ЛЕКЦИЯ №1. ОПРЕДЕЛЕНИЯ И ПОНЯТИЯ

Системы и системный подход, понятие системы, принципы управления, синергетика, классификация систем.

Основные положения современной системотехники и теории систем были разработаны в трудах таких ученых как Шеннон К.Э., Берталанфи К.Л., Винер. Н., Пригожин И., Эшби У.Р., Богданов А.А. [8, 12, 14, 7]. Рассмотрим ряд понятий и методологических концепций, лежащих в основе системотехники и анализа систем.

Системой называется некая совокупность элементов, обособленная от окружающей среды и взаимодействующая с ней как некое целое. Важной особенностью системы является невозможность выполнения отдельными элементами ее функций. Данная особенность получила наименование **эмерджентности**.

Система взаимодействует с внешней средой. Приведем основные виды взаимодействия:

- обмен энергией;
- обмен веществом;
- обмен информацией.

В общем виде можно выделить два класса систем:

- природные;
- искусственные.

Искусственные системы включают в себя:

- технические;
- автоматизированные;
- организационные.

Технические системы управления – это системы, которые содержат в качестве элементов технические устройства и могут в течение некоторого интервала времени функционировать без участия человека. Технические системы управления имеют следующие особенности:

- четко определенную единственную цель управления;
- отсутствие человека в контуре управления;
- достаточно высокую определенность исходных данных;
- характеризуются высокой степенью формализации процессов функционирования.

В технических системах легко выделить объект управления и управляющую систему.

Автоматизированные системы управления – это системы, включающие в себя в качестве элементов как технические системы, так и персонал, взаимодействующий с этими системами. Автоматизированные информационные системы функционируют в определенной предметной области. При этом предметную область можно определить как сферу интересов пользователей данной системы. Они обеспечивают автоматизацию процесса приема, хранения, обработки и передачи информации.

Организационные системы возникают в обществе. В качестве такой системы выступает коллектив людей, деятельность которых сознательно координируется для достижения определенной цели.

Системы, созданные человеком, возникают для удовлетворения определенной цели. При этом возникает причинно-следственная цепочка:

Потребность – Цель – Функционирование системы – Результат.

Для получения результата требуется определенный метод функционирования системы. Возникает потребность в управлении системой. Управление системой в общем виде может быть реализовано тремя способами:

- формирование управления с учетом изменения выходных сигналов;
- формирование управления с учетом изменения выходных сигналов и изменений в структуре системы;
- формирование управления с учетом изменений выходных сигналов, структуры и с учетом влияния внешней среды.

В системах управления любого типа выделяют следующие составные элементы:

- множество входных сигналов X;
- множество выходных сигналов Y;
- каналы воздействия на систему возмущений V.

$$X = \{x_1, x_2, \dots, x_i\},$$
$$Y = \{y_1, y_2, \dots, y_i\}, \qquad (1.1)$$
$$V = \{v_1, v_2, \dots, v_i\}.$$

Систему с учетом входных воздействий, выходных воздействий и возмущений можно представить в виде, показанном на рисунке 1.1.

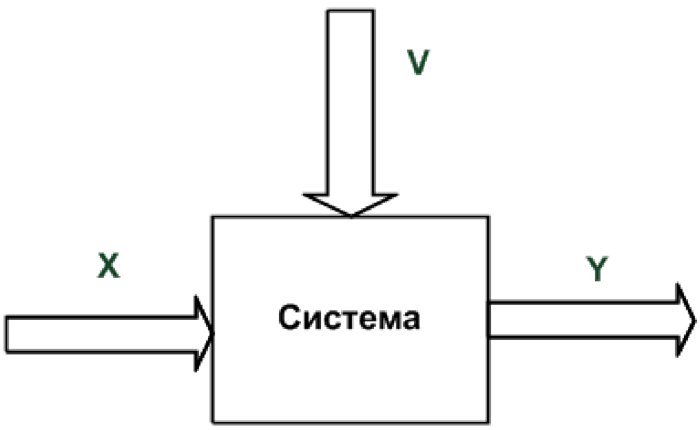


Рис. 1.1. Система, воздействие на систему, выход системы

Для контроля правильности работы системы и обеспечения ее устойчивости используют замкнутый контур управления. С наличием замкнутого контура связано понятие гомеостазиса. Гомеостазис – поддержание выходных параметров системы на определенном уровне. Реализуется гомеостазис отрицательной обратной связью (рис. 1.2).

Если между вектором входных величин X и вектором выходных величин Y существует однозначная функциональная связь, то вектор входных величин полностью определяет вектор выходных величин.

$$Y(t) = f[X(t)]. \tag{1.2}$$

Основой системы любой природы является ее элемент. Элементом системы называется ее неделимая часть. Структурой системой называется совокупность ее элементов, связанных определенным образом.

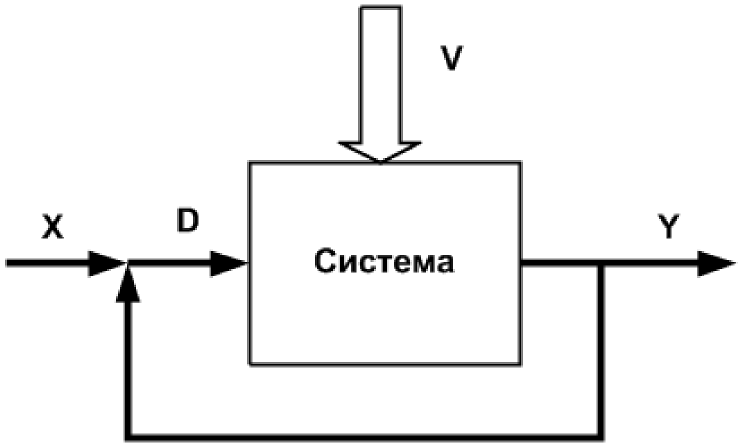


Рис. 1.2. Отрицательная обратная связь; D = X – Y

Структуру системы можно описать с помощью формулы:

$$S = \{e_{ij}\}, \tag{1.3}$$

где i и j изменяются от $1...n$; n – число связей в системе. Структуру системы отображают в виде направленного графа, так как это показано на рис. 1.3.

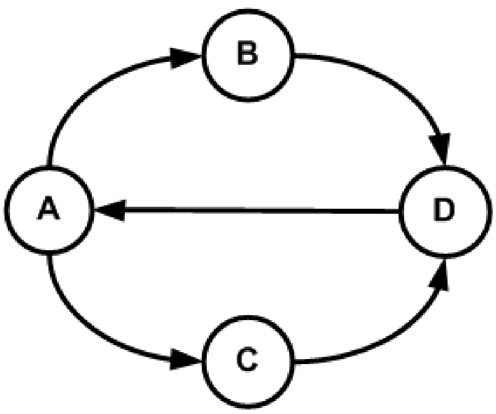


Рис. 1.3. Система, состоящая из четырех элементов A, B, C, D

Для анализа ее связей с помощью ЭВМ. Структура системы может быть также представлена в виде таблицы. Система, показанная на рис. 1.3. представлена в табл. 1.1.

Таблица 1.1

Структура системы

	A	B	C	D
A	0	1	1	0
B	0	0	0	1
C	0	0	0	1
D	1	0	0	0

Таблица состоит из нулей и единиц. Если между элементами имеется направленная связь, то на пересечении строки и столбца ставится 1, иначе 0.

Совокупность элементов системы характеризуется определенным набором параметров состояний, которые изменяются во времени.

$$Z = \{z_1, z_2, ..., z_i\}. \qquad (1.4)$$

Если в процессе управления системой параметры состояния должны обладать определенными постоянными значениями, то такое состояние системы называется состоянием с гомеостазисом.

С учетом состояния в системе имеет место два типа функциональных связей. Зависимость внутреннего состояния системы Z от вектора входных величин X и зависимость вектора выходных величин Y от множества внутренних параметров системы и вектора входных величин. Это можно выразить уравнением

$$Y(t) = f[X(t), Z(t)].$$

Для описания процесса движения динамических систем широко применяется метод, основанный на использовании фазового пространства (n-мерного Евклидова пространства). По осям откладываются значения всех n обобщенных координат рассматриваемой динамической системы. Размерность этого пространства определяется, числом фазовых координат. Это число отобранных для описания системы ее существенных параметров состояния.

Геометрическое место точек при переходе системы из состояния Z_1 в Z_2, представляющее собой некоторую траекторию в фазовом пространстве, отражает процесс движения системы. Число независимых параметров состояния системы называют числом степеней свободы.

На рисунке 1.4 показано двумерное фазовое пространство с двумя зонами в виде окружностей. Первая зона S_r – рабочая зона, вторая зона S_d зона предельных допустимых значений параметров состояния.

Процессы движения системы могут рассматриваться в непрерывном или дискретном времени.

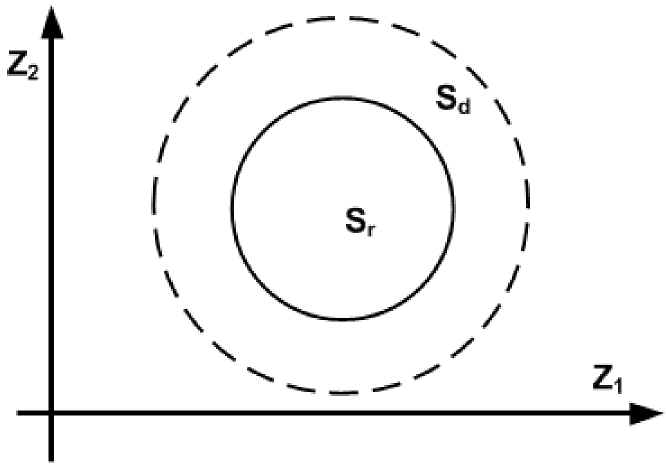


Рис. 1.4. Фазовое пространство

Непрерывное время: Фиксация состояния системы последовательностями изображающих точек, которые соответствуют отсчетам, осуществляемым через бесконечно малые интервалы времени Δt.

Дискретное движение: Фиксируются отсчеты, которые берутся через определенные (не бесконечно малые) интервалы времени Δt – такты (рис. 1.5).

Движение любой системы представляет собой некоторую последовательность изменений ее состояний. Характеризуя состояние системы в некоторый момент времени t_i вектором Z_i, а состояние в момент t_{i+1} – вектором Z_{i+1}, можно считать, что произошел переход $(Z_i, t_i) \Rightarrow (Z_{i+1}, t_{i+1})$. Закон, согласно которому каждому элементу некоторого множества состояний соответствует элемент другого множества состояний, называется оператором.

При переходе системы в новое состояние оператором будет называться закон, в соответствии с которым происходит этот переход.

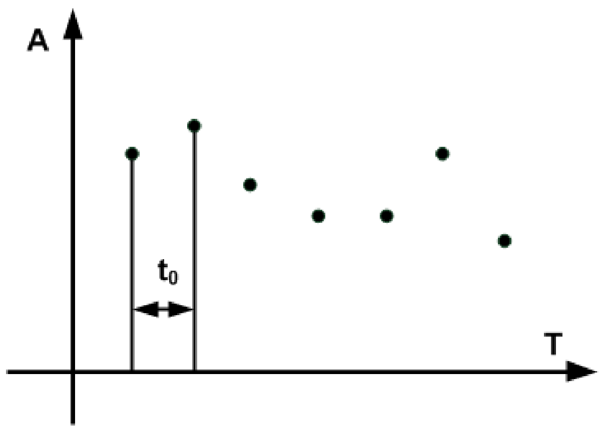


Рис. 1.5. Дискретный процесс: t_0 – такт дискретизации

Значение переменной величины, над которой совершается операция, вызывающая переход, называется операндом. Новое значение переменной, новое состояние, возникшее в результате воздействия оператора на операнд, называется *образом*. Преобразования, в которых каждому операнду соответствует только один образ, называются однозначными.

С точки зрения структуры системы говорят о:

- простых системах;
- больших системах;
- сложных системах.

Простые системы состоят из определенного набора элементов.

Большие системы состоят из отдельных подсистем.

Сложные системы это совокупность пространственно распределенных сложных систем. В сложных системах их структуру можно рассматривать как состояние, возникающее в результате многовариантного и неоднозначного поведения многоэлементных структур. Для таких систем характерны следующие особенности:

- они развиваются вследствие открытости;
- к ним поступает энергия извне;
- внутренние процессы нелинейные;
- в них присутствуют особые режимы с обострением и наличием более одного устойчивого состояния.

Режимы с обострением – точки бифуркации. Такие режимы наблюдаются при смене установившегося режима работы системы. Точка бифуркации – критическое состояние системы, при котором система становится неустойчивой. Обычно точка бифуркации имеет несколько веточек аттрактора (устойчивых режимов работы), по одному из которых пойдёт развитие системы. Однако заранее невозможно предсказать, какой новый аттрактор займёт система.

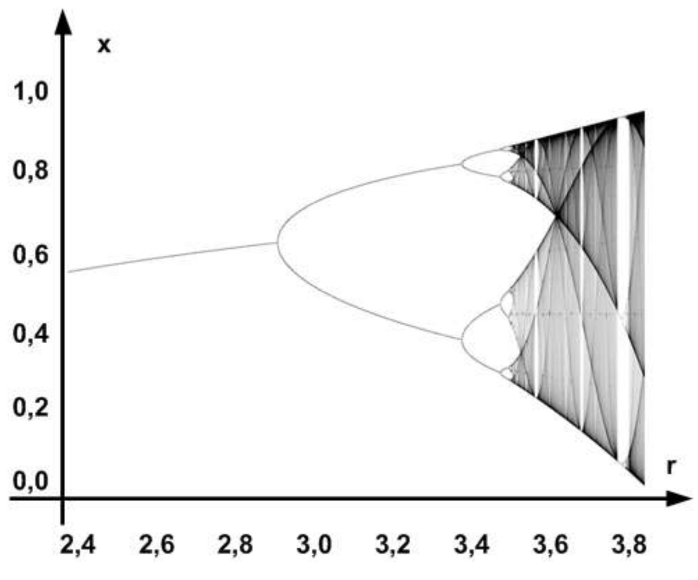


Рис. 1.6. Бифуркация в системе

Точка бифуркации носит кратковременный характер и разделяет более длительные устойчивые режимы системы. Возникает неопределенность: станет ли состояние системы хаотическим или она перейдет на новый, более дифференцированный и высокий уровень упорядоченности. Системы с бифуркацией предмет изучения синергетики. При этом говорят о синергетическом эффекте. **Синергетический эффект** – получение новых свойств системы в результате объединения в нее подсистем.

На рисунке 1.6 показан бифуркационный график развития системы, построенный на основе логистического уравнения:

$$x_{n+1} = r \cdot x_n \cdot (1 - x_n). \tag{1.5}$$

Из графика (рис. 1.6) видно, что при значении параметра *r* более 3,4 единиц система становится неустойчивой – это состояние хаоса.

Контрольные вопросы

1. Какие виды обмена можно выделить между системой и окружающей средой?
2. Дайте общую классификацию систем.
3. Что такое причинно-следственная цепочка?
4. Дайте определения эмерджентности и синергетического эффекта.
5. Что называется элементом системы?
6. Как можно отобразить структуру системы?
7. Как можно классифицировать системы с точки зрения их структуры?
8. Что называется фазовым пространством системы?
9. Что такое фазовая траектория?
10. Что такое точка бифуркации?
11. Как в системе поддерживается состояние гомеостазиса?

ЛЕКЦИЯ №2. ИНФОРМАЦИЯ В СИСТЕМАХ

Информация и ее особенности, измерение информации, управление системами, оценка устойчивости систем, проблемы исследования систем.

При анализе процессов в системах используют такие понятия как сигналы, данные, сообщения, информация [8]. Сигнал – физический процесс (явление), несущий сообщение о событии или состоянии объекта наблюдения. Данными называют факты, сведения, представленные в формализованном (закодированном) виде и занесенные на носители. Данные допускают обработку с помощью специальных технических средств. Данные несущие смысловую нагрузку называются знаниями.

Под информацией принято понимать некоторую упорядоченную последовательность сообщений, отражающих, передающих и увеличивающих наши знания. Это главная характеристика процессов, протекающих в системе. Рассмотрим способы оценки – измерения информации [8, 14].

Объемный подход

При реализации информационных процессов информация передается в виде сообщения, представляющего собой совокупность символов какого-либо алфавита. Каждый новый символ в сообщении увеличивает количество информации, представленной последовательностью символов данного алфавита. Кодирование символов выполняют в двоичной системе. При этом имеет место следующий ряд единиц измерения информации:

- 1 бит = 0 или 1;
- 1 байт = 8 бит;
- 1 килобайт (1 Кб) = 2^{13} бит;
- 1 мегабайт (1 Мб) = 2^{23} бит;
- 1 гигабайт (1 Гб) = 2^{33} бит;
- 1 терабайт (1 Тб) = 2^{43} бит;
- 1 петабайт (1 Пб) = 2^{53} бит;
- 1 эксабайт (1 Эб) = 2^{63} бит.

Энтропийный подход

Под информацией понимается количественная величина исчезнувшей в ходе какого-либо процесса (испытания, измерения и т.д.) неопределенности. Для измерения информации вводится мера Р. Хартли.

Пусть известны N состояний системы S (N опытов с различными, равновозможными – вероятными, последовательными состояниями системы). Если каждое состояние системы закодировать двоичными кодами, то длину кода d необходимо выбрать так, чтобы число всех различных комбинаций было бы не меньше, чем N:

$$2^d \geq N.$$

Логарифмируя это неравенство, можно записать:

$$d \geq lod_2 N, \qquad H = \log_2 N. \tag{2.1}$$

Наименьшее решение этого неравенства или мера разнообразия множества состояний системы называется формулой Р. Хартли.

Если во множестве $X = \{x_1, x_2, ..., x_n\}$ искать произвольный элемент, то для его нахождения необходимо иметь не менее $\log_a n$ (единиц) информации. Уменьшение H говорит об уменьшении разнообразия состояний N системы.

Формула Шеннона

Мера Хартли подходит лишь для идеальных, абстрактных систем, так как в реальных системах состояния системы неодинаково осуществимы (не равновероятны). Для таких систем используют меру К. Шеннона. Мера Шеннона оценивает информацию отвлеченно от ее смысла

$$I = -\sum_{i=1}^{n} p_i \log_2 p_i, \qquad (2.2)$$

где n – число состояний системы, p_i – вероятность (относительная частота) перехода системы в i-е состояние. Сумма всех p_i должна равняться 1.

Если все состояния рассматриваемой системы равновозможные, равновероятны, то $p_i = 1/n$, из формулы Шеннона можно получить (как частный случай) формулу Хартли $I = \log_2 n$.

Мера I, представляет собой неопределенность, приходящуюся в среднем на одно состояние, называют энтропией дискретного источника информации.

Используя формулы Хартли и Шеннона, можно определить избыточность D алфавита источника сообщений – A:

$$D = [H_{\max}(A) - H(A)] / [H_{\max}(A)], \qquad (2.3)$$

где $H_{\max}(A)$ – мера информации Хартли, $H(A)$ – мера информации Шеннона.

Избыточность показывает, насколько рационально применяются символы данного алфавита.

Семантическая мера информации

Для измерения смыслового содержания информации используется тезаурусная мера. Она связывает семантические свойства информации со способностью пользователя принимать поступившее сообщение.

Тезаурус – совокупность сведений, которыми располагает пользователь или система. Полное незнание предмета не позволяет извлечь полезную информацию из принятого сообщения об этом предмете.

В зависимости от соотношений между смысловым содержанием информации S и тезаурусом пользователя S_p изменяется количество семантической информации I_c, воспринимаемой пользователем и включаемой им в дальнейшем в свой тезаурус (рис. 2.1).

Анализируя рис. 2.1 можно сделать следующие выводы:

- если $S_p = 0$ пользователь системы не воспринимает (не понимает) поступающую информацию;
- если $S_p \to \infty$ пользователь системы «все знает», и поступающая информация ему не нужна.

Получение информации в системе может быть реализовано различными методами. Можно выделить:

- эмпирические методы или методы получения информации посредством эксперимента;

- теоретические методы или методы построения различных теорий;
- эмпирико-теоретические методы (смешанные) или методы построения теорий на основе полученных эмпирических данных об объекте, процессе, явлении.

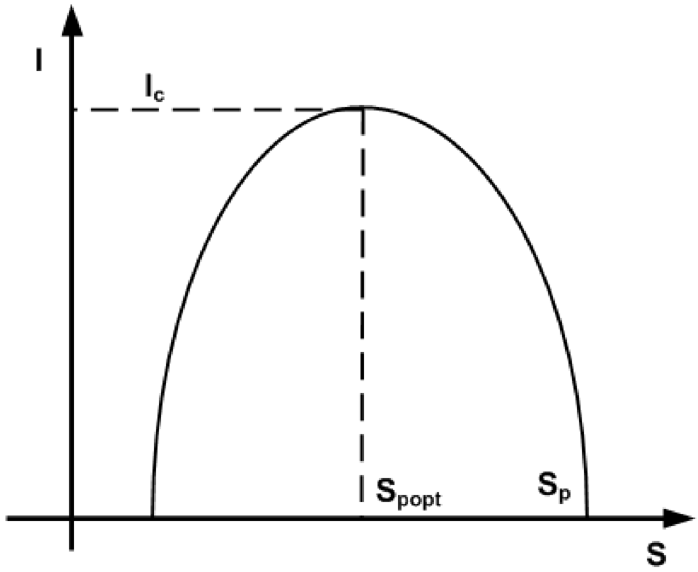


Рис. 2.1. Семантическая мера информации

Системы существуют в определенной информационной среде. Информационная среда – это среда из взаимодействующих информационных систем, включая и информацию, актуализируемую в этих системах. С учетом наличия информационной среды можно говорить об особом классе систем – информационных системах. Информационная систему можно определить как систему, в которой элементы, структура, цель, ресурсы рассматриваются на информационном уровне.

Информация играет важную роль в процессе управления системой. Управление в системе – внутренняя функция системы, осуществляемая независимо от того, каким образом, какими элементами системы она должна выполняться. Управление системой – выполнение внешних функций управления, обеспечивающих необходимые условия функционирования системы. Схема управления системой показана на рис. 2.2 [8, 14].

Управление системой используется для различных целей:

- увеличения скорости передачи сообщений;
- увеличения объема передаваемых сообщений;
- уменьшения времени обработки сообщений;
- увеличения степени сжатия сообщений;
- увеличения (модификации) связей системы;
- увеличения информации (информированности).

Рис. 2.2. Управление системой и информация

Реализация управления предусматривает выполнение определенного цикла. Цикл управления любой системой состоит из следующих шагов:

- сбор информации о системе;
- обработка и анализ информации;
- получение информации о траектории системы выявление управляющих параметров;
- определение ресурсов для управления;
- изменение траектории системы.

Цель управления – движение системы из одного целевого состояния в другое. Управление преследует цель обеспечить движение с максимальным приближением к идеальной траектории смены целевого состояния (рис. 2.3).

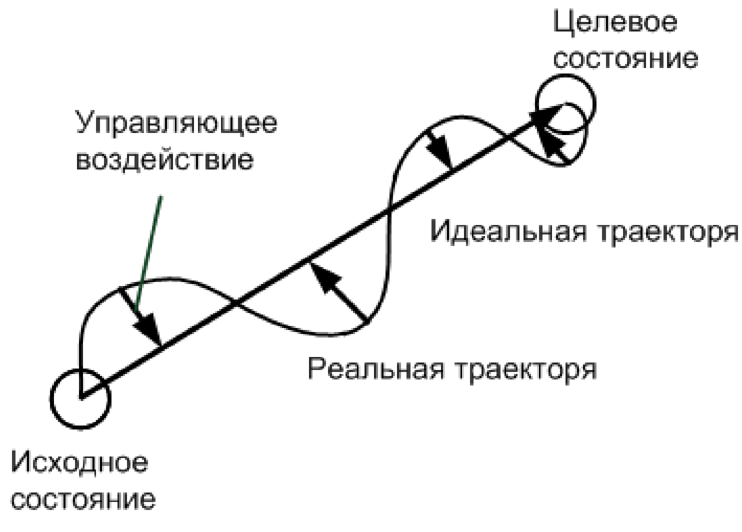


Рис. 2.3. Идеальная траектория и управление

Приведем основные правила организации информации для управления системой:

- выяснение формы и структуры исходной (входной) информации;
- выяснение средств, форм передачи и источников информации;
- выяснение формы и структуры выходной информации;
- выяснение надежности информации и контроль достоверности;
- выяснение форм использования информации для принятия решений.

