А.М. Инков

МОДЕЛИРОВАНИЕ И ИДЕНТИФИКАЦИЯ ОБЪЕКТОВ УПРАВЛЕНИЯ

КОНСПЕКТ ЛЕКЦИЙ



Шымкент, 2011

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН

ЮЖНО-КАЗАХСТАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ им. М.О.АУЕЗОВА

Инков А.М

МОДЕЛИРОВАНИЕ И ИДЕНТИФИКАЦИЯ ОБЪЕКТОВ УПРАВЛЕНИЯ

Конспект лекций
для студентов специальности
5В070200 (050702) «Автоматизация и управление»

Шымкент, 2011г.

УДК 658.012.011
ББК 6П7.1

Инков А.М. Моделирование и идентификация объектов управления: Конспект лекций для студентов специальности 5В070200 (050702). Шымкент, Южно-Казахстанский Государственный Университет им. М.О. Ауезова,. 2011 г. - 156 с.

ISBN _____

Конспект лекций разработан на основании государственного общеобязательного стандарта образования Республики Казахстан ГОСО РК ГОСО 03.08.328-2006 специальности 050702 – «Автоматизация и управление» и типовой программы дисциплины «Моделирование и идентификация объектов управления», разработанной Институтом автоматики и управления Казахского национального технического университета имени К.И.Сатпаева. (Сборник «Бакавлариат. Типовые учебные программы по специальности 050702-Автоматизация и управление. МОРК, Алматы, 2006г.»)/

Конспект лекций содержит материал, предусмотренные программой дисциплины и предназначен для студентов специальности 5В070200 (050702) «Автоматизация и управление» всех видов обучения и смежных специальностей. Содержит материл лекций и сведения, необходимые для выполнения СРС и СРСП.

Рецензенты:
д.т.н., профессор кафедры ПАОС ЮКГУ Ескендиров Ш.З.
к.т.н., доцент, руководитель филиала №89 НТЦ МОНРК Каланов С.М.

Учебно-методическое пособие рекомендовано к изданию Методическим советом ЮКГУ им. М.Ауезова, протокол № ___ от «____» _____ 2010 г.

© Южно-Казахстанский государственный университет им. М.Ауезова, 2011г.
© Инков А.М., 2011 г.

Содержание

Стр.

Модуль 1. Моделирование и идентификация статических характеристик объектов....5
Тема 1. Введение....5
Лекция 1 Вводная....5
Тема 2 Математические модели объектов идентификации....21
Лекция 2 Общие сведения о математических моделях и их классификация....21
Тема 3 Статические и динамические модели в форме уравнения регрессии....37
Лекция 3 Статические и динамические модели в форме уравнения регрессии (ч. 1)....37
Лекция 4 Статические и динамические модели в форме уравнения регрессии (ч. 2)....53
Лекция 5 Статические и динамические модели в форме уравнения регрессии (ч.3)....71
Лекция 6 Статические и динамические модели в форме уравнения регрессии (ч. 4)....83
Модуль 2. Моделирование и идентификация динамических характеристик объектов....95
Тема 4 Общие задачи идентификации....95
Лекция 7 Общие задачи идентификации. Основные определения.....95
Тема 5 Методы идентификации на основе простейших тестирующих сигналов....110
Лекция 8 Определение переходных характеристик....110
Лекция 9 . Определение частотных характеристик....128
Тема 6 Общая схема процесса идентификации....141
Лекция 10 Общая схема процесса идентификации....141
Тема 7 Проблема идентифицируемости....160
Лекция 11 Проблема идентифицируемости....160
Тема 8 Структурная статистическая идентификация....168
Лекция 12 Структурная статистическая идентификация....168
Тема 9 Оценивание параметров и состояния объектов....178
Лекция 13 Оценивание параметров и состояния объектов....178
Тема 10 Методы идентификации с настраиваемыми адаптивными моделями....201
Лекция 14 Методы идентификации с настраиваемыми адаптивными моделями....201
Тема 11 Методы идентификации нелинейных динамических характеристик....208
Лекция 15 Методы идентификации нелинейных динамических характеристик....208
Приложение А. Условные обозначения....217
Приложение Б. Глоссарий. Основная терминология....219
Приложение В. Матрица наблюдаемости и фундаментальная матрица....227
Методическое обеспечение дисциплины и ТСО....231

Модуль 1. Моделирование и идентификация статических характеристик объектов

Тема 1. Введение

Лекция 1 Вводная

В этой вводной лекции мы рассмотрим следующие проблемы и вопросы: краткая справка о развитии и формировании методов идентификации; философские аспекты моделирования; идентификация в процессах управления; основные понятия о моделях и методах их построения; неизбежность упрощения модели по сравнению с реальным объектом; отражение свойств объекта, существенных для цели моделирования; адекватность и критерии адекватности модели.

Цели и задачи дисциплины Дисциплина «Моделирование и идентификация объектов управления» является одним из обязательных направлений в современной подготовке специалистов по специальности 050702 - Автоматизация и управление к производственной и исследовательской работе по эксплуатации и разработке автоматизированных систем управления технологическими процессами, техническими системами и объектами управления и соответствует государственному общеобязательному стандарту высшего профессионального образования. Данное пособие составлено на основании государственного общеобязательного стандарта образования Республики Казахстан ГОСО РК ГОСО 03.08.328-2006 специальности 050702 – «Автоматизация и управление» и типовой программы дисциплины, разработанной Институтом автоматики и управления Казахского национального технического университета имени К.И.Сатпаева, Алматы, 2006г.

Целью преподавания дисциплины «Моделирование и идентификация объектов управления» является разработка математической модели и реализация процесса идентификации автоматизируемою технологического процесса (ТП), а основной задачей построение модели по результатаём наблюдений, которая являлась одной из основных задач теории автоматического управления.

В процессе обучения студенты должны получить знания: принципов системного подхода при решении задач автоматизации и управления; основных методов и алгоритмов; особенностей исследований непрерывных и дискретных технических систем и объектов управления; основных положений автоматизации как научных, так и промышленных исследований в задачах идентификации и моделирования технических систем. Студент должен приобрести умения: ориентироваться и обоснованно применять методы и алгоритмы; проводить исследования и обрабатывать результаты с целью получения математических моделей в рамках процесса проектирования и построения систем управления объектами различной физической природы. В результате специалист получает навыки решения творческих, исследовательских и производственных задач на основе самостоятельности освоения и проработки технического задания на разработку; изучения научно-технической литературы; применение вычислительной техники для реализации разрабатываемых или осваиваемых алгоритмов.

Основная задача изучения дисциплины. В соответствии с ГОСО 03.08.328-2006 «Специальности 050702 – Автоматизация и управление. Бакалавриат» студенты должны изучить следующее.

Классификация моделей и виды моделирования; примеры моделей систем; основные положения теории подобия; этапы математического моделирования; принципы построения и основные требования к математическим моделям систем; цели и задачи исследования математических моделей систем; общая схема разработки математических моделей; формализация процесса функционирования системы; построение математических моделей объектов и систем по экспериментальным данным; структурная и параметрическая идентификация; методы построения статических и динамических моделей объектов управления; описание модели при взаимодействии с внешней средой; модели возмущений; критерии и показатели качества идентификации; методы идентификации; условия идентифицируемости объектов; структурная идентификация; общая схема оценивания, Байесово оценивание, оценивание по методу максимального правдоподобия и методу наименьших квадратов; итеративные алгоритмы идентификации, программные средства моделирования объектов и систем.

Приведенное ниже содержание лекционного материала является основной для самостоятельной работы по изучению материала дисциплины при подготовке к лекции и отчетности. В процессе прослушивания лекции студент должен иметь при себе эти методические указания, используя их для пометок. Это позволяет значительно сократить время на конспектирование материала.

Задачу идентификации характеристик системы можно рассматривать как дуальную по отношению к задаче управления системой. Нельзя управлять системой, если она не идентифицирована либо заранее, либо в процессе управления. Например, мы не можем управлять автомобилем, пока не познакомимся с его реакцией на поворот руля, нажатие акселератора или тормоза, т. е. пока не ознакомимся со свойствами автомобиля. Этот процесс освоения автомобиля («привыкания» к нему) и представляет собой процесс идентификации. Таким образом, идентификацию реакции автомобиля мы осуществляем и в том случае, когда нам не известна система описывающих его дифференциальных уравнений. В общем случае, если необходимо перевести систему из состояния A в состояние B, то можно положиться либо на свое умение управлять системой, либо изучить реакции системы на одно или несколько управляющих воздействий. Если априори известно, что воздействие Uj переводит систему ближе к состоянию B, то следует прилагать именно это входное воздействие. В отсутствие такого априорного знания можно измерять реакции системы на ряд входных воздействий, выполняя таким образом по существу идентификацию. Знание результатов идентификации до начала процесса управления существенно влияет на его реализацию.

Выявление дифференциальных уравнений процесса представляет собой одну из возможных, но не единственную форму идентификации. Можно, например, составить таблицу возможных управляющих воздействий и соответствующих им откликов системы в заданном интервале времени. Из этой таблицы можно затем легко определить лучшие (с точки зрения преследуемой цели) процессы управления. Подобно этому могут быть сформированы идентификационные модели и на основании других методов описания процессов.

В дисциплине рассматриваются различные методы идентификации, основанные на разных подходах к форме задания идентификационных моделей (например, дифференциальные уравнения, разностные уравнения, передаточные функции, градиентные выражения и т. п.).

Ни один из обсуждаемых методов идентификации не годится для идентификации всех видов систем. Каждый из них имеет свою область или области применения. Это, однако, не означает, что на современном уровне идентификация должна рассматриваться как набор готовых рецептов для различных типов систем. Сейчас уже можно говорить о теории идентификации, имеющей дело с оцениванием параметров на основании измеренных текущих входных и выходных данных, причем качество идентификации повышается с увеличением числа измерений. Ошибки идентификации, естественно, приводят к ошибкам в управлении или в требуемом выходном параметре системы; эти ошибки могут быть использованы для дальнейшего улучшения идентификации. Следовательно, теория идентификации аналогична, точнее, дуальна теории управления, в которой ошибки управления (в предположении, что система идентифицирована) используются для улучшения последующего процесса управления. Аналогично теории управления в теории идентификации существует несколько подходов, применяемых ко многим ситуациям и случаям.

Теория идентификации распространена на случай оценивания параметров устройств предсказания и фильтров. Это объясняется тесной взаимосвязью задач предсказания и идентификации, поскольку идентификация проводится обычно в целях облегчения предсказания поведения идентифицируемой системы в будущем. Однако задача предсказания отличается от задачи идентификации тем, что последняя для предсказания поведения в будущем рассматривает соотношения входов и выходов системы при заданных параметрах и входных воздействиях системы. Предсказание временных рядов основано на анализе измеренных значений, однако входные воздействия часто недоступны измерению и полностью неизвестны. Поэтому идентификация параметров устройств предсказания основана лишь на использовании предшествующих измерений сигналов, значения которых в будущем необходимо предсказать (и которые рассматриваются как выход системы, чей вход недоступен измерению), а воспользоваться данными о соотношениях входов и выходов нельзя.

Вообще говоря, различают несколько характерных ситуаций, для которых необходимы различные методы исследования. Во-первых, различают системы линейные и нелинейные, причем линейные системы легче идентифицировать, поскольку они обладают свойствами суперпозиции. Во-вторых, различают системы стационарные и нестационарные (к последним относятся системы с изменяющимися во времени параметрами). Системы могут считаться стационарными, если их параметры меняются медленно по сравнению со временем, которое требуется для точной идентификации. В-третьих, системы часто делятся на дискретные и непрерывные, хотя преобразовать непрерывную формулировку задачи в дискретную обычно довольно просто. Четвертый вариант классификации различает методы идентификации для систем с одним или несколькими входными воздействиями. Это деление целесообразно вводить потому, что методы идентификации значительно упрощаются, если на систему подается лишь одно входное воздействие, по сравнению со случаем, когда на систему действует одновременно комбинация нескольких возмущений или входных воздействий. Пятый вариант классификации предусматривает возможность идентификации детерминированных или стохастических процессов. При идентификации последних ориентируются в основном на вероятностные представления о точном состоянии системы. (На практике все результаты измерений засорены шумом и для точной идентификации необходимо осуществить фильтрацию или сглаживание). При идентификации детерминированных систем обычно предполагается, что фильтрация уже была проведена. Шестой, и, возможно, наиболее важный, но трудно осуществимый, вариант классификации — классификация методов идентификации в зависимости от наличия априорной информации о системе. При классификации систем по признакам линейности или стационарности также используют априорную информацию. Эти признаки (линейность и стационарность), если они заранее неизвестны, конечно, могут быть установлены в процессе анализа результатов измерений. При любом методе идентификации очень важным является знание размерности вектора состояния и природы внутренних связей или нелинейностей.

В основу перечисленных способов классификации положена по существу степень сложности идентификации. Очевидно, идентифицировать детерминированный линейный стационарный процесс известного порядка с одним входом существенно проще, чем аналогичный стохастический процесс неизвестного порядка, который может быть нелинейным и нестационарным.

Краткая справка о развитии и формировании методов идентификации

Математические модели являются одним из основных инструментов познания человеком явлений окружающего мира. Под математическими моделями понимают основные закономерности и связи, присущие изучаемому явлению. Это могут быть формулы или уравнения, наборы правил или соглашений, выраженные в математической форме. Испокон веков в математике, механике, физике и других точных науках естествознания для описания изучаемых ими явлений использовались математические модели. Так, законы Ньютона полностью определяют закономерности движения планет вокруг Солнца. Используя основные законы механики, относительно нетрудно составить уравнения, описывающие движение космического аппарата, например, от Земли к Луне. Однако получить их решение в виде простых формул не представляется возможным. Для расчета траекторий космических аппаратов служат компьютеры.

Применение компьютеров для математического моделирования изменило само понятие "решить задачу". До этого исследователь удовлетворялся написанием математической модели. А если ему еще удавалось доказать, что решение (алгоритм) в принципе существует, то этого было достаточно, если априори полагать, что модель адекватно описывает изучаемое явление. Поскольку, как правило, нет простых формул, описывающих поведение модели, а стало быть и объекта, который описывается моделью, то единственный путь - свести дело к вычислениям, применению численных методов решения задач. В таком случае необходим конкретный алгоритм, указывающий последовательность вычислительных и логических операций, которые должны быть произведены для получения численного решения. С алгоритмами связана вся история математики. Само слово "алгоритм" является производным от имени средневекового узбекского ученого Аль-Хорезми. Еще древнегреческим ученым был известен алгоритм нахождения числа "пи" с высокой точностью. Ньютон предложил эффективный численный метод решения алгебраических уравнений, а Эйлер - численный метод решения обыкновенных дифференциальных уравнений. Как известно, модифицированные методы Ньютона и Эйлера до сих пор занимают почетное место в арсенале вычислительной математики. Ее предметом являются выбор расчетной области и расчетных точек, в которых вычисляются характеристики моделируемого объекта, правильная замена исходной математической модели ее аналогом, пригодным для расчета, т. е. некоторой дискретной моделью. Поскольку модели должны представлять изучаемые явления в необходимой полноте, понятно, что они становятся весьма сложными.

В модели входят множество величин, подлежащих определению, а сами эти величины зависят от большого числа переменных и постоянных параметров.

Наконец, модели реальных процессов оказываются нелинейными. Аппарат классической математической физики приспособлен для работы с линейными моделями. В этом случае сумма (суперпозиция) частных решений уравнения есть также его решение. Найдя частное решение уравнения для линейной модели, с помощью принципа суперпозиции можно получить решение в общем случае. На этом пути в традиционной математической физике были получены замечательные результаты. Однако она становится бессильной, если встречается с нелинейными моделями. Принцип суперпозиции здесь неприменим, и алгоритмов для построения общего решения не существует. Поэтому для нелинейных моделей законченных теоретических результатов получено немного.

Методология математического моделирования в кратком виде выражена знаменитой триадой "модель - алгоритм - программа", сформулированной академиком А. А. Самарским, основоположником отечественного математического моделирования. Эта методология получила свое развитие в виде технологии "вычислительного эксперимента", разработанной школой А. А. Самарского, - одной из информационных технологий, предназначенной для изучения явлений окружающего мира, когда натурный эксперимент оказывается слишком дорогим и сложным.

Во многих важных областях исследований натурный эксперимент невозможен, потому что он либо запрещен (например, при изучении здоровья человека), либо слишком опасен (например, при изучении экологических явлений), либо просто неосуществим (например, при изучении астрофизических явлений).

Вычислительный эксперимент в отличие от натурных экспериментальных установок позволяет накапливать результаты, полученные при исследовании какого-либо круга задач, а затем быстро и гибко применять их к решению задач в совершенно других областях. Этим свойством обладают используемые универсальные математические модели. Например, уравнение нелинейной теплопроводности пригодно для описания не только тепловых процессов, но и диффузии вещества, движения грунтовых вод, фильтрации газа в пористых средах. Изменяется только физический смысл величин, входящих в это уравнение.

Проведение вычислительного эксперимента можно условно разделить на два этапа. После первого этапа вычислительного эксперимента, если надо, модель уточняется как в направлении ее усложнения (учет дополнительных эффектов и связей в изучаемом явлении), так и упрощения (выяснение, какими закономерностями и связями в изучаемом явлении можно пренебречь). На последующих этапах цикл вычислительного эксперимента повторяется до тех пор, пока исследователь не убеждается, что модель адекватна тому объекту, для которого она составлена.

Информационные технологии, поддерживающие вычислительный эксперимент, включают в себя методы построения математических моделей силами конечных пользователей информационных систем (специалистов в своей предметной области, а не профессиональных математиков и программистов), информационную поддержку их деятельности для поиска и выбора алгоритмов и программ численного решения задач, методы и средства контроля точности производимых вычислений и правильности работы применяемых программ. При проведении вычислительного эксперимента исследователь может с помощью пользовательского интерфейса "играть" на модели, ставя интересующие его вопросы и получая ответы. Таким образом, исследователь получает мощный инструмент для анализа и прогноза поведения сложных нелинейных многопараметрических объектов и явлений, изучение которых традиционными методами затруднено или вообще невозможно.

Дата появления первых серьезных результатов вычислительного эксперимента в СССР зафиксирована вполне официально - 1968 год, когда Госкомитет СССР по делам открытий и изобретений засвидетельствовал открытие явления, которого на самом деле никто не наблюдал. Это было открытие, так называемого, эффекта Т-слоя (температурного токового слоя в плазме, которая образуется в МГД-генераторах). Свидетельство на это открытие было выдано академикам А. Н. Тихонову и А. А. Самарскому, члену-корреспонденту АН СССР С. П. Курдюмову, докторам физико-математических наук П. П. Волосевичу, Л. М. Дегтяреву, Л. А. Заклязьминскому, Ю. П. Попову (ныне директору ИПМ им. М. В. Келдыша РАН), В. С. Соколову и А. П. Фаворскому. В данном

случае вычислительный эксперимент предшествовал натурному. Натурные эксперименты "заказывались" по результатам математического моделирования. Через несколько лет в трех физических лабораториях на разных экспериментальных установках практически одновременно был надежно зарегистрирован Т-слой, после чего технологам и инженерам стал окончательно ясен принцип работы МГД-генератора с Т-слоем.

Плазма с ее нелинейными свойствами стала одним из важнейших объектов математического моделирования и вычислительного эксперимента. Заманчивая перспектива решения энергетической проблемы связана с управляемым термоядерным синтезом изотопов водорода, дейтерия и трития. Энергетическая проблема для человечества заключается в том, что нефти и газа при нынешнем темпе их потребления хватит всего на несколько десятков лет. А сжигать столь ценное химическое сырье в топках электростанций и двигателях внутреннего сгорания - это, по образному выражению Д. И. Менделеева, "почти все равно, что топить печь ассигнациями". С запасами угля дело обстоит гораздо лучше, но его добыча с каждым годом становится все труднее. Выходом может быть лазерный термоядерный управляемый синтез, исследование которого осуществляется с помощью вычислительного эксперимента. В 1974 г. коллектив сотрудников ФИАН и ИПМ АН СССР под руководством академиков Н. Г. Басова, А. Н. Тихонова и А. А. Самарского предложил принципиально новую концепцию лазерного термоядерного синтеза на основе результатов вычислительного эксперимента.

Еще одна область использования вычислительного эксперимента - это "вычислительная технология" - применение математического моделирования с помощью компьютеров не только для решения фундаментальных научных проблем, но и для разработки технологических процессов в промышленности. Для тех случаев, когда технологические процессы описываются хорошо известными математическими моделями, для расчета которых предложены эффективные вычислительные алгоритмы, разработаны пакеты прикладных программ, технология вычислительного эксперимента позволяет создавать новые программы и совершенствовать средства общения человека с компьютером. У технологов есть потребность в изучении новых промышленных технологий, например лазерно-плазменной обработки материалов (плазменной термохимии).

Основатель нобелевских премий Альфред Нобель, как известно, исключил математику из числа наук, за достижения в которых присуждается эта высшая научная награда. Вместе с тем, современное математическое моделирование охватывает области исследований, до недавнего времени недоступные математике. В последние годы ряд Нобелевских премий по химии, медицине, экономике, физике элементарных частиц были присуждены работам, методологическую основу которых составляло математическое моделирование.

Например, для дальнейшего исследования нелинейных процессов в микромире были разработаны соответствующие численные методы с применением компьютеров и компьютерных сетей (сетевых grid-технологий), ориентированные на решение задач физики элементарных частиц. Алгоритмы квантово-механических расчетов прогрессируют не менее быстрыми темпами, чем в других областях вычислительной математики.

Биология во многом остается экспериментальной и описательной дисциплиной, а история математического моделирования биологических процессов вряд ли насчитывает более 20 лет. И все-таки уже можно назвать биологические задачи, для которых вычислительный эксперимент становится определяющей методологией.

Математическое моделирование и вычислительный эксперимент - ведущие методологии изучения глобальных моделей процессов и явлений на Земле, например климата Земли. Проведение работ по глобальному моделированию стимулировалось деятельностью Римского клуба, неправительственной организации. Первую из таких моделей опубликовал в 1971 г. американский специалист по теории управления Д. Форрестер.

Компьютерные игры, проведенные Д. Форрестером с глобальной моделью, показали, что в середине XXI века человечество ждет кризис, связанный прежде всего с истощением природных ресурсов, падением численности населения и производства продуктов, ростом загрязнения окружающей среды.

Известны результаты глобального моделирования явления "ядерной зимы", выполненные в ВЦ АН СССР В. В. Александровым и Г. Л. Стенчиковым под руководством академика Н. Н. Моисеева. Эти результаты дали человечеству, в том числе политикам, неопровержимые аргументы против ядерной войны, даже так называемой "ограниченной ядерной войны".

Для математического моделирования и вычислительного эксперимента использовались, главным образом, универсальные цифровые вычислительные машины, доступные коллективам исследователей. В СССР в 70-80-х годах прошлого века это были БЭСМ-6 и модели ЕС ЭВМ, для которых разрабатывались библиотеки и пакеты прикладных программ вычислительной математики. С появлением персональных компьютеров стало возможно развитие информационной технологии вычислительного эксперимента, которая предусматривает поддержку пользовательского интерфейса и поиска нужных алгоритмов и программ с помощью персональных компьютеров (отечественного производства или импортных), а проведение расчетов на математических моделях - с помощью высокопроизводительных компьютеров БЭСМ-6, ЕС ЭВМ или суперкомпьютеров "Эльбрус".

Потребности вычислительного эксперимента при изучении явлений в наиболее сложных областях науки, таких, как проблемы физики элементарных частиц, молекулярной биологии (например, геном человека), геофизики (в частности, физики атмосферы) и др., оказались связанными с необходимостью обеспечить предельно возможные вычислительные мощности. Выход был найден в коллективном использовании вычислительных мощностей, доступных исследователям через компьютерные сети. В развитии так называемых grid-технологий, разрабатываемых мировым сообществом в настоящее время, участвуют и ведущие научные институты России: Объединенный институт ядерных исследований (г. Дубна), Научно-исследовательский институт ядерной физики МГУ, Институт физики высоких энергий РАН (г. Протвино), Институт биофизики РАН (г. Пущино), Институт прикладной математики им. М. В. Келдыша РАН и др. Идея организации распределенных вычислений в гетерогенной сетевой среде, называемая метакомпьютингом, образно выражается метафорой "grid (сеть)". Подобно тому, как мы подключаем к электросети бытовые приборы, не задумываясь об устройстве этой электросети, сетевые grid-технологии призваны предоставить исследователям требуемые вычислительные мощности как разделяемые ресурсы. В Европе такой сетью должна стать Data Grid, к которой будет подключен и российский сегмент.

Философские аспекты моделирования

В настоящее время нельзя назвать область человеческой деятельности, в которой в той или иной степени не использовались бы методы моделирования. Особенно это относится к сфере управления различными системами, где основными являются процессы принятия решений на основе получаемой информации. Остановимся на философских аспектах моделирования, а точнее общей теории моделирования.

Методологическая основа моделирования. Все то, на что направлена человеческая деятельность, называется объектом (лат. objection — предмет). Выработка методологии направлена на упорядочение получения и обработки информации об объектах, которые существуют вне нашего сознания и взаимодействуют между собой и внешней средой.

В научных исследованиях большую роль играют гипотезы, т. е. определенные предсказания, основывающиеся на небольшом количестве опытных данных, наблюдений, догадок. Быстрая и полная проверка выдвигаемых гипотез может быть проведена в ходе специально поставленного эксперимента. При формулировании и проверке правильности гипотез большое значение в качестве метода суждения имеет аналогия.

Аналогией называют суждение о каком-либо частном сходстве двух объектов, причем такое сходство может быть существенным и несущественным. Необходимо отметить, что понятия существенности и несущественности сходства или различия объектов условны и относительны. Существенность сходства (различия) зависит от уровня абстрагирования и в общем случае определяется конечной целью проводимого исследования. Современная научная гипотеза создается, как правило, по аналогии с проверенными на практике научными положениями. Таким образом, аналогия связывает гипотезу с экспериментом.

Гипотезы и аналогии, отражающие реальный, объективно существующий мир, должны обладать наглядностью или сводиться к удобным для исследования логическим схемам; такие логические схемы, упрощающие рассуждения и логические построения или позволяющие проводить эксперименты, уточняющие природу явлений, называются моделями. Другими словами, модель (лат. modulus — мера) — это объект-заместитель объекта-оригинала, обеспечивающий изучение некоторых свойств оригинала.

Определение моделирования. Замещение одного объекта другим с целью получения информации о важнейших свойствах объекта-оригинала с помощью объекта-модели называется моделированием. Таким образом, моделирование может быть определено как представление объекта моделью для получения информации об этом объекте путем проведения экспериментов с его моделью. Теория замещения одних объектов (оригиналов) другими объектами (моделями) и исследования свойств объектов на их моделях называется теорией моделирования [S, 36, 46].

Определяя гносеологическую роль теории моделирования, т. е. ее значение в процессе познания, необходимо прежде всего отвлечься от имеющегося в науке и технике многообразия моделей и выделить то общее, что присуще моделям различных по своей природе объектов реального мира. Это общее заключается в наличии некоторой структуры (статической или динамической, материальной или мысленной), которая подобна структуре данного объекта. В процессе изучения модель выступает в роли относительного самостоятельного квазиобъекта, позволяющего получить при исследовании некоторые знания о самом объекте.

Если результаты моделирования подтверждаются и могут служить основой для прогнозирования процессов, протекающих в исследуемых объектах, то говорят, что модель адекватна объекту. При этом адекватность модели зависит от цели моделирования и принятых критериев.

Обобщенно моделирование можно определить как метод опосредованного познания, при котором изучаемый объект-оригинал находится в некотором соответствии с другим объектом-моделью, причем модель способна в том или ином отношении замещать оригинал на некоторых стадиях познавательного процесса. Стадии познания, на которых происходит такая замена, а также формы соответствия модели и оригинала могут быть различными:

1) моделирование как познавательный процесс, содержащий переработку информации, поступающей из внешней среды, о происходящих в ней явлениях, в результате чего в сознании появляются образы, соответствующие объектам; 2) моделирование, заключающееся в построении некоторой системы-модели (второй системы), связанной определенными соотношениями подобия с системой-оригиналом (первой системой), причем в этом случае отображение одной системы в другую является средством выявления зависимостей между двумя системами, отраженными в соотношениях подобия, а не результатом непосредственного изучения поступающей информации.

Следует отметить, что с точки зрения философии моделирование — эффективное средство познания природы. Процесс моделирования предполагает наличие объекта исследования; исследователя, перед которым поставлена конкретная задача; модели, создаваемой для получения информации об объекте и необходимой для решения поставленной задачи. Причем по отношению к модели исследователь является, по сути дела, экспериментатором, только в данном случае эксперимент проводится не с реальным объектом, а с его моделью. Такой эксперимент для инженера есть инструмент непосредственного решения организационно-технических задач.

Надо иметь в виду, что любой эксперимент может иметь существенное значение в конкретной области науки только при специальной его обработке и обобщении. Единичный эксперимент никогда не может быть решающим для подтверждения гипотезы, проверки теории. Поэтому инженеры (исследователи и практики) должны быть знакомы с элементами современной методологии теории познания и, в частности, не должны забывать основного положения материалистической философии, что именно экспериментальное исследование, опыт, практика являются критерием истины.

Идентификация в процессах управления

Важнейшей задачей современной теории и практики управления является построение модели ОУ, т. е. формализация закономерностей функционирования объекта. На основе этой модели определяются структура, алгоритмы и параметры СУ, выбираются аппаратно-программные средства реализации системы. Одним из эффективных методов построения модели сложного объекта является идентификация.

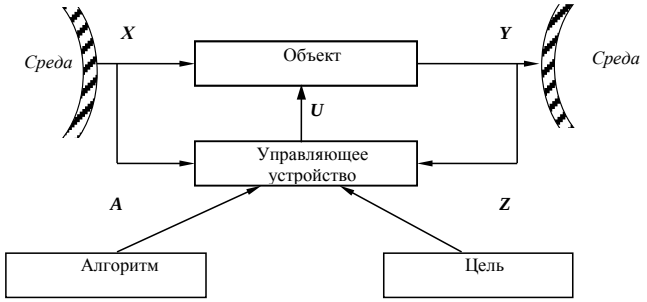
С проблемой моделирования мы сталкиваемся в двух случаях: во-первых, в процессах познания, когда стараются познавательные модели объектов и явлений, с которыми приходится сталкиваться человеку, и во-вторых, в процессах управления, связанных с целенаправленным изменением объекта, т.е. с достижением целей, постановленных человеком.

Нас интересует тип моделирования, который связан непосредственно с потребностями управления объектом. Действительно, чтобы управлять, нужно прежде всего знать, чем управляешь, т.е. иметь модель объекта, на которой можно «разыгрывать» последствия предполагаемого управления и выбрать наилучшее. Поэтому в процессе моделирования такого рода должна быть создана модель, которая, прежде всего обязана удовлетворять потребностям управления.

Следует отметить, что такая модель, синтезированная специально для потребностей управления, может и не отражать внутренних механизмов явления, что совершенно необходимо для познавательной модели. Ей достаточно лишь констатировать наличие определенной формальной связи между входом и выходом объекта.

В связи с этим целесообразно выяснить, что следует подразумевать под понятием «управление» и какие требования оно накладывает на модель управляемого объекта, получаемую в процессе моделирования.

Под управлением будем понимать процесс такого целенаправленного воздействия на объект, в результате которого объект оказывается в определенном смысле «ближе» к выполнению поставленных целей , чем до управления. На рисунке 1.1). показана общая схема управления объектом.



X- неуправляемая, но контролируемая составляющая;

U - управляемая составляющая; Y-информация о состоянии объекта, доступная управляющему устройству.

Рисунок 1.1 - Общая схема управления объектом

Для синтеза управления необходимо прежде всего определить цель Z, т.е. то, к чему должно «стремиться » управляющее устройство при воздействии на объект, каким должен быть объект с точки зрения управления. Однако, этого мало, необходимого еще иметь алгоритм управления A, который указывает, как достигнуть этой цели.

Таким образом управление реализуется четверкой

<U,I=<X,Y>,A,Z>,

где U-управляющее воздействие; I=<X,Y> - информация о состоянии среды и объекта; A- алгоритм; Z-цель управления.

Цель Z определяет требования, выполнение которых обеспечивается и организацией управляющего воздействия U с помощью алгоритма A и сбором информации по каналу Y. Не зная, как X и U влияют на состояние Y, т.е. не имея модели Y=F(X,U), нельзя определить управление U, достаточно лишь констатировать наличие определенной формальной связи.

Теория идентификации и моделирования – это научно-техническая дисциплина, которая занимается вопросами построения моделей объектов управления и систем управления и решает проблему оценки параметров этих моделей.

При рассмотрении проблемы идентификации различают **статический** подход, сущность которого в следующем: ставятся экспериментальные исследования, получают экспериментальную выборку, характеризующую динамику модели, на основании априорных данных о физических процессах в модели определяется структура самой модели, а по экспериментальной выборке определяются настроечные параметры модели.

Существует также **динамический** подход к проблеме идентификации: имеется некоторая замкнутая система, в этой системе специальным образом вводится дополнительный контур идентификации, который отслеживает изменение параметров модели в процессе реального функционирования системы, на основании некоторого критерия делаем оценку модели и при необходимости изменяем настроечные параметры объекта или системы в целом.

Предметом изучения курса являются методы построения моделей для объектов систем управления и методы определения динамических параметров этих моделей.

При изучении любых объектов (технических систем, процессов, явлений) основной задачей является построение их моделей. Как результат познания модель представляет собой отображение в той или иной форме свойств, закономерностей, физических и других характеристик, присущих исследуемому объекту. Характер модели определяется поставленными целями и может быть различным в зависимости от ее назначения. Модели разделяют на два основных класса: символические (словесные описания, схемы, чертежи, математические уравнения и т. д.) и вещественные (макеты, разного рода физические аналоги и электронные моделирующие устройства, имитирующие процессы в объектах).

При исследовании объектов, предназначенных для управления, применяют математические модели, входящие в класс символических, и вещественные. К математическим моделям относится такое математическое описание, которое адекватно отражает как статические, так и динамические связи между входными и выходными переменными объекта. Математическая модель может быть получена и аналитически (закономерности протекающих в объекте процессов полностью известны), и по результатам экспериментального исследования входных и выходных переменных объекта без изучения его физической сущности. Последний под-

ход особенно широко используется на практике, так как позволяет обойтись минимумом априорных сведений об объекте при построении его модели.

Для управления объектом необходимо иметь модель в виде математического описания, устанавливающего связь между входными и выходными переменными в форме, на основе которой может быть выбран закон управления, обеспечивающий заданное функционирование объекта. Получаемое описание должно давать правило преобразования воздействия на объект u в реакцию объекта y. Переменные u и y могут представлять собой функции одинаковых или разных аргументов. Преобразование одной функции в другую производится оператором, который определяет совокупность математических или логических операций, устанавливающих соответствие между ними: **x(t) = A{u(t)}.**

В качестве примера можно назвать операторы дифференцирования, интегрирования и т. п. Для стационарных линейных одномерных объектов оператор может быть задан в виде дифференциального уравнения или системы дифференциальных уравнений первого порядка, интегральной свертки, частотной характеристики (передаточной функции) объекта.

На практике объекты стремятся описывать линейными стационарными моделями, хотя в действительности все объекты в той или иной мере обладают свойствами нелинейности, нестационарности, распределенности, стохастичности. Использование более простых операторов следует рассматривать как попытку аппроксимации характеристик сложного объекта упрощенным приближенным описанием, но удобным для дальнейших расчетов. Описания могут быть заданы различным образом: аналитически, таблично, в виде разложения по какой-либо системе функций и т. д.

После формулировки целей управления необходимо выделить объект управления из среды, т. е. определить границы объекта и установить его взаимодействие со средой. Последнее характеризуется моделью возмущений. Далее строится структура и проводится идентификация параметров модели объекта. В процедуре синтеза управления, являющейся оптимизационной задачей, модель объекта выступает как ограничения. С помощью же модели возмущений можно оценить некоторые качественные показатели управления.

Когда решается задача управления сложным объектом, часто не удается получить описание, имеющее приемлемую точность. В этом случае используется ансамбль моделей, в котором каждая из них описывает отдельные стороны процесса. С упрощением моделей ослабляются и цели управления (например, в неопределенной ситуации ставится задача нахождения разумной стратегии управления без жестких качественных показателей). Часто такие модели реализуются как совокупность программ, имитирующих работу объекта и ориентированных на использование ЭВМ.

Выводы.

Построение математической модели объекта может производиться несколькими методами: аналитическим, экспериментальным и экспериментально-аналитическим.

Аналитический метод предусматривает получение математического описания объекта на основе законов физики, механики, химии и т. д. Такой подход дает положительный результат, если рассматриваемый объект достаточно прост по структуре и хорошо изучен. Если же объект изучен недостаточно или же настолько сложен, что аналитическое описание его математической моделью практически невозможно, прибегают к экспериментальным методам, суть которых сводится к статистической обработке технологических данных. При экспериментально-аналитическом методе априорная модель, полученная аналитическим путем, уточняется в соответствующих экспериментах.

Наиболее универсальную и достоверную математическую модель объекта можно найти аналитическим путем. Для этого необходимо располагать всесторонними сведениями об объекте (о конструкции, о законах, описывающих протекающие в нем процессы, об условиях функционирования и взаимодействия со средой). Однако часто из-за отсутствия достаточных данных получить решение задачи таким путем не удается. Трудности применения аналитических методов возникают и при описании реальных объектов, процессы в которых имеют сложный характер. Поэтому в подобных случаях эти методы дополняются экспериментальными исследованиями. Преимуществом моделей, полученных теоретическим путем, как правило, является их достаточно общий вид, позволяющий рассматривать поведение объектов в различных возможных режимах.

С практической точки зрения, более привлекательны экспериментальные методы, позволяющие находить модели объектов по результатам измерения их входных и выходных переменных. Хотя эти методы также предполагают наличие априорных сведений об изучаемом объекте, но их характер может быть не столь обстоятельным. Как правило, уровень априорных сведений должен быть достаточным лишь для выбора структуры модели и условий проведения эксперимента. Построение моделей объектов на основе такого подхода обычно и называют идентификацией.

В общем случае под идентификацией понимают определение структуры и параметров математической модели, которые обеспечивают наилучшую близость значений выходных переменных модели и объекта в смысле заданного критерия при одних и тех же входных воздействиях.

К выбору метода идентификации нельзя подойти однозначно, поскольку в самой постановке задачи заранее предполагается неопределенность (неполнота знаний об объекте, ограничения в наблюдениях объекта во времени, неточность измерения сигналов на входе и на выходе объекта и т. п.).

Комплекс задач при идентификации модели объекта обычно разделяется на три этапа:

- на первом этапе выбирается структура модели по результатам изучения объекта или по имеющимся априорным сведениям,
- на втором этапе - критерий близости (подобия) модели и объекта,
- на третьем этапе по экспериментальным данным определяются параметры модели исходя из выбранного критерия.

Следует заметить, что вследствие значительной сложности этап структурной идентификации часто сводят в значительной мере к эвристическому заданию структуры модели, опираясь на априорные данные об объекте. Очевидно, что в таких случаях эффективность последующей параметрической идентификации во многом определяется тем, насколько удачно была выбрана структура модели.

Особенности идентификации как оптимизационной задачи. Широкий спектр задач идентификации обусловливает, соответственно, большое разнообразие приемов и алгоритмов, применяемых для их решения. Однако практически все эти алгоритмы и их использование в процедурах идентификации имеют некоторую общую специфику, на которую важно обратить внимание.

Во-первых, при исследовании объектов с целью поиска их математического описания, как правило, применяют стандартные алгоритмы. Их необходимо хорошо понимать, чтобы правильно поставить задачу, подготовить исходные данные, выбрать параметры. Необходимость понимания существам алгоритма связана с тем, что необходимо хотя бы на качественном уровне прогнозировать результат и после вычислений либо принять его, либо повторить обработку, проведя необходимую коррекцию.

При использовании алгоритмов идентификации объектов необходимо иметь в виду, что в большинстве случаев они построены на основе принципа учета начальных представлений об исследуемом объекте. Однако начальные представления могут быть частично и даже полностью ошибочными, что может при очередном шаге реализации алгоритма не только улучшить, а даже ухудшить положение.

Начальные представления называют априорной (доопытной) информацией, в противоположность апостериорной (послеопытной) информации. Обычно апостериорная информация – это выборка (массивы, составленные из значений сигналов, измеренных на входе и выходе объекта); характер априорной информации может быть более разнообразен – представление о структуре оператора изучаемого объекта, частотной характеристике сигналов и объекта, функции распределения случайных величин и т.п.

Основная литература

1. Советов Б.Я., Яковлев С.А. Моделирование систем. – М.: Высшая школа. 2001
2. Авдеев П. Ф. Философия информационной цивилизации. — М.: ВЛАДОС, 1994

Дополнительная литература

3. Гроп Д. Методы идентификации систем. - М.: Мир, 1979.
4. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.

Тема 2 Математические модели объектов идентификации

Лекция 2 Общие сведения о математических моделях и их классификация

Данная лекция содержит обзорно-теоретический материал по следующим вопросам и проблемам: общие сведения о математических моделях и их классификация; множество моделей, структуры моделей; линейные модели и множества линейных моделей; семейство моделей передаточных функций; модели в пространстве состояний; модели с распределенными параметрами; временные характеристики; дискретные модели; дискретные модели в пространстве состояний; статические и динамические модели в форме управления регрессии.

Классификация математических моделей

Могут быть предложены различные методы классификации моделей объектов и систем, рассмотрим основные из них.

Физические модели. В основу классификации положена степень абстрагирования модели от оригинала. Предварительно все модели можно подразделить на две группы: материальные (физические) и абстрактные (математические). Физической моделью обычно называют систему, которая эквивалентна или подобна оригиналу, либо у которой процесс функционирования такой же, как у оригинала, и имеет ту же или другую физическую природу. Можно выделить следующие виды физических моделей: натурные, квазинатурные, масштабные и аналоговые.

Натурные модели — это реальные исследуемые системы. Их называют макетами и опытными образцами. Натурные модели имеют полную адекватность с системой-оригиналом, что обеспечивает высокую точность и достоверность результатов моделирования. Процесс проектирования ВС завершается зачастую испытанием опытных образцов. Квазинатурные модели представляют собой совокупность натурных и математических моделей. Этот вид моделей используется в случаях, когда математическая модель части системы не является удовлетворительной (например, модель человека-оператора) или когда часть системы должна быть исследована во взаимодействии с остальными частями, но их еще не существует, либо их включение в модель затруднено или дорого. Примерами квазинатурных моделей могут служить вычислительные полигоны, на которых отрабатывается программное обеспечение различных систем, или реальные АСУ, исследуемые совместно с математическими моделями соответствующих производств.

Масштабная модель — это система той же физической природы, что и оригинал, но отличающаяся от него масштабами. Методологической основой масштабного моделирования является теория подобия, которая предусматривает соблюдение геометрического подобия оригинала и модели и соответствующих масштабов для их параметров. При проектировании ВС масштабные модели могут использоваться для анализа вариантов компоновочных решений по конструкции системы и ее элементов.

Аналоговыми моделями называются системы, имеющие физическую природу, отличающуюся от оригинала, но сходные с оригиналом процессы функционирования. Обязательным условием при этом является однозначное соответствие между параметрами изучаемого объекта и его модели, а также тождественность безразмерных математических описаний процессов, протекаю-

щих в них. Для создания аналоговой модели требуется наличие математического описания изучаемой системы. В качестве аналоговых моделей' используются механические, гидравлические, пневматические системы, но наиболее широкое применение получили электрические и электронные аналоговые модели, в которых сила тока или напряжение являются аналогами физических величин другой природы. Особенностью аналоговых моделей является их гибкость и простота адаптации к изменению и измерению количественных значений параметров и характеристик моделируемой системы. Аналоговые модели используют при исследовании средств вычислительной техники на уровне логических элементов и электрических цепей, а также на системном уровне, когда функционирование системы описывается, например, дифференциальными или алгебраическими уравнениями.

Математические модели. Математическая модель представляет собой формализованное описание системы с помощью абстрактного языка, в частности с помощью математических соотношений, отражающих процесс функционирования системы. Для составления модели можно использовать любые математические средства — алгебраическое, дифференциальное и интегральное исчисление, теорию множеств, теорию алгоритмов и т. д. По существу вся математика создана для составления и исследования моделей объектов или процессов.

К средствам **абстрактного описания систем** относятся также языки химических формул, схем, чертежей, карт, диаграмм и т. п. Выбор вида модели определяется особенностями изучаемой системы и целями моделирования, так как исследование модели позволяет получить ответы на определенную группу вопросов. Для получения другой информации может потребоваться модель другого вида.

Цели моделирования и характерные черты оригинала определяют в конечном счете ряд других особенностей моделей и методы их исследования. Например, математические модели можно классифицировать на детерминированные и вероятностные (стохастические). Первые устанавливают однозначное соответствие между параметрами и характеристиками модели, а вторые — между статистическими значениями этих величин. Выбор того или иного вида модели обусловлен степенью необходимости учета случайных факторов. Среди математических моделей можно выделить по методу их исследования аналитические, численные и имитационные модели.

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

Численная модель характеризуется зависимостью такого вида, который допускает только частные численные решения для конкретных начальных условий и количественных параметров модели.

Имитационная модель — это совокупность описания системы и внешних воздействий, алгоритмов функционирования системы или правил изменения состояния системы под влиянием внешних и внутренних возмущений. Эти алгоритмы и правила не дают возможности использования имеющихся математических методов аналитического и численного решения, но позволяют имитировать процесс функционирования системы и производить измерения интересующих характеристик.

Многообразие систем и объектов, проявляющееся в многообразии их структурно-функциональной организации, определяет использование множества разных моделей, которые могут быть классифицированы в зависимости от:

1) характера функционирования исследуемой системы:
   - детерминированные, функционирование которых описывается детерминированными величинами;
   - стохастические или вероятностные, функционирование которых описывается случайными величинами.

2) характера протекающих в исследуемой системе процессов:
   - непрерывные, в которых процессы протекают непрерывно во времени;
   - дискретные, в которых процессы меняют свое состояние в дискретные моменты времени.

3) степени достоверности исходных данных об исследуемой системе:
   - с априорно известными параметрами;
   - с неизвестными параметрами.

4) режима функционирования системы:
   - стационарные, в которых характеристики не меняются со временем;
   - нестационарные, в которых характеристики изменяются со временем.

5) назначения:
   - статические или структурные, отображающие состав и структуру системы;
   - динамические или функциональные, отображающие функционирование системы во времени;
   - структурно-функциональные, отображающие структурные и функциональные особенности организации исследуемой системы.

6) способа представления (описания) и реализации:
   - концептуальные или содержательные, представляющие собой описание (в простейшем случае словесное) наиболее существенных особенностей структурно-функциональной организации исследуемой системы;
   - физические или материальные - модели, эквивалентные или подобные оригиналу (макеты) или процесс функционирования которых такой же как у оригинала и имеет ту же или другую физическую природу;
   - математические или абстрактные, представляющие собой формализованное описание системы с помощью абстрактного языка, в частности с помощью математических соотношений, отражающих процесс функционирования системы;
   - программные (алгоритмические, компьютерные), основанные на применении средств ВТ и представляющие собой обычно программный комплекс, и позволяющие наглядно и эффектно представить исследуемый объект посредством имитации или графического отображения математических зависимостей, описывающих искомый объект.

В дальнейшем нами основное внимание уделяется математическому моделированию, широко используемому при исследовании сложных технических систем.

Отметим, что одни и те же объекты могут быть описаны различными по сложности математическими моделями. Главным критерием выбора при этом является её **адекватность** исследуемому объекту.

Математические модели статических (ММС) объектов

Постановка задачи моделирования и идентификации статических характеристик объектов (см рисунок 2.1).

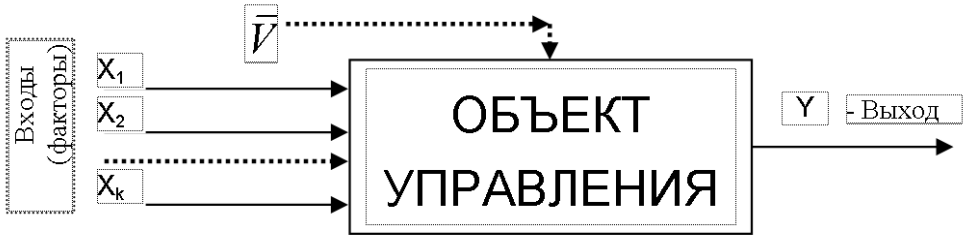


Рисунок 2.1 - Структура объекта статики

Статические характеристики широко применяются при расчете и исследовании систем в условиях установившихся состояний, когда все переходные процессы или завершились, или ими можно пренебречь. Характерной особенностью ММС является то, что в них отсутствуют производные по времени.

В общем виде ММС объекта – функция отклика, связывающая входные параметры X с входными $Y = \varphi(X_1, X_2 ... X_K)$ при наличии вектора возмущений V. При использовании статистических методов ММ статики обычно представляется в виде уравнения регрессии (полинома, отрезка ряда Тейлора, в который разлагается функция φ):

$$\hat{Y} = b_0 + \sum_{i=1}^{k} b_i \cdot X_i + \sum_{i=1}^{k} b_{ii} \cdot X_i^2 + \sum_{i=1}^{k} \sum_{j=i+1}^{k} b_{ij} \cdot X_i \cdot X_j \qquad (2.1)$$

где $\hat{Y}$ - расчетное значение выхода; X_i - входы;

$\bar{V}$ - вектор возмущений (шумы, помехи); N, n – количество опытов;

N0, n0 , – количество параллельных (дублирующих) опытов или опытов в центре плана; m – количество серий параллельных опытов;

k - количество входов (факторов, степень полинома);

N_{EX} – число экспериментальных точек плана;

L ; $L_{ЗН}$ – соответственно количество всех и значимых коэффициентов в уравнении регрессии;

$b_0, b_1, b_2, b_3, ...$ - выборочные коэффициенты регрессии, определяемые в результате идентификации;

b_0 - свободный член уравнения регрессии;

$b_1, b_2, b_3, ...$ - линейные эффекты; $b_{11}, b_{22}, b_{33}, ...$ - квадратичные эффекты;

$b_{12}, b_{23}, b_{13}, ...$ - эффекты парного взаимодействия;

$b_{111}, b_{222}, b_{333}, \ldots$ - эффекты тройного взаимодействия;

Например, чисто применяют вид уравнения для k=3 (L =10):

$$\hat{Y} = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + b_{11} \cdot X_1^2 + b_{22} \cdot X_2^2 + b_{33} \cdot X_3^2 + b_{12} \cdot X_1 \cdot X_2 + b_{13} \cdot X_1 \cdot X_3 + b_{23} \cdot X_2 \cdot X_3 + b_{123} \cdot X_1 \cdot X_2 \cdot X_3 \quad (2.2)$$

Могут применяться и другие виды ММС, в том числе и отражающие физическую природу изучаемого явления или объекта. Выбор модели вида (2.2) обусловлен его простотой, и , достаточно большой точностью описания исследуемых зависимостей. Дополнительные замечания по поводу выбора вида математической модели (уравнения регрессии). Традиционно используется вид (1), при необходимости (если уравнение не адекватно описывает эксперимент) могут быть добавлены члены вида $bi \cdot X_i^3$, $b_j \cdot X_1 \cdot X_2 \cdot X_3$ и т.д. Особых вычислительных трудностей при определении значений коэффициентов b_i уравнения регрессии b_i при этом не возникает, однако, характер изменения поведения графика функции высоких порядков сразу за пределами диапазона аппроксимации и даже между точками, найденными экспериментально может бать непредсказуем.

