под *идентификацией факторов* понимают их *ранжирование* по степеням влияния на значение наблюдаемой переменной. При этом все факторы делятся на две группы: *первичные* — факторы, которые являются предметом исследования, и *вторичные* — факторы, которые не являются предметом исследования, но влиянием которых нельзя пренебречь.

Выбор *уровней факторов* производится с учетом двух противоречивых требований:

- 1) уровни фактора должны заполнять весь возможный диапазон его изменения;
- 2) общее количество уровней по всем факторам не должно приводить к чрезмерному объему моделирования.

Поиск компромиссного решения, удовлетворяющего этим требованиям, и является задачей стратегического планирования эксперимента.

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется *полным факторным экспериментом*. Если l_i — число уровней i -го фактора, то общее число различных комбинаций уровней для k факторов определяется следующим образом: $N = l_1 \cdot l_2 \cdot \dots \cdot l_k$. Если число уровней для всех факторов одинаково и равно l , то $N = l^k$.

Для полного факторного эксперимента требуются большие временные затраты на подготовку и проведение, что является его недостатком. Поэтому полный факторный эксперимент целесообразно использовать только в том случае, если рассматривается взаимное влияние всех факторов в исследуемой системе.

Если предполагается, что факторы не оказывают взаимного влияния или им пренебрегают, то проводят *частичный факторный эксперимент*. Наиболее часто используются следующие методы построения планов в частичном факторном эксперименте.

1. *Рандомизированный план*. Сочетание уровней для каждого эксперимента задается случайным образом.

2. *Планы с ограничением на рандомизацию (неполноблочный сбалансированный план, латинские квадраты)*. Используются в том случае, когда полная рандомизация невозможна.

3. *Эксперимент с изменением факторов по одному*. Один из факторов «пробегает» все l уровней, а остальные $k-1$ факторов поддерживаются постоянными. Такой план обеспечивает исследование влияния каждого фактора в отдельности. Он требует всего $N = l_1 + l_2 + \dots + l_k$ экспериментов.

4. *Дробный факторный эксперимент*. Используется в том случае, когда реализуются не все сочетания уровней факторов, т.е. из программы

полного факторного эксперимента исключаются уровни факторов, априорно незначимо влияющие на поведение модели.

5. *Симплексные планы.* Насыщенные планы с минимально возможным числом экспериментов (сочетаний уровней и факторов).

5.5.2. Тактическое планирование имитационного эксперимента

Так как имитационное моделирование является чаще всего статистическим экспериментом, то при его проведении необходимо не только получать достоверные результаты, но и обеспечивать его «измерение» с заданной точностью. Точность оценки наблюдаемой переменной характеризуется ее дисперсией. В связи с этим основу методов тактического планирования экспериментов составляют так называемые методы понижения дисперсии.

Объем испытаний (число экспериментов), необходимых для получения оценок наблюдаемой переменной с заданной точностью, зависит от следующих факторов:

- 1) вида распределения наблюдаемой переменной y ;
- 2) коррелированности между собой элементов выборки;
- 3) наличия и длительности переходного режима функционирования моделируемой системы.

При отсутствии перечисленной информации имеется единственный способ повышения точности оценок истинных значений наблюдаемой переменной — многократное повторение *прогонов* модели для каждого сочетания уровней факторов, выбранного на этапе стратегического планирования экспериментов. Такой подход получил название *формирования простой случайной выборки*. В этом случае общее число прогонов модели равно $N_c N_\pi$, где N_c — число сочетаний уровней факторов по стратегическому планированию, N_π — число прогонов модели для каждого сочетания, вычисленное при тактическом планировании.

Основной подход к определению необходимого объема испытаний состоит в следующем. Если случайные значения наблюдаемой переменной не коррелированы и их распределения не изменяются от прогона к прогону, то выборочное среднее можно считать нормально распределенным. В этом случае число прогонов N_π , необходимых для того, чтобы истинное среднее y лежало в интервале $y \pm b$ с вероятностью $(1 - \alpha)$, определяется следующим образом:

$$N_\pi = \frac{Z^2 D_y}{b^2}, \quad (5.1)$$

где Z — значение нормированного нормального распределения при заданном уровне значимости $\frac{\alpha}{2}$; D_y — дисперсия; b — доверительный интервал.

Если значение дисперсии D_y до начала эксперимента неизвестно, то выполняется пробная серия из L прогонов и вычисляется на ее основе выборочная дисперсия

$$D_y = \frac{1}{L-1} \sum_{i=1}^L (y_i - y_L)^2,$$

где y_L — выборочное среднее по результатам L прогонов. Полученное значение D_y подставляется в (5.1) и определяется предварительная оценка числа прогонов N_{Π} .

Если наблюдаемая переменная является вектором, то оценку необходимого числа прогонов необходимо определять отдельно для каждого компонента. Наибольшее из полученных значений принимается в качестве числа прогонов N_{Π} .

Основной недостаток методов планирования, основанных на использовании простой случайной выборки, состоит в медленной сходимости выборочных средних к истинным средним с ростом объема выборки N_{Π} (пропорционально значению $\sqrt{N_{\Pi}}$). Это приводит к необходимости использования методов уменьшения ошибок, не приводящих к увеличению N_{Π} . Такие методы называются *методами понижения дисперсии*. Они делятся на три группы:

- 1) *активные*, предусматривающие формирование выборки специальным образом;
- 2) *пассивные*, применяемые после того, как выборка уже сформирована;
- 3) *косвенные*, в которых для получения оценок наблюдаемой переменной используются значения некоторых вспомогательных величин.

5.6. Обработка и анализ результатов моделирования

Решения, принимаемые по результатам имитационного моделирования, могут быть конструктивными только при выполнении двух условий:

- 1) полученные результаты обладают требуемой точностью и достоверностью;
- 2) исследователь способен правильно интерпретировать полученные результаты и знает, каким способом они могут быть использованы.

5.6.1. Оценка качества имитационной модели

Оценка качества модели является завершающим этапом ее разработки и преследует две цели:

- 1) проверить соответствие модели ее назначению;

2) оценить достоверность и статистические характеристики результатов, получаемых при проведении экспериментов с использованием построенной модели.

При аналитическом моделировании достоверность результатов определяется двумя основными факторами:

1) корректным выбором математического аппарата, используемого для описания исследуемой системы;

2) погрешностями методов, используемых для решения задач.

При имитационном моделировании на достоверность результатов влияет целый ряд дополнительных факторов, основными из которых являются:

1) моделирование случайных факторов, основанное на использовании датчиков случайных чисел, которые могут вносить «искажения» в поведение модели;

2) наличие нестационарного режима работы модели;

3) несогласованность погрешностей методов, используемых в рамках одной модели;

4) зависимость результатов моделирования от плана эксперимента;

5) необходимость синхронизации отдельных компонентов модели;

6) наличие модели *рабочей нагрузки* (моделей внешних воздействий, оказывающих влияние на эффективность применения данной системы в рамках проводимого эксперимента), качество которой зависит, в свою очередь, от тех же факторов.

Качество имитационной модели характеризуется тем, в какой степени она обладает *целевыми свойствами*: адекватностью; устойчивостью; чувствительностью.

Адекватность — степень соответствия модели реальному объекту или процессу, для которых она строится. При этом, поскольку модель строится, как правило, для исследования определенного подмножества свойств объекта, можно считать, что под адекватностью модели понимается ее соответствие целям моделирования. Одним из наиболее распространенных способов проверки такого соответствия является использование методов математической статистики, а именно проверки гипотез на основе некоторых статистических критериев. При этом следует иметь в виду, что статистические критерии не могут доказать ни одной гипотезы, они могут лишь указать на отсутствие опровержения. Процедура оценки адекватности модели основана на сравнении измерений на реальной системе и результатов экспериментов на модели и может проводиться различными способами. Наиболее распространенными из них являются следующие: по средним значениям откликов модели и системы; по дисперсиям отклонений откликов модели от среднего значения откликов системы; по максимальному значению относительных отклонений откликов модели от откликов системы.

П р и м е р

Рассмотрим процедуру проверки гипотезы о близости среднего значения наблюдаемой переменной Y среднему значению отклика реальной системы Y^* .

В результате N_O опытов на реальной системе получают множество значений (выборку) Y^* . Выполнив N_M экспериментов на модели, также получают множества значений наблюдаемой переменной Y . Затем вычисляются оценки математического ожидания и дисперсии откликов модели и системы, после чего выдвигается гипотеза о близости средних значений величин Y^* и Y (в статистическом смысле). Основой для проверки гипотезы является t -статистика (*распределение Стьюдента*). Ее значение, вычисленное по результатам испытаний, сравнивается с критическим значением $t_{кр}$, взятым из таблицы. Если выполняется неравенство $t < t_{кр}$, то гипотеза принимается.

Заметим, что статистические методы применимы только в том случае, если оценивается адекватность модели существующей системе. Для проектируемой системы в качестве эталонного объекта используется концептуальная модель. Тогда проверка адекватности программно-реализованной модели заключается в проверке того, насколько корректно она отражает концептуальную модель.

Устойчивость модели — это способность сохранять адекватность при исследовании эффективности системы на всем возможном диапазоне рабочей нагрузки, а также при внесении изменений в конфигурацию системы. Устойчивость результатов моделирования может быть оценена методами математической статистики, при этом проверяется гипотеза относительно свойств некоторого множества элементов, называемого *генеральной совокупностью*, оценивая свойства некоторого подмножества генеральной совокупности (*выборки*). В генеральной совокупности обычно интересует некоторый признак, который обусловлен случайностью и может иметь качественный или количественный характер. При этом устойчивость результатов моделирования можно рассматривать как признак, подлежащий оценке. Для проверки гипотезы об устойчивости результатов может быть использован *критерий Вилкоксона*, который служит для проверки того, относятся ли две выборки к одной и той же генеральной совокупности, т.е. обладают ли они одним и тем же статистическим признаком.

При статистической оценке устойчивости модели соответствующая гипотеза H может быть сформулирована следующим образом: при изменении входной (рабочей) нагрузки или структуры имитационной модели закон распределения результатов моделирования остается неизменным. Проверку указанной гипотезы H проводят для двух выборок $X = (x_1, \dots, x_n)$ и $Y = (y_1, \dots, y_m)$, полученных для различных значений рабочей нагрузки. При этом относительно законов распределения

X и Y никаких предположений не делается. Значения обеих выборок упорядочиваются вместе по возрастанию. Затем анализируется взаимное расположение x_i и y_j . В случае $y_j < x_i$ говорят, что пара значений (x, y) образует инверсию.

П р и м е р

Если при $n = m = 3$ после упорядочивания получилась последовательность значений $y_1, x_1, y_3, x_2, y_2, x_3$, то получаются следующие инверсии: (x_1, y_1) , (x_2, y_1) , (x_2, y_3) , (x_3, y_1) , (x_3, y_2) , (x_3, y_3) .

Подсчитывается полное число инверсий U . Если гипотеза верна, то U не должно сильно отклоняться от своего математического ожидания M :

$$M = \frac{nm}{2}.$$

От гипотезы отказываются, если $|U - M| > U_{кр}$ (значение $U_{кр}$ определяется по таблице для заданного уровня значимости).

Оценку чувствительности модели к изменению параметров рабочей нагрузки и внутренних параметров самой системы проводят по каждому параметру X_k в отдельности. Основана она на том, что обычно диапазон возможных изменений параметра известен. Одна из наиболее простых и распространенных процедур оценивания параметров состоит в следующем.

1. Вычисляется величина относительного среднего приращения параметра X_k :

$$\Delta X_k = \frac{2(X_{k \max} - X_{k \min})}{X_{k \max} + X_{k \min}} \cdot 100\%.$$

2. Проводится пара модельных экспериментов при значениях $X_{k \max}$, $X_{k \min}$ и средних фиксированных значениях остальных параметров. Определяются значения отклика модели $Y_{k_1} = f(X_{k \max})$ и $Y_{k_2} = f(X_{k \min})$.

3. Вычисляется относительное приращение наблюдаемой переменной Y_k :

$$\Delta Y_k = \frac{2|Y_{k_1} - Y_{k_2}|}{Y_{k_1} + Y_{k_2}} \cdot 100\%.$$

В результате для k -го параметра получается пара значений $(\Delta X_k, \Delta Y_k)$, характеризующая чувствительность модели по этому параметру. Аналогично формируются пары для остальных параметров модели, которые образуют множество пар $\{(\Delta X_k, \Delta Y_k)\}$ для всех значений k .

Данные, полученные при оценке чувствительности модели, могут быть использованы, в частности, при планировании экспериментов:

большее внимание должно уделяться тем параметрам, по которым модель является более чувствительной.

Если в результате проведенной оценки качества модели оказалось, что ее целевые свойства неудовлетворительны, то необходимо выполнить ее *калибровку*, т.е. коррекцию с целью приведения в соответствие предъявляемым требованиям. Процедура калибровки, как правило, состоит из трех шагов, каждый из которых является итеративным.

1. *Сравнение выходных распределений*. Цель — оценка адекватности модели. Критерии сравнения могут быть различны, в частности может использоваться величина разности между средними значениями откликов модели и системы. Установление различий на этом шаге основано на внесении глобальных изменений (введении новых процессов, изменении типов событий и т.д.).

2. *Балансировка модели*. Основная задача — оценка устойчивости и чувствительности модели. В результате производятся, как правило, локальные изменения (изменение некоторых законов распределения моделируемых случайных величин и т.д.).

3. *Оптимизация модели*. Цель — обеспечение требуемой точности результатов, которая может достигаться улучшением качества датчиков случайных чисел, снижением влияния переходного режима, применением специальных методов понижения дисперсии.

5.6.2. Подбор параметров распределений

При имитационном моделировании подбор законов распределений выполняется на основе статистических данных, полученных в ходе эксперимента. В основе такой процедуры лежит проверка статистических гипотез.

Статистическая гипотеза — это утверждение относительно значений одного или более параметров распределения некоторой величины или относительно самой формы распределения. Обычно выбирают две исходные гипотезы: основную — H_0 и альтернативную ей — H_1 .

Статистическая проверка гипотезы — это процедура выяснения, следует ли принять основную гипотезу H_0 или отвергнуть ее. Обычно выбирают некоторое значение α , которое называется *уровнем значимости*. Для выбранного значения α определяется так называемая критическая область B , удовлетворяющая условию

$$P \{Z \in B \mid H_0 \text{ верна}\} \leq \alpha,$$

где Z — критерий, представляющий собой некоторую функцию от выборки из генеральной совокупности некоторой случайной величины. Проверка гипотезы состоит в том, что производится выборка, на основании которой вычисляется z — частное значение критерия Z . Если $z \in B$, то от гипотезы H_0 отказываются, в противном случае

говорят, что полученные наблюдения не противоречат принятой гипотезе. При этом вид самого закона распределения достаточно часто определяется с помощью *гистограммы относительных частот*.

Для наиболее часто используемых статистических гипотез разработаны критерии, позволяющие проводить их проверку с наибольшей достоверностью.

t-критерий служит для проверки гипотезы о равенстве средних значений нормально распределенных случайных величин x и y в предположении, что дисперсии их равны (но неизвестны), при этом сравниваемые выборки из генеральной совокупности величин x и y могут иметь разный объем $n_1 \neq n_2$. В качестве критерия используют величину

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{n_1^{-1}D_x + n_2^{-1}D_y}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

где \bar{x}, \bar{y} — значения математического ожидания; D_x, D_y — дисперсии случайных величин x и y .

Величина T подчиняется *t-распределению Стьюдента*, а критическое значение $t_{кр}$ для *t-критерия* определяется по таблице для выбранного значения α и числа степеней свободы $k = n_1 + n_2 - 2$. Если вычисленное значение T удовлетворяет неравенству $T \geq t_{кр}$, то гипотезу H_0 отвергают.

Заметим, что по отношению к предположению о «нормальной распределенности» величин x и y *t-критерий* не очень чувствителен. Его можно применять, если распределения случайных величин не имеют нескольких вершин и не слишком асимметричны.

F-критерий служит для проверки гипотезы о равенстве дисперсий D_x и D_y при условии, что x и y распределены нормально. Гипотезы такого рода имеют большое значение в технике, так как дисперсия есть мера таких характеристик, как погрешности измерительных приборов, точность технологических процессов, точность наведения при стрельбе и т.д.

В качестве контрольной величины используется отношение дисперсий $F = D_x / D_y$ (или $F = D_y / D_x$ — большая дисперсия должна быть в числителе). Величина F подчиняется *F-распределению Фишера* с (m_1, m_2) степенями свободы ($m_1 = n_1 - 1, m_2 = n_2 - 1$).

Проверка гипотезы состоит в следующем. Для величины $\alpha / 2$ и величин m_1, m_2 по таблице *F-распределения* выбирают значение F_{α, m_1, m_2} . Если F , вычисленное по выборке, больше этого критического значения, то гипотеза должна быть отклонена с вероятностью ошибки α .

Критерии согласия — это критерии, с помощью которых проверяют, удовлетворяет ли рассматриваемая случайная величина данному закону распределения.

Рассмотрим проверку гипотезы H_0 о том, что $F_y(y) = F_0(y)$, где $F_y(y)$ — истинное распределение случайной величины y ; $F_0(y)$ — гипотетическое распределение.

Критерий Пирсона (χ^2). Проверка гипотезы производится следующим образом.

1. Область значений случайной величины y произвольно разбивается на k непересекающихся множеств (классов).
2. В результате n экспериментов формируется выборка (y_1, \dots, y_n) .
3. Вычисляется величина χ^2 :

$$\chi^2 = \left(\sum_{i=1}^k \frac{n_i^2}{nP_i} \right) - n,$$

где n_i^2 — число значений y , попавших в класс i ; P_i — теоретическая вероятность (для $F_0(y)$) попадания значения y в класс i .

4. По таблице χ^2 -распределений находят критическое значение χ_0^2 для уровня значимости α и $m = k - 1$ степеней свободы. Если $\chi^2 \geq \chi_0^2$, то гипотеза отвергается.

Критерий Колмогорова–Смирнова. Проверка гипотезы производится следующим образом.

1. Выборку (y_1, \dots, y_n) упорядочивают по возрастанию и строят следующую эмпирическую функцию распределения:

$$F_y(y) = \begin{cases} 0, & -\infty < y < y_1, \\ i/n, & y_i \leq y < y_{i+1}, \quad i = 1, n-1, \\ 1, & y_n \leq y < \infty. \end{cases}$$

2. По значениям этой функции вычисляется величина

$$\lambda = \max_{-\infty < y < +\infty} |F_y(y) - F_0(y)|.$$

3. Гипотеза $H_0 : F_y(y) = F_0(y)$ отвергается, если $\lambda < \frac{\lambda_\alpha}{\sqrt{n}}$, где величина λ_α выбирается из таблицы по уровню значимости α .

П р и м е р

Пусть имеется выборка (x_1, x_2, \dots, x_n) случайной величины X объема n . Вычислим оценки среднего значения и дисперсии

$$m = \frac{1}{n} \sum_{i=1}^n x_i, \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m)^2.$$

Тогда гипотеза о нормальности случайной величины X соответствует тому, что ее функция распределения имеет вид

$$F_0(x) = \Phi\left(\frac{x - m}{s}\right),$$

где

$$\Phi\left(\frac{x-m}{s}\right) = \int_{-\infty}^x \frac{1}{s\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2s^2}} dx.$$

Далее обычно поступают следующим образом. Задают значение α и из таблицы выбирают соответствующее значение λ_α (например $\alpha = 0,05$, $\lambda_\alpha = 1,36$). На одних осях строят графики функций

$$y(x) = F_0(x), \quad y(x) = F_0(x) + \lambda, \quad y(x) = F_0(x) - \lambda$$

где $\lambda = \frac{\lambda_\alpha}{\sqrt{n}}$. Затем на этом же графике строят эмпирическую функцию распределения $F_y(x)$. Если график эмпирической функции распределения нигде не выходит за пределы полосы $F_0(x) \pm \lambda$, то можно принимать гипотезу H_0 , т.е. считать, что опытные данные не противоречат тому, что функция распределения исследуемой случайной величины равна $F_0(x)$. Если график $F_y(x)$ где-нибудь выйдет за пределы этой полосы, то гипотеза H_0 отвергается.

5.6.3. Оценка влияния и взаимосвязи факторов

Как правило, количественная оценка степени влияния того или иного фактора на значение наблюдаемой переменной (показателя эффективности) вызывает значительную сложность, особенно при наличии взаимного влияния факторов. Наиболее простой и доступный способ решения этой проблемы состоит в использовании результатов оценки чувствительности модели.

Однофакторный дисперсионный анализ. Его суть сводится к определению влияния на результат моделирования одного выбранного фактора.

Формально постановка задачи однофакторного дисперсионного анализа состоит в следующем. Пусть интересующий фактор x имеет l уровней. Для каждого из них получена выборка значений наблюдаемой переменной $y: y_j(1), y_j(2), \dots, y_j(l)$, $j = \overline{1, n}$, n — объем выборки.

Необходимо проверить гипотезу H_0 о равенстве средних значений выборок, т.е. о независимости значений y от значений исследуемого фактора x . Уравнение однофакторного дисперсионного анализа имеет вид

$$y_{ij} = m + a_i + e_{ij},$$

где y_{ij} — j -е значение y в i -й серии экспериментов; m — генеральное среднее случайной величины y (т.е. среднее значение наблюдаемой переменной, обусловленное ее «сущностью»); a_i — неизвестный параметр, отражающий влияние фактора x («эффект» i -го значения фактора x); e_{ij} — ошибка измерения y .

Для проверки гипотезы H_0 используют F-критерий и переходят от проверки значимости различий средних к проверке значимости различий двух дисперсий: генеральной (обусловленной погрешностями измерений) — D_0 и факторной (обусловленной изменением фактора x) — D_x . Значение F-критерия вычисляется как отношение D_x / D_0 или D_0 / D_x (в числителе должна стоять большая из дисперсий). Критическое значение $F_{кр}$ находят по таблице F-распределений для заданного уровня значимости и числа степеней свободы $m_1 = n_1 - 1$, $m_2 = n_2 - 1$. Если $F > F_{кр}$, то гипотезу H_0 отвергают, т.е. различия являются значимым (фактор x влияет на значения y).

Многофакторный дисперсионный анализ позволяет оценивать влияние на наблюдаемую переменную произвольного числа факторов, т.е. позволяет выбрать из группы факторов, участвующих в эксперименте, те, которые действительно влияют на его результат.

П р и м е р

Рассмотрим методику проведения многофакторного дисперсионного анализа применительно к частичному факторному эксперименту, проводимому в соответствии с латинским планом.

Пусть в эксперименте рассматриваются один первичный фактор и два вторичных, каждый из которых имеет n уровней, т.е. объем испытаний равен $N = n^2$. Обозначим через y_{ijk} результат эксперимента при условии, что фактор a находится на уровне i , фактор b — на уровне j , фактор c — на уровне k . Множество значений, которые может принимать упорядоченная тройка (i, j, k) , обозначим через L .

В этом случае уравнение дисперсионного анализа выглядит следующим образом:

$$y_{ijk} = m + a_i + b_j + g_k + \epsilon_{ijk},$$

где m — генеральное среднее случайной величины y ; a_i , b_j , g_k — неизвестные параметры («эффекты» соответствующих факторов). Решение задачи дисперсионного анализа заключается в проверке гипотез о независимости результатов измерений от факторов a, b, g :

$$H_a: a_i = 0, \quad i = \overline{1, n};$$

$$H_b: b_j = 0, \quad j = \overline{1, n};$$

$$H_g: g_k = 0, \quad k = \overline{1, n}.$$

Оценки параметров m , a_i , b_j , g_k находят по методу наименьших квадратов из условия минимума функционала:

$$Ss = \sum_L (y_{ijk} - m + a_i + b_j + g_k)^2.$$

Затем по каждому фактору вычисляется F-статистика. Величина F есть мера потерь при принятии гипотезы H_0 . Чем больше F , тем

хуже модель, отвергающая влияние соответствующего фактора. Таким образом, если $F > F_{кр}$, то гипотеза отвергается; $F_{кр}$ определяется по таблице для заданного уровня значимости.

Заметим, что дисперсионный анализ может использоваться для оценки влияния факторов, имеющих как количественный, так и качественный характер, поскольку в уравнении дисперсионного анализа фигурируют не сами факторы, а их «эффекты». В том случае, если все факторы носят количественный характер, взаимосвязь между ними и наблюдаемой переменной может быть описана с помощью уравнения регрессии.

Корреляционный и регрессионный анализ. Два близких метода, которые обычно используются совместно для исследования взаимосвязи между двумя или более непрерывными переменными.

Методы корреляционного анализа позволяют делать статистические выводы о степени зависимости между переменными. Величина линейной зависимости одной переменной от другой измеряется посредством *простого коэффициента корреляции*, величина зависимости одной переменной от нескольких — посредством *множественного коэффициента корреляции*. В корреляционном анализе используется также понятие *частного коэффициента корреляции*, который измеряет линейную взаимосвязь между двумя переменными без учета влияния других переменных.

П р и м е р

Рассмотрим простой корреляционный анализ, т.е. метод определения взаимосвязи между двумя переменными x и y .

Независимо от способа получения выборки имеется два предварительных шага для определения существования и степени зависимости между x и y . Первый шаг заключается в графическом отображении точек (x_i, y_i) ($i = \overline{1, n}$, n — объем испытаний) на плоскости (x, y) , т.е. в построении *диаграммы рассеяния*. Анализируя диаграмму рассеяния, можно решить, допустимо ли предположение о линейной зависимости между x и y , т.е. о неравенстве нулю коэффициента корреляции r_{xy} .

Если предполагается, что $r_{xy} \neq 0$, то вычисляется его точное значение. Чем больше по абсолютному значению величина r_{xy} , тем сильнее линейная зависимость между переменными. При $|r_{xy}| = 1$ имеет место функциональная линейная зависимость между x и y вида $y = b_0 + b_1x$, причем если $r_{xy} = +1$, то говорят о положительной корреляции, т.е. большие значения одной переменной соответствуют большим значениям другой; при $r_{xy} = -1$ имеет место отрицательная корреляция; при $0 < r_{xy} < 1$ вероятна либо линейная корреляция с рассеянием, либо нелинейная корреляция.

Заметим, что при анализе результатов имитационного моделирования необходимо иметь в виду, что если даже удалось установить тесную взаимосвязь между двумя переменными, это еще не является прямым доказательством их причинно-следственной связи. Возможно, имеет место стохастическая зависимость, обусловленная, например, коррелированностью последовательностей псевдослучайных чисел, используемых в имитационной модели.

Поэтому результаты корреляционного анализа целесообразно уточнить, проведя регрессионный анализ, который позволяет решать две задачи:

1) устанавливать наличие возможной причинной связи между переменными;

2) предсказывать значения зависимой переменной по значениям независимых переменных, что особенно важно в тех случаях, когда прямые измерения зависимой переменной затруднительны.

Если предполагается, что есть линейная зависимость между x и y , то она может быть описана уравнением вида

$$y_i = b_0 + b_1 x_i + \varepsilon_i, \quad i = \overline{1, n},$$

которое называется *простой линейной регрессией* y по x . Величины b_0, b_1 являются неизвестными параметрами, а ε_i — случайными ошибками испытаний.

Цель регрессионного анализа — найти наилучшие в статистическом смысле оценки параметров b_0 и b_1 , т.е. \hat{b}_0 и \hat{b}_1 . Тогда можно найти оценку переменной y при $x = x_i$: $\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_i$. Разница между наблюдаемым и оцененным значением y при $x = x_i$, т.е. $d_i = y_i - \hat{y}_i$, называется *отклонением (остатком)*. Величины отклонений могут быть использованы для проверки адекватности полученной модели, для чего строятся графики $d = f(y)$ или $d = f(x)$, и по их виду делается предварительное заключение о степени адекватности регрессионной модели.

В случае нескольких независимых переменных имеет место *множественная линейная регрессия*

$$y = b_0 + b_1 x_1 + \dots + b_k x_k + \varepsilon$$

и для отыскания оценок b_i , $i = \overline{1, k}$, используется метод наименьших квадратов.

В случае нелинейной регрессии для построения регрессионной модели также используется метод наименьших квадратов, но для отыскания оценок b_j , $j = \overline{0, k}$, строится система нелинейных уравнений, которая решается итерационным методом.

Замечание

Эффективное использование процедур статистического анализа экспериментальных данных возможно только при наличии соответствующих инструментальных средств.

5.7. Достоинства и недостатки имитационного моделирования

Имитационное моделирование имеет следующие отличительные особенности:

1) при создании имитационных моделей сложных систем законы функционирования всей системы в целом могут быть неизвестны, достаточно знания алгоритмов, описывающих поведение отдельных элементов системы и связей между ними;

2) в имитационной модели связи между параметрами и характеристиками системы выявляются, а значения исследуемых характеристик определяются в ходе имитационного эксперимента на ЭВМ.

Областью применения имитационного моделирования является широкий класс систем практически любой сложности, используется оно в случаях, когда в силу сложности аналитической модели ее практическое использование невозможно.

Имитационное моделирование обладает следующими достоинствами:

- это часто единственно возможный метод исследования систем в силу их сложности;
- возможность исследования системы на различных уровнях ее детализации, определяемых целью исследования;
- возможность исследования динамики взаимодействия элементов системы во времени и пространстве параметров системы;
- возможность оценивания характеристик системы в определенные моменты времени.

Недостатки имитационного моделирования:

- дороговизна, разработка хорошей имитационной модели часто требует больших временных затрат;
- результаты имитационного моделирования обладают меньшей степенью общности по сравнению с аналитическим моделированием и не позволяют выявить общие закономерности функционирования классов систем;
- не существует надежных методов оценки адекватности имитационных моделей.

6. АГРЕГАТИВНОЕ МОДЕЛИРОВАНИЕ

Существует подход к построению математических моделей сложных систем, позволяющий облегчить процесс разработки и использования имитационных моделей исследуемых систем. Этот подход получил название *агрегативного моделирования* (А-моделирования).

Сущность А-моделирования состоит в следующем. При построении агрегативной модели (А-модели) сложная система разбивается на конечное число взаимосвязанных элементов (подсистем). Если некоторые из полученных подсистем оказываются, в свою очередь, достаточно сложными, то процесс их разделения продолжается до тех пор, пока не будут выделены элементы, допускающие стандартное математическое описание, называемое *агрегатом*. Таким образом, *под агрегатом понимается некоторое абстрактное математическое описание моделей различного типа* (детерминированных, вероятностных, непрерывных, дискретных).

Представление математических моделей сложных систем в виде А-моделей позволяет за счет использования одной и той же формы описания моделей различных типов унифицировать процесс построения имитационной модели и процесс обработки результатов имитационного моделирования для сложной системы, состоящей из большого числа разнотипных элементов.

П р и м е р

А-модели АСУ — АСУ предприятием, АСУ крупным аэропортом, АСУ нефтеперерабатывающим комплексом.

6.1. Основные характеристики агрегативной модели

Описание А-модели включает следующие характеристики.

1. Вектор управляющих сигналов (входных воздействий)

$$u = u(t) \in U \subset R^{n_1},$$

поступающих в случайные моменты времени $t_1 < t_2 < \dots$. В общем случае $u(t_1), u(t_2), \dots$ — случайная последовательность.

2. Вектор параметров

$$\theta \in \Theta \subset R^m,$$

от которого могут зависеть состояния x или их вероятностные характеристики.

3. Вектор воздействий внешней среды

$$v = v(t) \in V \subset R^{n_2},$$

поступающих в случайные моменты времени $t'_1 < t'_2 < \dots$. В общем случае $v(t'_1), v(t'_2), \dots$ образуют случайную последовательность с некоторым вероятностным законом распределения.

4. Вектор состояний

$$x \in X \subset R^{n_3},$$

где X — фазовое пространство агрегата. В каждый момент времени состояние А-модели меняется, образуя фазовую траекторию

$$x = x(t), \quad t \in [t_0, T].$$

5. Начальное состояние. В начальный момент времени ($t = t_0$) состояние А-модели есть случайный вектор $x(t_0) = x^0$, называемый *начальным состоянием* агрегата и имеющий закон распределения $L\{x^0\}$.

Оператор переходов, который определяет состояние агрегата для произвольного момента времени $t > 0$, $t \in [t_0, T]$, т.е. состояние агрегата в момент t определяется *оператором переходов* $H(\cdot)$:

$$x(t) = H(x^0, v^0, u^0, t),$$

где v^0, u^0 — начальные воздействия внешней среды и управление, $H(\cdot)$ — в общем случае случайная функция, т.е. для фиксированного t $x(t)$ — случайный вектор, распределение вероятностей которого зависит от $H(\cdot)$.

7. Вектор выходных сигналов

$$y = y(t) \in Y \subset R^{n_4},$$

где $y(t) = G(x(t), u(t))$, $t \in [t_0, T]$, $G(\cdot)$ — некоторая функция, называемая *оператором выходов*.

Связь описанных выше характеристик можно проиллюстрировать графически (рис. 6.1).

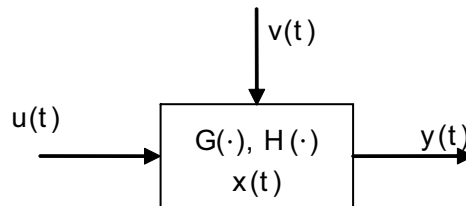


Рис. 6.1

6.2. Математическое описание процесса функционирования агрегата

Будем предполагать, что за конечный интервал времени на агрегат A поступает конечное число воздействий внешней среды, управляющих сигналов и выдается конечное число выходных сигналов. При этом происходит изменение состояния агрегата. Множество всех возможных состояний агрегата будем разделять на подмножества особых состояний и неособых состояний.

Принято говорить, что A в момент времени $t^* \in [t_0, T]$ находится в *особом состоянии* $x(t^*) \in X$, если t^* — либо момент поступления воздействия внешней среды $v(t^*) \in V$, либо момент поступления управляющего сигнала $u(t^*) \in U$, либо момент выдачи выходного сигнала $y(t^*) \in Y$. Все остальные состояния A называются *неособыми состояниями*.

Пусть $x(t^*)$, $t^* \in [t_0, T]$, — некоторое особое состояние; u^S , v^τ — последние перед моментом времени t^* управляющий сигнал и воздействие внешней среды, ($u^S \in U$, $v^\tau \in V$).

Обозначим

$$x(t+0) = \lim_{\delta \rightarrow +0} x(t+\delta),$$

т.е. $x(t+0)$ — сколь угодно близкий к t будущий момент времени.

В процессе функционирования агрегата происходят изменения его состояния в соответствии с законом функционирования, определяемым оператором $H(\cdot)$. В моменты наступления особых состояний функция $H(\cdot)$ изменяется скачкообразно, т.е. функция $H(\cdot)$ имеет разрывы. Во временных интервалах между моментами наступления особых состояний функция $H(\cdot)$ является непрерывной.

Пусть $t^* \in [t_0, T]$ — момент наступления одного из особых состояний. Тогда в момент времени $(t+0)$ состояние агрегата определяется следующим образом:

$$x(t^*+0) = \begin{cases} H_1(x(t^*), v, u^S), & \text{если } t^* \text{ — момент поступления } v, \\ H_2(x(t^*), v^\tau, u), & \text{если } t^* \text{ — момент поступления } u, \\ H_3(x(t^*), v, u), & \text{если } t^* \text{ — момент одновременного} \\ & \text{поступления } v, u, \\ H_4(x(t^*), v^\tau, u^S), & \text{если } t^* \text{ — момент выдачи } y. \end{cases} \quad (6.1)$$

Пусть $t^{**} > t^*$ — момент наступления следующего особого состояния. В интервале между особыми состояниями функция $H(\cdot)$ является непрерывной по t , т.е.

$$x(t) = H_0(x(t^* + 0), v^{\tau'}, u^{S'}, t), \quad t \in (t^* + 0, t^{**}), \quad (6.2)$$

где $v^{\tau'}$, $u^{S'}$ — последние перед моментом времени t^{**} управляющий сигнал и воздействие внешней среды.

Операторы H_0 , H_1 , H_2 , H_3 , H_4 в общем случае являются случайными и определяются некоторыми законами распределения вероятностей.

Выделим в фазовом пространстве X семейство критических областей. Критическая область $X_y \in X$ для $y \in Y$ — это область в пространстве X , при попадании фазовой траектории агрегата в которую выдается выходной сигнал y . Условие попадания фазовой траектории в X_y в момент времени $\tau \in [t_0, T]$ имеет вид

$$x(\tau) \in X_y, \quad x(\tau - \varepsilon) \notin X_y, \quad \varepsilon > 0. \quad (6.3)$$

При этом выходной сигнал определяется выражением

$$y = y(\tau) = G(x(\tau), u^S),$$

где u^S — последний перед моментом τ управляющий сигнал.

Опишем процесс функционирования A на интервале времени $[t_0, T]$.

Пусть

1) в начальный момент времени $t_0 = 0$ задано:

- начальное состояние x^0 ,
- начальное воздействие внешней среды v^0 ,
- начальное управляющее воздействие (управление) u^0 ;

2) в случайные моменты времени $t'_1 \leq t''_1 \leq t'''_1 \in [t_0, T]$ в агрегат поступило соответственно:

- первое воздействие внешней среды $v^1 = v(t'_1) \in V$,
- первое управляющее воздействие $u^1 = u(t''_1) \in U$,
- произошла выдача выходного сигнала $y^1 = y(t'''_1) \in Y$.

Рассмотрим иллюстрацию моментов перехода агрегата в особые состояния (рис. 6.2).

Моменты времени t'_1 , t''_1 , t'''_1 делят интервал моделирования $[0, t''']$ на фрагменты, определяющие поведение агрегата в интервалах времени $(0, t')$, (t', t'') , (t'', t''') .

Фрагмент 1. Закон функционирования агрегата определяется формулой (6.2)

$$x(t) = H_0(x^0, v^0, u^0, t), \quad t \in (0, t'_1)$$

и является непрерывным.

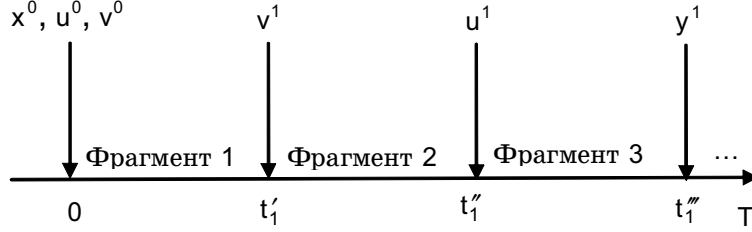


Рис. 6.2

В момент t'_1 :

- если $t'_1 \neq t''_1$, то поступает воздействие внешней среды v^1 и агрегат попадает в особое состояние (изменение состояния происходит скачкообразно), при этом закон функционирования агрегата, согласно (6.1), описывается выражением

$$x(t'_1 + 0) = H_1(x(t'_1), v^1, u^0);$$

- если $t'_1 = t''_1$, т.е. в агрегат одновременно поступают v^1 и u^1 , то закон функционирования агрегата описывается выражением

$$x(t'_1 + 0) = H_3(x(t'_1), v^1, u^1)$$

и имеет место фрагмент 3.

Фрагмент 2. Закон функционирования агрегата определяется формулой

$$x(t) = H_0(x(t'_1), v^1, u^0, t), \quad t \in (t'_1, t''_1),$$

и является непрерывным.

В момент t''_1 :

- если $t''_1 \neq t'''_1$, то поступает управляющий сигнал u^1 и состояние агрегата изменяется скачком по закону

$$x(t''_1 + 0) = H_2(x(t''_1), v^1, u^1);$$

- если $t''_1 = t'''_1$, т.е. в агрегат в один и тот же момент времени поступает управляющий сигнал u^1 , а фазовая траектория $x(t)$ попадает в критическое подмножество X_{y^1} (выполняется условие (6.3)), тогда состояние агрегата изменяется в данный момент дважды по закону

$$x(t''_1 + 0) = H_4(x(t''_1 + 0), v^1, u^1) = H_4(H_2(x(t''_1), v^1, u^1),$$

при этом выдается выходной сигнал

$$y^1 = G(x(t''_1), v^1, u^1, t), \quad t \in (t''_1, t'''_1),$$

и имеет место фрагмент 4, который наступит после фрагмента 3 при моделировании для $t > t'''_1$.

Фрагмент 3. Функционирование агрегата описывается непрерывным законом (6.2), т.е.

$$x(t) = H_0(x(t''_1), v^1, u^1, t), \quad t \in (t''_1, t'''_1).$$

В момент t_1''' фазовая траектория $x(t)$ попадает в критическую область X_{y_1} , т.е. выполняется условие (6.3), следовательно, состояние агрегата изменяется скачкообразно и выдается сигнал (6.4).

Таким же образом процесс функционирования агрегата продолжается далее до достижения момента завершения времени моделирования T .

6.3. Пример агрегативного моделирования

Рассмотрим систему S , описывающую производственный участок обработки деталей, который состоит из четырех обрабатывающих центров (ОЦ) и предназначен для обработки деталей двух типов. Процесс обработки состоит из двух последовательных операций, выполняемых для деталей:

- 1-го типа — на ОЦ 1, 2, 3;
- 2-го типа — на ОЦ 1, 2, 4.

Последовательности 1, 2, 3; 1, 2, 4 определяют технологические маршруты обработки деталей 1-го и 2-го типов. При каждом ОЦ имеется одно место для временного хранения детали. Интервалы поступления и выполнения операций, а также тип детали являются случайными величинами с заданными законами распределения.

Для того чтобы записать математическую модель системы, введем обозначения:

v — тип детали: $P\{v=1\} = p_1$, $P\{v=2\} = 1 - p_1$;

τ — интервал между поступлениями деталей, $M\{\tau\} = \mu$;

τ_k — время выполнения k -й операции, $M\{\tau_k\} = \mu_k$;

Законы распределения v , τ , τ_k известны.

A — модель системы S , состоит из 5 агрегатов (рис. 6.3).

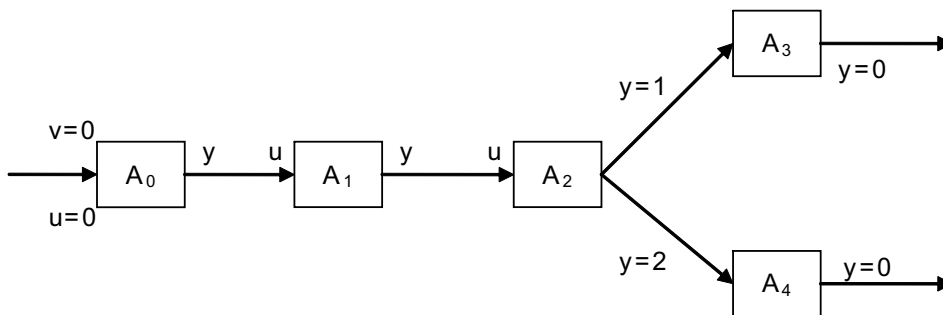


Рис. 6.3

Агрегат A_0 реализует следующие действия:

- моделирует интервалы поступления деталей в S со средним μ и заданным законом распределения;
- моделирует последовательность типов деталей.

Характеристика описания A_0 :

1) вектор состояний

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \in X_0, \quad t \in [0, T],$$

где $x_1(t) \in \{1, 2\}$ — тип детали; $x_2(t)$ — время в момент $t \in [0, T]$, оставшееся до поступления детали в S ;

2) вектор параметров

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ \mu \end{pmatrix};$$

3) $v = 0$, $u = 0$;

4) выходной сигнал

$$y = y(t) \in \{1, 2\},$$

характеризует тип детали.

Агрегат A_k , $k = 1, 4$, моделирует обработку деталей на ОЦ и имеет следующие характеристики:

1) вектор состояний в момент $t \in [0, T]$

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} \in X,$$

где $x_1(t) \in \{0, 1, 2\}$ — состояние ОЦ; $x_2(t) \in \{0, 1\}$ — состояние накопителя при ОЦ; $x_3(t)$ — время до окончания обработки детали;

2) вектор параметров

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix};$$

3) $v = 0$, управляющий сигнал u для A_k , $k = \overline{1, 4}$, совпадает с выходным сигналом из A_{k-1} и представляет собой тип детали;

4) выходные сигналы y для A_3, A_4 равны нулю.

7. ЧИСЛЕННЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ

Настоящее время характеризуется значительным расширением применения математики практически во всех сферах деятельности человека. Это во многом связано с быстрым развитием средств вычислительной техники. Расширение возможностей приложения математики обусловило математизацию различных разделов науки: химии, экономики, биологии, геологии, психологии, медицины, конкретных разделов техники и др. Процесс математизации состоит в построении математических моделей процессов и явлений и разработке методов их исследования. В физике или механике, например, построения математических моделей для описания различных явлений природы и изучения этих моделей с целью объяснения старых или предсказания новых эффектов явлений являются традиционными. Современные успехи в решении таких важных для общества проблем, как атомные, космические, экономические вряд ли были бы возможны без применения вычислительной техники и численных методов. Быстрое проникновение математики во многие области знания, в частности, объясняется тем, что математические модели и методы их исследования применимы ко многим явлениям, сходным по своей формальной структуре. Часто математическая модель, описывающая какое-либо явление, появляется при изучении других явлений или при абстрактных математических построениях задолго до конкретного рассмотрения данного явления.

Математические модели реальных объектов и процессов являются, как правило, достаточно сложными. Поэтому моделирование необходимо осуществлять с помощью ЭВМ с использованием вычислительных алгоритмов.

К вычислительным алгоритмам, которые используются для моделирования, предъявляются следующие основные требования:

- 1) реализуемость — возможность решения задачи за допустимое время;
- 2) экономичность — получение решения за минимальное время среди эквивалентных по точности алгоритмов;
- 3) возможность распознавания ошибок, приводящих к аварийным остановам ЭВМ в процессе вычисления;
- 4) сходимость итерационных алгоритмов, применяемых для решения конкретных задач;
- 5) вычислительная устойчивость — отсутствие накопления суммарной погрешности за счет влияния погрешности округления.

Для решения практически каждой задачи существует несколько численных методов, и выбор метода определяется конкретной задачей и условиями ее реализации. В настоящем учебном пособии будут рассмотрены лишь некоторые численные методы, позволяющие осуществлять моделирование с помощью вычислительной техники.

7.1. Элементы теории погрешности

Результаты вычислений, особенно достаточно большого объема, что характерно для моделирования реальных процессов, всегда являются неточными. Это вызвано накоплением различного рода погрешностей, влияние которых необходимо учитывать.

Основными источниками погрешностей являются:

- 1) исходные данные, которые для вычислений часто берутся из эксперимента, а каждый эксперимент дает результат с ограниченной точностью;
- 2) использование иррациональных величин, которые в ЭВМ представляются приближенно;
- 3) применение итерационных методов решения задач, которые дают приближенные результаты;
- 4) необходимость округления результатов при умножении и делении.

Общепринятой является следующая классификация погрешностей:

- 1) *неустраняемая погрешность* — возникает за счет неточности исходных данных;
- 2) *погрешность метода* — возникает в результате решения задачи;
- 3) *погрешность округления* — всегда присутствует в вычислениях.

В вычислительной математике большинство задач может быть записано в виде

$$y = A(x),$$

где y и x принадлежат заданным пространствам R_1 и R_2 ; $A(x)$ — некоторая заданная функция. Задача состоит либо в отыскании y по заданному x , либо в отыскании x по заданному y . Основным методом, при помощи которого в вычислительной математике решают поставленные задачи, является замена пространств R_1 и R_2 и функции A некоторыми другими пространствами \bar{R}_1 , \bar{R}_2 и функцией \bar{A} . При этом замена должна быть сделана таким образом, чтобы решение новой задачи

$$\bar{y} = \bar{A}(\bar{x}),$$

где $\bar{x} \in \bar{R}_1$; $\bar{y} \in \bar{R}_2$, было в каком-то смысле близким к точному решению исходной задачи и его можно было бы практически отыскать.

За счет погрешности исходных данных и округления фактически решенной является задача

$$\bar{y} = \bar{A}(\bar{x}).$$

При этом полная погрешность решения задачи может быть записана в виде

$$y - \bar{y} = (y - \bar{y}) - (\bar{y} - \bar{y}). \quad (7.1)$$

Первая скобка в (7.1) характеризует погрешность метода, а вторая — погрешность, возникающую за счет неточности исходных данных и округления, которая называется *вычислительной погрешностью*. Таким образом, полная погрешность результатов вычислений складывается из вычислительной погрешности и погрешности метода.

Обозначим точные значения некоторых величин через x^* , y^* , z^* , ..., а соответствующие им приближенные значения через x , y , z , ...

Абсолютной погрешностью величины x называется разность между точным и приближенным значением этой величины:

$$\alpha_x = x^* - x.$$

Предельной абсолютной погрешностью A_x величины x называется наименьшая из верхних границ $|\alpha_x|$, которая может быть найдена, исходя из способа получения числа x .

Относительной погрешностью величины x называется отношение модуля абсолютной погрешности к абсолютному значению приближенной величины x :

$$\delta_x = \frac{|\alpha_x|}{|x|}.$$

Предельной относительной погрешностью Δ_x величины x называется отношение предельной абсолютной погрешности A_x к абсолютному значению величины x :

$$\Delta_x = \frac{A_x}{|x|}.$$

Рассмотрим зависимость абсолютной и относительной погрешностей функции от соответствующих погрешностей ее аргументов. При этом функцию можно рассматривать как модель вычислительного процесса.

Пусть $y = f(x_1, x_2, \dots, x_n)$ — заданная функция n аргументов x_1, x_2, \dots, x_n , по значениям которых требуется определить значение y . Будем предполагать, что:

- 1) функция $f(x_1, x_2, \dots, x_n)$ непрерывно дифференцируема в области определения своих аргументов;
- 2) величины A_f и Δ_f необходимо определить с небольшой точностью;

3) погрешности аргументов настолько меньше значений соответствующих аргументов, что в сумме ими можно пренебречь.

По определению имеем

$$\begin{aligned}\alpha_f &= f(x_1^*, x_2^*, \dots, x_n^*) - f(x_1, x_2, \dots, x_n) = \\ &= \sum_{i=1}^n f'_{x_i}(\xi_1, \xi_2, \dots, \xi_n) \alpha_{x_i},\end{aligned}\quad (7.2)$$

где $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ — некоторая точка отрезка, соединяющего $(x_1^*, x_2^*, \dots, x_n^*)$ и (x_1, x_2, \dots, x_n) , а $f'_{x_i}(\xi)$ — частная производная функции f по x_i в точке ξ , $i = \overline{1, n}$. Последнее равенство в (7.2) получено согласно формуле конечных приращений. Учитывая предположение относительно малости погрешностей аргументов, заменим $f'_{x_i}(\xi)$ на $f'_{x_i}(x)$ и получим

$$\alpha_f \approx \sum_{i=1}^n f'_{x_i}(x) \alpha_{x_i} \quad \text{и} \quad A_f \approx \sum_{i=1}^n |f'_{x_i}(x)| A_{x_i}.$$

Относительные погрешности функции будут определяться следующим образом:

$$\begin{aligned}\delta_f &\approx \sum_{i=1}^n \left| \frac{f'_{x_i}(x)}{f(x)} \right| \alpha_{x_i} = \sum_{i=1}^n \left| \frac{x_i f'_{x_i}(x)}{f(x)} \right| \delta_{x_i}; \\ \Delta_f &\approx \sum_{i=1}^n \left| \frac{f'_{x_i}(x)}{f(x)} \right| A_{x_i} = \sum_{i=1}^n \left| \frac{x_i f'_{x_i}(x)}{f(x)} \right| \Delta_{x_i}.\end{aligned}$$

П р и м е р ы

1. Пусть $f(x) = x_1 + x_2 + x_3 + \dots + x_n$. Тогда

$$A_f \approx \sum_{i=1}^n A_{x_i}; \quad \Delta f \approx \sum_{i=1}^n \left| \frac{x_i}{f(x)} \right| \Delta x_i.$$

2. Пусть $f(x) = x_1 x_2 x_3 \dots x_n$. Тогда

$$A_f \approx \sum_{i=1}^n \left| \frac{f(x)}{x_i} \right| A_{x_i}; \quad \Delta f \approx \sum_{i=1}^n \Delta x_i.$$

Так как погрешности суммируются, то может случиться так, что через достаточно большое число операций погрешности станут столь большими, что полностью исказят результаты вычислений. В связи с этим необходимо всегда учитывать влияние погрешностей, возникающих при моделировании.

7.2. Приближение данных

При моделировании систем достаточно часто приходится осуществлять приближение данных, заданных таблично, непрерывными функциями. Такие задачи возникают, например, при моделировании систем управления для нестационарных моделей объектов (моделей движения самолета на взлете и посадке, движения судна при изменении курса и т.д.). Основными методами, которые используются для построения приближающих функций, являются интерполирование, сплайн-функции, аппроксимация методом наименьших квадратов.

7.2.1. Интерполирование

Пусть задана таблица значений неизвестной функции $y = f(x)$ в точках x_0, x_1, \dots, x_n ($x_0 < x_1 < \dots < x_n$), и необходимо определить значение этой функции в некоторой точке x , не совпадающей с x_i , $i = \overline{0, n}$. Это типичная постановка задач *интерполяции* (если $x_0 < x < x_n$) или *экстраполяции* (если $x < x_0$ или $x > x_n$). Точки x_0, x_1, \dots, x_n называются *узлами интерполирования*, x — *точкой интерполирования*.

В математике под *интерполированием* понимается задача нахождения неизвестного значения какой-либо величины по нескольким известным ее значениям и, может быть, по нескольким значениям других связанных с нею величин (например, ее производных). Без дополнительных предположений о функции $y = f(x)$ эта задача является неопределенной, и за $f(x)$ можно принять любое число, если только x не является одним из узлов таблицы.

Предположим, что рассматривается множество F функций f , обладающих некоторыми свойствами. Рассмотрим семейство функций $\{\varphi_n(x, \alpha_0, \alpha_1, \dots, \alpha_n)\}$, определенных на интервале $[a, b]$, содержащем все узлы интерполирования, и зависящих от параметров α_i , $i = \overline{0, n}$. Будем полагать, что

$$f(x) \approx \varphi_n(x, \alpha_0, \alpha_1, \dots, \alpha_n).$$

Относительно функций φ_n известны следующие предположения:

- 1) функции известны при всяких $n = \overline{0, \infty}$ и определены на отрезке $[a, b]$ для любых значений параметров $\alpha_0, \alpha_1, \dots, \alpha_n$;
- 2) с увеличением n семейство функций $\{\varphi_n\}$ расширяется;
- 3) при всяком n функции семейства $\{\varphi_n\}$ обладают теми же свойствами, что и функции F ;
- 4) для всякой функции $f \in F$ и любого числа $\varepsilon > 0$ существует такое n и такие значения параметров $\alpha_0^*, \alpha_1^*, \dots, \alpha_n^*$, что при всех $x \in [a, b]$ выполняется неравенство

$$|f(x) - \varphi_n(x, \alpha_0^*, \alpha_1^*, \dots, \alpha_n^*)| < \varepsilon$$

(это условие обеспечивает возможность равномерного на $[a, b]$ сколь угодно точного приближения любой функции $f \in F$ с помощью функций из последовательности семейств $\{\varphi_n\}$, при этом семейство функций $\{\varphi_n\}$ называется *полным* для множества F);

5) параметры $\alpha_i, i = \overline{0, n}$, выбираются из условия совпадения функций f и φ_n в узлах $x_i, i = \overline{0, n}$. Это приводит к системе уравнений

$$f(x_i) = \varphi_n(x_i, \alpha_0, \alpha_1, \dots, \alpha_n), \quad i = \overline{0, n}. \quad (7.3)$$

Если φ_n — нелинейная функция относительно параметров $\alpha_0, \alpha_1, \dots, \alpha_n$, то система (7.3) будет нелинейной и ее решение может быть только приближенным. Поэтому чаще всего задают функцию φ_n линейной относительно параметров $\alpha_0, \alpha_1, \dots, \alpha_n$, то есть в виде

$$\varphi_n(x, \alpha_0, \alpha_1, \dots, \alpha_n) = \alpha_0 u_0(x) + \alpha_1 u_1(x) + \dots + \alpha_n u_n(x),$$

где $u_i(x), i = \overline{0, n}$, — линейно независимые *координатные* функции, определенные на $[a, b]$. При этом система (7.3) примет вид

$$\left. \begin{aligned} \alpha_0 u_0(x_0) + \dots + \alpha_n u_n(x_0) &= f(x_0); \\ &\dots \\ \alpha_0 u_0(x_n) + \dots + \alpha_n u_n(x_n) &= f(x_n). \end{aligned} \right\} \quad (7.4)$$

Система (7.4) имеет единственное решение тогда и только тогда, когда ее определитель

$$\Delta = \begin{vmatrix} u_0(x_0) & \dots & u_n(x_0) \\ \dots & \dots & \dots \\ u_0(x_n) & \dots & u_n(x_n) \end{vmatrix} \neq 0. \quad (7.5)$$

Последовательности функций $u_i(x), i = \overline{0, n}$, для которых определитель (7.5) отличен от нуля при всяких несовпадающих значениях $x_i \in [a, b], i = \overline{0, n}$, называются *последовательностями Чебышева*.

Часто в качестве функций φ_n используются алгебраические многочлены

$$P_n(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_n x^n. \quad (7.6)$$

Система (7.4) в этом случае имеет вид

$$\left. \begin{aligned} \alpha_0 + \alpha_1 x_0 + \dots + \alpha_n x_0^n &= f(x_0); \\ &\dots \\ \alpha_0 + \alpha_1 x_n + \dots + \alpha_n x_n^n &= f(x_n), \end{aligned} \right\} \quad (7.7)$$

а ее определитель является *определителем Вандермонда*

$$\Delta = \begin{vmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^n \end{vmatrix},$$

который отличен от нуля для любых несовпадающих узлов $x_i, i = \overline{0, n}$. Таким образом, система (7.7) имеет единственное решение, и интерполирование функции $f(x)$ по ее значениям в узлах $x_i, i = \overline{0, n}$, с помощью многочлена $P_n(x)$ всегда возможно и единственно. Кроме того, согласно теореме Вейерштрасса если отрезок $[a, b]$ конечный и замкнутый и функция $f(x)$ непрерывна на нем вместе с производными первых m порядков, то для всякого $\varepsilon > 0$ существует алгебраический многочлен $P_n(x)$ некоторой степени n , для которого при любых $x \in [a, b]$ выполняются неравенства

$$|f^{(i)}(x) - P_n^{(i)}(x)| \leq \varepsilon, \quad i = \overline{0, m}.$$

Таким образом, если удачно расположить узлы на $[a, b]$, то можно получить достаточно хорошие результаты при вычислении значений функции $f(x)$ и ее производных в любой точке отрезка $[a, b]$.

Решим систему (7.7) по правилу Крамера. Тогда

$$\alpha_i = \frac{\Delta_i}{\Delta} = \sum_{j=0}^n f(x_j) \frac{\Delta_{ij}}{\Delta}, \quad i = \overline{0, n},$$

где Δ_i — определитель, полученный из Δ заменой i -го столбца столбцом свободных членов; Δ_{ij} — алгебраическое дополнение для j -го элемента в Δ_i при разложении по элементам i -го столбца.

Найденные таким образом параметры $\alpha_0, \alpha_1, \dots, \alpha_n$ подставим в (7.6), приведем подобные относительно $f(x_j), j = \overline{0, n}$, и получим

$$P_n(x) = \sum_{j=0}^n f(x_j) \Phi_j(x),$$

где $\Phi_j(x), j = \overline{0, n}$, — алгебраические многочлены n -й степени. Учитывая условие совпадения $f(x)$ и $P_n(x)$ в узлах интерполирования $x_i, i = \overline{0, n}$, получим

$$\Phi_j(x_i) = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Таким образом, $\Phi_j(x)$ есть многочлен n -й степени, равной нулю в n точках $x_0, \dots, x_{j-1}, x_{j+1}, \dots, x_n$, и, следовательно, представимый в виде

$$\Phi_j(x) = c_j (x - x_0) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_n).$$

Константу c_j определим из условия

$$\Phi_j(x_j) = 1.$$

Тогда

$$c_j = \frac{1}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_n)}$$

и

$$\Phi_j(x) = \frac{(x - x_0) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_n)}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_n)}.$$

Таким образом, получим функцию, которая называется *интерполяционным многочленом Лагранжа* и обозначается $L_n(x)$:

$$L_n(x) = \sum_{j=0}^n f(x_j) \frac{(x - x_0) \dots (x - x_{j-1})(x - x_{j+1}) \dots (x - x_n)}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_n)}.$$

Обозначим через $\omega_n(x)$ многочлен $(n+1)$ -й степени вида

$$\omega_n(x) = (x - x_0)(x - x_1) \dots (x - x_n).$$

Тогда многочлен Лагранжа запишется в виде

$$L_n(x) = \sum_{j=0}^n f(x_j) \frac{\omega_n(x)}{(x - x_j)\omega_n'(x_j)}.$$

Интерполяция многочленом Лагранжа имеет существенный недостаток: если выяснится, что полученная точность интерполирования недостаточна и для улучшения точности результата потребуется увеличить число узлов, то все вычисления придется проделать заново. От этого недостатка избавлен другой интерполяционный многочлен — многочлен Ньютона.

Введем следующие обозначения:

разностными отношениями первого порядка назовем величины

$$f(x_i, x_{i+1}) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i};$$

разностными отношениями второго порядка —

$$f(x_i, x_{i+1}, x_{i+2}) = \frac{f(x_{i+1}, x_{i+2}) - f(x_i, x_{i+1})}{x_{i+2} - x_i}$$

и т.д., и разностными отношениями k -го порядка —

$$f(x_i, x_{i+1}, \dots, x_{i+k}) = \frac{f(x_{i+1}, \dots, x_{i+k}) - f(x_i, \dots, x_{i+k-1})}{x_{i+k} - x_i}.$$

Разностные отношения являются симметрическими функциями своих аргументов:

$$f(x_0, x_1, \dots, x_n) = \sum_{j=0}^n \frac{f(x_j)}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_n)}.$$

Для построения многочлена Ньютона запишем многочлен Лагранжа $L_n(x)$ следующим образом:

$$L_n(x) = L_0(x) + [L_1(x) - L_0(x)] + \dots + [L_n(x) - L_{n-1}(x)]. \quad (7.8)$$

Разность $L_k(x) - L_{k-1}(x)$ есть многочлен k -й степени, равный нулю в точках x_i , $i = \overline{0, k-1}$, и, следовательно, представимый в виде

$$L_k(x) - L_{k-1}(x) = c_k(x - x_0) \dots (x - x_{k-1}).$$

Константу c_k определим, полагая $x = x_k$ и учитывая, что $L_k(x_k) = f(x_k)$, получим

$$\begin{aligned} c_k &= \frac{f(x_k)}{(x_k - x_0) \dots (x_k - x_{k-1})} - \\ &= \frac{\sum_{j=0}^{k-1} f(x_j) \frac{(x_k - x_0) \dots (x_k - x_{j-1})(x_k - x_{j+1}) \dots (x_k - x_{k-1})}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_{k-1})}}{(x_k - x_0) \dots (x_k - x_{k-1})} = \\ &= \sum_{j=0}^k \frac{f(x_j)}{(x_j - x_0) \dots (x_j - x_{j-1})(x_j - x_{j+1}) \dots (x_j - x_k)} = f(x_0, \dots, x_k). \end{aligned}$$

Тогда

$$L_k(x) - L_{k-1}(x) = (x - x_0) \dots (x - x_{k-1}) f(x_0, \dots, x_k).$$

Подставляя полученное выражение в (7.8) для $k = \overline{1, n}$, получим интерполяционный многочлен Ньютона

$$\begin{aligned} P_n(x) &= f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \\ &+ \dots + (x - x_0) \dots (x - x_{n-1})f(x_0, x_1, \dots, x_n). \end{aligned}$$

Несмотря на более сложное строение, формула Ньютона удобнее для вычислений, чем формула Лагранжа, так как позволяет не фиксировать заранее число узлов, а подключать новые узлы до тех пор, пока не будет достигнута требуемая точность вычислений.

Величина остатка или погрешности интерполирования

$$R(x) = f(x) - P_n(x)$$

зависит от свойств функции f , выбора узлов x_k , $k = \overline{0, n}$, и от положения точки x на $[a, b]$. Формула Лагранжа для $R(x)$ имеет вид

$$R(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_n(x),$$

где ξ — некоторая точка отрезка $[a, b]$. Если обозначить

$$M_{n+1} = \sup_{x \in [a, b]} |f^{(n+1)}(x)|,$$

то

$$R(x) \leq \frac{M_{n+1}}{(n+1)!} |\omega_n(x)|. \quad (7.9)$$

При этом точность экстраполирования практически всегда хуже точности интерполирования.

Для того чтобы погрешность в (7.9) была наименьшей, необходимо выбрать узлы интерполирования таким образом, чтобы значение

$\sup_{x \in [a,b]} |\omega_n(x)|$ было минимальным. Для этого можно воспользоваться *многочленами Чебышева*

$$T_n(t) = \cos[n \arccos(t)].$$

Многочлены Чебышева обладают следующими свойствами:

1) они являются алгебраическими многочленами соответствующей степени, так как $T_1(x) = 1$, $T_2(x) = x$, а для построения многочленов более высокой степени можно воспользоваться рекуррентной формулой

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x);$$

2) коэффициент при старшей степени x в $T_n(x)$ равен 2^{n-1} ;

3) области определения и изменения многочленов Чебышева равны $[-1, 1]$;

4) многочлен $T_n(x)$ имеет n корней на $[-1, 1]$, которые располагаются в точках

$$x_k = \cos \frac{(2k+1)\pi}{2n}, \quad k = \overline{0, n-1};$$

5) многочлен $T_n(x)$ достигает $n+1$ экстремум, равный ± 1 , на $[-1, 1]$ в точках

$$x_k = \cos \frac{k\pi}{n}, \quad k = \overline{0, n};$$

6) многочлен $T_n(x)$ является четным или нечетным в зависимости от n , так как

$$T_n(-x) = (-1)^n T_n(x);$$

7) многочлены Чебышева ортогональны с весом $\rho(x) = \frac{1}{\sqrt{1-x^2}}$ на $[-1, 1]$, так как

$$\int_{-1}^1 \rho(x) T_n(x) T_m(x) dx = \begin{cases} \pi, & n = m = 0, \\ \frac{\pi}{2}, & n = m \neq 0, \\ 0, & n \neq m; \end{cases}$$

8) графическое представление многочленов Чебышева совпадает с проекциями синусов, нанесенных на цилиндр;

9) многочлены Чебышева являются наименее отклоняющимися от нуля, так как

$$\sup_{x \in [-1, 1]} |T_n(x)| = \frac{1}{2^{n-1}},$$

при этом справедливо утверждение, что какой бы другой многочлен $P_n(x)$ степени n со старшим коэффициентом, равным единице, мы ни взяли,

$$\sup_{x \in [-1, 1]} |P_n(x)| \geq \frac{1}{2^{n-1}}.$$

Таким образом, для того чтобы значение $\sup_{x \in [-1, 1]} |\omega_n(x)|$ было наименьшим, необходимо в качестве функции

$$\omega_n(x) = (x - x_0)(x - x_1) \dots (x - x_n)$$

при $a = -1$, $b = 1$ взять приведенный многочлен Чебышева $\frac{1}{2^n} T_{n+1}(x)$, то есть в качестве узлов интерполирования взять корни многочлена Чебышева

$$x_i = \cos \frac{(2i + 1)\pi}{2n + 2}, \quad i = \overline{0, n}.$$

Тогда для оценки погрешности будет верно неравенство

$$R(x) \leq \frac{M_{n+1}}{2^n (n + 1)!}.$$

Если интерполирование производится на произвольном интервале $[a, b]$, то значения $x_i \in [-1, 1]$ с помощью линейного преобразования

$$t_i = \frac{1}{2} [(b - a)x_i + (b + a)], \quad i = \overline{0, n},$$

можно перевести в интервал $[a, b]$ и для оценки погрешности будет верно неравенство

$$R(x) \leq \frac{M_{n+1}(b - a)^{n+1}}{2^{2n+1} (n + 1)!}.$$

Достаточно часто при интерполировании приходится иметь дело с равноотстоящими узлами. Тогда полагают:

$$x_0 = a; \quad x_n = b; \quad h = \frac{b - a}{n}; \quad x_i = x_0 + ih; \quad i = \overline{0, n}.$$

В этом случае интерполяционные формулы можно значительно упростить, если ввести понятие *конечных разностей*.

Конечной разностью k -го порядка называется величина

$$\Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i, \quad k = 1, 2, \dots,$$

где

$$y_i = f(x_i); \quad \Delta^0 y_i = y_i, \quad i = \overline{0, n}.$$

Так как узлы, лежащие вблизи точки интерполирования x , оказывают большее влияние на результат, чем узлы, лежащие дальше, то для уменьшения погрешности при построении интерполяционной формулы целесообразно привлекать сначала ближний узел, а затем остальные в порядке их удаленности от точки x . При этом абсолютные величины слагаемых будут уменьшаться с увеличением их порядкового номера i , начиная с некоторого слагаемого, выйдут за пределы требуемой точности. В соответствии с этим построим конкретные интерполяционные формулы.

Формула Ньютона для интерполирования вперед

Пусть точка x находится ближе к левому концу отрезка $[a, b]$ или слева от него. Для построения интерполяционной формулы Ньютона за x_0 принимается ближайший к x узел интерполирования, а затем узлы привлекаются в следующем порядке:

$$x_0, \quad x_0 + h, \quad x_0 + 2h, \quad \dots, \quad x_0 + nh,$$

при этом, если это возможно, за x_0 выбирается лежащий слева от точки x узел, чтобы не попасть в ситуацию экстраполирования. Тогда многочлен Ньютона (7.5) запишется в виде

$$\begin{aligned} P(x) = & f(x_0) + (x - x_0)f(x_0, x_0 + h) + \\ & + (x - x_0)(x - x_0 - h)f(x_0, x_0 + h, x_0 + 2h) + \dots + (x - x_0)(x - x_0 - h) \dots \\ & \dots (x - x_0 - (n-1)h)f(x_0, x_0 + h, \dots, x_0 + nh). \end{aligned}$$

Если ввести новую переменную $t = \frac{x - x_0}{h}$, то

$$(x - x_0)(x - x_0 - h) \dots (x - x_0 - ih) = h^{i+1}t(t-1) \dots (t-i),$$

и учитывая, что

$$f(x_0, x_0 + h, \dots, x_0 + ih) = \frac{\Delta^i y_0}{i! h^i},$$

получим

$$\begin{aligned} P(x_0 + th) = & y_0 + \frac{t}{1!} \Delta y_0 + \frac{t(t-1)}{2!} \Delta^2 y_0 + \frac{t(t-1)(t-2)}{3!} \Delta^3 y_0 + \dots \\ & \dots + \frac{t(t-1) \dots (t-n+1)}{n!} \Delta^n y_0. \end{aligned}$$

Эта формула называется *формулой Ньютона для интерполирования вперед*.

Формула Ньютона для интерполирования назад

Пусть точка x находится ближе к правому концу отрезка $[a, b]$ или справа от него. За первый узел для построения интерполяционной формулы примем ближайший к x узел интерполирования и обозначим его x_n , при этом, если возможно, следует избегать ситуации экстраполирования. Далее узлы для построения интерполяционной формулы следует привлекать в следующем порядке:

$$x_n, \quad x_n - h, \quad x_n - 2h, \quad \dots, \quad x_n - nh.$$

Введем новую переменную $t = \frac{x - x_n}{h}$. Тогда, преобразовывая аналогично предыдущему, получим формулу Ньютона для интерполирования назад

$$P(x_n + th) = y_n + \frac{t}{1!} \Delta y_{n-1} + \frac{t(t+1)}{2!} \Delta^2 y_{n-2} + \\ + \frac{t(t+1)(t+2)}{3!} \Delta^3 y_{n-3} + \dots + \frac{t(t+1)\dots(t+n-1)}{n!} \Delta^n y_0.$$

Интерполяционные формулы Гаусса

Если точка интерполирования x находится в середине отрезка $[a, b]$ вблизи узла x_k , причем правее его, то есть $x > x_k$, то для построения интерполяционной формулы узлы привлекаются в следующем порядке:

$$x_k, \quad x_k + h, \quad x_k - h, \quad \dots, \quad x_k + ih, \quad x_k - ih.$$

Тогда, после введения новой переменной $t = \frac{x - x_k}{h}$ и соответствующих преобразований, получается формула Гаусса для интерполирования вперед, которая имеет вид

$$P(x_k + th) = y_k + \frac{t}{1!} \Delta y_k + \frac{t(t-1)}{2!} \Delta^2 y_{k-1} + \frac{t(t^2 - 1^2)}{3!} \Delta^3 y_{k-1} + \dots \\ \dots + \frac{t(t^2 - 1^2) \dots (t^2 - (i-1)^2)}{(2i-1)!} \Delta^{2i-1} y_{k-i+1} + \\ + \frac{t(t^2 - 1^2) \dots (t^2 - (i-1)^2)(t-i)}{(2i)!} \Delta^{2i} y_{k-i}.$$

Если точка x находится вблизи узла x_k , причем $x < x_k$, то для построения интерполяционной формулы узлы привлекаются в следующем порядке:

$$x_k, \quad x_k - h, \quad x_k + h, \quad \dots, \quad x_k - ih, \quad x_k + ih,$$

и, после введения переменной $t = \frac{x - x_k}{h}$ и соответствующих преобразований, получим формулу Гаусса для интерполирования назад

$$\begin{aligned} P(x_k + th) = & y_k + \frac{t}{1!} \Delta y_{k-1} + \frac{t(t+1)}{2!} \Delta^2 y_{k-1} + \frac{t(t^2-1^2)}{3!} \Delta^3 y_{k-2} + \dots \\ & \dots + \frac{t(t^2-1^2) \dots (t^2-(i-1)^2)}{(2i-1)!} \Delta^{2i-1} y_{k-i} + \\ & + \frac{t(t^2-1^2) \dots (t^2-(i-1)^2)(t+i)}{(2i)!} \Delta^{2i} y_{k-i}. \end{aligned}$$

Если функция $y(x)$ достаточно гладкая, то разности $\Delta^j y_i$, $j = 1, 2, \dots$ убывают с ростом j и для некоторого j они практически равны нулю. Но так как с ростом порядка разностей растет погрешность их вычисления, то привлечение таких разностей для построения интерполяционной формулы может существенно исказить результат. Существует следующее правило определения максимального порядка разностей, которые ведут себя *правильно*, то есть не вносят дополнительной погрешности в значение интерполяционного многочлена.

Пусть значения y_i , $i = \overline{0, n}$, вычислены с абсолютной погрешностью ε . Так как конечные разности определяются по формуле

$$\Delta^{j+1} y_i = \Delta^j y_{i+1} - \Delta^j y_i, \quad j = 0, 1, 2, \dots, \quad \Delta^0 y_i = y_i, \quad i = \overline{0, n},$$

то абсолютная погрешность разностей первого порядка равна 2ε , второго порядка — 4ε , ..., s -го порядка — $2^s \varepsilon$. Тогда, если одновременно выполняются неравенства

$$\max_i |\Delta^j y_i| \geq 2^j \varepsilon; \quad \max_i |\Delta^{j+1} y_i| < 2^{j+1} \varepsilon,$$

то максимальный порядок разностей, которые ведут себя *правильно*, равен j . Разности $(j+1)$ -го порядка уже меньше погрешности, поэтому их использование приведет к искажению результата и при построении интерполяционной формулы они отбрасываются.

7.2.2. Сплайн-функции

Пусть на отрезке $[a, b]$ задано разбиение $a = x_0 < x_1 < x_2 < \dots < x_n = b$, в узлах которого известны значения достаточно гладкой функции $y = f(x)$: $y_i = f(x_i)$, $i = \overline{0, n}$. Узлы разбиения делят отрезок $[a, b]$ на n отрезков $[x_0, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$.

Сплайном называется составная функция $P(x)$, которая вместе с производными нескольких порядков непрерывна на всем отрезке $[a, b]$,

а на каждом частичном отрезке $[x_{i-1}, x_i]$ в отдельности является составляющей функцией

$$P_i(x) = F(i, x_0, \dots, x_n, y_0, \dots, y_n), \quad i = \overline{1, n}.$$

Рассмотрим частный случай, когда функции $P_i(x)$ являются алгебраическими многочленами вида

$$P_i(x) = \sum_{k=0}^N \alpha_k^{(i)} (x - x_{i-1})^k, \quad i = \overline{1, n},$$

где $\alpha_k^{(i)}$ — коэффициенты, определяемые для каждого частичного отрезка.

Максимальная по всем частичным отрезкам степень многочлена называется *степенью сплайна*, а разность между степенью сплайна и порядком наивысшей непрерывной на $[a, b]$ производной — *дефектом сплайна*.

Среди существующих сплайнов наиболее широкое распространение получили линейные, параболические, кубические сплайны.

Рассмотрим алгоритмы построения коэффициентов перечисленных сплайнов.

Линейный сплайн

Сплайн $P(x)$ состоит из линейных многочленов вида

$$P_i(x) = a_i + b_i(x - x_{i-1}), \quad x \in [x_{i-1}, x_i], \quad i = \overline{1, n}. \quad (7.10)$$

Параметры сплайна a_i, b_i , $i = \overline{1, n}$, определим из условия непрерывности $P(x)$ на $[a, b]$ и требования совпадения значений сплайна с функцией $f(x)$ в узловых точках

$$P_i(x_i) = P_{i+1}(x_i), \quad i = \overline{0, n-1}; \quad (7.11)$$

$$P_i(x_i) = y_i, \quad i = \overline{0, n}. \quad (7.12)$$

Обозначим $h_i = x_i - x_{i-1}$. Тогда для каждого из многочленов можно записать

$$P_i(x_i) = a_i + b_i h_i = y_i, \quad i = \overline{0, n};$$

$$P_i(x_{i-1}) = a_i = y_{i-1}, \quad i = \overline{1, n},$$

и для определения коэффициентов линейного сплайна (7.10) получим уравнения

$$a_i = y_{i-1}, \quad i = \overline{1, n};$$

$$b_i = \frac{y_i - y_{i-1}}{h_i}, \quad i = \overline{1, n}.$$

Параболический сплайн

Сплайн $P(x)$ состоит из парабол, то есть многочлен $P_i(x)$ имеет вид

$$P_i(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2, \\ x \in [x_{i-1}, x_i], \quad i = \overline{1, n}.$$

Для определения коэффициентов сплайна дополнительно к условиям (7.11), (7.12) потребуем непрерывности первой производной сплайна на интервале $[a, b]$, то есть

$$P'_i(x_i) = P'_{i+1}(x_i), \quad i = \overline{1, n-1}. \quad (7.13)$$

Обозначив $h_i = x_i - x_{i-1}$, $i = \overline{1, n}$, и, учитывая, что

$$P'_i(x) = b_i + 2c_i(x - x_{i-1}),$$

в соответствии с формулами (7.11), (7.12) и (7.13) получим:

$$a_i = y_{i-1}, \quad (7.14)$$

$$a_i + b_i h_i + c_i h_i^2 = y_i, \quad (7.15)$$

$$b_{i+1} = b_i + 2c_i h_i. \quad (7.16)$$

Тогда коэффициенты a_i , $i = \overline{1, n}$, определяются согласно (7.14), а (7.15) можно записать в виде

$$b_i = \frac{y_i - y_{i-1}}{h_i} - h_i c_i, \quad i = \overline{1, n}. \quad (7.17)$$

Если теперь записать выражения для b_i и b_{i+1} в виде (7.17) и подставить их в (7.16), то получим

$$c_i h_i + c_{i+1} h_{i+1} = g_i, \quad i = \overline{1, n-1}, \quad (7.18)$$

где

$$g_i = \frac{(y_{i+1} - y_i)}{h_{i+1}} - \frac{(y_i - y_{i-1})}{h_i}, \quad i = \overline{1, n-1}.$$

Таким образом, уравнения (7.14), (7.17) и (7.18) образуют систему из $(3n-1)$ уравнений для определения $3n$ коэффициентов сплайна. Недостающее уравнение получается из дополнительного условия, которое накладывается на значение производной сплайна на конце интервала $[a, b]$, в виде

$$P'(x_n) = 0$$

или

$$b_n + 2c_n h_n = 0. \quad (7.19)$$

Если подставить в (7.19) выражение для b_n из (7.17), то формулу (7.19) можно переписать в виде

$$c_n h_n = g_n,$$

где $g_n = \frac{y_{n-1} - y_n}{h_n}$.
Тогда

$$c_n = \frac{g_n}{h_n} \quad (7.20)$$

и

$$c_i = \frac{g_i - c_{i+1}h_{i+1}}{h_i}, \quad i = \overline{n-1, 1}. \quad (7.21)$$

Таким образом, параметры параболического сплайна вычисляются в следующем порядке: сначала в обратном порядке вычисляются коэффициенты c_i , $i = \overline{n, 1}$, по (7.20), (7.21), затем b_i , $i = \overline{1, n}$, по (7.16) или (7.17), а затем a_i , $i = \overline{1, n}$, по (7.14).

Кубический сплайн

Сплайн $P(x)$ состоит из гипербол вида

$$P_i(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3, \\ x \in [x_{i-1}, x_i], \quad i = \overline{1, n}.$$

Для определения параметров сплайна потребуем дополнительно к (7.11), (7.12), (7.13) выполнения условия непрерывности второй производной

$$P_i''(x_i) = P_{i+1}''(x_i), \quad i = \overline{1, n-1}. \quad (7.22)$$

Первая и вторая производные отдельных многочленов сплайна соответственно равны:

$$P_i'(x) = b_i + 2c_i(x - x_{i-1}) + 3d_i(x - x_{i-1})^2, \\ P_i''(x) = 2c_i + 6d_i(x - x_{i-1}).$$

Тогда, обозначив $h_i = x_i - x_{i-1}$, получим:

$$P_i(x_{i-1}) = a_i = y_{i-1}, \quad i = \overline{1, n}, \quad (7.23)$$

$$P_i(x_i) = a_i + b_i h_i + c_i h_i^2 + d_i h_i^3 = y_i, \quad i = \overline{1, n}, \quad (7.24)$$

$$b_{i+1} = b_i + 2c_i h_i + 3d_i h_i^2, \quad i = \overline{1, n-1}, \quad (7.25)$$

$$c_{i+1} = c_i + 3d_i h_i, \quad i = \overline{1, n-1}. \quad (7.26)$$

Последние два уравнения получены из (7.13) и (7.22).

Уравнения (7.23)–(7.26) составляют систему из $(4n-2)$ уравнений для определения $4n$ параметров сплайна. Для получения двух недостающих уравнений потребуем выполнения дополнительных условий на концах интервала $[a, b]$:

$$P''(x_0) = 0, \quad (7.27)$$

$$P''(x_n) = 0. \quad (7.28)$$

Тогда из (7.27) следует

$$c_1 = 0, \quad (7.29)$$

а из (7.28) получим

$$c_{n+1} = c_n + 3d_n h_n = 0.$$

Из уравнения (7.26) выразим d_i :

$$d_i = \frac{c_{i+1} - c_i}{3h_i}, \quad i = \overline{1, n}, \quad (7.30)$$

и, подставляя его в (7.24), с учетом (7.23), получим

$$b_i = \frac{y_i - y_{i-1}}{h_i} - \frac{h_i(c_{i+1} + 2c_i)}{3}, \quad i = \overline{1, n}. \quad (7.31)$$

Если выражения для b_i , b_{i+1} и d_i подставить в (7.25), то получим

$$h_i c_i + 2(h_i + h_{i+1})c_{i+1} + h_{i+1}c_{i+2} = g_{i+1}, \quad (7.32)$$

где

$$g_{i+1} = 3 \left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right), \quad i = \overline{1, n-1}.$$

Уравнения (7.32) образуют систему из $n-1$ уравнений относительно неизвестных c_2, c_3, \dots, c_n . Эта система является системой с трехдиагональной матрицей вида

$$\begin{pmatrix} 2(h_1 + h_2) & h_2 & \dots & 0 & 0 \\ h_2 & 2(h_2 + h_3) & \dots & 0 & 0 \\ 0 & h_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & h_{n-1} & 2(h_{n-1} + h_n) \end{pmatrix}$$

и вектором свободных членов $(g_2, g_3, \dots, g_n)^T$, где T — символ транспонирования.

Решение таких систем осуществляется *методом прогонки*, согласно которому решение представляют в виде

$$c_{i+1} = \xi_{i+1}c_{i+2} + \eta_{i+1}, \quad i = \overline{n-1, 1}, \quad (7.33)$$

где ξ_{i+1}, η_{i+1} — неизвестные коэффициенты. Чтобы получить выражения для этих коэффициентов, запишем формулы для определения c_{i+1} и c_i согласно (7.33) и, подставив их в (7.32), сравним полученное выражение с (7.33). При сравнении получим, что выражения для определения неизвестных коэффициентов ξ_{i+1}, η_{i+1} будут иметь вид

$$\left. \begin{aligned} \xi_{i+1} &= \frac{-h_{i+1}}{h_i \xi_i + 2(h_i + h_{i+1})}; \\ \eta_{i+1} &= \frac{g_{i+1} - h_i \eta_i}{h_i \xi_i + 2(h_i + h_{i+1})}, \end{aligned} \right\} \quad i = \overline{1, n-1}. \quad (7.34)$$

При этом, учитывая (7.29), полагают $\eta_1 = 0$, $\xi_1 = 0$.

Таким образом, порядок вычисления коэффициентов кубического сплайна следующий: сначала определяют коэффициенты c_i , $i = \overline{2, n}$, для чего необходимо, осуществляя прямой ход метода прогонки по формулам (7.34), найти значения η_{i+1} , ξ_{i+1} , $i = \overline{1, n-1}$, при $\eta_1 = 0$, $\xi_1 = 0$, а затем, обратным ходом, считая $c_{n+1} = 0$, по формуле (7.33) вычислить c_n, c_{n-1}, \dots, c_2 . При этом, согласно (7.29), $c_1 = 0$. Остальные коэффициенты сплайна определяются по следующим формулам: a_i — (7.23), d_i — (7.30), b_i — (7.31), $i = \overline{1, n}$.

Полиномиальные сплайны имеют ряд существенных недостатков:

1) при увеличении степени составных многочленов вычисление их коэффициентов значительно усложняется из-за увеличения числа уравнений (условий непрерывности функций и их производных);

2) требуется достаточно большой объем памяти для хранения информации о сплайне (значений x_i , y_i , $i = \overline{0, n}$, и коэффициентов сплайнов).

От этих недостатков в значительной мере избавлен другой вид сплайнов, который основан на базисных функциях и называется *B-сплайном*.

7.2.3. Аппроксимация данных методом наименьших квадратов

Пусть x_j — узлы исходной таблицы данных, а $f(x_j)$, $j = \overline{0, n}$, — значения экспериментальных данных или некоторой неизвестной функции в узловых точках. Введем непрерывную функцию $\varphi(x)$ для аппроксимации дискретных значений $f(x_j)$ и обозначим

$$\xi_j = \varphi(x_j) - f(x_j), \quad j = \overline{0, n},$$

отклонения в узлах x_j . Тогда сумма квадратов отклонений аппроксимирующей функции $\varphi(x)$ от неизвестной функции $f(x)$ в узловых точках x_j , $j = \overline{0, n}$, запишется в виде

$$Q = \sum_{j=0}^n \xi_j^2 = \sum_{j=0}^n [\varphi(x_j) - f(x_j)]^2.$$

Метод построения аппроксимирующей функции $\varphi(x)$ из условия минимума суммы квадратов отклонений Q называется *методом наименьших квадратов* (МНК).