Принцип Эшби

Если число возможных состояний системы *S* равно *N*, то общее количество разнообразия системы Хартли равно

$$V(N) = \log_2 N. \qquad (2.4)$$

Управляемая система обладает разнообразием $V(N_1)$, управляющая – $V(N_2)$. Цель управляющей системы – уменьшить значение $V(N_1)$ за счет изменения $V(N_2)$. В свою очередь, изменение $V(N_1)$, как правило, влечет изменение и $V(N_2)$. Управляющая система может эффективно выполнять присущие ей функции управления лишь при условии, если верно неравенство 2.5 [14]

$$V(N_2) \geq V(N_1). \qquad (2.5)$$

Это неравенство определяет эффективность управления.

Устойчивость систем

Асимптотическая устойчивость системы состоит в возврате системы к равновесному состоянию при $t \to \infty$ из любого неравновесного состояния.

Пусть система *S* зависит от вектора факторов, переменных $x = (x_1, x_2, ..., x_n)$. Матрицей системы назовем матрицу $E = ||e_{ij}||$. Она состоит из 1 и 0 $e_{ij} = 1$ лишь тогда, когда переменная x_i оказывает влияние на x_j. Связная устойчивость состоит в асимптотической устойчивости системы при любых матрицах *E*.

Интеллектуальные методы исследования систем

Когнитивная схема (карта) ситуации представляет собой ориентированный взвешенный граф, который строится по правилам:

- вершины взаимно однозначно соответствуют выделенным факторам ситуации, в терминах которых описываются процессы в ситуации;
- выявляются и оцениваются (положительное влияние, отрицательное влияние) причинно-следственные связи выделенных факторов друг на друга.

На рисунке 2.4 приведена схема учета факторов в процессе потребления электроэнергии. Кроме когнитивных схем (схем ситуаций) могут использоваться когнитивные решетки (шкалы, матрицы), которые позволяют определять стратегии поведения (например, производителя на рынке). Решетка образуется с помощью системы факторных координат, где каждая координата соответствует одному фактору, показателю (например, финансовому) или некоторому интервалу изменения этого фактора. Каждая область решетки соответствует тому или иному поведению.

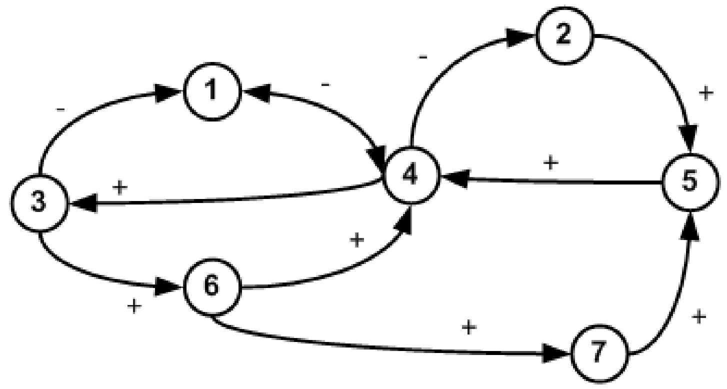


Рис. 2.4. Потребление энергии: 1– стоимость энергии;
2 – состояние окружающей среды; 3 – энергоресурсы; 4 –потребление;
5 – численность населения; 6 – число предприятий; 7 – число рабочих мест

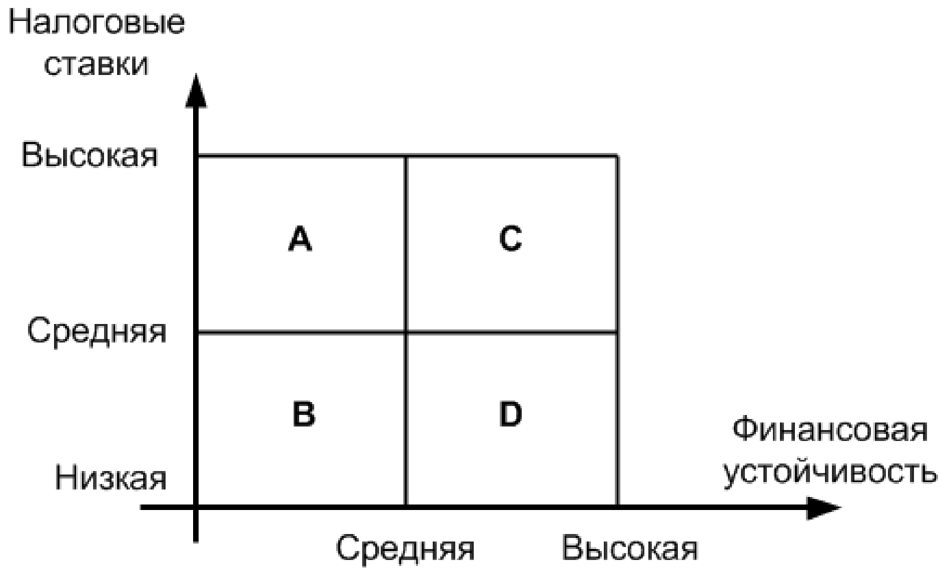


Рис. 2.5. Оценка налоговых ставок

Показатели могут быть относительными (например от 0 до 1), абсолютными (например от минимального до максимального) или биполярными («высокий или большой» – «низкий или маленький»). Пример когнитивной решетки показан на рис. 2.5. Оценивается налоговая ставка в зависимости от финансовой устойчивости фирмы (системы).

Интеллектуальные методы исследования требуют выполнения когнитивных операций:

- присвоение уникального имени свойству, отношению, объекту, процессу, системе;
- шкалирование и кластеризация, классификация;

- обобщение;
- сравнение;
- идентификация, узнавание объекта по его проявлениям;
- морфологический анализ (например, связей элементов);
- синтаксический анализ (например, атрибутов элементов и классов);
- семантический анализ (например, связей классов);
- верификация, сопоставление с опытом и заключение об обучении;
- планирование эксперимента;
- принятие решения.

Контрольные вопросы

1. Какие единицы измерения используются для оценки количества информации в системе?
2. Как вычислить количество информации в системе, если ее состояния равновероятны?
3. Как вычисляется количество информации в системе, если ее состояния не равновероятны?
4. Что такое информационный тезаурус?
5. Какие цели преследует организация управления системой?
6. В чем заключается принцип Эшби при организации управления системой?
7. Как оценить устойчивость систем?
8. Как создаются и используются когнитивные карты?
9. Как создаются и используются когнитивные решетки?
10. Перечислите интеллектуальные методы исследования систем.

ЛЕКЦИЯ №3. СИСТЕМНАЯ ИНЖЕНЕРИЯ

Универсальный язык моделирования систем UML, структурное моделирование, моделирование поведения систем, архитектурное моделирование.

Системная инженерия позволяет построить модели, как существующих систем, так и разрабатываемых систем. Выполняется моделирование с использованием универсального языка моделирования UML (Universal Model Language). Полностью метод системной инженерии с использованием языка UML приведен в источнике [6]. Рассмотрим базовые приемы моделирования с использованием методологии UML.

Структурное моделирование

Структурное моделирование предусматривает выделение в предметной области определенного набора классов определения взаимосвязей между ними. В результате формируется диаграмма классов (Static Structure Diagram). Основным элементом диаграммы является класс показанный на рис. 3.1.

Рис. 3.1. Класс и его описание

Элементами класса являются его атрибуты и операции. Класс должен обладать уникальным идентификатором – наименованием.

Атрибуты – это свойства класса, которые хранят информацию о наиболее существенных характеристиках реального объекта предметной области, выбранные разработчиком. Атрибуты помечаются специальными символами доступа, которые приведены в таблице 3.1.

Таблица 3.1

Символы доступа

Условное обозначение	Описание
–	Private (Закрытый). Доступ к атрибуту возможен только в операциях класса
+	Public (Глобальный). Атрибут доступен в операциях класса и в любом месте модели, где получен экземпляр класса
#	Protected (Защищенный). Доступ к атрибуту возможен в любой операции класса и подклассах

Атрибут должен быть определенного типа, который выбирается исходя из множества допустимых данных. Так атрибут может быть целого типа, вещественного типа, символьного типа. В качестве типа может указываться тип определенной системы разработки программного обеспечения. Таким образом, полное описание атрибута имеет вид:

Символ доступа Имя_атрибута: Тип.

Операция – выполнение действий с атрибутами. Задается в виде:

Доступ Наименование([параметры]):[Тип Возвращаемого значения]

Задание параметров операции имеет вид:

[Спецификация] Название: Тип

Спецификация параметров:

- in (в) – принимает значение;
- out (из) – возвращает значение;
- inout (в/из) – принимает и возвращает значение.

На рис. 3.2 приводится описание класса Worker(Работник). Типы атрибутов и параметров операций выбраны для языка программирования Visual Basic.

Диаграмма класса предполагает показ способа их объединения в систему. При этом различают такие способы объединения как зависимость, ассоциация, обобщение. Рассмотрим эти способы и правила их отображения.

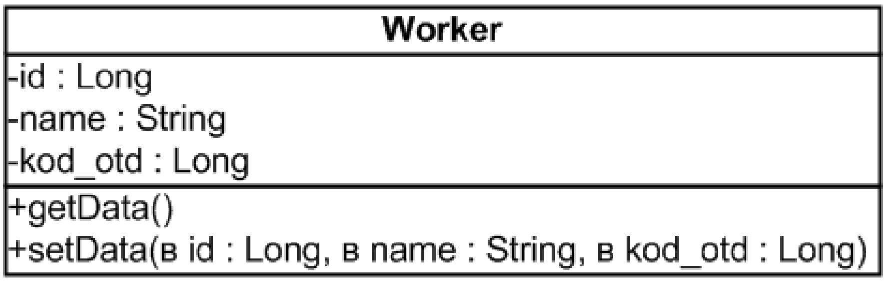


Рис. 3.2. Описание класса: id – код работника;name-фамилия работника; kod_otd – код отдела; getData – операция просмотра значений атрибутов; setData – операция задания значений атрибутам

Зависимость. Наличие одного класса в системе зависит от существования другого (рис. 3.3)

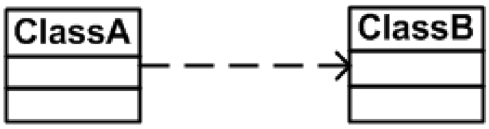


Рис. 3.3. Зависимость класса А от класса В

Ассоциация. Один объект содержит другой. Объекты, связанные отношением ассоциации, соединяются друг с другом. Ассоциация может быть указана со стрелкой или без нее. Существуют различные виды ассоциации.

Ролевое объединение

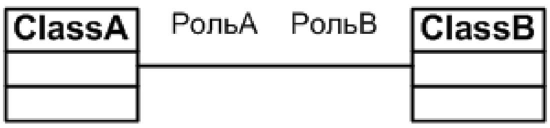


Рис. 3.4. Связь классов с указанием роли

Объединение с кратностью

Рис. 3.5. Указание кратности

В таблице 3.2 приводятся условные обозначения для указания кратности. В языке UML кратность – кардинальность.

Таблица 3.2

Условные обозначения кардинальности

Обозначение	Значение
1	Один объект
*	Любое число объектов
1...*	Не меньше одного объекта
x...*y*	От *x* до *y* объектов

Агрегация

Вид ассоциации, используется для моделирования «Целое/ Часть». В объединении участвуют равноправные классы. Изображается сплошной линией с не закрашенным ромбом наконце. Может использоваться с указанием кратности. Наличие такой связи говорит о том, что удаление класса из агрегатного объединения не разрушает систему.

Рис. 3.6. Агрегатное объединение

Пример. Отдел (ClassA) – Служащий (ClassB). Увольнение служащего не ведет к уничтожению отдела. После увольнения служащий может существовать как личность на рынке труда.

Композиция

Вид ассоциации, моделирующий отношение «Целое/Часть» между неравноправными классами. Часть, находящаяся в отношении композиции с целым, не является независимой от целого. Изображается композиция сплошной линией с закрашенным ромбом на конце (рис. 3.7).

Рис. 3.7. Композиция

Пример. Графическое окно (ClassA) – элемент диалога (ClassB). Элемент диалога не может существовать без контейнера – графического окна.

Обобщение. Такое отношение отражает связь между общим и конкретным. При наследовании выделяют родительский класс – суперкласс и подкласс (наследник). Подкласс наследует от суперкласса зачищенные и глобальные атрибуты и операции. Изображается обобщение сплошной линией с незакрашенной стрелкой. Стрелка должна указывать на суперкласс. На рисунке 3.8 показан пример наследования – обобщения.

Рис. 3.8. Обобщение: ClassA – суперкласс, ClassB – подкласс

Наличие обобщения при моделировании порождает иерархию классов. В главе иерархии находится общий предок – абстрактный класс. Имя этого класса записывается курсивом. Если в иерархии классов у класса нет подкласса, такой класс называется листовым. Пример иерархии классов показан на рис. 3.9.

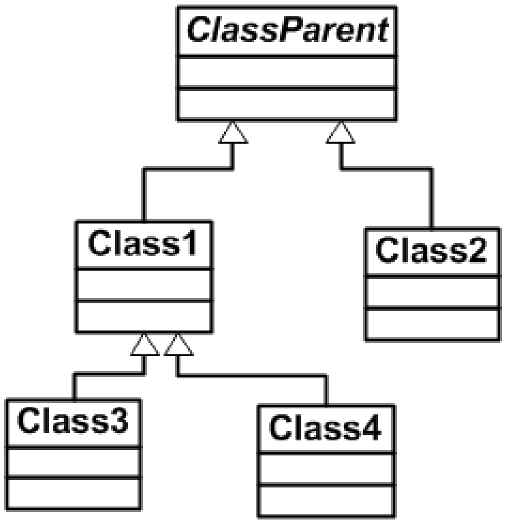


Рис. 3.9. Иерархия наследования: Class2, 3, 4 – листовые классы

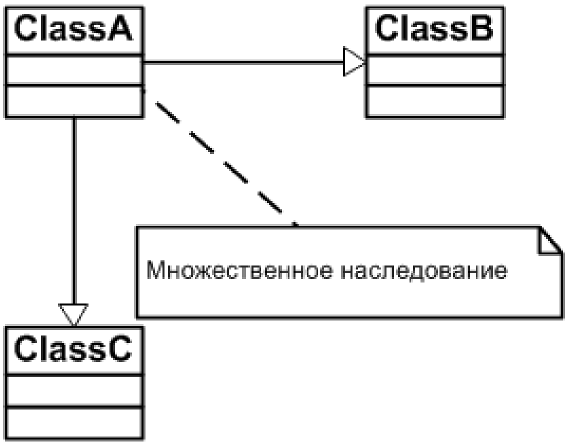


Рис. 3.10. Один подкласс и два суперкласса

При построении модели структуры системы возможно использование множественного наследования. Пример такого обобщения показан на рис. 3.10. Лист с загнутым углом – комментарий UML.

Объекты

При моделировании предметной области и информационной системы разработчик имеет дело не с самими классами, а их экземплярами – объектами. Объект – реализация данного класса.

Условное обозначение для экземпляра класса:

имяЭкземпляра: ИмяКласса.

Примеры объектов UML показаны на рис. 3.11.

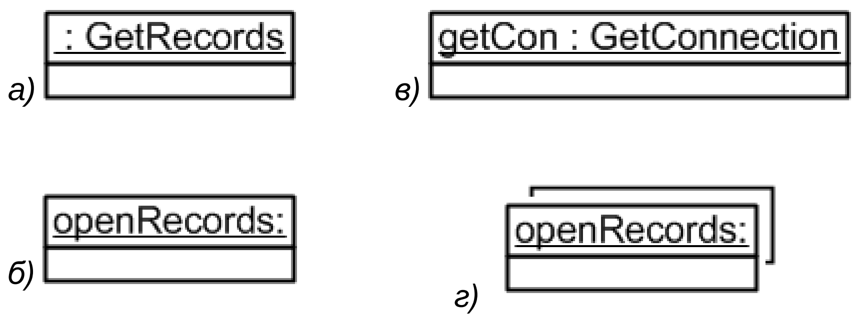


Рис. 3.11. Типы объектов: а – анонимный объект;
б – объект «сирота»; в – объект; г – коллекция объектов

При показе зависимости объекта от класса, связь может быть типизирована с помощью стереотипа. Стереотип указывается над связью в двойных кавычках. В UML для зависимости определено два стереотипа:

- instanceOf – объект является экземпляром класса;
- instantiate – объект создает экземпляры определенного класса.

Статические атрибуты и операции

Статические операции могут вызываться напрямую – без получения экземпляра класса. Статические атрибуты разделяются всеми экземплярами данного класса.

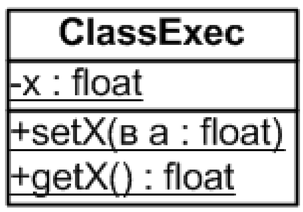


Рис. 3.12. Статические элементы класса

Таким атрибутам выделяется общий участок оперативной памяти. Если какой либо метод объекта данного класса вносит изменения в статический атрибут, то эти изменения автоматически распространяются на все объекты – экземпляра данного класса.

Статические элементы должны быть подчеркнуты, так как это показано на рис. 3.12.

Пакеты

Для группирования классов по определенному признаку используется механизм пакетов. Пакет позволяет также разграничить пространство имен классов. В модели могут быть классы с одинаковыми именами, но различающиеся по функциональному назначению. Принадлежность класса к пакету задается с помощью составного имени:

Наименование_пакета :: Название класса

Условное обозначение пакета показано на рисунке 3.13.

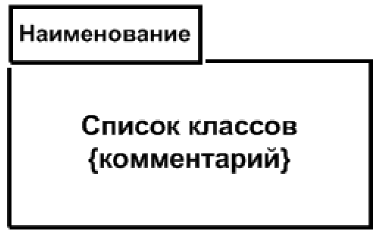


Рис. 3.13. Пакет

Между пакетами определена связь зависимости со стереотипом import. Один пакет может заимствовать классы с атрибутом глобального доступа из другого пакета. Импортировать каскадно классы нельзя. Импорт возможен только между двумя пакетами. Пример деления системы на пакеты показан на рис. 3.14.

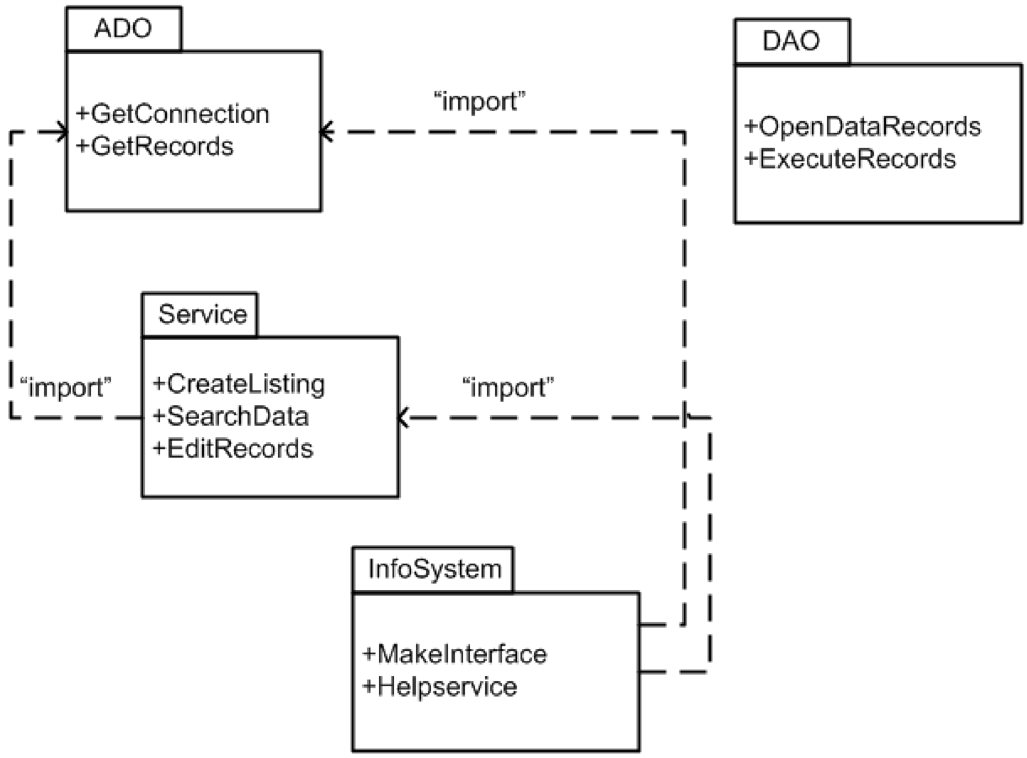


Рис. 3.14. Система с пакетами

При моделировании системы с помощью пакетов можно использовать стереотипные UML пакеты:

- facade (фасад) – определяет пакет, являющийся представлением другого пакета;
- framework (каркас) – расширяемый пакет, содержащий определенный набор готовых решений для предметной области;
- stub (заглушка) – заместитель другого пакета;
- subsystem (подсистема) – независимая часть моделируемой системы;
- system (система) – пакет, представляющий собой всю моделируемую систему.

Моделирование поведения систем

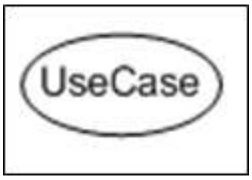
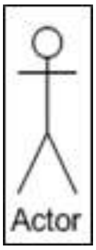
Поведение системы моделируется на базе трех диаграмм:

- прецедентов (Use Case Diagram);
- активности (Activity Diagram);
- взаимодействия (Sequence Diagram, Collaboration Diagram).

Прецедент – множество последовательностей действий для получения инициатором действий определенного результата. Инициатор действий это актер. Актер – связанное множество ролей, которые реализуются пользователями системы. Элементы диаграммы прецедентов показаны в табл. 3.3.

Таблица 3.3

Элементы диаграммы использования

Элемент	Описание
	Прецедент
	Инициатор действия

Между прецедентами устанавливается связь типа «зависимость» с указанием определенных стереотипов. В UML используется два стереотипа:

- include – включение одного прецедента в другой;
- extend – расширение прецедента.

Стереотип Include означает автоматическое исполнение связанного с ним прецедента. Стереотип extend – возможные дополнительные варианты работы пользователя с системой. Пример диаграммы прецедентов показан на рис. 3.15.

Диаграммы активности

Диаграммы активности моделируют динамический аспект поведения системы. Диаграмма документирует деятельность системы. Деятельность реализует определенную функцию системы, состоит из набора состояний:

- состояния действия;
- состояния деятельности.

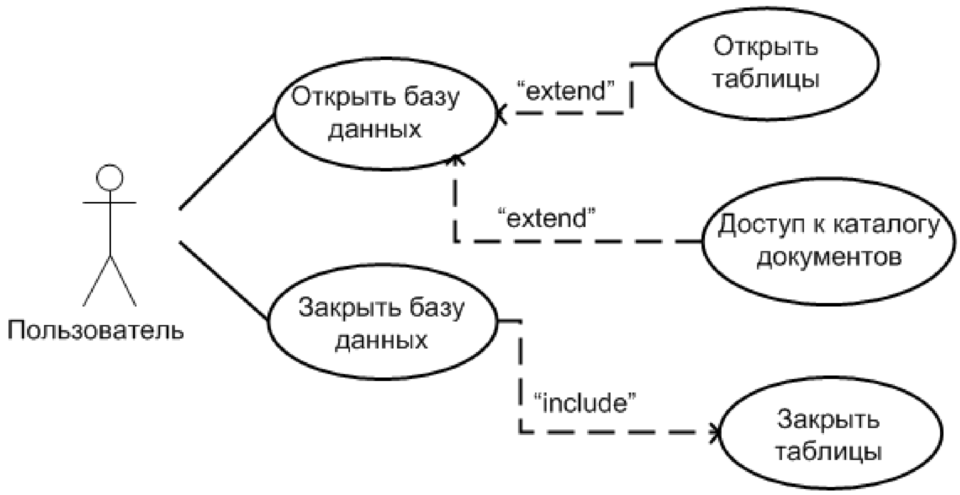
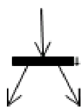
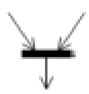
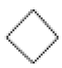
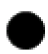


Рис. 3.15. Работа пользователя с системой

Состояния действия считаются атомарными. Они выполняются за малый промежуток времени и не подлежат дальнейшей декомпозиции. Элементы диаграммы приводятся в табл. 3.4.