Математические модели линейных динамических (ММД) объектов и систем и связь между ними. ММД.

Широко используются при анализе и синтезе различных систем автоматизации (см. рисунок 2.2). ММД описывают изменение параметров во времени, в их выражения обязательно присутствуют производные по времени или другие эквивалентные им параметры. Очень часто (в обоснованных случаях) используют линейные ММД.



Рисунок 2.2 – Объект динамики

Принято обозначать векторы входных воздействий буквой U (управления), векторы выходных воздействий буквой X (состояния) или Y (выходы), векторы возмущающих воздействий буквой V(на рисунке 2.2 – не показаны)

Все ММД можно разделить на классы.

1 Модели для описания непрерывных систем

Линейные **дифференциальные уравнения** (ДУ) это наиболее общая форма описания динамики, однако их практическое использование и инженерной практике часто бывает сложным.

Наиболее общий вид записи ДУ:

$$a_{n+1}\frac{d^n x}{dt^n}+\ldots+a_2\frac{dx}{dt}+a_1 x=b_{m+1}\frac{d^m u}{dt^m}+\ldots+b_2\frac{du}{dt}+b_1 u \qquad (2.3).$$

Передаточные функции (ПФ). В инженерной практике часто используются ММД в виде ПФ. Для объектов с самовыравниванием ПФ в общем виде:

$$W(p)=C\cdot\frac{b_1+b_2\cdot p+\ldots+b_{m+1}\cdot p^m}{a_1+a_2\cdot p+\ldots+a_{n+1}\cdot p^n}\cdot\exp(-p\cdot\tau)\,.\ (n\geq m) \qquad (2.4)$$

При использовании «метода площадей» (b1=a1=1) [ ]:

$$W(p)=C\cdot\frac{b_1+b_2\cdot p}{a_1+a_2\cdot p+a_3\cdot p^2}\cdot\exp(-p\cdot\tau)\text{, или} \qquad (2.5)$$

$$W(p)=C\cdot\frac{b_1}{a_1+a_2\cdot p+a_3\cdot p^2+a_4\cdot p^3}\cdot\exp(-p\cdot\tau) \qquad (2.6)$$

Другие виды, например:

$$W(p)=\frac{C\cdot\exp(-p\cdot\tau)}{\prod_{i=1}^{n}(1+T_i\cdot p)}\text{ и т.п.} \qquad (2.7)$$

Очень часто применяют упрощенное выражение ПФ в виде апериодического звена с запаздыванием:

$$W(p) = \frac{C}{1 + T \cdot p} \cdot \exp(-p \cdot \tau) \qquad (2.8)$$

Для объектов без самовыравнивания используется передаточная функция в общем случае используется в виде:

$$W(p) = \left[\frac{D}{p} - C \cdot \frac{b_1 + b_2 \cdot p + \ldots + b_{m+1} \cdot p^m}{a_1 + a_2 \cdot p + \ldots + a_{n+1} \cdot p^n} \right] \cdot \exp(-p \cdot t) \qquad (2.9)$$

(обращаем внимание на то, что здесь присутствует **интегральная** составляющая - величина 1/p)
или упрощенные выражения для объектов без самовыравнивания:

$$W(p) = \frac{C}{p \cdot T_Б \cdot (1 + T \cdot p)} \cdot \exp(-p \cdot \tau) \text{ или:} \qquad (2.10)$$

$$W(p) = \frac{C}{T \cdot p} \cdot \exp(-p \cdot \tau) \qquad (2.11)$$

Конкретный вид W(p) выбирается из условия обеспечения адекватности и удобства вычислений.

Адекватность математической модели может быть, например, оценена по формуле (2.12). Если значение δ, найденное по этой формуле, не превышает 3-7%, то модель считается адекватной:

$$\delta = \frac{100\%}{N} \cdot \sum_{i=1}^{N} \frac{X_i - \hat{X}_i}{X_i} \qquad (2.12)$$

Для уравнения вида (2.8) выход $\hat{X}_i$ определяется аналитически по формуле (см. пример на рисунке 8.7):

$$\hat{X}(t) = C \cdot \Delta U \cdot \left[1 - \exp(-\frac{t-\tau}{T})\right] \text{ (при } \quad t \geq \tau) \qquad (2.13)$$

$$\hat{X}(t) = 0 \text{ (при } \quad t < \tau) \qquad (2.13А)$$

а для W(p) вида (2.7) при n=2:

$$W(p) = \frac{C \cdot \exp(-p \cdot \tau)}{(1 + T_1 \cdot p) \cdot (1 + T_2 \cdot p)}$$

$$\hat{X}(t) = C \cdot \Delta U \cdot \left(1 - \frac{T_1}{T_1 - T_2} \cdot \exp(-\frac{t-\tau}{T_1}) + \frac{T_2}{T_1 - T_2} \cdot \exp(-\frac{t-\tau}{T_2})\right) \qquad (2.14)$$

В более общем случае в качестве критерия соответствия при решении данной задачи берут критерий вида:

$$\min(F_{ai}),\ F(a_i) = \sum_{j=1}^{n} \left(X_i - \hat{X}_i\right)^2 \qquad (2.15)$$

где X_i – экспериментальное значение;

$\hat{X}_i$ – расчётное значение.

Для нахождения коэффициентов a_i составляют уравнения:

$$\frac{\partial F}{\partial a_i} = 0 \qquad (2.16)$$

Таким образом, получается система уравнений, решая которую можно определить a_i.

Адекватность полученных математических моделей статики может проверяться проверялась по критерию Фишера [1] (см. также следующую лекцию). Дополнительно для этого может быть использован более удобный т.н. критерии пригодности приближения [5] R-квадрат (коэффициент детерминации), используемый для оценки точности нелинейных моделей. Критерий R-квадрат может принимать значения только от нуля до единицы и чем ближе он к единице, тем лучше параметрическая модель приближает исходные данные.

Для его определения вначале вычисляется критерий SSE (Sum of squares due to error) - сумма квадратов ошибок по формуле:

$$SSE = \sum_{k=1}^{n} w_k \cdot (y_k - \hat{y}_k)^2 ,$$

где w_k - веса (у нас они не заданы, и считаются равными единице);

y_k - экспериментальные (исходные) значения данных для каждого опыта;

$\hat{y}_k$ - расчетные (предсказанные) значения данных для каждого опыта, получены по формуле (1);

n - количество экспериментальных значений (например, n=20).

Критерий R-квадрат (обозначенный ниже как R) определяется как отношение суммы квадратов относительно регрессии SSR к полной сумме квадратов (SST), т.е.:

$$SSR = \sum_{k=1}^{n} w_k \cdot (\hat{y}_k - \bar{y}_k)^2; \qquad SST = \sum_{k=1}^{n} w_k \cdot (y_k - \bar{y}_k)^2; \qquad R = \frac{SSR}{SST} = 1 - \frac{SSE}{SST},$$

где $\bar{y}_k$ - среднее значение экспериментальных (исходных) значения данных.

Близость полученных значений критерия R-квадрат к единице говорит о высокой точности описания эксперимента, например выражением вида (2.2). Обычно приемлемыми для практики считают значения критерия R-квадрат выше 0,9.

Данный показатель является статистической мерой согласия, с помощью которой можно определить, насколько уравнение регрессии соответствует реальным данным.Коэффициент детерминации изменяется в диапазоне от 0 до 1. Если он равен 0, это означает, что связь между переменными регрессионной модели отсутствует, и вместо нее для оценки значения выходной переменной можно с таким же успехом использовать простое среднее ее наблюдаемых значений. Напротив, если коэффициент детерминации равен 1, это соответствует идеальной модели, когда все точки наблюдений лежат точно на линии регрессии, т.е. сумма квадратов их отклонений равна 0. На практике, если коэффициент детерминации близок к 1, это указывает на то, что модель работает очень хорошо (имеет высокую значимость), а если к 0, то это означает низкую значимость модели, когда входная переменная плохо "объясняет" поведение выходной, т.е. линейная зависимость между ними отсутствует. Очевидно, что такая модель будет иметь низкую эффективность. Кроме того о точности аппроксимации результатов эксперимента можно дополнительно судить по

значениям суммарных абсолютных и относительных ошибок и анализируя графики сравнения расчетных и экспериментально найденных значений выхода для каждого опыта.

Иногда показателям тесноты связи можно дать качественную оценку (шкала Чеддока):

Количественная мера тесноты связи	Качественная характеристика силы связи
0,1 - 0,3	Слабая
0,3 - 0,5	Умеренная
0,5 - 0,7	Заметная
0,7 - 0,9	Высокая
0,9 - 0,99	Весьма высокая

Функциональная связь возникает при значении равном 1, а отсутствие связи — 0. При значениях показателей тесноты связи меньше 0,7 величина коэффициента детерминации всегда будет ниже 50 %. Это означает, что на долю вариации факторных признаков приходится меньшая часть по сравнению с остальными неучтенными в модели факторами, влияющими на изменение результативного показателя. Построенные при таких условиях регрессионные модели имеют низкое практическое значение.

Частотные характеристики (ЧХ). При подаче на вход **линейной** системы сигнала:

$$U(t) = A_u \sin(\omega t) = A_u \cdot \exp(j\omega t) \qquad (2.17)$$

на выходе будет сигнал:

$$X(t) = A_x \sin(\omega t + \varphi) = A_x \cdot \exp(j\omega t + \varphi), \qquad (2.18)$$

а АФЧХ имеет вид:

$$W(j\omega) = C \cdot \frac{b_1 + b_2 \cdot j\omega + \ldots + b_{m+1} \cdot j\omega^m}{a_1 + a_2 \cdot j\omega + \ldots + a_{n+1} \cdot j\omega^n} \cdot \exp(-j\omega \cdot \tau) =$$

$$= A(\omega) \cdot \exp(j\phi(\omega)) = \mathrm{Re}W(\omega) + \mathrm{Im}W(\omega) \qquad (2.19),$$

Где:

$$A(\omega) = \sqrt{[\mathrm{Re}W(\omega)]^2 + [\mathrm{Im}W(\omega)]^2} \qquad (2.20)$$

$$\varphi(\omega) = \arctan \frac{[\mathrm{Im}W(\omega)]}{[\mathrm{Re}W(\omega)]} \qquad (2.21)$$

Используются также **весовые функции**:

$$x(t) = \int_0^t g(t-\tau)y(\tau)d\tau = \int_0^t g(\tau)y(t-\tau)d\tau \quad , \qquad (2.22)$$

И системы матричных **линейных уравнений пространства состояния**:

Модель в пространстве параметров состояния

$$\begin{aligned} \dot{x} &= Ax + bU \\ y &= C^T x \end{aligned} \quad , \qquad (2.23)$$

где

U - вектор входа; x - вектор переменных состояния; y - вектор выхода системы;

A - матрица динамики системы; B - матрица управления; C^T - матрица измерения (датчиков)

или

$$\begin{aligned} \dot{x} &= A \cdot x(t) + B \cdot u(t); \\ y(t) &= C \cdot x(t) + D \cdot u(t) \end{aligned} , \qquad (2.24)$$

2 Модели для описания дискретных систем

Линейные разностные уравнения

Математические модели в цифровых системах управления записываются в виде рекуррентных разностных уравнений .

Если математическая модель ТОУ представлена передаточной функцией вида (2.8) и использован **фиксатор нулевого порядк**а, то выход объекта в цифровом виде для момента времени T_j определяется как:

$$X_j = A \cdot X_{j-1} + B \cdot U_{j-D} \qquad (2.25)$$

где $A = -\exp(-T_O / T)$, $B = (1 - A) \cdot C$, $D = [\tau / T_O]$

Уравнение (2.25) является разностным эквивалентом непрерывного уравнения объекта (2.8) для дискретных моментов времени j = 1, 2, 3, ….

D - число, округленное до ближайшего большего целого, определяет запаздывание ТОУ, выраженное в целом числе периодов опроса T_O.

Для j < D:

$X_j = A \cdot X_{j-1}$ X_{j-1} - выход объекта в момент времени T_j - T_O, т.е. на прошлом шаге опроса.

U_{j-D} - управляющее воздействие (выход регулятора) в момент времени T_j -D.

Дискретные передаточные функции:

$$W(z) = \frac{y(z)}{U(z)} = \frac{b_m z^{-n} + b_{m-1} z^{m-1} + \ldots b_0}{a_n z^{-n} + a_{n-1} z^{-(n-1)} + \ldots a_0} \qquad (2.26)$$

Модель в пространстве параметров состояния

x(k+1)=A*x(k)+B*U(k) (2.27)
y(k)=Cт*x(k)

цифровые разностные рекуррентные уравнения:

$X_j = A \cdot X_{j-1} + B \cdot U_{j-D}$. (2.29)

3 Модели для описания нелинейных систем (см рисунок 2.3)

$u(t)=\delta(t)$ $y(t)=\omega(t)$

$$y(t) = \int_0^{\infty} \omega(\tau) \cdot u(t-\tau) d\tau$$

$$y(t) = \int_0^{\infty}\int_0^{\infty} \omega(\tau_1, \tau_2) \cdot u(t-\tau_1) \cdot u(t-\tau_2) d\tau_1 d\tau_2 \qquad (2.30)$$

$$y(t) = \int_0^{\infty}\int_0^{\infty}\int_0^{\infty} \omega(\tau_1\tau_2\tau_3) \cdot u(t-\tau_1) \cdot u(t-\tau_2) \cdot u(t-\tau_3) d\tau_1 d\tau_2 d\tau_3 \qquad (2.31)$$

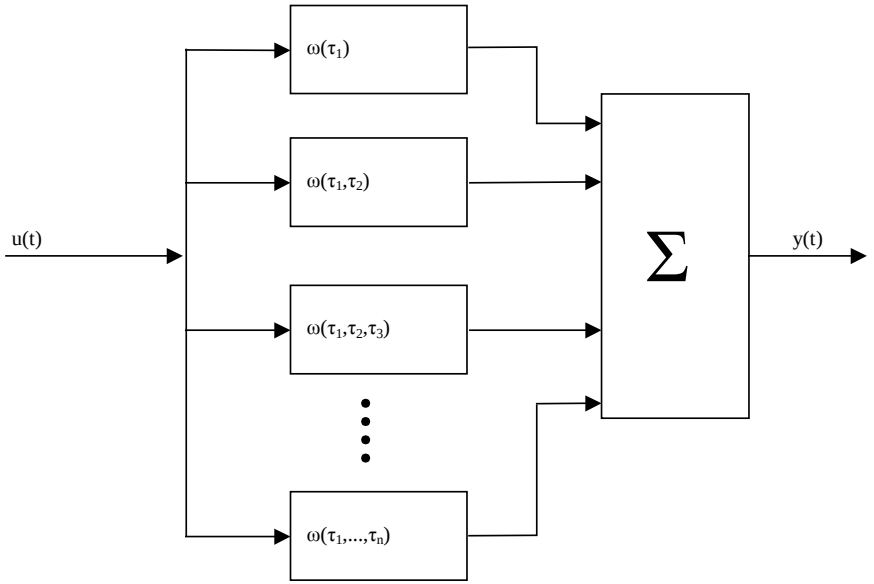


Рисунок 2.3 – нелинейные модели динамики

4 Стохастические модели. Модель нелинейной системы с использованием ядер Вольтера. При рассмотрении явлений в моделях с шумами принято оценивать влияние шумов на процесс идентификации путем использования понятий авто - и взаимно - корреляционной функций. Оценку влияния шумов можно производить, если процесс описания шумов описать следующим уравнением:

$$R_{uy}(\tau) = \int_0^{\infty} \omega(t) R_{uu}(t-\tau) dt \qquad (2.31)$$

R_{uu}- автокорреляционная функция входного сигнала (см. в лекции 6) ;

R_{uy}- взаимнокорреляционная функция входного и выходного сигнала.

Если u(t) – это случайный стационарный процесс и y(t) тоже, то, применяя эти понятия не учитывают, что R_{uu} и R_{uy} позволяют оценить величину случайной составляющей, то, решая это интегральное уравнение мы можем получать оценки с учетом помех входа и выхода. Задача имеет решение при условии, что входной сигнал можно измерять “абсолютно” точно, а выходной сигнал содержит все аддитивные составляющие помехи.

Таким образом по типу идентифицируемой модели можно выделить: линейная и нелинейная; детерминированная и стохастическая; с непрерывным и дискретным временем; стационарная и нестационарная; одномерная и многомерная; статическая и динамическая; с сосредоточенными и распределёнными параметрами. Студент должен уметь привести примеры таких объектов.

Свойства идентификации: управляемость, наблюдаемость, идентифицируемость.

Управляемость – система управляема, если для любого момента времени при любых состояниях существует такое управление u, которое переводит начальное состояние системы в конечное.

$$u_y = \left[B; AB; \ldots A^{n-1}B\right]$$

где n – порядок системы; A – матрица коэффициентов при x;

B – матрица коэффициентов при r.

Условием управляемости системы является то, чтобы $\det u_y$ не был равен нулю.

Наблюдаемость – система наблюдаема, если любое или все ее состояния можно непосредственно или косвенно определить по выходному вектору системы.

$u_н = \left[C^T; C^T A^T; \ldots C^T \left(A^{n-1}\right)^T\right]$, где C – матрица выхода, коэффициенты при y.

Хотя бы один минор не должен быть равен нулю, в этом случае система наблюдаема.

Идентифицируемость – система идентифицируема, если по изменениям координат состояния системы можно определить ее параметры.

$$ID = \left[V(0); A^n V(0); \ldots A^{nn} V(0)\right],$$

где V(0) – вектор начальных условий; Ап – матрица перехода.

$A^n = A^R + I$,

где AR – расширенная матрица; I – единичная матрица.

Система идентифицируема, если $\det \neq 0$.

Идентификация динамических систем. Допустим, динамическая система описана передаточной функцией следующего вида:

$$W(p) = \frac{b_0 p^m + b_1 p^{m-1} + ... + b_{m+1} p + b_m}{a_0 p^n + a_1 p^{n-1} + ... + a_{n-1} p + a_n}$$

Получаем систему дифференциальных уравнений первого порядка из n - уравнений

$$\frac{dx}{dt} = AX(t) + BU(t)$$

$$Y(t) = CX(t)$$

где Y – выходные переменные, U(t) – входные переменные, X – внутренние переменные.

Как получить эту систему?

I. а) $W(p) = \frac{y(p)}{u(p)}$ б) $p = \frac{d}{dt}$

в) Замена переменной $\frac{d^{n-1}y}{dt^{n-1}} = y_1, \quad \frac{d^{n-2}y}{dt^{n-2}} = y_2$

$\frac{d^{m-1}}{dt^{m-1}} = u_1, \quad \frac{d^{m-2}}{dt^{m-2}} = u_2$

г) получаем систему n – го порядка дифференциальных уравнений первого порядка.

С помощью методов пространства состояния.

2. От дифференциальных уравнений переходим к разностным уравнениям

$$\begin{cases} V(\kappa+1) = \Phi(T_0)V(\kappa) \\ Y(\kappa) = CV(\kappa) \end{cases}$$

где $V = [U \quad X]^T$ - обобщенный вектор.

3. получаем матрицу перехода:

$$\Phi(T_0) = V^T(\kappa+1)V^T(\kappa)\left[V(\kappa)V^T(\kappa)\right]^{-1},$$

при этом V(к) при е измерений примет значение:

$$V(\kappa) = \begin{bmatrix} v_1(1) & v_1(2) & \cdots & v_1(e) \\ \vdots & & & \vdots \\ v_n(1) & v_n(2) & \cdots & v_n(e) \end{bmatrix}$$

$$V(\kappa+1) = \begin{bmatrix} v_1(2) & v_2(2) & \cdots & v_n(2) \\ v_1(3) & v_2(3) & \cdots & v_n(3) \\ \vdots & & & \vdots \\ v_1(e+1) & v_2(e+1) & \cdots & v_n(e+1) \end{bmatrix}$$

Если $\Phi(T_0) = V^T(\kappa+1)V^T(\kappa)\left[V(\kappa)V^T(\kappa)\right]^{-1}$, то при

$$V(\kappa+1) = \begin{bmatrix} v_1(2) & v_2(3) & \cdots & v_n(e+1) \\ v_1(2) & v_2(3) & \cdots & v_n(e+1) \\ \vdots & & & \vdots \\ v_1(2) & v_2(3) & \cdots & v_n(e+1) \end{bmatrix}$$

Количество измерений определяется по аналогичной формуле для линейного регрессионного анализа.

Основная литература

1. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.
2. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.
3. Гроп Д. Методы идентификации систем. - М.: Мир, 1979

Дополнительная литература

4. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.

Тема 3 Статические и динамические модели в форме уравнения регрессии

Лекция 3 Статические и динамические модели в форме уравнения регрессии (ч. 1)

В этой лекции мы кратко рассмотрим основные теоретические разделы теории вероятностей и линейной алгебры. Эти сведения понадобятся нам в следующих лекциях и на выполнении лабораторных и практических занятиях. Отметим, что использование таких систем, как Mathcad и MATLAB значительно упрощают практическое применение данного материала.

Регрессионный анализ является в настоящее время классическим статистическим методом. Благодаря своим широким возможностям различные регрессионные процедуры давно и успешно используются в инженерной практике для идентификации процессов, однако их применение к идентификации многомерных процессов в реальном масштабе времени стало возможным только с развитием и внедрением быстродействующих компьютеров**.**

Методы идентификации, основанные на регрессионных процедурах с использованием метода наименьших квадратов, применимы как к линейным, так и к нелинейным процессам и облегчают проведение идентификации по нескольким входам одновременно. Более того, регрессионные методы позволяют осуществлять идентификацию в реальном масштабе времени, поскольку они основаны на измерениях входных и выходных сигналов, которые можно получить в процессе нормального функционирования системы.

В течение периода, пока выполняются измерения для регрессионной идентификации, параметры идентифицируемого процесса принимаются стационарными или квазистационарными. Этот период должен быть не менее mT, где T — интервал измерения,

Элементы теории вероятностей. Основные понятия и определения.

В теории вероятностей широко используются понятия события, вероятности, случайной величины.

Событие - всякий факт, который в результате опыта может произойти или не произойти.

Вероятность события есть численная мера степени объективной возможности этого события. Предположим, что рассматривается некоторый опыт или явление, в котором в зависимости от случая происходит или не происходит некоторое событие А.

Если условия опыта могут быть воспроизведены многократно, так что в принципе осуществима целая серия одинаковых и независимых друг от друга испытаний, то вероятность события А может быть вычислена по следующей формуле: P(A) = m / n, где n - общее число взаимно исключающих друг друга исходов; m – число исходов, которые приводят к наступлению события А.

Вероятность может принимать значения от 0 до 1. Событие, вероятность которого равна 0, называется невозможным, а событие, вероятность которого равна 1, называется достоверным. Несколько событий образуют полную группу событий, если в результате опыта должно непременно появиться хотя бы одно из них. Несколько событий называются несовместными в данном опыте, если никакие два из них не могут появиться вместе. Несколько событий называются равновозможными в данном опыте, если ни одно из этих событий не является объективно более возможным, чем другое. События называются независимыми, если появление одного из них не зависит от того, произошли ли другие события.

Случайной величиной называется величина, которая может принимать то или иное значение, неизвестное заранее.

Случайные величины могут быть двух типов:

- дискретные (прерывные), принимающие только отделенные друг от друга значения, которые можно пронумеровать;
- непрерывные (аналоговые), которые могут принимать любое значение из некоторого промежутка.

Иногда случайные величины, имеющие дискретную природу, рассматриваются как непрерывные. Такая замена оправдана в тех ситуациях, когда случайная величина принимает значения, которые незначительно отличаются друг от друга, так что замена дискретной случайной величины непрерывной практически не влияет на результаты расчетов. Например, время решения задачи в процессоре, определяемое как произведение среднего времени выполнения одной операции на число выполняемых операций, которое является дискретной случайной величиной, обычно рассматривается как непрерывная случайная величина, изменяющаяся в интервале от нуля до бесконечности.

Случайные величины часто обозначают большими буквами, а их возможные значения - соответствующими малыми буквами. Например, случайная величина X - число обращений к магнитным дискам в процессе решения задачи – может принимать значения x1=0, x2=1, x3=2, x4=3, ...

Законы распределения случайных величин. Рассмотрим дискретную случайную величину X, принимающую значения x1, x2,...,xn. Величина X может принять каждое из этих значений с некоторой вероятностью. Обозначим через pi вероятность того, что случайная величина X примет значение xi: pi = Pr(X = xi) (i = 1, n). Если в результате опыта величина X принимает только одно из этих значений, то имеем полную группу несовместных событий и сумма вероятностей всех возможных значений случайной величины равна единице:

Эта суммарная вероятность каким-то образом распределена между отдельными значениями. Случайная величина будет полностью описана с вероятностной точки зрения, если мы зададим это распределение, т.е. установим, так называемый, закон распределения.

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Про случайную величину говорят, что она подчинена данному закону распределения.

Дискретные законы распределения. Закон распределения дискретной случайной величины (дискретный закон распределения) может быть задан одним из следующих способов:

- аналитически в виде математического выражения, отражающего зависимость вероятности от значения случайной величины;
- таблично в виде ряда распределения случайной величины, в котором перечислены возможные значения случайной величины и соответствующие им вероятности;
- графически в виде многоугольника распределения, при котором по оси абсцисс откладываются возможные значения случайной величины, а по оси ординат - вероятности этих значений.

Числовые характеристики случайных величин. Числовые характеристики позволяют выразить в сжатой форме наиболее существенные особенности распределения случайной величины. Во многих случаях достаточно бывает указать только отдельные числовые параметры, до некоторой степени характеризующие существенные черты распределения случайной величины: например, какое-то среднее значение, около которого группируются возможные значения случайной величины; какое-либо число, характеризующее степень разбросанности этих значений относительно среднего; число, характеризующее асимметрию (или “скошенность”) распределения; число, характеризующее “крутость”, то есть островершинность или плосковершинность распределения и так далее.

В теории вероятностей используется большое количество числовых характеристик, имеющих различное назначение и различные области применения. Из них наиболее часто используются начальные и центральные моменты различных порядков, каждый из которых описывает то или иное свойство распределения. Начальные моменты рассматриваются относительно начала координат, а центральные моменты – относительно математического ожидания, то есть центра распределения.

Применение числовых характеристик существенно облегчает решение многих вероятностных задач. Особое значение это имеет при решении сложных задач, когда использование законов распределений приводит к громоздким выкладкам и не позволяет получить результаты в явном виде. Очень часто удается решить задачу до конца, оставляя в стороне законы распределения и оперируя одними числовыми характеристиками. Если в задаче фигурирует большое количество случайных величин, то в этих случаях по существу задачи для исчерпывающего суждения о результирующем законе распределения не требуется знать законов распределения отдельных случайных величин, фигурирующих в задаче; достаточно знать лишь некоторые числовые характеристики этих величин.

Пусть дискретная случайная величина X может принимать значения x1, x2,…,xk. Тогда: Частота появления события X= xi (частота = min) – отношение числа опытов mi в которых X приняла значение xi к общему числу опытов n.

Вероятность события X=xi [обозначается P(X=xi)] стремится к min при большом n: pi= P(X=xi) ≈ min

Аксиомы теории вероятностей Колмогорова.

1)Вероятность появления случайного события A – неотрицательное число: P(A)≥ 0

2)Вероятность достоверного события U – равна единице: P(U)=1; а невозможного события V – равна нулю: P(V)=0, т.е. $0 \le P \le 1$

3)Вероятность того, что наступит хотя бы одно из нескольких несовместимых событий A1 + A2 + … + An равно сумме вероятностей этих событий (теорема сложения вероятностей): P(A1 + A2 + … + An) = P(A1) + P(A2) + … + P(An)

Сумма событий (A1 + A2 + … + An) это событие, соответствующее появлению хотя бы одного из событий Ai, а произведение, это событие, соответствующее появлению всех событий Ai. Сумма вероятностей всех возможных значений случайной величины равно единице.

События независимые, если вероятность любого из них не зависит от того, произойдет или нет любое из остальных. Вероятность произведения нескольких независимых событий равна произведению вероятностей этих событий:

P(A1 · A2 · … · An) = P(A1) · P(A2) · … · P(An).

Событие A зависимо от события B, если вероятность события A меняется в зависимости появления события B. Вероятность события A, вычисленная при условии, что произошло события B – условная вероятность события A, обозначается P(A/B).

Для зависимых событий вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную при условии, что первое произошло. P(AB) = P(A) · P(B/A). Аналогично, если событие B предшествует событию A и влияет на него, то P(AB) = P(B) · P(A/B).

Пример. Требуется определить надежность системы автоматического управления, состоящей из трех последовательно соединенных элементов, каждый из которых может выйти из строя. Вероятности безотказной работы каждого из них соответственно равны: P(A1) = 0,9; P(A2) = 0,85; P(A3) = 0,82. В соответствии с теоремой умножения P(A1 · A2 · A3) = P(A1) · P(A2) · P(A3) = 0,9 ·0,85 ·0,82 = 0,6273

Математическое ожидание и дисперсия, их оценка и свойства. Оценка математического ожидания переменной X (обозначается mX или $\overline{X}$):

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad (3.1)$$

n – количество опытов

Оценка дисперсии переменной X (обозначается σ_X^2 или S_X^2):

$$S_X^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{f} \qquad (3.2)$$

Число степеней свободы f = [общее число измерений] – [число оценок, уже рассчитанным по этим измерениям и примененным в текущей формуле]. В данном случае уже рассчитана и используется величина $\overline{X}$, т.е. $f = n - 1$.

$S_X = \sqrt{S_X^2}$, среднее квадратичное отклонение (ошибка, стандарт).

Оценка дисперсии воспроизводимости Y (обозначается σ_Y^2 или S_Y^2):

$$S_{ВОСПР}^2 = \frac{\sum_{i=1}^{n_0}(Y_i - \overline{Y})^2}{n_0 - 1} \qquad (3.\ 3)$$

$$\text{или } S_{ВОСПР}^2 = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n_0}(Y_{ji} - \overline{Y}_j)^2}{m \cdot (n_0 - 1)} \qquad (3.3А)$$

$S_{ВОСП} = \sqrt{S_{ВОСП}^2}$ - ошибка опыта (ошибка воспроизводимости, среднеквадратическая ошибка, среднеквадратическое отклонение -СКО).(См [1] стр.37)

Оценка остаточной дисперсии Y (дисперсии адекватности):

$$S_{ОСТ}^2 = \frac{\sum_{i=1}^{n_0}(Y_i - \widehat{Y}_i)^2}{n_0 - L} \qquad (3.4)$$

$\widehat{Y}$ - расчетное значение выхода; L – количество коэффициентов в уравнении регрессии;

Нормальное распределение.

$$f(x) = \frac{1}{S_X \cdot \sqrt{2\pi}} \cdot e^{-\frac{(X - \overline{X})^2}{2 \cdot S_X^2}} \quad (-\infty < x < \infty) \qquad (3.5)$$

Критерий Стьюдента. Критерий позволяет определить значимость коэффициентов уравнения регрессии. b_i. (используется для проверки гипотезы значимости коэффициентов). Для этого для каждого из L коэффициентов b_i. Рассчитывают по формуле:

$$t_{b_i} = \frac{|b_i|}{S_{b_i}}, \qquad (3.6)$$

если $t_{b_i} > t_{TAБ}(f)$, то b_i считается значимым, в противном случае он не значимым и приравнивается к нулю.

$$S_{b_i} = \sqrt{\frac{S^2_{BOCПР}}{k}} = \frac{S_{BOCПР}}{\sqrt{k}}, \; S_{BOCПР} = \sqrt{S^2_{BOCПР}} \qquad (3.7)$$

S_{b_i} - среднеквадратическая ошибка в определении коэффициента регрессии b_i;

$t_{TAБ}(f)$ - табличное значение критерия Стьюдента при f = n0 – 1 или f =m· (n0 – 1);

Иногда незначимость b_i может быть вызвана и другими причинами, например, неверным интервалом варьирования при подготовке к эксперименту. Ниже приведен пример на Mathcad использования критерия Стьюдента. [1, стр.164].

Сделаем замечание относительно проблемы оценки значимости коэффициентов уравнения регрессии (2.2). Если какой либо коэффициент незначим, то он может быть принят равным нулю, т.к. его влияние на результат расчета по формуле (2.2) мало. Чтобы оценить допустимый предел требуемой точности определения расчетного значения выхода $\hat{Y}$ надо учесть, что в инженерной и научной практике обычно достаточно 3-5 значащих цифр при записи и использовании числовых расчетов. Для обоснованного отброса незначимых коэффициентов необходимо иметь информацию о дисперсии, присущей экспериментально найденным значениям выхода *Y*, а также значениям расчетным $\hat{Y}$. Строго говоря, экспериментально найденные значения Y являются случайными величинами. Дисперсия случайной величины это мера разброса данной случайной величины, то есть мера её отклонения от математического ожидания (т.е. среднеарифметического от ряда параллельных экспериментальных измерений). Если дисперсия эксперимента превышает допустимую величину, то доверять таким экспериментам нельзя, т.к. они невоспроизводимы. Это может происходить при низкой точности измерительных приборов, неправильной методики эксперимента, неучета дополнительных факторов, влияющих на результаты эксперимента, взаимного влияния входов друг на друга и т.п. Вообще говоря, отбрасывание незначимых коэффициентом актуально, когда уравнение (1) используется при ручном расчете, например с использованием калькуляторов. При использовании же при расчете компьютерных технологий отбрасывание незначимых коэффициентов можно не проводить.

Оценка значимости коэффициентов уравнения регрессии по критерию Стьюдента.

Исходные коэффициенты $b^T = (8.5 \;\; 2.5 \;\; -0.5 \;\; 3.5 \;\; -0.5 \;\; 0.5 \;\; -1.5 \;\; -0.5)$

Проведено 3 паралельных опыта $Y0_1 = 8$ $Y0_2 = 9$ $Y0_3 = 8.8$

$n0 = 3 \;\; i = 1 .. n0$

$L = 8 \;\; j = 1 .. L$

Матожидание Y0= $Y0cp = \dfrac{\left(\sum_{i=1}^{n0} Y0_i\right)}{n0}$ $Y0cp = 8.6$

Определение дисперсии воспроизводимости Sb

$$S2b = \frac{\left[\sum_{m=1}^{M} \left(Y0_m - y0\right)^2\right]}{f2}$$

$S = \sqrt{S2b}$ $Sb = \dfrac{S}{\sqrt{N}}$

$S2b = 0.28$ $S = 0.529$ $Sb = 0.187$

$Stud = \begin{pmatrix} 12.71 \\ 4.30 \\ 3.18 \\ 2.78 \\ 2.57 \\ 2.45 \\ 2.37 \\ 2.31 \\ 2.26 \\ 2.23 \end{pmatrix}$ <---при-f=2

$T_j = \dfrac{|b_j|}{Sb}$ $T^T = (45.434 \;\; 13.363 \;\; 2.673 \;\; 18.708 \;\; 2.673 \;\; 2.673 \;\; 8.018 \;\; 2.673)$

Число степеней свободы дисперсии воспроизводимости= $f = n0 - 1 \;\; f = 2$

$b_j = if\left[T_j < Stud_{(f)}, 0, b_j\right]$

Значимые коэффициенты $b^T = (8.5 \;\; 2.5 \;\; 0 \;\; 3.5 \;\; 0 \;\; 0 \;\; -1.5 \;\; 0)$

Таким образом, в этом примере Lзн = 4

Критерий Кохрена. Применяется для проверки воспроизводимости опытов (для проверки гипотезы воспроизводимости опытов):

$$G_P = \frac{S^2_{MAX}}{\sum_{i=1}^{n_0} S_i^2}, \qquad (3.8)$$

т.е. расчетное значение GP определяется как отношение наибольшей из оценок дисперсий к сумме всех найденных оценок дисперсий.

Если расчетное GP > GTАБ, то дисперсии неоднородны, т.е. значения Y не подчиняются нормальному закону распределения, а опыты невоспроизводимы). S^2_{MAX} - наибольшая из оценок выборочных дисперсий;

n0 – общее число сравниваемых дисперсий (количество параллельных (дублирующих) опытов);

$\sum_{i=1}^{n_0} S_i^2$ - сумма всех оценок дисперсий. Количество опытов в серии должно быть одинаковым, в противном случае можно воспользоваться критерием Бартлетта. Необходимо знать: n0 и число степеней свободы f = n0 - 1. Если опыты невоспроизводимы, то можно попытаться выявить и устранить источники невоспроизводимости (помехи), увеличить точность измерения и т.п. Если воспроизводимость не может быть обеспечена, то и результаты эксперимента не могут быть использованы для дальнейшей математической обработки. Например, (см. таб.) рассмотрим эксперимент, состоящий из трех серий (m = 3) по два параллельных опыта (n0 = 2). В них выход Y зависит от двух факторов X1 и X2.

номера серий опытов	Параллельные опыты				Y_i2средн%	Дисперсия Y= S_i^2
	Условия опытов		Выходы			
	X_1 гр.С	X_2 %	Y_i1 %	Y_i2 %		
1	24	45	35	36	35,5	0,5
2	24	55	39,3	38,1	38,7	0,72
3	25	45	31,8	33,4	32,6	1,28
					Сумма =	2,5

Расчетное значение критерия Кохрена GP = 1.28/2.5 = 0.51, в таб. (при m = 3 и f = n0 – 1 = 1) находим GTАБ = 0.967, т.к. GP < GTАБ, то опыты воспроизводимы, а оценки дисперсий можно считать однородными.

Вычислим также оценку дисперсии воспроизводимости:

$$S^2_{ВОСПР} = \frac{\sum_{i=1}^{n_0} (Y_i - \bar{Y})^2}{n_0 - 1} = (0.50 + 0.72 + 1.28)/3 = 0.83,$$ с ней связано число степеней свободы f = m·(n0 - 1) = 3·(2-1) = 3.

Критерий Фишера. Используется для проверки адекватности уравнения регрессии.

Расчетное значение критерия Фишера определяют как:

$$F_{PAC} = \frac{S^2_{OCT}}{S^2_{ВОСПР}} \qquad (3.9)$$

Оценка остаточной дисперсии (дисперсии адекватности) рассчитывают по формуле: $S^2_{OCT} = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n - L_{ЗН}}$ $\hat{Y}$ - значения Y, рассчитанные по уравнению регрессии на основе значимых коэффициентов. Если выполняется условие $F_{PAC} \le F_{ТАБ}$, уравнение регрессии адекватно описывает экспериментальные данные. $F_{ТАБ}$ определяется при известных значениях степеней свободы f1 для числителя и f2 для знаменателя.

f1 = n – L3H и f2 = n0 – 1.

Обычно используют уровень значимости p=0,05.

Выборочный коэффициент корреляции[См.1, стр.121]**:**

$$r = \frac{\sum_{i=1}^{n}\left[\left(X_i - \overline{X}\right)\cdot\left(Y_i - \overline{Y}\right)\right]^2}{\sqrt{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \cdot \sum_{i}^{n}\left(Y_i - \overline{Y}\right)^2 \cdot\right]}} \text{ или } r^*_{XY} = \frac{\sum_{i=1}^{n}\left[\left(X_i - \overline{X}\right)\cdot\left(Y_i - \overline{Y}\right)\right]^2}{(n-1)\cdot S_X \cdot S_Y}$$

Рассмотрим пример на Mathcad, для данных приведенный в [1, стр.164].

Проверка адекватности уравнения регрессии по критерию Фишера.

Значимые коэффициенты $b^T = (8.5 \quad 2.5 \quad 0 \quad 3.5 \quad 0 \quad 0 \quad -1.5 \quad 0)$

$$Yr_i = b_1 \cdot X_{i,1} + b_2 \cdot X_{i,2} + b_3 \cdot X_{i,3} + b_4 \cdot X_{i,4} + b_5 \cdot X_{i,5} + b_6 \cdot X_{i,6} + b_7 \cdot X_{i,7} + b_8 \cdot X_{i,8}$$

Число степеней свободы дисперсии воспроизводимости= $f2 = n0 - 1 \quad f2 = 2$

$$S2o = \frac{\left[\sum_{i=1}^{N} \left(Y_i - Yr_i\right)^2\right]}{f1}$$

Число степеней свободы остаточной дисперсии f1:

$f1 = N - 4 \qquad f1 = 4$

Остаточная дисперсия= $S2o = 2$

$$F = \frac{S2o}{S2b}$$

$$Fisher = \begin{pmatrix} 164.4 & 199.5 & 215.7 & 224.6 & 230.2 \\ 18.50 & 19.20 & 19.20 & 19.30 & 19.30 \\ 10.10 & 9.600 & 9.300 & 9.100 & 9.000 \\ 7.700 & 6.900 & 6.600 & 6.400 & 6.300 \\ 6.600 & 5.800 & 5.400 & 5.200 & 5.100 \end{pmatrix}$$

$$Yr = \begin{pmatrix} 1 \\ 6 \\ 4 \\ 9 \\ 11 \\ 16 \\ 8 \\ 13 \end{pmatrix} \qquad Y = \begin{pmatrix} 2 \\ 6 \\ 4 \\ 8 \\ 10 \\ 18 \\ 8 \\ 12 \end{pmatrix}$$

$Fisher_{f2,f1} = 19.3 \qquad F = 7.143$

Так как F<Fisher , то уравнение адекватно описывает эксперимент

Основы матричной алгебры (операции с матрицами)

Матрица - математический объект, записываемый в виде прямоугольной таблицы чисел (или элементов кольца) и допускающий алгебраические операции (сложение, вычитание, умножение) между ним и другими подобными объектами. Обычно матрицы представляются двумерными (прямоугольными) таблицам. Использование матричного аппарата позволяет значительно упростить вычислительные операции с использованием линейных математических моделей.

У каждого элемента матрицы есть два нижних индекса (a_{ij}) — первый «i» обозначает номер строки, в которой находится элемент, а второй «j» — номер столбца.

Говорят «матрица размера $m \times n$,» подразумевая, что в матрице m строк и n столбцов. В такой матрице индексы элементов удовлетворяют неравенствам $0 < i \leqslant m,$ (если индексы отсчитываются от единицы) или $0 \leqslant i < m,$ (если индексы отсчитываются от нуля).

Операции над матрицами. Пусть a_{ij} - элементы матрицы A, а b_{ij} - элементы матрицы B.

Линейные операции (это умножение на число, сложение и вычитание).

Умножение матрицы A на число λ (обозначение: λA) заключается в построении матрицы B, элементы которой получены путём умножения каждого элемента матрицы A на это число, то есть каждый элемент матрицы B равен: $b_{ij} = \lambda a_{ij}$

Сложение матриц A + B есть операция нахождения матрицы C, все элементы которой равны попарной сумме всех соответствующих элементов матриц A и B, то есть каждый элемент матрицы C равен: $c_{ij} = a_{ij} + b_{ij}$

Вычитание матриц A − B определяется аналогично сложению, это операция нахождения матрицы C, элементы которой: $c_{ij} = a_{ij} - b_{ij}$

Примеры:

$$A+B = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 0 \end{pmatrix} + \begin{pmatrix} 3 & 1 & 0 \\ 8 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2+3 & 0+1 & -1+0 \\ 1+8 & 3+2 & 0+3 \end{pmatrix} = \begin{pmatrix} 5 & 1 & -1 \\ 9 & 5 & 3 \end{pmatrix}$$

$$A-B = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 0 \end{pmatrix} - \begin{pmatrix} 3 & 1 & 0 \\ 8 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2-3 & 0-1 & -1-0 \\ 1-8 & 3-2 & 0-3 \end{pmatrix} = \begin{pmatrix} -1 & -1 & -1 \\ -7 & 1 & -3 \end{pmatrix}$$

Нелинейные операции (это умножение матриц, возведение в степень, транспонирование, обращение). Умножение матриц (обозначение: AB, реже со знаком умножения $A \times B$) — есть операция вычисления матрицы C, элементы которой равны сумме произведений элементов в соответствующей строке первого множителя и столбце второго, а именно:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

В первом множителе должно быть столько же столбцов, сколько строк во втором. Если матрица A имеет размерность $m \times n$, B - $n \times k$, то размерность их произведения AB = C есть $m \times k$. Схема этой процедуры показана на рисунке 3.1

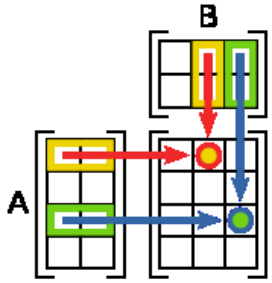


Рисунок 3.1 – Схема умножения матриц

Примеры умножения матриц:

$$FL = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} \times \begin{pmatrix} g & i & k \\ h & j & l \end{pmatrix} = \begin{pmatrix} (a \cdot g + d \cdot h) & (a \cdot i + d \cdot j) & (a \cdot k + d \cdot l) \\ (b \cdot g + e \cdot h) & (b \cdot i + e \cdot j) & (b \cdot k + e \cdot l) \\ (c \cdot g + f \cdot h) & (c \cdot i + f \cdot j) & (c \cdot k + f \cdot l) \end{pmatrix}$$

$$AB = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \times \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 2 \cdot (-1) + 3 \cdot (-2) & 2 \cdot 2 + 3 \cdot 3 \\ 5 \cdot (-1) + 7 \cdot (-2) & 5 \cdot 2 + 7 \cdot 3 \end{pmatrix} = \begin{pmatrix} -8 & 13 \\ -19 & 31 \end{pmatrix}$$

$$BA = \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} \times \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} -1 \cdot 2 + 2 \cdot 5 & -1 \cdot 3 + 2 \cdot 7 \\ -2 \cdot 2 + 3 \cdot 5 & -2 \cdot 3 + 3 \cdot 7 \end{pmatrix} = \begin{pmatrix} 8 & 11 \\ 11 & 15 \end{pmatrix}$$

$$A := \begin{pmatrix} 1 & 2 & 3 \\ -1 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix} \quad B := \begin{pmatrix} 5 & 4 & -3 \\ 0 & 9 & 6 \\ 5 & 8 & 3 \end{pmatrix} \quad A \cdot B = \begin{pmatrix} 20 & 46 & 18 \\ -5 & 41 & 33 \\ 20 & 46 & 18 \end{pmatrix} \quad B \cdot A = \begin{pmatrix} -2 & 24 & 6 \\ -3 & 57 & 18 \\ 0 & 56 & 24 \end{pmatrix}$$

Возводить в степень можно только квадратные матрицы, например:

$$A := \begin{pmatrix} 1 & 2 & 3 \\ -1 & 5 & 0 \\ 1 & 2 & 3 \end{pmatrix} \quad A^3 = \begin{pmatrix} -4 & 118 & 42 \\ -32 & 97 & -27 \\ -4 & 118 & 42 \end{pmatrix} \quad A \cdot A \cdot A = \begin{pmatrix} -4 & 118 & 42 \\ -32 & 97 & -27 \\ -4 & 118 & 42 \end{pmatrix}$$

Транспонирование матрицы (обозначение: A^T) — операция, при которой матрица отражается относительно главной диагонали, то есть: $a_{ij}^T = a_{ji}$

Если A — матрица размера $m \times n$, то A^T — матрица размера$n \times m$.

Свойства операций над матрицами

- ассоциативность сложения: A + (B + C) = (A + B) + C.
- коммутативность сложения: A + B = B + A.
- ассоциативность умножения: A(BC) = (AB)C.

Вообще говоря, умножение матриц не коммутативно: $AB \neq BA$

Дистрибутивность умножения относительно сложения:

A(B + C) = AB + AC; (B + C)A = BA + CA.

Свойства операции транспонирования матриц:

$(A^T)^T = A$; $(AB)^T = B^T A^T$: $(A^{-1})^T = (AT)^{-1}$, (если обратная матрица A^{-1} существует); $(A+ B)^T = A^T + B^T$; $detA = detA^T$

Если количество строк матрицы равно количеству столбцов, то такая матрица называется квадратной. Для квадратных матриц существует единичная матрица E (аналог единицы для операции умножения чисел) такая, что умножение любой матрицы на неё не влияет на результат, а именно: EA = AE = A. У единичной матрицы единицы стоят только по главной диагонали, остальные элементы равны нулю.

$$E := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Для некоторых квадратных матриц можно найти так называемую обратную матрицу. Обратная матрица A^{-1} такова, что если умножить матрицу на неё, то получится единичная матрица: $AA^{-1} = E$. Обратная матрица существует не всегда. Матрицы, для которых обратная существует, называются невырожденными (или регулярными), а для которых нет — вырожденными (или **сингулярными**). Матрица невырождена, если все ее строки (столбцы) линейно независимы как векторы. Максимальное число линейно независимых строк (столбцов) называется рангом матрицы. Определителем (детерминантом) матрицы называется значение нормированной кососимметрической (антисимметрической) полилинейной формы валентности $(p;\ 0)$ на столбцах матрицы. Квадратная матрица над числовым полем вырождена тогда и только тогда, когда ее определитель равен нулю. У сингулярной матрицы определитель равен нулю или близок к нулю. Обычно это происходит, когда элементы матрицы значительно (на много порядков) отличаются друг от друга.

Сингулярность матрицы иногда делает невозможность выполнять решение задчи регрессионного анализа и решения уравнений.