Наиболее часто аппроксимирующую функцию $\varphi(x)$ задают в виде

$$\varphi(x) = c_0\varphi_0(x) + c_1\varphi_1(x) + \dots + c_m\varphi_m(x),$$

где $\varphi_i(x)$, $i = \overline{0, m}$, $m \leq n$, — линейно независимые *базисные* функции; c_0, \dots, c_m — неизвестные коэффициенты, определяемые из условия минимума Q , т.е. из условий равенства нулю частных производных

$$\left. \begin{aligned} \frac{\partial Q}{\partial c_k} &= 2 \sum_{j=0}^n [c_0\varphi_0(x_j) + \dots + c_m\varphi_m(x_j) - f(x_j)] \varphi_k(x_j) = 0, \\ k &= \overline{0, m}. \end{aligned} \right\}$$

Таким образом, получаем систему линейных алгебраических уравнений вида $Ac=b$ для определения коэффициентов c_k , $k = \overline{0, m}$. Эта система называется системой *нормальных* уравнений. Матрица системы имеет вид

$$A = \begin{pmatrix} (\varphi_0, \varphi_0) & \dots & (\varphi_0, \varphi_m) \\ \dots & \dots & \dots \\ (\varphi_m, \varphi_0) & \dots & (\varphi_m, \varphi_m) \end{pmatrix}$$

и называется *матрицей Грама*.

Элементами матрицы Грама являются скалярные произведения базисных функций

$$(\varphi_i, \varphi_k) = \sum_{j=0}^n \varphi_i(x_j) \varphi_k(x_j), \quad i, k = \overline{0, m}. \quad (7.35)$$

Вектор свободных членов системы нормальных уравнений имеет вид

$$b = [(\varphi_0, f), (\varphi_1, f), \dots, (\varphi_m, f)]^T,$$

элементами этого вектора являются скалярные произведения

$$(\varphi_k, f) = \sum_{j=0}^n \varphi_k(x_j) f(x_j), \quad k = \overline{0, m}.$$

Матрица Грама обладает следующими основными свойствами:

1) она симметрична, что позволяет сократить объем вычислений при заполнении матрицы;

2) матрица является положительно определенной, поэтому при решении системы нормальных уравнений методом исключения Гаусса можно отказаться от процедуры выбора главного элемента, а использовать в качестве ведущего диагональный элемент;

3) определитель матрицы Грама отличен от нуля, если в качестве базиса выбраны линейно независимые функции $\varphi_k(x)$, $k = \overline{0, m}$.

Для аппроксимации экспериментальных данных, определенных с погрешностью ε в каждой узловой точке, сначала функцию $\varphi(x)$

задают в виде линейной комбинации из одной или двух базисных функций, и если после определения коэффициентов c_k окажется, что $\sqrt{Q} > \varepsilon$, то расширяют базис добавлением новых функций $\varphi_k(x)$ до тех пор, пока не выполнится условие $\sqrt{Q} \approx \varepsilon$.

Выбор конкретных базисных функций зависит от таких свойств аппроксимируемой функции $f(x)$, как периодичность, экспоненциальный или логарифмический характер, свойства симметрии, наличие асимптотики и т.д.

Аппроксимация алгебраическими полиномами

Пусть аппроксимирующая функция задана в виде алгебраического полинома степени m :

$$\varphi(x) = c_0 + c_1x + \dots + c_mx^m \quad (7.36)$$

или

$$\varphi_0(x) = 1, \quad \varphi_1(x) = x, \quad \dots, \quad \varphi_m(x) = x^m.$$

Степень полинома m выбирают обычно меньше n . Аппроксимирующая кривая в этом случае не проходит через экспериментальные точки, т.е. экспериментальные данные «сглаживаются» с помощью функции $\varphi(x)$. Если взять $m = n$, то $\varphi(x)$ совпадает с многочленом Лагранжа, аппроксимирующая кривая пройдет через все экспериментальные точки и $Q = 0$. Это обстоятельство часто используется для отладки и тестирования программ, реализующих алгоритмы МНК.

Матрица A системы нормальных уравнений и вектор свободных членов для функции $\varphi(x)$ вида (7.36) записываются следующим образом:

$$A = \begin{pmatrix} n+1 & \sum_{j=0}^n x_j & \sum_{j=0}^n x_j^2 & \dots & \sum_{j=0}^n x_j^m \\ \sum_{j=0}^n x_j & \sum_{j=0}^n x_j^2 & \sum_{j=0}^n x_j^3 & \dots & \sum_{j=0}^n x_j^{m+1} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{j=0}^n x_j^m & \sum_{j=0}^n x_j^{m+1} & \sum_{j=0}^n x_j^{m+2} & \dots & \sum_{j=0}^n x_j^{2m} \end{pmatrix}$$

$$b = \left(\sum_{j=0}^n f(x_j), \sum_{j=0}^n x_j f(x_j), \sum_{j=0}^n x_j^2 f(x_j), \dots, \sum_{j=0}^n x_j^m f(x_j) \right)^T.$$

Для решения систем уравнений с матрицей Грама разработаны *garfa* и *ncskoéžkĩcĩ sãëëřaķč*. Если же $m \leq 4$, то такие системы можно решать и более простым методом исключения Гаусса.

Система нормальных уравнений обычно бывает *плохо обусловленной* (способной накапливать погрешности). Это приводит к дополнительным сложностям при реализации МНК: вычисление с двойной (расширенной) точностью, введение весовых коэффициентов и т.д. Весовые коэффициенты выбираются из различных соображений, например, полагают

$$w_j = \frac{1}{f(x_j)}, \quad j = \overline{0, n}.$$

Такой выбор весовых коэффициентов называется *методом статистического взвешивания*.

В этом случае

$$Q = \sum_{j=0}^n w_j [\varphi(x_j) - f(x_j)]^2$$

и скалярные произведения в матрице A и векторе b будут соответственно иметь вид

$$(\varphi_i, \varphi_k) = \sum_{j=0}^n w_j \varphi_i(x_j) \varphi_k(x_j);$$

$$(\varphi_k, f) = \sum_{j=0}^n w_j \varphi_k(x_j) f(x_j).$$

Аппроксимация ортогональными полиномами

Лучшие по точности результаты при аппроксимации можно получить, если использовать в качестве базисных функций $\varphi_i(x)$ *классические ортогональные полиномы Чебышева, Лежандра, Лагерра, Якоби* и других.

Полиномы называются ортогональными, если существует некоторый интервал $[a, b]$, на котором

$$\int_a^b p(x) \varphi_j(x) \varphi_k(x) dx = 0, \quad j \neq k, \quad (7.37)$$

где $p(x)$ — весовая функция.

В случае большого количества узлов x_j на $[a, b]$ значения интегралов (7.37) будут близки к дискретным скалярным произведениям (7.35), так как интегрирование можно приближенно заменить суммированием. В этом случае недиагональные элементы матрицы Грама будут небольшими по абсолютной величине, что уменьшает погрешность решения системы нормальных уравнений.

Для наиболее гладкого представления экспериментальных данных (с минимальным числом и амплитудой выбросов) в качестве базисных функций $\varphi_k(x)$ выбирают ортогональные полиномы Чебышева $T_k(x)$,

которые определены и ортогональны на интервале $[-1, 1]$ с весовой

функцией $\rho(x) = \frac{1}{\sqrt{1-x^2}}$.

Для задания полиномов Чебышева используется рекуррентная формула

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x),$$

где $T_0(x) = 1$; $T_1(x) = x$.

Так как в многочленах Чебышева коэффициент при старших степенях x равен 2^k , это не всегда удобно при оценке вклада в аппроксимирующую функцию $\phi(x)$ старших степеней x по величине коэффициентов c_k . В этом случае полиномы Чебышева можно ввести по другой рекуррентной формуле, позволяющей построить приведенные многочлены Чебышева

$$T_{k+1}^*(x) = xT_k^*(x) - \frac{1}{4}T_{k-1}^*(x),$$

где $T_0^*(x) = 1$; $T_1^*(x) = x$.

Полиномы $T_k^*(x)$ ортогональны на интервале $[-1, 1]$ с такой же весовой функцией, что и $T_k(x)$.

Весовую функцию, равную единице на интервале $[-1, 1]$, имеют полиномы Лежандра, которые определяются по следующей рекуррентной формуле:

$$L_{k+1}(x) = \frac{1}{k+1}[(2k+1)xL_k(x) + kL_{k-1}(x)],$$

где $L_0(x) = 1$; $L_1(x) = x$.

Интервал $[x_0, x_n]$, где заданы узлы таблицы данных x_j , переводится в интервал $[-1, 1]$, где определены и ортогональны полиномы Чебышева и Лежандра с помощью линейного преобразования

$$z = 2 \frac{x - x_0}{x_n - x_0} - 1.$$

Аппроксимация ортогональными полиномами дискретной переменной

Если построить систему базисных функций $\phi_k(x)$ таким образом, чтобы обращались в нуль скалярные произведения на дискретном множестве узловых точек, то матрица Грама будет диагональной и можно избежать численного решения системы нормальных уравнений. В зависимости от распределения погрешности обрабатываемых данных можно построить ортогональные полиномы дискретной переменной с соот-

ветствующими дискретными весовыми функциями $p(x_j)$, $j = \overline{0, n}$. Из классических ортогональных полиномов дискретной переменной известны полиномы Хана, Мейкснера, Кравчука, Шарлье.

Рассмотрим алгоритм построения полиномов Чебышева $t_k(x)$ дискретной переменной, которые являются частным случаем полиномов Хана с единичной весовой функцией.

Полагаем:

$$t_0(x) = 1,$$

$$t_1(x) = x - a_{11},$$

и неизвестный коэффициент a_{11} определим из условия ортогональности $t_0(x)$ и $t_1(x)$, то есть

$$(t_0(x), t_1(x)) = 0$$

или

$$\sum_{j=0}^n 1 \cdot (x_j - a_{11}) = \sum_{j=0}^n x_j - a_{11} \sum_{j=0}^n 1 = \sum_{j=0}^n x_j - a_{11}(n+1).$$

Отсюда

$$a_{11} = \frac{1}{n+1} \sum_{j=0}^n x_j.$$

Полином второй степени также представляется в общем виде с нео-пределенными коэффициентами a_{21} и a_{20} :

$$t_2(x) = x^2 + a_{21}x + a_{20}.$$

Коэффициенты a_{21} и a_{20} найдем из условия ортогональности полиномов $t_0(x)$, $t_1(x)$, $t_2(x)$, то есть $(t_0, t_2) = 0$, $(t_1, t_2) = 0$ и т.д.

Для полиномов Чебышева дискретной переменной существует двух-слойная рекуррентная формула, по которой можно вычислить полином любой степени, зная $t_0(x)$ и $t_1(x)$:

$$t_{k+1}(x) = (x - a_{k+1})t_k - b_{k+1}t_{k-1}(x),$$

где

$$a_{k+1} = \frac{\sum_{j=0}^n x_j t_k^2(x_j)}{\sum_{j=0}^n t_k^2(x_j)}; \quad b_{k+1} = \frac{\sum_{j=0}^n t_k^2(x_j)}{\sum_{j=0}^n t_{k-1}^2(x_j)}.$$

Аппроксимирующая функция $\varphi(x)$ определяется, как и ранее, в виде линейной комбинации базисных функций, в качестве которых задаются полиномы Чебышева дискретной переменной $t_k(x)$:

$$\varphi(x) = \sum_{k=0}^m c_k t_k(x).$$

Тогда, так как матрица Грама является диагональной, коэффициенты c_k этой линейной комбинации определяются как частное от деления правых частей получающейся системы нормальных уравнений на диагональные элементы этой матрицы, то есть

$$c_k = \frac{\sum_{j=0}^n f(x_j) t_k(x_j)}{\sum_{j=0}^n t_k^2(x_j)}.$$

Заметим, что если для улучшения качества аппроксимации возникает необходимость в увеличении числа базисных функций, то нет необходимости пересчитывать коэффициенты c_k , определенные с меньшим значением m .

На практике достаточно часто при обработке экспериментальных данных можно ограничиться построением линейной аппроксимирующей функции (*линии регрессии*), то есть

$$\varphi(x) = a + bx.$$

Для коэффициентов a и b из общего алгоритма МНК получаются выражения

$$a = \bar{f} - b\bar{x}; \quad b = \frac{\sum_{j=1}^n (x_j - \bar{x})(f_j - \bar{f})}{\sum_{j=1}^n (x_j - \bar{x})^2},$$

где

$$\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j; \quad \bar{f} = \frac{1}{n} \sum_{j=1}^n f_j.$$

Погрешность вычисления коэффициентов a и b определяется по формулам

$$\Delta b = T_{\alpha_n} \sqrt{\frac{1}{n-2} \left[\frac{\sum_{j=1}^n (f_j - \bar{f})^2}{\sum_{j=1}^n (x_j - \bar{x})^2} - b^2 \right]};$$

$$\Delta a = \Delta b \sqrt{\bar{x}^2 + \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2},$$

где T_{α_n} — коэффициент Стьюдента для n измерений и уровня значимости α .

7.3. Численное дифференцирование

К численному дифференцированию приходится прибегать в том случае, когда функция $f(x)$, для которой нужно найти производную, задана таблично или функциональная зависимость $f(x)$ имеет очень сложное аналитическое выражение. В первом случае методы дифференциального исчисления просто неприменимы, а во втором случае их использование вызывает значительные трудности.

В тех случаях, когда численное дифференцирование неприменимо, вместо функции $f(x)$ рассматривают интерполяционный многочлен $P_n(x)$ и считают производную от $f(x)$ приближенно равной производной от $P_n(x)$. Естественно, что при этом производная от $f(x)$ будет найдена с некоторой погрешностью.

Пусть требуется найти производную функции $f(x)$ в некоторой точке x , если таблица значений функции $f(x)$ задана в произвольных точках x_j , $j = \overline{0, n}$. В этом случае наиболее удобным является использование в качестве $P_n(x)$ интерполяционного многочлена Ньютона. При этом для построения многочлена Ньютона в качестве первой точки привлекается ближайшая к x табличная точка, а затем остальные, в порядке их удаленности от точки x .

П р и м е р

Если x находится вблизи табличной точки x_k , то для вычисления значения первой производной функции $f(x)$ в точке x получаются следующие выражения:

$$f'(x_k) \approx \frac{f(x_k) - f(x_{k-1}))}{x_k - x_{k-1}}; \quad f'(x_k) \approx \frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k}.$$

Функцию $f(x)$ можно записать в виде

$$f(x) = P_n(x) + r_n(x)$$

и, дифференцируя k раз (в предположении, что $f(x)$ и $P_n(x)$ имеют производные k -го порядка), имеем

$$f^{(k)}(x) = P_n^{(k)}(x) + r_n^{(k)}(x).$$

Так как за приближенное значение $f^{(k)}(x)$ принимается $P_n^{(k)}(x)$, то погрешность дифференцирования есть $r_n^{(k)}(x)$. При замене $f(x)$ интерполяционным многочленом предполагается, что остаточный член $r_n(x)$ мал, но из этого вовсе не следует, что мало $r_n^{(k)}(x)$, так как производные от малой функции могут быть весьма велики. На самом деле практика показывает, что при таком способе вычисления производных

$f^{(k)}(x)$ получается достаточно большая погрешность, особенно при вычислении производных высших порядков.

Если табличные значения являются равноотстоящими, то есть $x_i = x_0 + ih$; $h = (x_n - x_0) / n$; $y_i = f(x_i)$, $i = 0, n$, то для вычисления значения первой производной в точке x , лежащей вблизи табличной точки x_k , получаются следующие выражения:

$$\left. \begin{aligned} f'(x_k) &= \frac{y_{k+1} - y_k}{h} + O(h); \\ f'(x_k) &= \frac{y_k - y_{k-1}}{h} + O(h); \\ f'(x_k) &= \frac{y_{k+1} - y_{k-1}}{2h} + O(h^2), \end{aligned} \right\} \quad (7.38)$$

где $O(h)$, $O(h^2)$ используются для обозначения выражений, имеющих порядок малости, превышающий величины h и h^2 соответственно. Несмотря на то что с помощью последнего выражения из (7.38) получается более точный результат, использовать это выражение для вычисления первой производной нужно с большой осторожностью, так как достаточно часто это приводит к потере устойчивости решения.

Для вычисления второй производной в случае, если точка x находится вблизи табличной точки x_k , используется следующее выражение:

$$f''(x_k) = \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} + O(h^2). \quad (7.39)$$

Аналогично можно получить выражения для вычисления производных и более высоких порядков. Кроме того, можно использовать и другие подходы для вычисления значений производных функций, заданных таблично, например *метод неопределенных коэффициентов*.

Заметим, что замена производных *разностными отношениями* типа (7.38) и (7.39) часто используется при моделировании задач математической физики.

7.4. Численное интегрирование

Будем рассматривать задачу вычисления интеграла при помощи некоторого числа значений интегрируемой функции. Достоинство этого метода состоит в простоте и универсальности.

Пусть $[a, b]$ есть любой конечный или бесконечный отрезок числовой оси и требуется найти приближенное значение интеграла

$$I = \int_a^b F(x) dx$$

по n значениям функции $F(x)$ в точках $x_i, i = \overline{0, n}$. Многие правила численного интегрирования основаны на замене интегрируемой функции $F(x)$ на всем отрезке $[a, b]$ или на его частях на более простую функцию, близкую к $F(x)$, легко интегрируемую точно и принимающую в точках $x_i, i = \overline{0, n}$, те же значения, что и $F(x)$. В качестве такой функции достаточно часто используют *алгебраический многочлен* или *рациональную функцию*. В том случае, если интегрируемая функция $F(x)$ является достаточно гладкой, то можно хорошо приблизить ее многочленом *невысокой степени* или *несложной* рациональной функцией. Если же сама функция $F(x)$ имеет особенности, то это затруднит такое приближение или сделает его вообще невозможным. В этом случае заранее освобождаются от этих особенностей путем их выделения. Для этого функцию $F(x)$ представляют в виде произведения двух функций:

$$F(x) = p(x)f(x),$$

где $p(x)$ имеет те же особенности, что и $F(x)$, и называется *весовой функцией* или *весом*, а $f(x)$ является достаточно гладкой функцией. Тогда задача заключается в вычислении интеграла вида

$$I = \int_a^b p(x)f(x)dx. \quad (7.40)$$

При этом используются правила вычисления интегралов, предназначенные для численного интегрирования функций, имеющих те же особенности, что и весовая функция $p(x)$. Поэтому при вычислении интеграла (7.40) функция $p(x)$ считается фиксированной функцией, а $f(x)$ — любой достаточно гладкой функцией на $[a, b]$.

7.4.1. Интерполяционные квадратурные формулы

Пусть $f(x)$ — достаточно гладкая функция, а интервал $[a, b]$ — конечный и замкнутый. Правило вычисления интеграла будем задавать в виде

$$\int_a^b p(x)f(x)dx = \sum_{k=0}^n A_k f(x_k) + R_n(f). \quad (7.41)$$

Такое правило называется *методом механических квадратур*, сумма — *квадратурной суммой*; $A_k, k = \overline{0, n}$, — *квадратурными коэффициентами*; $x_k, k = \overline{0, n}$, — *квадратурными узлами*.

Функцию $f(x)$ можно приблизить интерполяционным многочленом Лагранжа, который строится по заданным значениям $f(x_i), i = \overline{0, n}$, то есть

$$f(x) = L_n(x) + r_n(x),$$

где

$$L_n(x) = \sum_{k=0}^n f(x_k) \frac{\omega_n(x)}{(x-x_k)\omega'_n(x_k)};$$

$$r_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_n(x);$$

$$\omega_n(x) = (x-x_0)(x-x_1)\dots(x-x_n);$$

ξ — некоторая точка интервала $[a, b]$.

Тогда

$$\begin{aligned} \int_a^b p(x)f(x)dx &= \int_a^b p(x)L_n(x)dx + \int_a^b p(x)r_n(x)dx = \\ &= \sum_{k=0}^n A_k f(x_k) + R_n(f) \end{aligned}$$

и

$$A_k = \int_a^b p(x) \frac{\omega_n(x)}{(x-x_k)\omega'_n(x_k)} dx; \quad (7.42)$$

$$R_n(f) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \int_a^b p(x) \omega_n(x) dx.$$

Квадратурное правило, коэффициенты которого вычисляются согласно (7.42), называется *интерполяционным*. Оно является точным для всех алгебраических многочленов степени не выше n . Поэтому *степень точности* интерполяционного квадратурного правила равна n .

Частным случаем интерполяционного квадратурного правила является квадратурное правило Ньютона-Котеса. Это правило используется в том случае, когда $f(x)$ — достаточно гладкая функция и узлы x_k , $k = \overline{0, n}$, — равноотстоящие. В этом случае считают весовую функцию $p(x) = 1$ и квадратурное правило (7.41) записывают в виде

$$\int_a^b f(x)dx \approx (b-a) \sum_{k=0}^n B_k^n f(x_k). \quad (7.43)$$

Коэффициенты B_k^n в (7.43), учитывая выражение для A_k (7.42), вычисляются по формуле

$$B_k^n = \frac{A_k}{(b-a)} = \frac{(-1)^{n-k}}{nk!(n-k)!} \int_0^n \frac{t(t-1)\dots(t-n)}{(t-k)} dt$$

и имеют следующие конкретные значения:

$$n = 0; \quad B_0^0 = 1;$$

$$n = 1; \quad B_0^1 = B_1^1 = \frac{1}{2};$$

$$n = 2; \quad B_0^2 = B_2^2 = \frac{1}{6}; \quad B_1^2 = \frac{4}{6};$$

$$n = 3; \quad B_0^3 = B_3^3 = \frac{1}{8}; \quad B_1^3 = B_2^3 = \frac{3}{8}$$

и т.д.

Коэффициенты B_k^n вычислены до $n = 20$. Они являются рациональными числами и обладают следующими свойствами:

$$1) \quad \sum_{k=0}^n B_k^n = 1 \quad \text{при каждом } n, \text{ в чем легко убедиться, если в (7.43)}$$

положить $f(x) \equiv 1$;

$$2) \quad B_j^n = B_{n-j}^n;$$

3) при $n = 8$ и для всех $n \geq 10$ среди B_n^k встречаются отрицательные, причем абсолютные величины B_k^n быстро растут с ростом n .

Последнее свойство коэффициентов B_k^n является существенным при определении погрешности вычисления интеграла с помощью квадратурной суммы. Так, если значения подынтегральной функции $f(x_k)$, $k = \overline{0, n}$, известны с абсолютной погрешностью ε , то неустранимая погрешность вычисления интеграла в (7.43) может быть оценена величиной

$$\varepsilon(b-a) \sum_{k=0}^n |B_k^n|,$$

при этом значения $\sum_{k=0}^n |B_k^n|$ при увеличении n быстро растут. Например, при $n = 20$ эта сумма равна 560. Поэтому при больших значениях n незначительные погрешности в значениях функций $f(x_k)$, $k = \overline{0, n}$, могут привести к большой погрешности в квадратурной сумме (7.43). В связи с этим формулы Ньютона-Котеса используются только при малых значениях n . Для уменьшения погрешности результата отрезок $[a, b]$ разбивают на m интервалов. К каждому интервалу применяют квадратурную формулу с малым числом узлов и результаты суммируют. Погрешность метода для формулы Ньютона-Котеса можно представить в виде $(b-a)^p C(a, b)$, где $C(a, b)$ — медленно изменяющаяся функция на $[a, b]$. Погрешность той же формулы, примененной к

отрезку длины $\frac{b-a}{m}$, есть $\left(\frac{b-a}{m}\right)^p C(a, b)$. Тогда после суммирования погрешность результата примет вид $\frac{(b-a)^p}{m^{p-1}} C(a, b)$. Таким образом, в результате разбиения интервала интегрирования на m частей погрешность результата уменьшится в m^{p-1} раз.

Заметим, что если средняя точка интервала $[a, b]$ является узлом квадратурного правила, то алгебраическая точность правила увеличивается на единицу, то есть правило становится точным для многочленов степени $n+1$.

Приведем конкретные формулы Ньютона-Котеса для $n = 0, 1, 2$, которые используются чаще всего. При этом приведем сразу *обобщенные формулы*, полученные делением интервала $[a, b]$ на m частей и суммированием результатов. Кроме того, для записи обобщенных квадратурных формул будем использовать следующие обозначения: $f_i = f(x_i)$;

$h = \frac{b-a}{m}$; ξ — некоторая точка интервала $[a, b]$; $R_n^{ob}(f)$ — погрешность вычисления интеграла.

Квадратурные формулы прямоугольников ($n = 0$)

На интервале $[a, b]$ необходимо выбрать одну любую точку в качестве узла квадратурного правила. Обычно выбирают среднюю или крайние точки и в соответствии с этим получают следующие формулы:

1) *формула левых прямоугольников*:

$$\int_a^b f(x) dx \approx \frac{b-a}{m} [f_0 + f_1 + \dots + f_{m-1}];$$

$$R_{0\text{ лев}}^{ob}(f) = -\frac{(b-a)^2}{2m} f'(\xi);$$

2) *формула правых прямоугольников*:

$$\int_a^b f(x) dx \approx \frac{b-a}{m} [f_1 + f_2 + \dots + f_m];$$

$$R_{0\text{ пр}}^{ob}(f) = \frac{(b-a)^2}{2m} f'(\xi);$$

3) *формула средних прямоугольников*:

$$\int_a^b f(x) dx \approx \frac{b-a}{m} \left[f\left(a + \frac{h}{2}\right) + f\left(a + \frac{3h}{2}\right) + \dots + f\left(a + \frac{2m-1}{2}h\right) \right];$$

$$R_{0\text{ ср}}^{ob}(f) = \frac{(b-a)^3}{24m^2} f''(\xi).$$

Квадратурная формула трапеций ($n = 1$)

$$\int_a^b f(x)dx \approx \frac{b-a}{2m} [f_0 + 2(f_1 + f_2 + \dots + f_{m-1}) + f_m];$$

$$R_1^{of}(f) = -\frac{(b-a)^3}{12m^2} f''(\xi).$$

Квадратурная формула Симпсона ($n = 2$)

$$\int_a^b f(x)dx \approx \frac{b-a}{3m} [f_0 + f_m + 4(f_1 + f_3 + \dots + f_{m-1}) +$$

$$+ 2(f_2 + f_4 + \dots + f_{m-2})];$$

$$R_2^{of}(f) = -\frac{(b-a)^5}{180m^4} f^{(4)}(\xi).$$

При использовании квадратурной формулы Симпсона значение m должно быть четным.

В порядке убывания точности вычисления интеграла квадратурные формулы Ньютона-Котеса при фиксированном m располагаются в следующем порядке: формула Симпсона, средних прямоугольников, трапеций, крайних прямоугольников.

Обычно требуется вычислить значение интеграла с некоторой заданной точностью ε . В этом случае строится последовательность значений интеграла I_1, I_2, \dots , для которой справедливо соотношение

$$\lim_{n \rightarrow \infty} I_n \rightarrow I,$$

где I_n — значение интеграла, вычисленное на n -м шаге. Интеграл считается вычисленным с заданной точностью ε , если для некоторого шага k будет справедливо неравенство

$$|I_{k+1} - I_k| \leq \varepsilon.$$

Для формул Ньютона-Котеса I_1 соответствует значению интеграла, вычисленному при некотором значении m , I_2 — значению интеграла, вычисленному при увеличении значения m в два раза и т.д. Так как узлы квадратурного правила Ньютона-Котеса являются равноотстоящими, то увеличение m в два раза приведет просто к добавлению к старым узлам новых. Поэтому можно организовать процесс вычисления интеграла таким образом, чтобы не пересчитывать полученные ранее значения подынтегральной функции, а только добавлять в квадратурную сумму новые значения.

7.4.2. Квадратурные формулы наивысшей алгебраической степени точности. Квадратурные формулы Гаусса

Если вопрос о точности вычисления интеграла стоит очень остро, то можно воспользоваться другими квадратурными формулами, в частности *квадратурными формулами наивысшей алгебраической степени точности*, которая равна $2n-1$.

Пусть в квадратурном правиле

$$\int_a^b p(x)f(x)dx \approx \sum_{k=1}^n A_k f(x_k) \quad (7.44)$$

$[a, b]$ есть любой конечный или бесконечный отрезок и весовая функция $p(x)$ такова, что ее произведение на любую неотрицательную степень x абсолютно интегрируемо на $[a, b]$:

$$\int_a^b |p(x)x^i|dx < \infty.$$

Кроме того, будем считать функцию $p(x)$ не эквивалентной нулю, то есть

$$\int_a^b p(x)dx > 0.$$

Квадратурное правило (7.44) при фиксированном значении n содержит $2n$ параметров $x_k, A_k, k = \overline{1, n}$, и выбрать их можно так, чтобы равенство (7.44) выполнялось точно для всех алгебраических многочленов степени не выше $2n-1$ или, что равносильно, чтобы выполнялись равенства

$$\int_a^b p(x)x^i dx = \sum_{k=1}^n A_k x_k^i, \quad i = \overline{0, 2n-1}. \quad (7.45)$$

Равенства (7.45) образуют систему из $2n$ уравнений относительно $2n$ неизвестных $x_k, A_k, k = \overline{1, n}$, но в силу того, что данная система является нелинейной, ее решение весьма затруднительно.

Введем многочлен, корнями которого являются узлы квадратурного правила:

$$\omega(x) = (x - x_1)(x - x_2) \dots (x - x_n).$$

Для определения параметров квадратурного правила (7.44) доказаны следующие теоремы.

Теорема 1. Для того чтобы квадратурное правило (7.44) было точным для всех алгебраических многочленов степени не выше $2n-1$, необходимо и достаточно, чтобы выполнялись следующие условия:

1) правило (7.44) было интерполяционным, то есть коэффициенты A_k определялись по формулам

$$A_k = \int_a^b p(x) \frac{\omega(x)}{(x - x_k)\omega'(x_k)} dx; \quad (7.46)$$

2) многочлен $\omega(x)$ был ортогонален на $[a, b]$ по весу $p(x)$ ко всякому многочлену $Q(x)$ степени, меньшей n :

$$\int_a^b p(x) \omega(x) Q(x) dx = 0.$$

Теорема 2. Если весовая функция $p(x)$ не меняет знак на $[a, b]$, то существует и при этом единственный многочлен $\omega(x) = (x - x_1) \times \dots \times (x - x_n)$, ортогональный на $[a, b]$ по весу $p(x)$ ко всякому многочлену $Q(x)$ степени, меньшей n .

Теорема 3. Если весовая функция $p(x)$ не меняет знак на $[a, b]$ и многочлен $\omega(x)$ ортогонален на $[a, b]$ по весу $p(x)$ ко всякому многочлену $Q(x)$ степени, меньшей n , то все корни многочлена $\omega(x)$ действительные, различные и лежат внутри $[a, b]$.

Теорема 4. Если весовая функция $p(x)$ не меняет знак на $[a, b]$, то ни при каком выборе x_k и A_k равенство (7.44) не может быть верным для всех многочленов степени $2n$.

Приведенные теоремы доказывают справедливость следующего утверждения: если весовая функция $p(x)$ сохраняет знак на $[a, b]$, то квадратурное правило (7.44), верное для всех многочленов степени не выше $2n-1$, существует при всех n и является единственным для каждого n . При этом для знакопостоянной весовой функции $p(x)$ степень точности $2n-1$ является наивысшей возможной. Таким образом, для конкретной весовой функции $p(x)$ существует единственный многочлен $\omega(x)$, корни которого являются узлами квадратурного правила, а выражение (7.46) определяет коэффициенты этого правила.

Квадратурное правило наивысшей алгебраической степени точности для постоянной весовой функции $p(x) \equiv 1$ называется *квадратурным правилом Гаусса*. Отрезок интегрирования $[a, b]$ будем считать конечным. Всякий конечный отрезок $[a, b]$ линейной заменой переменных

$$x = \frac{b+a}{2} + \frac{b-a}{2}t$$

может быть преобразован в отрезок $[-1, 1]$. Тогда выражение (7.44) может быть преобразовано к виду

$$\begin{aligned} \int_a^b f(x)dx &= \frac{b-a}{2} \int_{-1}^1 f\left(\frac{b+a}{2} + \frac{b-a}{2}t\right)dt = \\ &= \frac{b-a}{2} \sum_{k=1}^n A_k f\left(\frac{b+a}{2} + \frac{b-a}{2}t_k\right). \end{aligned}$$

Систему многочленов, ортогональную на $[-1, 1]$, образуют *многочлены Лежандра*

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

Таким образом, в квадратурной формуле Гаусса узлы t_k , $k = \overline{1, n}$, должны располагаться в корнях многочлена Лежандра степени n .

Процесс вычисления интеграла методом Гаусса с точностью ε можно организовать следующим образом. Задать конкретное значение n и выписать значения узлов t_k и коэффициентов A_k для этого n . Затем строить последовательность значений интеграла $I_1, I_2, \dots, I_m, \dots$, где величина I_m получается путем суммирования значений интегралов, вычисленных на m отрезках, полученных делением отрезка $[a, b]$ на m частей. При этом каждый частичный отрезок переводится в $[-1, 1]$, и интеграл для каждого частичного отрезка вычисляется при использовании одних и тех же значений t_k , A_k , $k = \overline{1, n}$. Интеграл будем считать вычисленным с заданной точностью ε , если для некоторого m выполнится неравенство

$$|I_{m+1} - I_m| \leq \varepsilon.$$

В справочниках можно найти значения для узлов и коэффициентов квадратурного правила наивысшей алгебраической степени точности для различных весовых функций $p(x)$.

7.5. Решение нелинейных уравнений и систем

Во многих инженерных и научных задачах возникает необходимость решения уравнений или систем вида

$$F(x, p_1, p_2, \dots, p_k) = 0, \quad (7.47)$$

где F — заданная функция либо вектор-функция $F = (F_1, F_2, \dots, F_n)^T$; x — неизвестная переменная либо вектор $x = (x_1, x_2, \dots, x_n)$; p_1, p_2, \dots, p_k — параметры задачи. Как правило, при моделировании возникает необходимость определения решения в зависимости от параметров. При этом каждому фиксированному набору параметров, в соответствии с физическим смыслом конкретной задачи, соответствует либо конечное, либо бесконечное множество решений.

Так, например, в электродинамике при математическом моделировании электромагнитных волновых колебательных процессов в линиях передачи и резонаторах получают так называемое дисперсионное уравнение вида (7.47). В этом случае параметрами $\rho_1, \rho_2, \dots, \rho_k$ являются: частота колебаний, геометрические параметры системы и включения, пространственное распределение диэлектрической и магнитной проницаемостей в электродинамической структуре и т.д. В качестве неизвестного x могут быть выбраны коэффициенты распространения и затухания электромагнитных волн в линиях передачи либо собственные частоты и добротности колебательной системы. Бесконечное множество решений дисперсионного уравнения будет соответствовать бесконечному числу собственных типов волн (колебаний) в исследуемой системе.

Только для простейших нелинейных уравнений и систем удастся найти решение в аналитическом виде. В большинстве случаев приходится решать уравнения и системы вида (7.47) численными методами.

7.5.1. Решение нелинейных уравнений

Пусть дано уравнение

$$f(x) = 0, \quad (7.48)$$

где функция $f(x)$ определена и непрерывна на некотором конечном или бесконечном интервале $[a, b]$. Значение x^* , при котором $f(x^*) = 0$, называется *корнем уравнения* (7.48) или *нулем функции* $f(x)$.

Будем предполагать, что уравнение (7.48) имеет лишь *изолированные корни*, т.е. такие корни x^* , для которых существует окрестность, не содержащая других корней этого уравнения.

Численное решение уравнения (7.48) обычно проводят в два этапа:

- 1) отделение корней — определение таких интервалов изменения переменной x , где находится только один корень;
- 2) определение корня уравнения с заданной точностью ε .

Для отделения корней, т.е. для определения отрезка $[\alpha, \beta]$, содержащего только один корень, полезна следующая теорема математического анализа.

Теорема. Если непрерывная функция $f(x)$ принимает значения разных знаков на концах отрезка $[\alpha, \beta]$, т.е. $f(\alpha)f(\beta) < 0$, то внутри этого отрезка содержится по крайней мере один корень уравнения $f(x) = 0$, т.е. найдется хотя бы одно число $x^* \in [\alpha, \beta]$ такое, что $f(x^*) = 0$. Корень x^* будет единственным, если производная $f'(x)$ существует и сохраняет постоянный знак внутри этого отрезка.

Для определения начального приближения x_0 для реализации итерационных методов нахождения корней или для определения корня уравнения с невысокой точностью можно использовать следующие методы.

Табличный метод

Сначала определяются знаки функции $f(x)$ в граничных точках интервала $[a, b]$, затем определяют знаки функции $f(x)$ в ряде промежуточных точек этого интервала $\alpha_1, \alpha_2, \dots, \alpha_n$, выбор которых учитывает особенности функции $f(x)$. Если окажется, что для некоторой пары точек α_k, α_{k+1} выполняется неравенство

$$f(\alpha_k)f(\alpha_{k+1}) < 0,$$

то в силу приведенной выше теоремы в интервале $[\alpha_k, \alpha_{k+1}]$ имеется хотя бы один корень уравнения $f(x) = 0$. Затем нужно тем или иным способом убедиться, является ли корень единственным. После чего в качестве начального значения x_0 можно взять любую точку интервала $[\alpha_k, \alpha_{k+1}]$.

Метод половинного деления

В методе половинного деления интервал $[a, b]$, на концах которого функция $f(x)$ имеет значения разных знаков, делят пополам и из двух полученных интервалов $\left[a, \frac{a+b}{2}\right]$, $\left[\frac{a+b}{2}, b\right]$ выбирают тот, на концах которого $f(x)$ имеет значения разных знаков. Затем делят пополам выбранный интервал и т.д. Деление интервалов продолжается до тех пор, пока длина последнего интервала не станет меньше заданного значения ε . Число делений интервала k определяется формулой

$$k = \frac{\ln \frac{b-a}{\varepsilon}}{\ln 2},$$

и в качестве x_0 можно взять любую точку последнего выбранного интервала.

Метод пропорциональных частей

Этот метод применяют, если функция $f(x)$ на концах $[a, b]$ принимает значения разных знаков и, кроме того, $f'(x)$ и $f''(x)$ сохраняют постоянные знаки на $[a, b]$. В этом методе за точку деления отрезка $[a, b]$ берут точку

$$x_1 = a - \frac{f(a)}{f(b) - f(a)}(b - a),$$

которая является нулем функции

$$f_1(x) = f(a) + \frac{x-a}{b-a} [f(b) - f(a)].$$

Эта функция является линейным приближением функции $f(x)$ на $[a, b]$, а x_1 — это точка пересечения функцией $f_1(x)$ оси абсцисс. Затем рассматривают отрезки $[a, x_1]$ и $[x_1, b]$ и выбирают тот, на концах которого функция $f(x)$ имеет разные знаки. На выбранном отрезке строят функцию $f_2(x)$ и находят нуль этой функции и т.д. Если требуется найти корень уравнения с невысокой точностью ε , то деление отрезков продолжается до тех пор, пока длина очередного выбранного отрезка не будет меньше ε или не выполнится неравенство

$$|f(x_i)| < \varepsilon,$$

где x_i — очередная точка деления отрезка на два. Эти же критерии можно использовать и для определения начального приближения x_0 .

Графический метод

Начальное приближение x_0 можно определить графически как точку пересечения функцией $y = f(x)$ оси абсцисс. Если $f(x)$ — сложная функция, то $f(x)$ представляется в виде разности двух функций $\psi_1(x)$ и $\psi_2(x)$:

$$f(x) = \psi_1(x) - \psi_2(x),$$

каждую из которых можно достаточно просто изобразить графически. Тогда за x_0 принимают абсциссу точки пресечения функций $y = \psi_1(x)$ и $y = \psi_2(x)$.

Самыми распространенными методами решения нелинейных уравнений являются методы простых итераций и Ньютона.

Метод простых итераций

Сначала требуется привести заданное уравнение $f(x) = 0$ к канонической форме

$$x = \varphi(x),$$

причем для одного уравнения можно построить несколько канонических форм.

П р и м е р

Пусть исходное уравнение имеет вид

$$\cos(x) + x - \ln(x) = 0.$$

Функции $\varphi(x)$ канонических форм для этого уравнения будут следующими:

$$\varphi_1(x) = \ln(x) - \cos(x);$$

$$\varphi_2(x) = \arccos(\ln(x) - x);$$

$$\varphi_3(x) = \exp(\cos(x) + x).$$

Итерационное правило метода простых итераций имеет вид

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, \dots \quad (7.49)$$

Геометрически это правило означает следующее (рис. 7.1). Точное решение x^* является точкой пересечения кривой $y = \varphi(x)$ с биссектрисой $y = x$. За очередное приближение x_{n+1} берется точка, значение которой равно $\varphi(x_n)$. Это значение получается следующим образом: на кривой $y = \varphi(x)$ отмечается точка с координатами $(x_n, \varphi(x_n))$ и проводится прямая, параллельная оси абсцисс до пересечения с биссектрисой. Учитывая, что биссектриса — это множество точек, равноудаленных от осей, опускают перпендикуляр на ось абсцисс и находят точку x_{n+1} .

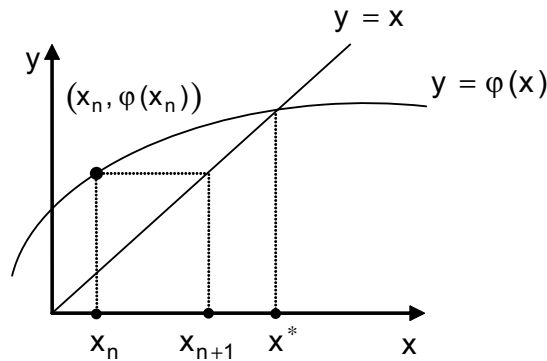


Рис. 7.1

Для проверки сходимости метода простых итераций используются следующие теоремы.

Теорема. Пусть выполняются условия:

- 1) функция $\varphi(x)$ определена на отрезке $[x_0 - \delta, x_0 + \delta]$,
- 2) непрерывна там и удовлетворяет условию Липшица с коэффициентом q , меньшим единицы, т.е. для любых точек $z_1, z_2 \in [x_0 - \delta, x_0 + \delta]$

$$|\varphi(z_1) - \varphi(z_2)| \leq q|z_1 - z_2|, \quad 0 \leq q < 1;$$

- 3) для начального значения x_0 верно неравенство

$$|x_0 - \varphi(x_0)| \leq m;$$

- 4) для чисел δ, q и m выполнено требование

$$\frac{m}{1-q} \leq \delta.$$

Тогда:

- 1) уравнение $x = \varphi(x)$ на отрезке $[x_0 - \delta, x_0 + \delta]$ имеет решение;
- 2) итерационная последовательность приближений x_n , $n = 1, 2, \dots$, может быть построена, принадлежит отрезку $[x_0 - \delta, x_0 + \delta]$ и является сходящейся, т.е. $\lim_{n \rightarrow \infty} x_n = x^*$, при этом x^* является решением уравнения $x = \varphi(x)$;
- 3) для x_n выполняется неравенство

$$|x^* - x_n| \leq \frac{m}{1-q} q^n,$$

которое характеризует скорость сходимости метода простых итераций.

Теорема (о единственности решения). Уравнение $x = \varphi(x)$ на всяком множестве точек, на котором для $\varphi(x)$ выполняется неравенство

$$|\varphi(z_1) - \varphi(z_2)| < |z_1 - z_2|, \quad z_1 \neq z_2,$$

может иметь не более одного решения.

При практическом применении метода простых итераций часто проверяют достаточные условия сходимости итерационного процесса.

Теорема. Пусть функция $\varphi(x)$ определена и дифференцируема на отрезке $[\alpha, \beta]$, причем все ее значения принадлежат этому отрезку. Тогда если существует q такое, что

$$|\varphi'(x)| \leq q < 1$$

для всех $x \in [\alpha, \beta]$, то

- 1) итерационный метод (7.49) сходится независимо от начального приближения $x_0 \in [\alpha, \beta]$;
- 2) предельное значение $x^* = \lim_{n \rightarrow \infty} x_n$ является единственным корнем уравнения $x = \varphi(x)$ на отрезке $[\alpha, \beta]$.

Замечания

1. Благодаря тому, что метод простых итераций сходится при любом выборе начального приближения $x_0 \in [\alpha, \beta]$, он является самоисправляющимся, т.е. отдельная ошибка в вычислениях, не выводящая за пределы $[\alpha, \beta]$, не повлияет на конечный результат, так как ошибочное значение можно рассматривать как новое начальное приближение x_0 .