Тблица 3.4

Элементы диаграммы активности

Элемент	Описание
	Расходящаяся последовательность состояния
	Сходящаяся последовательность состояний
	Состояние деятельности
	Блок принятия решений
	Начало последовательности состояний
	Завершение последовательности состояний

Пример диаграммы активности показан на рис. 3.16.

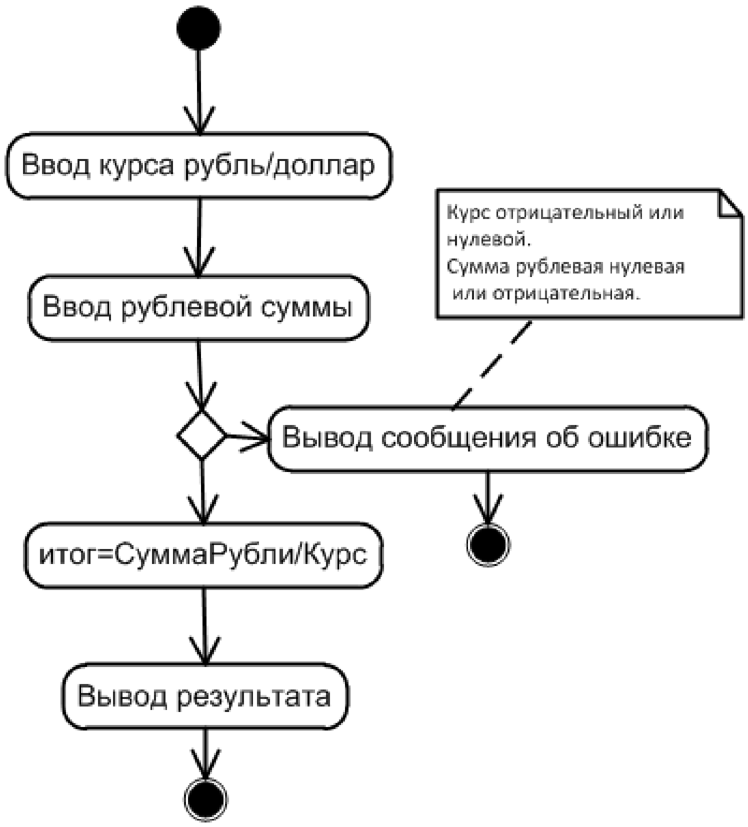


Рис. 3.16. Пересчет рублевой суммы в валюту

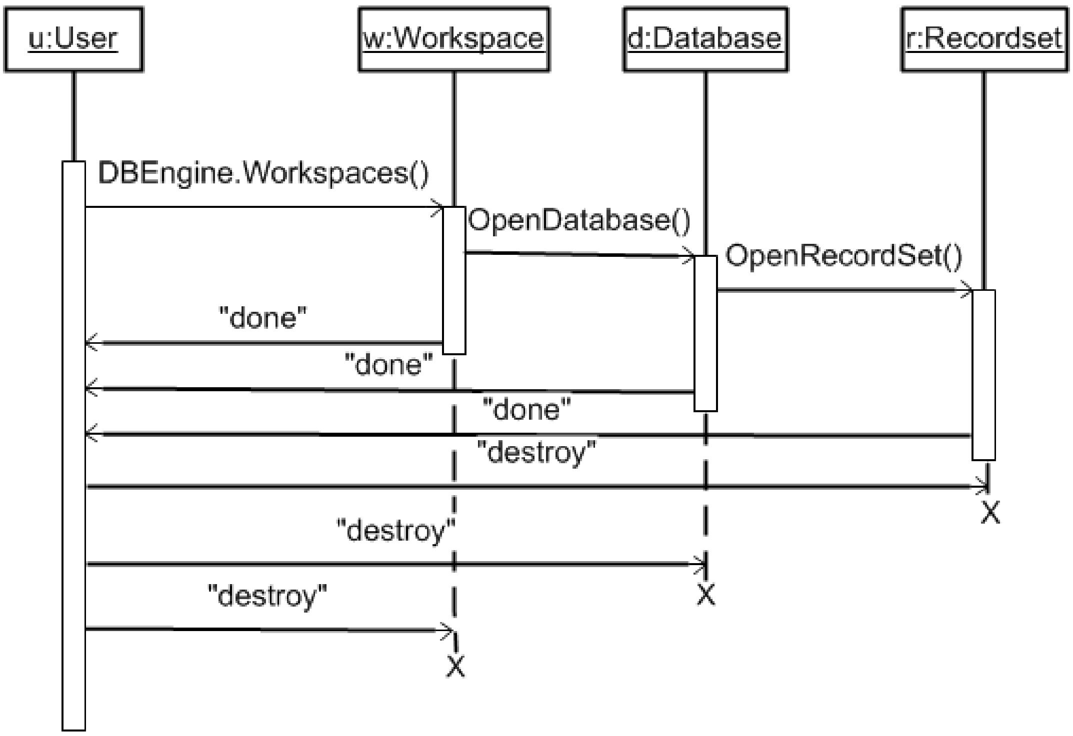


Рис. 3.17. Диаграмма последовательностей:
destroy – сигнал для уничтожения объекта;
done – сигнал, означающий, что объект выполнил требуемые от него функции

Диаграмма последовательностей и кооперации

Диаграмма последовательностей отображает процесс движения и возникновения сообщений в системе во времени. Условно по оси *X* показываются объекты в порядке их появления в системе. По оси *Y* отображается пунктирная линия «жизни» объекта до момента его уничтожения. Период жизни на линии жизни показывается прямоугольником.

Диаграмма кооперации отражает структурную организацию объектов порождающих и принимающих сообщения. Под сообщением понимается некоторая совокупность правил обмена данными между объектами. Сообщение предполагает, что объект, получив некоторую информацию, ответит на нее некоторой последовательностью действий.

При указании ассоциативных связей они могут снабжаться стереотипами local (локальная связь) и global (глобальная связь).

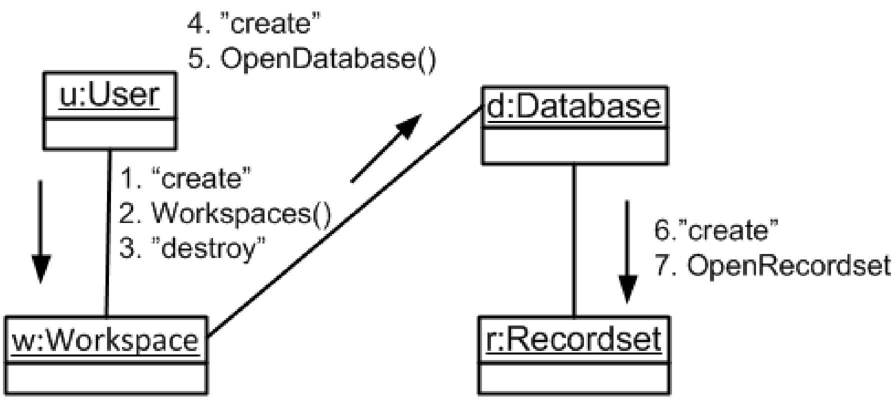


Рис. 3.18. Диаграмма кооперации

Локальная связь позволяет обмениваться сообщениями объектам, которые образуют пару источник – приемник. Глобальная связь определяет возможность передачи сообщения к любому объекту кооперации.

Сообщения, возникающие в кооперации, должны быть пронумерованы. Нумерация выполняется в возрастающем порядке. Сообщение с большим номером появляется позже во времени, чем сообщение с меньшим номером.

Архитектурное моделирование

Архитектурное моделирование предполагает использование двух видов диаграмм:

- диаграмма компонентов (Component Diagram);
- диаграмма развертывания (Deployment Diagram).

Данные диаграммы позволяют наглядно показать развертывание программного кода модели системы на ЭВМ. Строятся такие диаграммы с использованием трех элементов

- компонентов;

- узлов;
- интерфейсов.

Условные обозначения приводятся в табл. 3.5.

Таблица 3.5

Элементы диаграммы компонентов

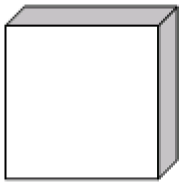
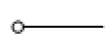
Элемент	Описание
	Компонент
	Узел вычислительной системы
	Интерфейс

Диаграмма компонентов

Компонент представляет собой физическую часть системы. В компьютерных информационных системах компонент это набор битов, который обрабатывается процессором. Компонент может представлять собой исполняемый программный код, либо электронный вариант документа. Описание компонентов приводятся в табл. 3.6.

Таблица 3.6

Описание компонентов

Компонент	**Описание**
Компонент развертывания	Представляет собой подключаемую динамическую библиотеку, либо исполняемую программу
Рабочий продукт	Файл с текстом программного кода, файл с исходными данными
Компонент исполнения	Результат работы системы. Например: новый объект, документ с результатами работы программы

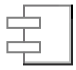
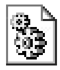
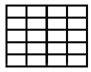
При показе компонентов могут быть использованы стереотипы UML:

- executable – исполняемый программный код;
- library – библиотека объектов;
- table – таблица базы данных;
- file – документ с текстом программы, данными;
- document – любой документ, полученный в результате работы программы.

Условные обозначения стереотипов приводятся в табл. 3.7.

Таблица 3.7

Пиктограммы стереотипов

Стереотип	Условное обозначение
executable	
library	
table	
file	
document	

Компоненты обычно реализуют некоторый интерфейс. Интерфейс это набор операций, которые описывают услуги, предоставляемые данным компонентом. На концептуальном уровне интерфейс реализуется определенным классом. После компилирования программы реализация интерфейса выполняется программным кодом – компонентом. Пример диаграммы компонентов показан на рис. 3.19.

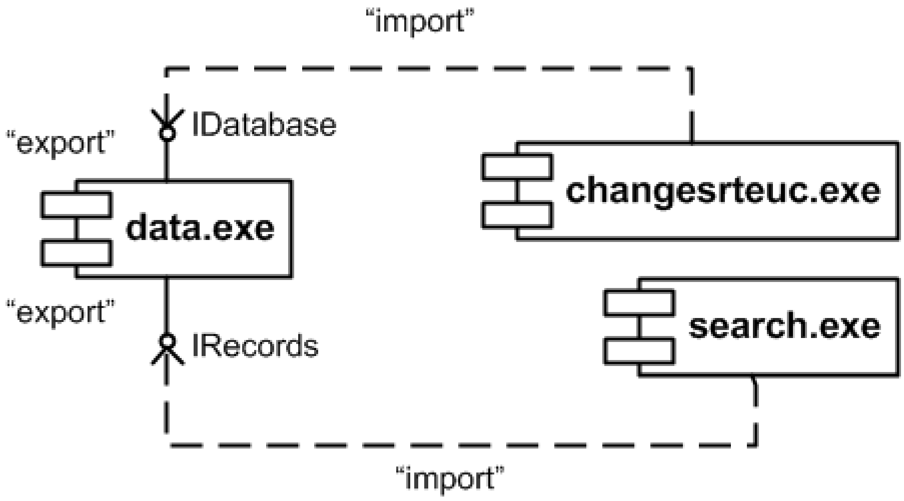


Рис. 3.19. Диаграмма компонентов

Компонент, предоставляющий интерфейс, связан с ним связью типа export, компонент использующий интерфейс связан с ним зависимостью типа import. Выделим следующие особенности компонентов:

- компонент представляет собой код;
- компонент может быть заменен другим, если он совместим с интерфейсом;

- компонент может экспортировать несколько интерфейсов;
- компонент это часть системы, взаимодействующая с остальными через набор интерфейсов.

Диаграмма развертывания

Диаграммы этого типа служат для моделирования развертывания программного обеспечения модели или моделируемой системы с учетом использования вычислительных ресурсов. Под вычислительными ресурсами понимается процессор вычислительной системы. Процессоры могут быть разделены с помощью общей среды передачи данных и представляют собой узлы обработки данных.

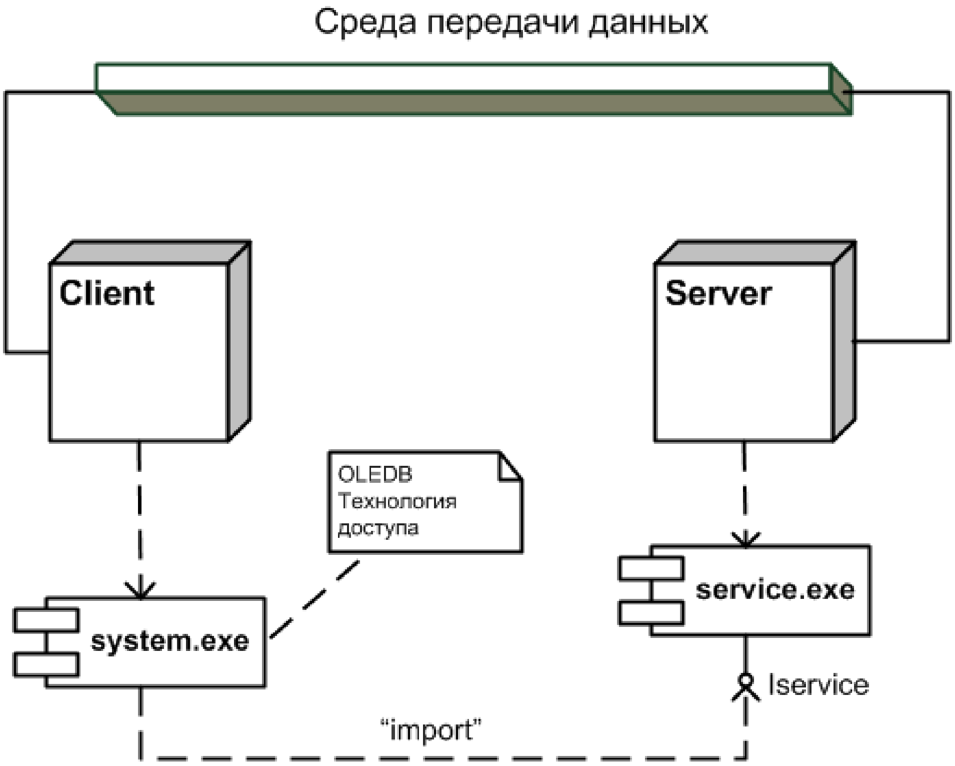


Рис. 3.20. Диаграмма развертывания

На диаграмме узел – это аппаратный элемент системы, обладающей памятью, и процессором для обработки данных. Узел должен обладать уникальным идентификатором – именем, который реализует выполнение компоненты. На диаграммах компонента связана с узлом с помощью отношения зависимости (рис. 3.19).

Контрольные вопросы

1. Как показывают класс и его элементы?
2. Перечислите и охарактеризуйте виды ассоциации?
3. Что такое обобщение и зависимость между классами?
4. Какие виды объектов можно выделить при моделировании предметной области модели?

5. Для чего используются статические элементы классов?
6. Для чего используются пакеты?
7. Перечислите правила построения диаграмм прецедентов.
8. Перечислите правила построения диаграмм активности.
9. Перечислите правила построения диаграмм взаимодействия.
10. Перечислите правила построения диаграмм развертывания и компонентов системы.

ЛЕКЦИЯ №4. СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ

Системы массового обслуживания (СМО) и Q схемы, Марковские цепи и процессы, определение параметров СМО.

В данной лекции использованы положения теории массового обслуживания, в разработку которой большой вклад внесли такие ученые и исследователи как Колмогоров А.Н., Марков А.А., Гнеденко Б.И., Советов Б.Я., Осипов Л.А. [1, 5, 11].

Системы массового обслуживания относят к непрерывно-стохастическим системам (*Q*-схемам). Для таких систем характерно случайное появление требований (заявок) на обслуживание. Завершение обслуживания происходит в случайные моменты времени. Основу системы массового обслуживания составляет канал обслуживания, показанный на рис. 4.1, где приняты условные обозначения:

- H_i – накопитель;
- u_i – поток обслуживание;
- w_i – поток заявок;
- K_i – канал обслуживания заявки.

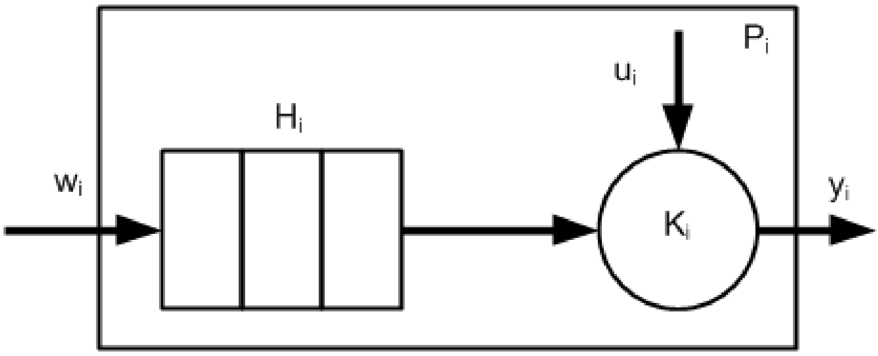


Рис. 4.1. Канал обслуживания

При моделировании реальных систем используют композиции из отдельных элементарных систем, которые называют *Q*-схемами. Если каналы отдельных приборов соединены параллельно, то имеет место многоканальное обслуживание. Если приборы и их параллельные композиции соединены последовательно, то имеет место многофазное обслуживание. Различают разомкнутые и замкнутые *Q*-схемы.

В разомкнутых Q-схемах обратные связи отсутствуют. В замкнутых Q-схемах есть обратные Q-связи, по которым заявки двигаются в направлении, обратном направлению «вход–выход».

Правило обслуживания жестко определяет, в какой последовательности будут обслужены заявки системы. К ним относятся:

- FIFO (FCFS) First In, First Out (First Come, First Served). Обслуживание выполняется в порядке поступления заявок;
- LIFO (LCFS) Last In, First Out (Last Come, First Served). Обслуживание выполняется в обратном порядке прибытия заявок;
- SIRO Selection in Random Order. Очередная заявка для обслуживания выбирается случайным образом;
- Non-preemptive priority. Относительный приоритет, некоторым заявкам отдается предпочтение по отношению к другим клиентам. Текущий процесс обслуживания не прерывается;
- Preemptive priority. Абсолютный приоритет. Если вновь прибывающая заявка обладает высшим приоритетом по сравнению с другими клиентами в системе, текущий процесс обслуживания прерывается и продолжается с новыми условиями. Прежние условия-требования отклоняются;
- RR (Round Robin). Каждая заявка может занять обслуживающий процесс на определенный временной интервал. Клиенты, чье обслуживание требует больше времени, должны несколько раз ставиться в очередь на обслуживание.

Поток событий – заявок

Плотность вероятности потока событий, при среднем значении a, выражается формулой

$$f(x) = \frac{1}{a} e^{-\frac{x}{a}} \text{ при } x > 0;\ f(x) = 0 \text{ при } x < 0, \quad (4.1)$$

где a – коэффициент плотности потока событий.

При моделировании СМО важную роль играет понятие простейшего потока событий. К такому потоку событий предъявляются следующие три требования:

- экспоненциальный закон распределения времени обслуживания и интервалов поступления заявок;
- стационарность (характеристики процесса не должен иметь существенных изменений во времени);
- ординарность (в один момент времени не должно быть более одного события);
- без последействия (будущий процесс зависит только от состояния системы в данный момент и не зависит от пути прихода в это состояние).

По последнему требованию процесс определяют как Марковский (более точно его можно назвать процессом без предыстории). Поток событий, удовлетворяющий перечисленным выше условиям, называется простейшим или Пуассоновским.

Описание систем массового обслуживания

Компактное описание СМО выполняют, используя нотацию Кендала: $A/B/m/n$.

Указывается распределение времени поступления заявок в систему; распределение времени, затрачиваемое на обслуживание заявки в процессоре СМО. Кроме этого указывается число процессоров системы, число мест в очереди ожидания. При этом используют следующие условные обозначения:

- A: закон распределения временных интервалов поступления заявок в систему;
- B: закон распределения временных интервалов в процессоре обслуживания;
- m: число процессоров в системе;
- n: число мест ожидания в очереди;
- M: экспоненциальное распределение;
- E_k: распределение Эрланга с параметром k (k = 1, 2, …);
- D: детерминированное распределение, постоянные временные интервалы;
- G: произвольное – неизвестное распределение.

Рассмотрим пример условного обозначения канала обслуживания: $M/D/3/5$. Данный канал состоит из трех процессоров. В очереди может размещаться пять заявок. Временные интервалы поступления заявок распределены экспоненциально. Временные интервалы обслуживания постоянные.

Построение формальной модели обслуживания может быть выполнено с помощью Марковских цепей и процессов.

Марковская цепь

Цепь Маркова описывает развитие системы во времени. При этом рассматривают состояния системы. Рассматривается текущее состояние системы для перехода в следующее состояние и при этом не учитываются предыдущие состояния системы.

Марковская цепь это последовательность случайных переменных в вероятностном пространстве:

$$(\Omega, P) : X_n : \Omega \to S. \tag{4.2}$$

Множество S называется пространством состояния Марковской цепи. Это пространство может быть конечным набором состояний или бесконечным.

$$S = \{1, 2, ..., m\} \text{ или } S = \{1, 2, 3, ...\}.$$

При этом справедливы следующие соотношения:

$$P(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_0 = x_0). \tag{4.3}$$

Условная вероятность события будет иметь вид:

$$P(X_n = x_n \mid X_{n-1} = x_{n-1}). \qquad (4.4)$$

Это вероятность перехода из одного состояния в другое, она не зависит от вероятностей предыдущих событий. Вероятность перехода из состояния X_{n-1} при временном интервале n–1 в состояние X_n при временном интервале n не зависит от предыдущих состояний.

Марковские цепи, в которых вероятность перехода из одного состояния в другое остается постоянной, называют гомогенными. В этом случае рассматривают только вероятность перехода из состояния X в состояние Y и не рассматривают временные интервалы.

$$p_{xy} = P(X_n = y \mid X_{n-1} = x). \qquad (4.5)$$

Вероятность смены состояния определяет матрица переходов. Матрицу переходов можно определить только для конечного пространства состояния S Марковской цепи

$$S = \{1, 2, ..., m\}.$$

В конечном пространстве состояний могут быть выделены m^2 вероятностей перехода p_{ij}.

$$\Pi = (p_{ij}) = \begin{vmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{vmatrix} \text{ при } 1 \le i,\ j \le m. \qquad (4.6)$$

Если рассматриваемая система находится в состоянии i, то она переходит в произвольное состояние j. Сумма вероятностей строки переходной матрицы Π равна единице.

$$\sum_{j=1}^{m} p_{ij} = 1 \text{ для всех } i \in S. \qquad (4.7)$$

Если это справедливо для каждой строки матрицы, то матрицу называют стохастической.

Рассмотрим систему с тремя состояниями. Смена состояний показана на рис. 4.2 в виде направленного графа. Этой системе соответствует матрица переходов:

$$\Pi = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}.$$

Система переходит с вероятностью равной одному (гарантированный переход) из состояния 1 в состояние 2, потом из состояния 2 в состояние 3 и наконец, из состояния 3 назад в состояние 1. Переход из одного состояния в другое происходит циклически, каждый переход гарантированно исполняется.

Рассмотрим систему, состоящую из двух подсистем, граф состояния которой показан на рис. 4.3. Этому графу соответствует матрица переходов:

$$\Pi = \begin{vmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \end{vmatrix}.$$

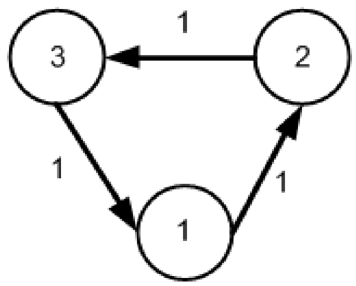


Рис. 4.2. Циклическая система

В зависимости от выбора начального состояния, система остается постоянно в состоянии 2 или циклически переходит из состояния 1 в состояние 3.

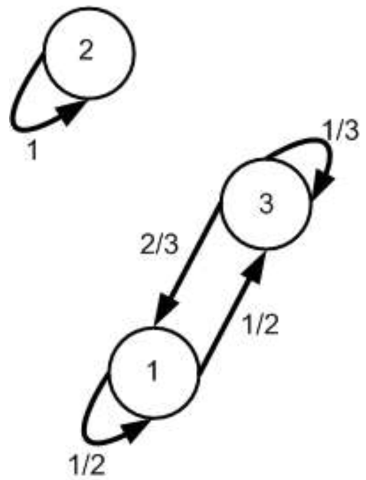


Рис. 4.3. Система с двумя независимыми подсистемами

На рисунке 4.4. приводится другой вариант системы. В системе есть переход из состояния 3 в состояние 2 с заданной вероятностью. Если система попала один раз в это состояние, она его больше не может покинуть. Такое состояние называются поглощающим.