Матрица как запись коэффициентов системы линейных уравнений. Систему из m уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

можно представить в матричном виде:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}; \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

и тогда всю систему можно записать так: AX = B, где A имеет смысл таблицы коэффициентов a_{ij} системы уравнений. Если m = n и матрица A **невырожденная**, то решение этого уравнения состоит в нахождении обратной матрицы A^{-1}, поскольку умножив обе части уравнения на эту матрицу слева: $A^{-1}AX = A^{-1}B$ $A^{-1}A$ — превращается в E (единичную матрицу). И это даёт возможность получить столбец корней уравнений: $X = A^{-1}B$.

Векторы. Нами будут использоваться понятие вектор в смысле - одномерная матрица, т.е. вектор столбец. Для удобства записи вектора в строчку иногда применяют запись вида $\mathbf{B^T}$**=** вместо записи **B=**. Для вектора справедливы все приведенные выше правила операций с матрицами, например:

$$X := \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad Y := \begin{pmatrix} 5 \\ 4 \\ 7 \end{pmatrix} \quad X \cdot Y = 34 \quad Y \cdot X = 34 \quad X \cdot Y^T = \begin{pmatrix} 5 & 4 & 7 \\ 10 & 8 & 14 \\ 15 & 12 & 21 \end{pmatrix} \quad X^T \cdot Y = 34$$

Основная литература

1. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.
2. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.

Дополнительная литература

3. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
4. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.

Лекция 4 Статические и динамические модели в форме уравнения регрессии (ч. 2)

В этой лекции мы рассмотрим методы регрессионного анализа, применяемые при получении статических и динамические модели в форме управления регрессии.

1. Корреляционный анализ результатов моделирования.

С помощью корреляционного анализа исследователь может установить, насколько тесна связь между двумя (или более) случайными величинами, наблюдаемыми и фиксируемыми при моделировании конкретной системы S. Корреляционный анализ результатов моделирования сводится к оценке разброса значений η относительно среднего значения $\bar{y}$, т. е. к оценке силы корреляционной связи. Существование этих связей и их тесноту можно для схемы корреляционного анализа $y = M[\eta/\xi = x]$ выразить при наличии линейной связи между исследуемыми величинами и нормальности их совместного распределения с помощью коэффициента корреляции.

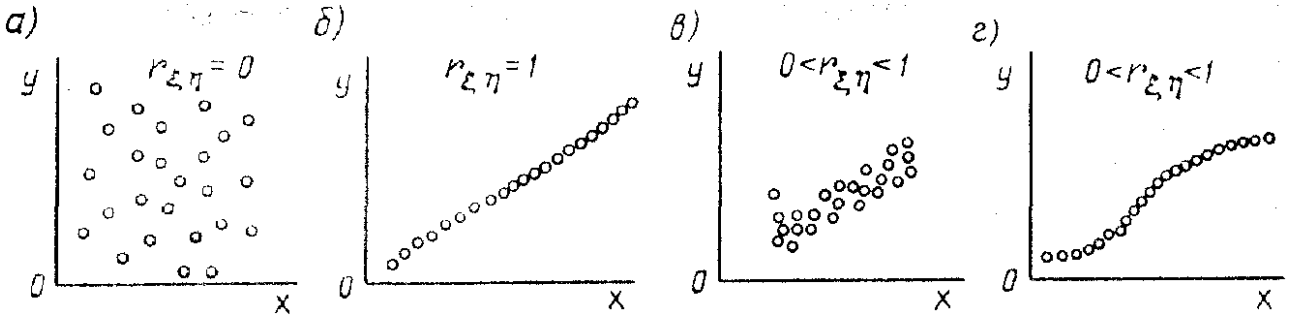


Рисунок 1. Различные случаи корреляции переменных

Для того чтобы оценить точность полученной при обработке результатов моделирования системы *S* оценки $r_{\xi\eta}$, целесообразно ввести в рассмотрение коэффициент

$w = \ln [(1+ r_{\xi\eta})/(1-r_{\xi\eta})]/2,$

причем *w* приближенно подчиняется гауссовскому распределению со средним значением и дисперсией:

$$\mu_w = \ln\left[\left(1+r_{\xi\eta}\right)/\left(1-r_{\xi\eta}\right)\right]/2$$

$$\sigma_w^2 = 1/(N-3)$$

Из-за влияния числа реализаций при моделировании *N* на оценку коэффициента корреляции необходимо убедиться в том, что $0 \le r_{\xi\eta} \le 1$ действительно отражает наличие статистически значимой корреляционной зависимости между исследуемыми переменными модели $M_м$. Это можно сделать проверкой гипотезы H_0: $r_{\xi\eta}=0$. Если гипотеза H_0 при анализе отвергается, то корреляци-

онную зависимость признают статистически значимой. Очевидно, что выборочное распределение введенного в рассмотрение коэффициента w при $r_{\xi\eta}= 0$ является гауссовским с нулевым средним $\mu_w = 0$ и дисперсией $\sigma_w^2 = (N-3)^{-1}$.

При анализе результатов моделирования системы S важно отметить то обстоятельство, что даже если удалось установить тесную зависимость между двумя переменными, то отсюда еще непосредственно не следует их причинно-следственная взаимообусловленность. Возможна ситуация, когда случайные ξ и η стохастически зависимы, хотя причинно они являются для системы S независимыми. При статистическом моделировании наличие такой зависимости может иметь место, например, из-за коррелированности последовательностей псевдослучайных чисел, используемых для имитации событий, положенных в основу вычисления значений x и y.

Таким образом, корреляционный анализ устанавливает связь между исследуемыми случайными переменными машинной модели и оценивает тесноту этой связи. Однако в дополнение к этому желательно располагать моделью зависимости, полученной после обработки результатов моделирования.

Регрессионный анализ результатов моделирования.

Регрессионный анализ дает возможность построить модель, наилучшим образом соответствующую набору данных, полученных в ходе машинного эксперимента с системой *S*. Под наилучшим соответствием понимается минимизированная функция ошибки, являющаяся разностью между прогнозируемой моделью и данными эксперимента. Такой функцией ошибки при регрессионном анализе служит сумма квадратов ошибок.

2. Дисперсионный анализ результатов моделирования.

При обработке и анализе результатов моделирования часто возникает задача сравнения средних выборок. Если в результате такой проверки окажется, что математическое ожидание совокупностей случайных переменных $\{y^{(1)}\}$, $\{y^{(2)}\}$, ..., $\{y^{(n)}\}$ отличается незначительно, то статистический материал, полученный в результате моделирования, можно считать однородным (в случае равенства двух первых моментов). Это дает возможность объединить все совокупности в одну и позволяет существенно увеличить информацию о свойствах исследуемой модели $M_м$, а следовательно, и системы *S*. Попарное использование для этих целей критериев Смирнова и Стьюдента для проверки нулевой гипотезы затруднено в связи с наличием большого числа выборок при моделировании системы. Поэтому для этой цели используется *дисперсионный анализ.*

дисперсионный анализ позволяет вместо проверки нулевой гипотезы о равенстве средних значений выборок проводить при обработке результатов моделирования проверку нулевой гипотезы о тождественности выборочной и генеральной дисперсий.

Возможны и другие подходы к анализу и интерпретации результатов моделирования, но при этом необходимо помнить, что их эффективность существенно зависит от вида и свойств конкретной моделируемой системы.

Регрессионный и корреляционный анализ широко используется при идентификация статических и динамических характеристик объектов.

Задача идентификации: задавшись видом уравнения регрессии (например, вида (2.2)), определить его неизвестные коэффициенты из условия, что заданная уравнением кривая будет с достаточной точностью описывать экспериментальную характеристику.

В качестве критерия соответствия при решении данной задачи берут критерий вида:

$$\min(\text{Fbi}),\ F(b_i) = \sum_{j=1}^{n} \left(y_{эj} - y_{pj}\right)^2 \qquad (4.1)$$

где $y_{эj}$ – экспериментальное значение; y_{pj} – расчётное значение.

Для нахождения коэффициентов bi составляют уравнения:

$$\frac{\partial F}{\partial b_i} = 0 \qquad (4.2)$$

Таким образом, получается система уравнений, решая которую можно определить b_i.

В конкретном случае, для выбора вида полинома используют графическое представление экспериментальной выборки, а также – априорные косвенные данные. Однако универсальных методик здесь нет.

Проиллюстрируем применение метода на примере для случая, когда уравнение регрессии выбрано в виде квадратного полинома: $\hat{y}_i = b_0 + b_1 \cdot x_i$.

Линейная регрессия от одного параметра.

С помощью этого метода ищется минимум функции (4.1), имеющей вид суммы квадратов разностей между экспериментальными yi и расчетными:

$$\hat{y}_i = F(x_i) = b_0 + b_1 \cdot x_i : \qquad (4.3)$$

$$I(b_i) = \sum_{i=1}^{n} \left[F(x_i) - y_i\right]^2 \rightarrow \min \qquad (4.4)$$

Минимизация осуществляется варьированием коэффициентов b , т.е. ищем такие b0 и b1, при которых I(b) будет минимальной. Необходимым условием минимума функции $I(b_i)$ является выполнение условий:

$$\frac{\partial \mathrm{I}}{\partial \mathrm{b}_0} = 2 \cdot \sum_{i=1}^{n} [F(x_i) - y_i] \cdot \frac{\partial F(x_i)}{\partial b_0} = 0 \qquad (4.5)$$

$$\frac{\partial \mathrm{I}}{\partial \mathrm{b}_1} = 2 \cdot \sum_{i=1}^{n} [F(x_i) - y_i] \cdot \frac{\partial F(x_i)}{\partial b_1} = 0$$

Решение этой система из двух уравнений с двумя неизвестными позволяет найти выражения для b0 и b1, при которых I(b) будет минимальной.

С учетом того, что $\frac{\partial F(x_i)}{\partial \mathrm{b}_0} = x_i$ и $\frac{\partial F(x_i)}{\partial \mathrm{b}_1} = 1$ система уравнений принимает вид:

$$2 \cdot \sum_{i=1}^{n} \left[(b_0 + b_1 \cdot x_i - y_i) \cdot x_i \right] = 0 \qquad \text{и} \qquad 2 \cdot \sum_{i=1}^{n} \left[b_0 + b_1 \cdot x_i - y_i \right] = 0$$

или (система нормальных уравнений, решая которую находим b0 и b1):

$$\begin{aligned} b_0 \cdot \sum x_i + b_1 \cdot \sum x_i^2 &= \sum x_i \cdot y_i \\ b_0 \cdot n + b_1 \cdot \sum x_i &= \sum y_i \end{aligned} \qquad (4.6)$$

откуда

$$b_0 = \frac{\sum y_i \cdot \sum x_i^2 - \sum x_i \cdot \sum x_i \cdot y_i}{n \cdot \sum x_i^2 - \left(\sum x_i \right)^2} \qquad (4.\ 7) \text{ и}$$

$$b_1 = \frac{n \cdot \sum x_i \cdot y_i - \sum x_i \cdot \sum y_i}{n \cdot \sum x_i^2 - \left(\sum x_i \right)^2} \qquad (4.8)$$

или проще сначала найти b1 , а затем $b_0 = \overline{Y} - b_1 \cdot \overline{X}$ из этого уравнения видно, что между b1 и b0 существует корреляционная зависимость. Для оценки силы **линейной** связи можно вычислить выборочный коэффициент корреляции:

$$r = b_1 \cdot \sqrt{\frac{n \cdot \sum x_i^2 - \left(\sum x_i \right)^2}{n \cdot \sum y_i^2 - \left(\sum y_i \right)^2}} \qquad (4.9)$$

Пример использования см. [1, стр. 130]

Метод множественной корреляции используется при идентификации объектом с несколькими входами (k - количество входов (факторов).. При k=1 – уравнение (2.1) график линии, k=2 – график плоскости, k=3 – график гиперповерхности.

Исходный статистический материал представлен в следующей таблице. В ней N строчек и k+2 столбцев.

Таблица 4.1

Исходные данные для получения математической модели с к - входами

№	X_0	X_1	X_2	X_3	…	Y
1	$x_{01}=1$	x_{11}	x_{12}	x_{13}	…	y_1
2	$x_{21}=1$	x_{21}	x_{22}	x_{23}	…	y_2
3	$x_{31}=1$	x_{31}	x_{32}	x_{33}	…	y_3
4	$x_{41}=1$	x_{41}	x_{42}	x_{43}	…	y_4
…	…	…	…	…	…	…
N	$x_{n1}=1$	x_{n1}	x_{n2}	x_{n3}	…	y_n

X_0 – фиктивная переменная, равная 1, она вводится для удобства записи. Столбик X_1 - значения входа номер 1, столбик X_2 - значения входа номер 2 и т.д. столбик Y - значения выходов в каждом из N опытов.

Отметим, что количество опытов N должно быть достаточно большим. Уменьшить количество опытов можно за счет использования методов планирования экспериментов (см. следующую лекцию).

Система нормальных уравнений, аналогично линейному случаю имеет вид (4.10).

Как видим, эта система уравнений имеет упорядоченное строение и может быть легко составлено для любого количества факторов k. В этой системе из k+1 уравнений с k+1 неизвестными k+1 слагаемых в левой части каждого уравнения. В том случае, если используется модель в виде полинома выше первого порядка, нелинейные члены уравнения регрессии рассматриваются как самостоятельные переменные

$$
\begin{aligned}
&b_0\sum X_0^2 + b_1\sum X_0X_1 + b_2\sum X_0X_2 + \cdots + b_k\sum X_0X_k = \sum X_0Y\\
&b_0\sum X_1X_0 + b_1\sum X_1^2 + b_2\sum X_1X_2 + \cdots + b_k\sum X_1X_k = \sum X_1Y\\
&b_0\sum X_2X_0 + b_1\sum X_2X_1 + b_2\sum X_2^2 + \cdots + b_k\sum X_2X_k = \sum X_2Y \qquad (4.10)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&b_0\sum X_kX_0 + b_1\sum X_kX_1 + b_2\sum X_kX_2 + \cdots + b_k\sum X_k^2 = \sum X_kY
\end{aligned}
$$

Рассмотрим пример, когда k=2 и ищутся коэффициенты модели уравнения вида:

$$\hat{Y} = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_{11} \cdot X_1^2 + b_{22} \cdot X_2^2 + b_{12} \cdot X_1 X_2. \qquad (4.11)$$

В качестве базовой берется модель:

1 2 3 4 5 6 <- номера коэффициентов

$$\hat{Y} = b_0 \cdot X_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_{11} \cdot X_3 + b_{22} \cdot X_4 + b_{12} \cdot X_5 \qquad (4.12)$$

Исходный статистический материал представлен в следующей таблице:

№ оп.	X_0	X_1	X_2	$X_3=X_1^2$	$X_4=X_2^2$	$X_5=X_1X_2$	Y
1	1	x_{11}	x_{12}	$x_{11}x_{12}$	x^2_{11}	x^2_{12}	y_1
2	1	x_{21}	x_{22}	$x_{21}x_{22}$	x^2_{21}	x^2_{22}	y_2
3	1	x_{31}	x_{32}	$x_{31}x_{32}$	x^2_{31}	x^2_{32}	y_3
4	1	x_{41}	x_{42}	$x_{41}x_{42}$	x^2_{41}	x^2_{42}	y_4
…	…	…	…	…	…	…	…
n	1	x_{n1}	x_{n2}	$x_{n1}x_{n2}$	x^2_{n1}	x^2_{n2}	y_n

Здесь выделенная часть таблицы – реальный эксперимент, остальная часть таблицы заполняется путем вычислений. Например:

№	X_0	X_1	X_2	X_3	X_4	X_5	Y
…	1	**2**	**4**	4	16	8	…
…	1	**3**	**6**	9	3	18	…

Система нормальных уравнений, аналогично предыдущему случаю имеет вид:

$$b_0\sum X_0^2 + b_1\sum X_0X_1 + b_2\sum X_0X_2 + b_{11}\sum X_0X_1^2 +$$
$$+ b_{22}\sum X_0X_2^2 + b_{12}\sum X_0X_1X_2 = \sum X_0Y$$
$$b_0\sum X_1X_0 + b_1\sum X_1^2 + b_2\sum X_1X_2 + b_{11}\sum X_1X_1^2 +$$
$$+ b_{22}\sum X_1X_2^2 + b_{12}\sum X_1^2X_2 = \sum X_1Y \qquad (4.13)$$

……

$$b_0\sum X_0X_1X_2 + b_1\sum X_1^2X_2 + b_2\sum X_1X_2^2 + b_{11}\sum X_1^3X_2 +$$
$$b_{22}\sum X_1X_2^3 + b_{12}\sum X_1^2X_2^2 = \sum X_1X_2Y$$

Коэффициенты bi находятся решением этой системы, что **гораздо** проще делать в матричной форме.

Регрессионный анализ в матричной форме. (См.[1] стр.146)

Использование матричной формы записи значительно упрощает как запись, так и решение задачи определения коэффициентов уравнения регрессии. Введем следующие обозначения (см. таблицу 4.1).

1).Матрица независимых переменных X – она содержит исходный статистический материал (см. таблицу в предыдущей теме, кроме последнего столбца) в ней n – строк и k+1 столбец:

2).Матрица (вектор) наблюдений Y (вектор – столбец): в нем n – строк.

X и Y известны в результате проведенного пассивного эксперимента.

3).Матрица (вектор) искомых коэффициентов уравнения регрессии B (вектор – столбец): в нем k+1 – строк.

Целью является определение вектора B по известным X и Y по формуле (4.15).

Таблица 4.1

Элементы матричной записи при регрессионном анализе

$X = \begin{bmatrix} x_{01} & x_{11} & \dots & x_{k1} \\ x_{02} & x_{12} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots \\ x_{0n} & x_{1n} & \dots & x_{kn} \end{bmatrix}$	$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$	$B = \begin{bmatrix} b_0 \\ b_1 \\ \dots \\ b_k \end{bmatrix}$	$(X^T \cdot X)^{-1} = \begin{bmatrix} c_{00} & c_{01} & \dots & c_{0k} \\ c_{10} & c_{11} & \dots & c_{1k} \\ \dots & \dots & \dots & \dots \\ c_{k0} & c_{k1} & \dots & c_{kk} \end{bmatrix}$
Матрица входов – факторы, независимые параметры	**Вектор выходов (наблюдений)**	**Вектор коэффициентов**	**Ковариационная матрица, матрица ошибок**

Матрицу $(X^T \cdot X)$ называют информационной матрицей (матрицей моментов), а $(X^T \cdot X)^{-1}$ матрицей ошибок или ковариационной матрицей.

В матричной форме система нормальных уравнений запишется:

$$X^T XB = X^T Y \qquad (4.14)$$

Решение этого уравнения имеет вид (**хорошо запомните эту формулу**!):

$$B = (X^T X)^{-1} X^T Y \qquad (4.15)$$

Уравнение (4.15) легко реализуется, например на Mathcad и широко применяется также и в методах планирования эксперимента. Однако вычисление вектора коэффициентов B иногда не может быть вычислено из-за **вырожденности** матрицы $(X^T \cdot X)$.

Это может происходить, например, если элементы матрицы X очень сильно отличаются друг от друга. Например, когда один из элементов равен 0.00005, в другой 100000.0.

Для определения остаточной дисперсии определяют матрицу столбец

$$\hat{Y} = X \cdot B \qquad (4.16)$$

Числитель остаточной дисперсии получают по формуле:

$$\left(Y - \hat{Y}\right)^T \cdot \left(Y - \hat{Y}\right) = \sum_{i=1}^{N} \left(Y - \hat{Y}\right)^2 \qquad (4.17)$$

Рассмотрим пример применения регрессионного анализа для построения математической модели с тремя входами и одним выходом в виде (2.2):

$$\hat{Y} = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + b_3 \cdot X_3 + b_{11} \cdot X_1^2 + b_{22} \cdot X_2^2 + b_{33} \cdot X_3^2 + b_{12} \cdot X_1 \cdot X_2 + b_{13} \cdot X_1 \cdot X_3 + b_{23} \cdot X_2 \cdot X_3 + b_{123} \cdot X_1 \cdot X_2 \cdot X_3 \qquad (2.2)$$

Или, что то же, в виде:

$$\hat{Y} = b_1 + b_2 \cdot X_1 + b_3 \cdot X_2 + b_4 \cdot X_3 + b_5 \cdot X_1^2 + b_6 \cdot X_2^2 + b_7 \cdot X_3^2 + b_8 \cdot X_1 \cdot X_2 + b_9 \cdot X_1 \cdot X_3 + b_{10} \cdot X_2 \cdot X_3 + b_{11} \cdot X_1 \cdot X_2 \cdot X_3 \qquad (2.2A)$$

Исходные данные для построения математической модели приведены в таблице 4.2. На основе эксперимента в ней заполнены столбец 2 (значения выхода – Y) и столбцы 4-6 (значения входов X_1, X_2 и X_3). Эти значения выделены в таблице жирным шрифтом.

Столбец 3 заполнен значениями равными 1, а столбцы 7-13 значениями, вычисленными на основе значений столбцов 4-4. Проведено 20 опытов.

Таблица 4.2

Исходные данные для примера построения математической модели

№	Y-		Входы - X			Условные входы, рассчитанные на основе входов X						
оп	Выход	X_0	X_1	X_2	X_3	X_1^2	X_2^2	X_3^2	X_1*X_2	X_1*X_3	X_2*X_3	$X_1X_2X_3$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	**43,0**	1	**90**	**3**	**120**	8100	9	14400	270,000	10800,000	360,000	32400
2	**44,0**	1	**110**	**3**	**120**	12100	9	14400	330,000	13200,000	360,000	39600
3	**40,0**	1	**90**	**7**	**120**	8100	49	14400	630,000	10800,000	840,000	75600
4	**38,0**	1	**110**	**7**	**120**	12100	49	14400	770,000	13200,000	840,000	92400
5	**45,0**	1	**90**	**3**	**150**	8100	9	22500	270,000	13500,000	450,000	40500
6	**43,0**	1	**110**	**3**	**150**	12100	9	22500	330,000	16500,000	450,000	49500
7	**44,0**	1	**90**	**7**	**150**	8100	49	22500	630,000	13500,000	1050,000	94500
8	**42,0**	1	**110**	**7**	**150**	12100	49	22500	770,000	16500,000	1050,000	115500
9	**44,0**	1	**83,18**	**5,00**	**135,**	6918,9	25	18225	415,900	11229,300	675,000	56146,5
10	**41,0**	1	**116,8**	**5,00**	**135**	13642	25	18225	584,000	15768,000	675,000	78840
11	**43,0**	1	**100**	**1,64**	**135**	10000	2,6766	18225	163,600	13500,000	220,860	22086
12	**37,0**	1	**100**	**8,36**	**135**	10000	69,955	18225	836,400	13500,000	1129,140	112914
13	**45,0**	1	**100**	**5,00**	**109,7**	10000	25	12049,45	500,000	10977,000	548,850	54885
14	**48,0**	1	**100**	**5,00**	**160,3**	10000	25	25673,65	500,000	16023,000	801,150	80115
15	**47,0**	1	**100**	**5**	**135**	10000	25	18225	500,000	13500,000	675,000	67500
16	**45,0**	1	**100**	**5**	**135**	10000	25	18225	500,000	13500,000	675,000	67500
17	**46,5**	1	**100**	**5**	**135**	10000	25	18225	500,000	13500,000	675,000	67500
18	**45,5**	1	**100**	**5**	**135**	10000	25	18225	500,000	13500,000	675,000	67500
19	**46,7**	1	**100**	**5**	**135**	10000	25	18225	500,000	13500,000	675,000	67500
20	**46,0**	1	**100**	**5**	**135**	10000	25	18225	500,000	13500,000	675,000	67500

Таким образом, заданы значения вектора Y(столбец 2) и матрицы X (столбцы 3-13) имеющие значения:

$Y^T =$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	43	44	40	38	45	43	44	42	44	41	43	37	45	48	47	45	46.5	45.5	46.7	46

$X =$

	1	2	3	4	5	6	7	8	9	10	11
1	1	90	3	120	8100	9	14400	270	10800	360	32400
2	1	110	3	120	12100	9	14400	330	13200	360	39600
3	1	90	7	120	8100	49	14400	630	10800	840	75600
4	1	110	7	120	12100	49	14400	770	13200	840	92400
5	1	90	3	150	8100	9	22500	270	13500	450	40500
6	1	110	3	150	12100	9	22500	330	16500	450	49500
7	1	90	7	150	8100	49	22500	630	13500	1050	94500
8	1	110	7	150	12100	49	22500	770	16500	1050	115500
9	1	83.18	5	135	6918.912	25	18225	415.9	11229.3	675	56146.5
10	1	116.8	5	135	13642.24	25	18225	584	15768	675	78840
11	1	100	1.636	135	10000	2.676	18225	163.6	13500	220.86	22086
12	1	100	8.364	135	10000	69.956	18225	836.4	13500	1129.14	112914
13	1	100	5	109.77	10000	25	12049.453	500	10977	548.85	54885
14	1	100	5	160.23	10000	25	25673.653	500	16023	801.15	80115
15	1	100	5	135	10000	25	18225	500	13500	675	67500
16	1	100	5	135	10000	25	18225	500	13500	675	67500
17	1	100	5	135	10000	25	18225	500	13500	675	67500
18	1	100	5	135	10000	25	18225	500	13500	675	67500
19	1	100	5	135	10000	25	18225	500	13500	675	67500
20	1	100	5	135	10000	25	18225	500	13500	675	67500

Далее вычислив, например, в системе Mathcad по формуле (4.15) $B = (X^T X)^{-1} X^T Y$ получаем вектор одиннадцати значений коэффициентов B:

$B^T =$

	1	2	3	4	5	6	7	8	9	10	11
1	-207.44742	3.94528	19.67478	0.74521	-0.01372	-0.56344	0.00019	-0.18750	-0.00875	-0.09583	0.00125

Можно произвести проверку, результаты которой сведены в таблицу 4.3. Расчетное значение выхода $\hat{Y}$ можно получить, применив формулу (2.2А) или её матричный аналог в виде (4.16): $\hat{Y} = X \cdot B$. Полученную нами математическую модель можно считать адекватной, т.к. Значение критерия $R^2 = 0{,}9724$ достаточно близко к единице.

Таблица 4.1

Результат проверки адекватности математической модели

№ опыта	Входные переменные			Выход		Погрешность (ошибка)	
	X_1	X_2	X_3	Y	$\hat{Y}$	абсолютная	относительная %
1	120,00	3,00	120,00	43,000	43,5514	-0,5514	-1,2823
2	120,00	3,00	120,00	44,000	44,3277	-0,3277	-0,7448
3	120,00	7,00	120,00	40,000	40,2128	-0,2128	-0,5321
4	120,00	7,00	120,00	38,000	37,9892	0,0108	0,0285
5	150,00	3,00	150,00	45,000	45,3582	-0,3582	-0,7959
6	150,00	3,00	150,00	43,000	43,1345	-0,1345	-0,3127
7	150,00	7,00	150,00	44,000	44,0196	-0,0196	-0,0446
8	150,00	7,00	150,00	42,000	41,7959	0,2041	0,4858
9	135,00	5,00	135,00	44,000	43,4885	0,5115	1,1625
10	135,00	5,00	135,00	41,000	41,0205	-0,0205	-0,0499
11	135,00	1,64	135,00	43,000	42,3519	0,6481	1,5072
12	135,00	8,36	135,00	37,000	37,1570	-0,1570	-0,4243
13	109,77	5,00	109,77	45,000	44,5247	0,4753	1,0562
14	160,23	5,00	160,23	48,000	47,9842	0,0158	0,0329
15	135,00	5,00	135,00	47,000	46,1306	0,8694	1,8497
16	135,00	5,00	135,00	45,000	46,1306	-1,1306	-2,5125
17	135,00	5,00	135,00	46,500	46,1306	0,3694	0,7943
18	135,00	5,00	135,00	45,500	46,1306	-0,6306	-1,3860
19	135,00	5,00	135,00	46,700	46,1306	0,5694	1,2192
20	135,00	5,00	135,00	46,000	46,1306	-0,1306	-0,2840
Суммарная ошибка =						1,0851E-07	-0,2330
Среднее значение ошибки =						5,4256E-09	**-0,0117**
Значение критерия Rквадрат =						**0,9724**	

Регрессионный метод идентификации линейных динамических систем (Метод наименьших квадратов)
Рассмотрим уравнение:

$$\frac{d^3x(t)}{dt^3} + a_1\frac{d^2x(t)}{dt^2} + a_2\frac{d\ x(t)}{dt} + a_3x(t) = b\sin(ct)\ , \qquad (4.21)$$

В уравнении заменим производную конечной разностью:

$$\frac{d^3x}{dt^3} = \frac{x_{i+2} - 3x_{i+1} + 3x_i - x_{i-1}}{\Delta t^3}\ , \qquad (4.22)$$

$$\frac{d^2x}{dt^2} = \frac{x_{i+1} - 2x_i + x_{i-1}}{\Delta t^2}\ , \qquad (4.23)$$

$$\frac{dx}{dt} = \frac{x_{i+1} - x_i}{\Delta t}\ , \qquad (4.24)$$

Подставим выражения (4.22)-(4.24) в уравнение (4.21):

$$\frac{x_{i+2} - 3x_{i+1} + 3x_i - x_{i-1}}{\Delta t^3} + a_1\frac{x_{i+1} - 2x_i + x_{i-1}}{\Delta t^2} + a_2\frac{x_{i+1} - x_i}{\Delta t} + a_3x_i = u\ , \qquad (4.25)$$

Приведя подобные, получим:

$$x_{i+2} = \left(3 - a_1\Delta t - a_2\Delta t^2\right)x_{i+1} - \left(3 - 2a_1\Delta t - a_2\Delta t^2 - a_3\Delta t^3\right)x_i + \left(1 - a_1\Delta t\right)x_{i-1} + u\Delta t^3\ , \qquad (4.26)$$

Представим (4.26) в виде:

$$x_{i+2} = \alpha_0x_{i+1} + \alpha_1x_i + \alpha_2x_{i-1} + \alpha_3u\ , \qquad (4.27)$$

где $\alpha_0 = \left(3 - a_1\Delta t - a_2\Delta t^2\right);\ \alpha_1 = \left(3 - 2a_1\Delta t - a_2\Delta t^2 - a_3\Delta t^3\right)$

$$\alpha_2 = (1 - a_1 \Delta t) \qquad \alpha_3 = \Delta t^3.$$

Рассматриваемый метод идентификации, основан на регрессионных процедурах с использованием метода наименьших квадратов.

Рассматриваем систему, которая задана уравнением:

$$x_{i+2} = \alpha_0 x_{i+1} + \alpha_1 x_i + \alpha_2 x_{i-1} + u$$

Минимизируемая функция имеет вид:

$$F = \sum_{j=1}^{n} (\tilde{x}_{i+2} - (\alpha_0 x_{i+1} + \alpha_1 x_i + \alpha_2 x_{i-1} + u))^2 \qquad (4.28)$$

Запишем систему уравнений для нахождения α_i , для этого найдем частные производные:

$$\begin{cases} \dfrac{\partial F}{\partial \alpha_0} = -2\sum_{i=1}^{n} (\tilde{x}_{i+2} - \alpha_0 x_{i+1} - \alpha_1 x_i - \alpha_2 x_{i-1} - \alpha_3 u) x_{i+1} = 0 \\ \dfrac{\partial F}{\partial \alpha_1} = -2\sum_{i=1}^{n} (\tilde{x}_{i+2} - \alpha_0 x_{i+1} - \alpha_1 x_i - \alpha_2 x_{i-1} - \alpha_3 u) x_i = 0 \\ \dfrac{\partial F}{\partial \alpha_2} = -2\sum_{i=1}^{n} (\tilde{x}_{i+2} - \alpha_0 x_{i+1} - \alpha_1 x_i - \alpha_2 x_{i-1} - \alpha_3 u) x_{i-1} = 0 \\ \dfrac{\partial F}{\partial \alpha_3} = -2\sum_{i=1}^{n} (\tilde{x}_{i+2} - \alpha_0 x_{i+1} - \alpha_1 x_i - \alpha_2 x_{i-1} - \alpha_3 u) u_i = 0 \end{cases} \qquad (4.29)$$

$$\begin{cases} \sum_{i=1}^{n} x_{i+2}x_{i+1} - \alpha_0 \sum_{i=1}^{n} x_{i+1}^2 - \alpha_1 \sum_{i=1}^{n} x_i x_{i+1} - \alpha_2 \sum_{i=1}^{n} x_{i-1}x_{i+1} - \alpha_3 \sum_{i=1}^{n} u x_{i+1} = 0 \\ \sum_{i=1}^{n} x_{i+2}x_i - \alpha_0 \sum_{i=1}^{n} x_{i+1}x_i - \alpha_1 \sum_{i=1}^{n} x_i^2 - \alpha_2 \sum_{i=1}^{n} x_{i-1}x_i - \alpha_3 \sum_{i=1}^{n} u x_i = 0 \\ \sum_{i=1}^{n} x_{i+2}x_{i-1} - \alpha_0 \sum_{i=1}^{n} x_{i+1}x_{i-1} - \alpha_1 \sum_{i=1}^{n} x_i x_{i-1} - \alpha_2 \sum_{i=1}^{n} x_{i-1}^2 x_{i+1} - \alpha_3 \sum_{i=1}^{n} u x_{i-1} = 0 \\ \sum_{i=1}^{n} x_{i+2}u_i - \alpha_0 \sum_{i=1}^{n} x_{i+1}u_i - \alpha_1 \sum_{i=1}^{n} x_i u_i - \alpha_2 \sum_{i=1}^{n} x_{i-1}u_i - \alpha_3 \sum_{i=1}^{n} u_i^2 = 0 \end{cases} \quad (4.30)$$

Запишем систему в матричной форме:

$$\begin{bmatrix} \sum_{i=1}^{n} x_{i+1}^2 & \sum_{i=1}^{n} x_i x_{i+1} & \sum_{i=1}^{n} x_{i-1}x_{i+1} & \sum_{i=1}^{n} u_i x_{i+1} \\ \sum_{i=1}^{n} x_{i+1}x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_{i-1}x_i & \sum_{i=1}^{n} u_i x_i \\ \sum_{i=1}^{n} x_{i+1}x_{i-1} & \sum_{i=1}^{n} x_i x_{i-1} & \sum_{i=1}^{n} x_{i-1}^2 & \sum_{i=1}^{n} u_i x_{i-1} \\ \sum_{i=1}^{n} x_{i+1}u_i & \sum_{i=1}^{n} x_i u_i & \sum_{i=1}^{n} x_{i-1}u_i & \sum_{i=1}^{n} u_i^2 \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_{i+2}x_{i+1} \\ \sum_{i=1}^{n} x_{i+2}x_i \\ \sum_{i=1}^{n} x_{i+2}x_{i-1} \\ \sum_{i=1}^{n} x_{i+2}u_i \end{bmatrix} \quad (4.31)$$

$$X \cdot \alpha = y \Rightarrow \alpha = X^{-1} \cdot y. \quad (4.32)$$

По полученным α найдем коэффициенты ai

$$a_1 = \frac{1-\alpha_2}{\Delta t}; \quad a_2 = \frac{2+\alpha_2-\alpha_0}{\Delta t^2}; \quad \alpha_3 = \frac{\alpha_0+\alpha_1+\alpha_2-1}{\Delta t^3}.$$

Основная литература

1. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.
2. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.
3. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.

Дополнительная литература

4. Толчеев В.О., Ягодкина Т.В. Методы идентификации линейных одномерных динамических систем. -М.: Изд-во МЭИ, 1997

Лекция 5 Статические и динамические модели в форме уравнения регрессии (ч.3)

В этой лекции нами будут рассмотрены элементы теории дисперсионного анализа, статистической оценки параметров распределения случайных величин и проверки статистических гипотез, используемых при практическом решении задач идентификации.

Перечислим кратко назначение методов анализа, применяемых при идентификации.

- **Дисперсионный** анализ применяется для исследования влияния одной или нескольких качественных переменных (факторов) на одну зависимую количественную переменную (отклик).
- **Анализ временных рядов** применим к одиночным или связанным временным рядам и позволяет выделять различные формы периодичности и взаимовлияния временных процессов, а также осуществлять прогнозирование будущего поведения временного ряда.
- Регрессионные процедуры позволяют рассчитать модель, описываемую некоторым уравнением и отражающую функциональную зависимость между экспериментальными количественными переменными, а также проверяют гипотезу об адекватности модели экспериментальным данным. По полученным результатам можно оценить природу и степень зависимости переменных и предсказать новые значения зависимой переменной.
- **Корреляционный анализ** – это группа статистических методов, направленная на выявление и математическое представление структурных зависимостей между выборками.
- **Кластерный анализ** осуществляет разбиение объектов на заданное число удаленных друг от друга классов, а также строит дерево классификаций объектов посредством иерархического объединения их в группы (кластеры).
- Основной задачей **факторного анализа** является нахождение в многомерном пространстве первичных переменных (значения которых регистрируются в эксперименте), сокращенной системы вторичных переменных (факторов). Метод факторного анализа первоначально был разработан в психологии с целью выделения отдельных компонентов человеческого интеллекта из многомерных данных по измерению различных проявлений умственных способностей.
- Методы **контроля качества** предназначены для контроля выпускаемой продукции с целью выявления нарушений и узких мест в организации производства и в технологических процессах, ведущих к снижению качества продукции.

Дисперсионный анализ (от латинского Dispersio – рассеивание / на английском Analysis Of Variance - ANOVA) применяется для исследования влияния одной или нескольких качественных переменных (факторов) на одну зависимую количественную переменную (отклик).

В основе дисперсионного анализа лежит предположение о том, что одни переменные могут рассматриваться как причины (факторы, независимые переменные), а другие как следствия (зависимые переменные). Независимые переменные называют иногда регулируемыми факторами именно потому, что в эксперименте исследователь имеет возможность варьировать ими и анализировать получающийся результат.

Основной целью дисперсионного анализа (ANOVA) является исследование значимости различия между средними с помощью сравнения (анализа) дисперсий. Разделение общей дисперсии на несколько источников, позволяет сравнить дисперсию, вызванную различием между группами, с дисперсией, вызванной внутригрупповой изменчивостью. При истинности нулевой гипотезы (о равенстве средних в нескольких группах наблюдений, выбранных из генеральной совокупности), оценка дисперсии, связанной с внутригрупповой изменчивостью, должна быть близкой к оценке межгрупповой дисперсии. Если вы просто сравниваете средние в двух выборках, дисперсионный анализ даст тот же результат, что и обычный t-критерий для независимых выборок (если сравниваются две независимые группы объектов или наблюдений) или t-критерий для зависимых выборок (если сравниваются две переменные на одном и том же множестве объектов или наблюдений).

Сущность дисперсионного анализа заключается в расчленении общей дисперсии изучаемого признака на отдельные компоненты, обусловленные влиянием конкретных факторов, и проверке гипотез о значимости влияния этих факторов на исследуемый признак. Сравнивая компоненты дисперсии друг с другом посредством F—критерия Фишера, можно определить, какая доля общей вариативности результативного признака обусловлена действием регулируемых факторов.

Исходным материалом для дисперсионного анализа служат данные исследования трех и более выборок: , которые могут быть как равными, так и неравными по численности, как связными, так и несвязными. По количеству выявляемых регулируемых факторов дисперсионный анализ может быть однофакторным (при этом изучается влияние одного фактора на результаты эксперимента), двухфакторным (при изучении влияния двух факторов) и многофакторным (позволяет оценить не только влияние каждого из факторов в отдельности, но и их взаимодействие).

Дисперсионный анализ относится к группе параметрических методов и поэтому его следует применять только тогда, когда доказано, что распределение является нормальным.

Дисперсионный анализ используют, если зависимая переменная измеряется в шкале отношений, интервалов или порядка, а влияющие переменные имеют нечисловую природу (шкала наименований).

Однофакторный дисперсионный анализ используется для проверки гипотезы о сходстве средних значений двух или более выборок, принадлежащих одной и той же генеральной совокупности. Этот метод распространяется также на тесты для двух средних (к которым относится, например, t-критерий).

Двухфакторный дисперсионный анализ с повторениями - Представляет собой более сложный вариант однофакторного анализа, включающее более чем одну выборку для каждой группы данных.

Двухфакторный дисперсионный анализ без повторения - Представляет собой двухфакторный анализ дисперсии, не включающий более одной выборки на группу. Используется для проверки гипотезы о том, что средние значения двух или нескольких выборок одинаковы (выборки принадлежат одной и той же генеральной совокупности). Этот метод распространяется также на тесты для двух средних, такие как t-критерий.

Теоретические основы. В любом эксперименте среднее значение наблюдаемых величин меняется в связи с изменением входных факторов, определяющих условия эксперимента, а также и случайных факторов (помех).Исследование влияния тех или иных факторов на изменчивость средних значений и является задачей дисперсионного анализа.

Дисперсионный анализ состоит в выделении и оценке отдельных факторов, вызывающих изменчивость изучаемой случайной величины. Для этого производится разложение суммарной выборочной дисперсии на составляющие, обусловленные независимыми факторами.

Для того чтобы определить значимо ли влияние данного фактора необходимо оценить значимость соответствующей выборочной дисперсии в соответствии с дисперсией воспроизводимости, обусловленной случайными факторами.

Предположим, что результат эксперимента зависит от некоторого одиночного фактора A, который принимает n различных значений (n-количество серий опытов). Для каждой серии опытов проводится m повторных наблюдений, результаты которых можно записать в следующем виде:

$Y_{11}\ Y_{12}\ Y_{13}\ ...\ Y_{1m}$
$Y_{21}\ Y_{22}\ Y_{23}\ ...\ Y_{2m}$
$Y_{31}\ Y_{32}\ Y_{33}\ ...\ Y_{3m}$
... ...
$Y_{n1}\ Y_{n2}\ Y_{n3}\ ...\ Y_{nm}$

На основе полученных статистических данных требуется проверить гипотезу о равенстве математических ожиданий для каждой конкретной серии. Если проверяемая гипотеза верна, то средние арифметические значения для всех серий практически не отличаются друг от друга, в противном случае предполагаемая гипотеза должна быть отвергнута.

Обозначим через $\overline{Y_i}$ среднее значение i-й серии опытов, а через $\overline{Y}$ общее среднее значение для всех наблюдений:

$$\overline{Y_j} = \frac{1}{m} \cdot \sum_{i=1}^{m} Y_{ji} \qquad (5.1)$$

$$\overline{Y} = \frac{1}{m \cdot n} \cdot \sum_{j=1}^{n} \sum_{i=1}^{m} Y_{ji}$$

$$Q = \sum_{i=1}^{n}\sum_{j=1}^{m}(Y_{ij} - \overline{Y})^2 = \left\{\sum_{i=1}^{n}\sum_{j=1}^{m}(Y_{ij} - \overline{Y}_i)^2\right\} + \left\{m \cdot \sum_{i=1}^{n}(\overline{Y}_i - \overline{Y})^2\right\} \qquad (5.2)$$

Сущность дисперсионного анализа состоит в разложении суммы квадратов отклонений отдельных Y_{ij} от общего среднего на две суммы:

Q - определяет общее отклонение значения каждого опыта (Y_{ij}) от среднего;

QA - характеризует рассеяние, вызванное фактором А (выражение во вторых фигурных скобках);

Qост - характеризует рассеяние вызванное случайными помехами (выражение в первых фигурных скобках).

Разделив суммы квадратов отклонений на соответствующие степени свободы получим следующие дисперсии:

$$\sigma^2 = Q/f$$
$$\sigma_A^2 = Q_A/f_1 \qquad (5.3)$$
$$\sigma_{ОСТ}^2 = Q_{ост}/f^2$$

Число степеней свободы f = m·n -1 f1 = n - 1 f2 = n·(m-1)

Проведение дисперсионного анализа состоит в сравнении оценок σ_A^2 и $\sigma_{ОСТ}^2$. Если гипотеза о том, что математические ожидания для каждой серии равны, верна, то σ_A^2 не должна существенно превышать $\sigma_{ОСТ}^2$, что проверяется по критерию Фишера:

$$F = \sigma_A^2/\sigma_{ОСТ}^2 \qquad (5.4)$$

Если $F < F_{кр}$, то различие между σ_A^2 и $\sigma_{ОСТ}^2$ можно считать несущественным, т.е. влияние фактора А сравнимо с влиянием случайных помех.

Если $F > F_{кр}$, то различие между σ_A^2 и $\sigma_{ОСТ}^2$ существенно, т.е. фактор А оказывает влияние на выходную величину.

Значение Fкр определяют по квантилям распределения Фишера, при уровне значимости α ("альфа") и степеням свободы f1 и f2:

$F_{кр} = f(\alpha, f1, f2)$

Статистическая оценка параметров распределения случайных величин Проверка гипотез

Статистическая гипотеза (statistical hypothesys) — это определённое предположение о распределении вероятностей, лежащем в основе наблюдаемой выборки данных.

Проверка статистической гипотезы (testing statistical hypotheses) — это процесс принятия решения о том, противоречит ли рассматриваемая статистическая гипотеза наблюдаемой выборке данных. Статистический тест или статистический критерий — строгое математическое правило, по которому принимается или отвергается статистическая гипотеза.

Статистическая гипотеза представляет собой некоторое предположение о законе распределения случайной величины или о параметрах этого закона, формулируемое на основе выборки. Примерами статистических гипотез являются предположения: генеральная совокупность распределена по экспоненциальному закону; математические ожидания двух экспоненциально распределенных выборок равны друг другу. В первой из них высказано предположение о виде закона распределения, а во второй – о параметрах двух распределений. Гипотезы, в основе которых нет никаких допущений о конкретном виде закона распределения, называют непараметрическими, в противном случае – параметрическими.

Гипотезу, утверждающую, что различие между сравниваемыми характеристиками отсутствует, а наблюдаемые отклонения объясняются лишь случайными колебаниями в выборках, на основании которых производится сравнение, называют *нулевой* (основной) гипотезой и обозначают H_0. Наряду с основной гипотезой рассматривают и *альтернативную* (конкурирующую, противоречащую) ей гипотезу H_1. И если нулевая гипотеза будет отвергнута, то будет иметь место альтернативная гипотеза.

Различают простые и сложные гипотезы. Гипотезу называют *простой*, если она однозначно характеризует параметр распределения случайной величины. Например, если λ является параметром экспоненциального распределения, то гипотеза H_0 о равенстве $\lambda = 10$ – простая гипотеза. Сложной называют гипотезу, которая состоит из конечного или бесконечного множества простых гипотез. Сложная гипотеза H_0 о неравенстве $\lambda > 10$ состоит из бесконечного множества простых гипотез H_0 о равенстве $\lambda = b_i$, где b_i – любое число, большее 10. Гипотеза H_0 о том, что математическое ожидание нормального распределения равно двум при неизвестной дисперсии, тоже является сложной. Сложной гипотезой будет предположение о распределении случайной величины X по нормальному закону, если не фиксируются конкретные значения математического ожидания и дисперсии.

Проверка гипотезы основывается на вычислении некоторой случайной величины – критерия, точное или приближенное распределение которого известно. Обозначим эту величину через z, ее значение является функцией от элементов выборки $z=z(x_1, x_2, \ldots, x_n)$. Процедура проверки гипотезы предписывает каждому значению критерия одно из двух решений – принять или отвергнуть гипотезу. Тем самым все выборочное пространство и соответственно множество значений критерия делятся на два непересекающихся подмножества S_0 и S_1. Если значение критерия z попадает в область S_0, то гипотеза принимается, а если в область S_1, – гипотеза отклоняется. Множество S_0 называется областью принятия гипотезы или областью допустимых значений, а множество S_1 – областью отклонения гипотезы или критической областью. Выбор одной области однозначно определяет и другую область.

Принятие или отклонение гипотезы H_0 по случайной выборке соответствует истине с некоторой вероятностью и, соответственно, возможны два рода ошибок. Ошибка первого рода возникает с вероятностью α тогда, когда отвергается верная гипотеза

H_0 и принимается конкурирующая гипотеза H_1. Ошибка второго рода возникает с вероятностью β в том случае, когда принимается неверная гипотеза H_0, в то время как справедлива конкурирующая гипотеза H_1. Доверительная вероятность – это вероятность не совершить ошибку первого рода и принять верную гипотезу H_0. Вероятность отвергнуть ложную гипотезу H_0 называется мощностью критерия. Следовательно, при проверке гипотезы возможны четыре варианта исходов, таблица. 5.1.

Таблица 5.1

Четыре варианта исходов

Гипотеза H_0	Решение	Вероятность	Примечание
Верна	Принимается	1–α	Доверительная вероятность
	Отвергается	α	Вероятность ошибки первого рода
Неверна	Принимается	β	Вероятность ошибки второго рода
	Отвергается	1–β	Мощность критерия

Например, рассмотрим случай, когда некоторая несмещенная оценка параметра θ вычислена по выборке объема *n*, и эта оценка имеет плотность распределения *f*(θ), рисунок. 5.1.



Рисунок 5.1 - Области и отклонения гипотезы

Предположим, что истинное значение оцениваемого параметра равно *T*. Если рассматривать гипотезу H_0 о равенстве θ =*T*, то насколько велико должно быть различие между θ и *T*, чтобы эту гипотезу отвергнуть. Ответить на данный вопрос можно в статистическом смысле, рассматривая вероятность достижения некоторой заданной разности между θ и *T* на основе выборочного распределения параметра θ .

Целесообразно полагать одинаковыми значения вероятности выхода параметра θ за нижний и верхний пределы интервала. Такое допущение во многих случаях позволяет минимизировать доверительный интервал, т.е. повысить мощность критерия провер-

ки. Суммарная вероятность того, что параметр θ выйдет за пределы интервала с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$, составляет величину α. Эту величину следует выбрать настолько малой, чтобы выход за пределы интервала был маловероятен. Если оценка параметра попала в заданный интервал, то в таком случае нет оснований подвергать сомнению проверяемую гипотезу, следовательно, гипотезу равенства $\theta = T$ можно принять. Но если после получения выборки окажется, что оценка выходит за установленные пределы, то в этом случае есть серьезные основания отвергнуть гипотезу H_0. Отсюда следует, что вероятность допустить ошибку первого рода равна α (равна уровню значимости критерия).

Если предположить, например, что истинное значение параметра в действительности равно $T+d$, то согласно гипотезе H_0 о равенстве $\theta = T$ – вероятность того, что оценка параметра θ попадет в область принятия гипотезы, составит β, рисунок 5.2.