2. Зависимость погрешности ε_{n+1} на $(n+1)$ -м шаге от погрешности ε_n на n -м шаге выражается соотношением

$$\varepsilon_{n+1} \approx \varphi'(x^*) \varepsilon_n,$$

поэтому говорят, что метод простых итераций сходится почти со скоростью геометрической прогрессии со знаменателем

$$q = \max_{x \in [\alpha, \beta]} |\varphi'(x)|.$$

3. Метод простых итераций является одношаговым, т.е. для реализации итерационного процесса достаточно одного приближения x_0 .

4. Можно построить каноническую форму таким образом, чтобы выполнялись условия сходимости метода простых итераций. Для этого представим функцию $\varphi(x)$ в виде

$$\varphi(x) = x - \lambda f(x),$$

где параметр λ нужно определить таким образом, чтобы выполнялось условие

$$\max_{x \in [\alpha, \beta]} |\varphi'(x)| = q < 1.$$

Введем следующие обозначения:

$$M = \max_{x \in [\alpha, \beta]} |f'(x)|, \quad m = \min_{x \in [\alpha, \beta]} |f'(x)|.$$

Тогда параметр λ можно определить следующим образом:

$$1) \lambda = \frac{1}{M} \operatorname{sign} [f'(x_0)], \text{ при этом } 0 \leq \varphi'(x) < 1;$$

$$2) \lambda = \frac{2}{M} \operatorname{sign} [f'(x_0)], \text{ при этом } -1 < \varphi'(x) < 1.$$

Действительно, так как $\varphi'(x) = 1 - \lambda f'(x)$, то в первом случае

$$q = 1 - \frac{m}{M} < 1, \text{ а во втором — } q = 1 - \frac{2m}{M} < 1.$$

Существуют видоизменения метода простых итераций. Наиболее известными, увеличивающими скорость сходимости, являются следующие методы:

1) *метод секущих* (правило линейной интерполяции):

$$x_{n+1} = \frac{x_{n-1}\varphi(x_n) - x_n\varphi(x_{n-1})}{\varphi(x_n) - x_n - \varphi(x_{n-1}) + x_{n-1}}, \quad n = 1, 2, \dots;$$

2) *метод Стеффенсена*:

$$x_{n+1} = \frac{x_n\varphi[\varphi(x_n)] - \varphi^2(x_n)}{\varphi[\varphi(x_n)] - 2\varphi(x_n) + x_n}, \quad n = 0, 1, \dots.$$

Метод Ньютона

Метод Ньютона применим к решению широкого класса нелинейных уравнений. Идея этого метода заключается в том, что он позволя-

ет решение нелинейного уравнения свести к решению последовательности линейных задач.

Пусть требуется найти точное решение x^* уравнения $f(x) = 0$ при заданном начальном приближении x_0 . Итерационное правило Ньютона имеет вид

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, \dots$$

Геометрически метод Ньютона означает следующее (рис. 7.2): точное решение x^* является точкой пересечения кривой $y = f(x)$ с осью абсцисс. За очередное приближение x_{n+1} принимается точка пересечения касательной к кривой в точке $(x_n, f(x_n))$ с осью абсцисс.

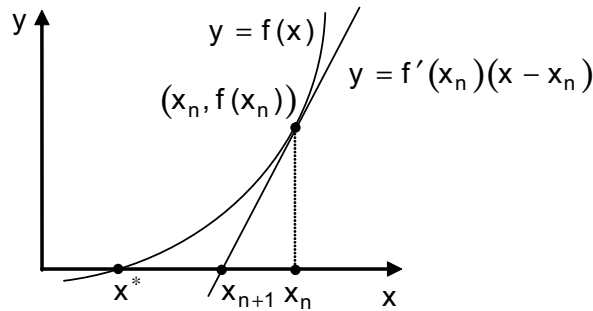


Рис. 7.2

Для проверки сходимости метода Ньютона используются следующие теоремы.

Теорема. Пусть выполнены условия:

1) функция $f(x)$ определена и дважды непрерывно дифференцируема на отрезке $[x_0 - \delta, x_0 + \delta]$, при этом $|f''(x)| \leq K$ для всех x на этом отрезке;

2) $f'(x_0) \neq 0$ и $\frac{1}{|f'(x_0)|} \leq B$;

3) $\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \eta$;

4) для B, K, η соблюдено условие $h = BK\eta \leq \frac{1}{2}$;

5) верно неравенство $\frac{1 - \sqrt{1 - 2h}}{h} \eta \leq \delta$.

Тогда:

- 1) последовательность $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, $n = 0, 1, \dots$, может быть построена и является сходящейся, т.е. $\lim_{n \rightarrow \infty} x_n = x^*$;
- 2) предельное значение x^* есть решение уравнения $f(x) = 0$;
- 3) верна оценка скорости сходимости

$$|x^* - x_n| \leq t^* - t_n,$$

где t_n — ньютонова последовательность приближений

$$t_{n+1} = t_n - \frac{P(t_n)}{P'(t_n)}$$

к меньшему корню t^* уравнения

$$P(t) = \frac{1}{2} K t^2 - \frac{1}{B} t + \frac{\eta}{B} = 0,$$

построенная при $t_0 = 0$.

Теорема. При соблюдении условий предыдущей теоремы о сходимости метода Ньютона для разности $x^* - x_n$ верна оценка

$$|x^* - x_n| \leq \frac{1}{2^{n-1}} (2h)^{2^{n-1}-1} \eta.$$

При практическом использовании метода Ньютона часто проверяют достаточные условия сходимости, определяемые следующей теоремой.

Теорема. Если $f(x)$ определена, дважды дифференцируема в интервале $[\alpha, \beta]$ и принимает значения разных знаков на концах интервала $[\alpha, \beta]$, причем $f'(x)$ и $f''(x)$ отличны от нуля и сохраняют постоянные знаки на $[\alpha, \beta]$, то, исходя из начального приближения $x_0 \in [\alpha, \beta]$, удовлетворяющего неравенству

$$f(x_0)f''(x_0) > 0,$$

можно вычислить методом Ньютона единственный корень x^* уравнения $f(x) = 0$ с любой степенью точности.

Замечание

Зависимость погрешности ε_{n+1} на $(n+1)$ -м шаге от погрешности ε_n на n -м шаге в методе Ньютона выражается соотношением

$$\varepsilon_{n+1} \approx -\frac{1}{2} \frac{f''(x^*)}{f'(x^*)} \varepsilon_n^2,$$

поэтому говорят, что сходимость метода Ньютона является почти квадратичной.

Наиболее известными видоизменениями метода Ньютона, которые уменьшают объем вычислений, являются следующие:

1) *метод секущих*

$$x_{n+1} = \frac{x_{n-1}f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad n = 1, 2, \dots;$$

2) *видоизменение с постоянным значением производной*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \quad n = 0, 1, 2, \dots$$

7.5.2. Метод Лобачевского при решении алгебраических уравнений

Особое место среди нелинейных уравнений занимают алгебраические уравнения, или полиномы n -й степени, которые можно представить в виде

$$P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n = 0. \quad (7.50)$$

Для нахождения корней алгебраического уравнения можно использовать методы простых итераций и Ньютона, но, во-первых, эти методы предназначены для отыскания действительных простых корней, и, во-вторых, даже если нужно найти только действительные корни, необходимо для каждого корня указать интервалы, содержащие изолированные корни и определить начальные приближения.

Для нахождения корней полиномов существует метод Лобачевского. Он не требует определения начальных приближений для корней и позволяет одновременно находить все корни полинома (7.50). Недостатком этого метода является то, что при вычислениях приходится иметь дело с числами, которые сильно различаются по порядкам величин.

Пусть коэффициенты полинома (7.50) a_0, a_1, \dots, a_n являются действительными числами, а корни пронумерованы в следующем порядке:

$$|x_1| \geq |x_2| \geq \dots \geq |x_n|. \quad (7.51)$$

В основании метода Лобачевского лежат следующие соотношения между корнями и коэффициентами алгебраического уравнения (равенства Виетта):

$$\left. \begin{aligned} x_1 + x_2 + \dots + x_n &= -\frac{a_1}{a_0}; \\ x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n &= \frac{a_2}{a_0}; \\ x_1x_2x_3 + \dots + x_{n-2}x_{n-1}x_n &= -\frac{a_3}{a_0}; \\ &\dots \\ x_1x_2x_3\dots x_n &= (-1)^n \frac{a_n}{a_0}. \end{aligned} \right\} \quad (7.52)$$

Будем говорить, что корни x_i , $i = \overline{1, n}$, сильно разделены в смысле отношения их модулей, если модуль предыдущего корня во много раз больше модуля последующего корня:

$$\left| \frac{x_i}{x_{i+1}} \right| \gg 1, \quad i = \overline{1, n-1}. \quad (7.53)$$

Рассмотрим случай, когда все корни являются действительными и различными.

Если выполняются соотношения (7.53), то равенства Виетта (7.52) значительно упрощаются. Так, первое уравнение в (7.52) можно записать в виде

$$x_1 \left(1 + \frac{x_2}{x_1} + \dots + \frac{x_n}{x_1} \right) = -\frac{a_1}{a_0},$$

и так как для корней выполняется соотношение (7.53), то все отношения, стоящие в скобках, будут величинами, пренебрежимо малыми в сравнении с единицей, и ими можно пренебречь. Тогда получим

$$x_1 \approx -\frac{a_1}{a_0}.$$

Аналогичное будет иметь место и для всех остальных равенств Виетта, и (7.52) можно заменить следующей системой приближенных равенств, верных лишь в принятой точности вычислений:

$$\left. \begin{aligned} x_1 &\approx -\frac{a_1}{a_0}; \\ x_1 x_2 &\approx \frac{a_2}{a_0}; \\ x_1 x_2 x_3 &\approx -\frac{a_3}{a_0}; \\ &\dots \\ x_1 x_2 \dots x_n &\approx (-1)^n \frac{a_n}{a_0}. \end{aligned} \right\} \quad (7.54)$$

Тогда из (7.54) следует, что

$$x_1 \approx -\frac{a_1}{a_0}, \quad x_2 \approx -\frac{a_2}{a_1}, \quad x_3 \approx -\frac{a_3}{a_2}, \quad \dots, \quad x_n \approx -\frac{a_n}{a_{n-1}}. \quad (7.55)$$

Таким образом найти сильно разделенные действительные корни алгебраического уравнения достаточно просто. Поэтому решение уравнения (7.50) необходимо начинать с разделения корней. Для этого можно воспользоваться процессом квадрирования, то есть построением последовательности таких полиномов, у которых корни последующего равны квадратам соответствующих корней предыдущего. Это равносильно вычислению коэффициентов для последовательности полиномов по следующим рекуррентным формулам:

$$\left. \begin{aligned} a_0^{(k+1)} &= [a_0^{(k)}]^2; \\ a_1^{(k+1)} &= [a_1^{(k)}]^2 - 2a_0^{(k)}a_2^{(k)}; \\ a_2^{(k+1)} &= [a_2^{(k)}]^2 - 2a_1^{(k)}a_3^{(k)} + 2a_0^{(k)}a_4^{(k)}; \\ &\dots \\ a_n^{(k+1)} &= [a_n^{(k)}]^2. \end{aligned} \right\}$$

Процесс квадрирования можно прекратить, если в пределах принятой точности выполняются соотношения

$$a_i^{(k+1)} \approx [a_i^{(k)}]^2, \quad i = \overline{0, n}.$$

Тогда для полинома

$$P^{(k)}(x) = a_0^{(k)}x^n + a_1^{(k)}x^{n-1} + \dots + a_n^{(k)}, \quad (7.56)$$

в силу разделенности его корней, выполняются соотношения, аналогичные (7.55), а модули приближенных значений корней исходного уравнения (7.50) можно определить из следующих равенств:

$$|x_1|^m = \left| \frac{a_1^{(k)}}{a_0^{(k)}} \right|, \quad |x_2|^m = \left| \frac{a_2^{(k)}}{a_1^{(k)}} \right|, \quad \dots, \quad |x_n|^m = \left| \frac{a_n^{(k)}}{a_{n-1}^{(k)}} \right|,$$

где $m = 2^k$, при этом знаки корней определяются подстановкой в исходное уравнение.

Если корни уравнения (7.50) все действительные и среди них есть равные по абсолютной величине, например x_2 и x_3 , то для уравнения (7.56), полученного после квадрирования, будут справедливы следующие приближенные равенства:

$$\left. \begin{aligned} x_1^m &\approx -\frac{a_1^{(k)}}{a_0^{(k)}}; \\ 2x_1^m x_2^m &\approx -\frac{a_2^{(k)}}{a_0^{(k)}}; \\ x_1^m x_2^{2m} &\approx -\frac{a_3^{(k)}}{a_0^{(k)}}; \\ &\dots \\ x_1^m x_2^{2m} \dots x_n^m &\approx (-1)^n \frac{a_n^{(k)}}{a_0^{(k)}}. \end{aligned} \right\} \quad (7.57)$$

Тогда из второго равенства в (7.57) следует, что

$$|x_2|^m = \frac{1}{2} \left| \frac{a_2^{(k)}}{a_0^{(k)}} \right|,$$

а из третьего —

$$|x_2|^{2m} = \left| \frac{a_3^{(k)}}{a_0^{(k)}} \right|,$$

и любое из этих равенств можно использовать для определения модулей корней x_2 и x_3 , знаки которых, как и ранее, определяются подстановкой в исходное уравнение.

Если уравнение имеет комплексные корни, то они будут попарно сопряжены, так как все коэффициенты уравнения действительные. Пусть, например, x_2 и x_3 — пара комплексно сопряженных корней. Эти корни можно представить в виде

$$x_2 = re^{i\varphi}, \quad x_3 = re^{-i\varphi}.$$

Так как $x_2^m + x_3^m = 2r^m \cos(m\varphi)$ и $x_2^m x_3^m = r^{2m}$, то для уравнения (7.56), полученного после квадрирования, будут справедливы следующие приближенные равенства:

$$\left. \begin{aligned} x_1^m &\approx -\frac{a_1^{(k)}}{a_0^{(k)}}; \\ 2x_1^m r^m \cos(m\varphi) &\approx \frac{a_2^{(k)}}{a_0^{(k)}}; \\ x_1^m r^{2m} &\approx -\frac{a_3^{(k)}}{a_0^{(k)}}; \\ x_1^m r^{2m} x_4^m &\approx \frac{a_4^{(k)}}{a_0^{(k)}}; \\ &\dots \\ x_1^m r^{2m} \dots x_n^m &\approx (-1)^n \frac{a_n^{(k)}}{a_0^{(k)}}. \end{aligned} \right\}$$

Тогда модуль комплексных чисел можно определить из соотношения

$$r^{2m} = \left| \frac{a_3^{(k)}}{a_1^{(k)}} \right|,$$

а для определения аргумента можно воспользоваться первым соотношением Виетта в (7.52), которое в данном случае будет иметь вид

$$x_1 + 2r \cos(\varphi) + x_4 + \dots + x_n = -\frac{a_1}{a_0}.$$

Отсюда определяется значение $\cos(\varphi)$, а $\sin(\varphi) = \sqrt{1 - \cos^2(\varphi)}$.

Таким образом, нахождение корней алгебраического уравнения методом Лобачевского осуществляется в несколько этапов.

1. Разделение корней путем квадрирования, при этом по поведению коэффициентов, получаемых в процессе квадрирования, делается вывод о том, какими являются корни уравнения:

- если все коэффициенты стремятся к квадратам соответствующих коэффициентов, полученных на предыдущем шаге процесса квадрирования, то все корни действительные и различные;
- если какой-то коэффициент стремится к квадрату соответствующего коэффициента, деленному на целое число, то номер этого коэффициента указывает на номер первого из равных по модулю корней

уравнения в соответствии с нумерацией (7.51), а целое число указывает на количество таких корней;

- если какой-то коэффициент меняет знак в процессе квадрирования, то это указывает на наличие комплексно сопряженных корней, причем номер этого коэффициента указывает на первый из таких корней.

2. Вычисление значений корней по формулам, которые соответствуют сделанным выводам о виде корней.

7.5.3. Решение систем нелинейных уравнений

Пусть требуется найти решение $x_1^*, x_2^*, \dots, x_n^*$ системы нелинейных уравнений

$$\left. \begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0; \\ f_2(x_1, x_2, \dots, x_n) &= 0; \\ &\dots \\ f_n(x_1, x_2, \dots, x_n) &= 0. \end{aligned} \right\} \quad (7.58)$$

Введем следующие обозначения:

$$x = (x_1, \dots, x_n)^T; \quad f(x) = (f_1(x), \dots, f_n(x))^T,$$

где T — символ транспонирования.

Тогда система (7.58) запишется в виде

$$f(x) = 0. \quad (7.59)$$

Для решения систем нелинейных уравнений используются методы простых итераций и Ньютона, которые являются естественным расширением этих методов для решения уравнений на многомерный случай.

Для метода простых итераций систему (7.59) необходимо привести к канонической форме

$$x = \varphi(x),$$

где $\varphi(x) = (\varphi_1(x), \dots, \varphi_n(x))^T$. Итерационное правило метода простых итераций записывается в виде

$$x^{(k+1)} = \varphi(x^{(k)}), \quad k = 0, 1, \dots$$

Итерационное правило метода Ньютона имеет вид

$$x^{(k+1)} = x^{(k)} - [f'(x^{(k)})]^{-1} f(x^{(k)}), \quad k = 0, 1, \dots,$$

где $[f'(x^{(k)})]^{-1}$ — обратная матрица к матрице Якоби вида

$$f'(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \dots & \frac{\partial f_2(x)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \dots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix}.$$

Для реализации этих методов необходимо отделить корни и задать начальное приближение $x^{(0)}$. Проверка сходимости методов достаточно сложна.

7.6. Решение задач матричной алгебры

В матричной алгебре рассматриваются четыре класса основных задач: решение систем линейных алгебраических уравнений, вычисление определителей, нахождение обратных матриц, определение собственных значений и собственных векторов матриц. Все эти задачи имеют важное прикладное значение при моделировании в различных областях науки и техники. Кроме того, задачи матричной алгебры являются вспомогательными при реализации многих алгоритмов вычислительной математики, математической физики, обработки результатов экспериментальных исследований.

7.6.1. Некоторые понятия матричной алгебры

Пусть $x = (x_1, \dots, x_n)^T \in R^n$ — вектор n -мерного пространства, $A = (a_{ij}) \in R^{n \times n}$ — матрица порядка n .

Одним из важных понятий матричной алгебры является понятие нормы векторов и матриц.

Нормой вектора x называется действительное число, обозначаемое $\|x\|$, удовлетворяющее условиям:

- 1) $\|x\| > 0$, если $x \neq 0$ и $\|0\| = 0$;
- 2) $\|c \cdot x\| = |c| \cdot \|x\|$ при любом численном множителе c ;
- 3) $\|x + y\| \leq \|x\| + \|y\|$.

Из такого определения нормы вектора непосредственно следует, что $\|x - y\| \geq \left| \|x\| - \|y\| \right|$.

Говорят, что последовательность векторов $x^{(k)}$ сходится к вектору x по норме, если $\lim_{k \rightarrow \infty} \|x - x^{(k)}\| = 0$.

Вводить норму вектора можно различными способами, но при этом должны выполняться условия 1–3 определения нормы.

Наиболее часто используются следующие определения нормы вектора.

1. Первая (кубическая) норма:

$$\|x\|_1 = \max_i |x_i|.$$

Множество векторов вещественного пространства, для которых $\|x\|_1 \leq 1$ заполняет единичный куб.

2. Вторая (октаэдрическая) норма:

$$\|x\|_{11} = \sum_{i=0}^n |x_i|.$$

Множество векторов вещественного пространства, для которых $\|x\|_{11} \leq 1$, заполняет n -мерный аналог октаэдра (восьмиугольника).

3. Третья (сферическая, или евклидова) норма:

$$\|x\|_{111} = \sqrt{(x, x)} = \sqrt{\sum_{i=0}^n |x_i|^2}.$$

Это длина вектора. Совокупность векторов, для которых $\|x\|_{111} \leq 1$, заполняет шар единичного радиуса.

Скалярное произведение векторов $x = (x_1, \dots, x_n)^T$ и $y = (y_1, \dots, y_n)^T$ вводится по формуле

$$(x, y) = \sum_{i=1}^n y_i x_i^*,$$

где x_i^* — компоненты вектора, комплексно сопряженного вектору x .

Норма вектора может быть определена многими способами в зависимости от условий задачи и целей исследования, но при всяком определении она должна удовлетворять трем вышеприведенным условиям, которые являются аксиомами нормы.

Нормой матрицы A называют число $\|A\|$, удовлетворяющее условиям:

- 1) $\|A\| = 0$, если $A \neq 0$ и $\|0\| = 0$;
- 2) $\|c \cdot A\| = |c| \cdot \|A\|$ для всякого числового множителя c ;
- 3) $\|A + B\| \leq \|A\| + \|B\|$;
- 4) $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

Для нормы матрицы верно неравенство

$$\|A - B\| \geq \left| \|A\| - \|B\| \right|.$$

В большинстве случаев приходится одновременно рассматривать матрицы и векторы, потому их нормы рационально вводить так, чтобы они были в какой-то мере согласованными. Обычно говорят, что норма матрицы A согласована с нормой вектора, если для всякого вектора x размерности n выполняется неравенство

$$\|Ax\| \leq \|A\| \|x\|. \quad (7.60)$$

Среди согласованных норм матрицы часто выбирают наименьшую. Так как случай нулевого вектора интереса не имеет, то неравенство

$$(7.60) \text{ можно записать в виде } \|Ay\| \leq \|A\|, \text{ где } y = \frac{x}{\|x\|}, \text{ т.е. } \|y\| = 1.$$

Последнее означает, что согласованная норма матрицы должна быть верхней границей норм векторов Ay при условии, что $\|y\| = 1$. Наименьшей согласованной нормой будет точная верхняя граница множества значений $\|Ay\|$ при $\|y\| = 1$:

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Эта норма называется нормой матрицы, подчиненной норме вектора.

Наиболее распространенными подчиненными нормами матриц являются следующие определения норм матриц.

1. Первая (кубическая) норма:

$$\|A\|_1 = \max_i \sum_{j=1}^n |a_{ij}|.$$

2. Вторая (октаэдрическая) норма:

$$\|A\|_{11} = \max_j \sum_{i=1}^n |a_{ij}|.$$

3. Третья (сферическая, или евклидова) норма:

$$\|A\|_{111} = \sqrt{\sum_{i=1}^n |a_{ij}|^2} = \sqrt{\mu_1},$$

где μ_1 — наибольшее по модулю собственное значение матрицы A^*A , где A^* — матрица, сопряженная A .

Важным понятием матричной алгебры является понятие собственного значения матрицы. Собственным значением (или характеристическим числом) квадратной матрицы A называется такое число λ , что для некоторого ненулевого вектора x имеет место равенство

$$Ax = \lambda x. \quad (7.61)$$

Любой ненулевой вектор x , удовлетворяющий этому равенству, называется собственным вектором матрицы A , соответствующим (или принадлежащим) собственному значению λ . Все собственные векторы матриц определены с точностью до числового множителя.

Перепишем систему (7.61) в виде

$$(A - \lambda I)x = 0, \quad (7.62)$$

где I — единичная матрица соответствующей размерности.

Условием существования для однородной системы ненулевого решения является равенство нулю ее определителя:

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Это уравнение обычно называют *вековым*, или *характеристическим, уравнением* матрицы A .

Левая часть векового уравнения

$$|A - \lambda I| = (-1)^n (\lambda^n - p_1 \lambda^{n-1} - \dots - p_n)$$

называется *характеристическим многочленом* матрицы A . Многочлен

$$P(\lambda) = \lambda^n - p_1 \lambda^{n-1} - \dots - p_n,$$

отличающийся от характеристического многочлена множителем $(-1)^n$, называется *собственным многочленом* матрицы A . Собственные значения матрицы являются корнями собственного многочлена. Совокупность всех собственных значений $\lambda_1, \lambda_2, \dots, \lambda_n$ матрицы A , где каждое собственное значение выписано столько раз, какова его кратность как корня собственного многочлена, называется *спектром* матрицы A . Собственными векторами x_i матрицы A являются нетривиальные решения однородной системы (7.62), в которой вместо λ подставлены собственные значения λ_i , $i = \overline{1, n}$.

Нахождение всех собственных значений и соответствующих собственных векторов матрицы называется решением *полной проблемы собственных значений*, нахождение части собственных значений — решением *частичной проблемы собственных значений*.

Полезно знать следующие соотношения между элементами матрицы A и ее собственными значениями:

$$\begin{aligned} \lambda_1 + \lambda_2 + \dots + \lambda_n &= \text{Sp}A = \text{tr}A; \\ \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n &= |A|, \end{aligned} \quad (7.63)$$

где $\text{Sp}A$ (или $\text{tr}A$) обозначает *след матрицы*, который равен сумме ее диагональных элементов. Из (7.63), в частности, следует, что матрица A имеет хотя бы одно собственное значение, равное нулю, если определитель этой матрицы равен нулю, т.е. если она *особенная*.

7.6.2. Обусловленность систем и матриц

Важное значение при решении задач матричной алгебры имеет понятие *обусловленности систем и матриц*.

Пусть δ — погрешность исходных данных, а ε — погрешность решения системы линейных алгебраических уравнений. Оценим зависимость погрешности решения от погрешности исходных данных неравенством

$$\varepsilon \leq k \delta. \quad (7.64)$$

Число k характеризует зависимость максимально возможных ошибок решения от ошибок исходных данных задачи, и зная его, можно разумно выбрать допустимую погрешность метода и необходимую точность проведения вычислений. Действительно, если исходные данные известны с погрешностью δ , то нецелесообразно искать решение задачи с точностью, намного превосходящей $k \delta$.

Если k «не очень велико», то будем говорить, что система *хорошо обусловлена*, если k «велико» — то система *плохо обусловлена*. Из (7.64) следует, что в плохо обусловленной системе малым погрешностям исходных данных отвечают большие погрешности в решении.

Пусть система линейных алгебраических уравнений задана в виде

$$Ax = b, \quad (7.65)$$

где A — невырожденная вещественная матрица.

Обозначим δA матрицу; δb — вектор малых погрешностей элементов матрицы A и вектора свободных членов b ; δx — вектор погрешностей решения. Тогда для относительной погрешности решения справедлива следующая оценка:

$$\frac{\|\delta x\|}{\|x\|} \leq \sqrt{\frac{\mu_1}{\mu_n}} \left\{ \frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|} \right\},$$

где μ_1 и μ_n — максимальное и минимальное собственные значения матрицы $A^* A$. Таким образом, обусловленность системы (7.65) зави-

сит от величины отношения $\sqrt{\frac{\mu_1}{\mu_n}}$.

Замечание

Если матрица A симметрическая, то собственные значения матрицы $A^* A$ равны квадратам собственных значений матрицы A , т.е.

$$\sqrt{\frac{\mu_1}{\mu_n}} = \frac{\lambda_1}{\lambda_n},$$

где λ_1, λ_n — соответственно максимальное и минимальное собственные значения матрицы A .

Пусть δ — величина *невязки* (результат подстановки найденного решения в исходную систему), т.е.

$$\delta = Ax - b. \quad (7.66)$$

Если (7.66) переписать в виде

$$Ax = b + \delta,$$

то, если система плохо обусловлена, малым δ может соответствовать большая погрешность в решении. Таким образом, оценка точности решения системы по невязкам применима только к хорошо обусловленным системам. В то же время проверка точности определения обратной матрицы не зависит от ее обусловленности. Действительно, если \tilde{A}^{-1} — приближенное значение обратной к A матрицы, тогда невязка решения задачи обращения матрицы δ равна

$$\delta = A\tilde{A}^{-1} - I$$

и

$$\frac{\|\delta A^{-1}\|}{\|A^{-1}\|} \leq \delta,$$

где $\delta A^{-1} = \tilde{A}^{-1} - A^{-1}$ — погрешность вычисления обратной матрицы.

Из геометрической интерпретации решения системы линейных алгебраических уравнений, где рассматривается решение системы как точка пересечения гиперплоскостей $L_i = \{A_i x - b_i = 0, i = \overline{1, n}\}$, следует, что если

$$|\det A| \approx \prod_{i=1}^n \|A_i\|,$$

то система хорошо обусловлена, если

$$\frac{\prod_{i=1}^n \|A_i\|}{|\det A|} \gg 1,$$

то система плохо обусловлена. Здесь A_i — вектор, являющийся i -й вектор-строкой матрицы A .

Для проверки обусловленности системы и матрицы A можно использовать *число обусловленности* матрицы A

$$v = \|A\| \cdot \|A^{-1}\|. \quad (7.67)$$

Из (7.67) следует, что если матрица A близка к особенной, то число v будет велико и матрица A будет плохо обусловленной, если v мало, то матрица A будет хорошо обусловленной. Как правило, система с плохо обусловленной матрицей будет плохо обусловленной системой и наоборот.

Значение числа ν зависит от выбора нормы матрицы A . Если A — вещественная матрица и используется третье определение нормы, то

$$\nu_{III} = \|A\|_{III} \cdot \|A^{-1}\|_{III} = \sqrt{\frac{\mu_1}{\mu_n}}.$$

Ввиду важности понятия обусловленности систем и матриц приведем численный пример.

П р и м е р

Рассмотрим систему

$$Ax = b,$$

которая является плохо обусловленной, где

$$A = \begin{pmatrix} 5 & 7 & 6 & 5 \\ 7 & 10 & 6 & 7 \\ 6 & 8 & 16 & 9 \\ 5 & 7 & 9 & 10 \end{pmatrix}; \quad b = \begin{pmatrix} 23 \\ 32 \\ 39 \\ 31 \end{pmatrix}.$$

Точное решение системы: $x^* = (1 \ 1 \ 1 \ 1)^T$.

Точное значение обратной матрицы:

$$A^{-1} = \begin{pmatrix} 12,065 & -8,097 & -0,548 & 0,129 \\ -8,097 & 5,645 & 6,323 & -0,194 \\ -0,548 & 6,323 & 0,101 & -0,097 \\ 0,129 & -0,194 & -0,097 & 0,258 \end{pmatrix}.$$

Если внести погрешность $\delta_1 = 0,01$ только в один элемент a_{11} матрицы A , то значение обратной матрицы будет равно

$$A^{-1} + \delta A^{-1} = \begin{pmatrix} 10,766 & -7,225 & -0,489 & 0,115 \\ -7,225 & 5,06 & 0,283 & -0,184 \\ -0,489 & 0,283 & 0,159 & 0,096 \\ 0,115 & -0,184 & 0,096 & 0,258 \end{pmatrix}.$$

Решение системы также изменится и будет равно

$$(A^{-1} + \delta A^{-1})b = (0,892 \ 1,072 \ 1,005 \ 0,995)^T.$$

Если внести погрешность $\delta_2 = 0,01$ в вектор свободных членов, т.е. $(b + \delta b) = (23 + \delta_2 \ 32 - \delta_2 \ 39 - \delta_2 \ 31 - \delta_2)^T$, то решение системы будет следующим:

$$A^{-1}(b + \delta b) = (1,206 \ 0,861 \ 0,991 \ 1,002)^T.$$

Собственные числа матрицы A равны: 5,468; 0,057; 3,08; 32,398,

где

$$b_{1j} = \frac{a_{1j}}{a_{11}}, \quad j = \overline{2, n}; \quad q_1 = \frac{b_1}{a_{11}}.$$

Исключим x_1 из остальных уравнений системы. Для этого умножим (7.68) последовательно на $a_{21}, a_{31}, \dots, a_{n1}$ и вычтем из остальных уравнений системы. Преобразованные уравнения будут иметь вид

[illegible]

где $a_{ij,1} = a_{ij} - b_{ij}a_{i1}$, $i, j = \overline{2, n}$; $b_{1,1} = b_1 - a_{11}q_1$.

К полученной системе применим такое же преобразование, т.е. выберем уравнение и неизвестное с коэффициентом, отличным от нуля, приведем этот коэффициент к единице, исключим неизвестное из прочих уравнений и так до тех пор, пока такие преобразования возможны.

В результате придем к одной из двух ситуаций.

1. После n шагов преобразований получим систему вида

[illegible]

Решение полученной системы осуществляется снизу вверх следующим образом:

$$\left. \begin{aligned} x_n &= q_n; \\ x_i &= q_i - \sum_{j=n}^{i-1} b_{ij} x_j, \quad i = \overline{n-1, 1}. \end{aligned} \right\} \quad (7.71)$$

2. После шага преобразований $m < n$ система приняла вид

[illegible]

Тогда если среди элементов $b_{m+1,m}, \dots, b_{n,m}$ есть отличные от нуля, то система (7.69) не имеет решения; если все $b_{j,m} = 0$, $j = \overline{m+1, n}$, то система имеет бесчисленное множество решений (неизвестные x_{m+1}, \dots, x_n могут принимать любые значения, а x_1, \dots, x_m выражают-

ся через них).

Приведение системы к виду (7.70) называется *прямым ходом метода Гаусса*, а нахождение ее решения (7.71) — *обратным*. Заметим, что на каждом шаге прямого хода метода Гаусса выбираются уравнение и неизвестное, подлежащие исключению из прочих уравнений. Это равносильно выбору коэффициента для очередного шага преобразований. Этот коэффициент называется *ведущим*, и он должен быть отличным от нуля. Во избежание большой потери точности рекомендуется осуществлять такую перестановку уравнений, чтобы ведущий коэффициент являлся либо максимальным по модулю коэффициентом во всей системе, либо максимальным по модулю коэффициентом в выбранном уравнении. Такая процедура называется *методом Гаусса с выбором главного элемента*.

Метод Гаусса применим к вычислению определителей и обратных матриц. Так, значение определителя Δ матрицы A равно произведению ведущих коэффициентов:

$$\Delta = a_{11}a_{22} \dots a_{nn},$$

а вычисление обратной матрицы осуществляется одновременным решением n систем:

$$Ax^{(k)} = e_k, \quad k = \overline{1, n}, \quad (7.72)$$

где e_k — единичный вектор (вектор, у которого все элементы, кроме k -го, равны нулю, а k -й равен единице), $x^{(k)}$ — k -й столбец матрицы A^{-1} . Невозможность вычисления обратной матрицы проявляется в невозможности решения какой-либо системы из (7.72).

7.6.4. Метод квадратного корня

Пусть матрица A системы $Ax = b$ является эрмитовой, т.е. $A = A^*$ ($a_{ij} = \overline{a_{ji}}$, $i, j = \overline{1, n}$). Алгоритм метода квадратного корня заключается в том, что матрицу A системы представляют в виде произведения диагональной и треугольных матриц:

$$A = S^*DS, \quad (7.73)$$

где

$$S = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ 0 & s_{22} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & s_{nn} \end{pmatrix}; \quad D = \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{nn} \end{pmatrix}.$$

Исходя из уравнения (7.73) и учитывая, что матрица A эрмитова, получим $\frac{n(n+1)}{2}$ независимых уравнений для определения $\frac{n(n+1)}{2} + n$ неизвестных элементов матриц S и D . Таким образом, имеется бесчисленное множество возможностей представить матрицу A в виде

произведения матриц S^* , D и S . Будем строить такое представление матрицы A , при котором все диагональные элементы матрицы S станут вещественными и положительными.

Запишем матричное равенство (7.73) поэлементно, причем, учитывая, что матрица A эрмитова, запишем равенства для элементов, стоящих на главной диагонали и выше ее. Уравнения при равенстве элементов, стоящих выше главной диагонали, имеют вид

$$\bar{s}_{1i} d_{11} s_{1j} + \bar{s}_{2i} d_{22} s_{2j} + \dots + \bar{s}_{ii} d_{ii} s_{ij} = a_{ij}, \quad i < j, \quad (7.74)$$

а при равенстве диагональных элементов —

$$|s_{1i}|^2 d_{11} + |s_{2i}|^2 d_{22} + \dots + |s_{ii}|^2 d_{ii} = a_{ii}, \quad i = j. \quad (7.75)$$

Из (7.74), (7.75) будем построчно определять элементы матриц S и D . Из уравнения (7.75) при $i = j = 1$ получим

$$|s_{11}|^2 d_{11} = a_{11}.$$

Тогда, полагая $d_{11} = \text{sign}(a_{11})$, элемент s_{11} будет определяться следующим образом:

$$s_{11} = \sqrt{|a_{11}|},$$

и из уравнения (7.74) при $j = \overline{2, n}$ получим остальные элементы первой строки:

$$s_{1j} = \frac{a_{1j}}{d_{11} s_{11}}.$$

Затем из уравнения (7.75) при $i = j = 2$ получим

$$|s_{12}|^2 d_{11} + |s_{22}|^2 d_{22} = a_{22},$$

откуда следует, что

$$d_{22} = \text{sign}(a_{22} - |s_{12}|^2 d_{11});$$

$$s_{2j} = \frac{a_{2j} - \bar{s}_{12} d_{11} s_{1j}}{d_{22} s_{22}}, \quad j = \overline{3, n}.$$

Таким образом, вычисление элементов матриц S и D осуществляется построчно по следующим формулам:

$$\begin{aligned} d_{11} &= \text{sign}(a_{11}); \quad s_{11} = \sqrt{|a_{11}|}; \quad s_{1j} = \frac{a_{1j}}{d_{11} s_{11}}, \quad j = \overline{2, n}, \\ d_{ii} &= \text{sign}\left(a_{ii} - \sum_{p=1}^{i-1} |s_{pi}|^2 d_{pp}\right); \quad s_{ii} = \sqrt{\left|a_{ii} - \sum_{p=1}^{i-1} |s_{pi}|^2 d_{pp}\right|}, \quad i > 1, \\ s_{ij} &= \frac{a_{ij} - \sum_{p=1}^{i-1} \bar{s}_{pi} d_{pp} s_{pj}}{d_{ii} s_{ii}}, \quad i < j; \quad j = i+1; i+2, \dots, n. \end{aligned} \quad (7.76)$$

Затем решение системы $Ax = b$ осуществляется в два этапа.

1. Решается система $Bu = \overline{b}$, где $B = S^*D$ — левая треугольная матрица с элементами β_{ij} , $i, j = \overline{1, n}$, $i \geq j$:

$$y_1 = \frac{b_1}{\beta_{11}}, \quad y_k = \frac{b_k - \sum_{p=1}^{k-1} \beta_{kp} y_p}{\beta_{kk}}, \quad k = \overline{2, n}.$$

2. Решается система $Sx = y$ с правой треугольной матрицей S :

$$x_n = \frac{y_n}{s_{nn}}, \quad x_k = \frac{y_k - \sum_{p=k+1}^n s_{kp} x_p}{s_{kk}}, \quad k = \overline{n-1, 1}.$$

Если матрица A системы вещественная и симметрическая, тогда ее можно разложить в произведение двух матриц:

$$A = S^T S,$$

где S — правая треугольная матрица. При этом формулы (7.76) упростятся и будут иметь вид

$$\begin{aligned} s_{11} &= \sqrt{a_{11}}; & s_{1j} &= \frac{a_{1j}}{s_{11}}, & j &= \overline{2, n}, \\ s_{ii} &= \sqrt{a_{ii} - \sum_{p=1}^{i-1} s_{pi}^2}, & i &> 1, \\ s_{ij} &= \frac{a_{ij} - \sum_{p=1}^{i-1} s_{pi} s_{pj}}{s_{ii}}, & i < j; & j = \overline{i+1, n}. \end{aligned}$$

Заметим, что элементы матрицы S будут вещественными и положительными только для положительно определенной матрицы A . В противном случае среди элементов матрицы S могут быть комплексные.

Метод квадратного корня эффективен при решении систем с положительно определенной эрмитовой матрицей. Такие задачи возникают, как правило, при минимизации положительно определенных квадратичных форм.

7.6.5. Итерационные методы решения систем линейных алгебраических уравнений

Эти методы дают решение системы в виде предела последовательности некоторых векторов, построение которых осуществляется посредством единообразного процесса, называемого процессом итераций.

В современной литературе описано большое количество итерационных методов, основанных на различных принципах. Как правило, вычислительные схемы таких методов просты и удобны для реализации на ЭВМ. Однако каждый итерационный процесс имеет свою ограниченную область применения, так как:

1) процесс итераций может оказаться расходящимся для данной системы;

2) сходимость процесса итераций для данной системы может быть настолько медленной, что практически невозможно достичь удовлетворительной близости к точному решению за приемлемое время.

Пусть дана система линейных алгебраических уравнений

$$Ax = f \quad (7.77)$$

с неособенной матрицей A . Точное решение системы x^* является пределом последовательности векторов $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$, которые строятся по рекуррентным формулам

$$x^{(k)} = x^{(k-1)} + H^{(k)} (f - Ax^{(k-1)}) \quad (7.78)$$

или

$$x^{(k)} = C^{(k)} x^{(k-1)} + z^{(k)}, \quad (7.79)$$

где $H^{(k)}, C^{(k)}$ — последовательности матриц; $z^{(k)}$ — последовательности векторов, $k = 0, 1, \dots$. При этом для линейных систем в качестве начального приближения $x^{(0)}$ можно взять любой вектор соответствующей размерности.

Итерационные процессы, протекающие по формулам (7.78) и (7.79), обладают тем свойством, что для каждого из них точное решение x^* является неподвижной точкой. Это значит, что если за начальное приближение $x^{(0)}$ взято x^* , то все последующие приближения будут также равны x^* .

Простейшими среди итерационных процессов являются стационарные, в которых матрицы в (7.78) и (7.79) не зависят от номера шага k .

Метод последовательных приближений

Под процессом последовательных приближений понимается следующий итерационный процесс. Система уравнений (7.77) записывается в виде

$$x = Bx + g,$$

где $B = I - A$; I — единичная матрица соответствующего порядка; $g = f$. Последовательные приближения вычисляются по формуле

$$x^{(k+1)} = Bx^{(k)} + g. \quad (7.80)$$

или в матричной форме:

$$x = Bx + g, \quad (7.82)$$

где

$$B = \begin{pmatrix} 0 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & 0 \end{pmatrix}; \quad g = \begin{pmatrix} \frac{f_1}{a_{11}} \\ \frac{f_2}{a_{22}} \\ \dots \\ \frac{f_n}{a_{nn}} \end{pmatrix}.$$

Преобразование исходной системы (7.81) в систему (7.82) равносильно матричному преобразованию

$$B = I - D^{-1}A; \quad g = D^{-1}f,$$

где $D = \text{diag}\{a_{11}, \dots, a_{nn}\}$ — диагональная матрица, на диагонали которой находятся элементы $a_{11}, a_{22}, \dots, a_{nn}$ матрицы A и

$$D^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{a_{nn}} \end{pmatrix}.$$

На основании приведенных выше теорем для сходимости метода простых итераций необходимо и достаточно, чтобы все собственные значения матрицы $B = I - D^{-1}A$ были по модулю меньше единицы или, что то же самое, чтобы корни уравнения

$$|A - D + tD| = \begin{vmatrix} a_{11}t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}t & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}t \end{vmatrix} = 0$$

были по модулю меньше единицы.

Достаточным условием сходимости является выполнение одного из условий:

$$\max_i \sum_{j=1, j \neq i}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1;$$

$$\max_j \sum_{i=1, i \neq j}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1;$$

$$\sum_{i \neq j=1}^n \left(\frac{a_{ij}}{a_{ii}} \right)^2 < 1.$$

Метод Зейделя

Пусть система (7.77) представлена в виде

$$x = Bx + g, \quad (7.83)$$

где $B = I - A$; $g = f$.

Одношаговый циклический процесс, называемый методом Зейделя, отличается от процесса последовательных приближений тем, что при вычислении k -го приближения для i -го компонента учитываются вычисленные ранее k -е приближения для компонентов $x_1^{(k)}, \dots, x_{i-1}^{(k)}$. Таким образом, вычисления ведутся по формулам

$$x_i^{(k)} = \sum_{j=1}^{i-1} b_{ij} x_j^{(k)} + \sum_{j=i}^n b_{ij} x_j^{(k-1)} + g_i, \quad i = \overline{1, n}. \quad (7.84)$$

Представим систему (7.83) в матричном виде

$$x = (M + N)x + g,$$

где

$$M = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ b_{21} & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn-1} & 0 \end{pmatrix}; \quad N = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n-1} & b_{1n} \\ 0 & b_{22} & \dots & b_{2n-1} & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & b_{nn} \end{pmatrix}.$$

Тогда итерационное правило (7.84) можно представить в матричной форме в виде

$$x^{(k)} = Mx^{(k)} + Nx^{(k-1)} + g$$

или

$$x^{(k)} = (I - M)^{-1} N x^{(k-1)} + (I - M)^{-1} g.$$

Для сходимости метода Зейделя необходимо и достаточно, чтобы все собственные значения матрицы $(I - M)^{-1} N$ были по модулю меньше единицы или, что то же самое, чтобы корни уравнения

$$|N - (I - M)t| = \begin{vmatrix} b_{11} - t & b_{12} & \dots & b_{1n-1} & b_{1n} \\ b_{21}t & b_{22} - t & \dots & b_{2n-1} & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ b_{n1}t & b_{n2}t & \dots & b_{nn-1}t & b_{nn} - t \end{vmatrix} = 0$$

были по модулю меньше единицы. Достаточное условие сходимости состоит в том, чтобы первая или вторая норма матрицы B были меньше единицы.