Переходная матрица такой системы имеет вид:

$$\Pi = \begin{vmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/3 & 1/3 & 1/3 \end{vmatrix}.$$

Марковский процесс

В Марковских процессах учитываются временные интервалы, которые могут постоянными и/или переменными, в отличие от дискретных интервалов в Марковских цепях.

Для каждого состояния вводится период пребывания системы в этом состоянии. Так же как и в Марковских цепях, система через определенный период времени переходит в новое состояние. Для СМО считают, что временные интервалы нахождения системы в определенном состоянии подчиняются экспоненциальному закону распределения.

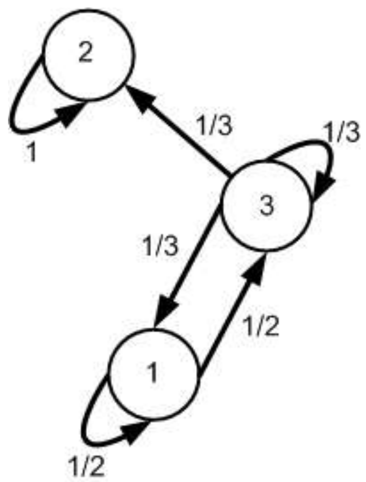


Рис. 4.4. Система с поглощающим состоянием 2

Одноканальная модель процесса обслуживания

В одноканальной СМО имеется один процессор обслуживания и одна очередь требований – заявок. Оценить эффективность работы такой СМО можно с помощью следующих параметров:

- $M[L_q]$ – математическое ожидание числа заявок, ожидающих обслуживания в очереди;
- $M[T_q]$ – математическое ожидание времени, которое заявка проводит в очереди, ожидая обслуживания;
- $M[L]$ – математическое ожидание числа заявок в СМО ожидающих обслуживания;
- $M[T]$ – математическое ожидание времени которое заявка проводит в СМО в ожидании обслуживания.

Эти параметры могут быть определены для СМО типа: $M/M/1/\infty$.

Предполагается, что очередь бесконечная. Все заявки – требования подлежат обслуживанию.

Для такой СМО Марковский процесс примет вид, показанный на рис. 4.5 где λ – интенсивность поступления заявок – требований в очередь, μ – интенсивность обслуживания заявок. Значения этих величин принимаются постоянными.

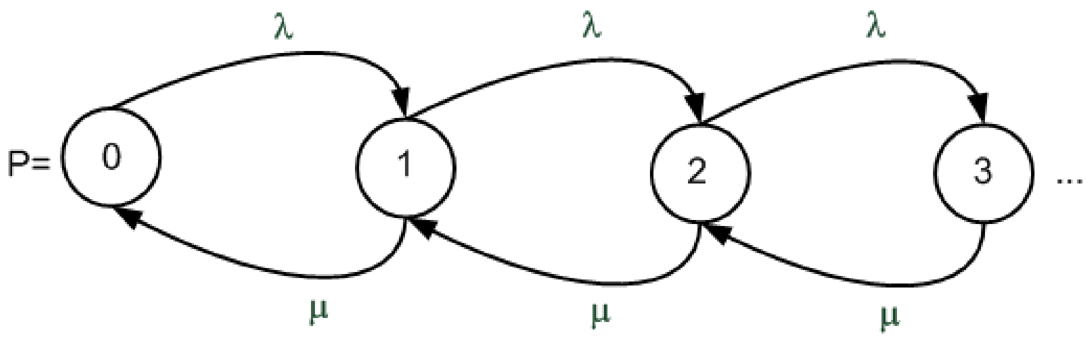


Рис. 4.5. Одноканальная СМО с бесконечной очередью

Работа СМО представляется последовательностью смены состояний с вероятностями p_n, n = 0, 1, 2 … для входа и выхода в состояние. В состоянии P_0 вход будет определен как $P_1 \cdot \mu$, а выход $P_0 \cdot \lambda$ (рис. 4.6).

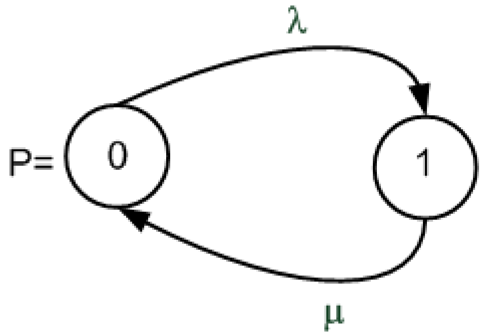


Рис. 4.6. Начальная фаза обслуживания

В итоге получим уравнение:

$$p_1 \cdot \mu = p_0 \cdot \lambda. \tag{4.8}$$

Для остальных состояний получим уравнения:

$$p_0 \cdot \lambda + p_2 \cdot \mu = p_1 \cdot \lambda + p_1 \cdot \mu, \text{ для } p_1; \tag{4.9}$$

$$p_1 \cdot \lambda + p_3 \cdot \mu = p_2 \cdot \lambda + p_2 \cdot \mu, \text{ для } p_2. \tag{4.10}$$

После преобразования получим:

$$-p_0 \cdot \lambda + p_1 \cdot \mu = 0;$$

$$p_0 \cdot \lambda - p_1 \cdot (\lambda + \mu) + p_2 \cdot \mu = 0;$$

$$p_1 \cdot \lambda - p_2 \cdot (\lambda + \mu) + p_3 \cdot \mu = 0;$$

$$\ldots$$

$$p_{i-1} \cdot \lambda - p_i \cdot (\lambda + \mu) + p_{i+1} \cdot \mu = 0; \tag{4.11}$$

$$\Rightarrow \begin{vmatrix} -\lambda & \mu & 0 & 0 \\ \lambda & -(\lambda+\mu) & \mu & 0 \\ 0 & \lambda & -(\lambda+\mu) & \mu \\ 0 & 0 & 0 & \text{и т.д.} \end{vmatrix} \cdot \begin{vmatrix} p_0 \\ p_1 \\ p_2 \\ . \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ 0 \\ . \end{vmatrix}. \tag{4.12}$$

Вероятность перехода из одного состояния в другое вычисляется согласно формуле:

$$p_n = (1-\rho)\cdot\rho^n, \qquad (4.13)$$

где n – номер текущего состояния. Согласно формуле значение вероятности p_0 – отсутствие требований в системе будет равно:

$$p_0 = (1-\rho). \qquad (4.14)$$

В формулах 4.13 и 4.14 $\rho = \lambda/\mu$.

Если в системе находится 0 или 1 заявка (n = 0, n = 1), заявка не будет ожидать обслуживания и сразу будет обработана, длина очереди будет равна 0. Если в СМО поступило две заявки, одна заявка будет обслуживаться, а другая в очереди будет ожидать обслуживания. Длина очереди будет равной 1 заявке. При поступлении трех заявок, одна заявка обслуживается, а две заявки будут ожидать обслуживания и т.д. Следовательно, можно сделать вывод, что при $n \geq 1$ всегда одна заявка будет обслужена, а n – 1 заявок будут ожидать обслуживания. Тогда число заявок в очереди будет равно

$$M[L_q] = \sum_{n=2}^{\infty} p_n \cdot (n-1). \qquad (4.15)$$

С другой стороны это значение может быть определено:

$$M[L_q] = (1-\rho)\cdot\rho\cdot\frac{\rho}{(1-\rho)^2} = \frac{\rho^2}{(1-\rho)}. \qquad (4.16)$$

Остальные параметры могут быть рассчитаны по формулам.

$$M[L] = M[L_\text{q}] + \rho; \qquad (4.17)$$

$$M[L] = \frac{\rho^2}{1-\rho} + \rho;$$
$$(4.18)$$
$$M[L] = \frac{\rho}{1-\rho};$$

$$M[T_\text{q}] = \frac{\lambda}{\mu\cdot(\mu-\lambda)}; \qquad (4.19)$$

$$M[T] = M[T_\text{q}] + \frac{1}{\mu}; \qquad (4.20)$$

$$M[T] = \frac{1}{\mu-\lambda}. \qquad (4.21)$$

В рассмотренной СМО имеют место соотношения:

$$M[L_\text{q}] = \lambda\cdot M[T_\text{q}]; \qquad (4.22)$$

$$M[L] = \lambda\cdot M[T]. \qquad (4.23)$$

Эти зависимости отвечают закону Литтла.

Система M/M/1/K

Общее число возможных состояний в такой системе $n = K + 2$. На рис. 4.7 показаны состояния СМО с очередью, в которой три места:

K = 3, n = 5. Получить систему уравнений для построения Марковского графа можно используя уравнения Колмогорова А.Н.

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t);$$

$$\frac{dp_i(t)}{dt} = \lambda p_{i-1}(t) - (\lambda + \mu) p_i(t) + \mu p_{i+1}(t), i = 1, ..., K; \qquad (4.24)$$

$$\frac{dp_{K+1}(t)}{dt} = \lambda p_M(t) - \mu p_{K+1}(t).$$

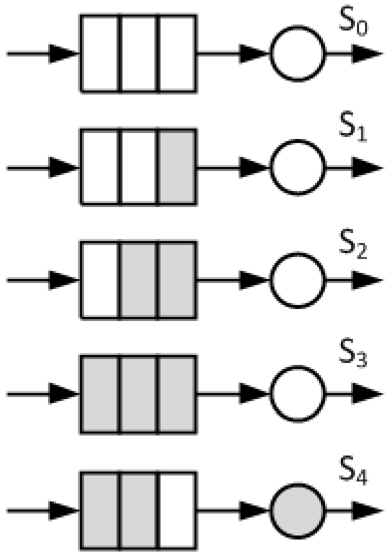


Рис. 4.7. СМО с ограниченной очередью и пять состояний S

Для примера СМО, показанного на рис. 4.7 уравнения примут вид:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) + \mu p_1(t);$$

$$\frac{dp_1(t)}{dt} = \lambda p_0(t) - (\lambda + \mu) p_1(t) + \mu p_2(t);$$

$$\frac{dp_2(t)}{dt} = \lambda p_1(t) - (\lambda + \mu) p_2(t) + \mu p_3(t); \qquad (4.25)$$

$$\frac{dp_3(t)}{dt} = \lambda p_2(t) - (\lambda + \mu) p_3(t) + \mu p_4(t);$$

$$\frac{dp_4(t)}{dt} = \lambda p_3(t) - \mu p_4(t).$$

В установившемся режиме работы СМО производные будут равны нулю и система дифференциальных уравнений будет соответствовать уравнениям 4.12.

Уравнения для определения вероятностей примут вид:

$$p_n = \rho^n p_0; \qquad (4.26)$$

$$p_0 = \frac{1}{\sum_{n=0}^{K} \rho^n}; \tag{4.27}$$

$$p_0 = \begin{cases} \dfrac{1-\rho}{1-\rho^{K+1}};\ \rho \neq 1; \\ \dfrac{1}{K+1};\ \rho = 1; \end{cases} \tag{4.28}$$

$$p_n = \begin{cases} \dfrac{(1-\rho)\cdot\rho^n}{1-\rho^{K+1}};\rho \neq 1; \\ \dfrac{1}{K+1};\rho = 1. \end{cases} \tag{4.29}$$

Число заявок в системе вычисляют по формулам:

$$\begin{cases} M[L] = \dfrac{\rho[1-(K+1)\rho^K + K\rho^{K+1}]}{(1-\rho^{K+1})\cdot(1-\rho)};\ \rho \neq 1; \\ M[L] = \dfrac{K}{2};\ \rho = 1. \end{cases} \tag{4.30}$$

Что бы вычислить время нахождения заявки в канале обслуживания нужно ввести приведенное значение интенсивности поступления заявок.

$$\lambda' = \lambda(1-p_K). \tag{4.31}$$

Время нахождения в системе будет вычислено:

$$M[T] = \frac{1}{\lambda'} M[L]. \tag{4.32}$$

Время нахождения заявки в очереди и длина очереди вычисляются по формулам:

$$M[T_q] = M[T] - \frac{1}{\mu}; \tag{4.33}$$

$$M[L_q] = M[L] - \rho\left(\frac{1-\rho^K}{1-\rho^{K+1}}\right); \tag{4.34}$$

$$M[L_q] = M[L] - (1-p_0).$$

Система M/M/m/∝

Такая СМО обладает неограниченной очередью и несколькими процессорами обслуживания:

$$\lambda_n = \lambda, \text{ для n} = 0, 1, 2, ...;$$

$$\mu_n = \begin{cases} n\mu;\ 0 \leq n \leq m; \\ m\mu;\ n > m. \end{cases}$$

Вероятности состояний такой системы вычисляются по формулам:

$$p_n = \begin{cases} \prod_{i=1}^{n} \frac{\lambda}{i\mu} p_0 = \frac{1}{n!}\rho^n p_0; \ 0 \le n \le m; \\ p_m \prod_{i=m+1}^{n} \frac{\lambda}{m\mu} = \frac{1}{m!m^{n-m}}\rho^n p_0; \ n > m; \end{cases} \quad (4.35)$$

$$p_0 = (1 + \sum_{n=1}^{m-1} \frac{1}{n!}\rho^n + \frac{1}{m!}\rho^m (\frac{1}{1-\frac{\rho}{m}}))^{-1}. \quad (4.36)$$

Длину очереди, время нахождения заявки в очереди, время нахождения заявки в системе и число заявок в системе определяют по формулам:

$$M[L_q] = \frac{\frac{\rho^{m+1}}{m}}{m!(1-\frac{\rho}{m})^2} p_0; \quad (4.37)$$

$$M[T_q] = M[L_q] \cdot \frac{1}{\lambda} = \frac{\frac{\rho^{m+1}}{m}}{\lambda m!(1-\frac{\rho}{m})^2} p_0; \quad (4.38)$$

$$M[T] = M[T_q] + \frac{1}{\mu} = \frac{\frac{\rho^{m+1}}{m}}{\lambda m!(1-\frac{\rho}{m})^2} p_0 + \frac{1}{\mu}; \quad (4.39)$$

$$M[L] = \lambda \cdot M[T] = \frac{\frac{\rho^{m+1}}{m}}{m!(1-\frac{\rho}{m})^2} p_0 + \rho. \quad (4.40)$$

С помощью формулы Эрланга второго рода вычисляют вероятность ожидания обслуживания клиентом:

$$p_r = \sum_{n=m}^{\infty} p_n = \sum_{n=m}^{\infty} \frac{1}{m!m^{n-m}} \rho^n p_0 = \\ = \frac{1}{m!}\rho^m \frac{1}{1-\frac{\rho}{m}} p_0 \quad (4.41)$$

Система *M*/*M*/*m*/*K*

В такой системе m процессоров обслуживания и K мест в очереди. Определить вероятность отсутствия заявок в системе можно по формулам:

$$p_0 = \begin{cases} \left(1 + \sum_{n=1}^{m-1} \frac{1}{n!}\rho^n + \frac{1}{m!}\rho^m \frac{1 - (\frac{\rho}{m})^{K-m+1}}{1 - \frac{\rho}{m}}\right)^{-1}; \ \frac{\rho}{m} \neq 1; \\ \left(1 + \sum_{n=1}^{m-1} \frac{1}{n!}\rho^n + \frac{1}{m!}\rho^m (K - m + 1)\right)^{-1}; \ \frac{\rho}{m} = 1. \end{cases} \quad (4.42)$$

Остальные параметры системы определяют по формулам:

$$M[L_q] = \frac{\rho^m (\frac{\rho}{m})}{m!(1 - \frac{\rho}{m})^2} p_0 \left[1 - (\frac{\rho}{m})^{K-m+1} - (1 - \frac{\rho}{m}) \cdot (K - m + 1) \cdot (\frac{\rho}{m})^{K-m}\right]; \quad (4.43)$$

$$\lambda' = \lambda(1 - p_K);$$

$$M[Tq] = \frac{1}{\lambda'} M[L_q]; \quad (4.44)$$

$$M[T] = M[T_q] + \frac{1}{\mu}; \quad (4.45)$$

$$M[L] = \lambda' M[T]. \quad (4.46)$$

Если в систему поступает число клиентов равное число процессоров $K = m$ для вероятностей можно использовать следующие соотношения:

$$p_n = \frac{1}{n!}\rho^n p_0; \quad (4.47)$$

$$p_0 = \sum_{n=0}^{m} \frac{1}{n!}\rho^n. \quad (4.48)$$

В противном случае вероятности будут равны:

$$p_n = \frac{\frac{\rho^n}{n!}}{\sum_{i=0}^{m} \frac{\rho^i}{i!}}; \ 0 \le n \le m. \quad (4.49)$$

Вероятность полной загруженности системы (заполненная очередь, все процессоры обслуживают заявки):

$$p_K = p_m = \frac{\frac{\rho}{m!}}{\sum_{i=0}^{m} \frac{\rho^i}{i!}}. \quad (4.50)$$

Контрольные вопросы

1. Из каких элементов состоит канал обслуживания?
2. Перечислите и охарактеризуйте правила обслуживания заявок.

3. Какие требования предъявляются к Пуассоновскому потоку событий?

4. Как используется нотация Кендала для описания канала обслуживания.

5. Что такое Марковская цепь?

6. В чем разница между Марковской цепью и процессом?

7. Как строится переходная матрица Марковского процесса?

8. Перечислите основные параметры канала обслуживания.

9. Как строится Марковский граф для канала обслуживания?

10. Как записываются уравнения Колмогорова А.Н. для канала обслуживания?

ЛЕКЦИЯ №5. СХЕМЫ МОДЕЛИРОВАНИЯ

5.1. Обзор современных схем моделирования

Сети Петри, автоматы, описание работы автомата, автомат Мили, автомат Мура, конечные детерминированные и недетерминированные автоматы, агрегатные модели.

Сети Петри (*N*-схемы)

В практике моделирования объектов часто приходится решать задачи, связанные с формализованным описанием и анализом причинно-следственных связей в сложных системах, где одновременно параллельно протекает несколько процессов [5].

Сеть Петри представляет собой двухдольный направленный граф. Позиции и переходы соединяются между собой направленными дугами. Позиции обозначаются окружностями, переходы показываются в виде прямоугольников (рис. 5.1).

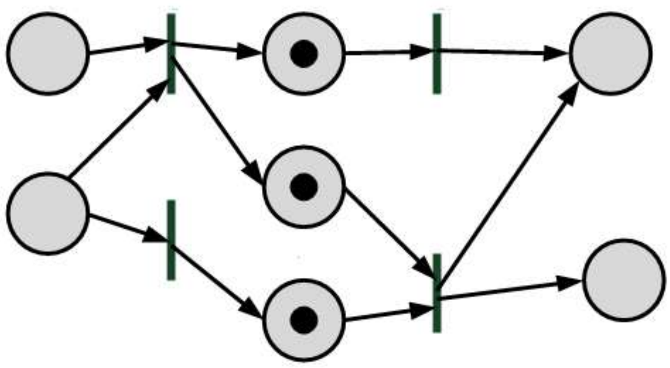


Рис. 5.1. Пример сети Петри

Позиция обладает определенной мощностью и в соответствии с мощностью может содержать несколько знаков (маркеров). Если значение мощности не задано, то оно принимается равной бесконечности или единице. Загрузка позиций маркерами, называется маркировкой и представляет собой состояние сети. Каждой дуге графа может быть поставлен в соответствие определенный «вес». Если это значение не указано, то принимается значение равное единице.

Переход считается активным или готовым к активации если во всех входных позициях находится такое число маркеров, которое соответствует весу перехода и все выходные позиции обладают достаточной мощностью – емкостью для того чтобы принять новые маркеры. В сети Петри маркеры не движутся по графу. Они удаляются и появляются в определенных позициях.

Маркированная (размеченная) *N-схема* задается кортежем [5]:

$$N = (B, D, I, O, M), \tag{5.1}$$

B – конечное множество символов – позиций; D – конечное множество символов – переходов; I – входная функция (прямая функция инцидентности); O – выходная функция (обратная функция инцидентности);

$$B \neq \varnothing;$$

$$D \neq \varnothing,\ B \cap D \neq \varnothing;$$

$$I = B \times D \to \{0,1\}; \tag{5.2}$$

$$O : D \times B \to \{0,1\}; \tag{5.3}$$

$$b_i \in D(d_j);$$

$$b_i \in I(d_j).$$

Входная функция I отображает переход d_j, в множество входных позиций, а выходная функция O отображает переход d_j, в множество выходных позиций.

Рассмотрим фазы работы сети Петри с тремя переходами, двумя позициями. Мощность позиции два маркера. Обозначим M – множество маркировок. Работа сети требует смены семи фаз {0…6}. Состояние сети показано на рис. 5.2–5.4.

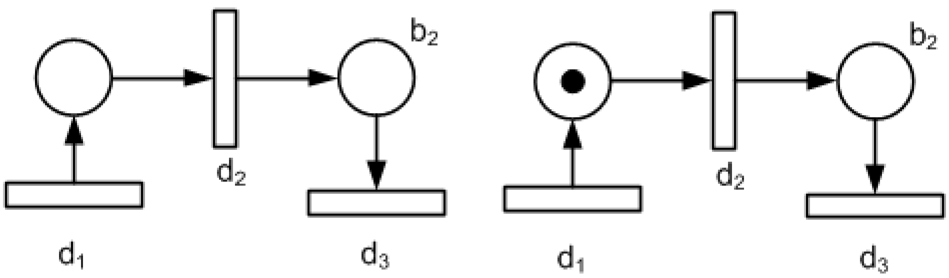


Рис. 5.2. Фаза 0 и фаза 1; M_0 = {0,0}, M_1 = {1,0}, активный переход d_1

Основными свойствами сети Петри являются:

- ограниченность – число меток в любой позиции сети не может превысить некоторого значения K;
- безопасность – частный случай ограниченности, K = 1;

• сохраняемость – постоянство загрузки ресурсов, величина постоянна. $\sum A_i N_i$, где N_i – число маркеров в i-ой позиции, A_i – весовой коэффициент;

• достижимость – возможность перехода сети из одного заданного состояния (характеризуемого распределением меток) в другое;

• живучесть – возможность срабатывания любого перехода при функционировании моделируемого объекта.

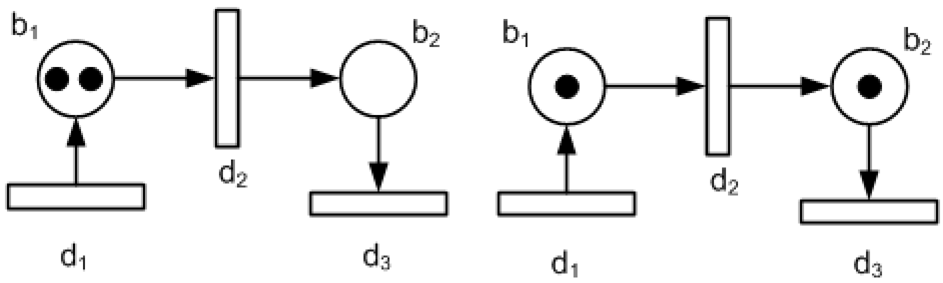


Рис. 5.3. Фаза 2 и фаза 3; M_2 = {2,0}, d_1; M_3 = {1,1}, d_2

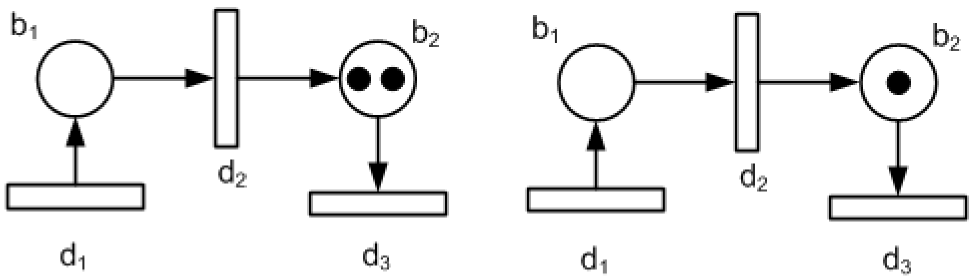


Рис. 5.4. Фаза 4 и фаза 5, фаза 6; M_4 = {0,2}, d_2; M_5 = {0,1}, d_3; M_6 = {0,0}, d_3

Дискретно-детерминированные модели (*F*-схемы)

Модели этого типа представляют собой автоматы. Автомат функционирует в моменты автоматного времени $t_0 = 0$, $t_1 = T_0$, $t_2 = 2T_0$

$$t_i = i \cdot T_0;\ i = \overline{0,\infty}.$$

Здесь T_0 период дискретизации модели автомата. В каждый момент $t_i \in T$, автомат может находиться в одном из конечного числа состояний $z_j \in Z$. Принято различать два типа автоматов [5]:

• автомат Мили;

• автомат Мура.