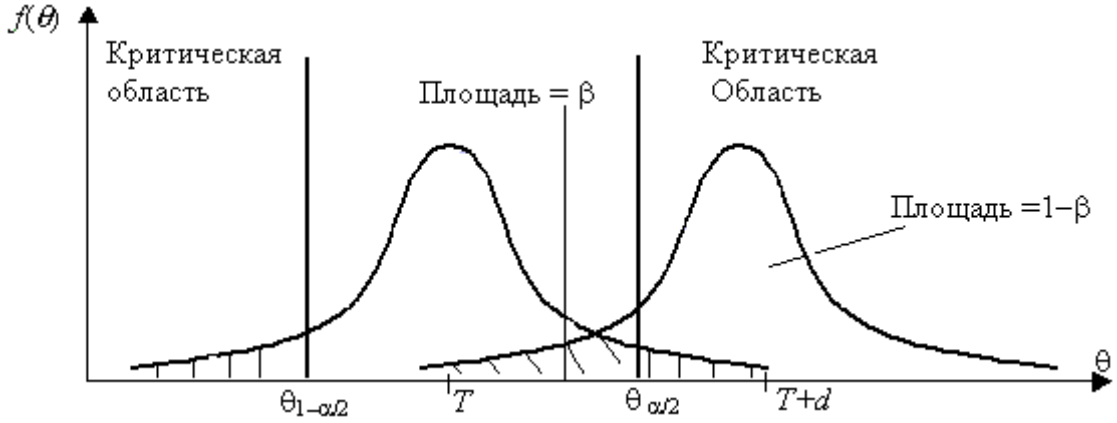


Рисунок 5.2 – Распределение областей

При заданном объеме выборки вероятность совершения ошибки первого рода можно уменьшить, снижая уровень значимости α. Однако при этом увеличивается вероятность ошибки второго рода β (снижается мощность критерия). Аналогичные рассуждения можно провести для случая, когда истинное значение параметра равно $T - d$.

Единственный способ уменьшить обе вероятности состоит в увеличении объема выборки (плотность распределения оценки параметра при этом становится более "узкой"). При выборе критической области руководствуются правилом Неймана – Пирсона: следует так выбирать критическую область, чтобы вероятность α была мала, если гипотеза верна, и велика в противном случае. Однако выбор конкретного значения α относительно произволен. Употребительные значения лежат в пределах от 0,001 до 0,2. В целях упрощения ручных расчетов составлены таблицы интервалов с границами $\theta_{1-\alpha/2}$ и $\theta_{\alpha/2}$ для типовых значений α и различных способов построения критерия.

При выборе уровня значимости необходимо учитывать мощность критерия при альтернативной гипотезе. Иногда большая мощность критерия оказывается существеннее малого уровня значимости, и его значение выбирают относительно большим, например 0,2. Такой выбор оправдан, если последствия ошибок второго рода более существенны, чем ошибок первого рода.

Например, если отвергнуто правильное решение "продолжить работу пользователей с текущими паролями", то ошибка первого рода приведет к некоторой задержке в нормальном функционировании системы, связанной со сменой паролей. Если же принято решения не менять пароли, несмотря на опасность несанкционированного доступа посторонних лиц к информации, то эта ошибка повлечет более серьезные последствия.

В зависимости от сущности проверяемой гипотезы и используемых мер расхождения оценки характеристики от ее теоретического значения применяют различные критерии. К числу наиболее часто применяемых критериев для проверки гипотез о законах распределения относят критерии хи-квадрат Пирсона, Колмогорова, Мизеса, Вилкоксона, о значениях параметров – критерии Фишера, Стьюдента.

Рассмотрим практическую методику использования метода статистической оценка параметров распределения. Состоятельные и несмещенные оценки основных параметров распределения случайной величины (СВ) (математического ожидания M_X и дисперсии σ_X^2) могут быть получены по формулам:

$$\overline{X} = \frac{1}{\mathrm{n}} \cdot \sum_{i=1}^{n} X_i \qquad (5.5)$$

$$\sigma_X{}^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \overline{X})^2 \qquad (5.6)$$

где n - объем выборки.

Оценку коэффициента корреляции между случайными величины X и Y определяют по формуле:

$$\mathrm{R}_{XY} = \frac{1}{\mathrm{n}-1} \cdot \frac{\sum_{i=1}^{n} (X_i - \overline{X}) \cdot (Y_i - \overline{Y})}{\sigma_X \cdot \sigma_Y} \qquad (5.7)$$

Так как оценки (5.5) - (5.6) определяют по выборке конечного объема, возникает вопрос об их статистической достоверности и точности.

Обозначим θ как оценку интересующего нас параметра. Тогда задача определения достоверности и точности оценки сводится к определению такого интервала (θ1, θ2), включающего параметр θ, что с вероятностью 1 - α (где α - достаточно малая величина, равная 0.1, 0.05, 0.01 …) можно утверждать, что неизвестное истинное значение параметра находится в этом интервале. Интервал (θ1, θ2) называют доверительным интервалом, а вероятность 1- α доверительной вероятностью.

Рассмотрим случай, когда величина X имеет нормальный закон распределения с плотностью вероятности:

$$f(x) = \frac{1}{\sigma_X \cdot \sqrt{2\pi}} \cdot e^{-\frac{(X - M_x)^2}{2 * \sigma_X^2}} \qquad (5.8)$$

Доверительный интервал для математического ожидания:
$\overline{X} - \varepsilon, \overline{X} + \varepsilon$, включающий M_x с вероятностью 1- α, находят из условия:
$P[\overline{X} - \varepsilon < M_X < \overline{X} + \varepsilon] = 1 - \alpha$

которое можно представить в виде:
$P[|X - M_X| < \theta] = 1 - \alpha$ (5.9)

Введем параметр:

$$t = [(\overline{X} - M_x)/\sigma_X] \cdot \sqrt{n} \qquad (5.10)$$

имеющий t-распределение Стьюдента с v = n - 1 степенями свободы. Тогда равенство (3.4) перепишется в виде:

$$P = \left[|\overline{X} - M_X| < t(\alpha, v) \cdot \frac{\sigma_X}{\sqrt{n}} \right] = 1 - \alpha \qquad (5.11)$$

где t(a, v) определяют по таблице распределения Стьюдента при вероятности α и степени свободы v = n - 1. Доверительный интервал для Mx, соответствующий доверительной вероятности 1 - α, есть:

$$\left[\overline{X} - \frac{t(\alpha, v) \cdot \sigma_X}{\sqrt{n}}; \overline{X} + \frac{t(\alpha, v) \cdot \sigma_X}{\sqrt{n}} \right] \qquad (5.12)$$

Чтобы определить доверительный интервал для дисперсии, необходимо найти границы интервала σ12и σ22, удовлетворяющие равенству:

$P[\sigma 12 < \sigma X2 < \sigma 22] = 1 - \alpha$ (5.13)

Для нормально распределенного X известен закон распределения величины со степенями свободы v = n - 1:

$\chi 2 = (n-1) \cdot \sigma X2 / \sigma 2$, (5.14)

где σX2-выборочная дисперсия, σ2-истинное значение σX2

После подстановки (5.14) в (5.13), при условии, что:
P[σX2<σ12]=P[σX2>σ22] = α/2, получим:

P[χ2(1-α/2, v) < (n-1)·σX2/σ2 < χ2(α/2, v)] = 1 - α.

Величину χ2(1-α/2,v)=(n-1)·σX2/σ22находят по таблице распределения Пирсона при вероятности 1-α/2 и числе степеней свободы
v=n-1, а χ2(α/2, v)=(n-1)·σX2/σ2определяют при вероятности α/2 и числе степеней свободы v = n - 1.
Следовательно, доверительный интервал для дисперсии σX2, соответствующий доверительной вероятности 1 - α, есть:

$$\left[\frac{(n-1)\cdot\sigma_X^{\,2}}{\chi^2(\alpha/2,v)};\frac{(n-1)\cdot\sigma_X^{\,2}}{\chi^2(1-\alpha/2,v)}\right] \qquad (5.14)$$

Алгоритм проверки статистических гипотез. Понятие статистической гипотезы означает предположение о виде распределения СВ или о некотором параметре ее распределения. Проверка гипотезы заключается в сопоставлении определенного статистического показателя (критерия значимости), вычисленного по данной выборке, с критерием значимости, найденным теоретически при условии, что проверяемая гипотеза верна.
1)При проверке гипотезы о том, что Mx = C, в качестве критерия используют величину:

$$t = (\overline{X} - C)\cdot\sqrt{n}/\sigma_X \qquad (5.15)$$

Эта величина при условии, что гипотеза верна, имеет t-распределение Стьюдента с v = n - 1 степенями свободы. Если вычисленное по соотношению (5.15) значение t по абсолютной величине не превышает критического значения $t_{кр}$=t(α, v), найденного по таблице t-распределения при уровне значимости α и числе степеней свободы v, то гипотеза о том, что Mx=C принимается, в противном случае она отвергается.
2)Проверку гипотезы о равенстве двух математических ожиданий Mx = My, вычисленных по двум выборкам случайных величин X и Y объемами n1 и n2 проводят по критерию:

t = (X-Y)/σX-Y (5.16)

$$\sigma_{X-Y} = \sqrt{\frac{(n_1+n_2)\cdot[(n_1-1)\cdot\sigma_X^{\,2}+(n_2-1)\cdot\sigma_Y^{\,2}]}{n_1\cdot n_2\cdot(n_1+n_2-2)}} \qquad (5.17)$$

Критерий t имеет t-распределение Стьюдента с числом степеней свободы v = n1 + n2 - 2. Проверку гипотезы проверяют также, как и в предыдущем случае, т.е. при |t|<=tкр гипотеза принимается, а при $|t|>t_{кр}$ отвергается.

3)Проверку гипотезы о равенстве дисперсий двух СВ X и Y, оценки которых σX2 и σY2 определены по двум выборкам объемом n1 и n2, проводят с использованием критерия:

$$F = \sigma X2/\sigma Y2, \qquad (5.18)$$

который имеет распределение Фишера со степенями свободы v1 = n1 - 1 для числителя и v2 = n2 - 1 для знаменателя. Полученное по критерию (5.18) значение сравнивают с критическим Fкр=F(α, v1, v2). Если F<Fkp нет оснований для того, чтобы нулевая гипотеза была отвергнута, в противном случае принимаем, что на генеральной совокупности σX2>σY2.

4)При проверке гипотезы об отсутствии корреляции между двумя СВ используют соотношение:

$$t = Rxy/\sigma R, \qquad (5.19)$$

где:

Rxy-оценка коэффициента корреляции найденная по (5.7),

$$\sigma R2 = [(1-Rxy2)/(n-2)]$$

Величина t имеет t-распределение Стьюдента с v = n - 2 степенями свободы. Если вычисленное по соотношению (5.19) значение t по абсолютной величине не превышает критического значения $t_{кр}=t(\alpha, v)$, найденного по таблице t-распределения при уровне значимости α и числе степеней свободы v, нет оснований для того, чтобы гипотеза об отсутствии корреляции на генеральной совокупности была отвергнута, в противном случае принимаем, что между величинами X и Y существует корреляция.

Основная литература

1. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.
2. Рузинов Л.П. Статистические методы оптимизации химических процессов. -М.: Химия, 1972

Дополнительная литература

3. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.

Лекция 6 Статические и динамические модели в форме уравнения регрессии (ч. 4)

В этой лекции нами будут рассмотрены методы планирования эксперимента, используемые при определении статические и динамические модели объектов в форме управления регрессии

Основные недостатки классического регрессионного анализа: корреляция между коэффициентами; трудности в оценке ошибки; необходимость проведения большого количества опытов; трудности при определении коэффициентов b_i вручную.

В основе методов статистического плана эксперимента (ПЭ) лежит использование упорядоченного плана расположения точек в факторном пространстве и использование новой безразмерной системы координат. Использование методов ПЭ позволяет устранить практически все недостатки регрессионного анализа.

Методы планирования эксперимента позволяют строить как статические, так и динамические модели. Те и другие могут быть определены аналитическими и экспериментально- статистическими методами. Теория оптимального эксперимента на начальном этапе развития имитационного моделирования применялась преимущественно при построении моделей статических объектов. Что касается построения динамических моделей, то методология построения базировалась в рамках пассивного эксперимента, когда идентификация велась в режиме нормального функционирования объекта. Однако, постепенно стали применять и активные методы идентификации с использованием специальных тестирующих сигналов. В качестве таких сигналов использовали псевдослучайные сигналы.

Основные понятия на которые надо обратить внимание: факторы (входы, воздействия – они не должны быть коррелированны), уровни (дискретные значения, принимаемые факторами), матрица ПЭ (собственно план), методы математической обработки.

Основные принципы планирования эксперимента (ПЭ).

Пассивным экспериментом называют эксперимент, в котором регистрация входных и выходных данных осуществляется в рабочем режиме, не используя дополнительных вмешательств. Он применяется тогда, когда структура модели хорошо известна и ее адекватность не вызывает сомнений (когда решаются задачи параметрической идентификации).

Активный эксперимент предполагает особую программу (план) проведения наблюдений таких, что позволяют по результатам исследований дополнительно оценить структуру модели.

Факторами активного эксперимента называют переменные, по которым возможно проводить управление и которые участвуют в построении модели (x_i).

Каждый из факторов может принимать различные значения, которые называются **уровнями**. На практике, количество уровней – это бесконечное количество или непрерывный ряд уровней $x_i \in \lfloor x_{0i},\ x_{ni} \rfloor$. В теории активного эксперимента этот ряд дискретизируется, и выбираются отдельные уровни: $\lfloor x_{i0},\ x_{i1},\ x_{i2} \rfloor$

Фиксированный набор уровней называется состоянием факторов.

План – это программа проведения эксперимента, позволяющая использовать все факторы на всех уровнях. Если план содержит всевозможные сочетания факторов и уровней, то такой план называют полным.

Если p – общее количество уровней; к – количество факторов, то полный план эксперимента будет включать в себя следующее количество экспериментов: $N = p^{\kappa}$.

методах ПЭ используются кодированные (безразмерные) параметров. Они соответствуют переносу начала координат в точку z_j^0 - центр плана (основной уровень).

$$z_j^0 = \frac{z_j^{\max} + z_j^{\min}}{2},\ j = 1,2..k. \qquad (6.1)$$

Интервал варьирования по оси z_j определяется по формуле:

$$\Delta z_j = \frac{z_j^{\max} - z_j^{\min}}{2} \qquad (6.2).$$

Для перехода от физических (натуральных) переменных z_j к безразмерным координатам x_j используют выражение:

$$x_j = \frac{z_j - z_j^0}{\Delta z_j} \qquad (6.3),$$

а для обратного перехода

$$z_j = z_j^0 + x_j \cdot \Delta z_j \qquad (6.5).$$

Например, пусть $z_1^{\max} = 200°C$ $z_1^{\min} = 100°C$, тогда

$\Delta z_1 = 50°C$, $z_1^0 = 150°C$ $x_1 = \dfrac{z_1 - z_1^0}{\Delta z_1}$,

при $z_1 = z_1^{\min} = 100°C$ $x_1 = \dfrac{100-150}{50} = -1$.

Для безразмерных переменных в ПФЭ верхний уровень равен +1, а нижний –1, координаты центра плана равны нулю. В некоторых планах (см. ниже) используются безразмерные значения факторов $x_j = \pm\alpha$, соответствующие им значения в натуральных координатах находят по выражению $z_j = z_j^0 + \alpha \cdot \Delta z_j$.

В приведенном примере при $\alpha = 1.4; z_1 = 150 + 1.4 \cdot 50 = 220°C$

Отметим, что введение безразмерной системы координат необходимо только для облегчения обработки результатов вручную. Если использовать для поиска решения матричное уравнение (4.15), то сразу получаем решение с найденными коэффициентами в натуральном (физическом) масштабе.

В полном факторном эксперименте (ПФЭ) обеспечивается проведение опытов при всех возможных сочетаниях значений факторов на всех уровнях варьирования, используются значения факторов только равные +1 и –1.

Необходимое число опытов ПФЭ: $N = (N_{УРОВНЕЙ})^K$, (6.6)

чаще используют $N_{УРОВНЕЙ} = 2$, тогда необходимое количество опытов см. таблицу 6.1):

Таблица 6.1

Необходимое количество опытов

Количество факторов –k	3	4	5	6	7
Количество опытов – N	8	16	32	64	128
Кол. коэф. линейной модели - L	4	5	6	7	8
Число степеней свободы - f	4	11	26	57	120

Матрица ПФЭ для 3-х факторов x_1 x_2 x_3 в безразмерном виде (y – выход, результат проведения опыта) имеет вид (см. таблицу. 6.2.):

Таблица 6.2

Матрица ПФЭ для 3-х факторов

№ опыта	x_0	x_1	x_2	x_3	y
1	+1	–1	–1	–1	y_1
2	+1	+1	–1	–1	y_2
3	+1	–1	+1	–1	y_3
4	+1	+1	+1	–1	y_4
5	+1	–1	–1	+1	y_5
6	+1	+1	–1	+1	y_6
7	+1	–1	+1	+1	y_7
8	+1	+1	+1	+1	y_8

План эксперимента понимается так. Например, для первого опыта:

$z_1 = z_1^{\min}$ $z_2 = z_2^{\min}$ $z_3 = z_3^{\min}$, результат опыта y_1, для второго опыта:

$z_1 = z_1^{\max}$ $z_2 = z_2^{\min}$ $z_3 = z_3^{\min}$, результат опыта y_2 и т.д.

Столбец x_0 соответствует т.н. – фиктивной переменной $x_0 = +1$, он необходим только для удобства вывода приведенных ниже выражений. Вычисление коэффициентов уравнения регрессии bi производят на основе известного выражения (5.14) $B = (X^T X)^{-1} X^T Y$. Однако приведенная в таблице матрица обладает рядом свойств, значительно упрощающих расчеты вручную.

$$\sum_{i=1}^{N} x_{ui} \cdot x_{ji} = 0; u \neq j; u, j = 0,1,...,k \text{ - свойство ортогональности;} \qquad (6.7)$$

$$\sum_{i=1}^{N} x_{ji} = 0;\ j = 1,...,k;\ j \neq 0 \qquad (6.8)$$

$$\sum_{i=1}^{N} x_{ji}^2 = N;\ j = 0,1,...,k; \qquad (6.9)$$

Ортогональность матрицы планирования приводит к диагональности матрицы $(X^T X)$, в результате чего легко вручную находится $(X^T X)^{-1}$ (см. стр. 162 в [1].), а коэффициенты уравнения регрессии b_i находятся по формуле:

$$b_j = \frac{1}{N}\sum_{i=1}^{N} x_{ji} \cdot y_i \qquad (6.10)$$

(напомним, что для можно сразу использовать уравнение (4.15).

Матрица ПЭ, приведенная в таблице 6.2. позволяет определить коэффициенты линейного уравнения регрессии:

$\hat{\hat{y}} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$.

В принципе имеющиеся степени свободы позволяют определить коэффициенты уравнения регрессии с учетом эффектов взаимодействия вида:

$$\hat{\hat{y}} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{123} x_1 x_2 x_3 \qquad (6.11)$$

Для этого матрицу ПЭ (см. таблицу 6.2.) надо дополнить фиктивными столбиками, содержащими информацию для x_1x_2, x_1x_3, x_2x_3 и $x_1x_2x_3$ (см. таблицу. 6.3.). Например $x_1 \cdot x_2 = x_1x_2 = (-1 \cdot -1) = +1$. При этом дополнительных опытов проводить не надо.

Таблица 6.3
План ПФЭ

№ оп.	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	2	3	4	5	6	7	8	9	10
1	+1	–1	–1	–1	+1	+1	+1	–1	y_1
2	+1	+1	–1	–1	–1	–1	+1	+1	y_2
3	+1	–1	+1	–1	–1	+1	–1	+1	y_3
4	+1	+1	+1	–1	+1	–1	–1	–1	y_4
5	+1	–1	–1	+1	+1	–1	–1	+1	y_5
6	+1	+1	–1	+1	–1	+1	–1	–1	y_6
7	+1	–1	+1	+1	–1	–1	+1	–1	y_7
8	+1	+1	+1	+1	+1	+1	+1	+1	y_8

Значимость коэффициентов и адекватность уравнения регрессии определяют, как описано выше. После того, как определен окончательный вид уравнения регрессии можно перейти от кодированных переменных к физическим переменным используя уравнения (6.5).

Дробный факторный эксперимент.

Другое название - метод дробных реплик. С увеличением количества факторов в ПФЭ резко возрастает число опытов и число степеней свободы. (см. таблицу 6.1.). Т.е. количество опытов в ПФЭ обычно превосходит число определяемых коэффициентов, т.е. ПФЭ обладает избыточностью опытов. В дробном факторном эксперименте (ДФЭ) в отличие от ПФЭ используются не все возможные сочетания уровней изучаемых факторов. Некоторые их сочетания пропускаются. Сокращение перебора уровней всегда приводит к потере части исходной информации. План ДФЭ строится так, чтобы обеспечить получение линейной модели, при этом необходимо, чтобы матрица планирования сохранила свойство ортогональности. Рассматриваемый метод заключается в том, что для получения математического описания используется часть ПФЭ. Например, при 4 факторах за основу берется план для трех факторов, а для четвертого фактора используется план, соответствующий одному из столбцов содержащих эффекты взаимодействия (столбцы 6-9 в таблице 6.3.). Если нет сведений о значимости взаимодействий факторов, то лучше выбрать столбец 9, соответствующий тройным взаимодействиям $x_4 = x_1 x_2 x_3$ или $x_4 = -x_1 x_2 x_3$, такие соотношения называют генерирующими соотношениями, так как они генерируют (создают) дробную реплику. Умножая обе части генерирующего соотношения на x_4 получаем $I = x_1 x_2 x_3 x_4$, называемое определяющим контрастом. Эксперимент проводят по плану (таблица 6.4):

Таблица 6.4
План ДФЭ

№ опыта	x_0	x_1	x_2	x_3	x_4	y
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	+1	–1	–1	–1	–1	y_1
2	+1	+1	–1	–1	+1	y_2
3	+1	–1	+1	–1	+1	y_3
4	+1	+1	+1	–1	–1	y_4
5	+1	–1	–1	+1	+1	y_5
6	+1	+1	–1	+1	–1	y_6
7	+1	–1	+1	+1	–1	y_7
8	+1	+1	+1	+1	+1	y_8

По этому плану, например, для первого опыта:

$z_1 = z_1^{\min}$ $z_2 = z_2^{\min}$ $z_3 = z_3^{\min}$ $z_4 = z_4^{\min}$, результат опыта y_1.

План содержит половину опытов ПФЭ 24 и носит название полуреплики. Используют также ¼ реплики, ⅛ реплики и т.д. Обозначаются ДПЭ в виде 2k-p, где k – количество факторов, p –число дробных реплик, в нашем примере это 24-1. Расчет коэффициентов регрессии, проверка их значимости производят по тем же формулам, что и для ПФЭ.

Ортогональное планирования эксперимента

Другое название - планы Бокса-Уилсона. Для построения ортогональных центральных композиционных планов (ЦКП) используется так называемое звездное плечо α (плечо звездных точек). Это опыты, которые проводятся дополнительно, например, к плану ПФЭ (при k<5) или к дробному плану(при k≥5). В таблице 6.5. приведено такое дополнение к плану в таблице 6.2.

Таблица 6.5

Дополнение к плану в таблице 6.2

№ опыта	x_0	x_1	x_2	x_3	y	
9	+1	–α	0	0	y9	Звездные точки (2k точек)
10	+1	+α	0	0	y10	
11	+1	0	–α	0	y11	
12	+1	0	+α	0	y12	
13	+1	0	0	–α	y13	
14	+1	0	0	+α	y14	
15	+1	0	0	0	Y15	Центр плана N0 точек

Общее количество опытов определяется по формуле:

$N = 2^k + 2 \cdot k + N_0$ (при k<5) (6.12)

Обычно звездные точки и точку в центре добавляют, если при ПФЭ не удается получить адекватную математическую модель и при исследовании области, близкой к экстремуму. При описании такой области наиболее широко используют полиномы второго порядка. ЦКП неортогональны, но легко сводятся к ортогональным (ОЦКП), путем выбора α. Значения α приведены в таблице 6.6. ([1]., стр.181, 183)

Таблица 6.6

N_0	k=2	k=3	k=4
1	α=1,000	α=1,476	α=2,000
2	α=1,160	α=1,650	α=2,164
3	α=1,317	α=1,831	α=2,390

Матрица планирования для ОЦКП, используемая для расчета коэффициентов дополняется столбцами x_1^*, x_2^*…Например при k=2 и N0=1 получаем матрицу, приведенную в таблице 6.7. В таблице $x_j^* = x_j^2 - \bar{x}_j^2 = x_j^2 - \frac{1}{N}\sum_{i=1}^{N} x_{ji}^2$, введено для получения ортогональной матрицы.

Таблица 6.7

№ опыта	x_0	x_1	x_2	x_1x_2	x_1^*	x_2^*	y
1	2	3	4	6	8	9	10
1	+1	+1	+1	+1	+0,333	+0,333	y_1
2	+1	+1	–1	–1	+0,333	+0,333	y_2
3	+1	–1	–1	+1	+0,333	+0,333	y_3
4	+1	–1	+1	–1	+0,333	+0,333	y_4
5	+1	+1	0	0	+0,333	–0,667	y_5
6	+1	–1	0	0	+0,333	–0,667	y_6
7	+1	0	+1	0	–0,667	+0,333	y_7
8	+1	0	–1	0	–0,667	+0,333	y_8
9	+1	0	0	0	–0,667	–0,667	y_9

Благодаря ортогональности матрицы планирования все коэффициенты рассчитываются независимо друг от друга по формуле:

$$b_j = \frac{\sum_{i=1}^{N} x_{ji} \cdot y_i}{\sum_{i=1}^{N} x_{ji}^2}, \qquad (6.13),$$

а дисперсии коэффициентов равны:

$$s_{b_j}^2 = \frac{s_{ВОСПР}^2}{\sum_{i=1}^{N} x_{ji}^2} \qquad (6.14)$$

В результате расчета по матрице в таблице 6.7 уравнение регрессии в ОЦКП получается в виде:

$$\hat{Y} = b_0^* + b_1 x_1 + b_2 x_2 + ... b_k x_k + b_{12} x_1 x_2 + ... + b_{(k-1)k} x_{k-1} x_k + ... + b_{11} x_1^* + ... + b_{kk} x_k^*$$

Чтобы перейти к обычной записи b_0 определяют по выражению:

$$b_0 = b_0^* - b_{11}\bar{x}_1^2 - ... - b_{kk}\bar{x}_k^2$$

и оценивают с дисперсией: $s_{b_0}^2 = s_{b_0^*}^2 + (\bar{x}_1^2)^2 s_{b_{11}}^2 + ... + (\bar{x}_k^2)^2 s_{b_{kk}}^2$

Зная дисперсию воспроизводимости, проверяют значимость коэффициентов и адекватность уравнения регрессии:

$$\hat{Y} = b_0 + b_1 x_1 + b_2 x_2 + ... b_k x_k + b_{12} x_1 x_2 + ... + b_{(k-1)k} x_{k-1} x_k + ... + b_{11} x_1^2 + ... + b_{kk} x_k^2$$

Адекватность проверяют по критерию Фишера, составляя отношение дисперсий: $F_{PAC} = \frac{S_{АД}^2}{S_{ВОСПР}^2}$, условие адекватности $F_{PAC} < F_{1-p}(f_1, f_2)$, где

$f_1 = f_{АД} = N - L \quad f_2 = f_{ВОСПР}$.

Значимость коэффициентов определяют по формуле:

$t_j = \frac{|b_j|}{S_{b_i}}$, если $t_j > t_{ТАБ}(f_2)$, то b_j считается значимым.

Т.к. b_j в ОЦКП определяются с разной точностью, то (при k<5):

$$s_{b_0} = \frac{s_{ВОСПР}}{\sqrt{N}},$$

$$s_{b_j} = \frac{s_{ВОСПР}}{\sqrt{2^k + 2\alpha^2}}, \ \mathrm{j} = 1,2,...,\mathrm{k}$$

$$s_{b_{uj}} = \frac{s_{ВОСПР}}{\sqrt{2^k}} \ \mathrm{u,j} = 1,2,..,\mathrm{k}, \ \mathrm{u} \neq \mathrm{j}$$

$$s_{b_{jj}} = \frac{s_{ВОСПР}}{\sqrt{2^k(1-\bar{x}_1^2)^2 + 2(\alpha^2 - \bar{x}_1^2)^2 + (N_0 + 2k - 2)(\bar{x}_1^2)^2}}, \ \mathrm{j} = 1,2,...,\mathrm{k}$$

Ротатабльное планирования эксперимента (РПЭ).

Другое название - планы Бокса-Хантера. Ортогональные планы не обладают свойством ротатабельности. РПЭ позволяет получать более точное математическое описание по сравнению с ЦКП. Это достигается увеличением числа опытов в центре плана и специальным выбором звездного плеча **α**. Характеристики РПЭ приведены в табл. 6.8.

Таблица 6.8.

Расчет числа опытов

Количество					Значение **α**
факторов - **k**	опытов				
	ПФЭ	в звездных точках	в центре плана – N_0	всего - N	
2	4	4	5	13	1,414
3	8	6	6	20	1,680
4	16	8	7	31	2,000
5	32	10	10	52	2,378

Коэффициенты уравнения регрессии коррелированны между собой, поэтому их можно найти, используя матричное выражение (5.14)

$B = (X^T X)^{-1} X^T Y$, т.е. надо выполнить обращение матрицы $(X^T X)$. Однако из-за особого вида этой матрицы при РПЭ можно получить аналитические выражения для расчета коэффициентов уравнения регрессии и их дисперсий.

Они имеют вид:

$$\left.\begin{aligned} b_0 &= a_1\sum_{i=1}^{N} y_i - a_2\sum_{j=1}^{k}\sum_{i=1}^{N} x_{ji}^2 y_i \quad b_j = a_3\sum_{i=1}^{N} x_{ji} y_i;\ \mathrm{j} = 1,2,...,\mathrm{k} \\ b_{uj} &= a_4\sum_{i=1}^{N} x_{ui} x_{ji} y_i;\ \mathrm{u} \neq \mathrm{j};\ \mathrm{j},\mathrm{u} = 1,2,...,\mathrm{k} \\ b_{jj} &= a_5\sum_{i=1}^{N} x_{ji}^2 y_i + a_6\sum_{j=1}^{k}\sum_{i=1}^{N} x_{ji}^2 y_i + a_7\sum_{i=1}^{N} y_i \end{aligned}\right\} \qquad (6.15)$$

$$s^2_{b_0} = a_1 s^2_y \;\; s^2_{b_j} = a_3 s^2_y \;\; s^2_{b_{uj}} = a_4 s^2_y \;\; s^2_{b_{jj}} = (a_5 + a_6) s^2_y, \qquad (6.66)$$

где значения a_i приведены в таблице в таблице 6.9.

Таблица 6.9.

k	N	N_0	α	a_1	a_2	A_3	a_4	a_5	a_6	a_7
2	13	5	1,412	0,2000	0,1000	0,1250	0,2500	0,1251	0,0187	0,1000
3	20	6	1,682	0,1663	0,0568	0,0732	0,1250	0,0625	0,0069	0,0568
4	31	7	2,000	0,1428	0,0357	0,0417	0,0625	0,0312	0,0037	0,0357
5	32	6	2,000	0,1591	0,0341	0,0417	0,0625	0,0312	0,0028	0,0341

Дисперсию воспроизводимости определяют по опытам в центре плана:

$$S^2_{BOCПР} = \frac{\sum_{u=1}^{N_0} (y_u^0 - \bar{y}^0)^2}{N_0 - 1} \quad \bar{y}^0 = \frac{\sum_{u=1}^{N_0} y_u^0}{N_0} \quad f_{BOCПР} = (N_0 - 1).$$

Остаточная дисперсия определяется по формуле:

$$S^2_{OCT} = \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N - L} \quad f_{OCT} = (N - L).$$

Адекватность уравнения регрессии определяют по критерию Фишера:

$F_{PAC} = \dfrac{S^2_{АД}}{S^2_{BOCПР}}$, где дисперсия адекватности:

$$S^2_{АД} = \frac{S^2_{OCT} f_{OCT} - S^2_{BOCПР} f_{BOCПР}}{f_{АД}} \quad f_{АД} = f_{OCT} - f_{BOCПР}$$

Уравнение адекватно, если $F_{PAC} < F_{1-p}(f_1, f_2)$, где $f_1 = f_{АД} \;\; f_2 = f_{BOCПР}$.

Значимость коэффициентов определяют по критерию Стьюдента, однако, если один из квадратичных критериев оказался незначимым, то после его исключения коэффициенты уравнения регрессии необходимо пересчитать заново.

Основная литература

1. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.
2. Инков А.М. Моделирование и идентификация объектов управления. Методические указания к выполнению лабораторных работ для студентов спец. 050702. Шымкент, ЮКГУ, 2010 г., -78 с.

Дополнительная литература

3. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.
4. Гроп Д. Методы идентификации систем. - М.: Мир, 1979

Модуль 2. Моделирование и идентификация динамических характеристик объектов

Тема 4 Общие задачи идентификации

Лекция 7 Общие задачи идентификации. Основные определения.

В данной лекции мы рассмотрим Основные характеристики времянных рядов, общие задачи идентификации, основные определения, идентификация объектов управления методом корреляционного анализа.

Основные характеристики времянных рядов

Подробнее (см. [1-4]). Основные понятия: случайная величина и случайный процесс (СП), времянной ряд (ВР), реализация, ансамбль, стационарный СП, эргодический СП, оценка. У стационарных СП матожидание и дисперсия постоянны.

Временной ряд (или ряд динамики) — это упорядоченная по времени последовательность значений некоторой произвольной переменной величины. Тем самым, временной ряд существенным образом отличается от простой выборки данных. Каждое отдельное значение данной переменной называется отсчётом (уровнем, элементов) временного ряда.

В то же время выборка это множество случаев (испытуемых, объектов, событий, образ-цов), с помощью определённой процедуры выбранных из генеральной совокупности для участия в исследовании.

Анализ временных рядов это совокупность математико-статистических методов анализа, предназначенных для выявления структуры временных рядов и для их прогнозирования. Сюда относятся, в частности, методы регрессионного анализа. Выявление структуры вре-мянного ряда необходимо для того, чтобы построить математическую модель того явления, которое является источником анализируемого временного ряда. Прогноз будущих значений временного ряда используется для эффективного принятия решений.

Временные ряды состоят из двух элементов:

- периода времени, за который или по состоянию на который приводятся числовые значения;
- числовых значений того или иного показателя, называемых уровнями ряда.

Временные ряды классифицируются по следующим признакам:

- по форме представления уровней:
  - ряды абсолютных показателей;
  - относительных показателей;
  - средних величин.
- по количеству показателей, для который определяются уровни в каждый момент времени: одномерные и многомерные временные ряды;
- по характеру временного параметра: моментные и интервальные временные ряды. В моментных временных рядах уровни характеризуют значения показателя по состоянию на определенные моменты времени. В интервальных рядах уровни характеризуют значение показателя за определенные периоды времени. Важная особенность интервальных временных рядов абсолютных величин заключается в возможности суммирования их уровней. Отдельные же уровни моментного ряда абсолютных величин содержат элементы повторного счета. Это делает бессмысленным суммирование уровней моментных рядов;
- по расстоянию между датами и интервалами времени выделяют равноотстоящие – когда даты регистрации или окончания периодов следуют друг за другом с равными интервалами и неполные (неравноотстоящие) – когда принцип равных интервалов не соблюдается;
- по наличию пропущенных значений: полные и неполные временные ряды;
- временные ряды бывают детерминированными и случайными: первые получают на основе значений некоторой неслучайной функции (ряд последовательных данных о количестве дней в месяцах); вторые есть результат реализации некоторой случайной величины.
- в зависимости от наличия основной тенденции выделяют **стационарные** ряды – в которых среднее значение и дисперсия постоянны и **нестационарные** – содержащие основную тенденцию развития.

Примеры временных рядов (см. рисунок 7.1)**.** Временные ряды, как правило, возникают в результате измерения некоторого показателя. Это могут быть как показатели (характеристики) технических систем, так и показатели природных, социальных, экономических и других систем (например, погодные данные). Типичным примером временного ряда можно назвать биржевой курс, при анализе которого пытаются определить основное направление развития (тенденцию или тренда).

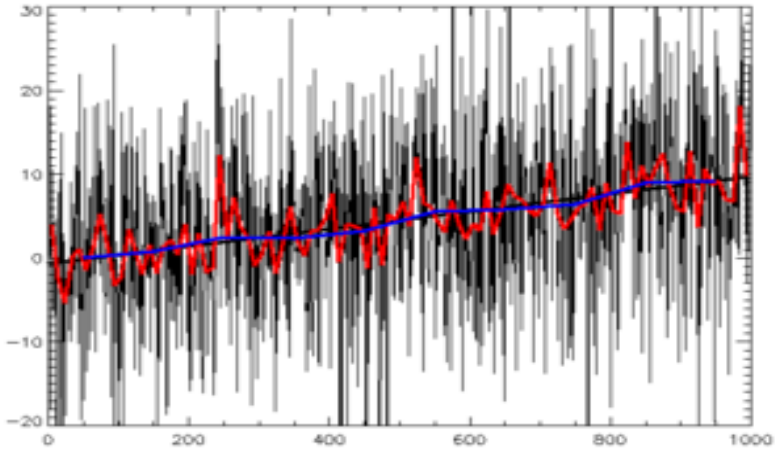


Рисунок 7.1 - Пример временного ряда

Основной чертой, выделяющей анализ временных рядов среди других видов статистического анализа, является существенность порядка, в котором производятся наблюдения. Если во многих задачах наблюдения статистически независимы, то во временных рядах они, как правило, зависимы, и характер этой зависимости может определяться положением наблюдений в последовательности. Природа ряда и структура порождающего ряд процесса могут предопределять порядок образования последовательности.

Иногда **нестационарный** СП g_i можно сделать стационарным $\tilde{y}_i$, удалив низкочастотную аддитивную составляющую (тренд, дрейф), например, применив фильтр верхних частот, использующий спектральное окно Тьюки:

$$\tilde{y}_i = \sum_{j=-m}^{m} h_i g_{i-j},$$

где $h_0 = 1 - \frac{1}{m+1}$ $\qquad h_j = h_{-j} = \frac{1}{2(m+1)}\left[1 + \cos\left(\frac{j\pi}{m+1}\right)\right]$, (7.1)

частота среза fc фильтра связана с его порядком m и частотой изменения времянного ряда $f_{ИЗ}$ $f_c = f_{ИЗ}/2m$.

Среднее значение (матожидание, МО) ВР можно определить как:**.**

$$m_x = M[x(t)] = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)dt ,$$

Оценка МО $\tilde{m}_x = \frac{1}{N}\sum_{i=1}^{N} x_i$ (7.2).

(Среднее по N измерениям, выборочное среднее, характеризует постоянное смещение ВР, неслучайная величина, для стационарного СП $\tilde{m}_x = const$).

В условиях АСУТП обычно стоит задача определения МО в режиме реального времени (online). Для этого может быть применена рекуррентная формула:

$$\bar{x}_i = \bar{x}_{i-1} + \frac{1}{i}(x_i - \bar{x}_{i-1}), \quad (7.3)$$

где $\bar{x}_i$ текущее значение МО, $\bar{x}_{i-1}$- МО определенное на предыдущем шаге, x_i текущее значение x (сигнал от датчика).

Другой вид формулы $\bar{x}_i = (1-\alpha)\bar{x}_{i-1} + \alpha x_i$, где 0<α<2.

Дисперсия ВР представляет собой МО квадрата отклонения x(t) от его среднего.

$$\tilde{D}_x^2 = \tilde{\sigma}_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{1}{N-1}\left[\sum_{i=1}^{N}(x_i^2) - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}\right], \qquad (7.4)$$

вторая из формул не требует предварительного определения x_i. Формула (7.4) – выборочная несмещенная дисперсия, если N-1 заменить на N получаем формулу для смещенной дисперсии, при больших N это одно и то же. Иногда дисперсию выражают в процентах. $\sigma_x = D_x = \sqrt{\sigma_x^2}$ - среднеквадратичная погрешность (формула Бесселя).

В АСУТП в режиме on line дисперсия может быть вычислена по рекуррентном формулам:

$$\bar{x}_i = (1-\alpha)\bar{x}_{i-1} + \alpha x_i \quad \bar{x}_i^2 = (1-\alpha)\bar{x}_i^2 + \alpha x_i^2 \quad \sigma_i^2 = \bar{x}_i^2 - (\bar{x}_i)^2 \qquad (7.5)$$

Автокорреляционная функция (АКФ) (см. рисунок 7.2а)

АКФ указывает степень связи последующих значений x(t) с предыдущими; определяет вероятность того, что функция x(t), имеющая в момент t значение x1 , будет в момент t+τ значение x2.

$$\tilde{R}_{xx} = \frac{1}{N-l}\int_0^{T-\tau} \dot{x}(t)\cdot\dot{x}(t+\tau) = \frac{1}{N-l}\sum_{i=1}^{N-l}\dot{x}_i\cdot\dot{x}_{i+l}\,, \qquad (7.6)$$

где $\dot{x}(t) = x(t) - m_x$, $\dot{x}_i = x_i - \bar{x}$ - центрированный ВР; l – смещение (лаг) на l шагов. В реальных процессах вероятность связи между xi и xi+l уменьшается, и Rxx имеет вид, показанный на рисунке 2.1. АКФ является четной функцией, чем круче опускается ее график, тем слабее связь между x(t) и x(t+τ) и наоборот; у белого шума АКФ δ-функция; Rxx = σ_x^2, при l=0.

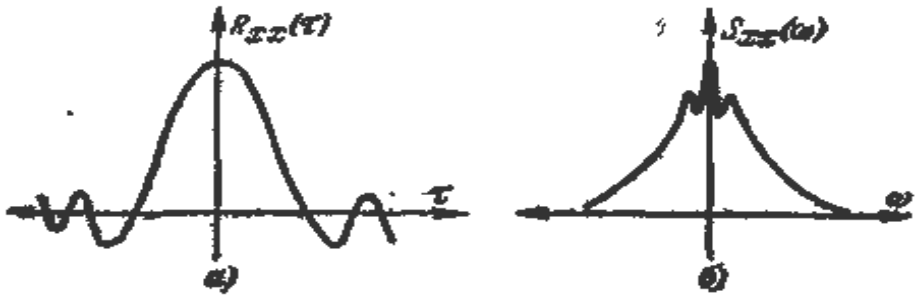


Рисунок.7.2- Графики автокорреляционной функции (а) и спектральной плотности (б).

Взаимная корреляционная функция (ВКФ)

ВКФ указывает степень связи значений входа x(t) в момент времени t со значениями выхода y(t-τ) в момент времени t-τ.

$$\tilde{R}_{xy} = \frac{1}{N-l}\int_0^{T-\tau} \dot{x}(t)\cdot\dot{y}(t+\tau) = \frac{1}{N-l}\sum_{i=1}^{N-l}\dot{x}_i\cdot\dot{y}_{i+l}\,, \qquad (7.7)$$

$R_{xy}(\tau)$ функция нечетная, график ее подобен графику АКФ, но смещен вправо на время запаздывания в объекте.

Спектральная плотность (СП)

Представляет собой преобразование Фурье от АКФ, может быть получена из АКФ или математической обработкой реализации СП, показывает, какая доля мощности СП приходится на данную частоту, позволяет судить о частотных свойствах объекта. Типичный график показан на рисунке 7.2.б. Часто $S_{xx}(\omega)$ получают аналитически, предварительно аппроксимировав R_{xx} каким-либо выражением.

$$S_{xx}(\omega) = 2\int_0^{\infty} R_{xx}(\tau)\cdot\cos(\omega\tau)d\tau = 2\int_0^{\infty} R_{xx}(\tau)\cdot e^{-j\omega\tau}d\tau \qquad (7.8)$$

$$R_{xx}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{xx}(\omega)\cdot\cos(\omega\tau)d\tau = \frac{1}{\pi}\int_0^{\infty} S_{xx}(\omega)\cdot e^{-j\omega\tau}d\tau \qquad (7.9)$$

Аналогично записываются выражения и для взаимной спектральной плотности $S_{xy}(\omega)$.

При идентификации используют выражения:

$$S_{xy}(j\omega) = W(j\omega)\cdot S_{xx}(j\omega) \qquad (7.10)$$

$$S_{yy}(j\omega) = S_{xx}(j\omega)\cdot|W(j\omega|^2 \qquad (7.11)$$

Если у процесса $S_{xx}(\omega)$ имеет одно и то же значение во всем диапазоне частот от -∞ до ∞ то это идеализированный СП, называемый белым шумом (БШ), у которого любое последующее значение x(t) не связано с предыдущими x(t). Для БШ $R_{xx}(\tau) = S_{xx}\cdot\delta(\tau)$, где $\delta(\tau)$- дельта-функция. Часто вместо БШ используют СП без постоянной и без периодической составляющей со статистическими характеристиками:

$$R_{xx}(\tau) = D_x e^{-\mu|\tau|} \text{ и } S_{xx}(\omega) = \frac{2\mu D}{\mu^2+\omega^2}. \qquad (7.12)$$

Получение графиков АКФ и ВКФ не представляют особого труда. Однако для расчетов при идентификации обычно используют аналитические выражения. Помимо использования аппроксимации полиномами Чебышева, на практике применяют аппроксимация АКФ показательными и другими функциями.

1). $R_{xx}(\tau) = D_x e^{-\alpha|\tau|}$ (7.13) (см. рисунке 7.2), где α = const, параметр затухания и D_x- дисперсия, получают из аппроксимацией графика. Аналитическое выражение для спектральной плотности получаем в виде:

$$S_{xx}(\omega) = \frac{2D_x}{\pi}\cdot\frac{\alpha}{\alpha^2+\omega^2} \qquad . (7.14).$$

В зависимости от вида графика корреляционной функции R_{uu} *(ф*)(Рисунок 4.10), ее аппроксимируют одним из следующих выражений.

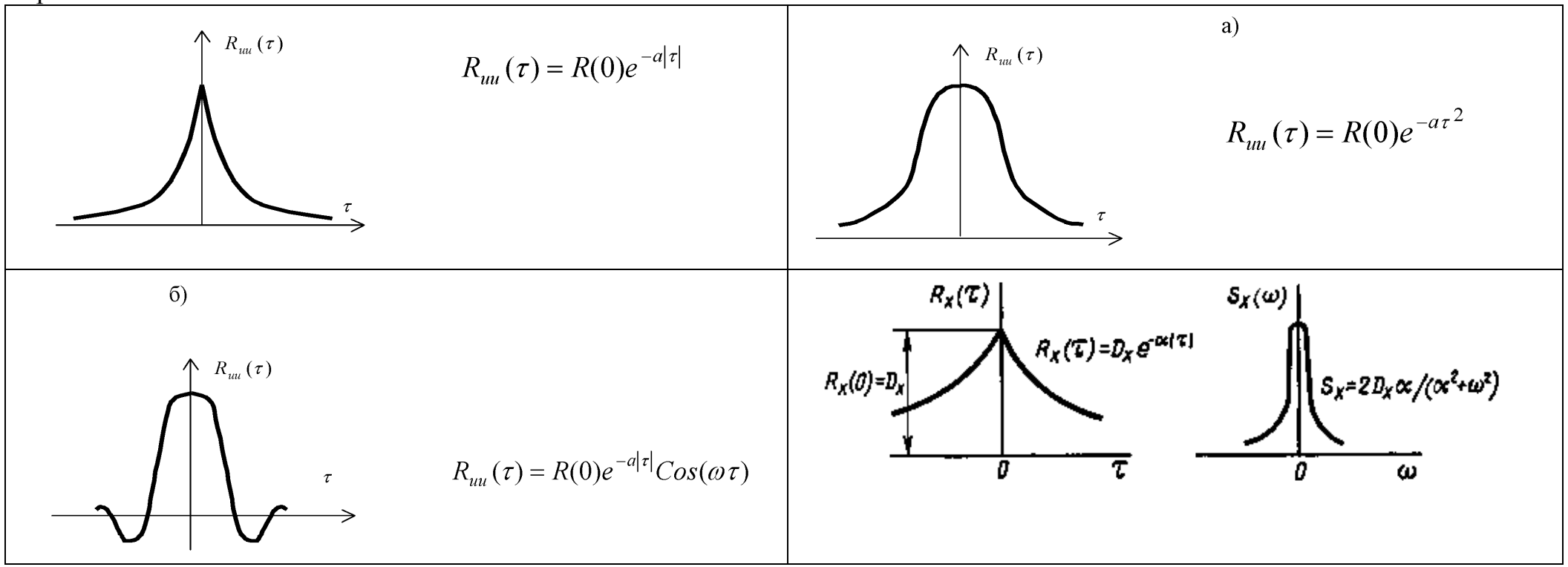


Рисунок 7.3 – аппроксимация АФХ и СП

2). $R_{xx}(\tau) = D_x e^{-\alpha|\tau|}\cos(\omega_0\tau)$ (7.15)

(см. рисунке2.2), где ω0 = const - резонансная частота, откуда:

$$S_{xx}(\omega) = \frac{2D_x}{\pi}\cdot\frac{\beta^2+\omega^2}{\beta^4+2(\alpha^2-\omega_0^2)\omega^2+\omega^4}, \beta = \alpha^2+\omega_0^2 \qquad (7.16)$$

3). $R_{xx}(\tau) = D_x e^{-\alpha|\tau|}(\cos(\omega_0\tau)+\gamma\cdot\sin(\omega_0|\tau|)$ (7.17)

Доказывается, что любая АКФ может быть с любой степенью точности описана линейной комбинацией уравнений (7.15, 2.16, 2.17).

Общие задачи статистической идентификации. В практике моделирования систем наиболее часто приходится иметь дело с объектами, которые в процессе своего функционирования содержат элементы стохастичности или подвергаются стохастическим воздействиям внешней среды. Поэтому основным методом получения результатов с помощью имитационных моделей таких стохастических систем является метод статистического моделирования на ЭВМ, использующий в качестве теоретической базы предельные теоремы теории вероятностей. Возможность получения пользователем модели результатов статистического моделирования сложных систем в условиях ограниченности машинных ресурсов существенно зависит от эффективности процедур генерации псевдослучайных последовательностей на ЭВМ, положенных в основу имитации воздействий на элементы моделируемой системы.

Общая характеристика метода статистического моделирования. На этапе исследования и проектирования систем при построении и реализации машинных моделей (аналитических и имитационных) широко используется метод статистических испытаний (Монте-Карло), который базируется на использовании случайных чисел, т. е. возможных значений некоторой случайной величины с заданным распределением вероятностей. Статистическое моделирование представляет собой метод получения с помощью ЭВМ статистических данных о процессах, происходящих в моделируемой системе. Для получения представляющих интерес оценок характеристик моделируемой системы S с учетом воздействий внешней среды E статистические данные обрабатываются и классифицируются с использованием методов математической статистики .