7.6.6. Метод Данилевского при решении полной проблемы собственных значений. Построение канонической формы Фробениуса

Метод Данилевского решения полной проблемы собственных значений основан на том, что преобразование подобия не изменяет характеристического многочлена матрицы. Целью преобразования подобия является приведение исходной матрицы A к канонической форме Фробениуса

$$\Phi = \begin{pmatrix} 0 & 0 & \dots & 0 & p_n \\ 1 & 0 & \dots & 0 & p_{n-1} \\ 0 & 1 & \dots & 0 & p_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & p_1 \end{pmatrix},$$

по виду которой можно достаточно просто построить ее собственный многочлен

$$P(\lambda) = \lambda^n - p_1\lambda^{n-1} - \dots - p_n,$$

так как коэффициенты собственного многочлена являются элементами последнего столбца матрицы Фробениуса.

Для получения матрицы Фробениуса строится последовательность матриц $A^{(1)}, A^{(2)}, \dots, A^{(n-1)}$ по следующему правилу:

$$A^{(i)} = S_i^{-1} A^{(i-1)} S_i, \quad i = \overline{1, n-1},$$

где каждая последующая матрица $A^{(i)}$ получается из предыдущей $A^{(i-1)}$ преобразованием подобия с помощью матрицы S_i :

$$S_i = \begin{pmatrix} 0 & 0 & \dots & 0 & a_{1n}^{(i-1)} \\ 1 & 0 & \dots & 0 & a_{2n}^{(i-1)} \\ 0 & 1 & \dots & 0 & a_{3n}^{(i-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a_{nn}^{(i-1)} \end{pmatrix}.$$

Обратная к S_i матрица имеет вид

$$S_i^{-1} = \begin{pmatrix} -\frac{a_{2n}^{(i-1)}}{a_{1n}^{(i-1)}} & 1 & 0 & \dots & 0 \\ -\frac{a_{3n}^{(i-1)}}{a_{1n}^{(i-1)}} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{nn}^{(i-1)}}{a_{1n}^{(i-1)}} & 0 & 0 & \dots & 1 \\ \frac{1}{a_{1n}^{(i-1)}} & 0 & 0 & \dots & 0 \end{pmatrix},$$

где $a_{jn}^{(i-1)}$, $j = \overline{1, n}$, — элементы последнего столбца матрицы $A^{(i-1)}$ и $A^{(0)} = A$. Тогда, если все элементы $a_{1n}^{(j)} \neq 0$, $j = \overline{0, n-2}$ (регулярный случай), то матрица

$$A^{(n-1)} = \begin{pmatrix} 0 & 0 & \dots & 0 & a_{1n}^{(n-1)} \\ 1 & 0 & \dots & 0 & a_{2n}^{(n-1)} \\ 0 & 1 & \dots & 0 & a_{3n}^{(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a_{nn}^{(n-1)} \end{pmatrix} \quad (7.85)$$

будет иметь форму Фробениуса.

Если при построении матрицы $A^{(j+1)}$ элемент $a_{1n}^{(j)} = 0$, то имеем *нерегулярный случай*, так как $A^{(j)}$ имеет вид

$$A^{(j)} = \left(\begin{array}{cccc|cccc} a_{1,1}^{(j)} & a_{1,2}^{(j)} & \dots & a_{1,n-j-1}^{(j)} & 0 & \dots & 0 & a_{1,n}^{(j)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-j-1,1}^{(j)} & a_{n-j-1,2}^{(j)} & \dots & a_{n-j-1,n-j-1}^{(j)} & 0 & \dots & 0 & a_{n-j-1,n}^{(j)} \\ \hline a_{n-j,1}^{(j)} & a_{n-j,2}^{(j)} & \dots & a_{n-j,n-j-1}^{(j)} & 0 & \dots & 0 & a_{n-j,n}^{(j)} \\ a_{n-j+1,1}^{(j)} & a_{n-j+1,2}^{(j)} & \dots & a_{n-j+1,n-j-1}^{(j)} & 1 & \dots & 0 & a_{n-j+1,n}^{(j)} \\ a_{n-j+2,1}^{(j)} & a_{n-j+2,2}^{(j)} & \dots & a_{n-j+2,n-j-1}^{(j)} & 0 & \dots & 0 & a_{n-j+2,n}^{(j)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n,1}^{(j)} & a_{n,2}^{(j)} & \dots & a_{n,n-j-1}^{(j)} & 0 & \dots & 1 & a_{n,n}^{(j)} \end{array} \right)$$

В этом случае возможны две ситуации.

1. Среди элементов $a_{2n}^{(j)}, a_{3n}^{(j)}, \dots, a_{n-j,n}^{(j)}$ имеется хотя бы один, отличный от нуля, например, $a_{kn}^{(j)} \neq 0$, $k = \overline{2, n-j}$. Тогда можно прийти к регулярному случаю, если в матрице $A^{(j)}$ поменять местами первую и k -ю строки и столбцы. Такое преобразование матрицы $A^{(j)}$ будет подобным, и, кроме того, столбцы, уже имеющие нужный вид, не изменятся.

2. Все элементы $a_{kn}^{(j)} = 0$, $k = \overline{2, n-j}$, т.е. матрица $A^{(j)}$ имеет блочный вид

$$A^{(j)} = \left(\begin{array}{c|c} B & 0 \\ \hline C & \Phi_{j+1} \end{array} \right),$$

где Φ_{j+1} — квадратная матрица порядка $j+1$, имеющая каноническую форму Фробениуса; B — квадратная матрица порядка $(n-j-1)$; 0 — нулевая матрица. Тогда, на основании теоремы Лапласа, можно записать:

$$|A^{(j)} - \lambda I| = |B - \lambda I_{n-j-1}| \cdot |\Phi_{j+1} - \lambda I_{j+1}|,$$

т.е. характеристический многочлен матрицы Φ_{j+1} является делителем характеристического многочлена матрицы A . Для отыскания характеристического многочлена матрицы A нужно найти характеристический многочлен матрицы B , что можно сделать путем приведения матрицы B к форме Фробениуса.

Таким образом, построив собственный многочлен матрицы A , можно найти собственные значения $\lambda_1, \lambda_2, \dots, \lambda_n$ этой матрицы, которые являются корнями ее собственного многочлена.

Если удалось привести A к виду (7.85), то собственный вектор $x^{(i)}$ матрицы A будет равен

$$x^{(i)} = Sy^{(i)}, \quad (7.86)$$

где $S = S_1 S_2 \dots S_{n-1}$ — произведение матриц преобразования, а $y^{(i)} = (y_1^{(i)}, y_2^{(i)}, \dots, y_n^{(i)})^T$ — собственный вектор матрицы (7.85), соответствующий собственному значению λ_i , компоненты которого $y_1^{(i)}, y_2^{(i)}, \dots, y_n^{(i)}$ определяются следующим образом:

$$\left. \begin{array}{l} y_n^{(i)} = 1; \\ y_{n-1}^{(i)} = \lambda_i - p_1; \\ y_{n-2}^{(i)} = \lambda_i^2 - p_1 \lambda_i - p_2; \\ \dots \dots \dots \\ y_1^{(i)} = \lambda_i^{n-1} - p_1 \lambda_i^{n-2} - \dots - p_{n-1}, \end{array} \right\},$$

где $p_j = a_{n-j+1,n}^{(n-1)}$, $j = \overline{1, n}$.

Если не удалось привести матрицу A к виду (7.85), то для отыскания собственных векторов необходимо решить системы линейных алгебраических уравнений

$$Ax^{(i)} = \lambda_i x^{(i)}, \quad i = \overline{1, n},$$

где $x^{(i)} = (x_1^{(i)}, \dots, x_n^{(i)})^T$. Так как собственные векторы определяются с точностью до постоянного множителя, то один из компонентов вектора $x^{(i)}$ можно приравнять некоторому числу.

7.6.7. Метод Фаддеева

Для заданной матрицы A порядка n строится таблица значений:

$$\begin{array}{lll} A_1 = A; & \text{Sp}A_1 = q_1; & B_1 = A_1 - q_1 I; \\ A_2 = AB_1; & \frac{\text{Sp}A_2}{2} = q_2; & B_2 = A_2 - q_2 I; \\ \dots\dots\dots & & \\ A_{n-1} = AB_{n-2}; & \frac{\text{Sp}A_{n-1}}{n-1} = q_{n-1}; & B_{n-1} = A_{n-1} - q_{n-1} I; \\ A_n = AB_{n-1}; & \frac{\text{Sp}A_n}{n} = q_n; & B_n = A_n - q_n I, \end{array}$$

где $\text{Sp}A_i$ — след матрицы A_i ; I — единичная матрица порядка n .

Для построенных значений справедливы следующие утверждения:

1) $q_i = p_i$, где p_i — коэффициенты собственного многочлена матрицы A : $P(\lambda) = \lambda^n - p_1 \lambda^{n-1} - \dots - p_n$;

2) B_n — нулевая матрица;

3) если матрица A неособенная, то $A^{-1} = \frac{B_{n-1}}{q_n}$.

Заметим, что матрица A будет особенной, если по крайней мере одно ее собственное значение равно нулю. В этом случае коэффициент $p_n = q_n$ собственного многочлена матрицы A должен быть равным нулю.

Таким образом, методом Фаддеева можно найти обратную матрицу и собственные значения для матрицы A . Причем в сравнении с методом Гаусса алгоритм нахождения обратной матрицы методом Фаддеева реализуется значительно проще.

7.7. Решение обыкновенных дифференциальных уравнений и систем

Среди задач, с которыми приходится иметь дело при моделировании, значительную часть составляют различные задачи для обыкновенных дифференциальных уравнений. Такие задачи возникают как непосредственно при моделировании многих реальных процессов, так и в качестве промежуточных при решении более сложных математических задач. При этом, как правило, точное решение рассматриваемой задачи не удается выразить через элементарные функции. Доля задач, решаемых в явном виде в случае обыкновенных дифференциальных уравнений, ничтожно мала. Обычно приходится прибегать к помощи приближенных методов решения подобных задач. Такие методы в зависимости от того, ищется ли приближенное решение в аналитическом виде или в виде таблицы значений, разделяют на *аналитические* и *численные*. Конкретный вид метода существенно зависит от типа решаемой задачи. В случае обыкновенных дифференциальных уравнений в зависимости от того, ставятся ли дополнительные условия в одной или в нескольких точках отрезка изменения независимой переменной, задачи обычно подразделяют на *одноточечные* (*задачи с начальными условиями*, или *задачи Коши*) и *многоточечные*. Среди многоточечных задач наиболее часто в прикладных вопросах встречаются так называемые *граничные*, или *краевые задачи*, когда дополнительные условия ставятся на концах рассматриваемого отрезка. При этом нередко решение многоточечных и, в частности, граничных задач сводят к решению задач с начальными условиями.

7.7.1. Методы Эйлера и Рунге — Кутты при решении задачи Коши для дифференциального уравнения первого порядка

Задача Коши для дифференциального уравнения первого порядка

$$y'(x) = f(x, y(x)) \quad (7.87)$$

состоит в нахождении функции $y(x)$, удовлетворяющей этому уравнению при начальном условии $y(x_0) = y_0$. При численном решении этой задачи обычно задают некоторый шаг h и находят решение в точках

x_1, x_2, \dots, x_n , где $x_i = x_0 + i h$, $i = \overline{0, n}$, $h = \frac{x_n - x_0}{n}$.

Обозначим разность между точными значениями функции $y(x)$ в двух рядом стоящих точках через

$$\Delta y(x_0) = y(x_0 + h) - y(x_0)$$

и представим приближенно эту же разность в виде некоторой линейной комбинации

$$\left. \begin{aligned} \varphi_r(0) &= 0; \\ \varphi'_r(0) &= y'(x_0) - p_1 f(x_0, y_0); \\ \varphi''_r(0) &= y''(x_0). \end{aligned} \right\}$$

При этом $\phi'_1(0) = 0$ тогда и только тогда, когда $p_1 = 1$, а $\phi'_1(0) = y''(x_0)$, вообще говоря, в нуль не обращается. Тогда приближенное решение в точке x_1 согласно (7.92) будет определяться следующим образом:

$$y(x_1) = y_0 + h f(x_0, y_0).$$

Формула Эйлера для определения решения в произвольной точке будет иметь вид

$$y_{i+1} = y_i + h f(x_i, y_i), \quad i = \overline{0, n},$$

где y_i — приближенное решение в точке x_i .

Формула Рунге — Кутта

Существует несколько формул Рунге — Кутта для решения дифференциального уравнения первого порядка, которые получены при различных значениях $r > 1$. Наиболее распространенной является формула Рунге — Кутта при $r = 4$, которая имеет следующий вид:

$$y_{i+1} = y_i + \frac{1}{6} (K_1^{(i)} + 2K_2^{(i)} + 2K_3^{(i)} + K_4^{(i)}),$$

где

$$\begin{aligned} K_1^{(i)} &= h f(x_i, y_i); \\ K_2^{(i)} &= h f\left(x_i + \frac{h}{2}, y_i + \frac{K_1^{(i)}}{2}\right); \\ K_3^{(i)} &= h f\left(x_i + \frac{h}{2}, y_i + \frac{K_2^{(i)}}{2}\right); \\ K_4^{(i)} &= h f\left(x_i + h, y_i + K_3^{(i)}\right) \end{aligned}$$

и $K_j^{(i)} = K_j(x_i, y_i)$, $j = \overline{1, 4}$.

Методы Эйлера и Рунге — Кутта можно использовать и для моделирования в неравноотстоящих точках.

7.7.2. Правило Рунге

Функция $\phi_r(h)$ в (7.89) имеет следующий смысл: это разность между точным и приближенным приращениями решения дифференциального уравнения в точке x_0 . Тогда, учитывая (7.90) и (7.92), можно записать уравнение для погрешности решения в точке следующим образом:

$$\varphi_r(h) = h^{s+1} \frac{\varphi_r^{(s+1)}(0)}{(s+1)!} + O(h^{s+1}),$$

где $h^{s+1} \frac{\varphi_r^{(s+1)}(0)}{(s+1)!}$ — главный член погрешности. Если главный член погрешности можно записать в виде $K(x)h^m$, то говорят, что порядок погрешности метода равен m . На практике в качестве погрешности приближенного решения берут главный член погрешности

$$R_r(x) = K(x)h^m.$$

Число шагов, которые необходимо сделать до произвольной точки, обратно пропорционально шагу h , т.е. $N = O(1/h)$, и погрешность в произвольной точке x будет иметь смысл произведения погрешности на шаге $\varphi_r(h)$ на число шагов N , т.е. $\varphi_r(h)N$.

Таким образом, рассмотренные методы имеют *порядок погрешности на шаге*, равный $s+1$, а *порядок погрешности метода* — s . Для метода Эйлера $s = 1$, а для метода Рунге-Кутты $s = 4$. Таким образом, при моделировании методом Эйлера результаты получаются с погрешностью порядка h , а методом Рунге-Кутты — с погрешностью порядка h^4 .

Оценка погрешности методов Эйлера и Рунге-Кутты затруднительна. Грубую оценку погрешности можно получить, используя *правило Рунге*.

Пусть $y(x_i)$ — точное решение в точке x_i ; $y_i^{(h)}$ — приближенное решение в точке x_i , найденное с шагом h ; $y_i^{(2h)}$ — приближенное решение в точке x_i , найденное с шагом $2h$, и пусть погрешность приближенного решения определяется формулой

$$R_r(x_i) = K(x_i)h^s.$$

Тогда справедливы следующие равенства:

$$y(x_i) = y_i^{(h)} + K(x_i)h^s;$$

$$y(x_i) = y_i^{(2h)} + K(x_i)(2h)^s.$$

Приравнивая правые части этих равенств, получим

$$K(x_i) = \frac{y_i^{(h)} - y_i^{(2h)}}{(2h)^s - h^s} = \frac{y_i^{(h)} - y_i^{(2h)}}{h^s(2^s - 1)}$$

и

$$R_r(x_i) = \frac{y_i^{(h)} - y_i^{(2h)}}{2^s - 1}.$$

Таким образом, грубую оценку погрешности на интервале $[x_0, x_n]$ можно получить методом *двойного счета* по формулам Эйлера и Рунге-Кутты и в качестве оценки погрешностей использовать величины:

- для метода Эйлера:

$$R_1(x) \approx \max_{i=1,n} |y_i^{(h)} - y_i^{(2h)}|;$$

- для метода Рунге-Кутты:

$$R_4(x) \approx \max_{i=1,n} \left| \frac{y_i^{(h)} - y_i^{(2h)}}{15} \right|.$$

Кроме того, для контроля правильности выбора шага h рекомендуется вычислять величину

$$\theta = \max_{i=1,n} \left| \frac{K_2^{(i)} - K_3^{(i)}}{K_1^{(i)} - K_2^{(i)}} \right|.$$

7.7.3. Методы Эйлера и Рунге — Кутта при решении задачи Коши для систем дифференциальных уравнений

Ограничимся для сокращения записи системой двух уравнений.

Задача Коши для решения системы из двух уравнений:

$$\begin{cases} y_1'(x) = f_1(x, y_1(x), y_2(x)); \\ y_2'(x) = f_2(x, y_1(x), y_2(x)) \end{cases} \quad (7.93)$$

состоит в определении на интервале $[x_0, x_n]$ решения этой системы, удовлетворяющего условиям:

$$y_1(x_0) = y_1^{(0)}; \quad y_2(x_0) = y_2^{(0)}.$$

Формулы Эйлера решения задачи Коши для системы (7.93) имеют вид

$$\begin{cases} y_1^{(i+1)} = y_1^{(i)} + h f_1(x_i, y_1^{(i)}, y_2^{(i)}); \\ y_2^{(i+1)} = y_2^{(i)} + h f_2(x_i, y_1^{(i)}, y_2^{(i)}), \end{cases}$$

где $h = x_{i+1} - x_i$; $y_1^{(i)}$, $y_2^{(i)}$ — приближенные решения в точке x_i , $i = \overline{0, n}$.

Формулы Рунге — Кутта решения задачи Коши для системы (7.93) имеют вид

$$\begin{cases} y_1^{(i+1)} = y_1^{(i)} + \frac{1}{6} [K_1^{(i)} + 2K_2^{(i)} + 2K_3^{(i)} + K_4^{(i)}]; \\ y_2^{(i+1)} = y_2^{(i)} + \frac{1}{6} [L_1^{(i)} + 2L_2^{(i)} + 2L_3^{(i)} + L_4^{(i)}], \end{cases}$$

где

$$\left. \begin{aligned} K_1^{(i)} &= h f_1(x_i, y_1^{(i)}, y_2^{(i)}); \\ L_1^{(i)} &= h f_2(x_i, y_1^{(i)}, y_2^{(i)}); \\ K_2^{(i)} &= h f_1\left(x_i + \frac{h}{2}, y_1^{(i)} + \frac{K_1^{(i)}}{2}, y_2^{(i)} + \frac{L_1^{(i)}}{2}\right); \\ L_2^{(i)} &= h f_2\left(x_i + \frac{h}{2}, y_1^{(i)} + \frac{K_1^{(i)}}{2}, y_2^{(i)} + \frac{L_1^{(i)}}{2}\right); \\ K_3^{(i)} &= h f_1\left(x_i + \frac{h}{2}, y_1^{(i)} + \frac{K_2^{(i)}}{2}, y_2^{(i)} + \frac{L_2^{(i)}}{2}\right); \\ L_3^{(i)} &= h f_2\left(x_i + \frac{h}{2}, y_1^{(i)} + \frac{K_2^{(i)}}{2}, y_2^{(i)} + \frac{L_2^{(i)}}{2}\right); \\ K_4^{(i)} &= h f_1(x_i + h, y_1^{(i)} + K_3^{(i)}, y_2^{(i)} + L_3^{(i)}); \\ L_4^{(i)} &= h f_2(x_i + h, y_1^{(i)} + K_3^{(i)}, y_2^{(i)} + L_3^{(i)}). \end{aligned} \right\}$$

7.7.4. Решение задачи Коши для систем линейных дифференциальных уравнений

Задача Коши для системы линейных дифференциальных уравнений, записанной в матричной форме, имеет вид

$$y'(t) = \bar{A}(t)y(t) + \bar{b}(t); \quad y(t_0) = y^{(0)}, \quad (7.94)$$

где $y(t) = (y_1(t), y_2(t), \dots, y_n(t))^T$; $y^{(0)} = (y_1^{(0)}, y_2^{(0)}, \dots, y_n^{(0)})^T$; $\bar{A}(t)$ — матрица размерности $n \times n$; $\bar{b}(t)$ — вектор с n элементами.

Системы такого вида обычно используются в теории автоматического управления для моделирования поведения объекта в пространстве состояний.

В силу линейности системы (7.94) методы Эйлера и Рунге-Кутты значительно упрощаются и принимают вид разностных уравнений

$$y(k+1) = A(k)y(k) + b(k); \quad y(0) = y^{(0)}.$$

При этом для формулы Эйлера

$$A(k) = I + \Delta t \bar{A}(t_k); \quad b(k) = \Delta t \bar{b}(t_k), \quad (7.95)$$

а для формулы Рунге-Кутта

$$A(k) = I + \sum_{i=1}^4 \frac{\Delta t^i}{i!} \bar{A}^i(t_k); \quad b(k) = \sum_{i=1}^4 \frac{\Delta t^i}{(i-1)!} \bar{A}^{i-1}(t_k) \bar{b}(t_k). \quad (7.96)$$

В (7.95), (7.96) k соответствует моменту времени $t_k = t_0 + k \Delta t$, где $\Delta t = t_{k+1} - t_k$, $k = 0, n-1$; I — единичная матрица соответствующего порядка.

Замечание

Для того чтобы воспользоваться методами Рунге — Кутты и Эйлера для решения дифференциальных уравнений и систем, содержащих производные высших порядков относительно искомых функций путем введения новых переменных, переходят к системе обыкновенных дифференциальных уравнений.

7.7.5. Методы Адамса при решении задачи Коши для обыкновенных дифференциальных уравнений

Методы Эйлера и Рунге — Кутты являются *одношаговыми*, так как при вычислении решения в точке $x_i + h$ они используют информацию о решении только в одной точке x_i . Существуют методы, которые позволяют часть полученной информации использовать повторно на нескольких шагах вычислительного процесса. Это так называемые *многошаговые* методы. При этом иногда оказывается целесообразным привлекать также и информацию с «забеганием» вперед.

Рассмотрим решение задачи Коши для дифференциального уравнения первого порядка (7.87).

Пусть:

- 1) узлы являются равноотстоящими: $x_{i+1} = x_i + h$, $i = \overline{0, n-1}$;
- 2) для нахождения решения в точке x_{i+1} требуется знание приближенного решения в $(k+1)$ предыдущих точках $x_i, x_{i-1}, \dots, x_{i-k}$.

Точки x_0, x_1, \dots, x_k образуют *начальный отрезок*. Решение на начальном отрезке может быть найдено, например, методами Рунге — Кутты или Эйлера.

Пусть известно решение в точках x_0, x_1, \dots, x_k и требуется найти решение в точке x_{k+1} . Проинтегрируем исходное уравнение $y' = f(x, y)$ на отрезке $[x_k, x_{k+1}]$:

$$\int_{x_k}^{x_{k+1}} y' dx = \int_{x_k}^{x_{k+1}} f(x, y) dx.$$

Тогда

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(x, y) dx. \quad (7.97)$$

Для подынтегральной функции $f(x, y)$ построим интерполяционный многочлен $P_n(x)$ и подставим его в (7.97). Так как значение интерполяционного многочлена вычисляется на отрезке $[x_k, x_{k+1}]$, то в качестве такого многочлена выберем формулу Ньютона для интерполирования назад, и, в зависимости от того, какая точка привлекается первой при построении интерполяционного многочлена, получаются различные формулы для решения дифференциального уравнения.

Если вместо функции $f(x, y)$ в (7.97) подставить $P_n(x)$, точные значения $y(x_i)$ заменить на приближенные y_i и проинтегрировать, то получим конкретные формулы Адамса. Наиболее распространенными являются формулы Адамса для $n = 4$.

Экстраполяционная формула Адамса

Если для построения интерполяционного многочлена $P_4(x)$ используются значения функции $f(x, y)$ в узлах $x_k, x_{k-1}, \dots, x_{k-4}$, то

$$P_4(x) = f(x_k, y(x_k)) + \frac{t}{1!} \Delta f(x_{k-1}, y(x_{k-1})) + \frac{t(t+1)}{2!} \Delta^2 f(x_{k-2}, y(x_{k-2})) + \dots + \frac{t(t+1)(t+2)(t+3)}{4!} \Delta^4 f(x_{k-4}, y(x_{k-4})).$$

Тогда для вычисления решения дифференциального уравнения (7.87) в точке x_{k+1} получим экстраполяционную формулу Адамса, которая имеет вид

$$y_{k+1} = y_k + \frac{h}{24} (55f_k - 59f_{k-1} + 37f_{k-2} - f_{k-3}).$$

Здесь $f_j = f(x_j, y_j)$, $j = k, k-3$.

Интерполяционная формула Адамса

Если для построения интерполяционного многочлена $P_4(x)$ используются значения функции $f(x, y)$ в узлах $x_{k+1}, x_k, \dots, x_{k-3}$, то

$$P_4(x) = f(x_{k+1}, y(x_{k+1})) + \frac{t}{1!} \Delta f(x_k, y(x_k)) + \frac{t(t+1)}{2!} \Delta^2 f(x_{k-1}, y(x_{k-1})) + \dots + \frac{t(t+1)(t+2)(t+3)}{4!} \Delta^4 f(x_{k-3}, y(x_{k-3}))$$

и для вычисления решения дифференциального уравнения (7.87) в точке x_{k+1} получим интерполяционную формулу Адамса, которая имеет вид

$$y_{k+1} = y_k + \frac{h}{24} (9f_{k+1} + 19f_k - 5f_{k-1} + f_{k-2}). \quad (7.98)$$

Заметим, что интерполяционная формула (7.98) представляет собой уравнение относительно y_{k+1} , так как $f_{k+1} = f(x_{k+1}, y_{k+1})$. Одним из методов решения этого уравнения является метод последовательных приближений:

$$y_{k+1}^{(i+1)} = y_k^{(i)} + \frac{h}{24} (9f_{k+1}^{(i)} + 19f_k - 5f_{k-1} + f_{k-2}),$$

где $f_{k+1}^{(i)} = f(x_{k+1}, y_{k+1}^{(i)})$; начальное значение $y_{k+1}^{(0)}$ задается, например, с помощью экстраполяционной формулы Адамса, $i = 0, 1, 2, \dots$.

Методы Адамса дают результат с погрешностью порядка h^4 . Практическая оценка погрешности приближенного решения может быть получена с помощью правила Рунге.

7.7.6. Метод сеток решения линейных краевых задач

Наряду с задачами Коши для обыкновенных дифференциальных уравнений достаточно часто рассматриваются также граничные задачи. В этих задачах дополнительные условия, которым должно удовлетворять решение, задаются в виде уравнений, содержащих комбинации значений решения и его производных, взятых в нескольких точках рассматриваемого отрезка $[a, b]$. Простейшими среди граничных задач являются *краевые задачи*, когда дополнительные условия задаются только на концах отрезка $[a, b]$. Если дифференциальное уравнение и уравнения, определяющие дополнительные условия, являются линейными, то граничная задача называется *линейной*.

Пусть на отрезке $[a, b]$ задана линейная краевая задача для дифференциального уравнения

$$L(y) = y'' + p(x)y' + q(x)y = f(x) \quad (7.99)$$

с условиями

$$\left. \begin{aligned} I_a(y) &= \alpha_0 y(a) + \alpha_1 y(b) = A; \\ I_b(y) &= \beta_0 y(a) + \beta_1 y(b) = B, \end{aligned} \right\} \quad (7.100)$$

где $L(y)$, $I_a(y)$, $I_b(y)$ — линейные операторы.

Метод сеток для решения краевой задачи (7.99), (7.100) состоит в следующем.

1. На отрезке $[a, b]$ выбирается некоторая система точек x_k . Совокупность этих точек называется *сеткой*, а точки x_k называются *узлами сетки*. Достаточно часто используют *равномерную сетку узлов*, когда $x_k = a + kh$; $h = \frac{b-a}{n}$, $k = \overline{0, n}$, n — заданное целое число.

2. Краевая задача (7.99), (7.100) на множестве узлов, принадлежащих сетке, заменяется некоторой *сеточной задачей*, то есть некоторыми соотношениями между приближенными значениями решения в узлах сетки.

3. Сеточная задача решается каким-либо численным методом. Это позволяет найти приближенные значения решения краевой задачи в узлах сетки.

Метод сеток является одним из наиболее универсальных методов решения как обыкновенных дифференциальных уравнений, так и уравнений в частных производных. Рассмотрим этот метод в применении к решению линейных краевых задач для дифференциальных уравнений второго порядка.

Пусть линейная краевая задача задана в виде

$$A(x)y'' + B(x)y' + C(x)y = f(x); \quad (7.101)$$

$$y(a) = \alpha; \quad y(b) = \beta. \quad (7.102)$$

На интервале $[a, b]$ зададим равномерную сетку узлов $x_i = a + i h$, $i = \overline{0, n}$, и обозначим y_i , значения приближенного решения в точках x_i . В дифференциальном уравнении (7.101) производные заменим разностными отношениями

$$y'(x_i) \approx \begin{cases} \frac{y_{i+1} - y_i}{h}; \\ \frac{y_i - y_{i-1}}{h}; \\ \frac{y_{i+1} - y_{i-1}}{2h}, \end{cases} \quad (7.103)$$

$$y''(x_i) \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}. \quad (7.104)$$

Тогда, записав уравнение (7.101) для точки x_i и заменив производные в этой точке согласно (7.103), (7.104), получим систему линейных алгебраических уравнений вида

$$\left. \begin{aligned} y_0 &= \alpha; \\ a_i y_{i-1} + b_i y_i + c_i y_{i+1} &= g_i, \quad i = \overline{1, n-1}; \\ y_n &= \beta, \end{aligned} \right\} \quad (7.105)$$

где значения a_i , b_i , c_i , g_i получены приведением подобных членов после замены производных. Система (7.105) является системой с трехдиагональной матрицей. Для решения таких систем используется *метод прогонки*. Согласно этому методу решение системы записывается в виде

$$y_i = L_i y_{i+1} + K_i,$$

где L_i и K_i — неизвестные коэффициенты, которые называются *прогночными*. Выражения для прогночных коэффициентов получаются следующим образом.

Так как при $i = 0$ имеем $y_0 = \alpha$ и

$$y_0 = L_0 y_1 + K_0,$$

то

$$K_0 = \alpha, \quad L_0 = 0.$$

При $i = 1$ имеем

$$y_1 = L_1 y_2 + K_1,$$

а из уравнения системы (7.105)

$$a_1 y_0 + b_1 y_1 + c_1 y_2 = g_1$$

или

$$y_1 = \frac{-c_1}{b_1} y_2 + \frac{g_1 - a_1 K_0}{b_1}.$$

Записав аналогично предыдущему выражения для $i=2$, получим рекуррентные соотношения для прогоночных коэффициентов:

$$L_i = -\frac{c_i}{a_i L_{i-1} + b_i}; \quad K_i = \frac{g_i - a_i K_{i-1}}{a_i L_{i-1} + b_i}, \quad i = \overline{1, n-1}.$$

Решение системы (7.105) находится в обратном порядке следующим образом:

$$\left. \begin{aligned} y_n &= \beta; \\ y_{i-1} &= L_{i-1} y_i + K_{i-1}, \quad i = n, n-1, \dots, 2; \\ y_0 &= \alpha. \end{aligned} \right\}$$

Для того чтобы получить решение задачи (7.101), (7.102) в аналитической форме, необходимо аппроксимировать значения y_0, y_1, \dots, y_n некоторой функцией.

7.7.7. Метод коллокации решения краевых задач

При решении задач физики и механики искомое решение иногда более выгодно находить в аналитическом виде. Часто в таких случаях не преследуют цель получить решение с большой точностью, а ограничиваются отысканием функции, точно удовлетворяющей граничным условиям и некоторым соотношениям, связанным с заданным дифференциальным уравнением. Эти соотношения строятся таким образом, чтобы функция, удовлетворяющая им, удовлетворяла приближенно с необходимой точностью также и заданному дифференциальному уравнению. При этом эти соотношения в различных методах строятся по-разному.

Решение краевой задачи

$$\left. \begin{aligned} L(y) &= y'' + p(x)y' + q(x)y = f(x); \\ I_a(y) &= A; \\ I_b(y) &= B \end{aligned} \right\} \quad (7.106)$$

ищется в виде

$$y_n(x) = \varphi_0(x) + \sum_{i=1}^n c_i \varphi_i(x), \quad (7.107)$$

где конечная система базисных функций $\{\varphi_i(x)\}$, $i = \overline{0, n}$, заданных на $[a, b]$, выбирается так, чтобы функция $\varphi_0(x)$ удовлетворяла неоднородным краевым условиям

$$I_a(\varphi_0) = A; \quad I_b(\varphi_0) = B, \quad (7.108)$$

а функции $\varphi_i(x)$ удовлетворяли бы однородным краевым условиям

$$I_a(\varphi_i) = I_b(\varphi_i) = 0, \quad i = \overline{1, n}. \quad (7.109)$$

Подставляя (7.107) в дифференциальное уравнение, получим невязку

$$R(x, c_1, \dots, c_n) = L[\varphi_0] + \sum_{i=1}^n c_i L[\varphi_i] - f(x).$$

Потребуем, чтобы невязка $R(x, c_1, \dots, c_n)$ обращалась в нуль на некоторой системе точек x_1, \dots, x_n отрезка $[a, b]$, называемых *точками коллокации*, причем число таких точек должно равняться числу коэффициентов c_i в выражении (7.107). Тогда, для определения значений c_1, \dots, c_n , получим систему уравнений

$$\left. \begin{aligned} R(x_1, c_1, \dots, c_n) &= 0; \\ R(x_2, c_1, \dots, c_n) &= 0; \\ &\dots \\ R(x_n, c_1, \dots, c_n) &= 0, \end{aligned} \right\} \quad (7.110)$$

определитель которой имеет вид

$$D = \begin{vmatrix} L(\varphi_1(x_1)) & L(\varphi_2(x_1)) & \dots & L(\varphi_n(x_1)) \\ & & \dots & \\ L(\varphi_1(x_n)) & L(\varphi_2(x_n)) & \dots & L(\varphi_n(x_n)) \end{vmatrix}.$$

Система (7.110) имеет единственное решение, если $D \neq 0$, в связи с этим функции $L(\varphi_i(x))$, $i = \overline{1, n}$, должны составлять *систему Чебышева*.

Для достижения хорошего качества приближения $y_n(x)$ к точному решению $y(x)$ функции $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$ должны удовлетворять следующим условиям:

- 1) быть линейно независимыми;
- 2) при непрерывных функциях $p(x), q(x), f(x)$ непрерывными на $[a, b]$ должны быть и функции $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$ вместе со своими производными до второго порядка включительно;
- 3) для того чтобы $L[y_n(x)]$ могли сколь угодно точно приближать значения $L[y(x)]$, система функций $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$ должна быть полной в классе дважды непрерывно дифференцируемых функций.

Существуют и другие методы решения граничной задачи (7.106) в виде (7.107), например метод Галеркина, наименьших квадратов. Однако метод коллокации приводит к более простым вычислениям.

Замечание

Для решения краевой задачи (7.106) в качестве функций $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$ можно выбрать, например, следующие функции:

1) $\varphi_0(x) = e + d \cdot x$, где e и d задаются таким образом, чтобы выполнялись неоднородные краевые условия (7.108);

2) $\varphi_1(x) = (x - a)(x - b)$, $\varphi_2(x) = (x - a)^2(x - b)$, ..., $\varphi_n(x) = (x - a)^n \times (x - b)$, которые будут удовлетворять однородным краевым условиям (7.109).

8. МОДЕЛИРОВАНИЕ СИСТЕМ УПРАВЛЕНИЯ

В современной теории автоматического управления доминирует концепция совмещенного синтеза следящих систем. Наиболее перспективными методами решения таких задач являются методы, основанные на оптимизации квадратичных критериев, а использование прогнозирующих моделей и скользящего интервала времени позволяет существенно расширить возможности синтеза адаптивных систем.

Рассмотрим методы синтеза систем управления при решении задач слежения для многомерных стохастических моделей объектов, функционирующих в условиях неполной информации о состоянии объекта и его параметров. При этом проектирование систем управления будем осуществлять путем постепенного добавления и усложнения используемых методов и алгоритмов: от оптимального управления для детерминированных моделей до адаптивных следящих систем для стохастических моделей при неполном измерении с ошибками.

В связи со сложностью рассматриваемых задач единственным способом их решения является имитационное моделирование.

8.1. Описание систем в пространстве состояний

Описание систем во временной области лежит в основе современной теории управления. *Временная область* — это область, в которой поведение системы рассматривается как функция переменной времени t . Анализ и синтез систем управления во временной области основан на понятии состояния системы. *Состояние системы* — это совокупность таких переменных, знание которых наряду с входными функциями и уравнениями, описывающими динамику системы, позволяет определить ее будущее состояние и выходную переменную. Для динамической системы ее состояние описывается набором переменных состояния $x_1(t), x_2(t), \dots, x_n(t)$. Это такие переменные, которые определяют будущее поведение системы, если известно ее текущее состояние и все внешние воздействия. Например, для системы, структурная схема которой изображена на рис. 8.1, переменные $x_1(t), x_2(t), \dots, x_n(t)$ имеют следующий смысл: если в момент времени t_0 известны начальные значения $x_1(t_0), x_2(t_0), \dots, x_n(t_0)$ и входные сигналы $U_1(t), \dots, U_m(t)$ для $t \geq t_0$, то этой информации достаточно, чтобы определить будущие значения всех переменных состояния и выходных переменных. Таким образом, *переменные состояния описывают поведение систем в будущем, если известны текущее состояние, внешние воздействия и уравнения динамики системы.*



Рис. 8.1

П р и м е р

Рассмотрим механическую систему «масса — пружина» с затуханием, изображенную на рис. 8.2.

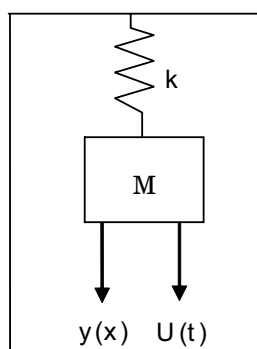


Рис. 8.2

Дифференциальное уравнение, описывающее поведение системы, обычно записывается в виде

$$M \frac{d^2 y(t)}{dt^2} + b \frac{dy(t)}{dt} + ky(t) = U(t). \quad (8.1)$$

Число переменных состояния, выбираемых для описания системы, должно быть по возможности минимальным, чтобы среди них не было излишних. Для данной системы вполне достаточно иметь две переменные состояния — положение и скорость движения массы:

$$\left. \begin{aligned} x_1(t) &= y(t); \\ x_2(t) &= \frac{dy(t)}{dt}. \end{aligned} \right\}$$

С учетом переменных состояния уравнение (8.1) примет вид

$$M \frac{dx_2(t)}{dt} + bx_2(t) + kx_1(t) = U(t).$$

Это уравнение можно представить в виде эквивалентной системы двух дифференциальных уравнений первого порядка

$$\left. \begin{aligned} \frac{dx_1(t)}{dt} &= x_2(t); \\ \frac{dx_2(t)}{dt} &= -\frac{k}{M}x_1(t) - \frac{b}{M}x_2(t) + \frac{1}{M}U(t). \end{aligned} \right\} \quad (8.2)$$

Уравнения (8.2) описывают поведение системы в терминах скорости изменения каждой переменной состояния.

Переменные состояния, описывающие систему, не являются единственными, и всегда можно выбрать альтернативную комбинацию таких переменных. Например, для приведенной выше системы «масса — пружина» в качестве переменных состояния можно выбрать любые две линейно-независимые комбинации $x_1(t)$ и $x_2(t)$. На практике в качестве переменных состояния часто выбирают такие физические переменные, которые легко могут быть измерены.

Переменные состояния характеризуют *динамику* системы. Инженеры в первую очередь интересуются физическими системами, в которых переменными состояниями являются напряжения, токи, скорости, перемещения, давления, температуры и другие физические величины. Однако понятие состояния применимо к анализу не только физических, но также биологических, социальных и экономических систем. Для этих систем понятие состояния не ограничивается рамками представлений об энергии, а подходит к переменным состояниям в более широком смысле, трактуя их как переменные любой природы, описывающие будущее поведение системы.

В общем случае систему обыкновенных дифференциальных уравнений, которая описывает систему управления в переменных состояниях, можно представить в виде

$$\dot{x}(t) = f(t, x(t), u(t)); \quad x(t_0) = x_0, \quad (8.3)$$

где $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ — вектор состояния;

$u(t) = (u_1(t), u_2(t), \dots, u_m(t))^T$ — вектор управляющих воздействий;

$f(t, x(t), u(t))$ — вектор-функция указанных аргументов, задающая динамику системы;

$x_0 = (x_{01}, \dots, x_{0n})^T$ — вектор начального состояния.

Если $x(t)$ и (или) $u(t)$ входят в $f(\cdot)$ нелинейно, то система (8.3) называется *нелинейной*, если систему (8.3) можно представить в виде

$$\dot{x}(t) = \bar{A}(t)x(t) + \bar{B}(t)u(t), \quad x(t_0) = x_0, \quad (8.4)$$

где $\bar{A}(t)$ — матрица динамических свойств модели объекта, размерности $n \times n$; $\bar{B}(t)$ — матрица влияния управляющих воздействий, раз-

мерности $n \times m$, то такая система называется *линейной*. Если элементы матрицы $\bar{A}(t)$ и (или) $\bar{B}(t)$ зависят от времени, то система (8.5) является *нестационарной*, если ни один из элементов матриц $\bar{A}(t)$ и $\bar{B}(t)$ не зависит от времени, то система (8.4) является *стационарной*. Если система управления является нелинейной, то достаточно часто ее *линеаризуют* и представляют в виде (8.4).

Заметим, что систему управления в переменных состояния можно представить и в дискретной форме

$$x(k+1) = f(k, x(k), u(k)); \quad x(0) = x_0, \quad (8.5)$$

где $x(k) = (x_1(k), \dots, x_n(k))^T$ — вектор состояния на k -м такте;

$u(k) = (u_1(k), \dots, u_m(k))^T$ — вектор управляющих воздействий;

$f(k, x(k), u(k))$ — вектор-функция дискретных аргументов, задающая динамику системы на k -м такте;

$x(0) = (x_{01}, \dots, x_{0n})^T$ — вектор начального состояния.

Аналогично (8.3) система (8.5) может быть нелинейной и линейной, представимой в виде

$$x(k+1) = A(k)x(k) + B(k)u(k); \quad x(0) = x_0. \quad (8.6)$$

Кроме того, системы (8.5), (8.6) могут быть как стационарными, так и нестационарными.

В дальнейшем будем считать, что математическая модель, описывающая поведение управляемого объекта, задана в виде (8.4) или (8.6). Использование линейных систем для описания моделей объектов обусловлено тем, что идеи и методы линейной теории автоматического управления с соответствующими оговорками широко используются и для других моделей объектов управления. Кроме того, математический аппарат матричной алгебры достаточно легко реализуется на ЭВМ.