Автомат Мили

Автомат может быть задан в виде кортежа:

$$A = (Q, \Sigma, \Omega, \delta, \lambda, q_0, F). \qquad (5.4)$$

Q является множеством возможных состояний. Вместо Q часто используется также условное обозначение в виде латинской буквы Z. Множество Σ является алфавитом ввода, множество Ω является алфавитом выхода.

Переходная функция имеет вид

$$\delta\colon\ Q \times \Sigma \to Q. \tag{5.5}$$

Функция выхода определяется выражением

$$\lambda\colon Q \times \Sigma \to \Omega. \tag{5.6}$$

В конечном итоге выбирается компактная нотация и обе функции сводятся к одной переходной функции состояния

$$\zeta\colon\ Q \times \Sigma \to \Omega \times Q. \tag{5.7}$$

Стартовое-начальное состояние: $q_0 \in Q$. Вместо q_0 могут быть использованы условные обозначения начального состояния z_0 и s_0.

Множество F это множество возможных доступных состояний

$$F \subseteq Q.$$

Схема работы:

$$\begin{aligned} q(t+1) &= \delta(q(t), \sigma(t)); \\ \omega(t) &= \lambda(q(t), \sigma(t)). \end{aligned} \tag{5.8}$$

Автоматы представляют в виде направленного графа. На рис. 5.5 показан пример автомата. Дуги помечены дробью σ_i/ω_i, а переход происходит в зависимости от значения из множества алфавита ввода и значения из множества алфавита вывода. Например, условное обозначение дуги 0/1, означает, что при задании 0 для смены состояний, на выходе появляется сигнал 1.

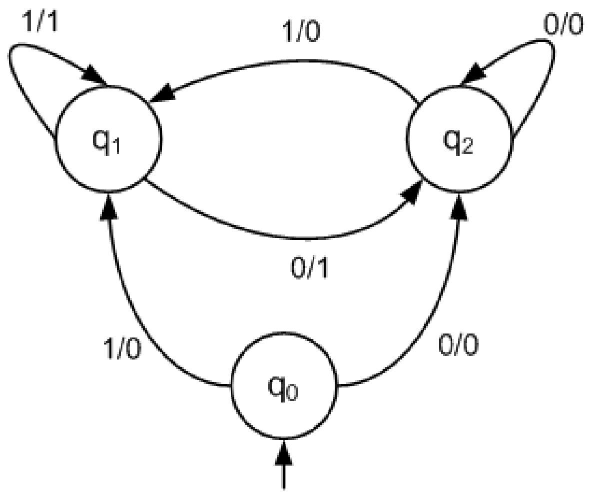


Рис. 5.5. Автомат Мили; $\Sigma = \{0,1\}$; $\Omega = \{0,1\}$

Автомат Мура

Автомат Мура получил свое наименование по имени математика Эдварда Ф. Мура. По сравнению с автоматом Миля вход такого автомата исключительно зависит от его состояния. При достижении определенного состояния формируется выходное значение, которое не зависит от перехода в это состояние.

Автомат описывается кортежем

$$A = (Q, \Sigma, \Omega, \delta, \lambda, q_0, F). \tag{5.9}$$

Здесь Q конечное множество состояний, множество Σ – входной алфавит автомата:

$$|\Sigma| < \infty, Q \cap \Sigma = 0.$$

Множество Ω – выходной алфавит автомата. Переходная функция автомата описывается выражением

$$\delta\colon Q \times \Sigma \to Q; \tag{5.10}$$

функция выхода имеет вид

$$\lambda\colon Q \to \Omega; \tag{5.11}$$

$$\begin{gathered} q(t+1) = \delta(q(t), \sigma(t)); \\ \omega(t) = \lambda(q(t)). \end{gathered} \tag{5.12}$$

Пример автомата Мура показан на рис. 5.6. Известны следующие значения для множеств автомата:

$$Q = \{q_0, q_1, q_2, q_3\},$$

$$\Sigma = \{x, y, z\},$$

$$\Omega = \{a, b, c\}.$$

Для иллюстрации работы автомата составим табл. 5.1. Стрелкой показан процесс смены состояния. При переходе из состояния q_0 в состояние q_1 требуется подать сигнал y, на выходе будет сигнал b. Аналогично можно проследить смену остальных состояний.

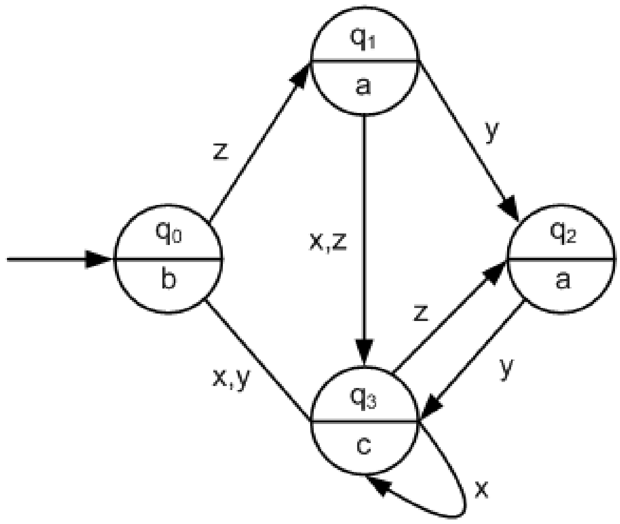


Рис. 5.6. Автомат Мура

Таблица 5.1

Алгоритм работы автомата

Переход (δ)	***x***	***y***	***z***	**Выходное значение (λ)**
$q_0 \to$	q_3	$\downarrow q_3$	q_1	b
q_1	q_3	q_2	q_3	a
q_2	–	q_3	–	a
q_3	q_3	–	q_2	c

Детерминированные и недетерминированные конечные автоматы

Рассмотрим автомат Миля с четырьмя состояниями. Множество входных воздействий автомата $\Sigma = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$, множество выходных сигналов автомата $\Omega = \{\omega_1, \omega_2, \omega_2, \omega_4\}$. Схема смены состояний автомата показана на рис. 5.7. На рисунке под номером 4 обозначено конечное состояние автомата. В зависимости от того какой сигнал будет получен автоматом из множества Σ автомат переходит однозначно в новое состояние, выдавая сигнал из множества Ω. Смена состояний и выдачи выходных сигналов происходят до тех пор пока автомат не перейдет в конечное состояние. Такое поведение автоматов называется детерминированным, а автоматы называют конечными и детерминированными.

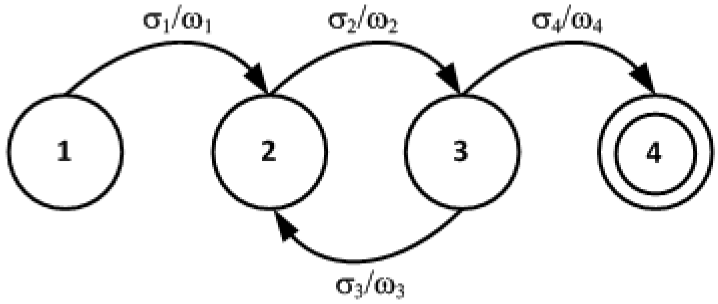
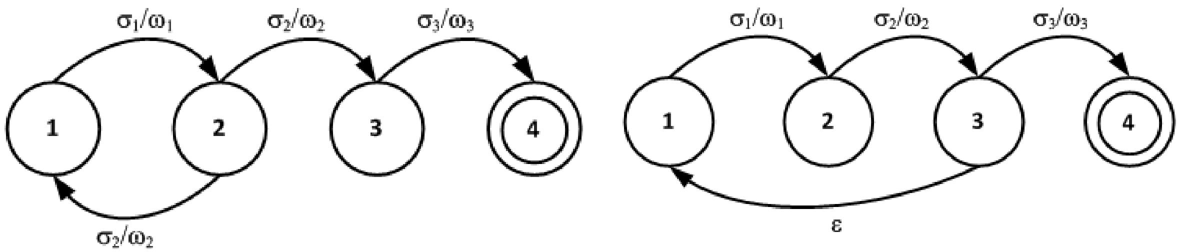


Рис. 5.7. Детерминированный автомат

На рисунке 5.8 показаны схемы поведения не детерминированных автоматов. В недетерминированных автоматах возникает неопределенность перехода из одного состояния в другое. Автомат может перейти с определенной вероятностью из одного состояния в другое и выдать один и тот же сигнал из множества Ω при появлении на входе сигнала из множества Σ.



а) *б)*

Рис. 5.8. Недетерминированный автомат:
а) неоднозначность смены состояний; б) неизвестный входной и выходной сигнал

Другой вариант неопределенности возникает, когда для перехода из одного состояния в другое выбирается с определенной вероятностью сигнал из множества Σ и на выходе появляется с определенной вероятностью сигнал из множества Ω.

Агрегаты (*A*-схемы)

Такие модели требуют определения следующих множеств [5]:

- *T* – множество моментов времени;
- *X* – множество входных сигналов;
- *Y* – множество выходных сигналов;
- *Z* – множество состояний.

Процессы перехода агрегата из одного состояния в другое происходит за малый промежуток времени, т.е. имеет место скачек состояний δz. В качестве элемента *A*-схемы выступает агрегат. Каждый агрегат A_n, состоит: $n = \overline{1, N_A}$. На рис. 5.9 показана схема модели, состоящая из двух агрегатов, на котором окружности – коммутационные точки агрегатов – полюса.

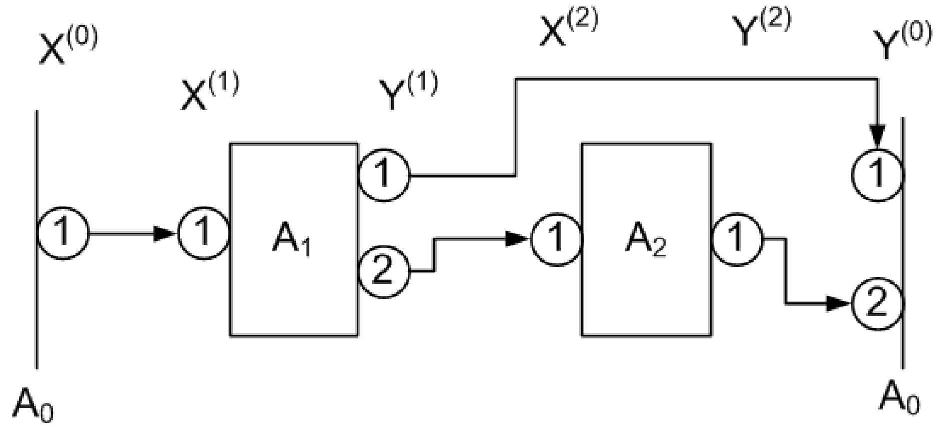


Рис. 5.9. Пример схемы агрегатной модели

Последовательность выходных сигналов, упорядоченную относительно времени выдачи называют выходным сообщением или *Y*-сообщением. Информация, которая циркулирует в *A*-схеме делиться на внутреннюю и внешнюю. Внешняя информация поступает от внешних объектов, а внутренняя вырабатывается агрегатами *A*-схемы. Обмен информации *A*-схемы с внешней средой происходит через агрегаты, которые называют полюсами.

Различают входные полюсы, на которые поступают *X*-сообщения. Выходные полюсы – это агрегаты, выходная информация которых представляет *Y*-сообщения.

Контрольные вопросы

1. Перечислите правила построения сетей Петри.
2. Перечислите элементы сети Петри.
3. Как используется маркировка при анализе работы сети Петри?

4. Дайте формальное описание принципа работы автомата Мили.
5. Как строится граф смены состояний автомата Мили?
6. Дайте формальное описание принципа работы автомата Мура.
7. Как строится граф смены состояний автомата Мура?
8. Приведите пример недетерминированного конечного автомата Мили.
9. Приведите пример недетерминированного конечного автомата Мура.
10. Перечислите правила составления агрегатной модели системы.

5.2. Динамические системы

Непрерывно-детерминированные модели, пространство состояния, регулятор состояния, наблюдатель состояния, настройка регулятора состояния и наблюдателя состояния.

Непрерывно-детерминированные модели (*D*-схемы)

Для описания поведения системы может быть использовано линейное дифференциальное уравнение с постоянными коэффициентами n-го порядка. Такое уравнение называется уравнением «вход-выход» и оно справедливо для системы, показанной на рис. 5.10.

Решение уравнения зависит [2]:

- от входного воздействия $u(t)$;
- от начальных условий.

Уравнение «вход–выход» представляет собой линейное дифференциальное уравнение с постоянными коэффициентами n-го порядка:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \ldots + a_0 y = K \cdot u, \qquad (5.13)$$

где K – коэффициент усиления системы, y – выходной сигнал системы, u – входной сигнал системы, a_n, ..., a_0 – коэффициенты уравнения.

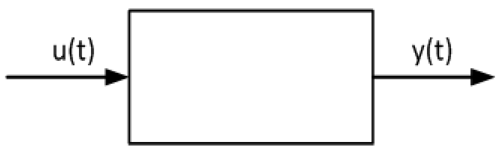


Рис. 5.10. Вход–выход системы

Введем в рассмотрение переменные состояния:

$$x_1 = y;\ x_2 = \frac{dy}{dt}, \ldots,\ x_n = \frac{d^{n-1} y}{dt^{n-1}}.$$

Для более компактного описания производную можно обозначить в виде:

$$\frac{dy}{dt} = y'.$$

С учетом введенных обозначений для переменных состояния уравнение (5.13) может быть представлено в матричном виде [10]:

$$\begin{cases} x' = A \cdot x + b \cdot u, \\ y = c \cdot x + d \cdot u. \end{cases} \tag{5.14}$$

Полученные уравнения – описание системы в пространстве состояний.

Матрица и вектора

$$\underset{(n\times n)}{A},\ \underset{(n\times 1)}{b},\ \underset{(1\times n)}{c},\ \underset{(1\times 1)}{d}.$$

Начальные условия

$$x_1(0),\ ...,\ x_n(0).$$

Вектор состояния

$$x = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}.$$

Матрица системы:

$$A = \begin{bmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \\ -\dfrac{a_0}{a_n} & ... & ... & ... & -\dfrac{a_{n-1}}{a_n} \end{bmatrix}.$$

Вектора коэффициентов входных воздействий:

$$b = \begin{bmatrix} 0 \\ 0 \\ ... \\ \dfrac{K}{a_n} \end{bmatrix};\ c = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \end{bmatrix};\ d = 0.$$

Пространство, координатами которого являются переменные $x_1, x_2, ..., x_n$, называется пространством состояний. Размерность пространства состояний равна порядку системы дифференциальных уравнений.

Если перейти от функции времени к преобразованию по Лапласу, то получают передаточную функцию системы. Для этого введем условное обозначение производной, используя ее изображение по Лапласу:

$$\frac{d}{dt} = s.$$

Получим передаточную функцию системы $W(s)$ на основе выражения (5.14):

$$\begin{cases} sx(s) = A \cdot x(s) + b \cdot u(s); \\ y(s) = c \cdot x(s) + d \cdot u(s). \end{cases}$$

Выполним преобразования:

$$(sE - A) \cdot x(s) = b \cdot u(s);$$

$$x(s) = (sE - A)^{-1} \cdot b \cdot u(s);$$

$$y(s) = [c(sE - A)^{-1} \cdot b + d] \cdot u(s).$$

В результате получим:

$$W(s) = \frac{y(s)}{u(s)} = c(sE - A)^{-1} \cdot b + d, \qquad (5.15)$$

где E – единичная матрица

$$E = \begin{bmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & ... & 1 \end{bmatrix}.$$

Передаточная функция определяет характер преобразования системой входного воздействия u в выходной сигнал y. В соответствии с выражением (5.15) для передаточной функции исследуемую систему можно представить в виде структурной схемы, показанной на рис. 5.11.

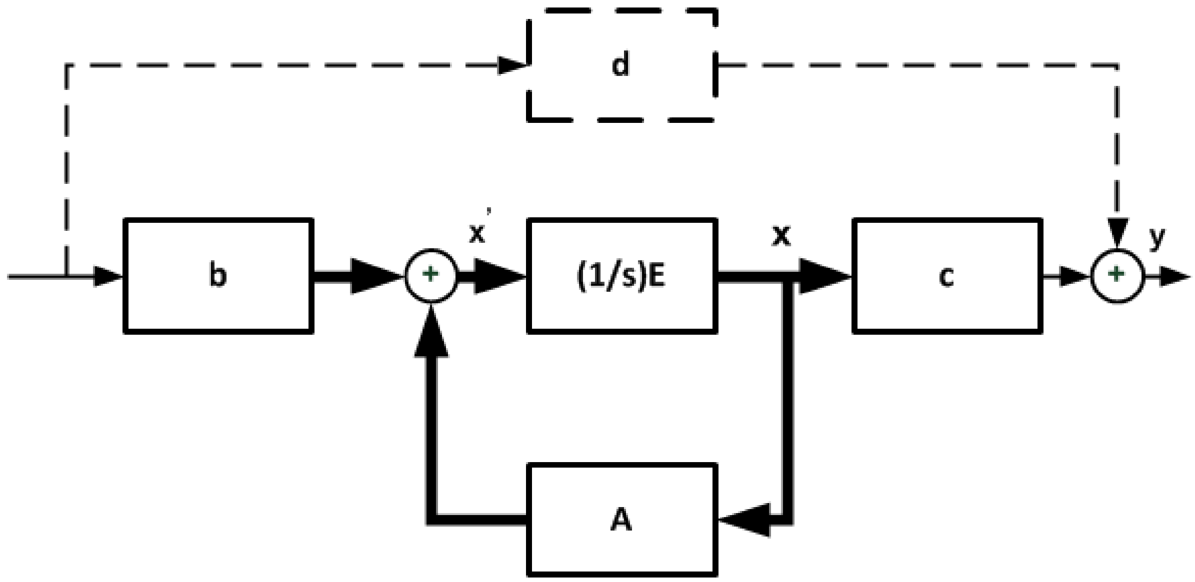


Рис. 5.11. Динамическая система:
d – коэффициент учета прямого воздействия входного сигнала на выходной сигнала; E – единичная матрица;
s – оператор Лапласа; x – вектор параметров состояния

Если перейти от матриц и векторов в выражении (5.15) к скалярным величинам получим выражение для передаточной функции в виде:

$$W(s) = \frac{K_\Sigma}{s^n + \alpha_{n-1}s^{n-1} + \alpha_{n-2}s^{n-2} + \ldots + \alpha_2 s^2 + \alpha_1 s + \alpha_0}, \qquad (5.16)$$

где

$$K_\Sigma = \frac{K}{a_n};\ \alpha_{n-1} = \frac{a_{n-1}}{a_n};\ \alpha_{n-2} = \frac{a_{n-2}}{a_n};\ \alpha_2 = \frac{a_2}{a_n};\ \alpha_1 = \frac{a_1}{a_n};\ \alpha_0 = \frac{a_0}{a_n}.$$

При известной передаточной функции системы, всегда можно перейти к описанию системы в пространстве состояний. Рассмотрим пример получения описания в пространстве состояний по передаточной функции:

$$W(s) = \frac{y(s)}{u(s)} = \frac{K}{T_2^2 s^2 + T_1 s + 1}.$$

Приведем передаточную функцию к виду:

$$W(s) = \frac{K_\Sigma}{s^2 + \alpha_1 s + \alpha_0};$$

$$\alpha_1 = \frac{T_1}{T_2^2};\ \alpha_0 = \frac{1}{T_2^2};\ K_\Sigma = \frac{K}{T_2^2}$$

Получим уравнение «вход–выход»:

$$W(s) = \frac{y(s)}{u(s)} = \frac{K_\Sigma}{s^2 + \alpha_1 s + \alpha_0};$$

$$s^2 y(s) + \alpha_1 y(s)s + \alpha_0 y(s) = K_\Sigma u(s);$$

$$s =';$$

$$(s) = (t);$$

$$y''(t) + \alpha_1 y'(t) + \alpha_0 y(t) = K_\Sigma u(t).$$

Введем параметры состояния:

$$x_1 = y(t);$$

$$x_2 = y'(t).$$

Получим уравнения системы в пространстве состояния:

$$x_1' = x_2;$$

$$x_2' + \alpha_1 x_2 + \alpha_0 x_1 = K_\Sigma u.$$

В окончательном виде уравнения примут вид:

$$x_1' = x_2;$$

$$x_2' = K_\Sigma u - \alpha_1 x_2 - \alpha_0 x_1.$$

Матрицы и вектора системы:

$$\begin{cases} x' = Ax + bu; \\ y = cx; \end{cases}$$

$$A = \begin{bmatrix} 0 & 1 \\ -\alpha_0 & -\alpha_1 \end{bmatrix};\ b = \begin{bmatrix} 0 \\ K_\Sigma \end{bmatrix};\ x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot x.$$

Регулирование по состоянию

Описание системы в пространстве состояний позволяет выполнить расчет регулятора состояния для обеспечения желаемого характера изменения выходного сигнала y при подаче на вход системы задающего сигнала [10] (рис. 5.12).

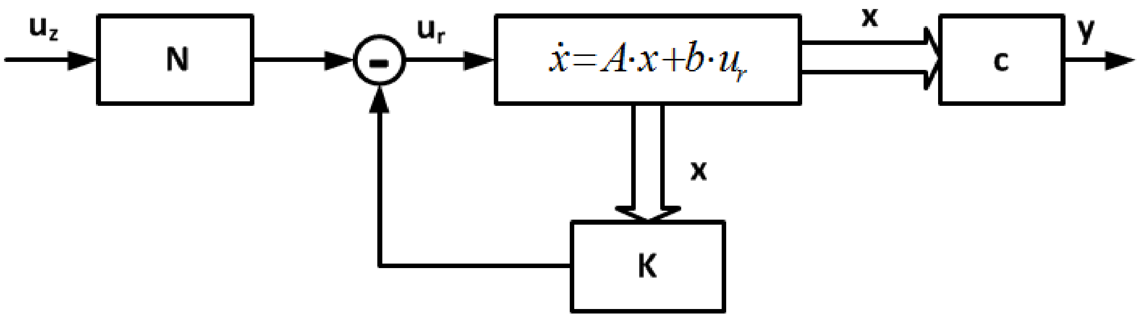


Рис. 5.12. Система с регулятором состояния: u_z – задающий сигнал; u_r – сигнал регулятора; k – вектор коэффициентов усиления регулятора; N – масштабный коэффициент усиления задающего сигнала

Регулятор состояния получает на вход вектор параметров состояния системы, каждый параметр масштабируется коэффициентом усиления вектора *k*. Полученное скалярное значение вычитается из масштабированного задающего сигнала. Таким образом, регулятор состояния формирует сигнал регулирования в виде [10]:

$$k = [k_1, k_2, \ldots, k_n];$$

$$u_r = u_z \cdot N - k \cdot x.$$

Такой регулятор состояния – регулятор состояния по полному вектору состояния системы. Следует иметь в виду, что регулятор состояния может быть рассчитан для систем с не вырожденной матрицей регулирования *R* [10]:

$$R = [b, Ab, A^2 b, \ldots, A^{n-1} b];$$

$$\mathrm{Det}\, R \neq 0.$$

Наблюдатель состояния

Практическая реализация регулятора состояния требует измерения всех параметров состояния системы. Обычно доступен для измерения только выходной сигнал системы. Для получения полной информации о векторе состояния используют наблюдатель [10]. Блок-схема системы с наблюдателем состояния показана на рис. 5.13.