Сущность метода статистического моделирования. Таким образом, сущность метода статистического моделирования сводится к построению для процесса функционирования исследуемой системы S некоторого моделирующего алгоритма, имитирующего поведение и взаимодействие элементов системы с учетом случайных входных воздействий и воздействий внешней среды E, и реализации этого алгоритма с использованием программно-технических средств ЭВМ.

Различают две области применения метода статистического моделирования:

1) для изучения стохастических систем;

2) для решения детерминированных задач.

Основной идеей, которая используется для решения детерминированных задач методом статистического моделирования, является замена детерминированной задачи эквивалентной схемой некоторой стохастической системы, выходные характеристики последней совпадают с результатом решения детерминированной задачи. Естественно, что при такой замене вместо точного решения задачи получается приближенное решение и погрешность уменьшается с увеличением числа испытаний (реализации моделирующего алгоритма) N.

В результате статистического моделирования системы S получается серия частных значений искомых величин или функций, статистическая обработка которых позволяет получить сведения о поведении реального объекта или процесса в произвольные моменты времени. Если количество реализации N достаточно велико, то полученные результаты моделирования системы приобретают статистическую устойчивость и с достаточной точностью могут быть приняты в качестве оценок искомых характеристик процесса функционирования системы S.

Теоретической основой метода статистического моделирования систем на ЭВМ являются предельные теоремы теории вероятностей. Множества случайных явлений (событий, величин) подчиняются определенным закономерностям, позволяющим не только прогнозировать их поведение, но и количественно оценить некоторые средние их характеристики, проявляющие определенную устойчивость. Характерные закономерности наблюдаются также в распределениях случайных величин, которые образуются при сложении множества воздействий. Выражением этих закономерностей и устойчивости средних показателей являются так называемые предельные теоремы теории вероятностей, часть из которых приводится ниже в пригодной для практического использования при статистическом моделировании формулировке. Принципиальное значение предельных теорем состоит в том, что они гарантируют высокое качество статистических оценок при весьма большом числе испытаний (реализации) N. Практически приемлемые при статистическом моделировании количественные оценки характеристик систем часто могут быть получены уже при сравнительно небольших (при использовании ЭВМ) N.

Идентификация объектов управления методом корреляционного анализа.

Метод корреляционного анализа используется для идентификации объектов управления в том случае, если входные и выходные сигналы являются случайными величинами. Рассмотрим пример определения АФЧХ на основе АКФ и ВКФ.



Рисунок 7.4 - Схема исследования объекта корреляционным методом

При корреляционном анализе используются описанные нами выше:

- автокорреляционная функция (АКФ) и
- взаимокорреляционная функция (ВКФ).

Напомним, что АКФ характеризует зависимость последующих значений случайной величины от предыдущих, находящихся на расстоянии $\Delta\tau$.

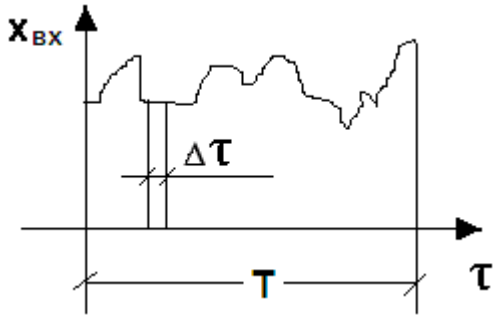


Рисунок 7.5 -. График изменения входной случайной величины – входного сигнала.

АКФ: $R_{x_{вх}}(\tau) = \frac{1}{T}\int_0^T x_{вх}(\tau + \Delta\tau) \cdot x_{вх}(\tau) \cdot d\tau$. При $\Delta\tau \to 0$ – точнее.

Взаимокорреляционная функция связывает две величины, отстоящие друг от друга на $\Delta\tau$.

ВКФ: $R_{x_{вх};x_{вых}}(\tau) = \frac{1}{T}\int_0^T x_{вых}(\tau + \Delta\tau) \cdot x_{вх}(\tau) \cdot d\tau$.

С АКФ и ВКФ связаны (через преобразование Фурье, когда входной-выходной сигнал раскладывается в ряд Фурье, состоящий из суммы синусоидальных колебаний с различной ω – ряд гармоник) спектральные плотности случайных величин.

$$S_{x_{вх}}(\omega) = \int_{-\infty}^{\infty} R_{x_{вх}}(\tau) \cdot e^{-i\omega\tau} \cdot d\tau$$ – для АКФ

$$S_{x_{вх};x_{вых}}(\omega) = \int_{-\infty}^{\infty} R_{x_{вх};x_{вых}}(\tau) \cdot \Delta\tau \cdot d\tau$$ – для ВКФ.

Физически $S_{x_{вх}}(\omega)$ показывает, какая доля мощности случайной величины приходится на данную частоту.

Через спектральную плотность находим искомую АФЧХ объекта:

$$АФЧХ = \overline{W}(i\omega) = \frac{S_{x_{вх};x_{вых}}(\omega)}{S_{x_{вх}}(\omega)}.$$

Аналитический метод идентификации. Аналитический метод вывода математической модели идентичной (совпадающей) по характеристикам с исследуемым объектом применим тогда, когда физико-химические процессы, происходящие в объекте, хорошо изучены. К таким объектам относятся механические системы, поведение которых в статике и динамике подчиняется законам Ньютона, некоторые химические реакторы с простыми химическими реакциями, протекающими в них. Простым примером такого объекта может служить бак, изображенный на рисунке 7.5.

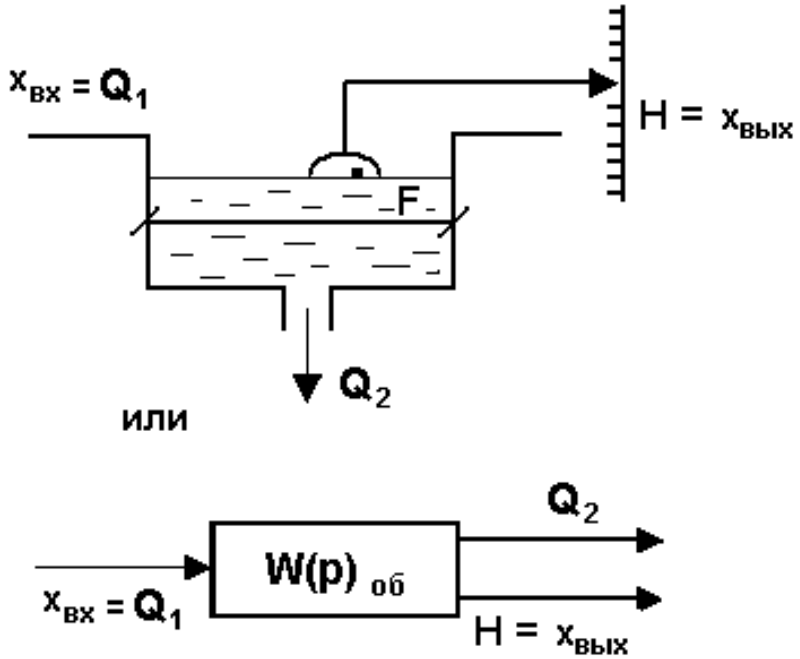


Рисунок 7.5 - Схема исследования объекта управления аналитическим методом.

Опишем явления, протекающие в этом объекте.

Статический режим: $Q_1^0 = Q_2^0$ (приток равен истоку);

Динамический режим: $\begin{cases} Q_1 = Q_1^0 + \Delta Q_1 \\ Q_2 = Q_2^0 + \Delta Q_2 \end{cases}$

$$Q_1 - Q_2 = Q_1^0 + \Delta Q_1 - Q_2^0 - \Delta Q_2 = \Delta Q_1 - \Delta Q_2 =$$
$$= \frac{d(\Delta V)}{dt} = \frac{d(F \cdot \Delta H)}{d\tau} = F \cdot \frac{d(\Delta H)}{d\tau}$$

Из гидравлики: $Q_2 = \alpha \cdot \sqrt{H}$ или для малых $\Delta Q_2 = a \cdot \Delta H$.

Тогда: $\Delta Q_1 - a \cdot \Delta H = F \cdot \frac{d(\Delta H)}{d\tau}$

или, переходя к бесконечно малым приращениям: $F \cdot \frac{dH}{d\tau} + a \cdot \Delta H = Q_1$

или $\frac{F}{a} \cdot \frac{dH}{d\tau} + H = \frac{1}{a} \cdot Q_1$

Обозначив в относительной размерности:

$\frac{H}{H_{ном}} = x_{вых}$; $\frac{Q}{Q_{ном}} = x_{вх}$; $\frac{F}{a} = T_0$; $\frac{1}{a} = k_{об}$

получим:

$T_0 \cdot \frac{dx_{вых}}{d\tau} + x_{вых} = k_{об} \cdot x_{вх}$

Это и есть искомая математическая модель, она носит универсальный характер (в отличии от статистических). При ее использовании требуется, чтобы все неизвестные переменные должны быть заданы.

Основная литература

1. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.
2. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.

Дополнительная литература

3. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.
4. Бенькович Е.С. и др. Практическое моделирование динамических систем. -СПб: БХВ-Петербург, 2002

Тема 5 Методы идентификации на основе простейших тестирующих сигналов

Лекция 8 Определение переходных характеристик

В лекции рассмотрены методы идентификации объектов управления на основе простейших тестирующих сигналов. Эти методы традиционно применяются в инженерной практике. Рассмотрено определение переходных характеристик при нанесении на вход объекта возмущений регулярной формы и методы. аппроксимация временных характеристик.

Методы идентификации с помощью синусоидальных, ступенчатых и импульсных сигналов

Первые реализованные в системах управления методы идентификации были основаны на использовании частотных, ступенчатых и импульсных воздействий. Большинство этих методов ограничивается применением для линейных процессов. Они могут быть также использованы и в линеаризованных системах, если уровни сигналов невелики (см. приложение 1). Эти методы требуют специальных входных сигналов, а именно ступенчатых сигналов для идентификации по переходной функции (ступенчатой переходной функции), импульсных входных сигналов для идентификации по импульсной переходной функции и синусоидальных входных сигналов с различными частотами для определения частотной характеристики. Поскольку вместо входных сигналов, соответствующих нормальному режиму работы, требуются указанные выше специальные сигналы, то очевидно, что эти методы предполагают идентификацию вне процесса управления. Поэтому указанные методы применимы только к линейным стационарным процессам, где отношения вход/выход, полученные для одного типа входных сигналов, сохраняются для всех других типов входных сигналов.

Из трех типов входных сигналов, о которых говорилось выше, ступенчатый входной сигнал является наиболее простым для применения (он соответствует, например, открыванию или закрыванию входного клапана либо включению или выключению входного напряжения), тогда как для подачи синусоидального входного сигнала требуется формирование синусоидальных воздействий и изменение частоты в соответствующем диапазоне. При идентификации по импульсному воздействию часто возникают технические трудности, связанные с формированием и использованием импульсных входных сигналов. Этот метод нельзя применить к линеаризованным системам, так как амплитуда импульса по определению не может быть малой.

Здесь нами рассматривается проблема идентификации математической модели динамики объекта. При использовании методов идентификации на основе простейших тестирующих сигналов на вход исследуемого объекта подается некоторое возмущающее воздействие определенного вида. После этого регистрируется отклик системы на этот сигнал как функция времени. Далее производится математическая обработка выходного сигнала. Суть такой обработки состоит в следующем. Допустим, что исследуемый объект описывается некоторой неизвестной математической моделью. Тогда можно считать, что полученный график переходного процесса представляет решение дифференциального уравнения описывающей объект при известных начальных условиях и при известной математической модели входного сигнала. Это так называемая **обратная** задача, по определению она относится к классу **некорректно** поставленных задач, т.к. имеет бесконечное множество решений. Такие задачи требуют регуляризации, например выбора определенной структуры и вида искомой математической модели исходя из некоторой априорной информации об объекте. На практике часто предполагают, что достаточно точно свойства объекта описываются обыкновенными **линейными** дифференциальными уравнениями с постоянными коэффициентами вида (2.3):

$$a_{n+1}\frac{d^n x}{dt^n}+\ldots+a_2\frac{dx}{dt}+a_1 x=b_{m+1}\frac{d^m u}{dt^m}+\ldots+b_2\frac{du}{dt}+b_1 u\,.$$

или считают, что адекватной математической моделью объекта может служить передаточная функция вида (2.8) или её разностный эквивалент вида (2.25):

$$W(p)=\frac{C}{1+T\cdot p}\cdot\exp(-p\cdot\tau) \qquad (*)$$

или более сложного вида (2.4) (если объект обладает самовыравниванием) или вида (2.9 – 2.11) для объекта без самовыравнивания.

Отметим что, несмотря на простоту выражения (*) оно достаточно точно описывает динамику широкого класса объектов в условиях малых отклонений. Существуют алгоритмы адаптации, которые осуществляют адаптивную т корректировку коэффициентов модели вида (*) или чаще (2.25) в условиях эксплуатации системы управления.

Для линейной системы справедлив принцип суперпозиции, заключающийся в том, что сумме любых возмущений соответствует сумма выходных реакций, каждая из которых определяется соответствующим воздействием; при любом изменении входного возмущения без изменения его формы выходная величина претерпевает такое же изменение, также не изменяя формы.

Принцип суперпозиции дает возможность выразить реакцию системы на любое возмущение через ее реакцию на определенный вид элементарных возмущений. Для этого достаточно представить произвольное возмущение элементарными воздействиями выбранного типа. В качестве типовых возмущений чаще всего применяют единичную скачкообразную функцию, единичную импульсную функцию, единичную линейную функцию, единичное гармоническое колебание, случайный двоичный сигнал и др..

1 Единичная скачкообразная функция описывает мгновенное изменение какого-то воздействия от 0 до 1 (рисунок .8.1).

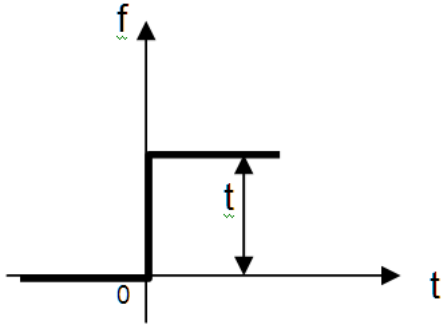


Рисунок .8.1 - . Единичная скачкообразная функция

Аналитически скачкообразную функцию записывают как :

$$f(t)=1(t)=\begin{cases} =0 \quad \text{при } t=-0 \text{ и } t<0 \\ \\ =1 \quad \text{при } t=+0 \text{ и } t>0 \end{cases} \qquad (8.1)$$

Т.е. это ступенчатое единичное воздействие вида:

$$f(t)=1(t)=\begin{cases} 1, t \geq 0 \\ 0, t \leq 0 \end{cases}$$

2 Единичная импульсная функция описывает кратковременное возмущение, имеющее характер кратковременного импульсного толчка (рисунок .8.2).

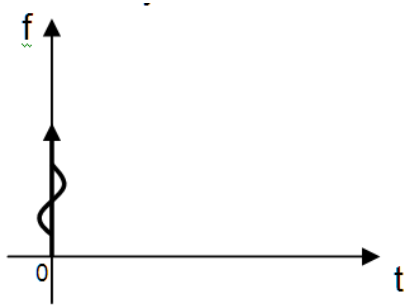


Рисунок .8.2 - . Единичная импульсная функция

Единичная импульсная функция, называемая δ - функцией Дирака (дельта - функцией), представляет собой первую производную от единичной скачкообразной функции:

$$\delta(t) = 1'(t) = \frac{d1(t)}{dt} \qquad (8.2)$$

и равна нулю везде, кроме $t = 0$, где она принимает бесконечное значение, причем при условии, что интеграл от нее по любому интервалу, содержащему $t = 0$, равен единице.

$$\int_{-\infty}^{+\infty} \delta(t)dt = 1 \qquad (8.3)$$

Т.е. $\delta(t) = \begin{cases} \delta(t) = \infty, t = t_0 \\ \delta(t) = 0, t \neq t_0 \end{cases}$.

Функцию, обладающую такими свойствами, можно получить как предел положительного прямоугольного импульса, имеющего единичную площадь, когда длительность этого импульса стремится к нулю (рисунок .8.3).

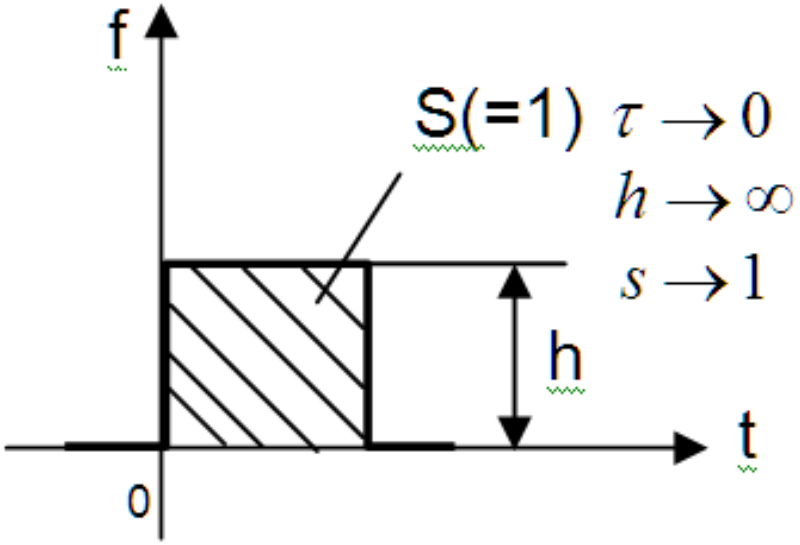


Рисунок .8.3 – К понятию единичная скачкообразная функция

3 Единичную линейную функцию $f(t) = kt$ при $k = 1$ называют еще рамповым возмущением (рисунок .8.4).

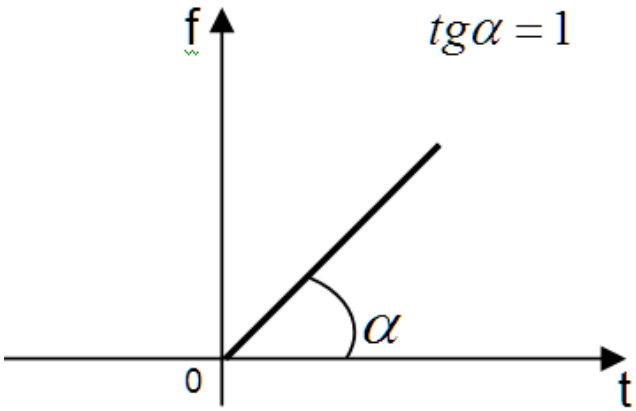


Рисунок .8.4 - . Единичная линейная функция

Такое возмущение является типичным для следящих систем регулирования.

4 Единичное гармоническое возмущение чаще всего записывают как функцию, изменяющуюся по синусоидальному закону (рисунок .8.5)

$$f(t) = \sin \omega t = e^{j\omega t} \qquad (8.5)$$

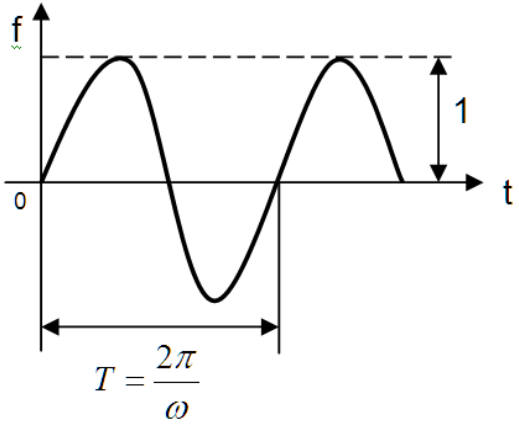


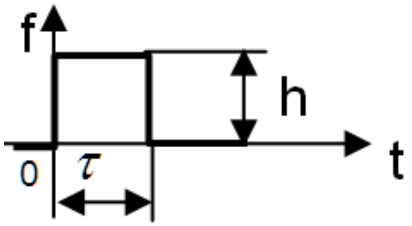
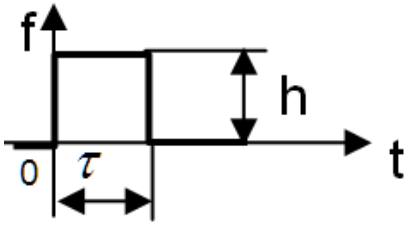
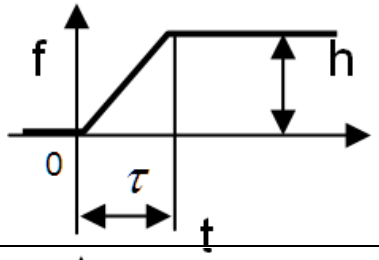
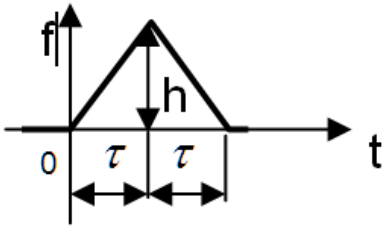
Рисунок .8.5 - Единичное гармоническое возмущение

Такой тип возмущений применяют при частотных методах анализа АСР.

В качестве возмущающих воздействий могут быть использованы и типовые функции, некоторые из которых приведены в таблице 8.1.

Таблица 8.1.

Типовые возмущающие воздействия

№п/п	Возмущение f(t)	График f(t)	Изображение F(p)
1	Ступенчатое при высоте ступени h		$\frac{h}{p}$ (8.6)
2	П-образное при высоте h и при длительности τ		$\frac{h}{p}(1-e^{-\tau p})$ (8.7)
3	Ступенчатое при высоте h и времени нанесения возмущения τ		$\frac{h}{\tau p^2}(1-e^{-\tau p})$ (8.8)
4	Треугольное при высоте h и времени нанесения возмущения 2τ		$\frac{h}{\tau p^2}(1-e^{-2\tau p})^2$ (8.9)

Простейшим входным сигналом, используемым при идентификации, является ступенчатый сигнал. Такой сигнал на входе системы может быть сформирован, например, путем внезапного открывания (или закрывания) входного клапана, включения (или выключения) управляющего напряжения или тока и т. д., так как это почти всегда возможно без применения специальной аппаратуры. У идеального ступенчатого сигнала время нарастания сигнала равно нулю, что физически невозможно, так как при этом скорость нарастания должна быть бесконечно большой. Следовательно, любой реальный ступенчатый входной сигнал является лишь аппроксимацией идеального ступенчатого сигнала. Однако если время нарастания сигнала гораздо меньше пе риода высшей гармоники, то ошибка идентификации становится незначительной. В процессах с помехами или в случаях, когда измерения содержат шум (что обычно имеет место на практике в той или иной степени), необходима соответствующая фильтрация шума.

Идентификация с помощью переходной функции проводится автономно, вне процесса управления, и поэтому применима только к стационарным процессам. Однако поскольку ступенчатые возмущения воздействуют на многие (если не на большинство) системы во время включения или в процессе нормальной работы, то переходные функцщ 'можно записать, не нарушая нормального режима работы системы. В этом заключается дополнительное преимущество рассматриваемого метода. Очевидно, при этом необходимо предположить, что система стационарная, так как результаты идентификации считаются достоверными и после приложения ступенчатого сигнала. Кроме того, предполагается, что в диапазоне амплитуд ступенчатого сигнала система линейна.

Для получения временной динамической характеристики объекта регулирования организуют специальный эксперимент.

В какой-то момент времени к объекту регулирования, находящемуся в установившемся режиме работы, прикладывают некоторое возмущающее воздействие и затем регистрируют отклонения регулируемой величины во времени до ее восстановившегося значения с помощью соответствующего измерительного прибора.

Перед началом эксперимента нужно убедиться в том, что объект регулирования находится в установившемся режиме работы. Для достоверности измеренных отклонений регулируемой величины необходимо один и тот же опыт повторять 2-3 раза.

В найденную путем эксперимента динамическую характеристику объекта включены характеристика собственно объекта и характеристика регистрирующего прибора.

В связи с этим желательно в качестве датчиков отклонений физических величин от их номинальных значений применять те, которые входят в измерительный блок регулирующего устройства, поддерживающего на каком-то значении данную физическую величину.

Для определения динамических свойств объекта на практике чаще всего используют методику снятия переходной характеристики. При определении динамических характеристик объекта по его переходной характеристике (кривой разгона) на вход подается или ступенчатый пробный сигнал или прямоугольный импульс – см.раздел 2.3. Во втором случае переходная характеристика (кривая отклика) должна быть достроена до соответствующей кривой разгона. Процесс получения передаточной функции объекта, исходя из данных о переходном процессе, называется идентификацией объекта.

При снятии переходной характеристики необходимо выполнить ряд условий, представленных в таблице 8.2:

Таблица8.2
Условия снятия переходной характеристики

No	Условия
1	Если проектируется система стабилизации технологического параметра, то переходная характеристика должна сниматься в окрестности рабочей точки процесса.
2	Переходные характеристики необходимо снимать как при положительных, так и отрицательных скачках управляющего сигнала. По виду кривых можно судить о степени асимметрии объекта. При небольшой асимметрии расчет настроек регулятора рекомендуется вести по усредненным значениям параметров передаточных функций. Линейная асимметрия наиболее часто проявляется в тепловых объектах управления.
3	При наличии зашумленного выхода желательно снимать несколько переходных характеристик (кривых разгона) с их последующим наложением друг на друга и получением усредненной кривой.
4	При снятии переходной характеристики необходимо выбирать наиболее стабильные режимы процесса, например, ночные смены, когда действие внешних случайных возмущений маловероятно.
5	При снятии переходной характеристики амплитуда пробного входного сигнала должна быть, с одной стороны, достаточно большой, чтобы четко выделялась переходная характеристика на фоне шумов, а, с другой стороны, она должна быть достаточно малой, чтобы не нарушать нормального хода технологического процесса.

Примечание к таблице 8.2: Начальные условия снятия переходной характеристики. В начальный момент необходимо, чтобы система управления находилась в покое, т.е. регулируемая величина **X** (например, температура в печи) и управляющее воздействие регулятора **Y** (выход регулятора на исполнительный механизм) не изменялись, а внешние возмущения отсутствовали. Например, температура в печи оставалась постоянной и исполнительный механизм не изменяет своего положения. Затем на вход исполнительного механизма подается ступенчатое воздействие, например, включается нагреватель. В результате состояние объекта начинает изменяться.

Определение динамических характеристик объекта управления с самовыравниванием по его переходной характеристике. Наиболее часто при определении динамических свойств объекта регулирования применяют ступенчатое возмущающее воздействие, т.е. экспериментально находят переходную функцию объекта (рисунок .8.6). Самовыравниванием процесса регулирования называется свойство регулируемого объекта после нарушения равновесия между притоком и расходом вернуться к этому состоянию самостоятельно, без участия человека или регулятора. Самовыравнивание способствует более быстрой стабилизации регулируемой величины и, следовательно, облегчает работу регулятора. Рассматривается объект с одним входом и одним выходом со свойствами: стационарности, линейности, сосредоточенности параметров. На вход подается ступенчатое воздействие и на выходе снимается кривая разгона. Необходимо решить **обратную** задачу: по известной кривой разгона определить коэффициенты уравнения.

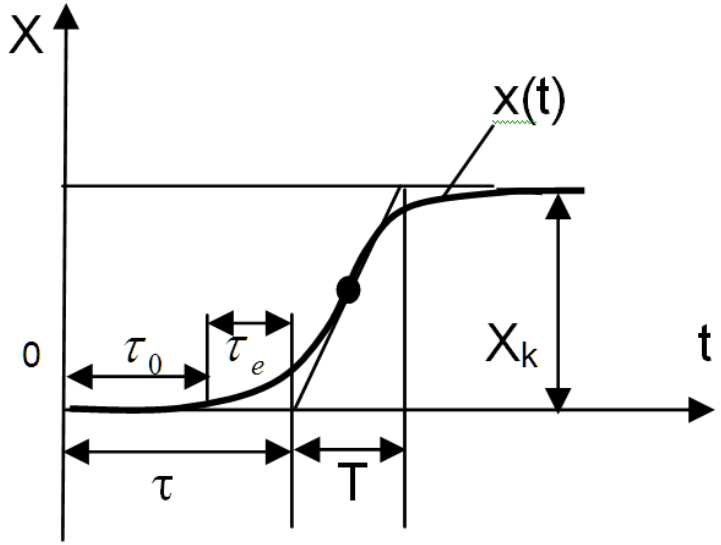


Рисунок .8.6 - . Единичная скачкообразная функция

В то случае, если нас устраивает вид уравнения (*), то его параметры находятся достаточно просто:

- коэффициент усиления - по формуле $C = \frac{X_K}{\Delta U}$, где X_K - установившееся значение выхода а ΔU – возмущение на входе объекта;
- постоянную времени Т – как проекцию касательной к кривой разгона в точке перегиба (см. рисунок .8.6);
- время запаздывания - τ как расстояния от точки 0 до пересечения касательной к кривой разгона в точке перегиба с осью времени (см. рисунок .8.6).

Проверка адекватности. Для решения этой проблемы надо выполнить моделирование функции выхода при передаточной функции объекта вида (*). Т.е. рассматривается объект с одним входом и одним выходом со свойствами: стационарности, линейности, сосредоточенности параметров. На вход подается ступенчатое воздействие и на выходе снимается кривая разгона. Необходимо решить прямую задачу по известной передаточной функции определить график кривой разгона. В данном простом примере достаточно воспользоваться уравнением (2.13)

$$\hat{X}(t) = C \cdot \Delta U \cdot \left[1 - \exp(-\frac{t-\tau}{T})\right] \text{(при} \quad t \geq \tau)$$

$$\hat{X}(t) = 0 \text{ (при} \quad t < \tau)$$

См. пример на следующей странице на рисунке 8.7, решенный с системе Mathcad. Если значение δ, найденное по следующей формуле, не превышает 3-7%, то модель считается адекватной:

$$\delta = \frac{100\%}{N} \cdot \sum_{i=1}^{N} \frac{X_i - \hat{X}_i}{X_i}$$

Если окажется, значение ошибки δ слишком велико, то надо изменить положение касательной на графике (не обязательно, чтобы она проходила через линию перегиба). Если и это не помогает, то надо применить более точный метод получения передаточной функции, например метод площадей (метод Симою).

Идентификация динамических характеристик объектов с самовыравниванием методом площадей [2-6]. Этот метод является одним из инженерных методов идентификации динамических характеристик объектов управления. Этот метод удобен как для реализации на ЭВМ, так и для ручного счета и обладает вполне удовлетворительной для практики точностью.

Метод предполагает идентификацию вне процесса управления, так как использует не результаты измерений при нормальной эксплуатации САУ, а требуется проведения специальных экспериментов по снятию кривой отклика объекта на ступенчатое возмущение на входе объекта. По этой экспериментально полученной кривой разгона определяется коэффициенты передаточной функции вида:

$$W(p) = C * W^* * e^{-p\tau_3} \qquad (8.10)$$

где:

$$W^* = \frac{b_1 + \sum_{i=1}^{M} b_{i+1} p^i}{a_1 + \sum_{i=1}^{N} a_{i+1} p^i} \qquad M \le N \qquad (8.11)$$

Коэффициент усиления объекта с самовыравниванием рассчитывается по формуле:

$$C = \frac{X_K}{\Delta U} \qquad (8.12)$$

Метод применим и для объекта без самовыравнивания .

$\begin{pmatrix} t \\ x \end{pmatrix} :=$

	Исходные данные		
№	t - время		x - выход
1	2,10	88,86	0,00
2	4,80	88,86	0,00
3	7,33	89,17	0,31
4	10,18	89,96	1,10
5	13,22	90,81	1,95
6	16,42	91,57	2,71
7	18,83	92,02	3,16
8	20,86	92,31	3,45
9	23,94	92,68	3,82
10	28,11	93,03	4,17
11	30,10	93,19	4,33
12	32,92	93,27	4,41
13	34,16	93,36	4,50
14	38,20	93,46	4,60
15	42,17	93,53	4,67
16	43,00	93,54	4,68
17	44,00	93,55	4,69
18	45,00	93,56	4,70
19	46,00	93,56	4,70
20	47,00	93,56	4,70

n -количество точек кривой разгона $n := 20$ $i := 1..n$

C - коэффициент усиления объекта $C := \frac{x_n}{10}$ $C = 0.47$

τ - время запаздывания объекта $\tau := 7.5$

T - постоянная времени объекта $T := 10$

Исходная кривая разгона аппроксимируется переходным процессом в апериодическом звене с передаточной функцией:

$$W(p) = \frac{C}{1 + T \cdot p} \cdot \exp(-p \cdot \tau)$$

и переходным процессом:

$$\hat{X}(t) = C \cdot \Delta U \cdot \left[1 - \exp\left(-\frac{t - \tau}{T}\right)\right] \quad \text{для} \quad t \geq \tau \quad \text{или}$$

$$X_i := \text{if}\left[t_i < \tau, 0, C \cdot 10 \cdot \left[1 - \exp\left[\frac{-(t_i - \tau)}{T}\right]\right]\right]$$

Проверка адекватности:

$$\delta := \left(\frac{100}{n}\right) \sum_{i=1}^{n} \frac{(x_i - X_i)}{x_i} \qquad \delta 1 := \left(\frac{100}{n}\right) \sum_{i=1}^{n} \frac{(|x_i - X_i|)}{x_i}$$

$\delta 1 = 6.773$

$\delta = 5.838$

Сплошная линия - исходная кривая разгона, пунктирная линия - результат аппроксимации

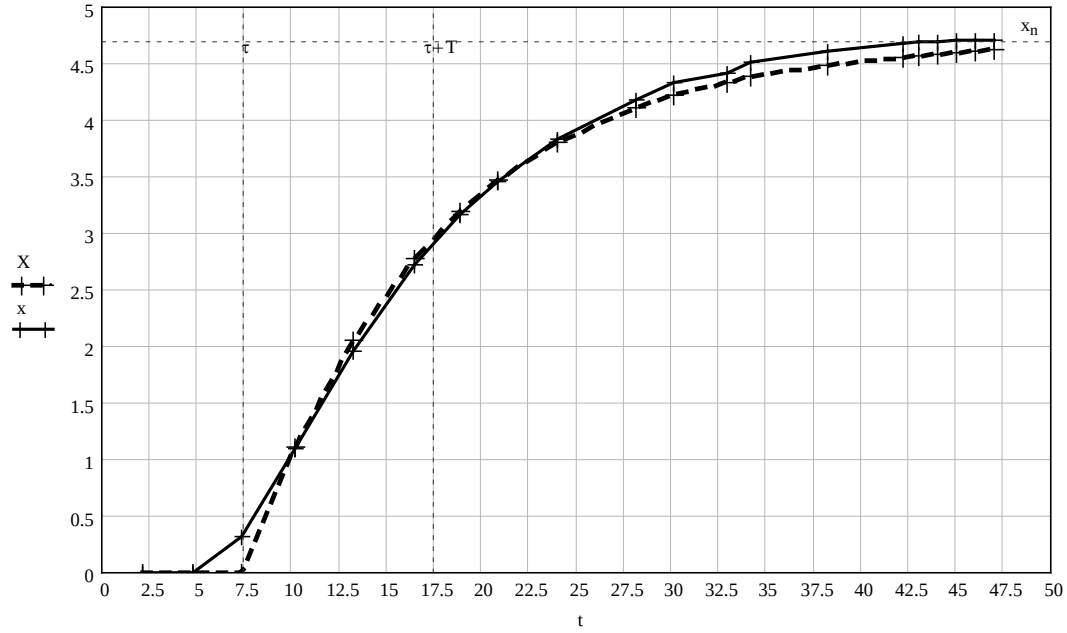


Рисунок .8.7 – Пример проверки адекватности

Время чистого запаздывания τ_3 определяется из графика кривой разгона как время, прошедшее с момента нанесения возмущения до появления реакции на выходе объекта.

Метод площадей позволяет определить коэффициенты ai, bi, M, N, входящие в (8.11).

Рассмотрим применение метода площадей для определения математической модели стационарного линейного объекта с одним входом и одним выходом, имеющего структурную схему вида:



На рисунке 8.8 показана полученная в результате эксперимента кривая разгона. Для обработки используется часть кривой разгона начина с момента времени t=τ_3, представленная дискретными значениями выхода X_i, задаваемыми с шагом по времени Δt. Точка t=τ_3 принимается за новое начало координат.

При расчетах используется кривая разгона в нормированном виде , которая получается из исходной по формуле :

$Z_i = 1-X_i/X_k$ (8.13)

Суть метода площадей сводится к разложению функции $(W^*)^{-1}$ в усеченный ряд по степеням p, то есть к разложению:

$$W^*(p)^{-1} = \frac{1}{W^*(p)} = 1 + \sum_{i=1}^{N} F_i p_i \qquad (8.14)$$

Интегральные площади F_i, входящие в(5) вычисляются по формулам:

$$F1 = \int_0^\infty Z(t)dt \qquad (8.15)$$

$$F2 = \int_0^\infty \int_0^t Z(t)dt^2 = F_1^2 \int_0^\infty Z(t)(1-\theta)d\theta \qquad (8.16)$$

$$F3 = \int_0^\infty \int_0^t \int_0^t Z(t)dt^3 = F_1^3 \int_0^\infty Z(t)(1-2\theta+\frac{\theta^2}{2})d\theta \qquad (8.17)$$

Величины F_i определяются методами численного интегрирования. Например, если воспользоваться методом трапеций, то:

$F_1 = \Delta t\ (S_1-0.5)$ (8.18)

$F_2 = F_1 \Delta t\ (S_2-0.5)$ (8.19)

$F_3 = F_1^{2}\,\Delta t\ (S_3\text{-}0.5)$ (8.20)

.....

$F_l = F_1^{l-1}\,\Delta t\ (S_l\text{-}0.5)$ (8.21)

$$S_1 = \sum_{i=1}^{k} Z_i \qquad (8.22)$$

$$S_2 = \sum_{i=1}^{k} Z_i (1-\theta_i) \qquad (8.23)$$

$$S_3 = \sum_{i=1}^{k} Z_i \left(1-2\theta_i+\frac{\theta_i^{2}}{2}\right) \qquad (8.24)$$

$$S_4 = \sum_{i=1}^{k} Z_i \left(\frac{F_3}{F_1^{3}}-\frac{F_2}{F_1^{2}}\theta_i+\frac{\theta_i^{2}}{2!}-\frac{\theta_i^{3}}{3!}\right) \qquad (8.25)$$

– – – – – – – – – – – – – – – – – – –

$$S_l = \sum_{i=1}^{k} Z_i \left(\frac{-\theta_i^{\,l-1}}{(l-1)!}+\frac{-\theta_i^{\,l-2}}{(l-2)!}+\sum_{j=0}^{l}\frac{F_{l-j-1}(-\theta^{i})}{F_1^{\,l-j-1}\,j!}\right) \qquad (8.26)$$

$\theta_i = t_i/F_1$ (8.27)

Таким образом, имеем значения Z_1, Z_2, Z_k, нетрудно рассчитать F_i.

Порядок передаточной функции N можно определить из условия, что если F_i мало по сравнению с F_{i-1}, или если $F_i < 0$, то N=i-1.

Величина M определяется из условий :

Если X(0)=0, а X’(0)#0, то M=N-1

Если X(0)=X’(0)=0, то M<=N-2

Если X(0)=X’(0)=X’’(0), то b2=b3=b4=...=0

Значения коэффициентов bi и ai находятся решением системы уравнений:

$a_1=1;\ b_1=1$

$a_2=F_1+b_2$

$a_3=F_2+b_3+b_2F_1$

$a_4=F_3+b_4+b_3F_1+b_2F_2$ (8.28)

.....

$$a_l = F_{l-1} + b_l + \sum_{j=2}^{l-1} b_j F_{l-j}$$

В этой системе необходимо подставить нули вместо каждого ai или bi при j>N+1 и j>M+1,а затем решить относительно ai и bi.

При ручном счете обычно ограничиваются вычислением F1,F2,F3 и принимают, что если F3<0, или если X'(0)#0, то M=1, N=2 , то есть передаточная функция W*(p) имеет вид :

$$W^*(p) = \frac{b_1 + b_2 p}{a_1 + a_2 p + a_3 p^2} \qquad \text{…. (8.29)}$$

где:

$b_1=1$; $b_2=-F_3/F_2$; $a_1=1$

$a_2=F_1+b_2$; $a_3=F_2+b_2F$ (21)

а если X'(0)=0 и $F_3>0$, то M=0 ,N=3 , а передаточная функция имеет вид :

$$W^*(p) = \frac{1}{a_1 + a_2 p + a_3 p^2 + a_4 p^3} \qquad (8.30)$$

где :

$b_1=1$; $a_1=1$; $a_2=F_1$; $a_3=F_2$; $a_4=F_3$ (8.31)

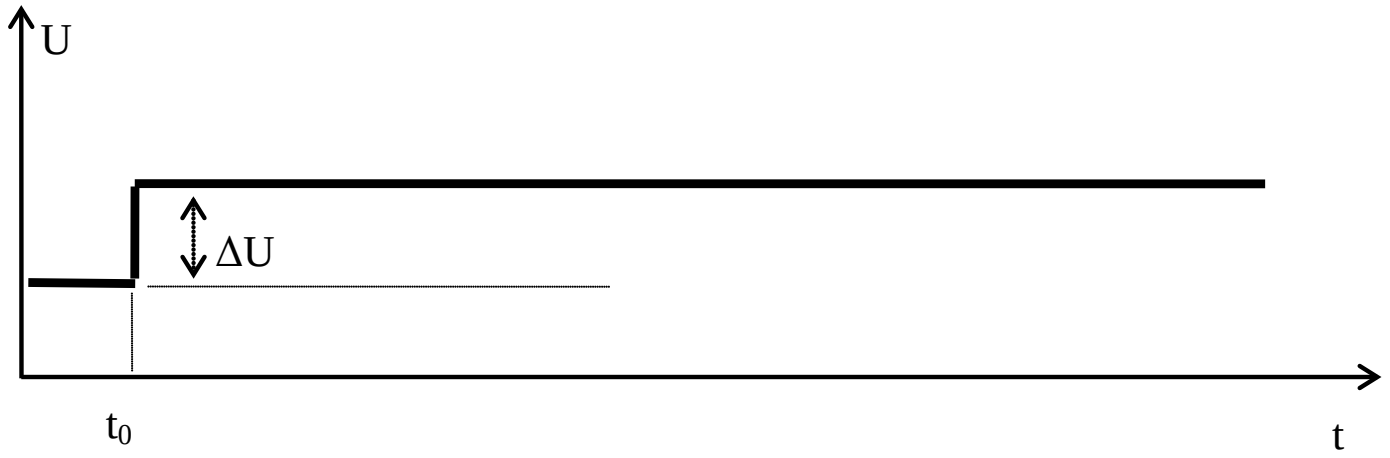


Рисунок. 8.8А - Возмущение на входе объекта

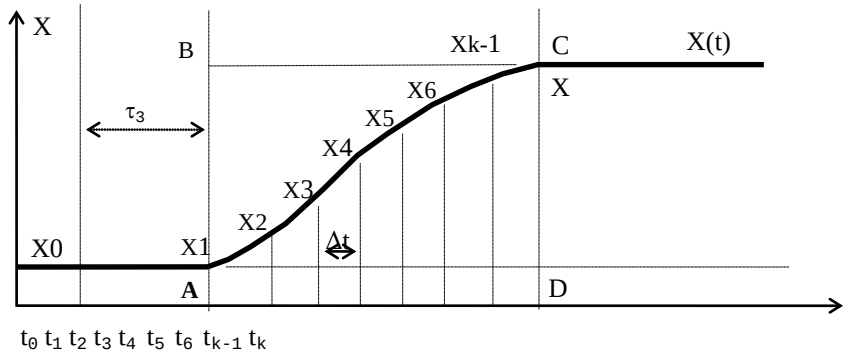


Рисунок 8.8б - Кривая разгона объекта с самовыравниванием

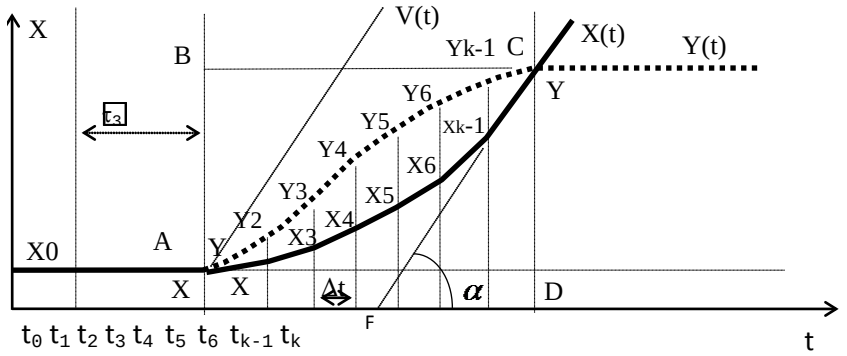


Рисунок 8.8в - Кривая разгона объекта без самовыравнивания

При ручном счете удобно заносить результаты в таблицу вида:

Таблица 8.4.

t_i	X_i	Z_i	θ_i	$1-\theta_i$	$Z_i(1-\theta_i)$	$1-2\theta_i+\theta_i^2/2$	$Z_i(1-2\theta_i+\theta_i^2/2)$
1	2	3	4	5	6	7	8

Эта таблица содержит К строк. Первые два столбца содержат исходные данные - точки времени и значения выхода. Сумма чисел столбца 4 даст значения S_1, сумма столбца 6- S_2, сумма столбца 8 - S_3.

См. также примеры, приведенные в лабораторной работе №5 [3], где рассмотрены и объекты без самовыравнивания.

Идентификация динамического объекта управления по импульсной характеристике. Иногда по технологическим условиям нельзя длительное время держать «единичный скачок» на входе объекта. Тогда подается возмущение типа «единичного импульса», длительность которого достаточна для заметного изменения выходного сигнала. Практически «единичный импульс» рассматривается как два последовательных «единичных скачка», только первый имеет значение (+1), а второй – (-1). Полученная на объекте экспериментальная импульсная характеристика – график изменения во времени выходного сигнала объекта путем несложных графических преобразований достраивается до экспериментальной кривой разгона и далее поиск математической модели – $W_{об}(p)$ идет по указанному выше пути. Перестройка импульсной характеристики объекта до экспериментальной кривой разгона идет так:

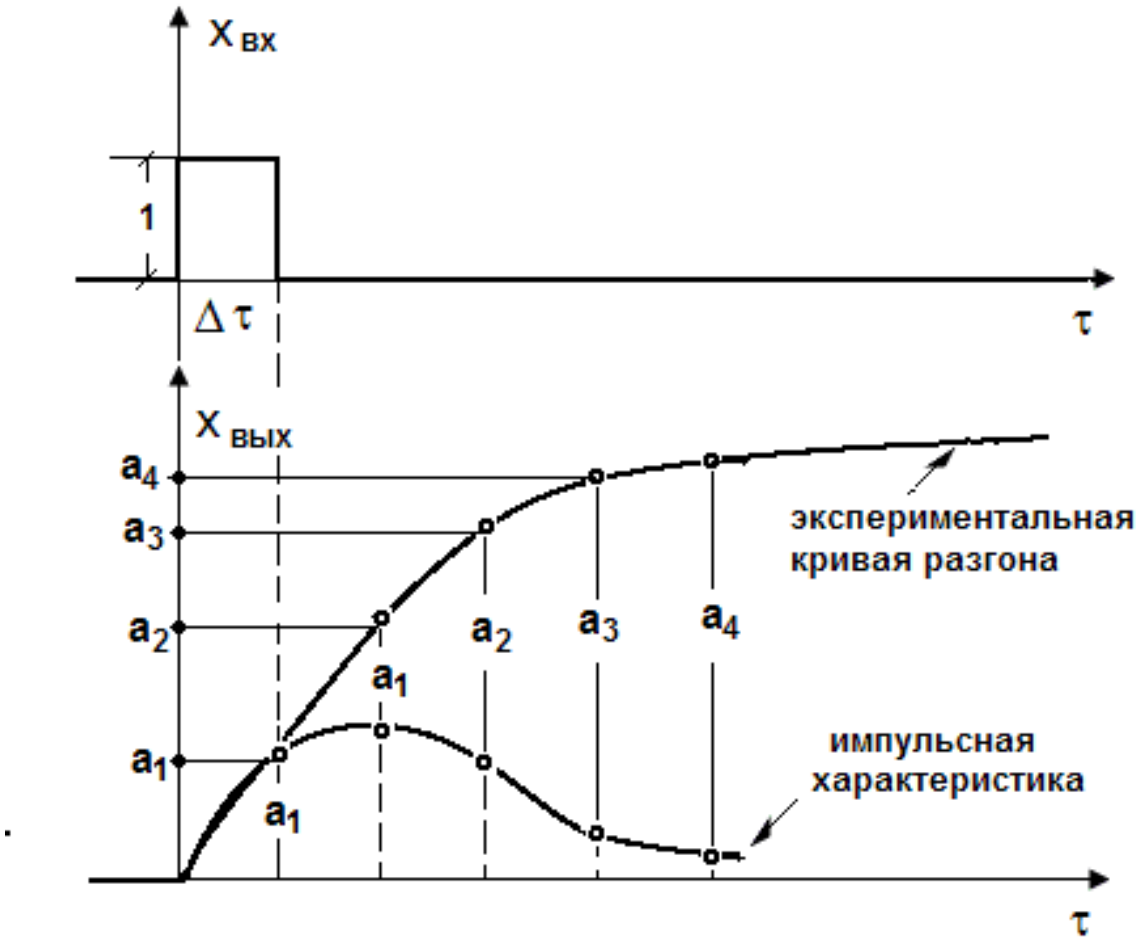


Рисунок .8.6 - Схема преобразования экспериментальной импульсной характеристики в кривую разгона

Основная литература

1. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.
2. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.

Дополнительная литература

3. Инков А.М. Моделирование и идентификация объектов управления. Методические указания к выполнению лабораторных работ для студентов спец. 050702. Шымкент, ЮКГУ, 2010 г., -78 с.

Лекция 9 . Определение частотных характеристик

В лекции рассмотрены методы идентификации объектов управления на основе простейших тестирующих сигналов. Эти методы традиционно применяются в инженерной практике. Рассмотрено определение частотных характеристик при нанесении на вход объекта возмущений периодического характера и аппроксимация экспериментальных частотных характеристик

Частотные характеристики динамического звена. Частотной характеристикой динамического звена называют функцию комплексного аргумента $j\omega$, полученную путем формальной замены p на $j\omega$ в выражении передаточной функции/

Рассмотренные выше формулы (2.19-2.21) описывают связь между разными видами частотных характеристик.