8.2. Аналитическое конструирование оптимальных регуляторов

Синтез систем управления — одна из важнейших технических задач. Целью синтеза является создание системы, которая удовлетворяла бы требуемым показателям качества. Качество системы управления можно охарактеризовать, например, интегральными оценками, и тогда синтез системы управления должен быть основан на минимизации оценки качества. Системы управления, в которых обеспечивается минимум соответствующей оценки качества, называются *оптимальными системами управления*.

Свойства оптимального управления в значительной мере определяются выбранным критерием оптимизации. Для методов синтеза управления объектами, динамические свойства которых заданы в

пространстве состояний, наиболее удобными являются интегральные показатели качества. При этом конкретная форма показателя качества тесно связана с методом синтеза. В общем случае такие показатели описываются функционалами вида

$$J = V_z(x(T), T) + \int_{t_0}^T Q_1(x(t), t) dt + \int_{t_0}^T Q_2(u(t), t) dt, \quad (8.7)$$

определенными на всех возможных траекториях $x(t)$ в рассматриваемом пространстве состояний для всех $t \in [t_0, T]$, где $Q_1(\cdot)$, $Q_2(\cdot)$ — заданные функции указанных аргументов, удовлетворяющие некоторым условиям. Первое слагаемое в (8.7) является *терминальным* членом функционала. Оно определяет вклад в функционал конечного состояния $x(T)$. При этом рассматривается задача со свободным правым концом, когда отсутствует требование прохождения вектора состояния через заданную точку в момент T . Второе слагаемое в (8.7) представляет собой интегральную оценку качества переходного процесса объекта управления на интервале $[t_0, T]$. Третье слагаемое — интегральная оценка «расходов» сигнала управления на интервале $[t_0, T]$. Если функции V_z , Q_1 , Q_2 являются положительно определенными, а их единственные нулевые значения соответствуют состоянию объекта, требуемому по условию решаемой задачи, то оптимальность синтезируемого управления можно понимать в смысле достижения минимума функционала (8.7).

Положительная определенность подынтегральных функций в (8.7) допускает использование для разработки конкретных критериев оптимизации широкого класса функций Q_1 и Q_2 . В то же время практика решения задач управления динамическими объектами показывает, что в большинстве случаев требования к качеству переходных процессов могут быть заданы квадратичной формой функции $Q_1(x(t), t)$:

$$Q_1(x(t), t) = \frac{1}{2} x^T(t) C x(t),$$

где C — неотрицательно определенная весовая матрица.

Среди методов автоматического управления объектами, математические модели которых в пространстве состояний описываются линейными нестационарными дифференциальными уравнениями, выделяется *метод аналитического конструирования оптимальных регуляторов* (АКОР), так как линейно-квадратичная задача является единственной, при которой решение получается в общем виде в форме обратной связи.

Термин «аналитическое конструирование регуляторов» был введен А.М. Летовым и означает синтез оптимальных систем управления, основанный на минимизации функционала, т.е. на решении вариационной задачи. Практически одновременно эта же задача была решена Р. Калманом. Метод аналитического конструирования позволяет в ана-

литическом виде определять структуру и параметры системы управления, достаточно полно учитывая при этом технические требования к качеству функционирования объекта.

Полнота общего решения обуславливается условиями, налагаемыми на динамические свойства управляемого объекта и структуру оптимизируемого критерия, отражающего предъявляемые к управлению требования. Эти условия состоят в следующем:

- 1) уравнение, описывающее движение объекта, должно быть линейным относительно вектора управляющих воздействий;
- 2) область возможных значений управлений должна быть незамкнутой;
- 3) все возможные переходные функции объекта управления должны быть непрерывно-дифференцируемыми в рассматриваемом пространстве состояний;
- 4) минимизируемый функционал должен быть квадратичным относительно вектора управления:

$$Q_2(u(t), t) = \frac{1}{2} u^T(t) D u(t),$$

где D — некоторая положительно определенная матрица, характеризующая «свободу» выбора управления.

8.2.1. Оптимальное управление при минимизации классического квадратичного функционала

Согласно методу аналитического конструирования Летова-Калмана оптимальное управление объектом (8.4), минимизирующее квадратичный функционал

$$J = \frac{1}{2} x^T(T) E x(T) + \frac{1}{2} \int_{t_0}^T [x^T(t) C x(t) + u^T(t) D u(t)] dt, \quad (8.8)$$

имеет вид

$$u(t) = -D^{-1} \bar{B}^T(t) \left(\frac{\partial V}{\partial x} \right)^T,$$

где функция $V = V(x(t), t)$ является решением нелинейного уравнения в частных производных

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \bar{A}(\tau) x(\tau) - \frac{1}{2} \frac{\partial V}{\partial x} \bar{B}(\tau) D^{-1} \bar{B}^T(\tau) \left(\frac{\partial V}{\partial x} \right)^T = \\ = -\frac{1}{2} x^T(\tau) C x(\tau), \end{aligned} \quad (8.9)$$

при граничном условии

$$V(T) = \frac{1}{2} x^T(T) E x(T). \quad (8.10)$$

В (8.8) E — заданная положительно определенная матрица.

Решение задачи (8.9), (8.10) ищется в виде квадратичной формы

$$V = \frac{1}{2} x^T(t) S(t) x(t),$$

где $S(t)$ — симметрическая матрица. Тогда

$$u(t) = -D^{-1} \bar{B}^T(t) S(t) x(t),$$

а матрица $S(t)$ является решением нелинейного дифференциального уравнения Риккати с переменными коэффициентами

$$\begin{aligned} \dot{S}(t) = & -S(t) \bar{A}(t) - \bar{A}^T(t) S(t) + \\ & + S(t) \bar{B}(t) D^{-1} \bar{B}^T(t) S(t) - C \end{aligned} \quad (8.11)$$

при граничном условии

$$S(T) = E. \quad (8.12)$$

Таким образом, при синтезе оптимального управления методом аналитического конструирования Летова-Калмана необходимо для каждого момента $t \in [t_0, T]$ формирования управляющих воздействий решать в обратном времени уравнение (8.11) при условии (8.12). Теория решения матричного уравнения Риккати достаточно хорошо разработана. Однако при практическом применении необходимость решения в обратном времени матричного дифференциального уравнения Риккати делает затруднительным использование этого метода синтеза в процессе функционирования объекта из-за больших затрат времени на получение результата. Кроме того, для нестационарных систем редко удастся указать обоснованные требования к состоянию объекта в фиксированный момент T , т.е. обоснованно назначить квадратичную форму $0,5 x^T(T) E x(T)$. Поэтому чаще всего рассматривают функционал (8.8) в виде

$$J = \frac{1}{2} \int_{t_0}^T [x^T(t) C x(t) + u^T(t) D u(t)] dt. \quad (8.13)$$

При этом граничное условие для решения уравнения Риккати (8.12) принимает вид

$$S(T) = 0.$$

В случае стационарной модели (8.4) ($\bar{A}(t) = \bar{A}$, $\bar{B}(t) = \bar{B}$) оптимальное управление, обеспечивающее минимум функционала

$$J = \frac{1}{2} \int_{t_0}^{\infty} [x^T(t) C x(t) + u^T(t) D u(t)] dt,$$

имеет вид

$$u(t) = -D^{-1}\bar{B}(t)Sx(t),$$

где S — положительно определенная матрица, удовлетворяющая матричному алгебраическому уравнению Риккати

$$\bar{A}^T S + S\bar{A} - S\bar{B}D^{-1}\bar{B}^T S + C = 0.$$

Решение алгебраического уравнения Риккати заменяется определением с заданной точностью установившегося решения дифференциального уравнения Риккати

$$\dot{S} = \bar{A}^T S + S\bar{A} - S\bar{B}D^{-1}\bar{B}^T S + C$$

при нулевом начальном условии.

Синтез управляющих воздействий, который осуществляется в реальном времени в процессе функционирования объекта, называется *совмещенным синтезом*. В этом случае большое значение имеет временная задержка, вызванная затратами времени на формирование управляющего сигнала. Время формирования управления при АКОРе Летова-Калмана зависит от двух факторов: величины интервала оптимизации $[t_0, T]$ и времени, которое затрачивается на решение уравнения Риккати. Целенаправленное изменение каждой из этих величин приводит к уменьшению вычислительной задержки при формировании управляющего сигнала.

8.2.2. Оптимальное управление при минимизации функционала обобщенной работы

Термин «критерий обобщенной работы» был введен А.А. Красовским. Минимизируемый функционал обобщенной работы задается в виде

$$J = \frac{1}{2}x^T(T)Ex(T) + \frac{1}{2} \int_{t_0}^T [x^T(t)Cx(t) + u^T(t)Du(t) + u_{оп}^T(t)Du_{оп}(t)] dt, \quad (8.14)$$

где $u_{оп}$ — оптимальное управление, доставляющее минимум функционалу (8.14), которое имеет вид

$$u(t) = u_{оп}(t) = -D^{-1}\bar{B}^T(t) \left(\frac{\partial V}{\partial x} \right)^T, \quad (8.15)$$

а $V = V(x, t)$ является решением линейного уравнения в частных производных

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} \bar{A}(\tau)x(\tau) = -\frac{1}{2}x^T(\tau)Cx(\tau) \quad (8.16)$$

при граничном условии

$$V(T) = \frac{1}{2} x^T(T) E x(T). \quad (8.17)$$

Если искать решение задачи (8.16)–(8.17) в виде квадратичной формы

$$V = \frac{1}{2} x^T(t) L(t) x(t),$$

то

$$u(t) = u_{оп}(t) = -D^{-1} \bar{B}^T(t) L(t) x(t),$$

где $L(t)$ является решением линейного матричного дифференциального уравнения Ляпунова

$$\left. \begin{aligned} \dot{L}(\tau) &= -L(\tau) \bar{A}(\tau) - \bar{A}^T(\tau) L(\tau) - C; \\ L(T) &= E. \end{aligned} \right\}$$

Вычислительные затраты при решении уравнения Ляпунова в силу его линейности значительно сокращаются по сравнению с решением уравнения Риккати. Недостатком этого метода является то, что управление (8.15) является оптимальным при «свободном» движении объекта, т.е. при $u \equiv 0$. При этом понятно, что поведение объекта при нулевом управлении (нулевом положении органов управления), может очень сильно отличаться от управляемого поведения на интервале оптимизации и приводить систему в область пространства состояний, далекую от реально достигаемой.

Лучшие результаты можно получить, если «свободным» считать поведение объекта при фиксированном положении органов управления. При этом задача управления формулируется как *управление скоростью перемещения регулирующих органов*, а модель объекта представляется в виде

$$\left. \begin{aligned} \dot{x}(t) &= \bar{A}(t)x(t) + B(t)u(t), \quad x(t_0) = x_0; \\ \dot{u}(t) &= v(t), \quad u(t_0) = u_0, \end{aligned} \right\} \quad (8.18)$$

где $u(t)$ — вектор положения органов управления, а $v(t)$ — вектор управления, характеризующий скорость перемещения органов управления.

Заметим, что (8.18) можно записать в традиционной форме, как задачу управления положением органов управления, если ввести расширенный вектор состояния $\tilde{x} = (x, u)^T$. Тогда (8.18) запишется в виде

$$\dot{\tilde{x}}(t) = \tilde{A}(t)\tilde{x}(t) + \tilde{B}(t)v(t), \quad \tilde{x}(t_0) = \tilde{x}_0, \quad (8.19)$$

где $\tilde{A}(t)$ и $\tilde{B}(t)$ — блочные матрицы вида

$$\tilde{A}(t) = \begin{pmatrix} \bar{A}(t) & \bar{B}(t) \\ 0 & 0 \end{pmatrix}; \quad \tilde{B}(t) = \begin{pmatrix} 0 \\ I_m \end{pmatrix}. \quad (8.20)$$

Здесь 0 — нулевые матрицы соответствующих размерностей; I_m — единичная матрица порядка m .

Кроме того, учитывая сложность определения терминального слагаемого, функционал обобщенной работы для системы (8.19) запишем в виде

$$J = \frac{1}{2} \int_{t_0}^T \left[\tilde{x}^T(t) \tilde{C} \tilde{x}(t) + v^T(t) D_1 v(t) + v_{оп}^T(t) D_1 v_{оп}(t) \right] dt, \quad (8.21)$$

где \tilde{C} — весовая матрица порядка $(n+m)$, которая имеет блочный вид

$$\tilde{C} = \begin{pmatrix} C & I & 0 \\ 0 & I & D_2 \end{pmatrix}, \quad (8.22)$$

а управление $v_{оп}(t)$ определяется следующим образом:

$$v_{оп}(t) = -D_1^{-1} \tilde{B}^T \left(\frac{\partial V}{\partial \tilde{x}} \right)^T = D_1^{-1} \left(\frac{\partial V}{\partial u} \right)^T. \quad (8.23)$$

Здесь C, D_2 — неотрицательно определенные весовые матрицы; D_1 — положительно определенная весовая матрица.

Минимизируемый функционал обобщенной работы для модели (8.18) запишется в виде

$$J = \frac{1}{2} \int_{t_0}^T \left[x^T(t) C x(t) + u^T(t) D_2 u(t) + v^T(t) D_1 v(t) + v_{оп}^T(t) D_1 v_{оп}(t) \right] dt, \quad (8.24)$$

а модель, описывающая «свободное» движение объекта на интервале оптимизации, которую будем называть *прогнозирующей моделью*, есть

$$\left. \begin{aligned} \dot{x}_M(\tau) &= \bar{A}(\tau) x_M(\tau) + \bar{B}(\tau) u_M(\tau), \quad x_M(t_0) = x_0; \\ \dot{u}_M(\tau) &= 0, \quad u_M(t_0) = u_0, \end{aligned} \right\} \quad (8.25)$$

где M указывает на принадлежность прогнозирующей модели.

Тогда

$$v_{оп}(t) = -D_1^{-1} \left(\frac{\partial V}{\partial u} \right)^T,$$

где функция $V = V(t, x(t), u(t))$ удовлетворяет уравнению

$$\dot{V} = -\frac{1}{2} \left[x_M^T(\tau) C x_M(\tau) + u_M^T(\tau) D_2 u_M(\tau) \right], \quad V(T) = 0. \quad (8.26)$$

Заметим, что (8.26) справедливо на траектории движения модели (8.25).

Распишем в (8.26) полную производную функции V :

$$\frac{\partial V}{\partial \tau} + \frac{\partial V}{\partial x_M} \dot{x}_M(\tau) = -\frac{1}{2} \left[x_M^T(\tau) C x_M(\tau) + u_M^T(\tau) D_2 u_M(\tau) \right] \quad (8.27)$$

и обозначим

$$W_1 = \left(\frac{\partial V}{\partial x_M} \right)^T; \quad W_2 = \left(\frac{\partial V}{\partial u_M} \right)^T, \quad (8.28)$$

n -мерный и m -мерный векторы-столбцы.

Тогда на траектории движения модели (8.25) полные производные для W_1 и W_2 будут равны

$$\left. \begin{aligned} \dot{W}_1 &= \frac{\partial W_1}{\partial \tau} + \frac{\partial W_1}{\partial x_M} \dot{x}_M; \\ \dot{W}_2 &= \frac{\partial W_2}{\partial \tau} + \frac{\partial W_2}{\partial x_M} \dot{x}_M. \end{aligned} \right\}$$

Если продифференцировать (8.27) последовательно по x_M , u_M и подставить уравнения для модели (8.25), то, с учетом обозначений (8.28), получим систему обыкновенных дифференциальных уравнений

$$\left. \begin{aligned} \dot{W}_1(\tau) &= -\bar{A}^T(\tau)W_1(\tau) - C x_M(\tau), \quad W_1(T) = 0; \\ \dot{W}_2 &= -\bar{B}^T(\tau)W_1(\tau) - D_2 u_M(\tau), \quad W_2(T) = 0. \end{aligned} \right\} \quad (8.29)$$

Решая систему (8.29) в обратном времени, найдем управление в момент t , которое будет равно:

$$v_{оп}(t) = -D_1^{-1}W_2(t).$$

8.2.3. Оптимальное управление при минимизации локального квадратичного критерия

Будем формировать управляющие воздействия по текущей информации об объекте. При этом минимизируемый функционал можно представить в виде

$$J(k) = \frac{1}{2} \left[x^T(k+1) C x(k+1) + u^T(k) D u(k) \right], \quad (8.30)$$

где C — неотрицательно определенная, а D — положительно определенная весовые матрицы; k соответствует дискретному моменту времени t_k , а модель объекта можно описать системой линейных разностных уравнений

$$x(k+1) = A(k)x(k) + B(k)u(k), \quad x(0) = x_0. \quad (8.31)$$

Управление $u(k)$ будем определять из условия минимума функционала (8.30), т.е.

$$\frac{\partial J(k)}{\partial u(k)} = 0.$$

Запишем функционал (8.30) для модели (8.31) (для простоты записи аргумент k временно опустим):

$$\begin{aligned} J &= \frac{1}{2}[(Ax + Bu)^T C(Ax + Bu) + u^T D u] = \\ &= \frac{1}{2}[x^T A^T C A x + x^T A^T C B u + u^T B^T C A x + u^T B^T C B u + u^T D u]. \end{aligned}$$

Воспользовавшись правилами дифференцирования билинейной формы $x^T A y$, т.е.

$$\frac{\partial x^T A y}{\partial x} = y^T A^T; \quad \frac{\partial x^T A y}{\partial y} = x^T A,$$

где x, y — векторы-столбцы; A — матрица соответствующего порядка, получим

$$\frac{\partial J}{\partial u} = x^T A^T C B + u^T B^T C B + u^T D = 0$$

и

$$(B^T C B + D) u + B^T C A x = 0.$$

Тогда управление, полученное на основе минимума функционала (8.30), имеет вид

$$u(k) = -(B^T(k)CB(k) + D)^{-1} B^T(k)CA(k)x(k).$$

8.2.4. Синтез следящей системы управления

Ранее рассматривались задачи формирования управляющих воздействий, когда целью управления являлось приведение системы в нулевое состояние. Существует большой круг задач, где выдвигается требование достижения объектом некоторого заданного состояния $x_z(t)$ с последующим удержанием состояния объекта в малой окрестности $x_z(t)$, причем задача оптимизации остается нетерминальной.

Имеются, вообще говоря, две возможности формализации требований к динамике переходных процессов управляемого движения:

1) $x_z(t)$ формируется почти мгновенно по результатам анализа решения навигационной задачи или положения органов управления;

2) используется эталонная модель движения объекта, возбуждаемая сигналами навигационного комплекса или датчиками органов управления.

В первом случае все требования сводятся к соответствующему подбору параметров функционала, во втором — параметры функционала подбираются из условия обеспечения наилучшего слежения за эталоном.

Будем считать, что $x_z(t)$ — вектор, компоненты которого заданы в каждый момент времени функционирования системы управления. В этом случае минимизируемые функционалы (8.13) и (8.24) запишутся с помощью вектора рассогласования $(x(t) - x_z(t))$, а (8.30) — $(x(k+1) - x_z(k))$.

Управление при минимизации классического квадратичного функционала

$$J = \frac{1}{2} \int_{t_0}^T [(x(t) - x_z(t))^T C(x(t) - x_z(t)) + u^T(t) D u(t)] dt \quad (8.32)$$

для модели объекта (8.4) запишется в виде

$$u(t) = -D^{-1} \bar{B}^T(t) S(t) (x(t) - x_z(t)).$$

Для функционала обобщенной работы

$$J = \frac{1}{2} \int_{t_0}^T [(x(t) - x_z(t))^T C(x(t) - x_z(t)) + u^T(t) D_2 u(t) + v^T(t) D_1 v(t) + v_{оп}^T(t) D_1 v_{оп}(t)] dt \quad (8.33)$$

для модели объекта (8.18) получим

$$v(t) = -D_1^{-1} W_2(t)$$

и

$$\left. \begin{aligned} \dot{W}_1(\tau) &= -\bar{A}^T(\tau) W_1(\tau) - C(x(\tau) - x_z(\tau)), \quad W_1(T) = 0; \\ \dot{W}_2(\tau) &= -\bar{B}^T(\tau) W_1(\tau) - D_2 u(\tau), \quad W_2(T) = 0. \end{aligned} \right\} \quad (8.34)$$

При формировании локально-оптимального управления на основе минимизации функционала

$$J(k) = \frac{1}{2} [(x(k+1) - x_z(k))^T C(x(k+1) - x_z(k)) + u^T(k) D u(k)] \quad (8.35)$$

для модели объекта (8.31) получим

$$u(k) = -(B^T(k) C B(k) + D)^{-1} B^T(k) C [A(k)x(k) - x_z(k)].$$

8.3. Моделирование систем оптимального управления

Развитие цифровых информационно-измерительных комплексов привело к тому, что все чаще управление объектом в процессе его функционирования осуществляется с помощью компьютера. При этом центр тяжести процесса проектирования системы управления приходится на алгоритмическое и программное обеспечение.

8.3.1. Основные понятия цифровых систем управления

Включение ЭВМ в контур управления неизбежно создает особые условия его реализации, основными из которых являются дискретность формируемого управления по времени и запаздывание подачи управляющего воздействия на приводы рулевых органов по отношению к моменту выдачи измерительной системой всей необходимой информации. Такие системы управления оперируют только с цифровой информацией и называются *цифровыми системами управления*.

Цифровую систему управления непрерывным объектом с замкнутой обратной связью схематично можно изобразить следующим образом (рис. 8.3).

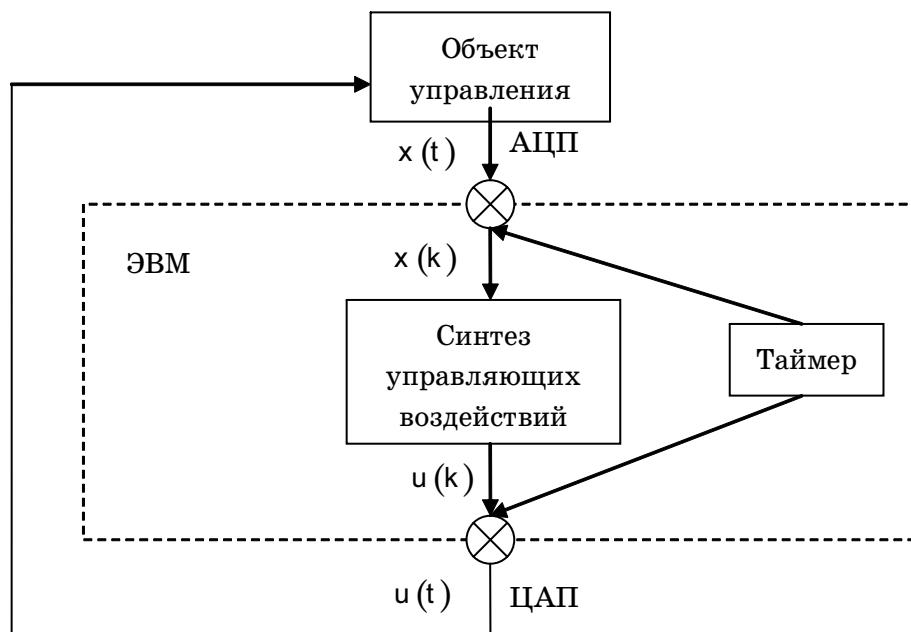


Рис. 8.3

Работа системы управления синхронизируется в ЭВМ таймером реального времени, который определяет момент поступления информации об объекте и момент воздействия управляющего сигнала на объект. При этом состояние объекта измеряется только в дискретные моменты времени. На входе в систему управления непрерывный вектор состояния $x(t)$ преобразуется в цифровую форму аналого-цифровым преобразователем (АЦП), а на выходе цифровой вектор управляющих воздействий $u(k)$ преобразуется в непрерывный сигнал $u(t)$ цифро-аналоговым преобразователем (ЦАП).

Момент времени, в который происходит преобразование непрерывной информации в дискретную, называется *моментом квантования* управляющего воздействия и обозначается k . Промежуток времени $\Delta t = t_{k+1} - t_k$ между двумя моментами квантования называется *периодом квантования*.

Квантование — неизбежный процесс в цифровых системах управления, обусловленный дискретной природой самих ЭВМ. Будем полагать, что для управления исходным непрерывным объектом используется цифровая система, формирующая кусочно-постоянный вектор управления $u(k)$ с моментом и периодом квантования, совпадающим с моментом и периодом поступления в ЭВМ информации об объекте. При этом непрерывность управляющего сигнала при воздействии на объект достигается путем сохранения его постоянной величины между формированиями. Таким образом, управление принадлежит к классу кусочно-постоянных непрерывных слева функций с фиксированным периодом квантования Δt . Если Δt достаточно мало, то цифровая система управления функционирует почти как непрерывная, так как потери информации за счет квантования будут незначительными. Но слишком малое время Δt может оказаться недостаточным для формирования управляющего сигнала, так как реализация систем управления в реальных условиях движения объекта характеризуется большой сложностью, при этом формирование управляющих воздействий должно осуществляться за время, пренебрежительно малое в сравнении со скоростью изменения внешней среды и самого объекта.

8.3.2. Моделирование поведения управляемого объекта

При проведении имитационных экспериментов для определения качества синтезируемого управления необходимо моделировать поведение управляемого объекта. Будем предполагать, что поведение объекта описывается системой обыкновенных линейных дифференциальных уравнений вида

$$\dot{x}(t) = \bar{A}(t)x(t) + \bar{B}(t)u(t), \quad x(t_0) = x_0. \quad (8.36)$$

Моделирование будем осуществлять в предположении, что шаг дискретизации Δt совпадает с периодом квантования управляющего сигнала, а момент дискретизации k — с моментом приложения управляющих воздействий. Причем управление является кусочно-постоянным на каждом интервале выдачи управляющих воздействий, то есть

$$u(t) = u(t_k), \quad t_k \leq t < t_{k+1}.$$

Тогда можно ограничиться описанием поведения объекта в моменты квантования и записать дискретный аналог для системы (8.36) в виде

$$\begin{aligned} x(t_{k+1}) &= \Phi_{\bar{A}}(t_{k+1}, t_k) x(t_k) + \\ &+ \int_{t_k}^{t_{k+1}} \Phi_{\bar{A}}(t_{k+1}, \tau) \bar{B}(\tau) u(\tau) d\tau, \quad x(t_0) = x_0, \end{aligned} \quad (8.37)$$

где $\Phi_{\bar{A}}(t_{k+1}, t_k)$ — фундаментальная матрица решений соответствующей однородной системы.

По свойству фундаментальной матрицы

$$\Phi_{\bar{A}}(t_{k+1}, t_k) = \exp[\bar{A}(t_k)(t_{k+1} - t_k)]. \quad (8.38)$$

Согласно свойству матричной экспоненты и учитывая, что $t_{k+1} - t_k = \Delta t$, (8.38) можно представить в виде ряда

$$\Phi_{\bar{A}}(t_{k+1}, t_k) = I_n + \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^i(t_k)}{i!}, \quad (8.39)$$

где I_n — единичная матрица порядка n .

Тогда, вычислив интеграл в (8.37) по формуле левых прямоугольников, дискретную систему (8.36) можно записать в разностной форме следующим образом:

$$x(k+1) = A(k)x(k) + B(k)u(k), \quad x(0) = x_0,$$

где

$$A(k) = I_n + \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^i(t_k)}{i!}; \quad (8.40)$$

$$B(k) = \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^{i-1}(t_k) \bar{B}(t_k)}{(i-1)!}; \quad (8.41)$$

k соответствует моменту времени $t_k = t_0 + k\Delta t$.

Представление матричной экспоненты в виде ряда (8.39) позволяет достаточно легко регулировать точность построения соответствующей дискретной системы, которую можно определять точностью вычисления фундаментальной матрицы для следующей стационарной системы:

$$\dot{x}(t) = \bar{A} x(t); \quad x(t_0) = x_0.$$

Дискретный аналог этой системы имеет вид

$$x(k+1) = A x(k); \quad x(0) = x_0,$$

где

$$A = I_n + \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^i}{i!}. \quad (8.42)$$

Если обозначить через A_N сумму N первых членов ряда (8.42), то матрица A_N аппроксимирует A с погрешностью порядка $O(\Delta t^N)$. При этом число слагаемых в A_N можно задавать заранее (вычисление с фиксированной точностью) или определять автоматически с помощью соотношения

$$\frac{\|A_N\| - \|A_{N-1}\|}{\|A_N\|} \leq \varepsilon,$$

где ε выбирается, например, из условия обеспечения максимальной точности вычислений на ЭВМ конкретного типа.

Если модель объекта задана в виде

$$\left. \begin{aligned} \dot{x}(t) &= \bar{A}(t)x(t) + \bar{B}(t)u(t), \quad x(t_0) = x_0; \\ \dot{u}(t) &= v(t), \quad u(t_0) = u_0, \end{aligned} \right\} \quad (8.43)$$

где $u(t)$ — вектор положения органов управления, а $v(t)$ — вектор управления, то дискретная система для (8.43) получается при введении расширенного вектора состояния $\tilde{x} = (x, u)^T$ и записи (8.43) в виде системы с блочными матрицами

$$\dot{\tilde{x}} = \tilde{A}(t)\tilde{x}(t) + \tilde{B}(t)v(t), \quad \tilde{x}(t_0) = \tilde{x}_0,$$

где $\tilde{A}(t)$, $\tilde{B}(t)$ определяются согласно (8.20), а $\tilde{x}_0 = (x_0, u_0)^T$.

Тогда система разностных уравнений для (8.43) примет вид

$$\left. \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) + B_1(k)v(k), \quad x(0) = x_0; \\ u(k+1) &= u(k) + \Delta t v(k), \quad u(0) = u_0, \end{aligned} \right\}$$

где $A(k)$ и $B(k)$ определяются согласно (8.40) и (8.41), а

$$B_1(k) = \sum_{i=2}^{\infty} \frac{\Delta t^i \bar{A}^{i-2}(t_k) \bar{B}(t_k)}{(i-1)!}.$$

Полученные соотношения для определения матриц дискретной системы являются общими. Так, если при вычислении матриц отбросить все слагаемые, порядок малости которых превышает $O(\Delta t^4)$, то полученные соотношения будут соответствовать моделированию поведения управляемого объекта методом Рунге-Кутты. А если отбросить слагаемые, имеющие порядок малости выше $O(\Delta t)$, то моделирование будет осуществляться методом Эйлера.

При малом значении Δt для моделирования достаточно часто используется метод Эйлера. В этом случае получаются достаточно простые соотношения для вычисления матриц дискретной системы:

$$A(k) = I_n + \Delta t \bar{A}(t_k); \quad B(k) = \Delta t \bar{B}(t_k); \quad B_1(k) = 0,$$

и система разностных уравнений для (8.43) примет вид

$$\left. \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k), \quad x(0) = x_0; \\ u(k+1) &= u(k) + \Delta t v(k), \quad u(0) = u_0. \end{aligned} \right\}$$

Моделирование поведения объекта в случае нестационарной модели (например, моделирование самолета на взлете) можно осуществлять следующим образом. На траектории движения объекта задаются значения матриц или элементов этих матриц, характеризующих модель объекта в некоторые моменты времени T_i , $i = \overline{1, N_R}$. Нестационарная модель в каждый момент времени $t \in [t_0, T]$ строится по заданным характеристикам, используя методы приближения данных: интерполирование, сплайн-функции, аппроксимацию по методу наименьших квадратов и т.д.

8.3.3. Синтез оптимального управления

Рассмотрим совмещенный синтез цифрового управления на основе оптимизации критериев классического квадратичного (8.32), обобщенной работы (8.33) и локального (8.35) при слежении за заданным состоянием $x_z(t)$. При этом если решается задача стабилизации, то полагаем $x_z(t) \equiv 0$.

Кроме того, будем предполагать, что:

- 1) все временные переменные кратны Δt ;
- 2) моменты и периоды квантования одинаковы для всех координат объекта;
- 3) цифровая система формирует вектор управления с моментом и периодом квантования, которые совпадают с моментом и периодом поступления информации об объекте;
- 4) управления являются кусочно-постоянными функциями на каждом интервале выдачи управляющих воздействий

$$u(t) = u(k), \quad t_k \leq t < t_{k+1}.$$

Пусть в управляющем компьютере модель объекта представлена в виде

$$x(k+1) = A(k)x(k) + B(k)u(k), \quad x(0) = x_0 \quad (8.44)$$

или

$$\left. \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k), \quad x(0) = x_0; \\ u(k+1) &= u(k) + \Delta t v(k), \quad u(0) = u_0, \end{aligned} \right\} \quad (8.45)$$

где

$$A(k) = I_n + \Delta t \bar{A}(t_k); \quad B(k) = \Delta t \bar{B}(t_k). \quad (8.46)$$

Форма представления модели объекта зависит от решаемой задачи управления: управление положением органов управления или управление скоростью отклонения органов управления.

Синтез управляющих воздействий в процессе функционирования объекта накладывает определенные ограничения на его реализацию. Основным ограничением является то, что при формировании управления в момент $t \in [t_0, T]$ известна информация о поведении объекта в предыдущие моменты времени и неизвестна — в последующие. В связи с этим будем рассматривать алгоритмы синтеза управляющих воздействий, учитывающие эти ограничения и предназначенные для совместного синтеза.

Синтез управляющих воздействий по классическому квадратичному критерию

Если модель объекта (8.44) является стационарной, то управление, формируемое на k -м такте (в момент $t_k = t_0 + k\Delta t$) на основе минимизации функционала (8.32), имеет вид

$$u(k) = -D_d^{-1} B^T S(x(k) - x_z(t_k)),$$

где матрица S является решением уравнения Риккати, которое путем замены переменной можно записать в прямом времени:

$$\left. \begin{aligned} \dot{S}(t) &= \bar{A}^T S(t) + S(t) \bar{A} - S(t) \bar{B} D_d^{-1} \bar{B}^T S(t) + C; \\ S(0) &= 0. \end{aligned} \right\} \quad (8.47)$$

При этом можно искать установившееся решение уравнения (8.47) по следующему итерационному алгоритму:

$$\left. \begin{aligned} S(i+1) &= A^T S(i) + S(i) A - S(i) B D_d^{-1} B^T S(i) - S(i) + C_d; \\ S(0) &= 0, \end{aligned} \right\} \quad (8.48)$$

и, при выполнении условия

$$\frac{\|S(i+1) - S(i)\|}{\|S(i+1)\|} \leq \varepsilon,$$

полагать $S = S(i+1)$. Здесь $D_d = \Delta t D$; $C_d = \Delta t C$; ε — заданная точность решения уравнения Риккати. Заметим, что уравнение (8.48) записано с помощью матриц дискретной системы, которая получена из непрерывной методом Эйлера.

Если модель объекта (8.44) является нестационарной, то запишем минимизируемый функционал с использованием скользящего интервала оптимизации, то есть:

$$J(t_k) = \frac{1}{2} \int_{t_k}^{t_k + l_p \Delta t} [(x(t) - x_z(t_k))^T C(x(t) - x_z(t_k)) + u^T(t) D u(t)] dt,$$

и поведение объекта на скользящем интервале оптимизации $[t_k, t_k + l_p \Delta t]$ будем описывать стационарной системой вида

$$\left. \begin{aligned} x_M(j+1) &= A(k)x_M(j) + B(k)u(j), \quad x_M(j=k) = x(k); \\ j &= k, k+1, \dots, k+l_p-1, \end{aligned} \right\}$$

которая является прогнозирующей моделью.

Тогда управление, формируемое на k -м такте, запишется в виде

$$u(k) = -D_d^{-1}B(k)S_k[x(k) - x_z(t_k)],$$

где S_k — решение уравнения Риккати, которое в силу описания модели объекта на интервале оптимизации с помощью стационарной прогнозирующей модели можно записать в прямом времени и решать итерационным методом:

$$\begin{aligned} S(i+1) &= A^T(k)S(i) + S(i)A(k) - \\ &- S(i)B(k)D_d^{-1}B^T(k)S(i) - S(i) + C_d, \quad S(0) = 0. \end{aligned}$$

При этом матрица S_k определяется следующим образом:

$$S_k = \begin{cases} S(i), & \text{если } \frac{\|S(i+1) - S(i)\|}{\|S(i+1)\|} \leq \varepsilon; \\ S(l_p), & \text{в противном случае.} \end{cases}$$

Синтез управляющих воздействий по квадратичному критерию обобщенной работы

Для модели объекта (8.45) функционал обобщенной работы запишем для скользящего интервала оптимизации $[t_k, t_k + l_p \Delta t]$:

$$\begin{aligned} J(t_k) &= \frac{1}{2} \int_{t_k}^{t_k + l_p \Delta t} [(x(t) - x_z(t_k))^T C(x(t) - x_z(t_k)) + \\ &+ u^T(t)D_2u(t) + v^T(t)D_1v(t) + v_{оп}^T(t)D_1v_{оп}(t)]dt, \end{aligned} \quad (8.49)$$

и будем описывать поведение объекта на интервале оптимизации с помощью стационарной прогнозирующей модели, характеризующей «свободное» движение объекта ($v(t) = 0$)

$$\begin{aligned} x_M(j+1) &= A(k)x_M(j) + B(k)u(k), \quad x_M(j=k) = x(k); \\ j &= k, k+1, \dots, k+l_p-1. \end{aligned} \quad (8.50)$$

Тогда управление будет иметь вид

$$v(k) = -D_1^{-1}W_2(k),$$

где $W_2(k)$ определяется решением в обратном времени системы обыкновенных дифференциальных уравнений (8.34) на интервале

оптимизации $[t_k, t_k + l_p \Delta t]$. Так как решение этой системы зависит от значений состояния прогнозирующей модели на интервале $[t_k, t_k + l_p \Delta t]$, то сначала, решая в прямом времени уравнение (8.50), определяется состояние прогнозирующей модели в конечный момент интервала оптимизации $x_M(k + l_p)$, а затем $W_2(k)$ определяется решением в обратном времени системы разностных уравнений вида

$$\left. \begin{aligned} x'_M(k + l_p - (j + 1)) &= 2x'_M(k + l_p - j) - A(k)x'_M(k + l_p - j) - \\ &\quad - B(k)u(k), \quad x'_M(k + l_p) = x_M(k + l_p); \\ W_1(k + l_p - (j + 1)) &= A^T(k)W_1(k + l_p - j) + \\ &\quad + \Delta t C(x'_M(k + l_p - j) - x_z(t_k)), \quad W_1(k + l_p) = 0; \\ W_2(k + l_p - (j + 1)) &= W_2(k + l_p - j) + B^T(k)W_1(k + l_p - j) + \\ &\quad + \Delta t D_2 u(k), \quad W_2(k + l_p) = 0, \end{aligned} \right\} \quad (8.51)$$

где первое уравнение описывает движение прогнозирующей модели (8.50) в обратном времени.

Синтез управляющих воздействий по локальному квадратичному критерию

Для модели объекта (8.44) локальный квадратичный критерий запишется следующим образом:

$$J(k) = \frac{1}{2} [(x(k + 1) - x_z(k))^T C(x(k + 1) - x_z(t_k)) + u^T(k) D u(k)]. \quad (8.52)$$

Управление, формируемое на основе минимизации этого критерия, имеет вид

$$u(k) = -(B^T(k)CB(k) + D)^{-1}B^T(k)C[A(k)x(k) - x_z(t_k)].$$

Качество функционирования системы управления зависит от весовых матриц рассматриваемых критериев, определение которых осуществляется на этапе предварительного проектирования путем коррекции элементов весовых матриц и анализа получаемых при этом переходных процессов и управлений.

Для выбора весовых матриц C и D локального критерия (8.52) можно воспользоваться методом случайного поиска или назначить их, исходя из каких-либо конструктивных соображений с учетом свойств управляемого объекта. Достаточно хорошо работает следующая методика.

Запишем для (8.44) систему вида

$$x(k+1) = \tilde{A}x(k) + \tilde{B}u(k); \quad x(0) = x_0,$$

где \tilde{A} , \tilde{B} — некоторые постоянные матрицы, например $\tilde{A} = A(0)$, $\tilde{B} = B(0)$. Используя принцип равного вклада максимальных отклонений, выбираются весовые матрицы для суммарного критерия

$$\sum_{k=0}^{\infty} [x^T(k+1)C_1x(k+1) + u^T(k)D_1u(k)]$$

в предположении, что требования к качеству функционирования системы управления для всех k задаются в виде неравенств для компонентов векторов состояния и управления:

$$|x_i(k)| \leq x_{\max_i}, \quad \overline{i = 1, n};$$

$$|u_i(k)| \leq u_{\max_i}, \quad \overline{i = 1, m}.$$

Тогда полагают

$$C_1 = \text{diag} \{x_{\max_1}^{-2}, \dots, x_{\max_n}^{-2}\},$$

$$D_1 = \text{diag} \{u_{\max_1}^{-2}, \dots, u_{\max_m}^{-2}\},$$

где $\text{diag} \{...\}$ — диагональная матрица соответствующей размерности с указанными элементами на главной диагонали. Весовые матрицы C и D локального критерия полагаются равными

$$C = \tilde{A}^T C \tilde{A} - \tilde{A}^T C \tilde{B} (\tilde{B}^T C \tilde{B} + D_1)^{-1} \tilde{B}^T C \tilde{A} + C_1; \quad (8.53)$$

$$D = D_1.$$

Алгебраическое уравнение Риккати (8.53) можно решить по итерационной схеме

$$C(i+1) = [\tilde{A}^T C(i) \tilde{A} - \tilde{A}^T C(i) \tilde{B} (\tilde{B}^T C(i) \tilde{B} + D)^{-1} \tilde{B}^T C(i) \tilde{A} + C_1 - C(i)] \Delta \eta + C(i), \quad C(0) = C_1, \quad (8.54)$$

где $\Delta \eta = \Delta t \eta$, η — параметр, который подбирается таким образом, чтобы уравнение (8.54) было сходящимся. При выполнении условия

$$\frac{\|C(i+1) - C(i)\|}{\|C(i+1)\|} \leq \varepsilon$$

матрица C полагается равной $C(i+1)$.

Структурная схема совмещенного синтеза цифрового управления объектом представлена на рис. 8.4.

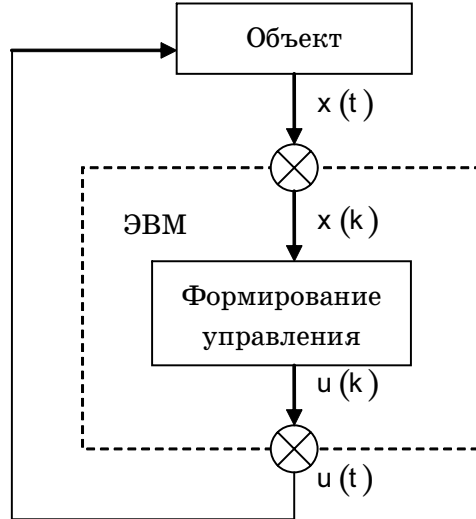


Рис. 8.4.

8.4. Моделирование систем управления при случайных внешних воздействиях

При синтезе управляющих воздействий в реальных условиях информация о состоянии объекта поступает с измерительного комплекса, который достаточно часто содержит неполную информацию об объекте, искаженную ошибками измерений. Кроме того, при описании модели объекта необходимо учитывать влияние внешних возмущений. В связи с этим будем рассматривать алгоритмы синтеза управляющих воздействий, учитывающие условия их реализации.

8.4.1. Моделирование поведения объекта при наличии внешних возмущений

Будем предполагать, что внешние возмущения являются аддитивными и описываются вектором гауссовских шумов с заданной матрицей влияния $\bar{F}(t)$. Тогда модель объекта можно записать в виде

$$\dot{x}(t) = \bar{A}(t)x(t) + \bar{B}(t)u(t) + \bar{F}(t)q(t); \quad x(t_0) = x_0, \quad (8.55)$$

где $q(t)$ — l_1 -мерный вектор гауссовских шумов с заданными характеристиками

$$M \{q(t)\} = \bar{q}(t);$$

$$M \left\{ (q(t) - \bar{q}(t))(q(\tau) - \bar{q}(\tau))^T \right\} = Q(t)\delta(t - \tau),$$

где $\delta(t - \tau)$ — дельта-функция Дирака; $\bar{F}(t)$ — матрица влияния внешних возмущений размерности $n \times l_1$.