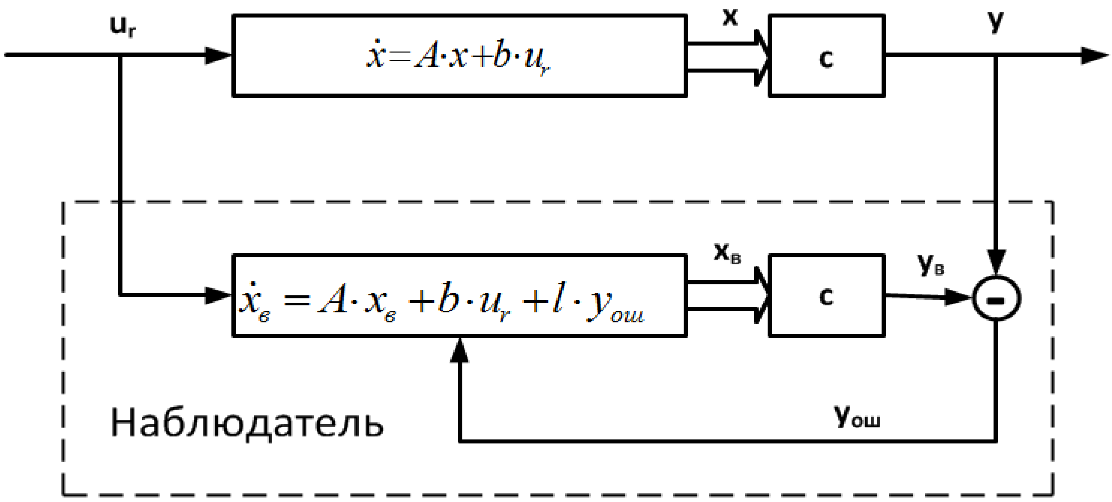


Рис. 5.13. Полный наблюдатель состояния:
$l(n\times 1)$ – вектор коэффициентов усиления наблюдателя; $y_{ош} = y - y_в$

Наблюдатель состояния – модель системы, которая работает параллельно с ней. На входы наблюдателя подают сигнал регулирования и выходной сигнал системы. Выходной сигнал системы y сравнивается с восстановленным сигналом выхода $y_в$ и вычисляется разность $y_{ош}$ – ошибка восстановления. Ошибка масштабируется коэффициентами усиления модели наблюдателя и добавляется в модель системы. Коэффициенты усиления наблюдателя выбираются таким образом, чтобы свести ошибку восстановления к минимуму. Регулятор состояния по восстановленному вектору состояния показан на рис. 5.14.

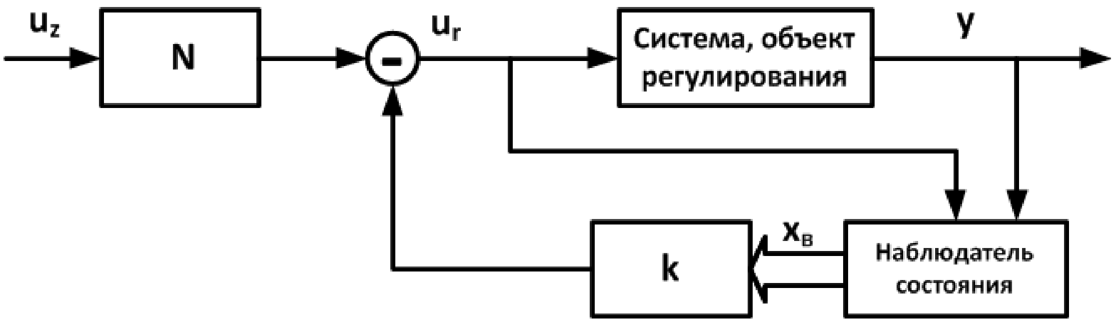


Рис. 5.14. Регулирование по восстановленному вектору состояния

Вычисление коэффициентов усиления регулятора состояния

Определение коэффициентов усиления требует проведения вычисления с помощью матричного уравнения 5.17.

$$k = k_r \cdot T = [g_0 - \alpha_0,\, g_1 - \alpha_1,\, ...,\, g_{n-1} - \alpha_{n-1}] \begin{bmatrix} p_n \\ p_n A \\ p_n A^2 \\ ... \\ p_n A^{n-1} \end{bmatrix}, \qquad (5.17)$$

где

$$R^{-1} = \begin{bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{bmatrix} \Rightarrow p_n \text{ – строка.}$$

Вычисления требуют задания характеристического полинома замкнутой системы по состоянию

$$Nr(s) = s^n + g_{n-1}s^{n-1} + \dots + g_2 s^2 + g_1 s + g_0.$$

При известном характеристическом полиноме передаточной функции 5.16 исходной системы

$$N(s) = s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_2 s^2 + \alpha_1 s + \alpha_0.$$

Замыкание системы по состоянию приводит к изменению масштаба отработки системой задающего сигнала в установившемся режиме. В системе имеет место статическая ошибка. Для ее устранения вычисляют коэффициент масштабирования N задающего сигнала. Выражение для значения N получают с помощью уравнения

$$(A - b \cdot k)x_\infty + b \cdot N \cdot u = 0. \qquad (5.18)$$

Уравнение справедливо для систем с регулятором состояния в установившемся режиме. Получим уравнение для вектора состояния

$$x_\infty = -(A - b \cdot k)^{-1} b \cdot N \cdot u. \qquad (5.19)$$

Выходной сигнал системы в установившемся режиме будет иметь вид

$$y_\infty = c x_\infty,$$

с учетом (5.19) получим

$$y_\infty = (-c(A - b \cdot k)^{-1} b \cdot N) \cdot u.$$

Учитывая, что выход системы в установившемся режиме будет равен входному сигналу $y_\infty = u$, получим уравнение для вычисления значения N

$$N = -(c(A - b \cdot k)^{-1} \cdot b)^{-1}. \qquad (5.20)$$

Вычисление коэффициентов усиления наблюдателя

Расчет коэффициентов возможен только для полностью наблюдаемых систем. У таких систем должна быть не вырождена матица наблюдаемости [10]

$$O = \begin{bmatrix} c \\ cA \\ cA^2 \\ \dots \\ cA^{n-1} \end{bmatrix},$$

$$\mathrm{Det}\, O \neq 0.$$

Вектор коэффициентов усиления наблюдателя вычисляют с помощью формулы

$$l = T^{-1}l_o = \left[t_n,\, At_n,\, A^2t_n,\, ...,\, A^{n-1}t_n\right] \begin{bmatrix} z_0 - \alpha_0 \\ z_1 - \alpha_1 \\ z_2 - \alpha_2 \\ ... \\ z_{n-1} - \alpha_{n-1} \end{bmatrix}, \tag{5.21}$$

где $O^{-1} = [t_1 \,|\, t_2 \,|\, ... \,|\, t_n] \;\Rightarrow\; t_n$ – столбец.

Вычисления производят на основе заданного характеристического полинома наблюдателя

$$No(s) = s^n + z_{n-1}s^{n-1} + ... + z_2s^2 + z_1s + z_0.$$

При известном характеристическом полиноме передаточной функции системы (5.16).

Канонические формы систем

В общем виде передаточная функция системы имеет вид:

$$\frac{y(s)}{u(s)} = \frac{b_m s^m + ... + b_2 s^2 + b_1 s + b_0}{s^n + \alpha_{n-1}s^{n-1} + ... + \alpha_2 s^2 + \alpha_1 s + \alpha_0}, \tag{5.22}$$

$$\frac{y(s)}{u(s)} = \frac{M(s)}{N(s)}.$$

Структурная схема такой системы показана на рис. 5.15.

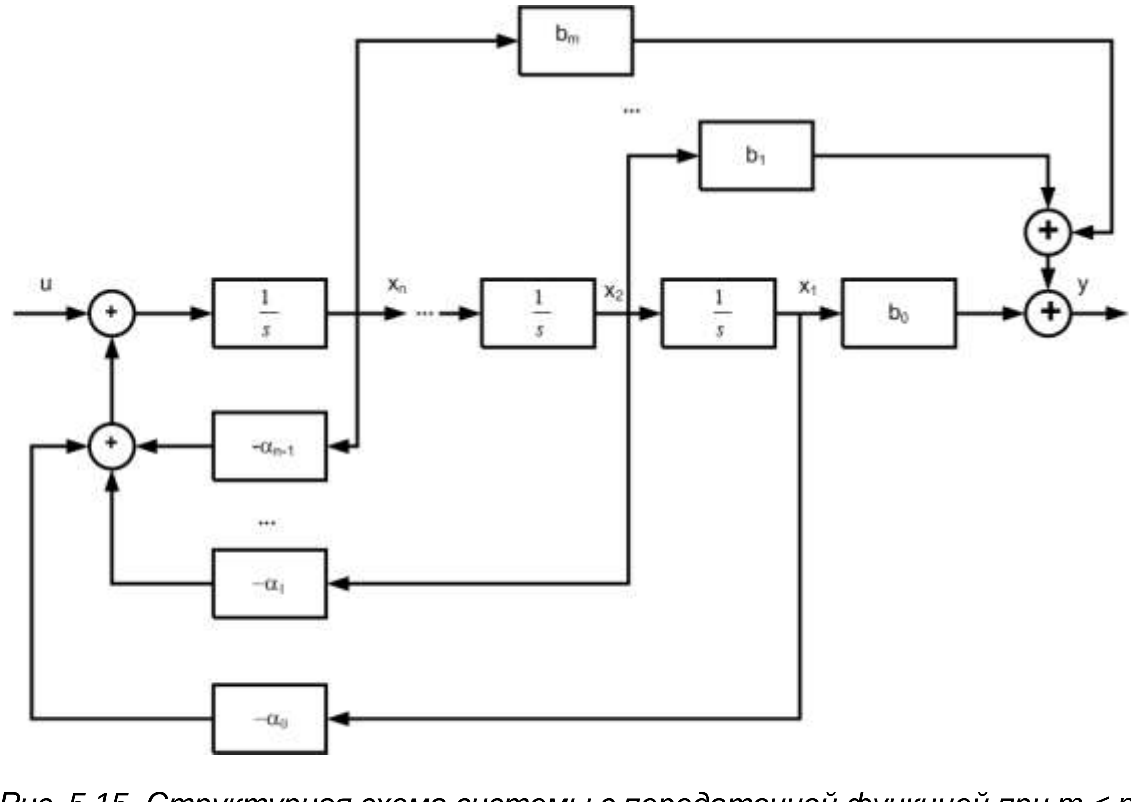


Рис. 5.15. Структурная схема системы с передаточной функцией при m < n; 1/s – изображение интеграла по Лапласу

Создать и исследовать можно только такие системы, для которых справедливо неравенство $n \geq m$ [10]. При $m = 0$; $b_0 = K$ получим упрощенный вариант передаточной функции:

$$\frac{y(s)}{u(s)} = \frac{M(s)}{N(s)} = \frac{K}{N(s)}.$$

Структурная схема такой системы показана на рис. 5.16.

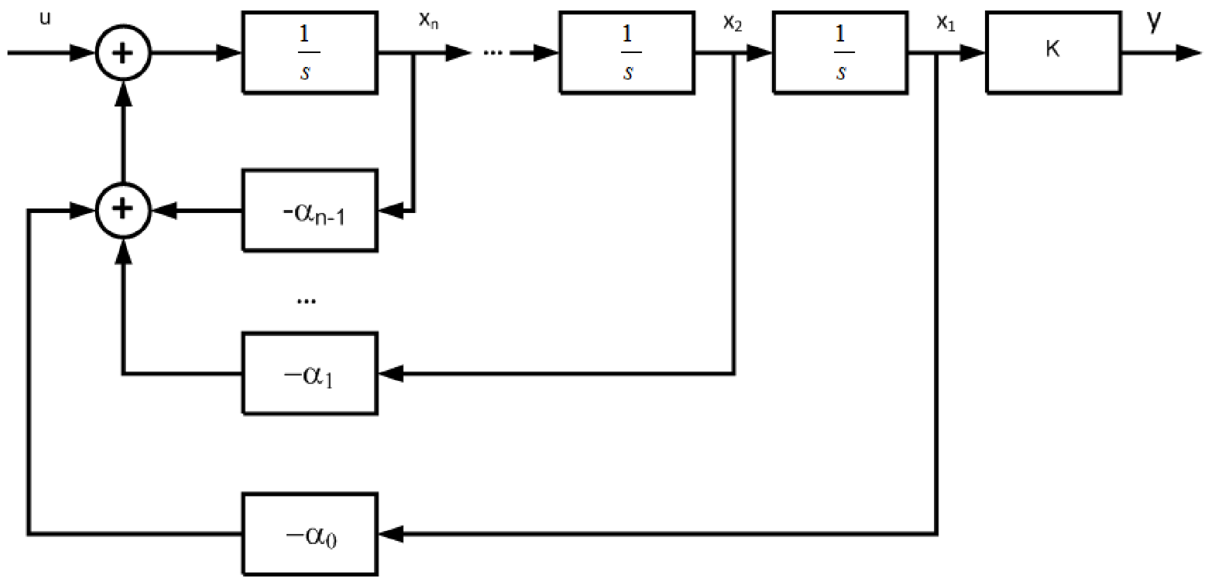


Рис. 5.16. Структурная схема системы с передаточной функцией с коэффициентом усиления в числителе

Если записать уравнение системы к канонической регулируемой форме с учетом выражения 5.22:

$$\begin{bmatrix} x_1' \\ x_2' \\ \ldots \\ x_{n-1}' \\ x_n' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \ldots & -\alpha_{n-1} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ 1 \end{bmatrix} \cdot u;$$

$$y = b_0, b_1, b_2, \ldots, b_{m-1} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_{n-1} \\ x_n \end{bmatrix},$$

то процедура вычисления коэффициентов усиления регулятора упрощается:

$$A - b \cdot k \Rightarrow k = g_0 - \alpha_0, g_1 - \alpha_1, \ldots, g_{n-1} - \alpha_{n-1} \,. \tag{5.23}$$

Если описание системы дано в канонической наблюдаемой форме, то можно упростить вычисление коэффициентов усиления наблюдателя.

$$\begin{bmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_{n-1} \\ x'_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \dots & 0 & -\alpha_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -\alpha_{n-1} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b_{m-2} \\ b_{m-1} \end{bmatrix} \cdot u,$$

$$y = \begin{matrix} 0 & 0 & 0 & \dots & 1 \end{matrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_{n-1} \\ x_n \end{bmatrix}.$$

Коэффициенты усиления наблюдателя определяются в виде:

$$A - l \cdot c => l = \begin{bmatrix} z_0 - \alpha_0 \\ z_1 - \alpha_1 \\ \dots \\ z_{n-1} - \alpha_{n-1} \end{bmatrix}. \tag{5.24}$$

Настройка регулятора состояния и наблюдателя

Отработка системой задающего сигнала зависит от расположения полюсов ее передаточной функции $W(s)$

$$p_i = -a \pm jb,$$

где a – вещественное значение, b – мнимое значение полюса. Полюса – корни характеристического полинома $N(s)$ (рис. 5.17). Чтобы система была устойчивой полюса должны располагаться в левой части комплексной плоскости. Характер расположения полюсов на комплексной плоскости определяет качество динамических процессов в системе:

- вид переходного процесса: монотонный или колебательный;
- быстродействие системы.

Замыкание системы по ее состоянию требует получение таких значений коэффициентов усиления k, чтобы замкнутая система получила передаточную функцию:

$$\frac{y(s)}{u(s)} = \frac{M(s)}{Nr(s)}.$$

Выбор $Nr(s)$ должен быть таким, чтобы полюса системы были расположены заданным образом [10].

Выбор полинома Nr удобно производить, используя стандартные формы требуемого вида. Стандартная форма – это эмпирически полученный полином, обеспечивающий определенный вид переходного процесса в системе. При этом под переходным процессом понимается реакция системы при подаче на ее вход ступенчатого задающего сигнала. Приведем примеры стандартных форм для n = 1–3.

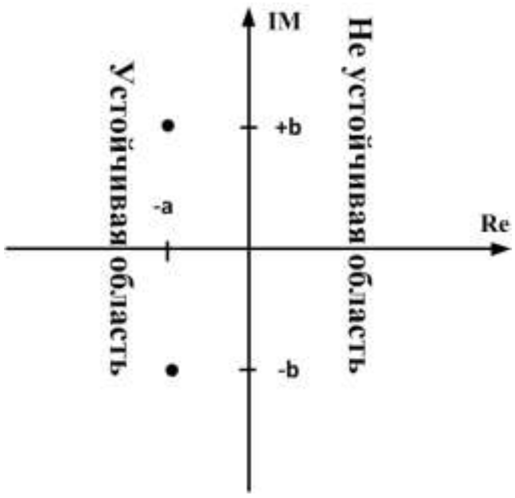


Рис. 5.17. Полюса системы, система второго порядка

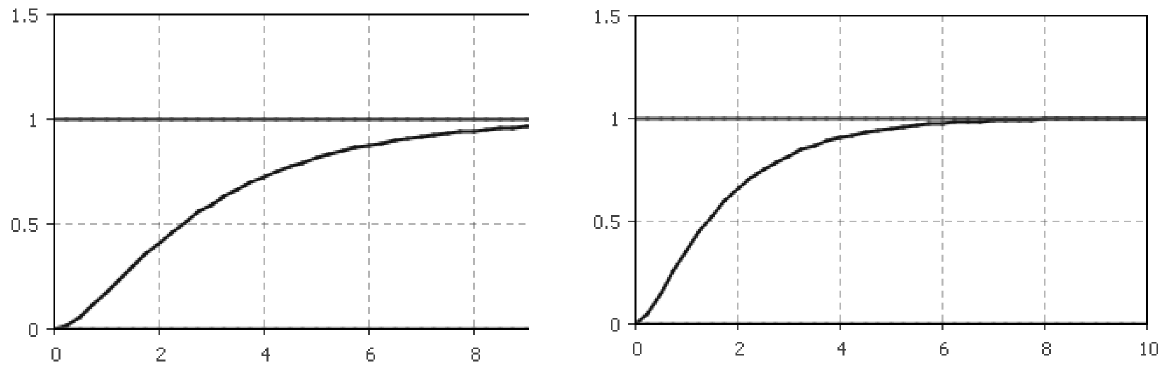
Биномиальная форма:

$$s+\omega_0;$$

$$s^2+2\omega_0 s+\omega_0^2;$$

$$s^3+3\omega_0 s^2+3\omega_0^2 s+\omega_0^3.$$

Обеспечивает монотонный характер протекания переходного процесса. Быстродействие системы зависит от параметра ω_0. Чем выше это значение, тем быстрее завершается переходный процесс. На рис. 5.18 показан переходный процесс в системе. Видно, что при ω_0 = 1 переходный процесс протекает быстрее.



ω_0 = 0,5 ω_0 = 1,0

Рис. 5.18. Переходные процессы, биномиальная форма; система третьего порядка

Стандартная форма, обеспечивающая минимум функционала:

$$I = \int_0^\infty t\,|\,e(t)\,|\,dt;$$

$$e(t) = u_z(t) - y(t);$$

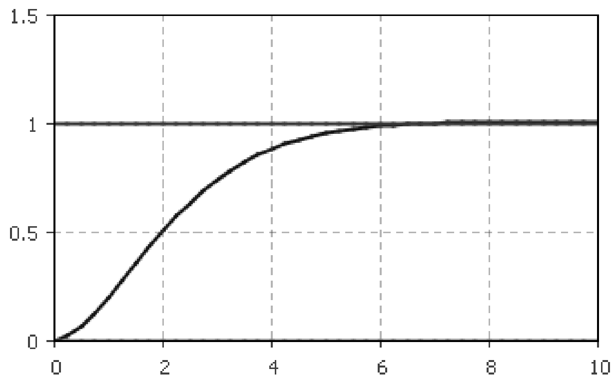
$$s + \omega_0;$$

$$s^2 + 1{,}4\omega_0 s + \omega_0^2;$$

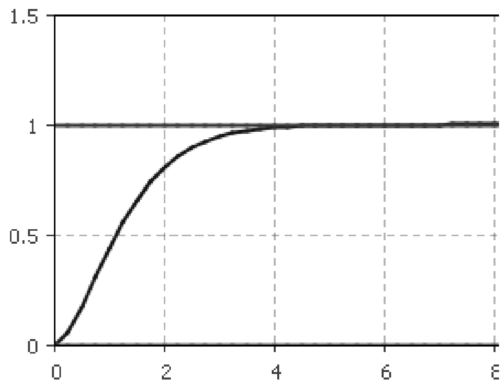
$$s^3 + 1{,}75\omega_0 s^2 + 2{,}15\omega_0^2 s + \omega_0^3.$$

Переходные процессы показаны на рис. 5.19. Сравнивая переходные процессы двух стандартных форм, можно сделать вывод, что вторая форма обеспечивает более быстрый характер протекания переходного процесса уже при значении ω_0 = 0,5.

При настройке наблюдателя выбирают стандартную форму для $No(s)$ со значением ω_0'. Для систем с наблюдателем $\omega_0' \geq \omega_0$.



ω_0 = 0,5



ω_0 = 1,0

Рис. 5.19. Переходные процессы; интегральный критерий; система третьего порядка

Контрольные вопросы

1. Что представляет собой уравнение «вход–выход» для системы с одним входом и выходом?
2. Как перейти к описанию системы в пространстве состояния по уравнению «вход-выход»?
3. Что такое передаточная функция системы?
4. Как перейти к описанию системы в пространстве состояния, если известна ее передаточная функция?
5. Как строится регулятор состояния системы?
6. Как строится наблюдатель состояния системы?
7. Для чего используют канонические формы описания работы систем?

8. Опишите алгоритм расчета регулятора состояния.
9. Опишите алгоритм расчета наблюдателя состояния.
10. Как выполняется настройка регулятора состояния?

ЛЕКЦИЯ №6. ЭКСПЕРИМЕНТ И МОДЕЛИРОВАНИЕ

Модели и их виды, методика проведения компьютерного эксперимента, статистическое моделирование классификация экспериментов, регрессионные модели.

Моделирование – замещение одного объекта (оригинала) другим (моделью) для фиксации и изучения свойств модели. Замещение производится с целью упрощения, удешевления, ускорения изучения свойств оригинала. В общем случае объектом-оригиналом может быть естественная или искусственная, реальная или воображаемая система. Она имеет множество параметров S и характеризуется определёнными свойствами. Количественной мерой свойств системы служит множество характеристик Y. Система проявляет свои свойства под влиянием внешних воздействий X.

Виды моделей

Физической моделью обычно называют систему, эквивалентную или подобную оригиналу, но возможно имеющую другую физическую природу.

Виды физических моделей [5]:

- квазинатуральные;
- масштабные;
- аналоговые;
- математические.

Квазинатуральные модели – совокупность натуральных и математических моделей.

Масштабная модель – это система той же физической природы, что и оригинал, но отличается от него масштабами.

Аналоговыми моделями называют системы, имеющие физическую природу, отличающуюся от оригинала, но сходные с оригиналом процессы функционирования. Для создания аналоговой модели требуется наличие математического описания изучаемой системы.

Математические модели представляют собой формализованное представление системы с помощью абстрактного языка, математических соотношений, отражающих процесс функционирования системы.

Математические модели можно классифицировать на:

- аналитические;
- численные;
- имитационные.

Аналитической моделью называется такое формализованное описание системы, которое позволяет получить решение уравнений, описывающих ее работу в явном виде, используя известный математический аппарат.

Численная модель характеризуется такими уравнениями, которые допускают только частные решения для конкретных начальных условий и количественных параметров моделей.

Имитационная модель – это совокупность описания системы и внешних воздействий, алгоритмов функционирования системы или правил изменения состояния системы под влиянием внешних и внутренних возмущений.

Математическую схему можно определить как звено при переходе от содержательного к формализованному описанию процесса функционирования системы с учётом воздействия внешней среды.

Имеет место цепочка: описательная модель > математическая схема > имитационная модель. Процесс функционирования системы S описывается оператором F_S [5]:

$$Y(t) = F_s(X, V, H, t); \tag{6.1}$$

- совокупность X – входных воздействий на S $x_i \in X$, $i = 1 \ldots n_X$;
- совокупность воздействий внешней среды $v_i \in V$, $i = 1 \ldots n_V$;
- совокупность внутренних (собственных) параметров системы $h_k \in H$, $k = 1 \ldots n_h$;
- совокупность выходных характеристик системы $y_j \in Y$, $j = 1 \ldots n_y$.

Также соотношения в ряде случаев могут быть получены через свойства системы в конкретные моменты времени, называемые состояниями. Совокупность всех возможных значений состояний называется пространством состояний объекта моделирования Z

$$Z(t) = \Phi(z, X, V, h, t). \tag{6.2}$$

Состояния системы S характеризуются вектором

$$Z(z_1, \ldots, z_k).$$

Статистическое моделирование

Статистическое моделирование – базовый метод моделирования, заключающийся в том, что модель испытывается множеством случайных сигналов с заданной плотностью вероятности, целью которого является статистическое определение выходных результатов. В основе статистического моделирования лежит метод Монте-Карло.

Пусть требуется найти значение интеграла

$$y = \int_{x_1}^{x_2} f(x)dx. \tag{6.3}$$

График функции $f(x)$ показан на рис. 6.1. Вычислить значение интеграла этой функции – это нахождение площади под кривой графика.