Напомним, что частотную характеристику динамического звена можно определить как отношение спектра (преобразования Фурье) выходного сигнала к спектру входного сигнала.

Знание частотной характеристики звена позволяет определить выходной спектр по входному.

На рисунке 9.1 показаны частотные характеристики некоторого динамического звена.

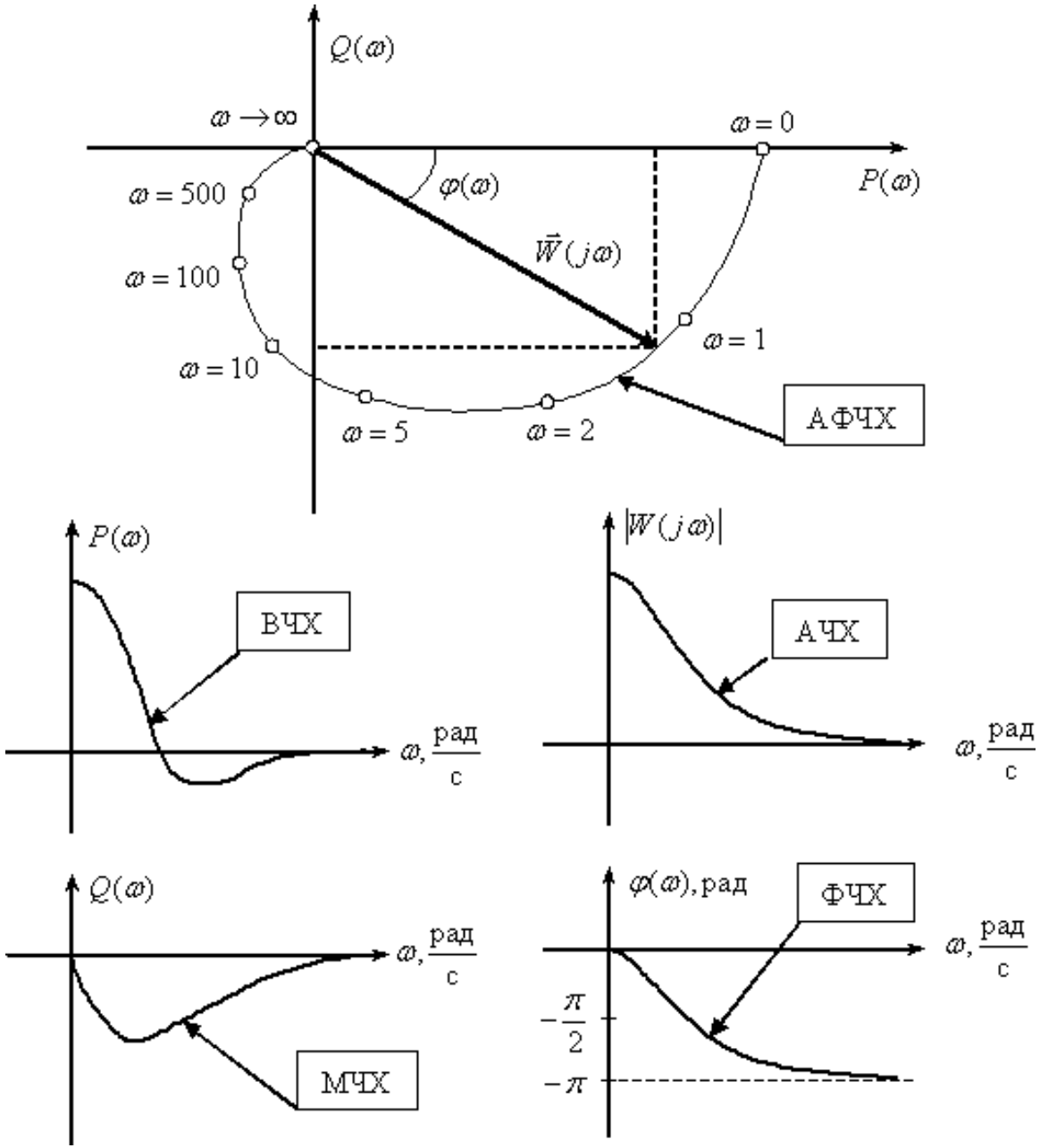


Рисунок 9.1 – Частотные характеристики

При подаче на вход **линейной** системы сигнала:

$$U(t) = A_u \sin(\omega t) = A_u \cdot \exp(j\omega t) \qquad (9.1)$$

на выходе будет сигнал:

$$X(t) = A_x \sin(\omega t + \varphi) = A_x \cdot \exp(j\omega t + \varphi), \qquad (9.2)$$

Отсюда следует простой алгоритм экспериментального определения частотной характеристики линейного динамического звена, объекта или системы управления для конкретной частоты ω_1:

- подать на вход объекта синусоидальный сигнал частоты ω_1 и постоянной амплитуды;
- дождаться затухания свободной составляющей переходного процесса;
- измерить амплитуду выходного сигнала и сдвиг его по фазе относительно входного сигнала.

Отношение амплитуды выходного установившегося сигнала к амплитуде входного сигнала определит модуль частотной характеристики при частоте ω_1.

Сдвиг фазы выходного сигнала относительно входного сигнала определит угол (аргумент) частотной характеристики при частоте ω_1.

Применяя данный алгоритм для частот от нуля до бесконечности, можно экспериментальным путем определить частотную характеристику конкретного устройства.

Функциональная схема экспериментальной установки для снятия частотных характеристик имеет вид, показанный на рисунке 9.2..

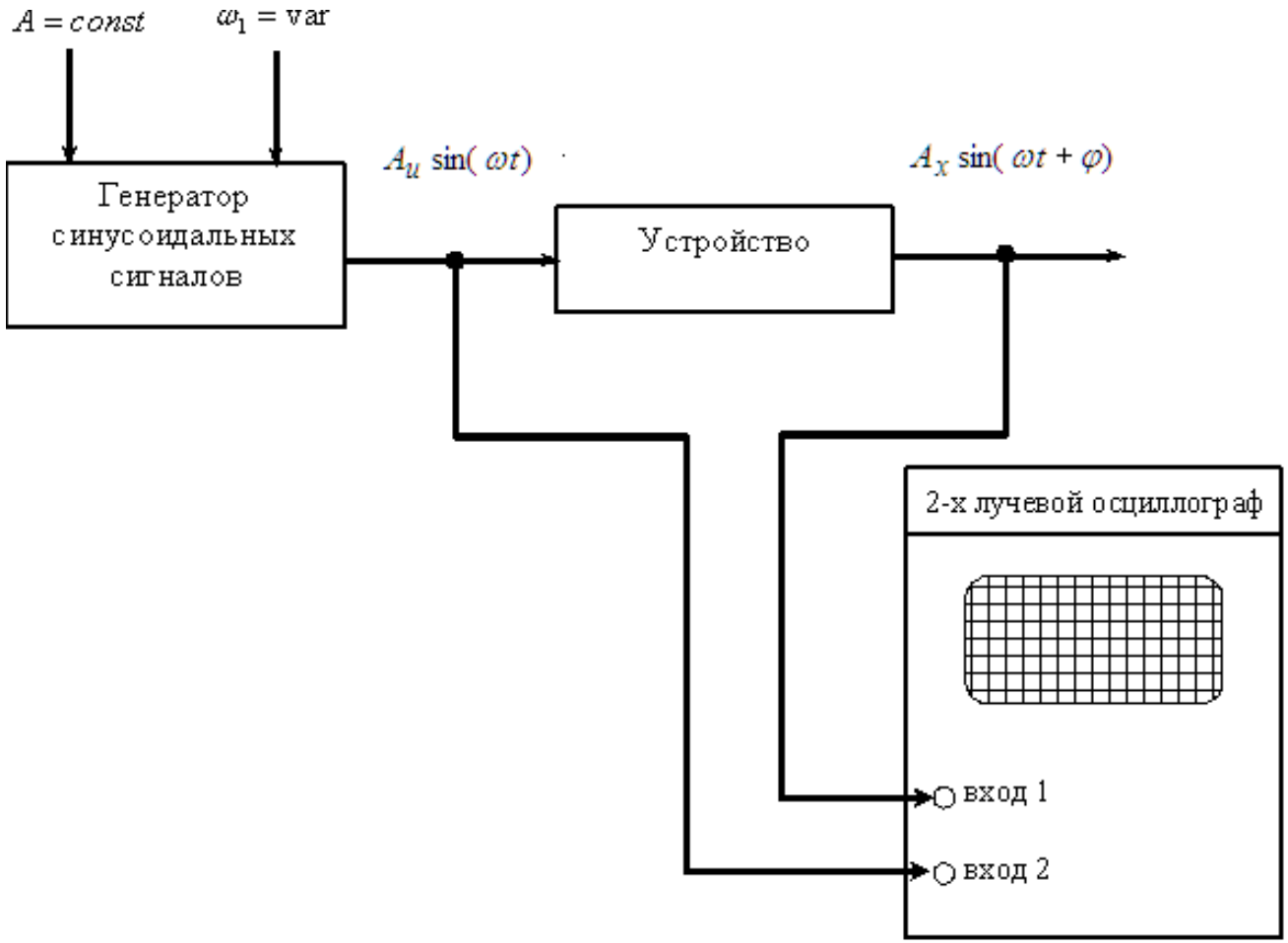
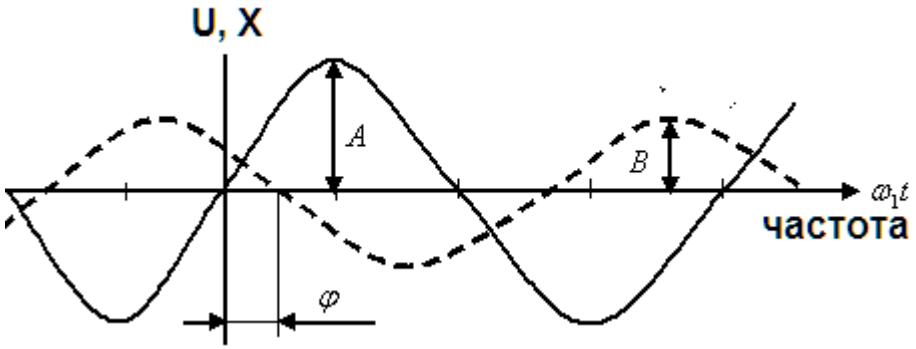


Рисунок 9.2 – Снятие частотных характеристик

При частоте ω_1 на экране осциллографа получаем после затухания свободной составляющей следующую картину, показанную на рисунке 9.3. Отметим, что процедура снятия частотных характеристик - длительная процедура, т.к. в одном опыте – получается только одна точка графика 9.3



A – амплитуда входного сигнала $A_u \sin(\omega t)$;

B - амплитуда выходного сигнала $A_x \sin(\omega t + \varphi)$;

ϕ - сдвиг по фазе

Рисунок 9.3 – Входные и выходные сигналы при снятии частотных характеристик

На основании полученных данных можно построить на комплексной плоскости точку, принадлежащую частотной характеристике устройства, а совокупность точек при изменении частоты от нуля до величины, когда амплитуда выходного установившегося сигнала станет пренебрежимо мала, будет представлять собой амплитудно-фазовую частотную характеристику (АФЧХ). Как видно из рисунка, по этим данным может быть построена любая необходимая частотная характеристика устройства.

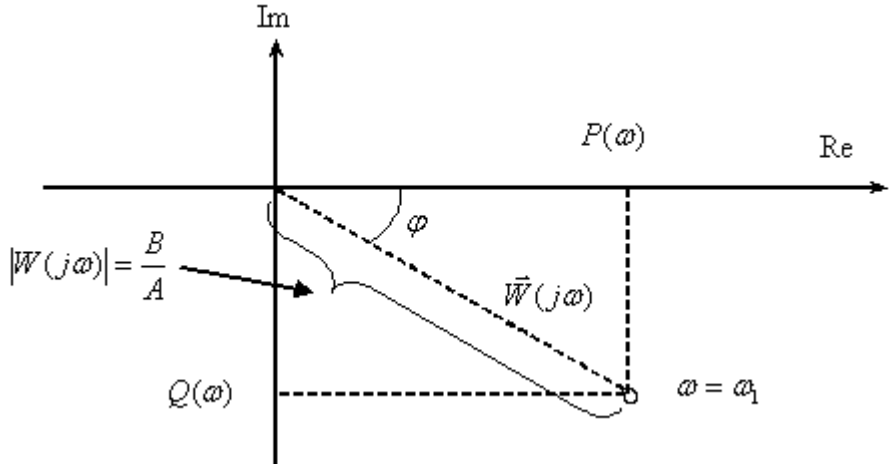


Рисунок 9.4 – График АФЧХ

Для экспериментального получения частотных характеристик различных объектов в инженерной практике используют специализированные приборы, а в последнее время широко используют для таких целей персональные компьютеры, оснащенные специализированными платами ввода-вывода и пакетами прикладных программ.

Учитывая все вышеизложенное, становится ясным и физический смысл частотной характеристики. Она показывает, во сколько раз изменяет динамическое звено (устройство), работающее в установившемся режиме, амплитуду входной синусоиды частоты ω, и на какой угол сдвигает входную синусоиду по фазе.

Определение передаточной функции объекта по частотным характеристикам. Существует ряд достаточно сложных для вычисления методов, и даже ряд вычислительных устройств для обработки частотных характеристик.

Рассмотрим один из методов. Этот метод идентификации относится к непараметрическим методам, т. к. вначале экспериментально снимают частотные характеристики объекта, а затем уже по полученным экспериментальным характеристикам вычисляют передаточную функцию. Основные трудности при проведении экспериментов заключаются в определении рабочего диапазона частот и дрейфе оси колебаний на выходе объекта. Чаще всего область рабочих частот задается ориентировочно и наибольшее внимание уделяется диапазону, в котором сдвиг по фазе между входным и выходным гармоническими сигналами составляет 180^0.

При снятии частотных характеристик используют различные методы воздействия на объект.

Метод синусоидальной волны предполагает подачу на вход объекта гармонических колебаний. На каждой из выбранных в пределах рабочего диапазона частот проводится отдельный опыт. На входе исследуемого объекта возбуждаются колебания выбранной частоты. Процессу колебаний дают установиться и, когда ось колебаний, их форма и амплитуда станут неизменными, измеряют амплитуды входных и выходных колебаний и фазовый сдвиг между ними. Частное от деления амплитуды выходных колебаний на амплитуду входных колебаний дает амплитуду частотной характеристики на взятой частоте, сдвиг по фазе – ординату фазовой частотной характеристики.

Основным затруднение при использовании этого метода является необходимость возбуждения колебаний большой мощности, имеющих правильную синусоидальную форму.

Поэтому чаще применяют метод «прямоугольной» волны. При этом по прибору, измеряющему входную величину объекта, градуируются три положения регулирующего органа, изменяющего эту величину. Среднее положение соответствует значению входной величины, при котором стабилизируется режим работы объекта перед началом испытаний (ось колебаний); два других равноотстоят от среднего положения. Перед началом опыта регулирующий орган устанавливается в среднее положение и удерживается там до установления в объекте стабильного режима. Затем регулирующее устройство через равные промежутки времени, соответствующие полупериоду избранной частоты, переводится из одного крайнего положения в другое и обратно. Эти переключения продолжаются до тех пор, пока на выходе объекта колебания y(t) не примут установившейся формы. После чего производят регистрацию этих колебаний.

На основе полученных осциллограмм колебаний на выходе объекта проводят их гармонический анализ, ограничиваясь вычислением амплитуд и фаз первой и третьей гармоник

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} y(t)Cos\left(k\frac{2\pi}{T}t\right)dt;$$
$$c_k = \frac{1}{\pi}\int_{-\pi}^{\pi} y(t)Sin\left(k\frac{2\pi}{T}t\right)dt;$$
$$a_k = \sqrt{b_k^2 + c_k^2};$$
$$\varphi_k = arctg\left(\frac{b_k}{c_k}\right), \qquad (9.3)$$

где T – период колебаний; к – номер гармоники.

В том случае если значения выходной величины известны только в дискретные, равноотстоящие моменты времени интегралы в (9.3) заменяются суммами

$$b_k = \frac{1}{N}\sum_{i=0}^{N-1} y_i Cos\left(2\pi\frac{ik}{N}\right);$$
$$c_k = \frac{1}{N}\sum_{i=0}^{N-1} y_i Sin\left(2\pi\frac{ik}{N}\right), \qquad (9.4)$$

где N – число дискрет выходного сигнала.

Гармонический анализ входного прямоугольного сигнала приводит к такому выражению

$$u(t) = \frac{4A}{\pi}\left[Sin\left(\frac{2\pi}{T}t\right) + \frac{1}{3}Sin\left(3\frac{2\pi}{T}t\right) + \frac{1}{5}Sin\left(5\frac{2\pi}{T}t\right) + ...\right], \qquad (9.5)$$

где A – амплитуда прямоугольной волны.

Вычислив амплитуды и фазы входных и выходных гармонических составляющих можно вычислить значения амплитудно-частотной характеристики на выбранной частоте, как отношение амплитуд гармонических составляющих на выходе и входе объекта и значения фазо - частотной характеристики, как соответствующий фазовый сдвиг $\phi_к$.

Для повышения точности определения частотных характеристик рекомендуется использовать при гармоническом анализе только первую гармонику.

Для того чтобы не допускать больших отклонений регулируемой величины в процессе проведения опытов переключения регулирующего органа осуществляют в моменты времени, когда регулируемая (выходная) величина достигает заранее выбранные граничные значения. Для реализации такого режима используют двух позиционный регулятор с регулируемой зоной нечувствительности. Изменяя величину зоны нечувствительности и знак обратной связи можно изменять частоту возникающих в такой системе автоколебаний.

Эксперимент обычно начинают на частоте, при которой фазовый сдвиг между входным и выходным сигналами равен π . Это достигается при нулевой зоне нечувствительности позиционного регулятора и отрицательной обратной связи. Увеличивая зону нечувствительности при отрицательной обратной связи, уменьшают частоту автоколебаний. Увеличивая зону нечувствительности при положительной обратной связи, увеличивают частоту автоколебаний.

Результаты экспериментов обрабатываются так же, как и в методе прямоугольной волны.

Основным недостатком рассмотренных методов является длительное время эксперимента, затрачиваемое в основном на ожидание установившегося режима колебаний и получения достаточного для аппроксимации частотных характеристик значений их ординат. Для ускорения экспериментов иногда на вход объекта подается сумма гармонических составляющих разных частот. Установившиеся колебания выходной величины также подвергают гармоническому анализу и сразу находят несколько ординат частотной характеристики. Однако в этом случае требуется специальный источник полигармонического воздействия и линейность объекта.

Определение аналитического выражения передаточной функции по частотным характеристикам может быть осуществлено несколькими способами. Рассмотрим один из них позволяющий аналитически вычислить выражение для передаточной функции по дискретным ординатам вещественной частотной характеристики.

Будем искать выражение для передаточной функции в виде ряда

$$W(p)=\sum_{k=0}^{\infty}A_k\left(\frac{1-p}{1+p}\right)^k. \qquad (9.6)$$

Тогда делая подстановку

$$tg\left(\frac{\varphi}{2}\right)=\omega, \qquad (9.7)$$

выражение для вещественной частотной характеристики можно записать в виде

$$U(\varphi)=\sum_{k=0}^{\infty}A_k Cos(k\varphi). \qquad (9.8)$$

Если теперь провести гармонический анализ экспериментально полученной вещественной частотной характеристики, то в результате будут получены неизвестные коэффициенты передаточной функции $A_к$.

Ограничиваясь значащими членами ряда (9.8) можно записать выражение передаточной функции объекта в виде (9.6).

Ограничение на использование данного метода вычисления передаточной функции относятся к самовыравнивающимся, устойчивым и неминимально - фазовым объектам. В том случае если объект не отвечает этим требованиям, то при вычислении вещественной частотной характеристики необходимо учесть влияние интегрирующих звеньев и звеньев запаздывания на вид этой характеристики и проводить ее расчет с учетом этого влияния.

Компенсация запаздывания осуществляется выражением

$$U_з(\omega) = U(\omega)Cos\omega\tau - V(\omega)Sin\omega\tau, \qquad (9.9)$$

а компенсация интегрирующих звеньев –

$$U_и(\omega) = \frac{V(\omega)}{\omega}, \qquad (9.10)$$

где U (ω и)V (ω -) экспериментальные вещественная и мнимая частотные характеристики.

Пример. Проведем идентификацию объекта с передаточной функцией

$$W(p) = \frac{2,5}{(p+0,1)(p^2+6p+25)}. \qquad (9.11)$$

Идентификация проводилась программой MATLAB, приведенной ниже.

```
k=2.5;p1=-.1;p2=-3+4*i;p3=-3-4*i;
p=[p1 p2 p3];
wo=zpk([],p,k);
%wo=tf(1,[1 2 1]);
f=0:180/30:180;
w=tan(pi*f/360);
H=freqresp(wo,w);% Вычисление амплидудно - частотной характеристики объекта
H=squeeze(H);
```

```
U=real(H);% Вычисление вещественной характеристики объекта
n=length(U);
u=[U(n:-1:2);U];
ab=fft(u)/(n-1); % Фурье- преобразование вещественной характеристики
f=angle(ab);
a=abs(ab);
a(1)=a(1)/2;
w0=tf([-1 1],[1 1]);
ws=a(1);nun=a(1);
pp=[1 1];pm=[-1 1];
den=1;d=1;
% Вычисление передаточной функции объекта по (9.6)
for j=2:(n+1)/2
 den=conv(den,pp);
 d=conv(d,pm);
 nun=conv(nun,pp)+a(j)*d;
 ww=tf(nun,den)
end
ww=minreal(ww)% Получение минимальной реализации передаточной функции
step(ww,wo)
pause
bode (ww,wo)
pause
[wb,g]=balreal(ww); )% Получение сбалансированной реализации передаточной функции
wm=modred(wb,[4:(n-1)/2]);
step(wm,ww,wo)
ww=zpk(tf(wm))
```

На рисунке 9.5 показана полученная вещественная частотная характеристика объекта для относительных значений частоты ф =2a rctg(ω)

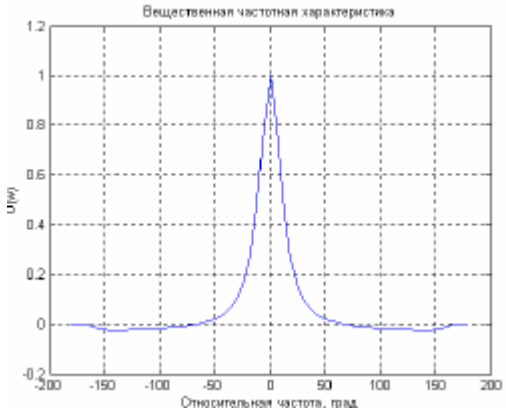


Рисунок 9.5 – Полученная вещественная частотная характеристика объекта

На рисунке 9.6 приведены результаты гармонического анализа вещественной частотной характеристики (рисунок 9.6), позволившие найти амплитуды гармоник Ак.

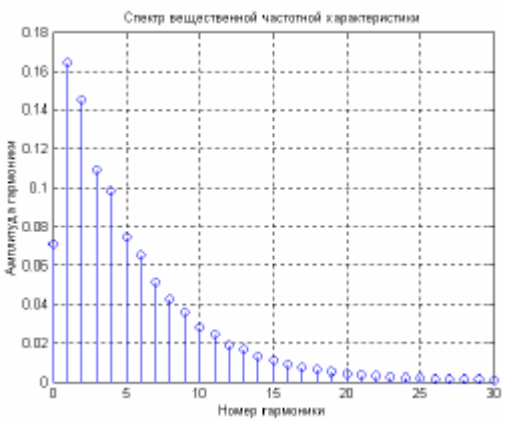


Рисунок 9.5 – Результаты гармонического анализа полученной вещественной частотной характеристики

Расчет передаточной функции по формуле (9.6) с удержанием трех составляющих сбалансированной реализации модели (9.6) дает следующее выражение передаточной функции

$$W(p) = \frac{-0.038723\,(s - 2.923)\,(s^2 + 0.2386s + 0.04178)}{(s + 0.1553)\,(s^2 + 0.2452s + 0.03147)} \qquad (9.12)$$

Переходные характеристики исходного объекта и его модели, полученной в результате идентификации, изображены на рисунке 9.6. Следует отметить значительные погрешности идентификации рассмотренного метода, особенно при определении коэффициента передачи объекта.

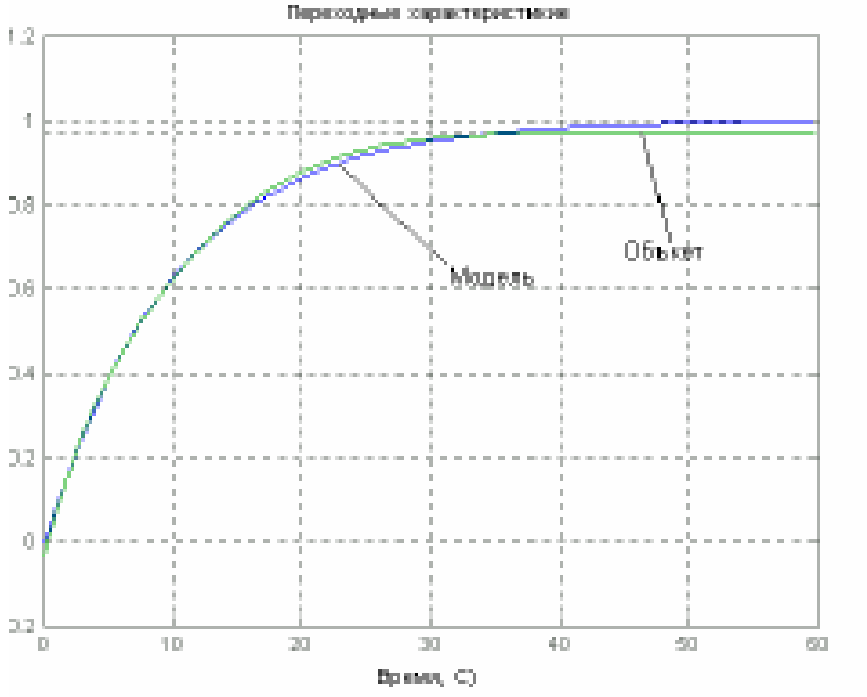


Рисунок 9.3 – Переходные характеристики исходного объекта и его модели

Основная литература

1. Современные методы идентификации систем. Под ред. П. Эйкгоффа, М.: Мир, 1983
2. Ордынцев В. М. Математическое описание объектов автоматизации. – М: Машиностроение, 1965. – 360 с.
3. Гроп Д. Методы идентификации систем. – М: Мир, 1979. – 302 с.

Дополнительная литература

4. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.
5. Семенов А. Д., Артамонов Д. В., Брюхачев А. В. Идентификация объектов управления: Учебн. пособие. - Пенза: Изд-во Пенз. гос. ун-та, 2003.- 211 с.
6. Инков А.М. Моделирование и идентификация объектов управления. Методические указания к выполнению лабораторных работ для студентов спец. 050702. Шымкент, ЮКГУ, 2010 г., -78 с.

Тема 6 Общая схема процесса идентификации

Лекция 10 Общая схема процесса идентификации

В лекции рассмотрены основные этапы идентификации; и такие понятия как априорная и апостериорная информация; критерии и показатели качества идентификации. Проведена классификация методов идентификации. Рассмотрены принципы структурной и параметрической идентификация; активной и пассивной идентификации.

Рассмотрим перечень примерных этапов идентификации. Технология математического моделирования технических систем в общем случае предполагает выполнение следующих основных этапов:

1. Формулировка целей. В основе всякой задачи, проблемы моделирования лежит информация о том, какие зависимости нас интересуют, каковы его цели его использования. Именно эта информация определяет объект. Эти представления могут быть очень приближенными, но всегда отражают некоторые его свойства, достаточные для эффективной формулировки целей моделирования. Обычно в задачах моделирования цель достигается путем максимизации или минимизации некоторого критерия, задаваемого в виде целевой функции.

2. Изучение объекта. При этом требуется понять происходящий процесс, определить границы объекта с окружающей его средой, если таковые имеются. Кроме того, на данном этапе определяются перечень всех входных и выходных параметров объекта исследования и их влияние на достижение целей моделирования.

3. Описательное моделирование (разработка концептуальной модели) - установление и словесная фиксация основных связей входных и выходных параметров объекта.

4. Математическое моделирование (разработка математической модели)- перевод описательной модели на формальный математический язык. Цель записывается в виде функции, которую обычно называют целевой. Поведение объекта описывается с помощью соотношений, входные и выходные параметры объекта на данном этапе в зависимости от сложности исследуемой проблемы могут возникать ряд задач чисто математического характера. Такими задачами являются задачи математического программирования, линейной алгебры, задачи дифференциального и интегрального исчисления и многие другие.

5. Выбор (или создание) метода решения задачи. На данном этапе для возникшей математической задачи подберется подходящий метод. При выборе такого метода необходимо будет обратить внимание на сложность метода и потребляемые вычислительные ресурсы. Если подходящего метода по предъявленным критериям не окажется, то требуется разработать новый метод решения задачи. Мы делаем упор на разработку новых эффективных методов, не уступающих известным методам по основным вычислительным характеристикам.

6. Выбор или написание компьютерных программы. На данном этапе выбирается подходящая программа, реализующая выбранный метод решения. Если такая программа отсутствует, то необходимо написать такую программу.

7. Решение задачи на компьютере. Вся необходимая информация для решения задачи вводится в память компьютера вместе с программой. С использованием подходящей программы производится обработка целевой информации и получение результатов решения задач в удобной форме.

8. Проверка адекватности модели (верификация модели).

9. Анализ получаемого решения. Анализ решения бывают двух видов: формальный (математический), когда проверяется соответствие полученного решения построенной математической модели (в случае несоответствия проверяется программа, исходные данные, работа ЭВМ и др.) и содержательный (экономический, технологический и т.п.), когда проверяется соответствие полученного решения тому объекту, который моделировался. В результате такого анализа в модель могут быть внесены изменения или уточнения, после чего весь рассмотренный процесс повторяется. Модель считается построенной и завершенной, если она с достаточной точностью характеризует деятельность объекта по выбранному критерию. Только после этого модель можно использовать при расчетах.

Перечисленные этапы вытекают из общей методологии моделирования систем. При моделировании различных систем трудоемкости одних и тех же этапов могут быть разными. В процессе моделирования конкретной системы могут иметь место некоторые изменения технологии. В частности, может быть заранее предопределен метод моделирования или выбрано конкретное средство моделирования. Математическая модель окажется настолько простой, что не потребуется проведения машинных экспериментов, разработка программной модели исключит необходимость создания математической модели.

Следует обратить внимание на первоочередную необходимость постановки, формулирования цели моделирования. В этом вопросе должно быть достигнуто взаимопонимание между заказчиком, ответственным за создание или модернизацию системы, и разработчиком модели. Важность корректного выполнения этого этапа определяется тем, что все последующие этапы проводятся с ориентацией на определенную цель моделирования.

На этом же этапе конкретизируется, в каких единицах измерения (относительных или абсолютных) должны быть представлены результаты моделирования. Под относительными единицами здесь понимаются качественные градации, сравнительные оценки разных вариантов системы (типа «лучше—хуже», «больше— меньше»). При необходимости представления результатов в абсолютных единицах должен быть решен вопрос о точности измерения. Этот вопрос зачастую не имеет однозначного ответа, но крайне важен для выполнения всех этапов моделирования.

Проверка адекватности указана выше в виде одного из этапов моделирования. Не надо это понимать буквально, так как на адекватность модели оказывает влияние качество выполнения практически всех этапов. Поэтому проверка адекватности должна проводиться в том или ином виде, начиная от разработки концептуальной модели и кончая анализом результатов моделирования.

Под разработкой математической модели подразумевается создание полностью формализованного описания динамики функционирования системы. Однако не для всех систем, внешних условий и целей моделирования может быть подобран известный метод формализации или конструктивный математический аппарат. Тем не менее и для таких систем следует разработать одно-

значные зависимости выходных характеристик от параметров и воздействий для каждой составляющей системы, алгоритмы взаимодействия между составляющими, логические условия изменения состояний.

Результаты машинного моделирования должны быть проанализированы с целью проверки их достоверности и выработки рекомендаций о способах повышения качества исследуемой системы. На всех этапах моделирования следует обратить особое внимание на документирование принимаемых решений, допущений, ограничений и выводов.

Из организационных аспектов моделирования следует выделить необходимость непосредственного участия в работе квалифицированных представителей заказчика на первых этапах (вплоть до разработки математической модели) и на этапе анализа результатов моделирования. Ответственный за систему заказчик должен четко понимать цели моделирования, разработанную концептуальную модель, программу исследований и полученные результаты. Это будет способствовать внедрению выработанных рекомендаций.

Математическое моделирование технических систем предполагает применение следующих методов:

1. аналитических;
2. численных;
3. статистических (имитационных);
4. аналитико-имитационных.

Выбор конкретного метода моделирования зависит от многих факторов, в том числе от:

- целей моделирования;
- сложности исследуемой системы;
- сложности модели, определяемой выбранным уровнем ее детализации;
- требований к номенклатуре исследуемых характеристик;
- требований к точности получаемых результатов;
- требований к общности получаемых результатов;
- требований к затратам времени на моделирование;
- требований к материальным затратам;
- наличия специальных технических средств для проведения моделирования;
- квалификации специалиста, проводящего моделирование и т.д.

Результаты сравнительного анализа методов моделирования, выполненного на качественном уровне, представлены в таблице10.1 (фигурными скобками отмечены наилучшие значения каждого показателя).

Таблица 10.1

Результаты сравнительного анализа методов моделирования

Метод моделирования	Сложность метода	Общность резтов	Точность резтов	Затраты времени	Матер. затраты	Задачи синтеза
Аналитический	{+}	{++++}	+	{+}	{+}	{+}
Численный	++	+++	++	++	++	++
Имитационный	+++	+	{++++}	++++	++++	++++
Комбинированный	++++	++	+++	+++	+++	+++

Для решения задач математического моделирования и идентификации необходима исходная информация. Её условно разделяют на два вида - априорную и апостериорную информацию. (**априорная** информация [aprior information] и **апостериорная** информация [aposterior information]). Соответственно, это данные, имевшиеся до проведения какого-либо опыта или другого действия, и сведения, полученные после его выполнения. Обычно приходиться работать в условиях неполноты априорной и текущей информации об объекте и при наличии неконтролируемых возмущений различного происхождения.

В том случае, если объект рассматривается как «черный ящик», то как правило используют только апостериорную информацию.

«Чёрный ящик» — термин, используемый в точных науках (в частности, системотехнике, кибернетике и др.) для обозначения системы, механизм работы которой очень сложен, неизвестен или неважен в рамках данной задачи. Такие системы обычно имеют некий «вход» для ввода информации и «выход» для отображения результатов работы. Состояние выходов обычно функционально зависит от состояния входов.

Использование априорной информации об объекте исследования (если она доступна) во многих случаях облегчает решение проблемы идентификации.

Критерии и показатели качества идентификации. Формирование критерия качества, характеризующего адекватность модели реальному объекту, является одним из основных этапов идентификации.

Переход от исследования объекта к исследованию модели и подтверждение пригодности модели для решения задач моделирования требует оценку качества полученной модели, т.е. проверку адекватности модели и объекта. Никогда нельзя говорить об абсолютной адекватности, при которой модель по всем свойствам соответствует объекту, так как в зависимости от цели исследования могут строиться различные модели объекта. Например, при исследовании влияния размещения пассажиров на центровку самолета моделью человека может служить мешок с песком, для конструктора одежды - манекен, а для медико-биологических исследований - животные. Таким образом, всякая модель имеет характер проекции и отражает отдельные свойства объекта. В связи с этим основное подтверждение модели заключается в том, чтобы убедиться в возможности использования полученной модели для решения той задачи, ради которой эта модель и строилась. Поэтому адекватность предполагает воспроизведение моделью с необходимой полнотой всех свойств объекта, существенных для целей данного исследования.

Количественно степень адекватности модели и объекта можно оценить путем сравнения их выходных сигналов при подаче одинаковых входных воздействий на объект и его модель. Такое сравнение предпочтительное на основе новой информации, отличной от того множества данных, которое использовалось в процессе идентификации объекта.

Структурная схема вычисления оценки ошибки модели статического объекта приведена на рисунке 10.1.

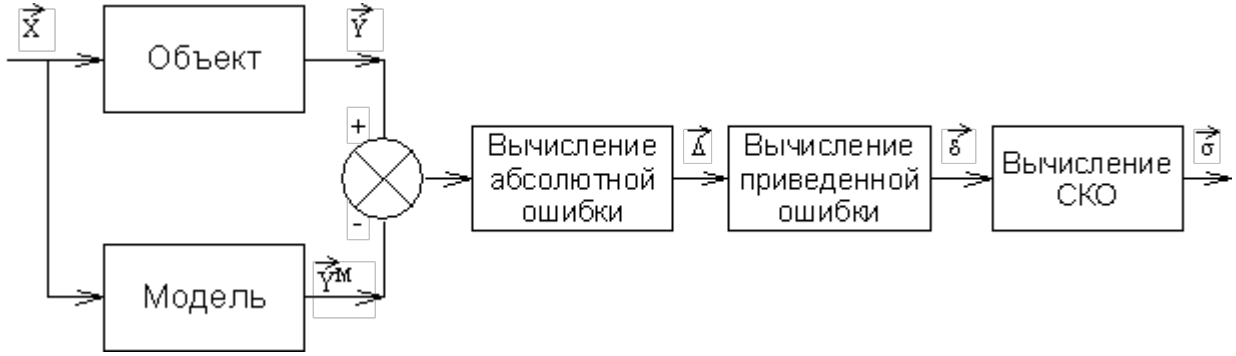


Рисунок 10.1 - Структурная схема вычисления оценки ошибки модели

Пусть проведено *l* опытов при различных уровнях входных воздействий $\vec{x}_j = (x_{1j}, x_{2j}, \ldots, x_{nj})$ из области D_x их допустимых значений и получены реализации выходов объекта $\vec{y}_j = (y_{1j}, y_{2j}, \ldots, y_{mj})$ и выходы модели $\vec{y}_j^{M} = (y_{1j}^{M}, y_{2j}^{M}, \ldots, y_{mj}^{M})$, $(j = \overline{1,l})$.

Ошибки модели $\vec{\Delta} = (\Delta_1, \Delta_2, \dots \Delta_m)$ и $\vec{\delta} = (\delta_1, \delta_2, \dots \delta_m)$ для оценки ее адекватности вычисляются по формулам:

$$\Delta_i = \max_{j=\overline{1,l}} |y_{ij} - y_{ij}^M|, \ (i = \overline{1,m});$$

$$\delta_i = \frac{\Delta_i}{\Delta y_i}, \ (i = \overline{1,m});$$

$$\sigma_i = \sqrt{\frac{1}{l}\sum_{j=1}^{l}(y_{ij} - y_{ij}^M)^2}, \ (i = \overline{1,m}),$$

где $\Delta_i, \delta_i, \sigma_i$ - абсолютная, приведенная и среднеквадратичная ошибки модели по i-му выходу $(i = \overline{1,m})$; y_{ij}, y_{ij}^M - значение i-го выхода объекта и модели в j-ом опыте $(i = \overline{1,m}, j = \overline{1,l})$; Δy_i - максимальное изменение i-го выхода объекта $(i = \overline{1,m})$ при допустимых значениях входов $x_k \ (k = \overline{1,n})$ из области D_X.

Если величины этих ошибок меньше некоторого заданного положительного числа, то модель адекватна объекту.

Теперь рассмотрим оценку модели динамического объекта. Положим, что после идентификации получена модель одномерного объекта в форме линейного дифференциального уравнения вида.

$$\sum_{k=0}^{n} a_k \frac{d^k y_M(t)}{dt^k} = \sum_{k=0}^{m} b_k \frac{d^k x_M(t)}{dt^k}, \qquad (10.1)$$

где $x_M(t)$ - входной сигнал модели;

$y_M(t)$ - выходной сигнал модели;

n,m - наивысшие порядки производных;

$y_M(t)$ и $x_M(t)$ соответственно $(m \le n)$.

Пусть получены реализации входа $x(t)$ и выхода $y(t)$ объекта на отрезке времени $t \in [0,T]$, где T - длина реализации (время наблюдения). Теперь качество модели можно оценить путем сравнения $y_M(t)$ и $y(t)$ либо непосредственно на графике (визуально),либо введя некоторую формальную меру расстояния между этими сигналами.

Выходные сигналы объекта и модели при одном и том же входном сигнале различаются, так как их дифференциальные уравнения и начальные состояния неидентичны. Для оценки адекватности модели и объекта введем критерий их близости по разности выходных сигналов, т.е. реакций на один и тот же входной сигнал $x(t)$, например следующего вида:

$$I_x = \int_0^T F(y(t) - y_M(t))dt \quad , \qquad (10.2)$$

где $F(\bullet)$ - некоторая выпуклая функция.
В частности:

$$F(y(t) - y_M(t)) = (y(t) - y_M(t))^2. \qquad (10.3)$$

В общем случае оценка адекватности проводится для различных форм входного сигнала $x(t)$. Отсюда следует идея необходимости усреднения по входным сигналам и начальным условиям, т.е. введения операции математического ожидания оценки I_x:

$$I = M[I_x] = M[\int_0^T F(y(t) - y_M(t))dt] \quad . \qquad (10.4)$$

Выражение выходного сигнала имеет довольно сложный вид, что затрудняет аналитическое исследование зависимости I от коэффициентов модели, поэтому вводятся и другие критерии. В частности, если уравнение модели имеет вид

$$\sum_{k=0}^{n} a_k \frac{d^k y_M(t)}{dt^k} = x_M(t) \quad , \qquad (10.5)$$

то для оценки близости модели и объекта удобным оказывается функционал от разности входных сигналов $(x_M(t) - x(t))$ модели и объекта, обеспечивающих один и тот же выходной сигнал:

$$I = M[\int_0^T F(x_M(t) - x(t))dt] \quad , \qquad (10.6)$$

при условии, что $y_m(t) = y(t)$. В этом случае выходной сигнал модели и объекта будем обозначать $y(t)$.
Тогда, подставляя (10.5) в (10.6), имеем

$$I = M[\int_0^T F(\sum_{k=0}^{n} a_k \frac{d^k y(t)}{dt^k} - x(t))dt] \quad , \qquad (10.7)$$

т.е. функционал в явном виде зависит от коэффициентов модели, что удобно для аналитического исследования.
Развивая эту идею, можно формализовать удобный функционал для общего случая модели (10.1)

$$I = M[\int_0^T F(\sum_{k=0}^{n} a_k \frac{d^k y(t)}{dt^k} - \sum_{k=0}^{m} b_k \frac{d^k x(t)}{dt^k})dt] \quad (10.8)$$

Выражение:

$$\Delta = \sum_{k=0}^{n} a_k \frac{d^k y(t)}{dt^k} - \sum_{k=0}^{m} b_k \frac{d^k x(t)}{dt^k}, \quad (10.9)$$

называется обобщенной ошибкой модели. В качестве функции $F(\bullet)$, как правило, принимают квадрат обобщенной ошибки

$$I = M[\int_0^T \Delta^2(t)dt] \quad (10.11)$$

Этот функционал удобен тем, что в явном виде зависит от параметров модели и от доступных измерению входного и выходного сигналов объекта. Однако при вычислении этого функционала возникают определенные трудности, связанные с дифференцированием сигналов $x(t)$ и $y(t)$, а также с необходимостью выполнения операции математического ожидания. Структурная схема вычисления обобщенной ошибки и оценки критерия I представлена на на рисунке 10.2, где $p = \frac{d}{dt}$ - оператор дифференцирования.

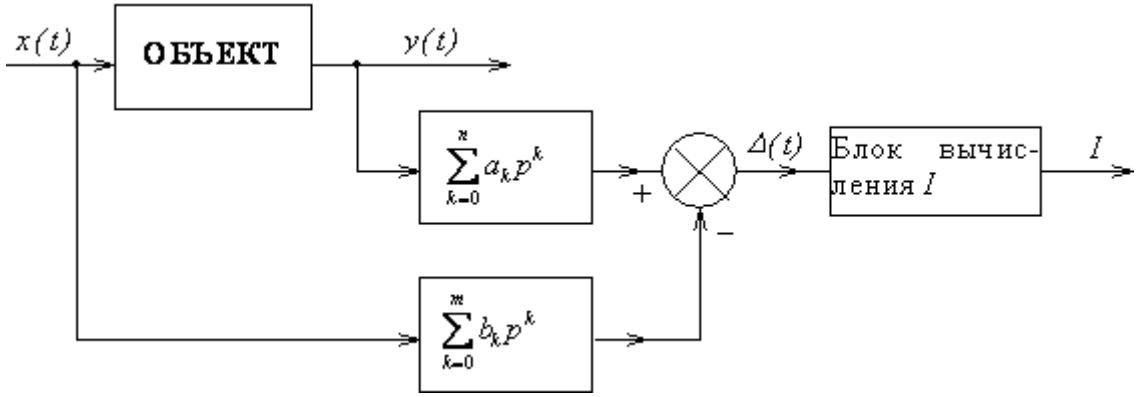


Рисунок 10.2 - Структурная схема вычисления обобщенной ошибки и оценки критерия I

Однако по условиям физической реализуемости можно создавать лишь устройства, порядок числителя которых меньше(или равно) порядка знаменателя, т.е. можно реализовать устройства с операторами

$\sum_{k=0}^{n} a_k p^k \Big/ D(p)$ и $\sum_{k=0}^{m} b_k p^k \Big/ D(p)$,

где $D(p)$ - многочлен степени больше или равно n; $m \le n$.

Тогда структурная схема вычисления обобщенной ошибки $\tilde{\Delta}(t)$ и оценки критерия $\tilde{I}$ будет иметь вид, представленной на рисунке 10.3.

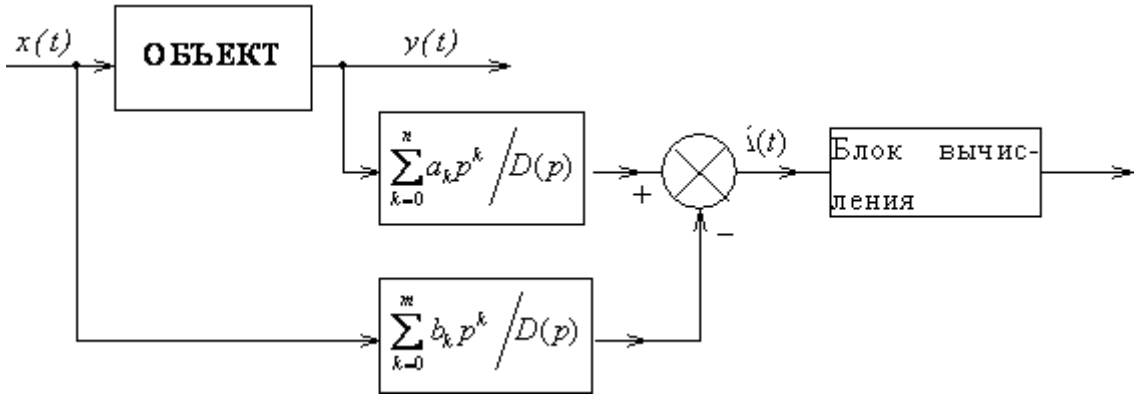


Рисунок 10.3 - Структурная схема вычисления обобщенной ошибки

$$\tilde{\Delta}(t) = (\sum_{k=0}^{n} a_k p^k \Big/ D(p))y(t) - (\sum_{k=0}^{m} b_k p^k \Big/ D(p))x(t); \qquad (10.12)$$

$$\tilde{I} = M[\int_0^T \tilde{\Delta}^2(t)dt]. \qquad (10.13)$$

Структурной схеме, изображенной на рисунке 10.3, эквивалентна схема, приведенная на на рисунке 10.4.

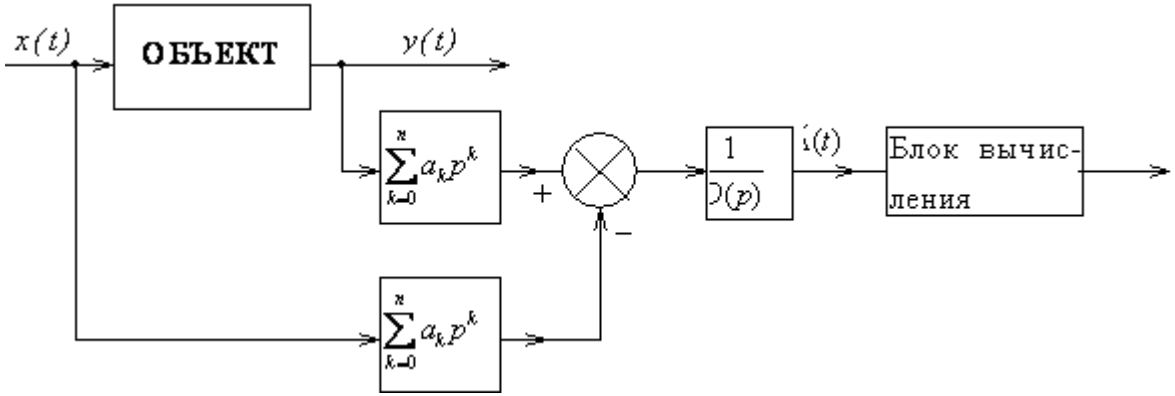


Рисунок 10.4 – Эквивалентная структурная схема вычисления обобщенной ошибки

Таким образом, обобщенная ошибка $\tilde{\Delta}(t)$, измеряемая с помощью физически реализуемых устройств, отличается от обобщенной ошибки $\Delta(t)$ тем, что $\tilde{\Delta}(t)$ является результатом преобразования $\Delta(t)$ фильтром с передаточной функцией $1/D(p)$. В силу конечности полосы пропускания этого фильтра происходят искажения сигнала обобщенной ошибки. Эти искажения будут тем меньше, чем больше полоса пропускания фильтра.

Если величина ошибок модели и оценок критериев приближения удовлетворяют требованиям к качеству модели, то модель считается адекватной объекту и может быть использована для решения задач моделирования, оптимизации и управления. В противном случае модель необходимо усовершенствовать путем изменения структуры и введения в нее неучтенных ранее факторов.