Для моделирования поведения объекта строится разностный аналог уравнения (8.55) с помощью фундаментальных матриц соответствующей однородной системы и в предположении, что $q(t)$ является кусочно-постоянной функцией

$$q(t) = q(t_k), \quad t_k \leq t < t_{k+1}.$$

Тогда разностное уравнение, соответствующее (8.55), имеет вид

$$x(k+1) = A(k)x(k) + B(k)u(k) + F(k)q(k), \quad x(0) = x_0, \quad (8.56)$$

где $A(k)$, $B(k)$ и $F(k)$ вычисляются следующим образом:

$$A(k) = I_n + \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^i(t_k)}{i!}; \quad (8.57)$$

$$B(k) = \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^{i-1}(t_k) \bar{B}(t_k)}{(i-1)!}; \quad (8.58)$$

$$F(k) = \sum_{i=1}^{\infty} \frac{\Delta t^i \bar{A}^{i-1}(t_k) \bar{F}(t_k)}{(i-1)!}. \quad (8.59)$$

В (8.56) $q(k)$ — последовательность гауссовских величин, при этом

$$M \{q(k)\} = \bar{q}(k) = \bar{q}(t_k);$$

$$M \left\{ (q(k) - \bar{q}(k))(q(j) - \bar{q}(j))^T \right\} = Q(k) \delta_{kj};$$

$$Q(k) = \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} M \left\{ (q(t) - \bar{q}(t))(q(\tau) - \bar{q}(\tau))^T \right\} dt d\tau = Q(t_k) \Delta t,$$

где δ_{kj} — символ Кронекера.

Для того чтобы внешние возмущения $q(t)$ и $q(k)$ имели одни и те же статистические характеристики, матрицу $F(k)$ представим в виде

$$F(k) = \sqrt{\Delta t} \sum_{i=1}^{\infty} \frac{\Delta t^{i-1} \bar{A}^{i-1}(t_k) \bar{F}(t_k)}{(i-1)!}. \quad (8.60)$$

Если для моделирования используется метод Эйлера, то

$$F(k) = \sqrt{\Delta t} \bar{F}(t_k),$$

а при использовании метода Рунге-Кутты в (8.60) в сумме остаются только четыре слагаемых.

Если рассматривается задача управления скоростью отклонения органов управления, то есть

$$\left. \begin{aligned} \dot{x}(t) &= \bar{A}(t)x(t) + \bar{B}(t)u(t) + \bar{F}(t)q(t), \quad x(t_0) = x_0; \\ \dot{u}(t) &= v(t), \quad u(t_0) = u_0, \end{aligned} \right\} \quad (8.61)$$

то эквивалентная дискретная система имеет вид

$$\left. \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) + B_1(k)v(k) + F(k)q(k), \quad x(0) = x_0; \\ u(k+1) &= u(k) + \Delta t v(k), \quad u(0) = u_0, \end{aligned} \right\}$$

где $A(k)$, $B(k)$ и $F(k)$ вычисляются согласно (8.57), (8.58), (8.60), а

$$B_1(k) = \sum_{i=2}^{\infty} \frac{\Delta t^i \bar{A}^{i-2}(t_k) \bar{B}(t_k)}{(i-1)!}.$$

Если для построения дискретной системы используется метод Эйлера, то (8.61) имеет вид

$$\left. \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) + F(k)q(k), \quad x(0) = x_0; \\ u(k+1) &= u(k) + \Delta t v(k), \quad u(0) = u_0. \end{aligned} \right\} \quad (8.62)$$

При использовании метода Эйлера для построения дискретных систем (8.56) и (8.62) матрицы $A(k)$, $B(k)$ и $F(k)$ имеют вид

$$A(k) = I_n + \Delta t \bar{A}(t_k); \quad B(k) = \Delta t \bar{B}(t_k); \quad F(k) = \sqrt{\Delta t} \bar{F}(t_k).$$

8.4.2. Описание математической модели измерительного комплекса

Использование компьютеров для формирования управляющих воздействий приводит к тому, что измерения поступают в дискретные моменты времени t_1, t_2, \dots , а управления являются кусочно-постоянными непрерывными слева функциями на каждом интервале выдачи управляющих воздействий, то есть

$$u(t) = u(k); \quad t_k \leq t < t_{k+1}; \quad t_{k+1} - t_k = \Delta t.$$

Будем предполагать, что математическая модель измерительного комплекса имеет вид

$$y(k) = Hx(k) + r(k),$$

где $y(k)$ — l -мерный вектор измерений; H — матрица канала измерений размерности $l \times n$, состоящая из нулей и единиц, нулевые столбцы которой соответствуют неизмеряемым компонентам вектора состояния; $r(k)$ — дискретный гауссовский шум с характеристиками

$$M\{r(k)\} = 0; \quad M\left\{ \begin{bmatrix} r(k) \\ r(j) \end{bmatrix} \right\} = R \delta_{kj};$$

$x(k)$ — состояние преобразованной в дискретную форму непрерывной системы, описывающей модель объекта.

8.4.3. Оценивание состояния модели объекта

В реальных условиях управление объектом реализуется по результатам измерений, при этом все измерения осуществляются с погрешностью, и, кроме того, в модели объекта могут существовать компоненты вектора состояния, недоступные измерению. Поэтому для формирования управляющих воздействий необходимо оценивать вектор состояния. Так как информация о модели объекта поступает с измерительного комплекса в дискретные моменты времени, то будем строить оценку $\hat{x}(k)$ вектора состояния $x(k)$ по результатам текущих измерений $y(k)$.

Основные подходы к построению оценок векторов состояния состоят в следующем. Предположим, что для некоторой системы состояние $x(k)$ недоступно непосредственному измерению, а получены последовательные измерения $y(1), y(2), \dots, y(j)$. Обозначим оценку $x(k)$, полученную на основе этих измерений, через $\hat{x}(k/j)$ и определим ее как n -мерную вектор-функцию измерений

$$\hat{x}(k/j) = \varphi_k \{y(i), i = \overline{1, j}\}.$$

Задача построения оценки для состояния $x(k)$ представляет собой задачу определения функции φ_k некоторым рациональным и обоснованным способом.

Если $k > j$, то задача построения оценки называется *задачей предсказания*, если $k = j$, то такая задача называется *задачей фильтрации*, если $k < j$, то получаем *задачу сглаживания* или *интерполяции*.

Рассмотрим функцию ошибки оценки

$$\tilde{x}(k/j) = x(k) - \hat{x}(k/j)$$

и обозначим через $L = L[\tilde{x}(k/j)]$ функцию *потерь*, или *штрафов*. Эта функция должна обладать следующими свойствами:

- 1) быть скалярной функцией;
- 2) $L(0) = 0$;
- 3) быть неубывающей функцией расстояния от начала координат;
- 4) быть симметричной относительно нуля.

Функция L , обладающая такими свойствами, называется *допустимой функцией потерь*.

Определим критерий качества оценивания как среднее значение функции потерь:

$$J[\tilde{x}(k/j)] = M \{L[\tilde{x}(k/j)]\}.$$

Говорят, что оценка $\hat{x}(k/j)$, минимизирующая $J[\tilde{x}(k/j)]$, является *наилучшей*, или *оптимальной*. Заметим, что оптимальная оценка минимизирует среднее значение функции потерь, то есть является *оптимальной в среднем*.

Пусть

$$L[\tilde{x}(k/j)] = \tilde{x}^T(k/j)\tilde{x}(k/j).$$

Соответствующий этой функции потерь критерий качества оценки называется *среднеквадратической ошибкой*, так как $M\{\tilde{x}^T(k/j)\tilde{x}(k/j)\}$ есть среднее значение квадрата евклидовой меры вектора ошибки.

Для построения оценки вектора состояния используются различные подходы. Рассмотрим рекуррентный алгоритм статистической обработки типа фильтра Калмана, осуществляющий оценивание состояния по результатам текущих измерений.

Пусть математические модели объекта и канала измерений имеют вид

$$x(k+1) = A(k)x(k) + B(k)u(k) + F(k)q(k), \quad x(0) = x_0; \quad (8.63)$$

$$y(k) = Hx(k) + r(k), \quad (8.64)$$

где $q(k)$ и $r(k)$ — независимые гауссовские последовательности с характеристиками:

$$\left. \begin{aligned} M\{q(k)\} &= \bar{q}(k), \quad M\{(q(k) - \bar{q}(k))(q(j) - \bar{q}(j))^T\} = Q\delta_{kj}; \\ M\{r(k)\} &= 0, \quad M\{r(k)r^T(j)\} = R\delta_{kj}; \\ M\{q(k)r^T(j)\} &= 0. \end{aligned} \right\}$$

Кроме того, предполагается, что априорное распределение вектора x_0 является гауссовским:

$$M\{x_0\} = \bar{x}_0, \quad M\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} = P_{x_0},$$

где P_{x_0} — дисперсионная матрица ошибок оценивания начального состояния.

Обозначим оценку фильтрации, полученную по результатам измерений $y(1), y(2), \dots, y(k)$, через $\hat{x}(k)$. При принятых предположениях задача оптимальной фильтрации может быть сформулирована следующим образом. Располагая последовательностью векторов измерений $y(1), y(2), \dots, y(k)$, моделью которых является система уравнений (8.63), (8.64), требуется определить оценку $\hat{x}(k)$ вектора состояния $x(k)$, удовлетворяющую требованиям несмещенности

$$M\{\hat{x}(k)\} = M\{x(k)\} \quad (8.65)$$

при минимальной дисперсии ошибок оценки.

Пусть $\varepsilon(k) = \hat{x}(k) - x(k)$ — ошибка оценки, а $P(k) = M\{\varepsilon(k)\varepsilon^T(k)\}$ — дисперсионная матрица ошибки оценки.

Тогда критерий оптимальности запишется в виде

$$J(k) = M\{\varepsilon^T(k)\varepsilon(k)\} = M\{\text{tr}(\varepsilon(k)\varepsilon^T(k))\} = \text{tr}(P(k)), \quad (8.66)$$

где $\text{tr}(\cdot)$ обозначает след матрицы.

Будем строить оптимальную оценку $\hat{x}(k+1)$ по текущему измерению $y(k+1)$ в виде

$$\hat{x}(k+1) = \hat{x}(k/k+1) + K(k)[y(k+1) - H\hat{x}(k/k+1)], \quad (8.67)$$

где вектор

$$\hat{x}(k/k+1) = A(k)\hat{x}(k) + B(k)u(k) + F(k)\bar{q}(k), \quad \hat{x}(0) = \bar{x}_0 \quad (8.68)$$

является экстраполированной оценкой вектора состояния $x(k+1)$; $K(k)$ — матрица коэффициентов усиления фильтра.

Таким образом, задача построения оптимальной оценки дискретной системы (8.63) сводится к задаче определения такой матрицы $K(k)$, при которой оценка $\hat{x}(k+1)$ удовлетворяет условиям (8.65) при ограничениях (8.63)–(8.64), (8.67)–(8.68) и обеспечивает минимум критерия (8.66).

Заметим, что выражение (8.67) объединяет множество фильтров, различающихся законом изменения матриц коэффициентов усиления, и характеризует рекуррентную форму выработки оценок состояния.

Получим уравнение для вектора ошибки оценки $\varepsilon(k+1)$. Для этого подставим в (8.67) выражение (8.68) и вычтем (8.63) с учетом (8.64). В дальнейшем для сокращения записи опустим аргумент k везде, кроме векторов ошибок и дисперсионных матриц, тогда

$\varepsilon(k+1) = (A - KHA)\varepsilon(k)(A - KHA)^T - (I - KH)F\bar{q} + (KH - I)Fq + Kr$
и дисперсионная матрица ошибки оценки $\varepsilon(k+1)$ будет равна

$$P(k+1) = (A - KHA)P(k)(A - KHA)^T - (KH - I)FQF^T(KH - I)^T + KRK^T, \quad (8.69)$$

где I — единичная матрица соответствующего порядка.

В соответствии с критерием оптимальности (8.66) матрица усиления фильтра K определяется из условия

$$\frac{dJ(k)}{dK} = \frac{d \operatorname{tr} P(k+1)}{dK} = 0,$$

или

$$\begin{aligned} & -2AP(k)A^TH^T + 2KHAP(k)A^TH^T - 2FQF^TH^T + \\ & + 2KHFQF^TH^T + 2KR = 0. \end{aligned}$$

Из последнего равенства, после приведения подобных членов, получается

$$KH(AP(k)A^T + FQF^T)H^T + KR = (AP(k)A^T + FQF^T)H^T. \quad (8.70)$$

Обозначим через

$$P(k+1/k) = AP(k)A^T + FQF^T \quad (8.71)$$

прогноз дисперсионной матрицы на один шаг вперед. Тогда выражение (8.70) с учетом (8.71) примет вид

$$KHP(k+1/k)H^T + KR = P(k+1/k)H^T,$$

и для определения матрицы K получается выражение

$$K = P(k+1/k)H^T (HP(k+1/k)H^T + R)^{-1}. \quad (8.72)$$

Выражение для дисперсионной матрицы $P(k+1)$ можно упростить, если в (8.69) подставить (8.72) и учесть (8.71). Тогда получим

$$P(k+1) = (I - KH)P(k+1/k).$$

Таким образом, построение оценки $\hat{x}(k+1)$ по результатам измерений $y(k+1)$ осуществляется с помощью следующего рекуррентного алгоритма:

$$\left. \begin{aligned} \hat{x}(k+1) &= \hat{x}(k+1/k) + K(k)[y(k+1) - H\hat{x}(k+1/k)]; \\ \hat{x}(k+1/k) &= A(k)\hat{x}(k) + B(k)u(k) + F(k)\bar{q}(k), \hat{x}(0) = \bar{x}_0; \\ K(k) &= P(k+1/k)H^T [HP(k+1/k)H^T + R]^{-1}; \\ P(k+1/k) &= A(k)P(k)A^T(k) + F(k)QF^T(k); \\ P(k+1) &= [I - K(k)H]P(k+1/k); \\ P(0) &= P_{x_0}. \end{aligned} \right\} \quad (8.73)$$

8.4.4. Синтез управляющих воздействий по оценкам состояния

Для классического критерия оптимальности и критерия обобщенной работы доказаны *теоремы разделения*, согласно которым для линейного управляемого объекта, математическая модель которого описывается уравнениями типа (8.55) или (8.61), при линейном наблюдении с аддитивным белым гауссовским шумом минимизация математического ожидания функционалов приводит к системе, состоящей из фильтра Калмана (8.73) и алгоритма управления, структура которого совпадает с законом оптимального управления, построенным для детерминированной полностью известной модели объекта.

Таким образом, управление для модели объекта (8.55), синтезируемое на основе минимизации математического ожидания классического квадратичного функционала, будет формироваться по оценкам состояния, то есть вместо значений вектора состояния $x(k)$ в соответствующих формулах будет использоваться его оценка $\hat{x}(k)$. Аналогичная ситуация складывается при формировании управления на основе минимизации математического ожидания функционала обобщенной

работы (8.49). В этом случае оценка состояния $\hat{x}(k)$ заменит вектор состояния в уравнении прогнозирующей модели, в которой будем использовать постоянные внешние возмущения $\bar{q}(k)$, равные математическому ожиданию шума в модели объекта в начальный момент интервала оптимизации:

$$\left. \begin{aligned} x_M(j+1) &= A(k)x_M(j) + B(k)u(k) + F(k)\bar{q}(k); \\ x_M(j=k) &= \hat{x}(k), \quad j = k, k+1, \dots, k+l_p-1. \end{aligned} \right\}$$

Постоянные внешние возмущения $\bar{q}(k)$ должны учитываться и в уравнении, описывающем движение прогнозирующей модели в обратном времени, то есть первое уравнение в системе (8.51) должно иметь вид

$$\begin{aligned} x'_M(k+l_p-(j+1)) &= 2x'_M(k+l_p-j) - A(k)x'_M(k+l_p-j) - \\ &\quad - B(k)u(k) - F(k)\bar{q}(k); \\ x'_M(k+l_p) &= x_M(k+l_p). \end{aligned}$$

При формировании управления для модели объекта вида

$$x(k+1) = A(k)x(k) + B(k)u(k) + F(k)q(k), \quad x(0) = x_0 \quad (8.74)$$

запишем математическое ожидание для локального критерия:

$$J(k) = \frac{1}{2} M \left\{ \left(x(k+1) - x_z(t_k) \right)^T C \left(x(k+1) - x_z(t_k) + u^T(k) D u(k) \right) \right\}. \quad (8.75)$$

Для того чтобы найти оптимальное управление по критерию (8.75), запишем значение этого критерия, учитывая уравнение для модели объекта (8.74) и свойства операции $\text{tr}(\cdot)$ (для сокращения записи аргумент k опустим):

$$\begin{aligned} J(k) &= \frac{1}{2} [\text{tr}(CFQF^T) + u^T(B^T C B + D)u + \\ &\quad + (Ax + Fq - x_z)^T C B u + u^T B^T C (Ax + F\bar{q} - x_z)]. \end{aligned}$$

Оптимальное управление определим из условия

$$\frac{\partial J(k)}{\partial u} = (B^T C B + D)u + B^T C (Ax + F\bar{q} - x_z) = 0.$$

Тогда получим

$$u(k) = -(B^T(k)CB(k) + D)^{-1} B^T(k) C [A(k)\hat{x}(k) + F(k)\bar{q}(k) - x_z(t_k)].$$

Структурная схема совмещенного синтеза управляющих воздействий по оценкам состояния приведена на рис. 8.5.

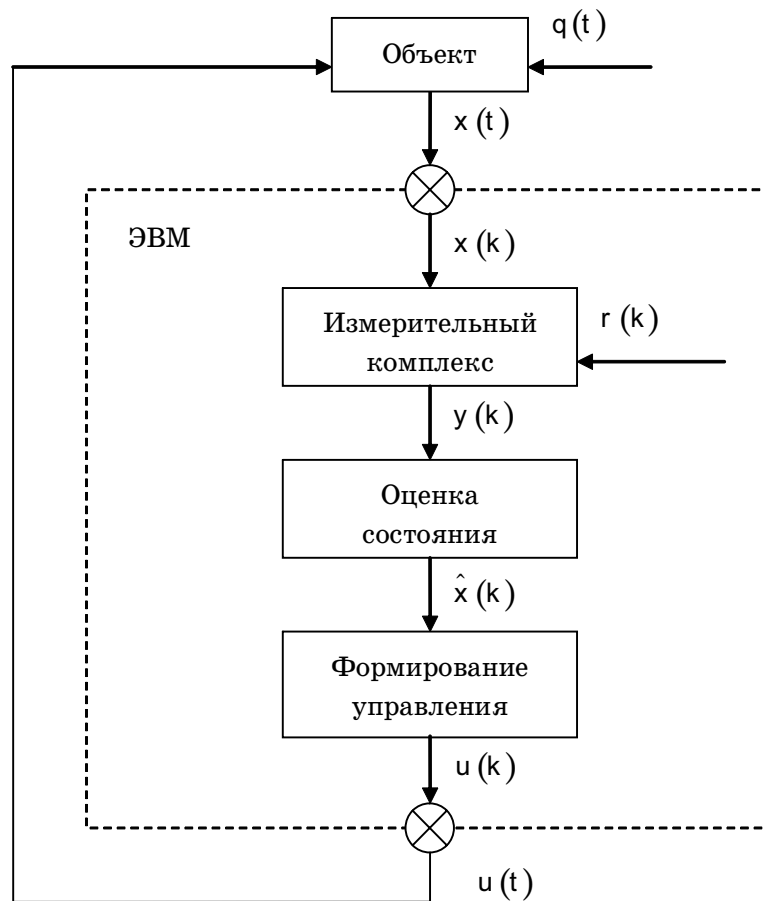


Рис. 8.5

8.5. Синтез адаптивной следящей системы

Процесс проектирования систем управления всегда предполагает наличие четко сформулированной цели управления и априорной информации об объекте управления и о характере действующих на него возмущений.

Объем априорной информации при этом может быть различным и за редким исключением не является исчерпывающим. Однако в данном случае принципиальным является вопрос о достаточности или недостаточности имеющейся априорной информации об объекте для достижения сформулированной цели управления. Все системы управления, построенные с использованием априорной информации, достаточной для достижения поставленной цели, относятся к *неадаптивным*, или

традиционным, системам управления независимо от реализуемого принципа управления, наличия обратной связи, случайности или детерминированности возмущений, используемых вычислительных средств и т.д. Если же объем располагаемой априорной информации о свойствах объекта не может обеспечить достижения сформулированной цели управления, то речь идет об *адаптивных* системах управления.

Существует достаточно много вариантов различных определений адаптивных систем. Но основной смысл при этом остается неизменным. Одним из наиболее известных является определение Я.З. Цыпкина: *под адаптивными системами управления будем понимать такие системы, которые, используя информацию о внешних воздействиях, условиях работы и выходных величинах, изменяют структуру или параметры регулятора с целью обеспечения оптимального или заданного функционирования управляемой системы при изменяющихся условиях ее работы.*

Достаточно известным является определение, данное В.Н. Буковым: *адаптивными называются такие системы управления, которые предназначены для функционирования в условиях априорной неопределенности и которые в процессе функционирования автоматически приспосабливаются к непредвиденным изменениям свойств объекта управления и внешней среды.*

Если изменяется структура регулятора, то адаптивная система управления называется *самоорганизующейся*, если структура остается неизменной, но изменяются параметры регулятора, то система управления называется *самонастраивающейся с параметрической адаптацией*.

8.5.1. Основные понятия адаптивных систем управления

Наиболее существенными являются следующие признаки деления принципов адаптации.

1. По уровню априорной неопределенности:

а) параметрическая адаптация, при которой априорная неопределенность заключается в недостаточном знании параметров (коэффициентов) управляемого объекта;

б) непараметрическая адаптация, при которой априорная неопределенность не связана непосредственно с какими-либо параметрами.

В обоих случаях неопределенность уменьшается на основе последовательных наблюдений входных и выходных данных в процессе управления.

2. По организации процесса адаптации:

а) поисковые, для которых характерны процессы итеративного движения к достижению требуемого качества управления;

б) беспойсковые, основанные на использовании некоторых необходимых или достаточных условий требуемого качества управления.

У систем поисковой адаптации формируются специальными устройствами детерминированные или случайные пробные сигналы или

создаются условия для возбуждения в объекте незатухающих колебаний, которые используются как поисковые. Наличие пробных движений является основным недостатком поисковой адаптации, так как они не всегда допустимы по условиям функционирования объекта.

3. По целям организации адаптации:

а) системы со специальными свойствами, в результате функционирования которых управляемый процесс приобретает некоторые обязательные свойства (устойчивость, чувствительность к каким-либо возмущениям или ошибкам априорной информации, заданное расположение корней характеристического уравнения и т.д.);

б) оптимальные системы, обеспечивающие минимизацию некоторых функционалов, отражающих качество управляемого движения.

Наиболее перспективными для создания цифровых систем управления являются оптимальные беспоисковые системы с параметрической адаптацией, в которых на основе полученной в процессе функционирования объекта информации о самом объекте и возмущающих воздействиях осуществляется автоматическая настройка параметров модели объекта. При этом используется «принцип разделения», согласно которому формирование адаптивного управления состоит из трех этапов:

1) выбор алгоритма управления в предположении, что параметры модели объекта известны точно (оптимальное управление);

2) оценивание неизвестных параметров (идентификация) модели объекта;

3) формирование адаптивного управления посредством замены точных значений параметров на их оценки.

8.5.2. Одновременное оценивание состояния и параметров модели объекта

Для синтеза управлений в реальных условиях функционирования объекта необходимо учитывать как влияние внешних возмущений, так и наличие переменных во времени неизвестных параметров.

Пусть для формирования управляющих воздействий используется математическая модель объекта в виде

$$\begin{aligned} x(k+1) = A(k, \theta(k))x(k) + B(k, \theta(k))u(k) + \\ + F(k)q(k), \quad x(0) = x_0, \end{aligned} \quad (8.76)$$

где $\theta(k)$ — N -мерный вектор переменных во времени неизвестных параметров. При этом предполагается, что элементы матриц A и B линейно зависят от компонентов вектора θ , что соответствует дискретизации по методу Эйлера, и априорные распределения векторов начальных условий $\theta(0)$ и x_0 являются гауссовскими:

$$\left. \begin{aligned} M\{\theta(0)\} &= \bar{\theta}_0, \quad M\{(\theta(0) - \bar{\theta}_0)(\theta(0) - \bar{\theta}_0)^T\} = P_{\theta_0}; \\ M\{x_0\} &= \bar{x}_0; \quad M\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} = P_{x_0}. \end{aligned} \right\}$$

Здесь P_{θ_0} и P_{x_0} — ковариационные матрицы ошибок начальных условий векторов параметров и состояния модели объекта.

При совмещенном синтезе требуется одновременно оценивать состояние и параметры системы, описывающей поведение объекта в моменты квантования. Это обычно осуществляется с помощью рекуррентного алгоритма статистической обработки типа фильтра Калмана по результатам текущих измерений. При этом оценивается обобщенный вектор, формально объединяющий векторы состояния и параметров. Такой подход сопровождается резким возрастанием трудоемкости вычислений.

Алгоритмы оценивания параметров (идентификация) в реальном масштабе времени, используемые в адаптивных системах управления, строятся, как правило, на основе гипотезы квазистационарности характеристик объекта управления, то есть $\dot{\theta} = 0$. В соответствии с этой гипотезой оцениваемые параметры или постоянны во времени, или изменяются с незначительной скоростью, пренебрежение которой практически не ухудшает оценок, получаемых на ограниченном временном интервале наблюдения.

Таким образом, можно записать дискретную модель изменения параметров в виде

$$\theta(k+1) = \theta(k), \quad \theta(0) = \bar{\theta}_0.$$

Измерительный комплекс

$$y(k+1) = Hx(k+1) + r(k+1)$$

является общим для состояния и параметров модели объекта. Действительно,

$$\begin{aligned} y(k+1) &= HA(k, \theta(k))x(k) + HB(k, \theta(k))u(k) + HF(k)q(k) + r(k+1) = \\ &= H\Phi(x(k), u(k))\theta(k) + Hf(x(k), u(k)) + \tilde{q}(k), \end{aligned}$$

где матрица $\Phi(\cdot)$ размерности $N \times n$ и вектор-столбец $f(\cdot)$ с n элементами получаются в результате представления системы (8.76) следующим образом:

$$x(k+1) = \Phi(x(k), u(k))\theta(k) + f(x(k), u(k)) + F(k)q(k).$$

Канал измерений вектора параметров можно приближенно представить в виде

$$y(k+1) = H\Phi(\hat{x}(k), u(k))\theta(k) + Hf(\hat{x}(k), u(k)) + \tilde{q}(k).$$

Будем строить оценку вектора параметров $\hat{\theta}(k+1)$ по текущему измерению $y(k+1)$ в виде

$$\hat{\theta}(k+1) = \hat{\theta}(k) + L(k)[y(k+1) - H\Phi(\hat{x}(k), u(k)) - Hf(\hat{x}(k), u(k))],$$

где $L(k)$ — коэффициент усиления фильтра, выражение для которого можно получить аналогично тому, как был получен коэффициент усиления $K(k)$ фильтра для оценки состояния $\hat{x}(k+1)$.

Таким образом, оценивание состояния объекта и идентификацию его параметров можно осуществлять с помощью двух параллельно работающих дискретных фильтров Калмана. Рекуррентный алгоритм для оценки состояния будет иметь следующий вид:

$$\left. \begin{aligned} \hat{x}(k+1) &= \hat{x}(k+1/k) + K(k)[y(k+1) - H\hat{x}(k+1/k)]; \\ \hat{x}(k+1/k) &= A(k, \hat{\theta}(k))\hat{x}(k) + B(k, \hat{\theta}(k))u(k) + \\ &\quad + F(k)\bar{q}(k), \quad \hat{x}(0) = \bar{x}_0; \\ K(k) &= P_x(k+1/k)H^T [HP_x(k+1/k)H^T + R]^{-1}; \\ P_x(k+1/k) &= A(k, \hat{\theta}(k))P_x(k)A^T(k, \hat{\theta}(k)) + F(k)QF^T(k); \\ P_x(k+1) &= [I_n - K(k)H]P_x(k+1/k); \\ P_x(0) &= P_{x_0}. \end{aligned} \right\}$$

Идентификацию параметров будем осуществлять с помощью следующего фильтра Калмана:

$$\left. \begin{aligned} \hat{\theta}(k+1) &= \hat{\theta}(k) + L(k)[y(k+1) - H\Phi(\hat{x}(k), u(k))\hat{\theta}(k) - \\ &\quad - Hf(\hat{x}(k), u(k))], \quad \hat{\theta}(0) = \bar{\theta}_0; \\ L(k) &= P_\theta(k)\Phi^T(\hat{x}(k), u(k))M^{-1}; \\ M(k) &= H\Phi(\hat{x}(k), u(k))P_\theta(k)\Phi^T(\hat{x}(k), u(k))H^T + \\ &\quad + HF(k)QF^T(k)H^T + R; \\ P_\theta(k+1) &= [I_N - L(k)H\Phi(\hat{x}(k), u(k))]P_\theta(k); \\ P_\theta(0) &= P_{\theta_0}. \end{aligned} \right\}$$

При постановке задачи идентификации параметров в большинстве случаев подразумевается, что приближение оценок параметров $\hat{\theta}$ к их истинным значениям θ обеспечивает приближение формируемого на основе оценок оптимального управления к искомому оптимальному управлению для реального объекта. Такое предположение подкрепляется теоремой разделения для предельной точности оценивания параметров и состояния объекта. Однако для случая ограниченной точности оценивания параметров эта теорема не дает конструктивных рекомендаций по организации совместного оценивания и оптимизации

управления. В то же время опыт моделирования процессов адаптивного управления показывает не очень сильную связь между точностью оценок $\hat{\theta}$ и качеством поведения управляемого объекта. Поэтому необходимо использовать следующий подход к текущей идентификации, выполняемой при адаптивном управлении, — идентификацию параметров осуществлять таким образом, чтобы в первую очередь достигался достаточно высокий уровень качества управления поведением объекта и в меньшей степени — точность идентификации параметров объекта как таковая.

Структурная схема синтеза адаптивного управления представлена на рис. 8.6.

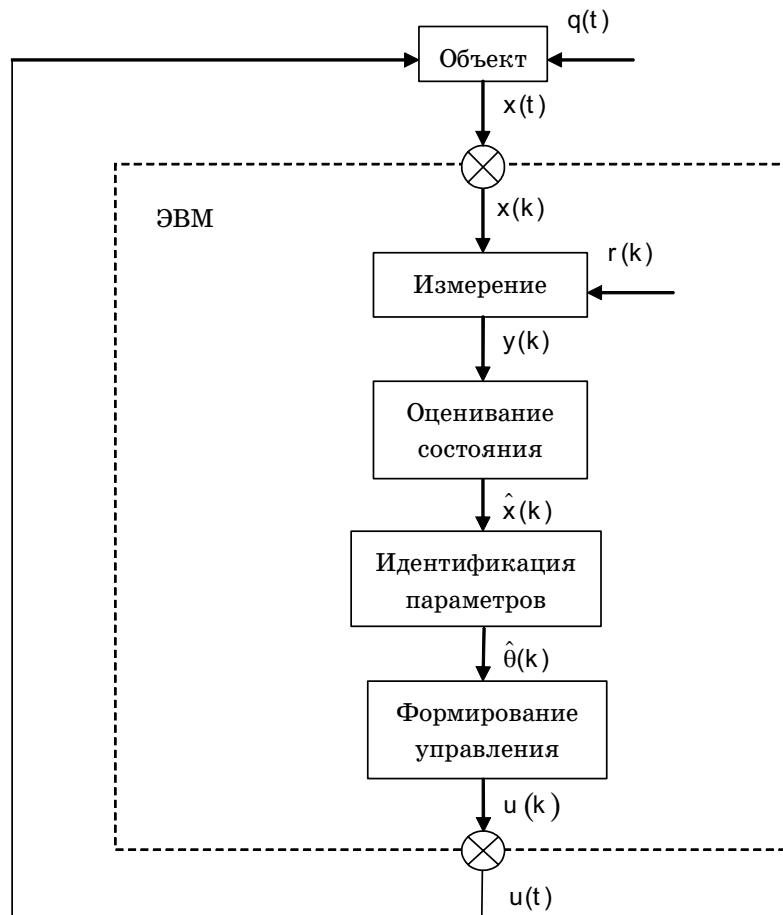


Рис. 8.6

8.6. Учет ограничений и запаздываний по управлению

При синтезе управлений на величину управляющего воздействия обычно накладываются ограничения. Достаточно часто эти ограничения накладываются в форме неравенств

$$U \min_i \leq u_i(k) \leq U \max_i, \quad i = \overline{1, m},$$

где $U \min_i$, $U \max_i$ — заданные величины. Например, для автопилота в самолете $U \min_i$, $U \max_i$ определены по условию безопасности полета.

Тогда, учитывая ограничения, будем полагать:

$$\bar{u}_i(k) = \begin{cases} U \min_i, & \text{если } u_i(k) < U \min_i; \\ u_i(k), & \text{если } U \min_i \leq u_i(k) \leq U \max_i; \\ U \max_i, & \text{если } u_i(k) > U \max_i. \end{cases}$$

Включение ЭВМ в контур управления неизбежно создает особые условия его реализации, одним из основных условий является запаздывание подачи управляющего сигнала на рулевые органы по отношению к моменту выдачи измерительной системой информации об объекте. Кроме того, запаздывание вызвано затратами времени на расчет оценок $\hat{x}(k)$ и $\hat{\theta}(k)$ и формирование управляющего сигнала $\bar{u}(k)$.

Управление с учетом запаздывания будем задавать уравнением

$$\bar{u}_i(k) = \bar{u}_i(k - l_z), \quad i = \overline{1, m},$$

где $l_z \Delta t$ — время запаздывания.

8.7. Общая схема синтеза адаптивных систем управления

Проектирование любой системы управления должно осуществляться с учетом конкретных условий ее применения. В данном случае рассматривается проектирование системы адаптивного управления для совмещенного синтеза, т.е. для ситуации, когда управление формируется в процессе функционирования объекта. Здесь большое значение приобретают алгоритмы и способы, позволяющие уменьшить вычислительную задержку, вызванную затратами времени на формирование управления. Один из таких способов — введение дискретности решения задач оценивания состояния, идентификации параметров и формирования управляющих воздействий, т.е. $k_{\text{оц}}, k_{\text{ид}}, k_{\text{упр}}$. При этом $k_{\text{оц}} \geq k$, $k_{\text{ид}} \geq k$, $k_{\text{упр}} \geq k$ и $k_{\text{оц}} \geq k_{\text{упр}}$, $k_{\text{ид}} \geq k_{\text{упр}}$. Тогда общая схема синтеза адаптивного управления будет иметь следующий вид (рис. 8.7). При этом предполагается, что моменты и периоды квантования одина-

ковы для всех координат объекта и система управления формирует управляющее воздействие с моментом и периодом квантования, которые совпадают с моментом и периодом поступления информации об объекте. Кроме того, учет запаздывания осуществляется на время, кратное Δt .

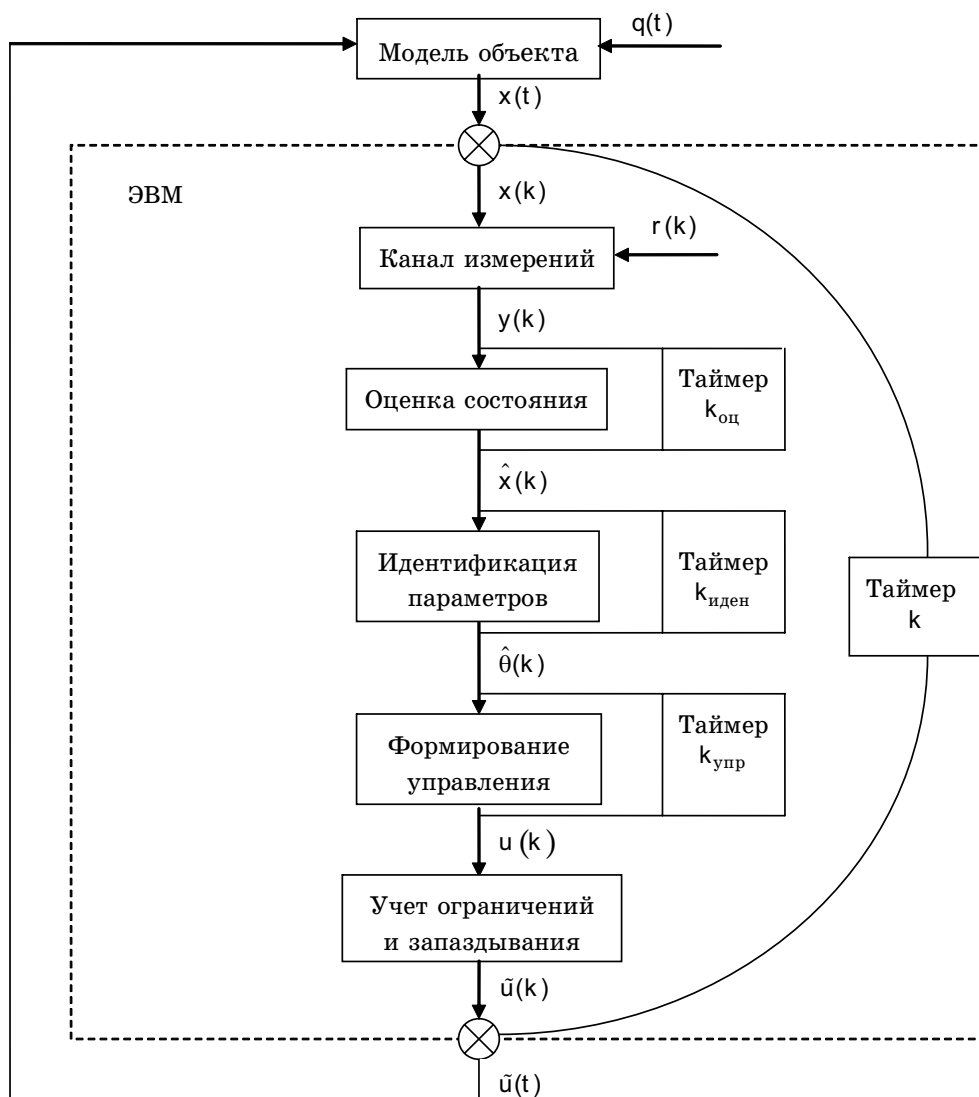


Рис. 8.7

Рассмотренные алгоритмы синтеза адаптивного управления все еще требуют достаточно высокой производительности управляющих ЭВМ,

что ограничивает их применение в реальном времени функционирования объекта. Известны два подхода к преодолению этих трудностей.

Первый подход связан с построением для системы, описывающей движение объекта, системы пониженного порядка и использованием *централизованной* структуры управления. При этом исходная система заменяется при синтезе другой, меньшего порядка, которая отражает лишь некоторые существенные особенности исходной системы. Применение такой схемы управления приводит к субоптимальному функционированию реального объекта и является эффективным для слабосвязанных систем, или систем, имеющих быстрые и медленные составляющие.

Второй подход связан с разбиением системы управления на ряд подсистем, каждая из которых имеет меньшую размерность. Это приводит к построению *распределенной децентрализованной* или *иерархической децентрализованной* структуры управления. В последнем случае используются обычно два основных вида структур управления, основанных на многослойной и многоуровневой декомпозиции. При таком подходе трудоемким является построение структуры децентрализованного управления, что требует достаточно сложной программной реализации. Следует отметить, что существенное сокращение времени синтеза достигается, в основном, при параллельных вычислениях.

9. ПРИМЕРЫ ПОСТРОЕНИЯ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

К объектам управления, поведение которых описывается обыкновенными дифференциальными уравнениями, относятся все механические, электромеханические объекты, а также многие технологические процессы, подчиняющиеся законам механики и электричества. Кроме того, такими уравнениями можно описать и многие макро- и микроэкономические процессы.

При проектировании систем управления одной из первых и основных задач является построение математических моделей. Качество функционирования системы управления существенно зависит от того, насколько адекватной является построенная модель.

9.1. Построение математических моделей прямолинейного и вращательного движения

Основной закон механики заключается в том, что движение твердого тела состоит из движения его центра масс, которое подчиняется второму закону Ньютона:

$$m\ddot{x} = F, \quad (9.1)$$

и движения тела вокруг этого центра масс, которое описывается уравнением

$$J\ddot{\varphi} = M, \quad (9.2)$$

где m — масса тела; J — момент инерции; F — силы, действующие на тело и приведенные к центру масс; M — момент сил; x — координаты центра масс; φ — углы положения тела.

9.1.1. Математическая модель прямолинейного движения

Примером прямолинейного движения может служить механическое колебательное звено (рис. 9.1).

Пусть тело G движется вдоль оси координат OX . На тело G действуют следующие силы:

1) вес

$$F_g = mg,$$

где m — масса тела; g — ускорение силы тяжести;

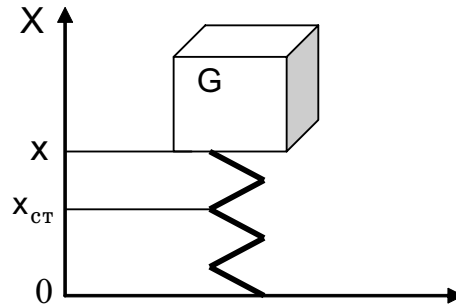


Рис. 9.1

2) реакция пружины, которая пропорциональна отклонению размеров пружины от ненагруженного состояния:

$$F_{\text{пр}} = -k_1(x - x_0),$$

где x_0 — длина ненагруженной пружины; k_1 — коэффициент упругости, $k_1 > 0$;

3) реакция амортизатора, которая пропорциональна скорости движения штока:

$$F_{\text{амор}} = -k_2\dot{x},$$

где k_2 — коэффициент, зависящий от вязкости жидкости, залитой в амортизатор, $k_2 > 0$.

Результирующая всех сил равна

$$F = -mg - k_1(x - x_0) - k_2\dot{x}. \quad (9.3)$$

Если ввести обозначения $x_1 = x$, $x_2 = \dot{x}$, то уравнение (9.1) с учетом (9.3) можно записать в виде системы уравнений, описывающей динамическую модель рассматриваемого процесса:

$$\left. \begin{aligned} \dot{x}_1 &= x_2; \\ \dot{x}_2 &= -g - \frac{k_1}{m}(x_1 - x_0) - \frac{k_2}{m}x_2. \end{aligned} \right\} \quad (9.4)$$

Стационарное решение системы можно найти, если приравнять правые части нулю. В результате получается *стационарная модель*:

$$x_{1\text{ст}} = x_0 - \frac{mg}{k_1}; \quad x_{2\text{ст}} = 0.$$

Очевидно, что значение $x_{1\text{ст}}$ должно быть положительным. Поэтому коэффициент упругости k_1 должен быть больше значения $\frac{mg}{x_0}$.

Если ввести разность $\Delta x_1 = x_1 - x_{1\text{ст}}$, которая характеризует отклонение координаты $x(t)$ от стационарного значения $x_{1\text{ст}}$, то согласно (9.4) эта разность удовлетворяет системе уравнений

$$\left. \begin{aligned} \Delta \dot{x}_1 &= \Delta x_2; \\ \Delta \dot{x}_2 &= -\frac{k_1}{m} \Delta x_1 - \frac{k_2}{m} \Delta x_2 \end{aligned} \right\} \quad (9.5)$$

Система (9.5) является моделью прямолинейного движения в отклонениях.

9.1.2. Математическая модель вращательного движения

Многие технологические процессы связаны с вращением механических систем. Довольно часто задача сводится к необходимости подавать электрический ток нужной силы на вход электродвигателя, на валу которого насажена некоторая масса m . При этом необходимо задавать такую силу тока, чтобы обеспечить заданное вращательное движение массы. Примерами таких ситуаций может быть суточное слежение наземных антенн за спутниками связи, равномерное движение наматывающих барабанов в бумагоделательных машинах и т.п.