Кривую ограничивают сверху, справа и слева. Случайным образом распределяют точки в прямоугольнике поиска. Определяют:

N_1 – количество точек, принятых для испытаний (то есть попавших в прямоугольник);

N_2 – количество точек под кривой, то есть попавших в закрашенную площадь под функцией.

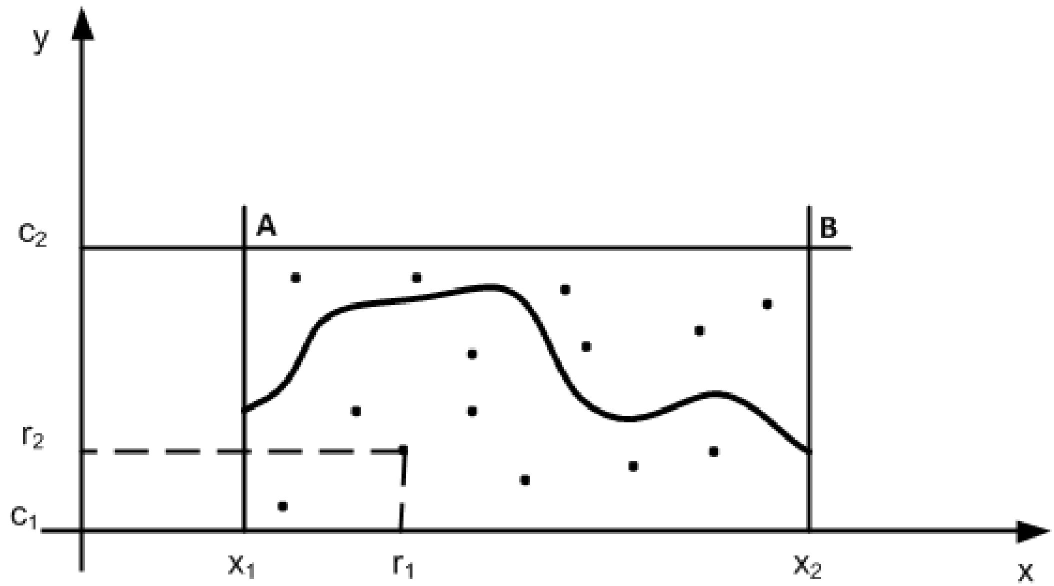


Рис. 6.1. Методика статистического исследования методом Монте-Карло

Количество точек, попавших под кривую по отношению к общему числу точек пропорционально площади под кривой (величине интеграла).

$$\frac{N_2}{N_1} = \frac{y}{(x_2 - x_1)\cdot(c_2 - c_1)}. \qquad (6.4)$$

Это верно при большом числе испытуемых точек. Значения r_1 и r_2 на рисунке 6.1 являются равномерно распределенными случайными числами из интервалов (x_1 и x_2) и (c_1, c_2) соответственно. Эксперименты показывают: чтобы увеличить точность в 10 раз, необходимо объем выборки увеличить в 100 раз. Точность примерно пропорциональна корню квадратному из объема выборки:

$$\text{точность} = \sqrt{\text{объем выборки}}.$$

Планирование эксперимента

В планировании эксперимента различают входные (изогенные) и выходные (эндогенные) переменные

$$y_1, y_2 \ldots, y_m.$$

Входные переменные x_1, x_2, …, x_k называют факторами, *а* выходные – реакциями. Каждый фактор x_i, i = 1, 2, …, k может принимать в эксперименте одно или несколько значений, называемых уровнями.

Фиксированный набор уровней факторов определяет одно из возможных состояний рассматриваемой системы. Одновременно этот набор представляет собой условия проведения одного из возможных экспериментов.

Каждому фиксированному набору уровню факторов соответствует определённая точка в многомерном пространстве, называемая факторным пространством.

Эксперименты не могут быть реализованы во всех точках факторного пространства, а лишь в точках, принадлежащих допустимой области.

План эксперимента обычно используется для определения экстремальной характеристики объекта. Поэтому планирование эксперимента называется экстремальным.

Стратегическое планирование – ставит своей целью получение необходимой информации о системе с помощью модели, реализованной на ЭВМ. Тактическое планирование – определяет способы проведения каждой серии испытаний машинной модели.

Эксперимент может быть пассивным или активным. При пассивном эксперименте информация об исследуемом объекте накапливается путем наблюдения, то есть информацию получают в условиях обычного функционирования объекта.

Активный эксперимент проводится с применением искусственного воздействия на объект по специальной программе. Примером пассивного эксперимента может быть анализ работы схемы, которая не имеет входов, а только выходы и повлиять на ее работу невозможно.

Активный эксперимент предполагает возможность воздействия на ход процесса и выбора в каждом опыте уровней факторов. Совокупности факторов должны отвечать требованиям совместимости и независимости.

Соблюдение первого требования означает, что все комбинации факторов осуществимы и безопасны. Второго – возможность установления фактора на любом уровне независимо от уровней других факторов.

Схемы эксперимента

Рис. 6.2. Объект с одним входом и выходом

Однофакторный пассивный эксперимент проводится путем выполнения *n* пар измерений в дискретные моменты времени единственного входного параметра *X* и соответствующих значений выходного параметра *Y* (рис. 6.2).

Рис. 6.3. Несколько факторов

Многофакторный пассивный эксперимент проводится при контроле значений нескольких входных параметров X_i. Его целью является установление зависимости выходного параметра от двух или более переменных $y = F(x_1, x_2, \ldots)$ (рис. 6.3).

Объект с несколькими входами и возмущением. Параметры:

- управляющие (входные x_i);
- параметры состояния (выходные Y);
- возмущающие воздействия (W_i).

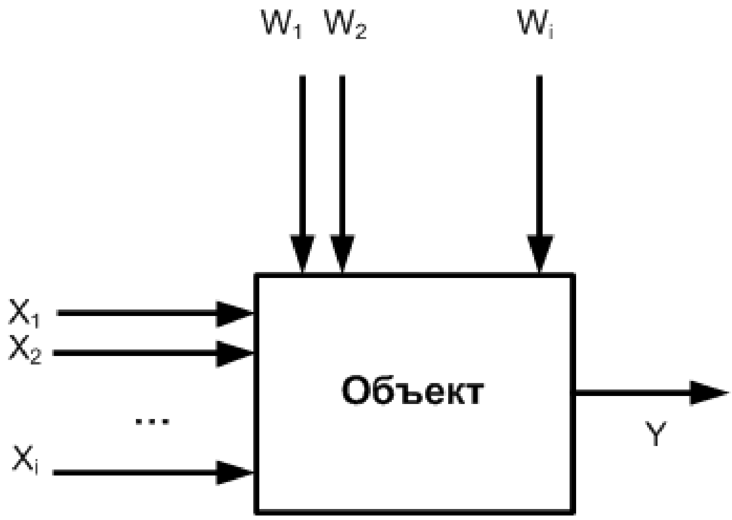


Рис. 6.4. Полнофакторный эксперимент

Полнофакторный эксперимент предполагает возможность управлять объектом по одному или нескольким независимым каналам (рис. 6.4). Полнофакторный эксперимент характеризуется тем, что при фиксированных возмущающих воздействиях W_i минимальное число уровней каждого фактора равно двум. Необходимое число N опытов в полном факторном эксперименте для реализации всех возможных сочетаний уровней факторов $N = 2^k$, где k – число факторов.

Эксперимент и концепция черного ящика

В целях исследований часто бывает удобно представить исследуемый объект в виде «черного ящика», имеющего входы и выходы, не рассматривая детально его внутреннюю структуру (рис. 6.5).

По степени информированности исследователя об объекте существует деление объектов на три типа «ящиков»:

- «белый ящик»: об объекте известно все;
- «серый ящик»: известна структура объекта, неизвестны количественные значения параметров;
- «черный ящик»: об объекте неизвестно ничего.

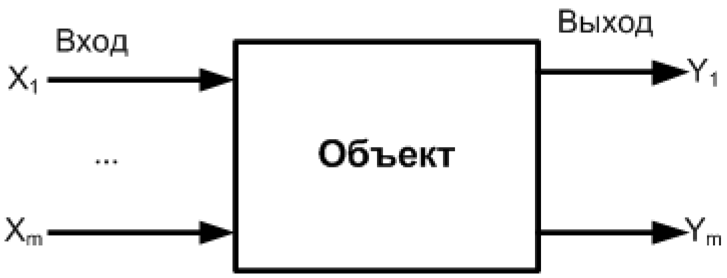


Рис. 6.5. Объект с несколькими входами и выходами

Значения на входах и выходах черного ящика можно наблюдать и измерять. Содержимое ящика неизвестно.

Задача состоит в том, чтобы, зная множество значений на входах и выходах, построить модель, то есть определить функцию ящика, по которой вход преобразуется в выход. Такая задача называется задачей регрессионного анализа.

В зависимости от того, доступны входы исследователю для управления или только для наблюдения, можно говорить про активный или пассивный эксперимент с ящиком. Рассмотрим правила выполнения регрессионного анализа.

Один вход и один выход

Исследователь вносит гипотезу о структуре ящика. Рассматривая экспериментально полученные данные, предположим, что они подчиняются линейной гипотезе, то есть выход Y зависит от входа X линейно, и гипотеза имеет вид

$$y = A_1 \cdot X + A_0. \tag{6.5}$$

Нужно определить неизвестные коэффициенты A_0 и A_1 модели. Ошибки E_i, для всех n точек наблюдений следует сложить. Чтобы положительные ошибки не компенсировали в сумме отрицательные, каждую из ошибок возводят в квадрат и складывают их значения в суммарную ошибку F уже одного знака.

$$\begin{gathered} E_i^2 = (Y_i - A_0 - A_1 \cdot X_i)^2,\ i = 1, \ldots, n; \\ F(A_0, A_1) = \sum_{i=1}^{n} E_i^2. \end{gathered} \tag{6.6}$$

Цель метода – минимизация суммарной ошибки F за счет подбора коэффициентов A_0, A_1. Необходимо найти такие коэффициенты A_0,

A_1 линейной функции $Y = A_1X + A_0$, чтобы ее график проходил как можно ближе одновременно ко всем экспериментальным точкам (рис. 6.6). Данный метод называется так же методом наименьших квадратов.

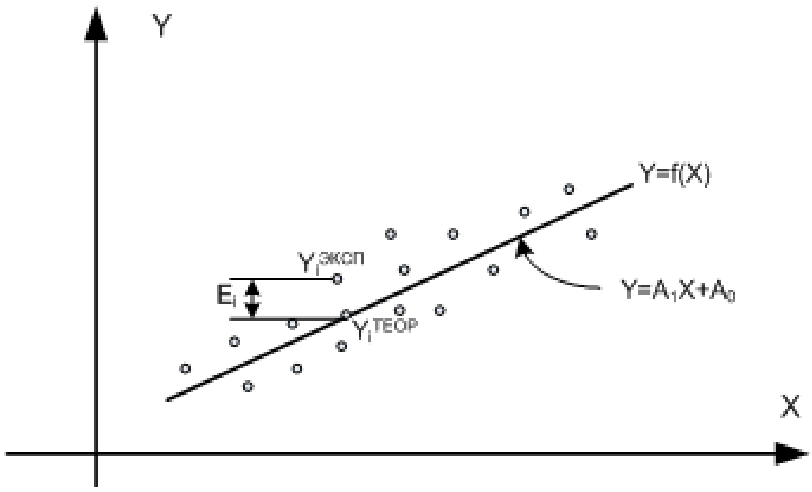


Рис. 6.6. Схема линейной регрессии

Суммарная ошибка F является функцией двух переменных A_0 и A_1, $F(A_0, A_1)$, меняя переменные, можно влиять на величину суммарной ошибки:

$$E_i = (Y_i^{\text{эксп}} - Y_i^{\text{теор}});$$
$$F(A_0, A_1) = \sum_{i=1}^{n} E_i^2; \qquad (6.7)$$
$$F(A_0, A_1) \Rightarrow \min.$$

Для нахождения коэффициентов A_0, A_1 используют матричное уравнение Крамера:

$$\begin{pmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{pmatrix} \cdot \begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{pmatrix}. \qquad (6.8)$$

Оценки приемлемости рассчитанных значений A_0 и A_1 выполняют путем определения числа точек эксперимента попавших в полосу допуска (рис. 6.7):

$$Y^{\text{теор}} + S;\ Y^{\text{теор}} - S,$$

где

$$S = \frac{\sigma}{\cos(\operatorname{arctg}(A_1))}; \qquad (6.9)$$
$$\sigma = \sqrt{\frac{F}{n}}.$$

Если в полосу допуска попадает более 70% экспериментальных точек, то выдвинутая гипотеза принимается.

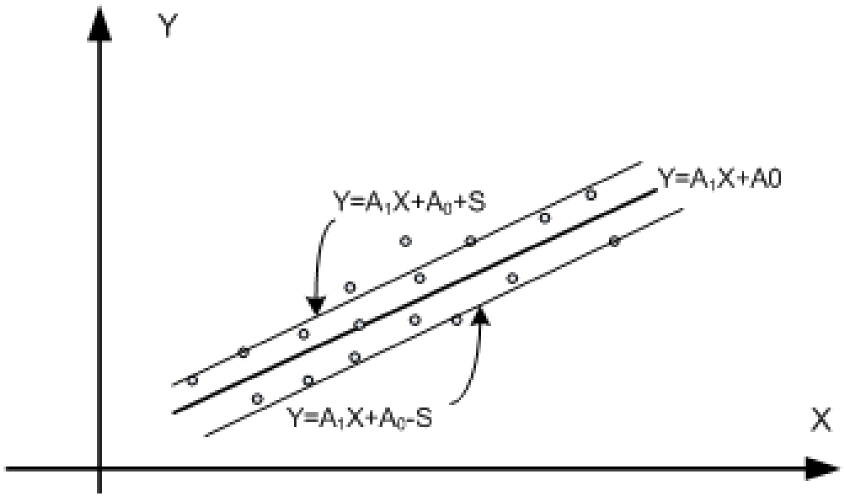


Рис. 6.7. Оценка качества регрессионной модели

Система с несколькими входами и одним выходом

Линейная регрессия для объекта с одним входом и несколькими выходами выполняется с помощью выражений 6.10, 6.11 и 6.12:

$$Y = A_0 + A_1 \cdot X_1 + \ldots + A_m \cdot X_m, \tag{6.10}$$

$$F(A_0, A_1, \ldots, A_m) = \sum_{i=1}^{n} E_i^2 \Rightarrow \min, \tag{6.11}$$

$$\begin{bmatrix} n & \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{2i} & \ldots & \sum_{i=1}^{n} X_{mi} \\ \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{1i}X_{1i} & \sum_{i=1}^{n} X_{2i}X_{1i} & \ldots & \sum_{i=1}^{n} X_{mi}X_{1i} \\ \sum_{i=1}^{n} X_{2i} & \sum_{i=1}^{n} X_{1i}X_{2i} & \sum_{i=1}^{n} X_{2i}X_{2i} & \ldots & \sum_{i=1}^{n} X_{mi}X_{2i} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sum_{i=1}^{n} X_{mi} & \sum_{i=1}^{n} X_{1i}X_{mi} & \sum_{i=1}^{n} X_{2i}X_{mi} & \ldots & \sum_{i=1}^{n} X_{mi}X_{mi} \end{bmatrix} \cdot \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \ldots \\ A_m \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} Y_i X_{1i} \\ \sum_{i=1}^{n} Y_i X_{2i} \\ \ldots \\ \sum_{i=1}^{n} Y_i X_{mi} \end{bmatrix} \tag{6.12}$$

Ошибка F зависит от выбора параметров A_0, A_1, …, A_m. По аналогии с одномерной моделью, для каждой точки вычисляется ошибка E_i затем находится суммарная ошибка F и значения σ и S и выполняется проверка гипотезы о линейности многомерного черного ящика.

Нелинейные регрессионные модели

Два входа, один выход

Если зависимость выхода от входов напоминает квадратичную, то целесообразно выбрать такую гипотезу:

$$Y = A_0 + A_1 \cdot X_1 + A_2 \cdot X_2 + A_3 \cdot X_1 \cdot X_2 + \\ + A_4 \cdot X_1 \cdot X_1 + A_5 \cdot X_2 \cdot X_2, \qquad (6.13)$$

$$Z_1 = X_1 \cdot X_2;\ Z_2 = X_1 \cdot X_1;\ Z_3 = X_2 \cdot X_2. \qquad (6.14)$$

Обозначим: $Z_1 = X_1 \cdot X_2$; $Z_2 = X_1 \cdot X_1$; $Z_3 = X_2 \cdot X_2$ и подставим эти выражения в предыдущую формулу:

$$Y = A_0 + A_1 \cdot X_1 + A_2 \cdot X_2 + A_3 \cdot Z_1 + A_4 \cdot Z_2 + A_5 \cdot Z_3. \qquad (6.15)$$

Таким образом, данная задача сведена к линейной множественной модели (рис. 6.8).

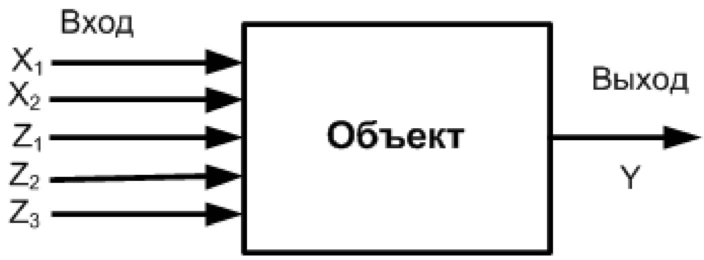


Рис. 6.8. Приближенная линейная модель

Несколько входов – один выход

Если выход системы можно описать зависимостью

$$Y = A_0 \cdot X_1^{A_1} \cdot X_2^{A_2} \cdot \ldots \cdot X_m^{A_m}, \qquad (6.16)$$

то логарифмируя левую и правую части данного уравнения получим:

$$\ln(Y) = \ln(A_0) + A_1 \cdot \ln(X_1) + A_2 \cdot \ln(X_2) + \ldots + A_m \cdot \ln(X_m). \qquad (6.17)$$

Введем обозначения:

$$W = \ln(Y), B_0 = \ln(A_0), Z_1 = \ln(X_1), Z_2 = \ln(X_2), \ldots, Z_m = \ln(X_m). \qquad (6.18)$$

Получим линейное уравнение:

$$W = B_0 + A_1 \cdot Z_1 + A_2 \cdot Z_2 + \ldots + A_m \cdot Z_m. \qquad (6.19)$$

Обратная модель

Пусть выход системы описывается зависимостью:

$$Y = \frac{k}{A_0 + A_1 \cdot X_1 + \ldots + A_m \cdot X_m}. \qquad (6.20)$$

Заменим:

$$W = \frac{1}{Y},\ a_j = \frac{A_j}{k}.$$

Выполним переход к линейной модели:

$$W = a_0 + a_1 \cdot X_1 + \ldots + a_m \cdot X_m. \qquad (6.21)$$

Экспоненциальная модель

Выход системы описывается выражением:

$$Y = e^{B_0 + B_1 \cdot X_1 + B_2 \cdot X_2 + \ldots + B_m \cdot X_m}. \qquad (6.22)$$

Прологарифмируем левую и правую части уравнения:

$$\ln(Y) = B_0 + B_1 \cdot X_1 + B_2 \cdot X_2 + \ldots + B_m \cdot X_m. \qquad (6.23)$$

Выполним замену $W = \ln(Y)$ и получим линейную модель:

$$W = B_0 + B_1 \cdot X_1 + B_2 \cdot X_2 + ... + B_m \cdot X_m. \qquad (6.24)$$

Контрольные вопросы

1. Перечислите виды физических моделей систем.
2. Что такое математическая и имитационная модели?
3. Опишите алгоритм использования метода Монте-Карло?
4. Как выполняется планирование и проведение эксперимента?
5. В чем разница между пассивным активным экспериментом?
6. Дайте классификацию экспериментов.
7. Как выполнятся линейная регрессия для систем с одним входом и выходом?
8. Как выполняется линейная регрессия для систем с нелинейным выходом?

ЛЕКЦИЯ №7. СЛУЧАЙНАЯ ВЕЛИЧИНА

Понятие случайной величины, характеристики случайной величины, описательная статистика, статистические параметры, законы распределения случайной величины.

Случайная величина (СВ) – это численная характеристика, измеряемая по ходу опыта и зависящая от случайного исхода. Каждая СВ задает распределение вероятностей. Вероятность показывает степень возможности осуществления данного события, явления, результата. Вероятность невозможного события равна нулю, достоверного события – единице (100%). Вероятность любого события лежит в пределах от 0 до 1 – в зависимости от того, насколько это событие случайно.

Потребность в понятии вероятности и ее вычисления возникнет, только тогда, когда событие наблюдается не каждый раз, либо событие может произойти, а может не произойти. И в том и другом случае полезно использовать понятие частоты появления события

$$f(A) = \frac{n}{m}. \qquad (7.1)$$

Это отношение числа случаев появления n (благоприятных исходов) события к общему числу наблюдений m.

Существует два вида выборок СВ: зависимые и независимые. Если результаты измерения некоторого свойства у объектов первой выборки не оказывают влияния на результаты измерения этого свойства у объектов второй выборки, то такие выборки считаются независимыми.

В тех случаях, когда результаты одной выборки влияют на результаты другой выборки, выборки считают зависимыми. Событие A

не зависит от события B, если вероятность события A не зависит от того произошло или нет событие B. События A и B независимы, если

$$P(AB) = P(A)P(B). \qquad (7.2)$$

Случайная величина бывает дискретной (можно пронумеровать ее возможные значения), например, выпадение игральной кости = 4, 6, 2, 1, 5, 3. Непрерывной (ее функция распределения $F(x)$ – непрерывна), например, время службы осветительной лампы. Случайная величина имеет две характеристики.

Математическое ожидание – числовая характеристика СВ, приближенно равная среднему значению СВ

$$M(x) = x_1p_1 + x_2p_2 + ... + x_np_n, \qquad (7.3)$$

где p_i – вероятность появления значения случайной величины x_i.

Дисперсия случайной величины

$$D(x) = p_1(x_1 - \bar{x})^2 + p_2(x_2 - \bar{x})^2 + ... + p_n(x_n - \bar{x})^2, \qquad (7.4)$$

где $\bar{x}$ – арифметическое среднее значение СВ.

Для непрерывно распределенной случайной величины X большую роль в ее описании играет функция, обозначаемая обычно $p(x)$, называемая плотностью вероятности (или дифференциальным законом распределения).

Содержательный смысл $p(x)$ заключается в том, что для всякой точки $x_0 \in [a; b]$ и взятого около нее малого приращения dx произведение $p(x_0)dx$ является вероятностью того, что случайная переменная примет значение, заключенное между x_0 и $x_0 + dx$.

Поскольку какое-нибудь значение из $[a; b]$ случайная величина примет наверняка, то

$$\int_a^b p(x)dx = 1 \qquad (7.5)$$

Это условие нормировки для $p(x)$.

Оценивая выборку, следует предварительно сделать заключение о том, какой вид имеет функция распределения величины X, либо каковы значения наиболее часто используемых параметров распределения (таких, как математическое ожидание, дисперсия).

Правдоподобную гипотезу о форме функции $p(x)$ (плотности вероятности) можно попытаться сформулировать по столбчатой диаграмме (гистограмме), построенной с помощью случайной выборки.

Допустим, что случайная величина X ограничена отрезком $[a; b]$ (если a, b заранее неизвестны, то в качестве них можно принять наименьшее и наибольшее из выборочных значений x). Разделим отрезок $[a; b]$ на m равных частей и подсчитаем n_i – число членов выборки, попадающих в i-й участок (при этом m берется таким, чтобы в каждую часть попало много членов выборки, т.е. заведомо m много меньше n). На рис. 7.1 показан типичный вид гистограмммы.