Структурная и параметрическая идентификация. Моделирование (в широком смысле) является основным методом исследований во всех областях знаний и научно обоснованным методом оценок характеристик сложных систем, используемым для принятия решений в различных сферах инженерной деятельности. Существующие и проектируемые системы можно эффективно исследовать с помощью математических моделей (аналитических и имитационных), реализуемых на современных ЭВМ, которые в этом случае выступают в качестве инструмента экспериментатора с моделью системы.

Одна из проблем современной науки и техники — разработка и внедрение в практику проектирования новейших методов исследования характеристик сложных информационно-управляющих и информационно-вычислительных систем различных уровней (например, автоматизированных систем научных исследований и комплексных испытаний, систем автоматизации проектирования, АСУ технологическими процессами, а также интегрированных АСУ, вычислительных систем, комплексов и сетей, информационных систем, цифровых сетей интегрального обслуживания и т. д.). При проектировании сложных систем и их подсистем возникают многочисленные задачи, требующие оценки количественных и качественных закономерностей процессов функционирования такихсистем, проведения структурного алгоритмического и параметрического их синтеза.

Характеризуя проблему моделирования в целом, необходимо учитывать, что от постановки задачи моделирования до интерпретации полученных результатов существует большая группа сложных научно-технических проблем, к основным из которых можно отнести следующие: идентификацию реальных объектов, выбор вида моделей, построение моделей и их машинную реализацию, взаимодействие исследователя с моделью в ходе машинного эксперимента, проверку правильности полученных в ходе моделирования результатов, выявление основных закономерностей, исследованных в процессе моделирования. В зависимости от объекта моделирования и вида используемой модели эти проблемы могут иметь разную значимость.

В одних случаях наиболее сложной оказывается идентификация, в других — проблема построения формальной структуры объекта. Возможны трудности и при реализации модели, особенно в случае имитационного моделирования больших систем. При этом следует подчеркнуть роль исследователя в процессе моделирования. Постановка задачи, построение содержательной модели реального объекта во многом представляют собой творческий процесс и базируются на эвристике. И в этом смысле нет формальных путей выбора оптимального вида модели. Часто отсутствуют формальные методы, позволяющие достаточно точно описать реальный процесс. Поэтому выбор той или иной аналогии, выбор того или иного математического аппарата моделирования полностью основывается на имеющемся опыте исследователя и ошибка исследователя может привести к ошибочным результатам моделирования.

Идентификацией называется нахождение оптимальной в некотором смысле модели Am (или оценки F оператора) объекта являющейся функцией входных (U) и выходных (Y или X) параметров системы, построенной по результатам наблюдений над входными и выходными переменными объекта.

Полученная модель, (в случае ее адекватности исследуемому объекту) в основном, предназначена для замены реального объекта в задачах управления, прогноза, конструирования, поиска оптимальных режимов и условий, имитации явлений и устройств и т.п.

Задачей идентификации называется **обратная** задача системного синтеза.

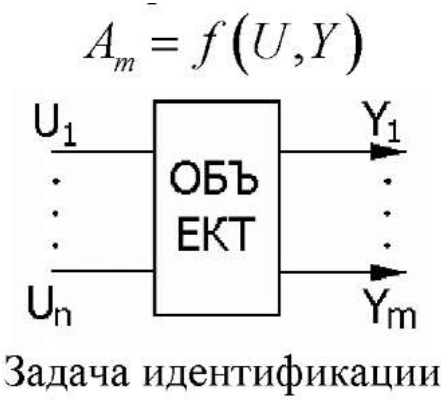


Рисунок 10.2 – Задача идентификации

Среди задач идентификации выделяют два типа (проблемы):

1. **структурная** идентификация (в широком смысле слова);
2. **параметрическая** идентификация (идентификация в узком смысле слова).

Структурная идентификация подразумевает построение модели типа «черный ящик», т.е. информация об объекте отсутствует полностью или частично. Главная задача структурной идентификации - определение структуры модели (см. рисунок 10.5).

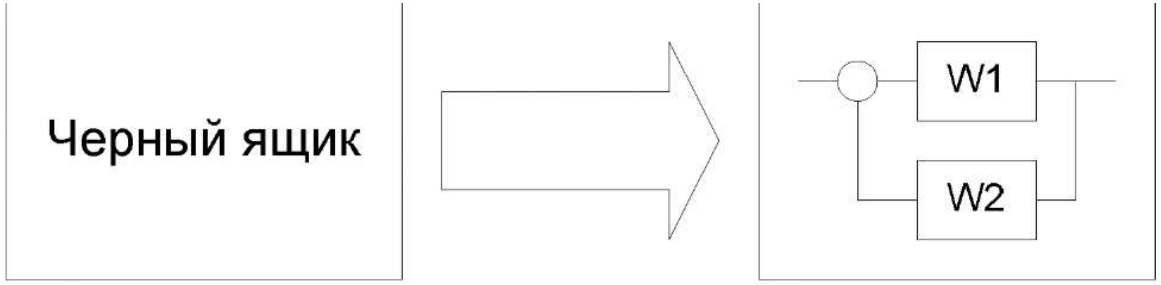
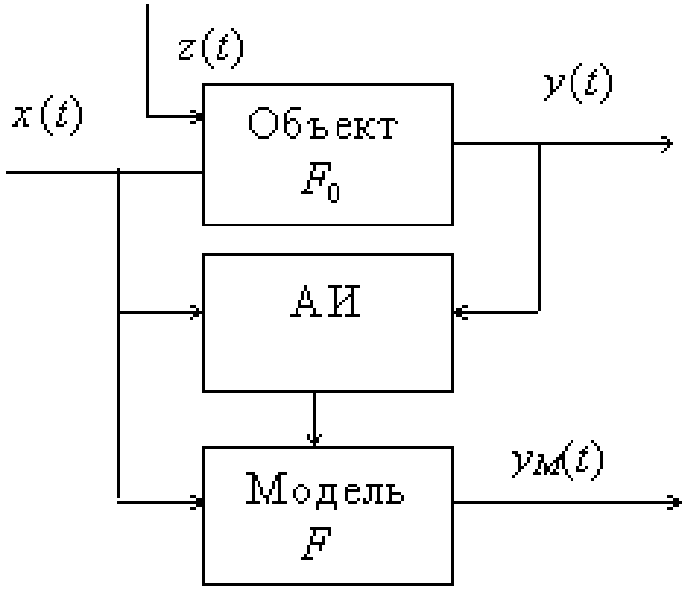


Рисунок 10.5- Структурная идентификация

Первая проблема - **структурная идентификация** является, по существу, основной проблемой всего процесса моделирования состоящего из следующих четырех основных этапов:

- постановка задачи
- выбор структуры модели и математическое описание ее блоков;
- исследование модели;
- экспериментальная проверка модели.

Структура модели ещё не сама модель, и для определения ее параметров необходимо располагать измерениями. Задачу определения параметров модели по наблюдениям работы объекта при заданной структуре модели называют идентификацией в узком смысле или параметрической идентификацией. Например, известна система уравнений, описывающая некоторый объект. Необходимо определить только коэффициенты уравнений. Процедура структурной идентификации показана на рисунке 10.6



АИ – алгоритм идентификации

Рисунок 10.6 – Процедура идентификации

Вторая проблема (**параметрическая** идентификация) при заданной структуре модели поддается формализации и смыкается с четвертым этапом моделировании.

Таким образом, по отношению к многоэтапному процессу моделирования в целом идентификация выступает как инструмент проверки гипотез о соответствии структуры или параметров объекта и модели на основе экспериментальных данных о его функционировании. Характер и степень несоответствия используются при этом для принятия содержательных или формализованных решений по корректировке модели (см. рисунки 10.7 и 10.8)

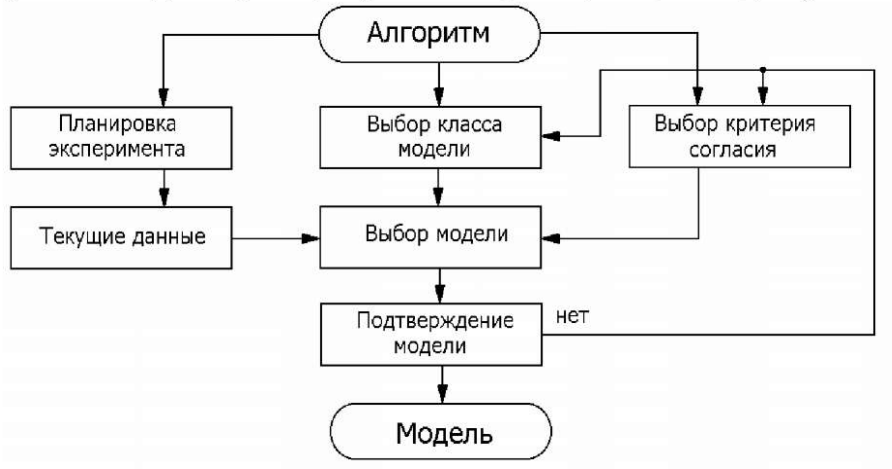


Рисунок 10.7 - Общая схема идентификации модели

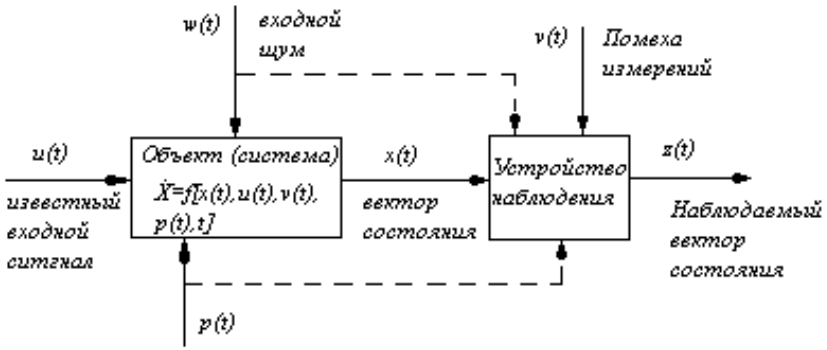


Рисунок 10.8 - Структурная схема идентификации модели

На этом рисунке - текущие данные могут быть получены в результате пассивного или активного эксперимента. Пассивный эксперимент, когда исследователь не влияет на процедуру регистрации (изменения) данных. Активный эксперимент, когда исследователь формирует программу эксперимента.

Методы планирования эксперимента позволяют эффективно организовать эксперимент. Выбор класса модели и модели – достаточно сложная процедура при решении которых обычно приходится делать компромиссный выбор между сложностью модели и её точностью. Выбор критерия согласия - это выбор критерия точности математического описания. Часто используют для этого метод наименьших квадратов. Метод наименьших квадратов не требует никакой априорной информации. В отдельных практических задачах автоматического управления в качестве мер сравнения можно принимать различные характеристики (временные, частотные и т.д.) объекта и модели. Критерием идентификации в этом случае является рассогласование этих характеристик. Однако, если модель используется в самонастраивающейся САУ, настройка модели по динамическим характеристикам требует наличие измерителей динамических характеристик объекта и модели, что приводит к конструктивному усложнению САУ и уменьшению быстродействия контуров самонастройки.

Построение моделей опирается в основном на данные наблюдений. Существует два способа (а также комбинации) формирования математических моделей.

В первом способе исследуемая система расчленяется на такие подсистемы, свойства которых очевидны из ранее накопленного опыта. По существу, это означает, что мы опираемся на известные законы природы и другие надежные соотношения, основанные на ранее проведенных экспериментальных исследованиях. Формальное математическое объединение этих подсистем становится моделью всей системы. Такой подход называется моделированием или аналитическим методом построения моделей. В его рамках проведение натурного эксперимента не обязательно. Конкретный вид процедуры моделирования сильно зависит от прикладной задачи и часто определяется традиционными и специфическими средствами из рассматриваемой прикладной области. Основной прием сводится к структуризации процесса в виде блок-схемы, блоки которой состоят из более простых элементов. Процесс восстановления системы по этим простым блокам чаще всего выполняется с помощью ЭВМ и приводит не к математической, а к машинной модели системы.

В другом способе построения моделей непосредственно используются экспериментальные данные. В этом случае ведётся регистрация входных и выходных сигналов системы, и модель формируется в результате обработки соответствующих данных. Этот способ называется идентификацией.

Таким образом, задача идентификации формулируется следующим образом: по результатам наблюдений за входными и выходными переменными объекта построить оптимальную в некотором смысле его модель. При этом объект находится в нормальном режиме функционирования (т. е. в обстановке случайных возмущений и помех). Иными словами, если объект описывается некоторым неизвестным оператором F_0, то имея измеренные значения входа и выхода необходимо построить оценку F, оператора объекта, оптимальную в смысле некоторого критерия.

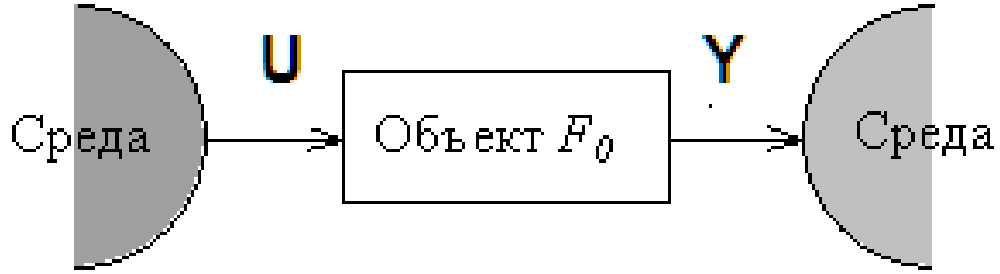


Рисунок 10.9 - Взаимодействие идентифицируемого объекта со средой

Рисунок 10.9 иллюстрирует взаимодействие идентифицируемого объекта со средой. Это взаимодействие происходит по каналам U и Y. По каналу U – вход среда воздействует на объект, а по каналу Y (выход) объект воздействует на среду. Задача идентификации сводится к определению оператора модели F, связывающего вход и выход объекта Y = F(U).

Так как часто отсутствует модель среды, воздействующей на объект, то его вход естественно рассматривать как случайную функцию времени, статистические свойства которой в общем случае неизвестны. Однако известны наблюдения входа и выхода объекта, т.е. реализации функций U и Y. На объект может воздействовать ненаблюдаемые факторы V(t), которые рассматриваются как случайные помехи.

Таким образом, идентификация — это синтез оптимального в смысле некоторого критерия согласия (точности) модельного оператора F исследуемого объекта с использованием результатов наблюдений за его входными и выходными переменными.

Классификация методов идентификации.. В соответствии с современной теорией можно предложить следующую классификацию идентификации:

1) по конечному результату идентификации (структурная и параметрическая);

2) по способу изучения объекта идентификации (активная и пассивная)

3)по типу идентифицируемой модели (линейная и нелинейная; детерминированная и стохастическая; с непрерывным и дискретным временем; стационарная и нестационарная; одномерная и многомерная; статическая и динамическая; с сосредоточенными и распределёнными параметрами).

Активная и пассивная идентификация. Успех идентификации объекта существенно зависит от соотношения двух факторов: объема априорной информации о структуре объекта и объема измерительной информации. Априорные сведения помогают определить структуру модели, т.е. ее вид (число входов и выходов, характер связи между ними). При **активном** способе идентификации реализация входа формируется самим исследователем путем подачи на вход объекта испытательного сигнала желаемой формы (скачкообразного сигнала, импульсного сигнала, сигнала в виде гармонических, прямоугольных, трапецеидальных, треугольных колебаний и др.). Реализацией выхода объекта является его реакция на испытательный сигнал. При этом в современной теории идентификации широко применяются методы оптимального планирования эксперимента. При **пассивном** способе идентификации в качестве реализаций входа и выхода объекта принимают естественные их изменения в процессе нормального функционирования объекта.

На рисунке 10.10 показана связь между априорной информацией (о структуре) и апостериорной информацией об измерениях) при построении модели. Верхняя часть рисунка иллюстрирует процесс построения модели как конкретный пример использования физических законов с последующей линеаризацией и преобразованием к системе обыкновенных дифференциальных уравнений. Получаю щиеся уравнения определяют структуру модели. На каж- дом шаге возникают ошибки. В нижней части рисунка иллюстрируется процедура оценивания, основанная на **измерениях и** включающая обработку данных **и** алгоритмы **оценивания.** Здесь также следует учесть различные виды ошибок.

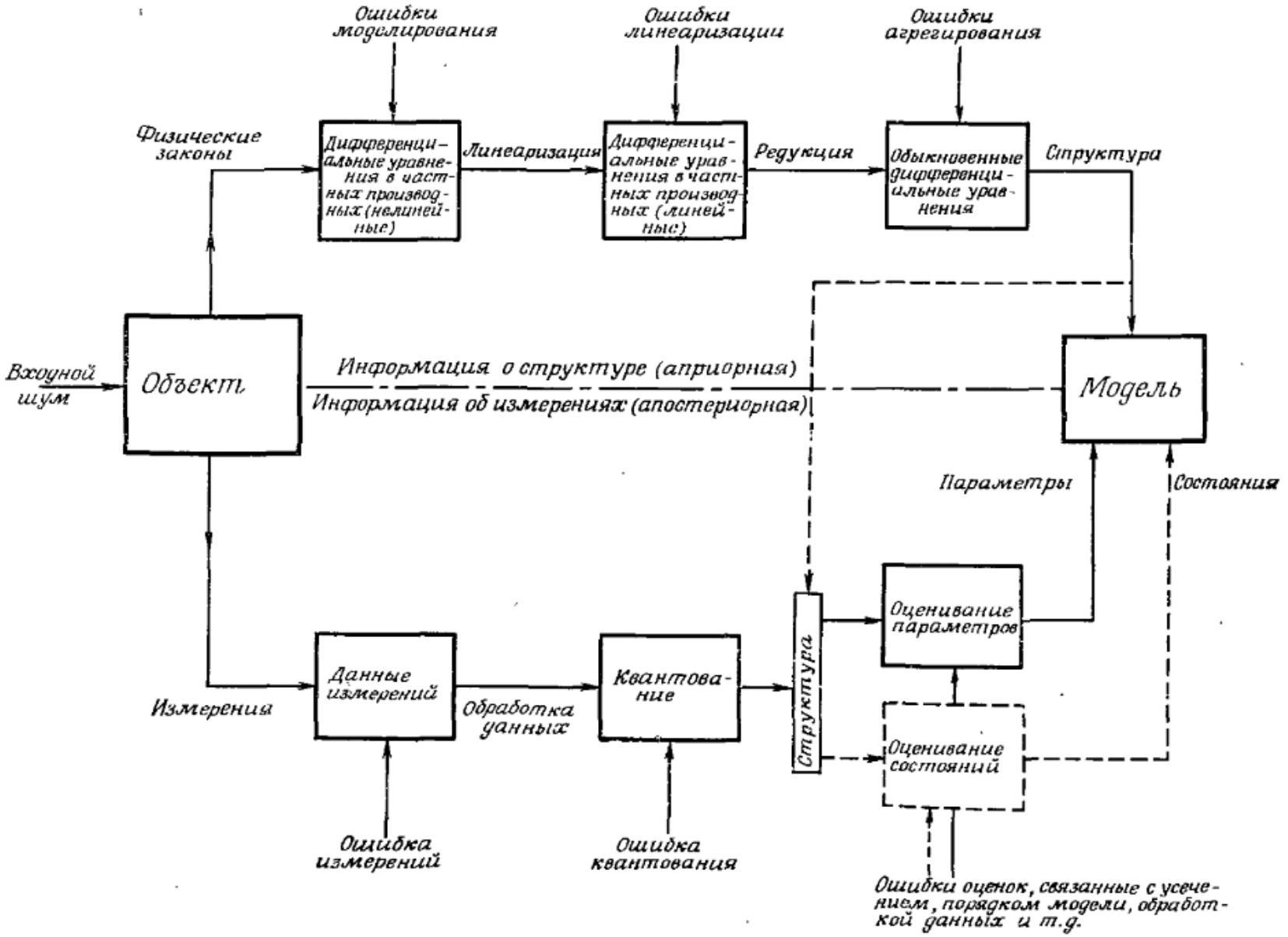


Рисунок 10.10 - Связь между априорной информацией (о структуре) и апостериорной информацией об измерениях)

Основная литература

1. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
2. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.

Дополнительная литература

3. Бокс Дж, Дженкинс. Анализ временных рядов. Прогноз и управление. Выпуск 2. –М.: Мир, 1974, 193с
4. Толчеев В.О., Ягодкина Т.В. Методы идентификации линейных одномерных динамических систем. -М.: Изд-во МЭИ, 1997

Тема 7 Проблема идентифицируемости

Лекция 11 Проблема идентифицируемости

Здесь нами будут рассмотрены проблема идентифицируемости; системные условия идентифицируемости в разомкнутых системах; системные условия идентифицируемости линейных замкнутых систем; понятие вероятностной идентифицируемости.

Условия идентифицируемости линейных динамических систем.

Рассмотрим задачу идентификации параметров математической модели, заданной в виде системы обыкновенных линейных дифференциальных уравнений:

$$\dot{x} = Ax\,.$$

Состояние системы задается вектором x(t) n- мерного эвклидового пространства и известны путем измерений следующие векторы

$$x(t), \dot{x}(t), \ddot{x}, \dddot{x}, \ldots, \frac{d^{n-1}x(t)}{dt^{n-1}} \qquad (11.1)$$

Задача состоит в нахождении такой матрицы A, чтобы выполнялись условия

$$\begin{aligned} &\dot{x} = Ax \\ &\ddot{x} = A\dot{x} \\ &\ldots\ldots\ldots\ldots\ldots \\ &\frac{d^n x}{dt^n} = A\frac{d^{n-1}x}{dt^{n-1}} \end{aligned} \qquad (11.2)$$

Состояние системы, заданное известными векторами (11.1), назовем идентифицируемым, если для него существует матрица A, удовлетворяющая равенствам (11.2).

Систему уравнений (11.2) для каждой строки матрицы А $a_j^{\tau} = (a_{j1}, a_{j2}, ..., a_{jn})$ несложно записать системой уравнений:

$$\begin{bmatrix} x^{\tau} \\ \dot{x}^{\tau} \\ \vdots \\ \dfrac{d^{n-1}x^{\tau}}{dt^{n-1}} \end{bmatrix} \quad a_j = \begin{bmatrix} \dfrac{dx_j}{dt} \\ \dfrac{d^2x_j}{dt^2} \\ \vdots \\ \dfrac{d^nx_j}{dt^n} \end{bmatrix} \text{ для j =1,2,..., n} \qquad (11.3)$$

Условие существования решения задачи идентификации, т.е. систем (3), можно записать в виде

$$\det(x, \frac{dx}{dt}, ..., \frac{d^{n-1}x}{dt^{n-1}} \neq 0 \qquad (11.4)$$

Из соотношений (11.3) параметры модели можно определить по формулам

$$a_j = \begin{bmatrix} x^{\tau} \\ \dot{x}^{\tau} \\ \vdots \\ \dfrac{d^{n-1}x^{\tau}}{dt^{n-1}} \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{dx_j}{dt} \\ \dfrac{d^2x_j}{dt^2} \\ \vdots \\ \dfrac{d^nx_j}{dt^n} \end{bmatrix}; \text{ для j =1,2,..., n} \qquad (11.5)$$

Подставим соотношения (11.2) в условие (11.4), получим неравенство:

$$\det (x, Ax, ... , An\text{-}1x) \neq 0 \qquad (11.6)$$

Укажем на связь между задачами идентификации и управляемости. Эта связь формулируется так:

- для существования решения задачи идентификации в виде математической модели $\dot{x} = Ax$
- при наблюдении вектора состояния x(t) достаточно, чтобы матрица A и вектор x(t) удовлетворяли условия вполне управляемости
  det (b, Ab, ... , An-1b) ≠0
  где b = x(t).

Поскольку матрица A наперед неизвестна, то на практике условия идентификации проверяют на основании условия (4).

Для дискретных линейных систем управления определение **идентифицируемости** формулируется следующим образом: если по известным значениям векторов x(k), x(k+1), ... , x(k+n) состояние дискретной линейной стационарной системы n-го порядка

x(k+1) = Ax(k)

можно восстановить матрицу A, то систему называют идентифицируемой.

Теорема. Решение задачи идентификации дискретной линейной системы имеет место, если
det (x(k), x(k+1), ... , x(k+n-1)) ≠0

Понятие вероятностной идентифицируемости.

Отправная точка в суждениях о потенциальной идентифицируемости системы безотносительно к методу восстановления ее параметров - как и управляемость или наблюдаемость - это свойство сугубо потенциальное, состоит в доказательстве отсутствия другой какой-либо другой системы на выбранном классе систем, которая ведет себя точно также, как и исходная система, и, следовательно, такие системы принципиально неразличимы в выбранных для идентификации условиях. Восстановить однозначно параметры такой системы нельзя. Этот подход позволяет доказать ряд теорем и вывести критерии, сходные с критериями Калмана и Ли. Существуют и принципиально неидентифицируемые системы, для которых невозможно выбрать подходящие условия, ведущие к их идентификации. Всегда находится какой-либо смежник. Покажем это.

Идентифицируемость автономных систем. Модель линейной автономной динамической системы имеет вид:
$dx/dt=Ax$,

где x - вектор состояния (пространства R^n), $x_0=x(0)$.

Определение. Линейная автономная динамическая система называется полностью идентифицируемой по состоянию, если при заданном векторе начальных условий x_0 матрица ее параметров A может быть однозначно восстановлена по временной последовательности $x=x(t)$.

Иначе, тоже следуя калмановской традиции, пара (A,x_0) полностью идентифицируема тогда и только тогда, когда множество пар, объединенных общностью интегральной кривой $x=x(t)$, вырождается в точку. В противном случае указанная пара неидентифицируема. Публикуемый ниже критерий идентифицируемости напоминает критерии управляемости и наблюдаемости.

Теорема 1. Необходимое и достаточное условие полной идентифицируемости пары (A,x_0) состоит в следующем:

$$\mathrm{rank}[x_0\ Ax_0\ A^2x_0\ \dots\ A^{n-1}x_0] = n,$$

где n - порядок системы. Матрицу, выписанную в квадратных скобках, будем называть матрицей идентифицируемости и обозначать W.

Доказательство. Опирается на разложение матричной экспоненты в конечную сумму слагаемых
$e^{At}=\sum_0^{m-1}\alpha_k(t)A^k$,
где m - степень минимального аннулирующего полинома A, $\alpha_k(t)$ - коэффициенты интерполяционного полинома Лагранжа-Сильвестра для экспоненциальной функции, определенной на спектре A.

У пар (A,x_0) и (B,x_0), объединенных общностью интегральной кривой, равны и производные процессов, т.е. $Ax(t)=Bx(t)$, где $x(t)=e^{At}x_0$. Функции $\alpha_k(t)$ линейно независимы между собой на любом интервале времени, что позволяет перейти к матричному уравнению:

$$(A-B)[x_0\ Ax_0\ A^2x_0\ \dots\ A^{m-1}x_0]=0.$$

Матрица A является корнем своего минимального аннулирующего полинома, любая ее степень, которая выше m, выражается через предыдущие, поэтому ранг матрицы справа равен рангу матрицы идентифицируемости W. Множество возможных решений сводится в точку B=A тогда и только тогда, когда ранг матрицы идентифицируемости полон. Доказательство окончено.

Дискретные системы. Идентифицируемость непрерывных систем не сводится к вопросу об идентифицируемости дискретных систем, рассмотренных Ли. Матрицу непрерывной системы в принципе нельзя восстановить однозначно по матрице системы, появляющейся в итоге дискретизации: A по значениям e^{AL} и L. Варианты восстановлений рассматривались в работе А.В. Прасолова. Между тем, проверки критериев идентифицируемости соответствующих непрерывной и дискретной систем дадут одинаковый результат, поскольку замена A в W матричной экспонентой e^{AL} для любого L (т.е. матрицей дискретной системы) не сказывается на ее ранге. Критерий идентифицируемости проверить можно даже тогда, когда параметры системы при частном (равномерном) способе составления выборки данных для идентификации найти невозможно.

Идентифицируемость - это свойство системы, а не метода восстановления ее параметров. Это системное свойство, такое же, как управляемость и наблюдаемость. Невозможность построения одного работоспособного регулятора или наблюдающего устройства, не означает невозможность построения другого. Но для неуправляемых и ненаблюдаемых систем решить задачи управления или наблюдения в калмановской их трактовке невозможно. То же самое касается построения идентификаторов. В теоретически разрешимом случае, необходимо либо изменить способ составления выборки, либо измерить производные, что близко один к другому, ввиду известных методов вычисления значений производных.

Идентифицируемость неавтономных систем. Наше рассмотрение будет неполным, без охвата общих описаний уравнений систем в пространстве состояний, а они таковы:

$dx/dt=Ax+bu$, $y=cx(t)$,

где по прежнему x - вектор состояния (пространства R^n), $x_0=x(0)$.

Определение. Линейная динамическая система называется полностью идентифицируемой по состоянию, если при заданном векторе начальных условий x_0 существует управление, при котором параметры A,b,c можно восстановить с точностью до параметров ее канонической формы (наблюдаемости) по временной последовательности x=x(t).

Иначе, пара $((A,b,c),x_0)$ полностью идентифицируема тогда и только тогда, когда существует управление, при котором множество пар, объединенных общностью интегральной кривой x=x(t), вырождается при каноническом ее описании в точку. В противном случае указанная пара неидентифицируема.

Теорема 2. Необходимое и достаточное условие полной идентифицируемости пары $((A,b,c),x_0)$ состоит в следующем:

$$\mathrm{rank}[x_0\ Ax_0\ A^2x_0\ \dots\ A^{n-1}x_0\ b\ Ab\ A^2b\ \dots\ A^{n-1}b] = n,$$

где n - порядок системы. Наблюдаемость системы при этом подразумевается, поэтому запишем иначе Rank W_o=n, Rank [W W_c]=n.

Доказательство. Ассиметричное по отношению к матрице наблюдаемости вхождение матрицы управляемости в критерий идентифицируемости диктуется разным значением проблем управляемости и наблюдаемости для нужд идентификации. Ненаблюдаемые системы содержат части, параметры которых восстановить невозможно, поэтому критерий наблюдаемости выписывается особняком как предварительное условие. Вектор состояния канонической формы наблюдаемости можно восстановить, не зная параметров системы, например, при импульсном воздействии. Расширительное толкование принципа дуальности Калмана сводит проблему управляемости системы с парой (A,b) к проблеме идентифицируемости пары (A,b), поскольку импульсный сигнал переводит систему в состояние b. Аддитивность влияний исходного начального состояния и состояния, в которое систему по истечении некоторого любого конечного времени идентификации импульсный сигнал переводит, позволяет рассматривать общие уравнения идентификации систем по нескольким запускам процесса, которые аналогичны автономному случаю, но оперируют указанной составной матрицей. Доказательство окончено.

Если вектор состояния доступен непосредственному измерению при любом управлении, то есть, уравнения выхода отсутствуют, можно говорить об идентифицируемости пары $((A,b),x_0)$ с точностью до всех параметров при условии, что Rank $[W\ W_c]=n$. У автономных систем имеем Rank $W=n$.

Нестационарные системы. Если характер нестационарности неизвестен, параметры нестационарной линейной динамической системы восстановить невозможно, поскольку уравнения по одной точке процесса заведомо неразрешимы. В данном случае имеет смысл указать на смысл грамиана идентифицируемости, аналогичному грамианам управляемости и наблюдаемости. В стационарном случае все более или менее прозрачно - все равно, что проверять, грамиан или системную матрицу.

Обыденная постановка задачи идентификации предусматривает попытку описания автономной системы с нестационарными параметрами автономной стационарной системой, причем грамиан связывает между собой параметры обоих систем, интеграл от нормы невязки фазовых скоростей которых на участке идентификации минимален. Это свойство можно положить в основу определения квазиидентифицируемости (как бы идентифицируемости), и тогда невырожденность грамиана будет гарантировать также, как и в тривиальном случае рассмотрения стационарных систем, однозначную разрешимость задачи. Квазиидентифицируемость подразумевается в задачах на восстановление медленно изменяющихся параметров, но есть ньюансы.

Квазиидентифицируемость вовсе не гарантирует близость параметров нестационарной и стационарной систем. Она гарантирует едиственность восстановления. Существуют контрпримеры, показывающие, что на классе нестационарных систем с гармоническими процессами, аналитически точно описываемыми стационарными уравнениями, "замороженные" параметры далеки от параметров стационарной системы. В таких случаях алгоритмы идентификации приводят к странным, на первый взгляд, результатам, поскольку их оценки однозначны и не меняются, а решение лежит в стороне от ожидаемого при сколь угодно медленном темпе изменения параметров. Эта особенность нестационарных систем поясняет, почему бессмысленно приводить графики "отслеживания" параметров.

Подведем некоторый итог. **Управляемость.** Понятие наблюдаемости и дуальное ему понятие управляемости были впервые введены Калманом в 1960 г. Хотя при обсуждении методов идентификации понятие наблюдаемости важнее понятия управляемости, оба они ввиду их дуальности рассматриваются совместно.

Говорят, что система является управляемой, если она может быть переведена из любого состояния $x(t_o)$ при $t=t$o в любое другое желаемое состояние $x(t_i)$ за конечный интервал времени τ ($\tau= t_i - t_o$) путем приложения кусочно-непрерывного входного воздействия/

Система является **неуправляемой**, если управляющее входное воздействие u(t) влияет не на все переменные состояния.

Кроме того, управляемая замкнутая линейная система может иметь произвольные собственные значения независимо от собственных значений соответствующей разомкнутой системы.

В литературе описаны критерии анализа управляемости (и соответственно наблюдаемости) систем. Все они основаны на рассмотрении канонического уравнения состояния и на полиномиальном разложении $e^{м}$.

Наблюдаемость. Понятие наблюдаемости дополняет понятие управляемости. Если управляемость требует, чтобы каждое состояние системы было чувствительно к воздействию входного сигнала, то наблюдаемость требует, чтобы каждое состояние системы влияло на измеряемый выходной сигнал.

Система наблюдаема, если все ее состояние можно непосредственно или косвенно определить по выходному вектору системы. Поэтому, когда определенное состояние (или изменение этого состояния) не влияет на выходной вектор, система **ненаблюдаема,** точно так же как отсутствие влияния вектора выходного сигнала на определенное состояние означает, что система неуправляема. Кроме того, ненаблюдаемая система не может быть идентифицирована; в терминах ее полной модели в пространстве состояний, очевидно, невозможна идентификация параметров, относящихся к ненаблюдаемым состояниям.

Основная литература

1. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
2. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.

Дополнительная литература

3. Семенов А. Д., Артамонов Д. В., Брюхачев А. В. Идентификация объектов управления: Учебн. пособие. - Пенза: Изд-во Пенз. гос. ун-та, 2003.- 211 с

Тема 8 Структурная статистическая идентификация

Лекция 12 Структурная статистическая идентификация

В лекции рассмотрены: структурная статистическая идентификация; статистические критерии тесноты связи; критерии и методы ориентации причинно-следственных отношений координат модели; организация статистической процедуры принятия решений на этапе идентификации структуры модели.

Прежде чем применять методы параметрической идентификации, необходимо определить структуру модели. Это одна из основных проблем теории идентификации. Как правило, в большинстве работ по синтезу математических моделей структура постулируется априори с точностью до некоторого множества неизвестных параметров. В дальнейшем это множество является основным объектом исследования. Это не означает, что методам идентификации структурной идентификации не уделялось внимания. Доминирующим подходом при выборе структуры является статистический подход и решение принимается на заданном классе моделей-претендентов. Но какие-либо формализованные подходы и методы, позволяющие выбрать структуру модели на основе доступного для наблюдения информационного множества объекта, отсутствуют. Это объясняется тем, что некоторые элементы структуры модели не поддаются адекватной математической трактовке. Поэтому структурное множество часто сужают до таких математических категорий и объектов, которые можно описать на существующем математическом языке и, следовательно, задать на классе функций из заданного множества.

Наиболее сложной и наименее изученной является проблема оценки структуры нелинейных систем. Если для класса регрессионных моделей предложен ряд методов, основанных на статистическом подходе, то для динамических систем какие-либо формализованные процедуры отсутствуют.

Структурная идентификация сводится к выбору математической модели, описывающей процессы в исследуемом объекте. Каких-либо формализованных процедур априорного выбора структуры модели до настоящего времени не существует. Объясняется это тем, что:

1) в системах управления используются математические модели, отражающие наиболее существенные стороны протекающих процессов. Поэтому в системах идентификации доминирует концепция черного ящика;

2) попытка использовать физические законы для описания природы процессов может существенно усложнить процедуру получения математического описания. Поэтому выбор структуры математической модели может превратиться в нереализуемую мечту, так как потребуется описывать процессы, которые не нашли отражения в физических законах или имеют очень сложное математическое представление;

3) идентификация новых (недостаточно изученных) объектов базируется только на концепции черного ящика, а это, в свою очередь, приводит к появлению априорной неопределенности, которую необходимо каким-то образом преодолевать;

4) любой реальный объект всегда связан с внешней средой и поэтому на этапе выбора структуры модели необходимо учитывать реальные ограничения. Это существенно влияет на сам подход к получению математического описания. Математическая модель должна быть конкретизирована только до того уровня, какой допускает имеющееся информационное множество системы.

С учетом многообразия методов структурной идентификации, очевидно, наиболее приемлемой основой для такой классификации являются свойства объектов, так как это дает возможность целенаправленно выбирать методы, позволяющие рациональным путем решить поставленные задачи для конкретных объектов.

Следует различать, как уже отмечалось выше, проблему идентификации в широком смысле – структурную идентификацию – и идентификацию в узком смысле (в малом) – параметрическую идентификацию.

Структурная идентификация на основных этапах моделирования.

Первая проблема (структурная идентификация) является, по существу, основной проблемой всего процесса моделирования, состоящего из рассмотренных в гл. I следующих четырех основных этапов:

1. постановка задачи
2. выбор структуры модели и математическое описание ее блоков;
3. исследование модели;
4. экспериментальная проверка модели.

По крайней мере, с этой проблемой смыкаются три первых этапа моделирования. Полная формализация этой проблемы вряд ли возможна из-за большого многообразия и практически неисчерпаемой сложности реальных объектов. Здесь велика роль профессионализма исследователя, знание физического механизма процессов, правильной формулировки цели и постановки задачи.

Параметрическая идентификация. Вторая проблема (параметрическая идентификация) при заданной структуре модели поддается формализации и смыкается с четвертым этапом моделирования, ей посвящено **большинство публикаций** по методам идентификации.

Таким образом, по отношению к многоэтапному процессу моделирования в целом идентификация выступает как инструмент проверки гипотез о соответствии структуры или параметров объекта и модели на основе экспериментальных данных о его функционировании. Характер и степень несоответствия используются при этом для принятия содержательных или формализованных решений по корректировке модели. Остановимся несколько подробнее на вопросах структурной и параметрической идентификации.

Организация статистической процедуры принятия решений на этапе идентификации структуры модели.

Этапы структурной идентификации. Задачи вскрытия структуры объекта. Роль структуры модели трудно переоценить, неудачный выбор ее сводит на нет и все результаты параметрической идентификации.

Среди задач вскрытия структуры объекта можно отметить следующие:

1. выделение объекта из среды;
2. ранжирование входов и выходов объекта по степени их влияния на конечный целевой показатель;
3. определение рационального числа входов и выходов объекта, учитываемых в модели;
4. определение характера связи между входом и выходом модели объекта, т. е. вида оператора системы

Рассмотрим коротко каждую из этих задач.

Выделение объекта из среды. Описание процесса Процесс выделения объекта из среды, прежде всего, определяется целями, для которых строится модель. Цель по отношению, например, к управлению имеет внешний характер. Она формулируется на более высоком иерархическом уровне и выражает его требования к объекту управления. В то же время определение целей связано с представлениями об объекте, в котором должны быть реализованы эти цели, т. е. нельзя эффективно сформулировать цель, не имея какой-то модели объекта управления. Следовательно, еще до формулирования цели должна быть некоторая, хотя бы приближенная, модель, которая используется для определения объекта управления. Выделение объекта из среды или разделение на подобъекты (в случае сложного объекта) должно осуществляться таким образом, чтобы он имел минимум связей со средой или с другими подобъектами (звеньями). При этом одновременно удовлетворяется также требование создания наилучших условий для использования присущих большинству объектов свойств саморегулирования. Процесс выделения объекта из среды может осуществляться как последовательный переход от простейших форм объекта к более сложным. В качестве простейшей формы можно рассматривать такую часть среды, которая несет информацию, необходимую для проверки выполнимости поставленной цели. Далее происходит расширение объекта за счет присоединения части среды таким образом, чтобы лучше удовлетворялась цель управления, например, за счет расширения ресурса управления. Этот процесс может повторяться до тех пор, пока не будет эффективно достигаться цель управления или будет показано, что она не достижима.

Пример выделения объекта из среды. Рассмотрим в качестве примера процесс нагрева металла в камерной печи. Допустим, целью управления является получение равномерно прогретой до определенной температуры заготовки (в смысле близости температур поверхности и центра). Если при этом в качестве объекта управления выделить лишь саму заготовку, то можно легко убедиться, что из-за отсутствия соответствующих ресурсов управления заданная цель не достижима. Входным параметром этого объекта является температура поверхности заготовки, теснейшим образом связанная с температурой рабочего пространства, а выходным – температура центра. Здесь в лучшем случае, можно за счет фактора времени добиться выравнивания этих температур при их заранее неопределенном уровне. Проблема существенно не меняется и в том случае, если за счет расширения объекта включить в его состав температуру рабочего пространства, поскольку последняя не является независимым управляющим воздействием. Поставленная задача решается лишь в том случае, когда в качестве объекта управления принимается все рабочее пространство печи вместе с топливосжигающими устройствами и датчиками для измерения расходов топлива и окислителя. При таком выделении объекта в нашем распоряжении имеются необходимые ресурсы управления. Если, наряду с указанной, поставить более широкую цель, например экономии топлива и уменьшения загрязнения окружающей среды, то в состав объекта следует включить устройства утилизации тепла (рекуператоры, регенераторы, котлы-утилизаторы и т. д.), газоочистку и дымовую трубу, а в качестве выходных параметров кроме температуры металла рассматривать химический состав, температуру и запыленность отходящих газов. Таким образом, даже на таком относительно простом примере можно видеть, насколько важно правильно выделить объект управления или исследования. Естественно, что это самым непосредственным образом сказывается на выборе структуры модели, которая строится для решения той или иной задачи.

Ранжирование входов и выходов и определение их рационального числа. Важное значение для определения структуры модели, которая на первых этапах исследования может быть представлена в виде многополюсника, является отбор входов и выходов объекта, которые будут включены в модель. Для этого сначала определяются все входы и выходы, состояние которых в какой-то степени влияет на выполнение цели в объекте (например, цели управления). Затем среди них отбираются наиболее существенные, которые и образуют многополюсник модели с размерностью $П \times Т$ Отбор существенных факторов является непростой процедурой. Для этого используются методы экспертных оценок (непосредственное ранжирование, метод парных сравнений и др.), наблюдение за функционированием реального объекта и деятельностью оператора на нем, специально спланированный и организованный эксперимент на объекте. Для принятия решения о структуре модели (в смысле числа входов и выходов) могут сравниваться несколько конкурирующих моделей. При этом в качестве критерия выбора предпочтительной модели может служить ее точность с одной стороны и сложность или реализуемость с другой.

Если уже имеется некоторая хотя бы приближенная модель или теория, то решение рассмотренной задачи отбора существенных параметров значительно упрощается.

Определение характера связи между входом и выходом. Целенаправленное определение характера связи между входами и выходами модели (вида оператора), т. е. внутренней структуры многополюсника $П \times Т$, возможно лишь на основе некоторых теоретических представлений о механизме процессов, протекающих в объекте. В противном случае остается лишь простой перебор структур, что практически нереально. Имеются попытки формализации синтеза структур, например, путем выделения типовых звеньев, ячеек или, так называемых, условно – элементарных операторов: идеального смешения, вытеснения, диффузии, химической кинетики, теплопередачи и др. Из таких ячеек (звеньев), обменивающихся между собой потоками вещества или энергии, предлагается синтезировать соответствующие структуры моделей сложных объектов. Такой переход, безусловно, облегчает решение рассматриваемой задачи, но полностью исчерпать все многообразие свойств реальных объектов, естественно, не может. Поэтому практически гипотезы о структуре ставятся с учетом физических, физико – химических и других теоретических представлений о конкретных объектах, а для проверки этих гипотез используются экспериментально-статистические методы.

Содержательный анализ остатков как метод проверки гипотез об адекватности структуры. Исследование ошибок и остатков. Под остатками ε_i понимается разность между фактически измеренными y_i и предсказанными с помощью модели (например, регрессионного уравнения) $\hat{y}_i$ n значениями выходного параметра. Это величины, которые не удается объяснить с помощью регрессионного или какого-либо другого уравнения, т. е. остаточные ошибки модели.

Относительно ошибок делаются следующие предположения: ошибки независимы, имеют нулевые средние, постоянную дисперсию и подчиняются нормальному закону распределения. Последнее необходимо в случае использования -критерия.

Если подбираемая модель находится в удовлетворительном соответствии с объектом, остатки должны проявлять тенденцию к подтверждению сделанных нами предположений или, по меньшей мере, не должны противоречить им.

Формулируется следующий вопрос: «Не доказывают ли остатки, что наши предположения ошибочны?»

После исследования остатков можно прийти к одному из следующих выводов:

1. предположения, по-видимому, нарушены (в некотором смысле)
2. предположения, по-видимому, не нарушены.

Последнее не означает, что мы пришли к выводу о правильности предположения, мы не имеем основания для утверждения о неправильности.

Графические процедуры исследования ошибок и остатков. Основные виды графиков остатков. Процедуры исследования остатков с целью проверки модели носят графический характер, что дает возможность осуществлять не только количественный, но и качественный (содержательный) анализ степени идентичности модели и объекта.

Основные виды графиков остатков:

1. общий; например, гистограмма распределения;
2. в зависимости от времени или номера опытов, если известна их последовательность;
3. в зависимости от предсказываемых значений $\hat{y}_i$;
4. в зависимости от входных факторов x_{ji};
5. любой вид графика, который целесообразен для данной конкретной задачи.

Рассмотрим конкретнее некоторые из указанных видов графиков.

Гистограмма распределения. Гистограмма распределения. Она строится следующим образом. Весь диапазон изменения переменной, в нашем случае ошибки модели, разбивается на ряд равных интервалов (обычно на 10 – 15), которые откладываются на оси абсцисс, а на оси ординат отмечается частота попадания ошибки в каждый из этих интервалов (число случаев). Например, гистограмма, изображенная на рисунке 12.1,а, имеет симметричный характер и не дает каких-либо оснований для суждений о неправильности наших предположений.

Несимметричный характер гистограммы или наличие второго “горба” (рисунок. 12.1,б), может свидетельствовать о том, что в модели не учтена какая-то неслучайная составляющая и требуется более глубокий анализ ошибки модели.

График временной последовательности.

График временной последовательности. Возможно, несколько характерных случаев зависимости остаточной ошибки от времени или других перечисленных выше факторов. Рассмотрим их сначала для временной зависимости, пример которой приведен на Рисунок 12.2, а.

В случае (б) эффект времени не влияет на ошибку и не дает каких-либо оснований для принятия решения о целесообразности дальнейшего совершенствования модели.

В случае (в) **дисперсия** не постоянна, а растет со временем, что вызывает необходимость использования взвешенного **метод наименьших квадратов.**

В случае (г) целесообразно включить в **модель** линейный член от времени.

В случае (д) в **модель** должны быть включены линейный и квадратичный члены от времени.

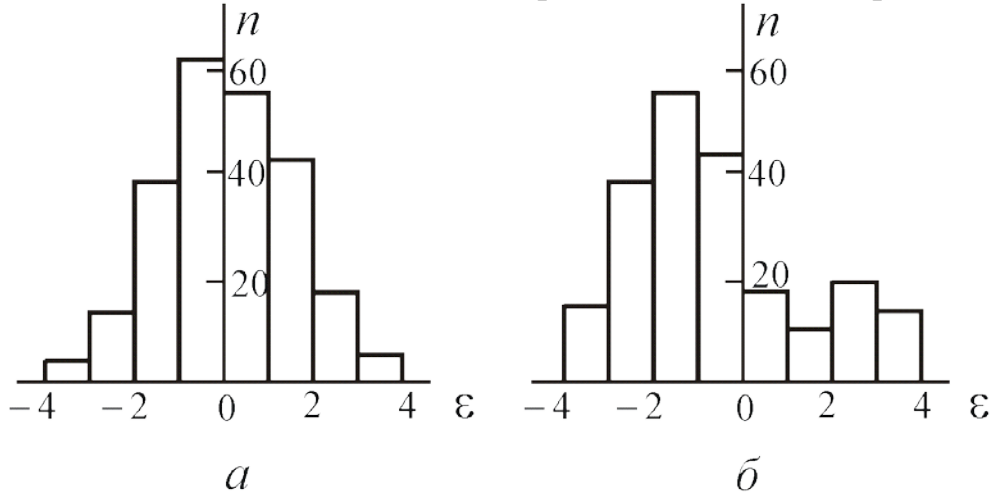


Рисунок 12.1 - Гистограммы распределения

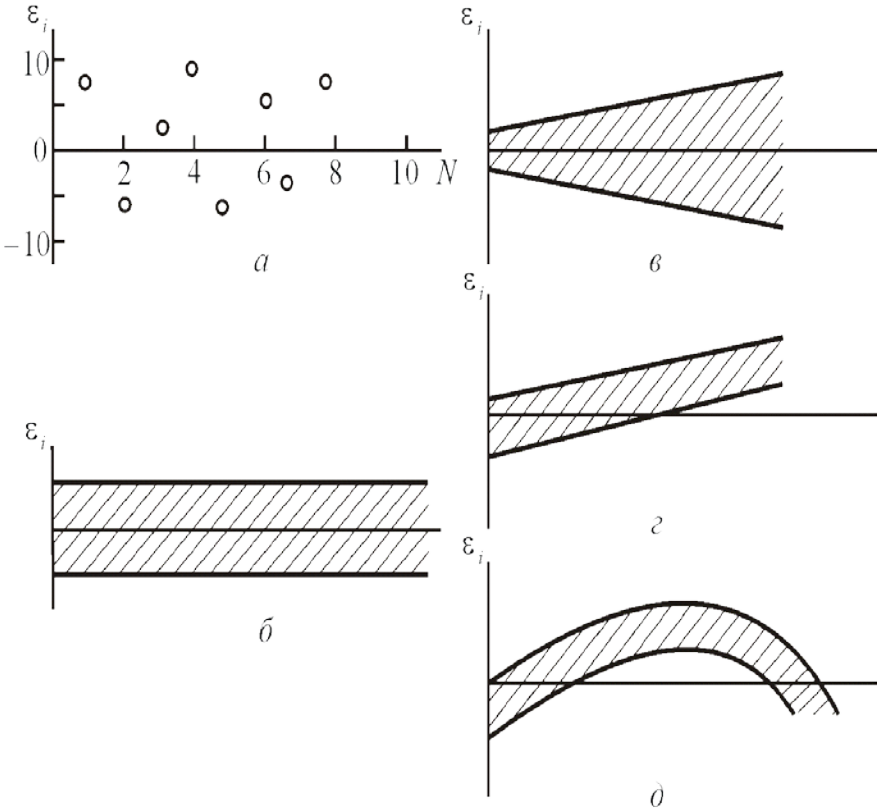


Рисунок 10.2 - Характерные случаи распределения остаточной ошибки

Возможны также различные сочетания рассмотренных случаев.