Математически вращательное движение массы описывается уравнением (9.2). Если обозначить через x_1 угол φ , а через x_2 — угловую скорость $\omega = \dot{\varphi}$, то это уравнение можно переписать в виде

$$\dot{x}_1 = x_2; \quad \dot{x}_2 = \frac{M}{J}.$$

Момент сил M складывается из следующих составляющих:

- 1) $k_1 u$ — сила вращения, создаваемая электродвигателем, на вход которого поступает ток силой u ;
- 2) $k_2 x_2$ — сила сухого трения, пропорциональная скорости вращения и обратная ей по знаку;
- 3) $f(t)$ — сила внешних нагрузок.

В результате уравнения вращательного движения, описывающие динамическую модель рассматриваемого процесса, примут вид

$$\left. \begin{aligned} \dot{x}_1 &= x_2; \\ \dot{x}_2 &= -\frac{k_2}{J} x_2 + \frac{k_1}{J} u + \frac{f(t)}{J} \end{aligned} \right\}$$

Здесь k_1 и k_2 — некоторые положительные коэффициенты.

9.2. Макроэкономическая модель динамики фондов производственного накопления и потребления

Пусть $x_1(t)$ — фонд производственного накопления; $x_2(t)$ — фонд потребления.

Уравнение баланса имеет вид

$$\dot{x}_1(t) = b_1 \dot{x}_1(t) + b_2 \dot{x}_2(t),$$

где b_1 и b_2 — коэффициенты приростной капиталоемкости фондов накопления и потребления соответственно.

Пусть $L(t) = L_0 e^{\gamma t}$ — динамика изменения численности населения; L_0 — численность населения в начальный момент времени; γ — темп роста населения.

Тогда уравнения, характеризующие фонд потребления на душу населения и изменение душевого фонда потребления, будут соответственно равны

$$\frac{x_2(t)}{L(t)} = \frac{x_2(t)}{L_0} e^{-\gamma t};$$

$$\frac{d}{dt} \left[\frac{x_2(t)}{L(t)} e^{-\gamma t} \right] = (\dot{x}_2(t) - \gamma x_2(t)) \frac{e^{-\gamma t}}{L_0}.$$

Если в качестве управления использовать скорость роста душевого фонда потребления

$$u(t) = \dot{x}_2(t) - \gamma x_2(t),$$

то

$$\dot{x}_2(t) = \gamma x_2(t) + u(t). \quad (9.6)$$

Подставляя (9.6) в (9.1), получим

$$x_1(t) = b_1 \dot{x}_1(t) + b_2 (\gamma x_2(t) + u(t)).$$

Таким образом, динамическая модель изменения фондов накопления и потребления описывается системой обыкновенных дифференциальных уравнений вида

$$\left. \begin{aligned} \dot{x}_1(t) &= \frac{1}{b_1} x_1(t) - \gamma \frac{b_2}{b_1} x_2(t) - \frac{b_2}{b_1} u(t), \quad x_1(t_0) = x_1^{(0)}; \\ \dot{x}_2(t) &= \gamma x_2(t) + u(t), \quad x_2(t_0) = x_2^{(0)}. \end{aligned} \right\} \quad (9.7)$$

Заметим, что модель (9.7) является линейной по состоянию и управлению.

Если обозначить $x(t) = (x_1(t), x_2(t))^T$; $x^{(0)} = (x_1^{(0)}, x_2^{(0)})^T$, то система (9.7) в матричной форме запишется следующим образом:

$$\dot{x}(t) = \bar{A} x(t) + \bar{B} u(t), \quad x(t_0) = x^{(0)},$$

где

$$\bar{A} = \begin{pmatrix} \frac{1}{b_1} & -\gamma \frac{b_2}{b_1} \\ 0 & \gamma \end{pmatrix}; \quad \bar{B} = \begin{pmatrix} -\frac{b_2}{b_1} \\ 1 \end{pmatrix}.$$

9.3. Построение математических моделей производства, хранения и сбыта товара повседневного спроса

Введем следующие обозначения:

X — общее количество единиц товара, выпущенных предприятием;

u — темп производства;

Z — количество товара на рынке;

V — количество товара у потребителя;

W — доход;

Y — потенциальный спрос (полное количество товара, способное мгновенно удовлетворить спрос в условиях отсутствия ажиотажного спроса);

c — коэффициент, характеризующий превышение цены над себестоимостью ($c > 1$, так как себестоимость считается равной единице);

k_1 — коэффициент, характеризующий темп потребления;

k_2 — коэффициент, характеризующий плату за хранение товара;

$n(c)$ — коэффициент, характеризующий скорость продажи товара, например $n(c) = n_0 e^{1-c}$ отражает распределение по достатку.

Параметры Y , k_1 , k_2 считаются постоянными. Величины u и c назначаются производителем и могут быть как постоянными, так и переменными. Переменные X , Z , V , W — функции времени, которые могут изменяться в соответствии с законами рынка.

Статическая модель описывает один цикл производства и сбыта: выработанный за расчетную единицу времени товар поступает на рынок и продается, в результате чего получается определенный доход. Статическая модель строится, исходя из предположения, что в равновесном состоянии система характеризуется двумя соотношениями баланса:

1) поступившее на рынок количество товара равно проданному;

2) сколько товара продано, столько и потреблено.

Эти соотношения с учетом введенных обозначений запишутся следующим образом: за единицу времени прирост товара на рынке составляет u единиц; количество товара на рынке — Z единиц; у потребителя — V единиц; в единицу времени продается $n(c)(Y-V)Z$ единиц товара, где $(Y-V)$ — текущий спрос. При этом скорость продажи пропорциональна спросу и количеству товара на рынке. Тогда статическая модель процесса производства, хранения и сбыта товара запишется в виде

$$\left. \begin{aligned} u - n(c) (Y - V) Z &= 0; \\ n(c) (Y - V) Z - k_1 V &= 0; \\ W &= cn(c) (Y - V) Z - \frac{u}{c} - k_2 Z. \end{aligned} \right\} \quad (9.8)$$

Последнее уравнение показывает, что доход в единицу времени складывается из выручки от продаж $n(c)(Y-V)Z$ единиц товара, расходов на производство и затрат на хранение.

Динамическая модель строится следующим образом. Так как за время dt прирост произведенного товара равен $u(t)dt$, а продается $n(c)(Y-V(t))Z(t)dt$ единиц товара, то скорость прироста товара на рынке равна

$$\frac{dZ(t)}{dt} = u(t) - n(c)(Y - V(t))Z(t).$$

Прирост товара у потребителя за время dt равен количеству проданного товара минус количество потребленного, т.е. скорость прироста товара у потребителя описывается уравнением

$$\frac{dV(t)}{dt} = n(c)(Y - V(t))Z(t) - k_1 V(t).$$

Таким образом, динамическая модель, описывающая процесс производства, хранения и сбыта, имеет вид

$$\left. \begin{aligned} \frac{dX(t)}{dt} &= u(t); \\ \frac{dZ(t)}{dt} &= u(t) - n(c)(Y - V(t))Z(t); \\ \frac{dV(t)}{dt} &= n(c)(Y - V(t))Z(t) - k_1 V(t); \\ \frac{dW(t)}{dt} &= cn(c)(Y - V(t))Z(t) - \frac{u(t)}{c} - k_2 Z(t). \end{aligned} \right\} \quad (9.9)$$

Заметим, что статическая модель (9.8) может быть получена из динамической (9.9) путем приравнивания нулю правых частей в (9.9). Поэтому статическая модель может рассматриваться как частный случай динамической, отражающей состояние динамического равновесия.

Перепишем динамическую модель (9.9) в терминах теории автоматического управления.

Пусть $x(t) = (x_1(t), x_2(t), x_3(t), x_4(t))^T$ — вектор состояния, где $x_1(t) = X$, $x_2(t) = Z$, $x_3(t) = V$, $x_4(t) = W$; $u(t)$ — управление, задающее темп производства. Тогда система (9.9) запишется следующим образом:

$$\dot{x}(t) = f(t, x(t), u(t)), \quad (9.10)$$

где $f(t, x(t), u(t))$ — вектор-функция вида

$$f(x(t), u(t)) = \begin{pmatrix} u(t) \\ u(t) - n(c)Yx_2(t) + n(c)x_2(t)x_3(t) \\ n(c)Yx_2(t) - n(c)x_2(t)x_3(t) - k_1x_3(t) \\ cn(c)Yx_2(t) - cn(c)x_2(t)x_3(t) - \frac{u(t)}{c} - k_2x_2(t) \end{pmatrix}.$$

Система (9.10) является *нелинейной нестационарной детерминированной моделью* процесса производства, сбыта и хранения товара.

Для того чтобы получить *линейную* систему, (9.10) запишем в виде

$$\dot{x}(t) = \left. \frac{\partial f(\cdot)}{\partial x} \right|_{x=x^H(t)} x(t) + \left. \frac{\partial f(\cdot)}{\partial u} \right|_{u=u^H(t)} u(t)$$

или

$$\dot{x}(t) = \bar{A}(t)x(t) + \bar{B}(t)u(t), \quad (9.11)$$

где $x^H(t)$, $u^H(t)$ — номинальные траектория и управление (например, характеристики, соответствующие плановым, расчетным, прогнозируемым показателям).

Учитывая смысл второго и третьего компонентов вектора состояния в рассматриваемой модели, построим функции $\varphi_Z(t)$ и $\varphi_V(t)$ с помощью которых моделируется ситуация на рынке и поведение покупателей. Кроме того, построим функцию $\varphi_Y(t)$, которая характеризует потенциальный спрос. Тогда матрицы динамики и влияния управляющих воздействий будут иметь вид

$$\bar{A}(t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -n(c)\varphi_Y(t) + n(c)\varphi_V(t) & n(c)\varphi_Z(t) & 0 \\ 0 & n(c)\varphi_Y(t) - n(c)\varphi_V(t) & -n(c)\varphi_Z(t) - k_1 & 0 \\ 0 & cn(c)\varphi_Y(t) - cn(c)\varphi_V(t) - k_2 & -cn(c)\varphi_Z(t) & 0 \end{pmatrix};$$

$$\bar{B}(t) = \begin{pmatrix} 1 & 1 & 0 & -\frac{1}{c} \end{pmatrix}^T.$$

Для учета колебаний рынка, поведения покупателей, погрешности линеаризации и так далее добавим в систему (9.11) вектор гауссовских шумов $q(t)$ с матрицей влияния $\bar{F}(t)$, компоненты которой, например, можно определить для конкретного предприятия экспериментально. Тогда математическая модель, описывающая процесс производства, сбыта и хранения товара, запишется в виде системы *линейных нестационарных стохастических дифференциальных уравнений*

$$\dot{x}(t) = \bar{A}(t)x(t) + \bar{B}(t)u(t) + \bar{F}(t)q(t), \quad (9.12)$$

где $M \{q(t)\} = \bar{q}(t)$; $M \left\{ (q(t) - \bar{q}(t)) (q(t) - \bar{q}(t))^T \right\} = Q(t) \delta(t - \tau)$;
 $\delta(t - \tau)$ — δ -функция Дирака.

В терминах теории автоматического управления система (9.12) сформулирована для решения задачи управления положением рулей, в данном случае темпом производства. Можно решать задачу управления скоростью отклонения рулей. В данном случае это будет задача управления изменением темпа производства, математическая модель которой может быть описана следующей системой линейных нестационарных стохастических дифференциальных уравнений:

$$\left. \begin{aligned} \dot{x}(t) &= \bar{A}(t) x(t) + \bar{B}(t) u(t) + \bar{F}(t) q(t); \\ \dot{u}(t) &= v(t), \end{aligned} \right\} \quad (9.13)$$

где $x(t) = (Z, V, W)^T \in R^3$ — вектор состояния; $u(t)$ — темп производства; $q(t)$ — вектор внешних возмущений; $v(t)$ — управление, задающее изменение темпа производства;

$$\bar{A}(t) = \begin{pmatrix} -n(c)\varphi_Y(t) + n(c)\varphi_V(t) & n(c)\varphi_Z(t) & 0 \\ n(c)\varphi_Y(t) - n(c)\varphi_V(t) & -n(c)\varphi_Z(t) - k_1 & 0 \\ cn(c)\varphi_Y(t) - cn(c)\varphi_V(t) - k_2 & -cn(c)\varphi_Z(t) & 0 \end{pmatrix};$$

$$\bar{B}(t) = \begin{pmatrix} 1 & 0 & -\frac{1}{c} \end{pmatrix}^T;$$

$\bar{F}(t)$ — матрица соответствующей размерности, описывающая влияние случайных факторов в модели (9.13).

9.4. Построение математической модели вертикального движения ракеты

Пусть ракета движется вертикально в поле тяготения какой-то планеты. На основании законов механики можно записать следующие уравнения движения:

$$m\ddot{x} = F; \quad \dot{m} = -c, \quad (9.14)$$

где m — масса ракеты; x — высота над поверхностью планеты; c — секундный расход массы за счет горения топлива; F — вертикальная составляющая сил, действующих на ракету. Обозначим

$$x_1 = x; \quad x_2 = \dot{x} = v; \quad x_3 = m,$$

где v — скорость ракеты. Пусть оси Ox_1 и Ox_2 направлены вверх от поверхности планеты. Управление ракетой осуществляется за счет тяги двигателя, что связано с расходом топлива c , т.е. $u = c$.

С учетом введенных обозначений уравнения (9.14) можно переписать

сать следующим образом:

$$\dot{x}_1 = x_2; \quad \dot{x}_2 = \frac{F(x, u)}{x_3}; \quad \dot{x}_3 = -u.$$

Функция $F(x, u)$ имеет вид

$$F(x, u) = -G + P(u) - Q \operatorname{sign}(v),$$

где $G = mg$ — сила тяжести (g — ускорение свободного падения); $P(u)$ — тяга двигателя, которая с точностью до знака есть известная функция u ; $\operatorname{sign}(v)$ — знаковая функция, равная плюс 1, если $v \geq 0$, и минус 1, если $v < 0$ (скорость v отрицательна, если ракета опускается); Q — аэродинамическое сопротивление, равное

$$Q = \rho(x_1) \frac{v^2}{2} S C_x.$$

Здесь $\rho(x_1)$ — плотность атмосферы на высоте x_1 ; S — поперечное сечение ракеты; C_x — аэродинамический коэффициент. Последние две величины можно считать постоянными. С учетом введенных обозначений выражение для функции $F(x, u)$ можно переписать в виде

$$F(x, u) = -gx_3 + P(u) - \rho(x_1) \frac{x_2^2}{2} S C_x \operatorname{sign}(x_2).$$

Подставляя это выражение в (9.14), получим математическую модель движения ракеты:

$$\left. \begin{aligned} \dot{x}_1 &= x_2; \\ \dot{x}_2 &= \frac{F(x, u)}{x_3} - g + \frac{P(u)}{x_3} - \rho(x_1) \frac{x_2^2}{2x_3} S C_x \operatorname{sign}(x_2); \\ \dot{x}_3 &= -u. \end{aligned} \right\} \quad (9.15)$$

Начальные условия для системы (9.15) зависят от постановки задачи управления.

Если решается задача взлета ракеты, то начальные условия будут следующими:

$$x_1(t_0) = 0; \quad x_2(t_0) = 0; \quad x_3(t_0) = m_0,$$

где t_0 — момент взлета; $m_0 = m_p + m_T$, m_p — масса тела ракеты; m_T — начальная масса топлива.

Если решается задача посадки ракеты, то начальные условия будут следующими:

$$x_1(t_0) = h_0; \quad x_2(t_0) = v_0; \quad x_3(t_0) = m_0,$$

где h_0 , v_0 — расстояние до планеты и скорость ракеты в момент начала посадки; $m_0 = m_p + m_o$, m_o — остаток топлива в момент начала посадки.

10. ПРОГРАММНЫЕ СРЕДСТВА МОДЕЛИРОВАНИЯ

Выделим следующие основные требования к программным системам, предназначенным для моделирования:

- наличие математических функций;
- точность вычислений;
- обработка ошибок;
- возможность пополнения библиотек;
- система отображения;
- система разработки интерфейса;
- возможности коммуникации с другими программами.

Учитывая вышеперечисленные требования, рассмотрим различные системы моделирования — как специализированные системы, предназначенные для математических расчетов, так и универсальные, широко применяемые инструментальные системы. Рассмотрение программных средств моделирования начнем именно с универсальных систем.

Прежде чем перейти к описанию систем программирования, построенных на базе языков C/C++, следует понять разницу между языком программирования и средой разработки. Язык — это некоторый формальный способ описания задачи, решаемой машиной. Транслятор языка — это программа, которая переводит текст задачи с заданного языка на язык машины. А система программирования — это ряд программ и системных библиотек, которые призваны облегчить программисту разработку программ. В состав систем программирования, помимо транслятора языка (иногда даже нескольких), обычно включается интегрированная среда разработки, различные системные библиотеки и т.д. Например, есть язык C++, есть множество различных реализаций компилятора этого языка, в том числе реализация от фирмы Microsoft, и есть система программирования Microsoft Visual C++, которая строится на этом компиляторе и большом количестве дополнительных библиотек и инструментальных средств.

10.1. Языки программирования

10.1.1. Язык Fortran

Язык Fortran (от англ. Formula Translation) — один из первых

языков программирования, появившийся еще в 1950-х годах. В то время было разработано множество библиотек, в том числе математических, ориентированных на этот язык. Несмотря на то, что язык оказался не слишком удобным в использовании, он довольно прочно занял свою нишу для разработки математических программ, в первую очередь для моделирования. С тех пор язык претерпел несколько редакций (последняя — в 1995-м году), и существует большое множество различных версий компилятора этого языка. Среди современных реализаций для персональных компьютеров следует отметить Microsoft Visual Fortran.

Как следует из названия языка, основное его предназначение — перевод формул. И с этой работой он справляется достаточно успешно. К недостаткам этого языка следует отнести слабые возможности по описанию абстрактных типов данных, неудобную организацию циклов и других языковых конструкций.

10.1.2. Язык C

Язык C, появившийся в первой половине 1970-х годов в лаборатории Bell Labs компании AT&T (Керниган&Ритчи), является в настоящее время де-факто стандартом для разработки программ системного назначения малого и среднего размеров. В настоящее время реализация этого языка есть практически для всех вычислительных систем от микроконтроллеров до сверхбольших вычислительных машин. Понятие «среднего размера» весьма относительно и, как правило, соответствует проектам от 10 до 100 тысяч строк исходного текста. Под «системными» понимаются программы, предназначенные, в основном, для нужд операционной системы, в том числе и сама операционная система является системной программой. Однако этот язык вполне подходит и для разработки прикладных программ, связанных с математическими вычислениями. Изюминка языка C — в его относительной близости к вычислительной системе, что позволяет разрабатывать высокоэффективный код (с минимальными накладными расходами, а соответственно, максимальный по скорости выполнения и минимальный по расходу памяти).

Язык C обладает набором из 50 операций (как целочисленных, так и вещественных), распределённых по 15 уровням приоритетов. Среди математических операций языка собраны все операции, которые могут выполнить центральный процессор и математический сопроцессор. Часть операций применима как к целым типам данных, так и к вещественным.

Язык не обладает встроенным понятием модуля, но, как правило, среда разработки (или по-другому — окружение языка) позволяет ввести это понятие искусственно и использовать как большое количество

стандартных библиотек, так и свои собственные разработанные библиотеки. Более того, язык не обладает собственными средствами ввода/вывода, все эти средства реализуются внешними библиотеками, что позволяет разрабатывать более гибкий исходный код программ, применимый для различного операционного окружения языка и исполняемой программы.

Недостатки языка являются обратной стороной его достоинств. К ним можно отнести отсутствие встроенных типов данных высокой степени абстракции (например, «строка», часто примененного для организации диалога с пользователем), отсутствие средств диагностики (контроль выхода за пределы массива, переполнения стека и числа и т.д.). Такой контроль можно обеспечить, лишь связавшись на низком уровне с операционной системой, при условии, что операционная система поддерживает такой контроль.

Таким образом язык *C* хорошо подходит для разработки низкоуровневых системных библиотек и для математического моделирования.

10.1.3. Язык C++

Язык C++, разработанный в первой половине 1980-х годов как объектно-ориентированное (ОО) развитие языка *C* Бьярном Страуструпом, занимает несколько иную нишу. В отличие от языка *C*, программы, скомпилированные с языка C++, обладают, как правило, немного худшей производительностью и большим объемом исполняемого кода. Основная ниша данного языка — разработка больших (обычно прикладных) программных систем сравнительно большой группой разработчиков, распределённой в пространстве и/или времени. Основное преимущество данного языка — языковая поддержка трех базисных конструкций ОО программирования: инкапсуляция, полиморфизм и наследование.

Под *инкапсуляцией* понимается разделение реализации объекта (модуля) и его интерфейса. Как правило, скрываются данные, а в интерфейсе остаются только методы их обработки. Поэтому инкапсуляцию иногда трактуют как объединение данных и методов их обработки в единой языковой конструкции. При решении задач моделирования это позволяет создавать типы данных высокой степени абстракции. Такие, например, как вектора, матрицы, генераторы, фильтры и т.д. Фактически, любое понятие предметной области (ПО) можно рассматривать как объект с точки зрения ОО программирования и строить соответствующую ему модель с детерминированным набором свойств и поведением. В языке C++ для поддержки инкапсуляции вводят понятия «объект» и «класс», который представляет собой обобщенное описание объекта.

Язык C++ позволяет, используя еще одну языковую конструкцию —

наследование, создавать обобщенные (базовые) классы, описывая в них общую составляющую всех дочерних классов. Это упрощает процесс доработки программ, который теперь можно не сводить к переделке всего разработанного ранее класса. Для изменения набора свойств и/или расширения функциональности можно описать новый класс, наследованный от созданного ранее.

Еще одной принципиальной с точки зрения ОО программирования возможностью языка является *полиморфизм*. Это понятие определяется как возможность указания разной семантической нагрузки для конструкций с одинаковым синтаксисом. Например, пусть есть классы «вектор» и «матрица». Оба этих класса имеют метод «абсолютное значение» со схожим синтаксисом (на входе объект, на выходе вещественное число), но принципиально различной внутренней реализацией этих методов. Тем не менее вполне вероятно, что некоторая функция может работать с любым типом данных, если тот обладает методом вычисления абсолютного значения. Можно разработать такой метод для всех типов данных, с которыми будет работать программа, но при этом может возникнуть большая степень дублированности кода. С другой стороны, можно разработать некоторый базовый класс, в котором перечислены основные свойства и методы большой группы классов. При этом часть методов могут быть описаны как виртуальные, что позволяет давать этим методам разную реализацию в дочерних классах.

Еще одной возможностью языка C++ является возможность описания шаблонов функций и классов — так называемое параметризованное описание. Параметром такого описания является, как правило, тип или набор типов, с которым работает данная функция или класс. Это позволяет создавать более высокоэффективные (по сравнению с наследованием) программы.

Кроме того, язык C++, один из немногих, позволяет перегружать смысл своих операций для абстрактных типов данных. Таким образом, можно получить набор типов (классов), максимально приближенный к аналогичным объектам в математике, например класс «матрица» или «вектор». Совместное использование всех языковых конструкций позволило создать весьма эффективную, но в то же самое время очень гибкую стандартную библиотеку шаблонов (*standard template library* — STL), включающую все основные способы хранения данных и базовые алгоритмы обработки этих данных. Существует построенная на этой библиотеке, например, библиотека MTL (*matrix template library* — библиотека шаблонов матрицы), которая позволяет достаточно эффективно обрабатывать матрицы произвольной размерности с произвольным типом хранимых данных.

К сожалению, богатая история развития языка C++ привела к появлению огромного количества разночтений и, фактически, различных диалектов этого языка. Усилия консорциума разработчиков при-

вели к тому, что в 1998 году был утвержден стандарт этого языка. Однако далеко не все существующие компиляторы языка C++ совместимы с этим стандартом.

10.1.4. Языки Pascal и Object Pascal

Процедурный язык Pascal был разработан Николаусом Виртом исключительно как лабораторный язык. Этот язык отлично подходил для обучения студентов программированию, но не предполагался использоваться, как серьезный язык для разработки прикладных программ. Однако некоторые фирмы (например, Borland) смогли сделать из этого языка серьезную систему разработки. Позже Николаус Вирт разработал языки Modula, Modula-2, Modula-3 и язык Oberon. Лучшие идеи, реализованные в этих языках, были использованы в коммерческих разработках языка Pascal.

Стандартных операций в языке Pascal несколько меньше, чем, например, в языке C, но их вполне достаточно для того, чтобы покрыть все основные потребности при разработке математических программ. В коммерческих реализациях Pascal присутствует понятие модуля, что позволяет разрабатывать различные библиотеки.

Позже был разработан язык Object Pascal, который добавил к языку Pascal возможности объектно-ориентированности. Объектная модель языка Object Pascal беднее, чем C++, например, отсутствует возможность перегрузки смысла операций и нет шаблонов, что несколько уменьшает применимость данного языка.

10.2. Системы разработки программного обеспечения

10.2.1. Microsoft Visual C++ и MFC

Microsoft Visual C++ является составной частью пакета Microsoft Visual Studio. Как уже было сказано, система разработки ПО Microsoft Visual C++ базируется на компиляторе языка C++ от фирмы Microsoft, который, к сожалению, очень далек от стандарта языка. Тем не менее Visual C++ является де-факто стандартом при разработке программного обеспечения для ОС Windows, в том числе и программ моделирования. Ограниченность персональных компьютеров не позволяет решать на них серьезные задачи моделирования, что, впрочем, не умаляет их роли и роли системы Visual C++ для разработки программ-интерфейсов.

Именно для разработки интерфейсов с пользователем эта система программирования предоставляет максимальные возможности. Можно назвать богатую по возможностям библиотеку MFC, есть средства дос-

тупа к базам данных ADO, средства по созданию и использованию ActiveX компонентов — ATL и многое другое.

Однако есть также и масса недостатков. Главный недостаток — неполная совместимость компилятора со стандартом языка. Наибольшие отличия наблюдаются в реализации шаблонов, что не позволяет эффективно использовать STL, MTL и ряд других известных шаблонных библиотек.

10.2.2. Borland Delphi

В основе этой системы программирования лежит сильно модифицированный фирмой Borland язык Object Pascal. Язык изменен настолько сильно, что правильнее его называть языком Delphi. Первая версия этой среды разработки была ориентирована на ОС Windows 3.x и в то время стала прорывом в области разработок программ с интерфейсом пользователя.

Основная ниша этой среды разработки — разработка клиентских программ для баз данных. Для этого в системе разработки программ предусмотрены богатые возможности по разработке интерфейса и несколько различных вариантов доступа к базам данных.

10.2.3. Borland C++Builder

Эта система разработки распространена значительно меньше предыдущей, однако она имеет ряд принципиальных преимуществ, которые позволяют о ней говорить как о реально используемой платформе. В частности, компилятор C++, на базе которого построена эта система, очень близок к стандарту языка. Вообще говоря, начиная с 5-й версии, Borland C++Builder получил сертификат соответствия стандарту, хотя на самом деле небольшие отличия есть. Все специализированные средства языка, которых нет в стандарте, являются расширениями и не противоречат стандарту.

В основу положена та же библиотека, которая лежит в основе системы Delphi, к возможностям которой добавляются возможности, присущие стандарту языка C++.

К недостаткам следует отнести сравнительно большой объем исполняемого кода и несколько худшие по сравнению с Visual C++ скоростные показатели для программ графического интерфейса.

10.3. Специализированные математические пакеты

В 1980-х годах возникло и получило бурное развитие новое фунда-

ментальное научное направление — *компьютерная математика* — на стыке математики и информатики. Первыми серьезными средствами для автоматизированного выполнения массовых научно-технических расчетов были программируемые микрокалькуляторы. Предвестниками и систем компьютерной математики стали специализированные программы для математических численных расчетов, работающие в среде Microsoft MS-DOS. Это Eureka, Mercury, первые версии системы MathCAD и MatLAB под операционную систему MS-DOS. Вслед за этим на основе достижений компьютерной математики появились новейшие программы символьной математики, или компьютерной алгебры. Среди них особенно большую известность получили системы MathCAD под Windows, Derive, Mathematica и Maple. Созданные для проведения символьных (аналитических) преобразований математических выражений, эти системы были в поразительно короткое время доведены до уровня, позволяющего резко облегчить, а подчас и заменить труд самой почтенной научной элиты мира — математиков-теоретиков и аналитиков.

В последнее время появившиеся тенденции к объединению передовых разработчиков математических систем привели к тому, что лидеры в данной области, такие, как Maple, MathCAD, Mathematica, MatLAB, интегрировались и теперь стоят приблизительно на одной ступени развития. Различия между возможностями этих систем сохранились лишь в каких-то базовых элементах и визуализации программных документов.

10.3.1. Система Maple

Системы компьютерной математики класса Maple были созданы корпорацией Waterloo Maple как система компьютерной алгебры с расширенными возможностями в области символьных вычислений. Система содержит средства для выполнения быстрых численных расчетов, лежащих в основе математического моделирования различных явлений окружающего мира, систем и устройств различного назначения. Все это сочетается с новейшими и весьма эффективными средствами визуализации вычислений.

Maple — типичная интегрированная система. Она объединяет в себе мощный язык программирования, редактор для подготовки и редактирования документов и программ, многооконный пользовательский интерфейс с возможностью работы в диалоговом режиме, справочную систему со множеством примеров, ядро алгоритмов и правил преобразования математических выражений, численный и символьный процессоры, систему диагностики, библиотеки встроенных и дополнительных функций, пакеты функций сторонних производителей и поддержку

некоторых других языков программирования и программ. Ко всем этим средствам имеется полный доступ прямо из системного окна. Система реализована на больших ЭВМ, ПК класса IBM PC, Macintosh и др.

Ядро системы Maple используется целым рядом систем компьютерной математики, например MatLAB и MathCAD.

10.3.2. Система Mathematica

Система Mathematica разработана фирмой Wolfram Research. Основная идея разработчиков — объединить все известные понятия и методы математики в единую универсальную систему, способную функционировать на любой вычислительной платформе. Эта система дает возможность решать большое количество достаточно сложных задач, не вдаваясь в тонкости программирования, что привело к её широкому использованию в физике, биологии, экономике и т.д.

Система Mathematica состоит из двух частей — ядра, которое и проводит вычисления, выполняя заданные команды, и интерфейсного процессора, фактически задающего внешнее оформление и характер взаимодействия с пользователем и программой. Пользователь записывает все выкладки в основном рабочем документе программы — notebook («записной книжке»). Внешний вид рабочего документа на экране монитора в большей или меньшей степени зависит от типа платформы (Windows, Macintosh, Unix) и определяется интерфейсным процессором для каждой платформы.

Notebook — полностью интерактивный документ, содержащий текст, таблицы, графики, вычисления и другие элементы. Документы могут быть представлены на экране с различным качеством исполнения, могут включать в себя всевозможные математические или специальные символы. Однако всегда внутреннее представление документов, с которыми работает ядро программы, — не форматируемый текст в виде печатаемых (printable) 7-битовых ASCII-символов. Это позволяет легко переносить документы с одной платформы на другую.

Язык программирования Mathematica рассматривается как единый универсальный язык программирования и математики. Он содержит уникальную комбинацию математических и вычислительных обозначений. Основная унифицирующая идея языка — представление любого объекта в виде символьного выражения. Это касается не только привычных структур данных, но и графиков, звуков, ячеек и даже самого документа, что позволяет легко дополнять к системе Mathematica новые конструкции языка и изменять интерфейс программы.

Язык включает в себя представление широкоизвестных развитых методов программирования и добавляет множество новых. На языке программирования можно всегда написать программу в наиболее есте-

ственном виде, так как Mathematica включает в себя множество парадигм программирования: процедурное программирование; основанное на операциях со списками (list-based); основанное на операциях со строками (string-based); функциональное; объектно-ориентированное; программирование, задающее правила преобразования выражений («правила переписывания термов»).

В виде встроенных функций реализовано большое множество высокоинтеллектуальных математических операций. В состав Mathematica входят стандартные дополнения, включающие в себя подпрограммы (пакеты), которые значительно расширяют функциональные возможности системы в таких областях, как алгебра, аналитические и численные расчеты, графика, дискретная математика, геометрия, теория чисел, статистика и др.

10.3.3. Система MatLAB

Лидером в области численных и матричных расчетов, а также реализации техники имитационного и ситуационного моделирования является система MatLAB с ее многочисленными пакетами расширения. Однако в области аналитических вычислений она сильно уступает таким системам, как Maple и Mathematica.

Система MatLAB (MATrix LABoratory — Матричная лаборатория) разработана фирмой Math Works. Это интерактивная система, ориентированная, в первую очередь, на обработку массивов данных, базируется на матричных математических операциях. Даже одиночное число MatLAB рассматривает как матрицу, что позволило существенно повысить скорость выполнения вычислений.

С точки зрения пользователя MatLAB представляет собой богатейшую библиотеку функций. Для облегчения поиска библиотека функций разбита на разделы. Те из функций, которые носят более общий характер и используются наиболее часто, входят в состав ядра MatLAB. Те функции, которые являются специфическими для конкретной области, включены в состав соответствующих специализированных разделов, называемых MatLAB Toolboxes (*Инструменты*). Каждый из них имеет свое собственное название, отражающее его предназначение. Полная комплектация системы MatLAB содержит более 30 инструментальных приложений. В их число входят как достаточно стандартные для математических пакетов средства, так и нетрадиционные — средства цифровой обработки изображений, идентификации, синтеза и моделирования систем управления (оптимального, робастного, с прогнозирующими моделями), средства поиска решений на основе нечеткой логики, аппарат построения и анализа нейронных сетей, средства финансового анализа и др. Кроме того, имеются средства взаимодействия с офисными продуктами фирмы Microsoft — MS Word и MS Excel.

К несомненным достоинствам системы MatLAB следует отнести тот факт, что она базируется на современных математических подходах, имеющих непосредственный выход в практику проектирования технических систем, например таких, как теория робастности и др.

Особое место среди инструментальных приложений занимает система визуального моделирования SIMULINK. В определенном смысле SIMULINK можно рассматривать как самостоятельный продукт фирмы Math Works, однако он работает только при наличии ядра MatLAB и использует многие функции, входящие в его состав.

MatLAB является платформно-независимой системой, так как может работать под управлением нескольких операционных систем: Windows, UNIX, MacOS. При этом технология моделирования с помощью SIMULINK остается неизменной.

10.3.4. Система MathCAD

Система MathCAD разработана фирмой MathSoft. Она существенно отличается от аналогичных прикладных систем MatLAB, Mathematica, Maple тем, что является единственной прикладной системой, в которой описания математических задач и их решений задаются с помощью обычных в математике символов, формул и операторов, а документ MathCAD выглядит как страницы учебника или научной статьи.

MathCAD — это популярная система компьютерной математики, предназначенная для решения математических задач в самых разных областях науки, техники и образования. Ранние версии MathCAD вообще не имели средств обычного программирования, а имели лишь средства визуально-ориентированного программирования в виде шаблонов математических операций, из которых составлялись математические выражения. Возможность задания программных модулей появилась, начиная с версии MathCAD PLUS 6.0. Программные модули, в сущности, являются функциями, описанными с применением программных средств.

Последние версии MathCAD допускают применение внешних расширений: электронных книг, пакетов расширений и библиотек. Внешние библиотеки чаще всего являются справочниками. Пакеты расширений и сопровождающие их электронные книги имеют различное назначение и ориентированы, как правило, на выполнение узкоспециальных вычислений. Пакеты расширений предоставляют следующие возможности: осуществлять вейвлет-преобразования, анализ временных рядов, финансовые расчеты, обработку сигналов и изображений, расчеты по прикладной статистике, механике, дифференциальным уравнениям и т.д. Существенными дополнениями к системе являются пакеты расширения графических возможностей, численных методов.

В MathCAD возможна интеграция с текстовым процессором Word,

электронными таблицами Excel, графической системой Axum, пакетом научной и инженерной графики Visio, с пакетом SmartSketch LE, позволяющим включать в состав документов MathCAD довольно сложные конструкторские чертежи, с пакетом VisSim — моделирующей программой, в которой объекты задаются блоками, между которыми указываются дополнительные связи.

Привычный вид математических формул, встроенный язык программирования, широкие возможности в отображении графической информации, представление текстовых комментариев с использованием любых символов, доступных в Windows, возможность проводить расчеты любой сложности и готовить любые документы — все это характеризует систему MathCAD.

10.3.5. Система STATISTICA

Программная система STATISTICA разработана фирмой Copyright StatSoft. Это современный пакет для статистического анализа данных, в котором реализованы все новейшие компьютерные и математические методы анализа данных — описательные статистики; анализ многомерных таблиц; многомерная и нелинейная регрессии; подгонка распределений; дискриминантный, кластерный, факторный, дисперсионный и ковариационный анализы; структурные модели; прогнозирование временных рядов; непараметрическая статистика; анализы надежности предпочтений, Монте-Карло, выживаемости и т.д. Система STATISTICA подходит для применения в любой области — в маркетинге, финансах, страховании, бизнесе, промышленности, медицине и пр.

10.3.6. Система EXCEL

Табличный процессор EXCEL входит в самый популярный пакет автоматизации офисной деятельности Microsoft Office. EXCEL — одна из самых мощных и гибких систем обработки электронных таблиц. Эта система может работать не только с двумерными, но и с трехмерными таблицами.

EXCEL широко используется для подготовки иллюстрированных финансово-экономических и других документов. Этот процессор содержит сотни математических и экономических функций, что позволяет решать множество задач в области естественных и технических наук.

Возможности системы EXCEL можно использовать для анализа внешних данных в Microsoft FoxPro, Access, Paradox, dBASE, SQL Server, а также как базы данных сторонних производителей, поддерживающих технологию OLE (Object Linking and Embedding связывания и внедрения объектов).

В системе EXCEL можно создавать макросы и приложения с помощью Visual Basic for Applications (VBA). Visual Basic for Applications — это среда разработки приложений на базе программ, входящих в пакет Microsoft Office. Одно из главных достоинств языка Visual Basic for Applications заключается в том, что созданные средствами EXCEL VBA-макросы можно без труда использовать в других программах фирмы Microsoft.

ЛИТЕРАТУРА

1. Андреев, Ю.Н. Управление конечномерными линейными объектами / Ю.Н. Андреев. — М. : Наука, 1976. — 424 с.
2. Антонов, А.В. Системный анализ : учеб. для вузов / А.В. Антонов. — М. : Высш. шк., 2004. — 454 с. — ISBN 5-06-004862-4.
3. Браммер, К. Фильтр Калмана-Бьюси / К. Браммер, Г. Зиффлинг. — М. : Наука, 1982. — 200 с.
4. Буков, В.Н. Адаптивные прогнозирующие системы управления полетом / В.Н. Буков. — М. : Наука, 1987. — 232 с.
5. Бусленко, Н.П. Моделирование сложных систем / Н.П. Бусленко. — М. : Наука, 1978. — 400 с.
6. Бусленко, Н.П. Метод статистических испытаний (Монте-Карло) и его реализация на цифровых вычислительных машинах / Н.П. Бусленко, Ю.А. Шрейдер. — М. : Физматгиз, 1961. — 226 с.
7. Волкова, В.Н. Основы теории систем и системного анализа : учеб. для студентов вузов, обучающихся по специальности «Системный анализ и управление» / В.Н. Волкова, А.А. Денисов. — СПб. : Изд-во СПбГТУ, 1997. — 510 с. — ISBN 5-7422-0026-9.
8. Горский, А.А. Динамическая модель процесса производства, хранения и сбыта товара повседневного спроса / А.А. Горский, И.Г. Колпакова, Б.Я. Локшин // Изв. РАН. Теория и системы управления. — 1998. — № 1. — С. 144–148 — ISSN 0002-3388.
9. Гультияев, А.К. MATLAB 5.2. Имитационное моделирование в среде Windows : практ. пособие / А.К. Гультияев. — СПб. : КОРОНА принт, 1999. — 288 с. — ISBN 5-7931-0053-9.
10. Дорф, Р. Современные системы управления / Р. Дорф, Р. Бишоп. — М. : Лаборатория базовых знаний, 2002. — 832 с. — ISBN 5-93208-119-8.
11. Ермаков, С.М. Курс статистического моделирования / С.М. Ермаков, Г.А. Михайлов. — М. : Наука, 1976. — 320 с.
12. Красовский, А.А. Универсальные алгоритмы оптимального управления непрерывными процессами / А.А. Красовский, В.Н. Буков, В.С. Шендрик. — М. : Наука, 1977. — 272 с.
13. Крылов, В.И. Вычислительные методы. В 2 т. Т. I / В.И. Крылов, В.В. Бобков, П.И. Монастырный. — М : Наука, 1976. — 304 с.
14. Крылов, В.И. Вычислительные методы. В 2 т. Т. II / В.И. Крылов, В.В. Бобков, П.И. Монастырный. — М : Наука, 1976. — 400 с.

15. Летов, А.М. Аналитическое конструирование регуляторов / А.М. Летов // Автоматика и телемеханика. – 1960. – Ч. 1. – № 1. – С. 436–441 ; ч. 2. – № 5. – С. 561–568 ; ч. 3. – № 6. – С. 661–665 ; ч. 4. – № 4. – С. 425–435 ; 1962. – ч. 5. – № 11. – С. 1405–1413.
16. Медич, Дж. Статистически оптимальные линейные оценки и управление / Дж. Медич. – М. : Энергия, 1973. – 440 с.
17. Мудров, А.Е. Численные методы для ПЭВМ на языках Бейсик, Фортран и Паскаль / А.Е. Мудров. – Томск : МП «РАСКО», 1991. – 272 с. – ISBN 5-256-00602-9.
18. Острем, К. Системы управления с ЭВМ / К. Острем, Б. Виттенмарк. – М. : Мир, 1987. – 480 с.
19. Параев, Ю.И. Теория оптимального управления: учеб. пособие / Ю.И. Параев. – Томск : Изд-во НТЛ, 2004. – 168 с. – ISBN 5-89503-225-7.
20. Перегудов, Ф.И. Основы системного анализа: учеб. / Ф.П. Перегудов, Ф.П. Тарасенко. – 2-е изд., доп. – Томск : Изд-во НТЛ, 1997. – 396 с. – ISBN 5-89503-004-1.
21. Решетникова, Г.Н. Синтез и моделирование дискретных адаптивных систем (гос. регистр. № 50880000594) / Г.Н. Решетникова // Алгоритмы и программы : информ. бюл. – 1989. – № 1. – С. 10.
22. Решетникова, Г.Н. Синтез следящей системы адаптивного управления темпом производства по локальному критерию / Г.Н. Решетникова // Вестник ТГУ. – 2004. – Вып. 284. – С. 166–168. – ISSN 1561-7793.
23. Седов, Л.И. Методы подобия и размерности в механике / Л.И. Седов. – М. : Наука, 1987. – 430 с.
24. Смагин, В.И. Синтез следящих систем управления по квадратичным критериям / В.И. Смагин, Ю.И. Параев. – Томск : Изд-во Том. ун-та, 1996. – 171 с. – ISBN 5-7511-0830-2.
25. Советов, Б.Я. Моделирование систем: учеб. для вузов / Б.Я. Советов, С.А. Яковлев. – 3-е изд., перераб. и доп. – М. : Высш. шк., 2001. – 343 с. – ISBN 5-06-003860-9.
26. Справочник по теории автоматического управления / под ред. А.А. Красовского. – М. : Наука, 1987. – 712 с.
27. Харин, Ю.С. Основы имитационного и статистического моделирования : учеб. пособие / Ю.С. Харин, В.И. Малюгин, В.П. Кирлица, В.И. Лобач, Г.А. Хацкевич. – Мн. : Дизайн ПРО, 1997. – 288 с. – ISBN 985-6182-31-X.
28. Шеннон, Р. Имитационное моделирование систем: искусство и наука / Р. Шеннон. – М. : Мир, 1978. – 418 с.

Учебное издание

Решетникова Галина Николаевна

МОДЕЛИРОВАНИЕ СИСТЕМ

Учебное пособие

Редактор Н.С. Голикова
Компьютерная верстка Г.В. Чернова
Корректор О.В. Полещук

Подписано в печать 02.08.05. Формат 70×100/16.
Бумага офсетная. Гарнитура School. Печать офсетная.
Усл. печ. л. 21,04. Учет.-изд. л. 15,90. Тираж . Заказ .

Томский государственный университет
систем управления и радиоэлектроники.
634050, Томск, пр. Ленина, 40.
Тел. 8 (3822) 533018.