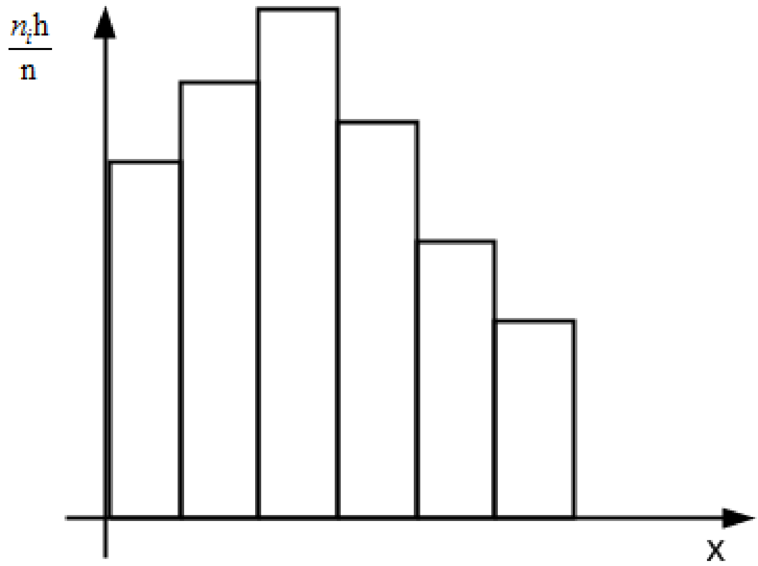


Рис. 7.1. Гистограмма $h = \frac{b-a}{m}$

Описательная статистика

Раздел математической статистики предназначен для представления данных в удобном виде и описания информации в терминах математической статистики и теории вероятностей. Основной величиной в статистических измерениях является единица статистической совокупности. Единица статистической совокупности характеризуется набором признаков или параметров.

Значения каждого параметра или признака могут быть различными и в целом образовывать ряд случайных значений x_1, x_2, …, x_n.

Переменная – это параметр измерения, который можно контролировать или которым можно манипулировать в исследовании.

Относительное значение параметра – это отношение числа объектов, имеющих этот показатель, к величине выборки. Выражается относительным числом или в процентах (процентное значение).

Удельное значение признака – это расчетная величина, показывающая количество объектов с данным показателем, которое содержалось бы в условной выборке, состоящей из 10, или 100, 1000 и т.д. объектов.

Рассмотрим показатели описательной статистики, используемые для оценки результатов эксперимента.

Минимальное и максимальное значения переменной.

Среднее (оценка среднего, выборочное среднее) – сумма значений переменной, деленная на *n* (число значений переменной).

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{x_1 + x_2 + \ldots + x_N}{N}. \quad (7.6)$$

Выборочное среднее. Это значение случайной величины сумма отклонений наблюдений от которой равна 0. Формально это записывается следующим образом:

$$(\bar{x}-x_1)+(\bar{x}-x_2)+...+(\bar{x}-x_n)=0. \qquad (7.7)$$

Для оценки степени разброса (отклонения) какого-то показателя от его среднего значения, наряду с максимальным и минимальным значениями, используются понятия дисперсии и стандартного отклонения.

Дисперсия выборки или выборочная дисперсия – это мера изменчивости переменной. Термин впервые был введен Фишером в 1918 году. Выборочная дисперсия вычисляется по формуле:

$$\sigma^2=\frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{N}. \qquad (7.8)$$

Дисперсия меняется от нуля до бесконечности. Крайнее значение 0 означает отсутствие изменчивости, когда значения переменной постоянны.

Стандартное отклонение, среднее квадратическое отклонение вычисляется как корень квадратный из дисперсии и обозначается σ.

Чем выше дисперсия или стандартное отклонение, тем сильнее разбросаны значения переменной относительно среднего.

Медиана разбивает выборку на две равные части. Половина значений переменной лежит ниже медианы, половина значений лежит выше. Медиана дает общее представление о том, где сосредоточены значения переменной, иными словами, где находится ее центр. Для нахождения медианы измерения записывают в ряд по возрастанию значений. Если число измерений N нечетное, то медиана численно равна значению этого ряда, стоящему точно в середине, или на $(N + 1)/2$ месте.

Например, медиана пяти измерений:

10, 17, 21, 24, 25 – равна 21. Это значение, стоящее на третьем месте $(N + 1)/2 = (5 + 1)/2 = 3$.

Если число измерений четное, то медиана численно равна среднему арифметическому значений ряда, стоящих в середине, или на $N/2$ и $N/2 + 1$ местах.

Например, медиана восьми измерений:

5, 5, 6, 7, 8, 8, 9, 9 – равна 7,5 (7 + 8)/2 = 7,5. Это среднее арифметическое значение ряда, стоящее на четвертом и пятом местах ряда. Так как $N/2 = 8/2 = 4$ и $N/2 + 1 = 4 + 1 = 5$.

Квартили (от слова кварта – четверть) представляют собой значения, которые делят две половины выборки разбитые медианой еще раз пополам.

Различают верхнюю квартиль, которая больше медианы и делит пополам верхнюю часть выборки (значения переменной больше медианы), и нижнюю квартиль, которая меньше медианы и делит пополам нижнюю часть выборки.

Нижнюю квартиль часто обозначают символом 25%, это означает, что 25% значений переменной меньше нижней квартили.

Верхнюю квартиль часто обозначают символом 75%, это означает, что 75% значений переменной меньше верхней квартили.

Мода представляет собой максимально часто встречающееся значение переменной (иными словами, наиболее «модное» значение переменной). Редкая совокупность имеет единственную моду. (Например: 2, 6, 6, 8, 9, 9, 9, 10 – мода = 9).

Если распределение имеет несколько мод, то говорят, что оно мульти модально или много модально (имеет два или более «пика»).

Асимметрия – это свойство распределения выборки, которое характеризует несимметричность распределения СВ. На практике симметричные распределения встречаются редко и чтобы выявить и оценить степень асимметрии, вводят следующую меру:

$$A_{33} = \frac{\sum (x_i - \bar{x})^3 / n}{\sigma^3}. \tag{7.9}$$

Асимметрия бывает положительной и отрицательной. Положительная асимметрия сдвигает влево кривую распределения относительно математического ожидания, а отрицательная – вправо.

Эксцесс – это мера крутости кривой распределения. Эксцесс равен:

$$E_x = \frac{\sum (x_i - \bar{x})^4 / n}{\sigma_x^4} - 3. \tag{7.10}$$

Виды распределений, дискретные законы распределения

Пусть некоторая СВ является дискретной, т.е. может принимать лишь фиксированные значения x_i. В этом случае ряд значений вероятностей $p(x_i)$ для всех (i = 1…n) допустимых значений этой величины называют её законом распределения.

Биномиальное распределение (распределение Бернулли). Возникает в тех случаях, когда ставится вопрос: сколько раз происходит некоторое событие в серии из определенного числа независимых наблюдений (опытов), выполняемых в одинаковых условиях.

Пусть известна величина p – вероятность того, что вошедший в магазин посетитель окажется покупателем, тогда $(1 - p) = q$ – вероятность того, что вошедший в магазин посетитель не окажется покупателем.

Если x – число покупателей из общего числа n посетителей, то вероятность того, что среди n посетителей оказалось k покупателей равна:

$$P_{(X=k)} = \frac{n!}{k!(n-k)!} p^k q^{n-k} = C_n^k p^k q^{n-k}, \tag{7.11}$$

где k = 0, 1, … n.

Распределение Пуассона

Распределение Пуассона играет важную роль в ряде вопросов физики, теории связи, теории надежности, теории массового обслуживания где в течение определенного времени может происходить случайное число каких-то событий (радиоактивных распадов, телефонных вызовов, отказов оборудования, несчастный случаях и т.п.).

Случайное число событий, происшедших за время от 0 до T, распределено по закону Пуассона с параметром $\lambda = aT$, где $a > 0$ – параметр задачи, отражающий среднюю частоту событий [1, 5, 11].

$$P_{(Z=k)} = \frac{\lambda^k}{k!} e^{-\lambda}. \tag{7.12}$$

Вероятность k событий в течение большого интервала времени.

Непрерывные законы распределения

Нормальное распределение занимает центральное место в теории и практике вероятностно-статистических исследований. Непрерывная случайная величина X называется распределенной по нормальному закону, если ее плотность распределения равна:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2, \tag{7.13}$$
$$-\infty < x < \infty,\ \sigma > 0,\ -\infty < \mu < \infty.$$

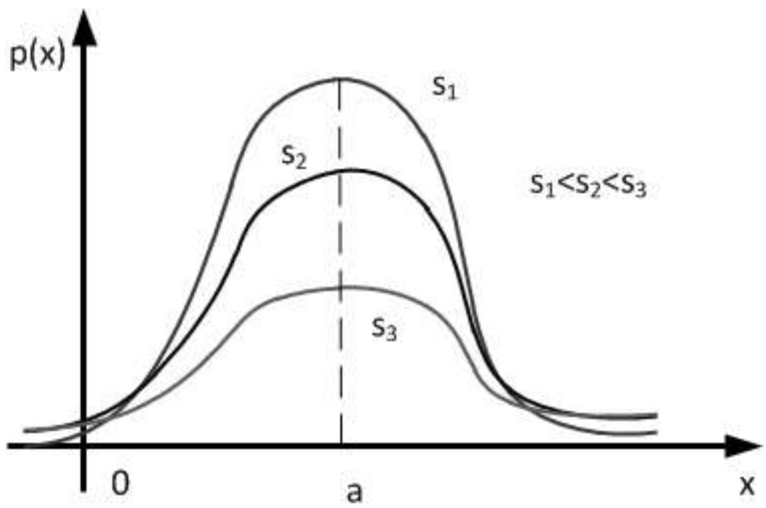


Рис. 7.2. Плотность вероятности для нормального распределения; S_1, S_2, S_3 – стандартное отклонение

Математическое ожидание случайной величины, $s = \sigma$ стандартное квадратичное отклонение. Нормальное распределение с параметрами $\mu = 0$ и $\sigma = 1$ называется нормированным

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}. \tag{7.14}$$

Равномерное распределение

Плотность вероятности равномерного распределения равна:

$$p(x) = \frac{1}{N}. \qquad (7.15)$$

Равномерное распределение вероятностей является простейшим и может быть как дискретным, так и непрерывным. Дискретное равномерное распределение – это такое распределение, для которого вероятность каждого из значений СВ одна и та же.

Распределение вероятностей непрерывной СВ X, принимающие все свои значения из отрезка [a; b] называется равномерным, если ее плотность вероятности на этом отрезке постоянна, а вне его равна нулю (рис. 7.3):

$$p(x) = \begin{cases} 0 \text{ при } x < a, \\ \dfrac{1}{b-a} \text{ при } a \le x \le b, \\ 0 \text{ при } x > b. \end{cases} \qquad (7.16)$$

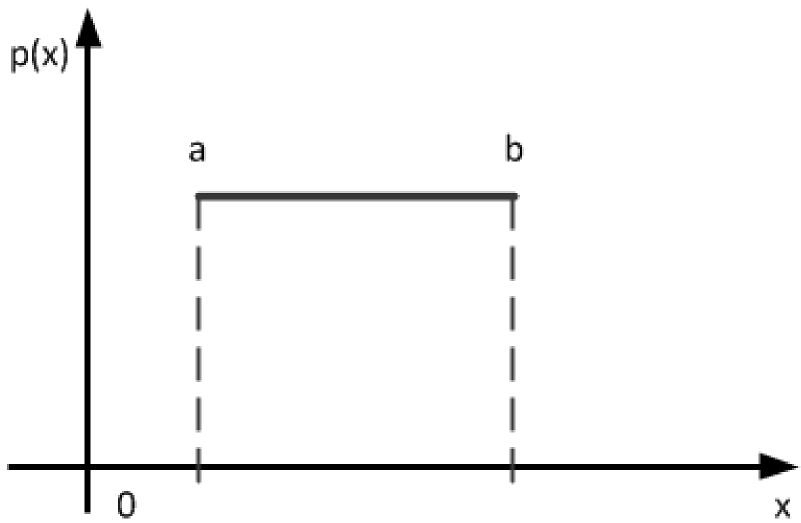


Рис. 7.3. Плотность вероятности равномерного распределения

Равномерно распределенное случайное число можно получить, если использовать стандартный генератор случайных чисел ЭВМ. Для генерации числа используется формула:

$$p = (n_2 - n_1 + 1) \times r + n_1. \qquad (7.17)$$

Здесь $n_2 > n_1$ границы диапазона генерации случайного числа r – равномерно-распределенное случайное число, полученное от ГСЧ [3, 4].

Генерация случайного числа в коде Java:

```
double n2=20,n1=5;
double r=(n2-n1+1)*Math.random()+n1;
```

Контрольные вопросы

1. Охарактеризуйте понятие случайной величины.
2. В чем разница между вероятностью и частотой появления события?
3. Что такое математическое ожидание и дисперсия случайной величины?
4. Что такое плотность распределения случайной величины?
5. Как получить закон распределения случайной величины?
6. Перечислите параметры описательной статистики.
7. Приведите примеры законов дискретных законов распределения случайной величины.
8. Приведите примеры непрерывных законов распределения случайной величины.
9. Как выполняется генерация равномерно случайного числа путем использования встроенного генератора случайных чисел ЭВМ?

ЛЕКЦИЯ №8. ПРОВЕРКА СТАТИСТИЧЕСКОЙ ГИПОТЕЗЫ

Понятие статистической гипотезы, нулевая гипотеза, ошибки при проверке гипотезы, методы проверки статистической гипотезы.

Статистическая гипотеза – это предположение о свойствах случайных величин или событий, которое нужно проверить по имеющимся данным.

Нулевая гипотеза – это основное проверяемое предположение, которое обычно формулируется как отсутствие различий, отсутствие влияние фактора, отсутствие эффекта, равенство нулю значений выборочных характеристик.

Например: имеется механическое устройство, выходное звено которого совершает вращательное движение.

Гипотеза: увеличение скорости вращения выходного звена не влияет на устойчивость процессов внутри механизма.

Другое проверяемое предположение (не всегда строго противоположное или обратное первому) называется конкурирующей или альтернативной гипотезой.

Нулевая гипотеза обозначается H_0, а альтернативная как H_1. Выдвинутая гипотеза может быть правильной или неправильной, поэтому возникает необходимость проверить ее. Так как проверку производят статистическими методами, то данная проверка называется статистической.

При проверке статистических гипотез возможны ошибки (ошибочные суждения) двух видов:

$$H_0 = TRUE;$$
$$H_1 = FALSE.$$

Можно отвергнуть нулевую гипотезу, когда она на самом деле верна (так называемая ошибка первого рода);

Если имеет место:

$$H_0 = FALSE;$$
$$H_1 = TRUE.$$

Можно принять нулевую гипотезу, когда она на самом деле не верна (так называемая ошибка второго рода).

Ошибка, состоящая в принятии нулевой гипотезы, когда она ложна, качественно отличается от ошибки, состоящей в отвержении гипотезы, когда она истинна. Эта разница очень существенна вследствие того, что различна значимость этих ошибок.

Допустимая вероятность ошибки первого рода ($P_{кр}$) может быть равна 5% или 1% (0.05 или 0.01). Уровень значимости – это вероятность ошибки первого рода при принятии решения (вероятность ошибочного отклонения нулевой гипотезы). Альтернативные гипотезы принимаются тогда и только тогда, когда опровергается нулевая гипотеза.

Риск ошибки отвергнуть нулевую гипотезу и принять альтернативную не превышает одного из трех принятых уровней значимости статистического вывода:

- первый уровень – 5% (p = 5%), где допускается риск ошибки в выводе в пяти случаях из ста теоретически возможных таких же экспериментов;
- второй уровень – 1%, т.е. соответственно допускается риск ошибиться только в одном случае из ста;
- третий уровень – 0,1%, т.е. допускается риск ошибиться только в одном случае из тысячи.

Статистика критерия

Статистика критерия (T) – некоторая функция от исходных данных, по значению которой проверяется нулевая гипотеза. Всякое правило, на основе которого отклоняется или принимается нулевая гипотеза, называется критерием для проверки данной гипотезы.

Статистический критерий – это случайная величина, которая служит для проверки статистических гипотез. Критическая область – совокупность значений критерия, при котором нулевую гипотезу отвергают.

Анализ двух выборок может производиться с помощью параметрических критериев. Данные методы основываются на предположении о том, что распределение выборок подчиняется нормальному закону распределения. Рассмотрим два критерия: критерий Стьюдента и критерий Фишера.

Критерий Стьюдента (*t*-критерий)

Критерий позволяет найти вероятность того, что оба средних значения в выборке относятся к одной и той же совокупности. Данный критерий наиболее часто используется для проверки гипотезы: «Средние двух выборок относятся к одной и той же совокупности».

В первом случае его применяют для проверки гипотезы о равенстве генеральных средних двух независимых, несвязанных выборок (так называемый двух выборочный t-критерий). В этом случае есть контрольная выборка и экспериментальная выборка, количество значений в выборках различно. Статистика критерия вычисляется по формуле (8.1). Стандартное отклонение вычисляется по формулам (8.2) или (8.3). Формула (8.3) используется, если $n_1 = n_2 = n$.

$$t_{\text{эмп}} = \frac{\bar{x} - \bar{y}}{\sigma_{x-y}}; \tag{8.1}$$

$$\sigma_{x-y} = \sqrt{\frac{\sum (x_i - \bar{x}) + (y_i - \bar{y})^2}{n_1 + n_2 - 2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}; \tag{8.2}$$

$$\sigma_{x-y} = \sqrt{\frac{\sum (x_i - \bar{x}) + (y_i - \bar{y})^2}{(n-1) \cdot n}}. \tag{8.3}$$

Подсчет числа степеней свободы осуществляется по формуле

$$k = n_1 + n_2 - 2. \tag{8.4}$$

При численном равенстве выборок

$$k = 2n - 2. \tag{8.5}$$

Во втором случае, когда одна и та же группа объектов порождает числовой материал для проверки гипотез о средних, используется так называемый парный t-критерий. Выборки при этом называют зависимыми, связанными. Расчет ведется по формулам (8.6) и (8.7).

$$t_{\text{эмп}} = \frac{\bar{d}}{S_d}; \quad d_i = x_i - y_i; \tag{8.6}$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n}}{n \cdot (n-1)}}. \tag{8.7}$$

Далее необходимо сравнить полученное значение $t_{\text{эмп}}$ с теоретическим значением t – распределения Стьюдента. Если $t_{\text{эмп}} < t_{\text{крит}}$, то гипотеза H_0 принимается, в противном случае нулевая гипотеза отвергается и принимается альтернативная гипотеза.

Критерий Фишера

Критерий Фишера позволяет сравнивать величины выборочных дисперсий двух независимых выборок. Для вычисления статистики критерия $F_{\text{эмп}}$ нужно найти отношение дисперсий двух выборок, причем так, чтобы большая по величине дисперсия находилась бы в числителе, а меньшая – в знаменателе.

$$F_{\text{ЭМП}} = \frac{\sigma_x^2}{\sigma_y^2}. \tag{8.8}$$

Так как, согласно условию критерия, величина числителя должна быть больше или равна величине знаменателя, то значение $F_{\text{ЭМП}}$ всегда будет больше или равно единице.

Число степеней свободы определяется по формулам:

$k_1 = n_1 - 1$ для первой выборки (величина дисперсии которой больше);

$k_2 = n_2 - 1$ для второй выборки.

Если $t_{\text{эмп}} < t_{\text{крит}}$, то нулевая гипотеза принимается, в противном случае принимается альтернативная.

Непараметрический анализ выборок

Непараметрические критерии анализа выборок позволяют проводить анализ, не учитывая закон распределения. Такой анализ может быть проведен с помощью критерия знаков (G-критерий). Критерий предназначен для сравнения состояния некоторого свойства у членов двух зависимых выборок на основе измерений, сделанных по шкале не ниже ранговой.

Пусть имеется две серии наблюдений над случайными переменными X и $У$, полученные при рассмотрении двух зависимых выборок. На их основе составлено N пар вида (x_i, y_i), где x_i, y_i – результаты двукратного измерения одного и того же свойства у одного и того же объекта. Элементы каждой пары x_i, y_i сравниваются между собой по величине, и паре присваивается знак:

«+», если $x_i < y_i$,

«–», если $x_i > y_i$,

«0», если $x_i = y_i$.

Нулевая гипотеза формулируется следующим образом: в состоянии изучаемого свойства нет значимых различий при первичном и вторичном измерениях.

Альтернативная гипотеза: законы распределения величин X и $У$ различны, т.е. состояния изучаемого свойства существенно различны в одной и той же совокупности при первичном и вторичном измерениях этого свойства.

Нулевая гипотеза принимается на уровне значимости 0,05, если наблюдаемое значение $T < n - ta$, где значение $n - ta$ определяется из статистических таблиц для критерия знаков.

Пример. Студенты выполняли контрольную работу, направленную на проверку усвоения некоторого понятия. Пятнадцати студентам затем предложили электронное пособие, составленное с целью формирования данного понятия у студентов с низким уровнем обучаемости. После изучения пособия студенты снова выполняли ту же контрольного работу, которая оценивалась по пятибалльной системе. Результаты проверки приведены в табл. 8.1.

Таблица 8.1

Статистика критерия

Студент №	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Первое выполнение, x_i	2	2	2	2	2	3	3	3	3	3	2	2	3	3	3
Второе выполнение, y_i	2	3	3	4	3	2	3	4	4	3	4	3	2	4	4
Знак разности	0	+	+	+	+	–	0	+	+	0	+	+	–	+	+

Подсчитаем значение статистики критерия T, равное числу положительных разностей отметок, полученных студентами. Согласно данным таблицы $T = 10$. Исключим нулевые пары, в результате получим $n = 12$ не нулевых пар. Для уровня значимости $a = 0{,}05$ при $n = 12$ табличное значение $n - ta = 9$. Следовательно выполняется неравенство $T > n - ta = 10 > 9$.

В соответствии с правилом принятия решения нулевая гипотеза отклоняется на уровне значимости 0,05 и принимается альтернативная гипотеза, что позволяет сделать вывод об улучшении знаний студентов после самостоятельного изучения пособия.

Контрольные вопросы

1. Что такое статистическая гипотеза?
2. Что называется нулевой гипотезой?
3. Какие ошибки можно выделить при проверке статистической гипотезы?
4. Опишите алгоритм использования критерия Стьюдента для проверки статистической гипотезы.
5. Опишите алгоритм использования критерия Фишера для проверки статистической гипотезы.
6. Опишите алгоритм использования критерия знаков для проверки статистической гипотезы.

СПИСОК ЛИТЕРАТУРЫ

Основная

1. Гнеденко, Б.В. Введение в теорию массового обслуживания / Б.В. Гнеденко. – М.: «Наука», 2007. – 430 с.
2. Ким, Д.П. Теория автоматического управления. В 2 ч. Ч. 1. Линейные системы / Д.П. Ким. – М.: «ФИЗМАТЛИТ», 2007. – 312 с.
3. Мезенцев, К.Н. Моделирование систем в среде AnyLogic 6.4.1. В 2 ч. Ч. 1: учебное пособие / К.Н. Мезенцев; под ред. А.Б. Николаева. – М.: МАДИ, 2011. – 105 с.
4. Мезенцев, К.Н. Моделирование систем в среде AnyLogic 6.4.1. В 2 ч. Ч. 2: учебное пособие / К.Н. Мезенцев; под ред. А.Б. Николаева. – М.: МАДИ, 2011. – 112 с.
5. Советов, Б.Я. Моделирование систем: учебник / Б.Я. Советов. – М.: «Высшая школа», 2009. – 343 с.
6. Язык UML. Руководство пользователя / пер. с англ. Н. Мухин; Г. Буч [и др.]. – 2-е изд. – М.: «ДМК Пресс», 2006. – 496 с.

Дополнительная

7. Богданов, А.А. Тектология. Всеобщая организационная наука / А.А. Богданов. – М.: «Финансы», 2003. – 496 с.
8. Винер, Н. Кибернетика, или Управление и связь в животном и машине / пер. с англ. И.В. Соловьева и Г.Н. Поварова; Н. Винер; под ред. Г.Н. Поварова. – 2-е изд. – М.: «Наука», 1983. – 344 с.
9. Карпов, Ю. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5 / Ю. Карпов. – Спб.: «БХВ», 2005 – 400 с.
10. Красовский, А.А. Справочник по теории автоматического управления / А.А. Крассовский. – М.: «Наука», 1987. – 711 с.
11. Осипов, Л.А. Проектирование систем массового обслуживания / Л.А. Осипов. – М.: «Адвансед Солюшнз», 2011. – 111 с.
12. Пригожин, И. Порядок из хаоса: Новый диалог человека с природой / пер. с англ. общ. ред. В.И. Аршинова, Ю.Л. Климонтовича и Ю.В. Сачкова; И. Пригожин. – М.: «Прогресс», 1986. – 432 с.
13. Рассел, С. Искусственный интеллект: современный подход / пер. с англ.; С. Рассел. – 2-е изд. – М.: «Издательский дом Вильямс», 2006. – 1408 с.
14. Эшби, У.Р. Введение в кибернетику / пер. с англ. Делир Лахути; У.Р. Эшби. – М.: «Либроком», 2009. – 432 с.