График зависимости остатков от $\hat{y}_i$ (рисунок. 12.3).

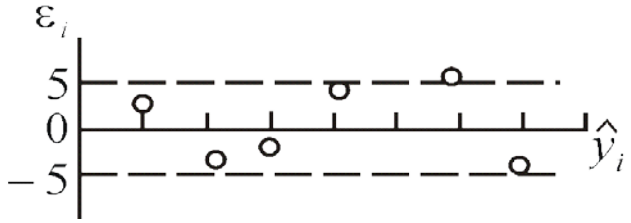


Рисунок 12.3 - Зависимость остатков от предсказанного значения выхода

Попадание большинства данных в горизонтальную полосу свидетельствует, что наши предположения, по-видимому, оправданы. Графики для случаев, когда предположения оказываются неоправданными, могут иметь вид аналогичный рассмотренным выше (рисунок. 12.2, в, г, д).

В случае (в) дисперсия вопреки предположению непостоянна и зависит от $\hat{y}_i$, что приводит к необходимости использования взвешенного метода наименьших квадратов и преобразования наблюдений $\hat{y}_i$. В случае (г) отклонения от полученного уравнения (модели) носят систематический характер, что может свидетельствовать о том, что в модели ошибочно пропущен свободный член. Случай (д) – модель неадекватна, необходимо ввести в модель квадратичные члены и взаимодействия.

Графики остатков по каждой из независимых переменных x_j строятся и анализируются аналогичным образом. Здесь случай (г) может быть следствием либо ошибок в вычислениях, либо необоснованного исключения линейного члена от x_j, а случай (д) – свидетельствовать о необходимости введения в модель квадратичного члена от x_j.

Прочие методы, эффективность, выбросы.

Среди других видов графиков можно отметить группировку остатков для разных агрегатов, печей, смен, бригад, времен года, на серии положительных и отрицательных остатков и т. д.

Эффективным при последовательном совершенствовании моделей может оказаться введение в рассмотрение новой переменной. Строится график зависимости ε_i от новой переменной, не включенной в рассматриваемую **модель**. Если такая зависимость обнаруживается, целесообразно ввести в **модель** соответствующие члены для учета этой переменной.

Большой интерес при исследовании остатков могут представить выбросы – значительные отклонения параметров от установленного **закона распределения**. С точки зрения получения устойчивых средних значений по большому ансамблю данных выбросы за зону шириной $\pm 3\sigma$, где σ – среднее квадратическое отклонение, рекомендуется не учитывать, делая предположения, что они являются результатами каких-либо промахов в постановке и проведении экспериментов.

Если ставится задача выяснения причин, вскрытия внутреннего механизма явлений, то выбросы должны подвергаться особенно тщательному анализу. При этом можно получить такую информацию, которую другие данные дать не могут, поскольку выброс (если он не является результатом промаха) связан с необычной комбинацией условий, являющейся жизненно важной и интересной с точки зрения постановки гипотез о направлениях дальнейших исследований.

Выше рассмотрены лишь основы исследования остатков, что, однако, достаточно для понимания важности этого метода при содержательном анализе моделей и определении направлений их структурного совершенствования.

В настоящее время интенсивно развивается еще целый ряд методов структурной идентификации, имеются примеры их применения в химии и других отраслях.

Основная литература

1. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
2. Эйкхоф П. Основы идентификации систем управления. М.: Мир, 1975. 683 с.

Дополнительная литература

3. Карабутов Н. Н. Структурная идентификация систем: Анализ динамических структур М.: МГИУ, 2008. – 160 с.
4. Карабутов Н. Н. Структурная идентификация систем: Анализ информационных структур. МГИУ,2009. 176 с.
5. Материалы с сайта http://www.matematicheskoe-modelirovanie.ru/kniga

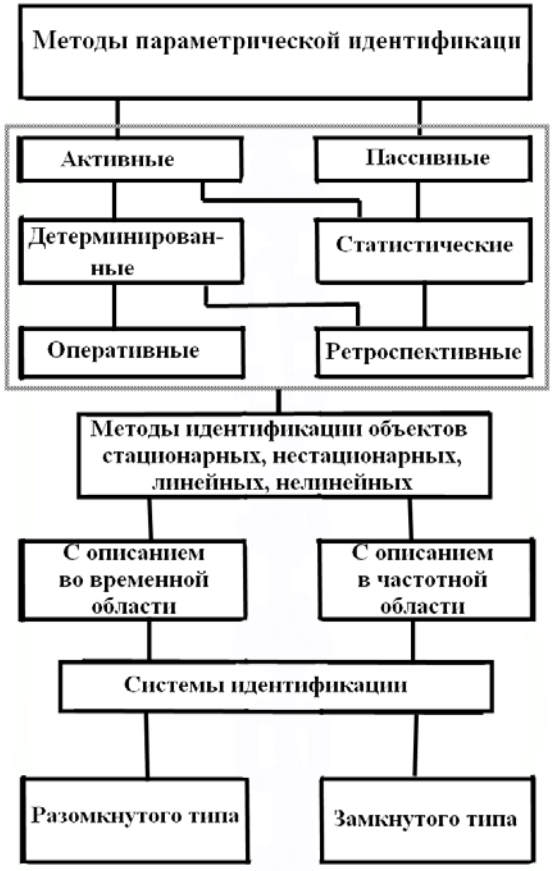
Тема 9 Оценивание параметров и состояния объектов

Лекция 13 Оценивание параметров и состояния объектов

Материал темы лекции содержит следующие вопросы: оценивание параметров и состояния объектов; фильтр Калмана-Бьюси; одновременное оценивание параметров и состояний; методы квазилинеаризации.

Оценивание параметров и состояния объектов производится на стадии параметрической идентификации.

Для решения задач параметрической идентификации разработано большое число методов, учитывающих особенности объектов, условия их функционирования, способ тестирования и математическую основу анализа экспериментальных данных, вид получаемых моделей и т. п. Методы параметрической идентификации можно характеризовать различными признаками. По способу тестирования исследуемого объекта методы идентификации делятся на активные и пассивные.



Применение активных методов предполагает подачу на вход объекта специально сформированных воздействий - детерминированных или случайного характера. Среди активных методов идентификации широкое распространение получили частотные методы, основанные на измерении установившихся выходных сигналов исследуемого объекта, вызванных гармоническим входным воздействием. Для идентификации линейных объектов используют и другие периодические воздействия (прямоугольные, треугольные), а также апериодические воздействия в виде ступенчатых, импульсных и других сигналов. В качестве случайных тестовых сигналов особенно широко применяют псевдослучайные двоичные последовательности, что объясняется простотой их получения и удобством обработки с помощью средств вычислительной техники. Достоинство активной идентификации заключается в нежесткости требований к априорным данным об объекте. Основываясь на методах планирования эксперимента, такую идентификацию можно осуществлять целенаправленно, что позволяет ускорять выявление закономерностей в зависимостях между переменными объекта и сокращать тем самым временные и материальные затраты на его испытания.

При использовании пассивных методов идентификации объект находится в условиях нормального функционирования. При этом параметры его модели ищутся по результатам статистической обработки наблюдений естественных изменений величин на его входе и выходе. При пассивной идентификации применяют статистические принципы обработки данных измерений (методы корреляционного и регрессионного анализов, стохастической 319 аппроксимации и др.). Преимущество пассивных методов состоит в том, что для их применения достаточна регистрация переменных только в режимах рабочего функционирования объекта. Это особенно важно при идентификации реальных промышленных процессов с непрерывным производством дорогостоящего продукта. С другой стороны, пассивные методы идентификации сопряжены со значительными затратами времени на накопление и обработку информации. Кроме того, их применение возможно лишь в том случае, если воздействие на входе объекта обладает достаточно широким частотным спектром (по крайней мере, не меньше, чем полоса частот, в пределах которой необходимо оценить динамическую характеристику объекта). Снижение точности идентификации, обусловленное последней причиной, можно существенно уменьшить, если добавить к естественному входному воздействию объекта специального случайного сигнала небольшого уровня с заданными статистическими характеристиками. В этом случае ослабляется также и влияние обычно действующих на объект неконтролируемых помех.

При идентификации объектов применяют как детерминированные, так и статистические методы обработки данных наблюдений, что определяется характером анализируемых сигналов. Детерминированные методы могут быть использованы только при активной идентификации в случае, когда сигналы на входе и выходе объекта имеют детерминированную форму. Однако в реальных условиях из-за сильного влияния помех такие сигналы зачастую сильно зашумлены. Требуемая точность анализа здесь может быть достигнута дополнением детерминированных алгоритмов обработки статистическим усреднением (сглаживанием) получаемых результатов. Очевидно, что для этого необходимы более длительные испытания объекта с целью накопления данных.

По признаку временных затрат методы идентификации разделяются на оперативные и ретроспективные. При оперативной идентификации обеспечивается отслеживание меняющихся параметров объекта. С этой целью применяют рекуррентные алгоритмы обработки данных, которые удается реализовать аппаратными средствами в темпе, близком к скорости протекания процес-

сов в изучаемом объекте. При ретроспективной идентификации значительно упрощают условия решения задачи идентификации. В этом случае можно многократно обращаться к накопленным экспериментальным данным и подбирать наиболее эффективные алгоритмы их анализа.

Методы идентификации часто различают также по признаку, указывающему на их приспособленность к исследованию динамических объектов того или иного класса. Важной особенностью при идентификации является наличие или отсутствие процедур сравнения получаемой модели с объектом. Это определяет соответственно две возможные структуры построения систем идентификации: по разомкнутой и замкнутой схемам.

Результатом решения задачи идентификации является математическая модель, представленная во временной или частотной области. При этом полученная модель адекватна объекту по поведению, т. е. по динамическим свойствам, в соответствии с выбранным при идентификации критерием подобия.

Рассмотрим некоторые основные понятия из области оценивания параметров и состояния объектов.

Винеровское оценивание — задача нахождения импульсной характеристики линейной стационарной системы, которая минимизирует среднюю квадратическую ошибку между реальным y(t) и желаемым d(t) выходными сигналами при бесконечном времени наблюдения. На вход системы подается сигнал f(t), выходной сигнал определяется выражением:

$$y(t) = \int_{-\infty}^{+\infty} w(\tau) f(t-\tau) d\tau$$

Предполагается, что условия применения, характер сигналов и помех остаются достаточно стабильными, их статистические характеристики меняются мало. Если же условия переменны и помехи в процессе работы систем изменяются существенно, то возникает необходимость автоматической оптимизации параметров систем. Это осуществляется в различного рода экстремальных, адаптивных, обучаемых системах. Во время второй мировой войны перед американским математиком Н. Винером встала задача отделения полезного сигнала от шума при решении задач автоматизации систем противовоздушной обороны, использующих радиолокационную технику. В 1942 г. Н. Винер решил эту задачу, допустив что искомая система должна быть линейной с постоянными параметрами, время наблюдения бесконечно, входной и желаемый выходной сигналы системы являются стационарными и стационарно связанными случайными процессами, система минимизирует среднюю квадратическую ошибку между желаемым и реальным выходными сигналами.

Уравнение Винера – Хопфа. В основе статистических методов определения импульсной переходной функции объекта лежит интегральным уравнением Фредгольма $(-\infty < t < \infty)$ или интегральное **уравнение Винера-Хопфа** (t>0):

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} R_{xx}(\tau - t)g(t)dt \approx \int_{0}^{T} R_{xx}(\tau - t)g(t)dt, \qquad \text{(УВХ)}$$

где $g(t)$ - импульсная характеристика объекта, весовая функция.

Определенная из этого уравнения $g(t)$ является наилучшей в смысле среднеквадратичной оценки. В уравнении (УВХ) обозначено *Rxy* - взаимная корреляционная функция выхода и входа; *Rxx* - корреляционная функция входного сигнала.

Решение уравнения (УВХ) с целью определения $g(t)$ проводят различными методами, например: методом подбора ординат импульсной функции на управляемом фильтре (требует специальную аппаратуру); методом сведения уравнения (УВХ) к системе линейных алгебраических уравнений (см. ниже); методом преобразования Фурье.

УВХ было получено Винером и Хопфом при решении задачи радиационного равновесия внутри звезд. Также используется в кибернетике, при решении задачи выделения, фильтрации полезного сигнала из его смеси с шумом.

Напомним, что выход $x(t)$ линейного объекта, имеющего вход $y(t)$ определяется интегралом свертки:

$$x(t) = \int_{0}^{t} g(t-\tau)y(\tau)d\tau = \int_{0}^{t} g(\tau)y(t-\tau)d\tau \qquad \text{(УС)}$$

Фильтр Винера .Во время второй мировой войны Винер выполнил свои фундаментальные, ставшие теперь классическими исследования по интерполяции, экстраполяции и сглаживанию временных рядов. Фильтр Калмана позволяет получить решение, сформулированное в пространстве состояний. В отличие от метода Винера фильтр Калмана легко распространяется на случай нестационарных сигналов

Фильтр Калмана-Бьюси. В 1960 году . Рудольф Калман предложил алгоритм решения фдачи линейной оптимальной фильтрации для дискретных нестационарных гауссовских случайных процессов. А в 1961 году совместно с Р. Бьюси он опубликовал алгоритм решения задачи для непрерывного времени]. Р. Калман несколько изменил постановку задачи линейной оптимальной фильтрации. Он вместо математического ожидания и ковариации для описания оцениваемого случайного процесса x(t, ω) применил формирующий фильтр - динамическую систему. возбуждаемую белым гауссовским шумом. В результате был получен рекуррентный алгоритм решения задачи оценивания в случае наличия белого шума в измерениях. Алгоритм удобен для реализации на ЭВМ. Это привело к его широкому распространению, несмотря на то. что в общем случае замена описания случайного процесса с помощью математического среднего и ковариации на описание формирующим фильтром сталкивается с еще не решенной проблемой нахождения параметров динамической системы.

Винеровские фильтры лучше всего подходят для обработки процессов или отрезков процессов в целом (блочная обработка). Для последовательной обработки требуется текущая оценка сигнала на каждом такте с учетом информации, поступающей на вход фильтра в процессе наблюдения. При винеровской фильтрации каждый новый отсчет сигнала потребовал бы пересчета всех весовых коэффициентов фильтра. В настоящее время широкое распространение получили адаптивные фильтры, в которых поступающая новая информация используется для непрерывной корректировки ранее сделанной оценки сигнала (сопровождение цели в радиолокации, системы автоматического регулирования в управлении и т.д).

Проблема построения формирующего фильтра полностью решена лишь для стационарных случайных процессов, спектральная плотность которых допускает факторизацию. Но все же для большинства практических задач теории управления эту задачу удается успешно решать, а во многих случаях формирующий фильтр известен уже на стадии проектирования системы управления.

Фильтр Калмана — эффективный (имеющий способ гарантированно достигать результат за конечное число действий) рекурсивный фильтр, оценивающий вектор состояния динамической системы, используя ряд неполных и зашумленных измерений.. Фильтр Калмана широко используется в инженерных и эконометрических приложениях: от радаров и систем технического зрения до оценок параметров макроэкономических моделей. Калмановская фильтрация является важной частью теории управления, играет большую роль в создании систем управления. Совместно с линейно-квадратичным регулятором фильтр Калмана позволяет решить задачу линейно-квадратичного гауссовского управления. Фильтр Калмана и линейно-квадратичный регулятор – возможное решение большинства фундаментальных задач в теории управления. В большинстве приложений, количество параметров, задающих состояние объекта, больше, чем количество наблюдаемых параметров, доступных для измерения. При помощи модели объекта по ряду доступных измерений фильтр Калмана позволяет получить оценку внутреннего состояния. Фильтр Калмана предназначен для рекурсивного дооценивания вектора состояния априорно известной динамической системы, то есть для расчёта текущего состояния системы необходимо знать текущее измерение, а также предыдущее состояние самого фильтра. Таким образом фильтр Калмана, как и множество других рекурсивных фильтров, реализован во временном представлении, а не в частотном.

Наглядный пример возможностей фильтра – получение точных, непрерывно обновляемых оценок положения и скорости некоторого объекта по результатам временного ряда неточных измерений его местоположения. Например, в радиолокации стоит задача сопровождения цели, определения её местоположения, скорости и ускорения, при этом результаты измерений поступают постепенно и сильно зашумлены. Фильтр Калмана использует вероятностную модель динамики цели, задающую тип вероятного движения объекта, что позволяет снизить воздействие шума и получить хорошие оценки положения объекта в настоящий, будущий или прошедший момент времени.

Рассмотрим второй пример. Пусть про старый тихоходный автомобиль известно, что его разгон до 60 миль в час занимает не менее 10 секунд. Представим, что его спидометр дает очень зашумленные измерения, которые находятся в разбросе в 40 миль в час вокруг истинного значения скорости автомобиля. Из неподвижного положения, которое можно определить точно, так как при этом не вращаются колеса автомобиля, водитель до упора нажимает педаль акселератора. Через пять секунд спидометр показывает 70 миль в час. Водитель, зная, что автомобиль не может разгоняться так быстро, и используя информацию о погрешности спидометра, может прийти к выводу, что более вероятная скорость автомобиля около 40 миль в час. Схожим образом поступает и фильтр Калмана, благодаря информации о характеристиках шума и динамике системы он уменьшает влияние шума измерений.

Одновременное оценивание параметров и состояния

Задача оценивания параметров тесно связана с задачами фильтрации и оценивания состояния. Схемы обработки сигналов, которые позволяют отделить полезные сигналы от нежелательных возмущений (шумов), называются фильтрами.

Упомянутый выше фильтр Калмана можно описать как метод (аппаратуру, программу вычислений) для обработки результатов измерений, искаженных помехами; эта обработка должна осуществляться таким образом, чтобы получалась оптимальная оценка некоторой переменной. Этот фильтр может применяться в целом ряде задач:

а) фильтрация, интерполяция, сглаживание, экстраполяция (предсказание) временных рядов и непрерывных сигналов ;

б) оценивание состояния, необходимое для реализации алгоритмов оптимального управления в случаях, когда некоторые или даже все компоненты вектора состояния объекта не поддаются непосредственному измерению;

в) объединение результатов измерений некоторой переменной несколькими измерительными приборами, различающимися типом ошибок (например, в задачах навига ции — акселерометры, гироскопы, радиолокаторы с измерением эффекта Допплера и т. д.);

г) оценивание параметров путем перехода от вектора состояния к вектору параметров.

Методы квазилинеаризации.

Метод квазилинеаризации впервые ввели Беллман и Калаба [1, 2] для решения краевых задач в теории нелинейных дифференциальных уравнений. Применение этого метода для идентификации параметров нелинейных систем было рассмотрено в основном в работах [3—6].

Метод квазилинеаризации по существу представляет собой метод преобразования нелинейной многоточечной краевой задачи, являющейся в основном стационарной, в линейную нестационарную задачу. Этот метод применим как к непрерывным, так и к дискретным процессам. Предполагается, что идентифицируемые параметры постоянны и, так же как и во всех других методах идентификации нелинейных систем, должен быть задан тип нелинейности, по крайней мере в виде аппроксимации. Может быть учтена и нестационарность параметров, если они изменяются медленно в сравнении со скоростью сходимости процедуры идентификации. Сходимость процедуры будет довольно высокой, если имеется близкое начальное приближение к величинам параметров, которые необходимо идентифицировать. Метод по своей сути является итерационным; он не требует введения специальных пробных воздействий и поэтому применим для использования в реальном масштабе времени.

Так как процедуры идентификации методом квазилинеаризации сходятся к истинным значениям параметров только тогда, когда начальные приближения величин параметров оказываются внутри области сходимости, для указанных процедур требуется определенная априорная информация о диапазоне значений параметров; в этом нет необходимости, если используются последовательные подходы, рассмотренные в гл. 6 и 7. Метод квазилинеаризации имеет особое значение в тех случаях, *когда различные переменные состояния системы не могут быть измерены одновременно во все моменты измерений.* В этом случае описание системы в пространстве состояний само по себе является многоточечной задачей, для которой применение метода квазилинеаризации естественно. Метод квазилинеаризации по своей сути является методом идентификации, основанным на фиксированном числе измерений, а не на последовательно возрастающем объеме измерений, как в гл. 6, 7. В том случае, когда имеется достаточное число измерений некоторых состояний, а измерения других состояний не проводятся, метод квазилинеаризации может дать оценки этих недостающих состояний и параметров одновременно

При построении системы управления не всегда можно с достаточной точностью описать объект управления линеаризованной моделью. Если объект управления не удается линеаризовать, приходится решать задачу идентификации нелинейного объекта.

Одним из наиболее мощных методов идентификации нелинейных систем является метод квазилинеаризации. Этот метод успешно применяется и тогда, когда различные компоненты вектора состояния измеряются в разное время, или некоторые компоненты вектора состояния вообще не измеряются.

Метод квазилинеаризации представляет собой преобразование нелинейной стационарной многоточечной краевой задачи в линейную нестационарную задачу.

Применение метода квазилинеаризации позволяет одновременно получить оценку вектора параметров и вектора состояния системы на интервале идентификации.

Рассмотрим нелинейную систему, описываемую уравнением:

$$\dot{x} = f(x,u,q), \qquad (13.1)$$

где: x – вектор состояния размерности n,;
u – вектор входных воздействий размерности m,;
q – вектор параметров системы размерности k,;

f – вектор-функция размерности n.

Предполагается, что векторы u и x измеряются (x может измеряться частично), а вид компонентов вектор-функции f известен. Вектор параметров q неизвестен и подлежит идентификации.

Полагаем, что компоненты вектора параметров q неизменны на интервале идентификации, то есть,

$$\dot{q} = 0. \qquad (13.2)$$

Введем новый вектор состояния

$$\chi = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ q_1 \\ \vdots \\ q_k \end{bmatrix}. \qquad (13.3)$$

Для этого вектора справедливо уравнение

$$\dot{\chi} = \Psi(\chi,u), \qquad (13.4)$$

где

$$\Psi = \begin{bmatrix} f_1(\chi,u) \\ \vdots \\ f_n(\chi,u) \\ \hline 0 \\ \vdots \\ 0 \end{bmatrix}. \qquad (13.5)$$

Предположим, что уже найдена *N*-я итерация оценки вектора состояния (13.4) $\hat{\chi}_N$. Тогда, раскладывая уравнение (13.5) в ряд Тейлора относительно $\hat{\chi}_N$ и учитывая только члены первого порядка малости, получим *(N+1)*-ю оценку вектора:

$$\dot{\hat{\chi}}_{N+1} = \hat{\Psi}_{N+1} = \hat{\Psi}_N + \frac{\partial \hat{\Psi}_N}{\partial \chi}(\hat{\chi}_{N+1} - \hat{\chi}_N), \qquad (13.6)$$

где: $\dfrac{\partial \hat{\Psi}_N}{\partial \chi} = \dfrac{\partial \hat{\Psi}}{\partial \chi}\Big|_{\chi=\chi_N}$, $\dfrac{\partial \hat{\Psi}}{\partial \chi} = \left[\dfrac{\partial \hat{\Psi}}{\partial x} \vdots \dfrac{\partial \hat{\Psi}}{\partial q}\right]$,

$$\frac{\partial \hat{\Psi}}{\partial x} = \begin{bmatrix} \frac{\partial \hat{\Psi}_1}{\partial x_1} & \frac{\partial \hat{\Psi}_1}{\partial x_2} & \dots & \frac{\partial \hat{\Psi}_1}{\partial x_n} \\ \frac{\partial \hat{\Psi}_2}{\partial x_1} & \frac{\partial \hat{\Psi}_2}{\partial x_2} & \dots & \frac{\partial \hat{\Psi}_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{\Psi}_n}{\partial x_1} & \frac{\partial \hat{\Psi}_n}{\partial x_2} & \dots & \frac{\partial \hat{\Psi}_n}{\partial x_n} \end{bmatrix}, \qquad \frac{\partial \hat{\Psi}}{\partial q} = \begin{bmatrix} \frac{\partial \hat{\Psi}_1}{\partial q_1} & \frac{\partial \hat{\Psi}_1}{\partial q_2} & \dots & \frac{\partial \hat{\Psi}_1}{\partial q_k} \\ \frac{\partial \hat{\Psi}_2}{\partial q_1} & \frac{\partial \hat{\Psi}_2}{\partial q_2} & \dots & \frac{\partial \hat{\Psi}_2}{\partial q_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{\Psi}_n}{\partial q_1} & \frac{\partial \hat{\Psi}_n}{\partial q_2} & \dots & \frac{\partial \hat{\Psi}_n}{\partial q_k} \end{bmatrix}.$$

Уравнение (13.6) линейно относительно $\hat{\chi}_{N+1}$, поэтому его можно записать в виде:

$$\dot{\hat{\chi}}_{N+1} = \hat{A}_N \hat{\chi}_{N+1} + \hat{V}_{N+1}, \qquad (13.7)$$

где:

$$\hat{A}_N = \frac{\partial \hat{\Psi}}{\partial \chi}\Big|_{\chi=\hat{\chi}_N}, \quad \hat{V}_N = \hat{\Psi}\Big|_{\chi=\ddot{\chi}_N} - \frac{\partial \hat{\Psi}}{\partial \chi}\Big|_{\chi=\hat{\chi}_N} = \hat{\Psi}_N - \hat{A}_N \hat{\chi}_N.$$

Уравнение (13.7) линейное нестационарное. Его решение имеет вид:

$$\hat{\chi}_{N+1}(t) = X_N(t,t_0)\hat{\chi}_{N+1}(t_0) + \int_{t_0}^{t} X_N(t,\tau)V_N(\tau)\,d\tau, \qquad (13.8)$$

или

$$\hat{\chi}_{N+1}(t) = X_N(t,t_0)\hat{\chi}_{N+1}(t_0) + \int_{t_0}^{t} X_N(t,\tau)\hat{\Psi}_N(\tau) - \int_{t_0}^{t} X_N(t,\tau)\hat{A}_N(\tau)\hat{\chi}_N(\tau)d\tau, \qquad (13.9)$$

где $X_N(t,\tau)$ – фундаментальная матрица однородного уравнения:

$$\dot{\hat{\chi}}_{N+1} = \hat{A}_N\hat{\chi}_{N+!}. \qquad (13.10)$$

Для нахождения вектора χ размерности $n+k$ необходимо иметь $n+k$ измерений. Предположим, что измеряются компоненты вектора состояния $x_l(t_j)$, то есть, j-е компоненты вектора состояния в t_j моменты времени. Тогда вектор $\hat{\chi}_{N+1}(t_0)$ удовлетворяет многоточечному граничному условию

$$x_j(t_j) = X_{j,N}(t_j,t_0)\hat{\chi}_{N+1}(t_0) + \rho_{j,N}(t_j), \qquad (13.11)$$

$$\rho_{j,N}(t_j) = \int_{t_0}^{t_j} X_{j,N}(t_j,\tau)V_N(\tau)\,d\tau. \qquad (13.12)$$

В (13.11) и (13.12) через $X_{j,N}()$ обозначена j-я строка фундаментальной матрицы $X_N()$. Компоненты вектора состояния $x_j(t_j)$ могут быть одними и теми же или разными.

Так как имеется $n+k$ измерений, то уравнений (13.11) тоже $n+k$. По сути уравнения (13.11) являются линейными алгебраическими с $n+k$ переменными. Таким образом, $n+k$ измерений дает $n+k$ линейных алгебраических уравнений с $n+k$ неизвестными, составляющими вектор $\hat{\chi}_{N+1}(t_0)$. Этот вектор ищется, и последние k его компонентов представляют собой искомый вектор параметров q.

Блок-схема алгоритма вычислений при идентификации методом квазилинеаризации приведена на рисунке 13.1. В блоке *2* производится ввод $n+k$ измерений $x_j(t_j)$, начального приближения вектора параметров $\hat{q}_0$ и начальных условий для вектора состояния $x_0(t_0)$.

В блоке 4 производится решение уравнения с начальными приближениями $\hat{q}_{N+1}, x_N(t_0)$, которыми на первом шаге являются $\hat{q}_0, x_0(t_0)$.Решение производится на интервале $[t_0, t_{max}]$, где t_{max} –максимальное из времен t_j, в которые производятся измерения $x_j(t_j)$.

В блоке *5* решение *x(t)*, полученное в блоке *4*, подставляется в уравнение (13.9) для получения оценок из уравнения (13.11).

В блоке *6* решается система из $n+k$ линейных алгебраических уравнений (13.11) и ищется $\hat{\chi}_{N+1}(t_0)$, то есть, $\hat{q}_{N+1}$и $\hat{x}_{N+1}(t_0)$.

В блоке *7* оценивается норма разности приближений вектора параметров на последних двух шагах. Если эта норма не больше заданной величины ε, то процесс идентификации завершается, а в блоке *9* производится вывод найденного значения вектора параметров. В противном случае происходит переход к следующей итерации.

Процедура идентификации методом квазилинеаризации сходится к истинным значениям только в том случае, если начальные приближения параметров достаточно близки к истинным значениям. Для получения начального приближения можно использовать результаты идентификации линеаризованной модели или применить эвристические методы идентификации.

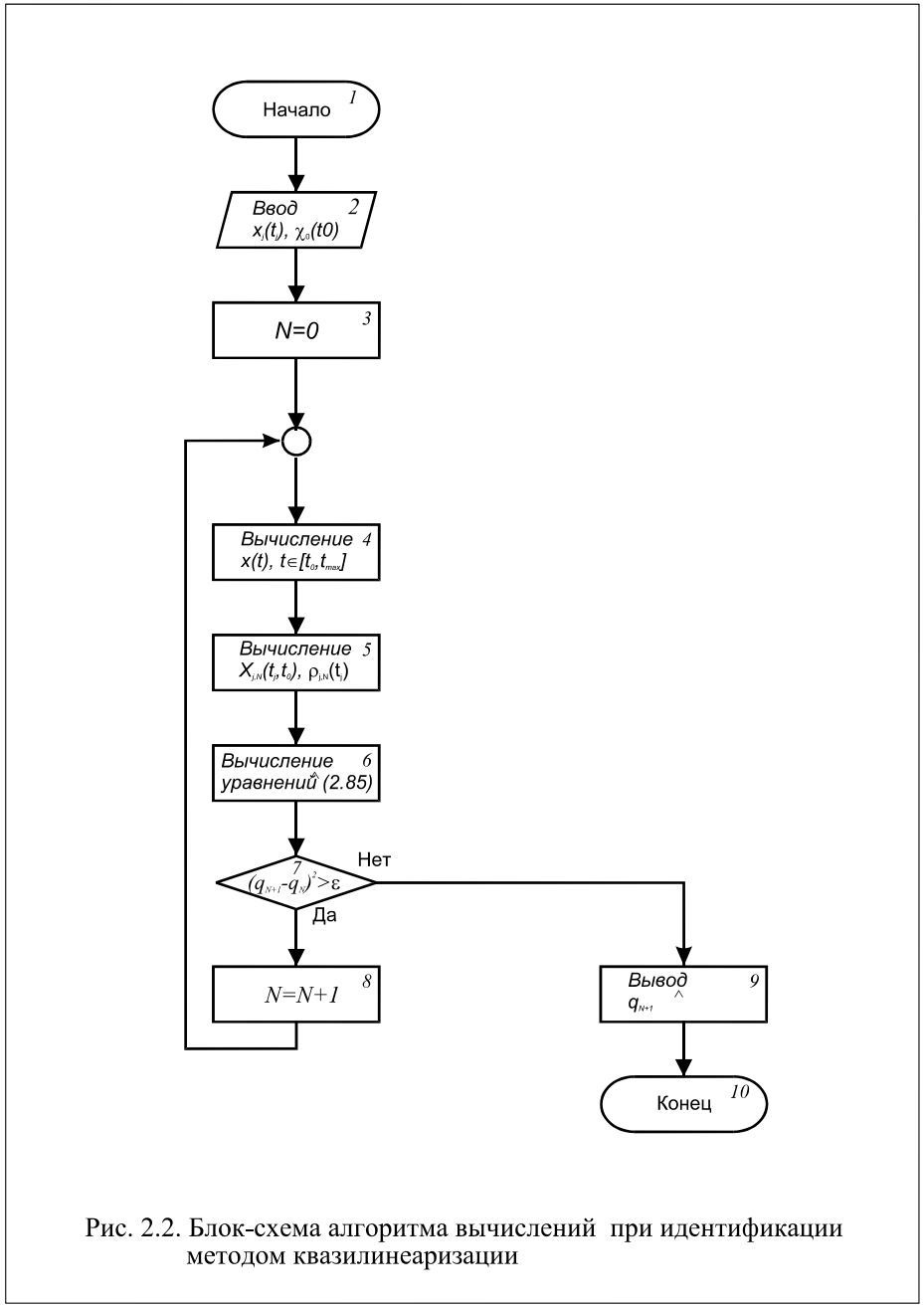


Рис. 2.2. Блок-схема алгоритма вычислений при идентификации методом квазилинеаризации

Рисунок 13.1 – блок-схема алгоритма вычислений при идентификации методом квазилинеаризации

Рассмотрим реализацию метода квазилинеаризации на примере.

Пример

Рассмотрим достаточно простую систему:

$$\dot{x} = ax^3 + bu(t). \qquad (13.13)$$

Здесь: $u(t)$ – входная измеряемая скалярная функция,
x – скаляр,
a, b – необходимо найти.

Имеется три измерения переменной состояния $x(t_1)$, $x(t_2)$, $x(t_3)$. Обозначим:

$$\chi = \begin{bmatrix} x \\ a \\ b \end{bmatrix}.$$

В соответствии с (13.6), $(N+1)$-е приближение вектора χ имеет вид:

$$\dot{\hat{\chi}}_{N+1} = \hat{\Psi}\Big|_{\hat{\chi}=\hat{\chi}_N} + \frac{\partial\hat{\Psi}}{\partial\chi}\Big|_{\hat{\chi}=\hat{\chi}_N} (\hat{\chi}_{N+1} - \hat{\chi}_N), \qquad (13.14)$$

$$\hat{\Psi} = \begin{bmatrix} \hat{a}\hat{x}^3 + \hat{b}u \\ 0 \\ 0 \end{bmatrix}, \quad \frac{\partial\hat{\Psi}}{\partial\chi} = \begin{bmatrix} 3\hat{a}\hat{x}^2 & x^3 & u \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad (13.15)$$

В соответствием с уравнение (13.10), будет:

$$\begin{bmatrix} \dot{\hat{x}} \\ \dot{\hat{a}} \\ \dot{\hat{b}} \end{bmatrix}_{N+1} = \begin{bmatrix} 3\hat{a}\hat{x}^2 & x^3 & u \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_N \begin{bmatrix} \hat{x} \\ \hat{a} \\ \hat{b} \end{bmatrix}_N. \qquad (13.16)$$

Из уравнения (13.16) ищется фундаментальная матрица $X_N(t,t_0)$. Согласно (13.7)

$$\hat{V}_N = \hat{\Psi}_{\chi=\hat{\chi}_N}\Big| - \hat{A}_N\hat{\chi}_N = \\ = \begin{bmatrix} \hat{a}\hat{x}^3 + \hat{b}u \\ 0 \\ 0 \end{bmatrix}_N - \begin{bmatrix} 3\hat{a}\hat{x}^2 & \hat{x}^3 & u \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_N \begin{bmatrix} \hat{x} \\ \hat{a} \\ \hat{b} \end{bmatrix}. \qquad (13.17)$$

По найденной $X_N(t,t_0)$ с учетом (13.17) из (13.13) ищутся $\rho_{j,N}(t_j)$. Затем из уравнений

$$\begin{cases} x(t_1) = X_{11,N}(t_1,t_0)\hat{x}_{N+1}(t_1) + X_{12,N}(t_1,t_0)\hat{a}_{N+1}(t_1) + \\ \qquad + X_{13,N}(t_1,t_0)\hat{b}_{N+1}(t_1) + \rho_{1,N}(t_1), \\ x(t_2) = X_{11,N}(t_2,t_0)\hat{x}_{N+1}(t_2) + X_{12,N}(t_2,t_0)\hat{a}_{N+1}(t_2) + \\ \qquad + X_{13,N}(t_2,t_0)\hat{b}_{N+1}(t_2) + \rho_{1,N}(t_2), \\ x(t_3) = X_{11,N}(t_3,t_0)\hat{x}_{N+1}(t_3) + X_{12,N}(t_3,t_0)\hat{a}_{N+1}(t_3) + \\ \qquad + X_{13,N}(t_3,t_0)\hat{b}_{N+1}(t_3) + \rho_{1,N}(t_3). \end{cases} \qquad (13.18)$$

ищутся $\hat{x}_{N+1}(t_0)$, $\hat{a}_{N+1}$, $\hat{b}_{N+1}$. При необходимости процедура повторяется.

Идентификация линейных динамических систем. Определение весовой функции из уравнения свертки

По наблюдениям входного и выходного сигналов линейной стационарной системы на конечном промежутке времени нужно определить ее весовую функцию (импульсная переходная функция).

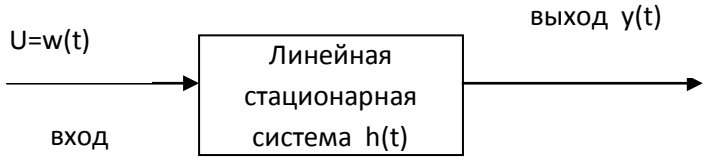


Рисунок 13.2 - Линейная стационарная системы

Выходной сигнал системы при входе и нулевых начальных условиях выражается известным интегралом свертки:

$$y(t) = \int_0^t h(t-\tau)w(\tau)d\tau \qquad (13.23)$$

предполагается, что $w(\tau) = 0$ при $r < 0$, $w(0) \neq 0$.

Введем теперь аппроксимацию входной функции времени $w(t)$ кусочно-постоянной функцией в N точках с шагом Δ, причем $N\Delta = T$

$$w(t) \approx w(n\Delta) \text{ при } n\Delta < t < (n+1)\Delta \qquad (13.24)$$

h(t) примем постоянной между точками разбиения.

$$h(t) \approx h\left(\frac{2n+1}{2}\Delta\right) \text{ при } n\Delta<t<(n+1)\Delta \qquad (13.15)$$

В терминах ступенчатой аппроксимации $w(t)$ и $h(t)$ интеграл (13.23) при $t = n\Delta$ приближенно запишется в виде

$$y(n\Delta) = \Delta\sum_{i=0}^{n-1} h\left(\frac{2n-1}{2}\Delta - i\Delta\right)w(i\Delta) \qquad (13.25)$$

Обозначим вектор наблюдения выхода (размерности N) через

$Y^T(T) = [y(\Delta)\ y(2\Delta)\ \ldots\ y(N\Delta)]$

и значений весовой функции

$$h^T(T) = \left[h\left(\frac{\Delta}{2}\right) h\left(\frac{3\Delta}{2}\right) \ldots h\left(\frac{2N-1}{2}\Delta\right) \right].$$

Перепишем уравнение (13.25) в векторно- матричном виде

$$Y(T) = \Delta W h(T) \qquad (13.26)$$

Матрица W определяется равенством

$$W = \begin{bmatrix} w(0) & 0 & 0 & 0 & \ldots & 0 \\ w(\Delta) & w(0) & 0 & 0 & \ldots & 0 \\ w(2\Delta) & w(\Delta) & w(0) & 0 & \ldots & 0 \\ \ldots\ldots\ldots\

где $h_n = h\left(\frac{2n-1}{2}\Delta\right), \qquad h_1 = \frac{Y(\Delta)}{\Delta w(0)}$

Достоинством рассмотренного подхода является возможность использовать любые входные сигналы. Поскольку нет необходимости применять специальные тестовые сигналы, можно использовать реализации, полученные в процессе нормальной эксплуатации системы.

Если входной сигнал является функцией единичного скачка, алгоритм (13.28) заметно упрощается. В этом случае w(iΔ) = 1 для всех i, (13.28) примет вид

$$h_n = \frac{y(n\Delta)}{\Delta} - \sum_{i-1}^{n-1} h_{n-i}.$$

Определив величину $H_n = \sum_{i-1}^{n-1} h_{n-i}$,

получим $h_n = \frac{y(n\Delta)}{\Delta} - H_n$ при этом $H_n = H_{n-1} + h_{n-1}$.

Оценивание весовой функции по методу наименьших квадратов

Рассмотрим систему с одним входом и одним выходом показанную на рисунке 13.2.. Объект предполагается линейным и стационарным. Выход системы запишем в виде (см. рисунок 13.3):

$$z(t) = \int_0^{T_S} h(\tau)x(t-\tau)d\tau + n(t) \qquad (13.29)$$

h(τ) - весовая функция (импульсная переходная функция)

x(t-τ)- вход, n(t) - невязка (иногда называют шумом),

T_S - время установления, определяется как min интервал времени, измеренный от момента подачи импульсного сигнала до момента, когда реакция системы составит 5% пикового значения .

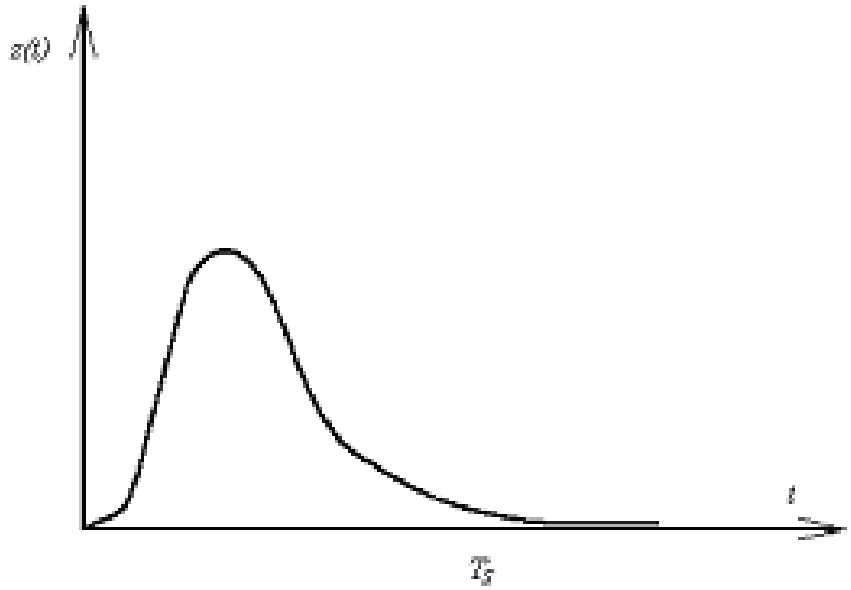


Рисунок 13.3 - Зависимость выхода системы z(t) от времени – уравнение 13.29

Входная и выходная переменные представлены в формуле (13.29) в виде отклонений от своих математических ожиданий, т.е.

$x(t) = x - m_x, \qquad z(t) = z - m_z.$

Представим уравнение (13.29) в дискретном виде

$$z(i\Delta) = \sum_{j=0}^{N_S-1} h(j\Delta)x[(i-j)\Delta]\Delta + n_i, i = 0,1,...,N_{m-1}$$

или

$$z_i = \sum_{j=0}^{N_S-1} h_j x_{i-j}\Delta + n_i,$$

(13.30)

Здесь $T_S = N_S\Delta$ - время установления, $T_m = N_m\Delta$ - время измерения выхода, N_i - содержит не только невязку в моменты времени n(iΔ), но и ошибки аппроксимации функции x(t-τ).

В результате аппроксимации задача оценивания непрерывной функции h(τ) заменяется (параметризуется) оцениванием конечного множества параметров $h_0,..., h_{Ns-1}$ (называется дискретной импульсной переходной функцией)

Для упрощения представления запишем уравнения (2) в матричном виде:

$$\begin{bmatrix} z_o \\ \vdots \\ z_{N_m-1} \end{bmatrix} = \begin{bmatrix} x_{0,0} \cdots x_{0,N_S-1} \\ \cdots \\ x_{N_m-1,0} \cdots x_{N_m-1,N_S-1} \end{bmatrix} \begin{bmatrix} h_o\Delta \\ \vdots \\ h_{N_S-1}\Delta \end{bmatrix} \begin{bmatrix} n_o \\ \vdots \\ n_{N_m-1} \end{bmatrix}$$

(13.31)

Запишем матричное представление в символическом виде

$$z = A\beta + n$$

(13.32)

Задача сводится к определению вектора параметров β при заданной матрице A и вектора измерений z.

Критерием при оценивании вектора параметров β является выбор таких β, которые минимизируют сумму квадратов невязок на интервале измерений.

Положим $J = \sum_{i=1}^{N_m-1} n_i^2$ в матричном виде $J = n^T n$ (13.33)

подставляя (13.32) в (13.33) получим $J=(z-A\beta)^T (z-A\beta)$.

Необходимо определить β^*, удовлетворяющие условию

$$J^* = \min_{\beta} J = J /_{\beta=\beta^*}.$$

Необходимым условием вычисления J является выполнение условия экстремума

$$\frac{\partial y}{\partial \beta} \Big|_{\beta=\beta^*} = 0 \quad .$$

Запишем уравнения (13.21) в виде сумм

$$J = \sum_{i=0}^{N_m-1} \left(z_i - \sum_{j=0}^{N_S-1} a_{ij}\beta_j \right) \left(z_i - \sum_{k=0}^{N_S-1} a_{ik}\beta_k \right) \qquad (13.34)$$

Продифференцируем (13.24) J по компонентам вектора β:

$$\frac{\partial J}{\partial \beta_m} /_{\beta_m=\beta_m^*} = \sum_{i=0}^{N_m-1} \left[-a_{im} \left(z_i - \sum_{k=0}^{N_S-1} a_{ik}\beta_k^* \right) + \left(z_i - \sum_{j=0}^{N_S-1} a_{ij}\beta_j \right)(-a_i) \right]_=$$

$$= -2 \sum_{i=0}^{N_m-1} a_{im} \left(z_i - \sum_{k=0}^{N_S-1} a_{ik}\beta_k^* \right) \quad \text{при } m = 0,1,..., N_S - 1 \qquad (13.35)$$

Формулу (13.35) представим в матричном виде:

$$\frac{\partial J}{\partial \beta} /_{\beta=\beta^*} = -2A^T (z - A\beta^*) = 0 \qquad (13.36)$$

это уравнение является необходимым условием экстремума J.

Достаточным условием при расчете min J является положительная определенность квадратной матрицы

$$\frac{\partial}{\partial \beta} \left(\frac{\partial J}{\partial \beta} \right)^T /_{\beta=\beta^*}$$

Если формулу (13.26) продифференцируем еще раз по β, то получим

$$\frac{\partial}{\partial\beta}\left(\frac{\partial J}{\partial\beta}\right)^{T} /_{\beta=\beta^{*}} = 2A^{T}A \qquad (13.37)$$

Если матрица $A^{T}A$ - неособенная, и также правая часть не зависит от β в (13.37), то условие экстремума (13.36) является необходимым и достаточным условием минимума.

Перепишем (13.36) в виде:

$A^{T}A\,\beta^{*} = A^{T}z$, отсюда $\beta^{*} = (A^{T}A)^{-1}A^{T}z$ (13.38)

Напомним, что

$$A=\begin{bmatrix} x\dots x \\ x\dots x \end{bmatrix},\quad \beta^{*T}=\left[h_0^{*}\Delta,\dots,h_{N_i-1}^{*}\Delta\right],\quad z^{*T}=\left[z_0\dots z_{N_m-1}\right].$$

В непрерывной форме уравнение (13.38) принимает вид

$$\int_0^{T_S}\left[\int_0^{T_m} x(t-\tau)x(t-\theta)dt\right]h^{*}(\theta)d\theta=\int_0^{T_m} x(t-\tau)z(t)dt \qquad (13.39)$$

Уравнение (13.39) – **уравнение Винера-Хопфа (УВХ)** и может быть переписано в виде:

$$\int_0^{T_S} R_{xx}(r-\theta)h^{*}(\theta)d\theta=R_{xz}(r)$$

где $R_{xx}(\tau)$ - автокорреляционная функция $\frac{1}{T_m}\int_0^{T_m} x(t-\tau)x(t)dt$ и $R_{xz}(\tau)$ - взаимная корреляционная функция.

Автокорреляционной функцией случайного процесса X(t) называется неслучайная функция двух аргументов $K_{xx}(t,t')$, которая при каждой паре значений t,t‘ равна корреляционному моменту соответствующему случайному сечению функции X(t).

$$K_{xx}(t,t') = M\begin{bmatrix} 0 & 0 \\ X(t) & X(t') \end{bmatrix}.$$

Оценивание состояния фильтром Калмана.

Для учёта погрешности измерения, погрешности моделирования вектора состояния и влияния на систему случайных факторов воспользуемся дискретным фильтром Калмана для оценки вектора состояния.

Текущие значения оценок вектора состояния $\hat{x}(k)$ определяются по итерационному правилу:

$$\begin{aligned}
\hat{x}(k+1) &= \hat{x}(k+1|k) + K(k)\cdot\left[y(k+1) - H\cdot\hat{x}(k+1|k)\right],\\
\hat{x}(k+1|k) &= A(k)\cdot\hat{x}(k) + B(k)\cdot u(k) + F(k)\cdot\bar{q}(k), \quad \hat{x}(0) = \bar{x}_0,\\
K(k) &= P_x(k+1|k)\cdot H^T\cdot\left[H\cdot P_x(k+1|k)\cdot H^T + R\right]^{-1},\\
P_x(k+1|k) &= A(k)\cdot P_x(k|k)\cdot A(k)^T + F(k)\cdot Q(k)\cdot F(k)^T,\\
P_x(k+1|k+1) &= \left[I - K(k)\cdot H\right]\cdot P_x(k+1|k),\\
P_x(0|0) &= P_{x_0},
\end{aligned} \qquad (13.41)$$

где y(.) – канал измерений ( ), H – матрица канал измерений, I – единичная матрица соответствующей размерности, $P_{x_0} = M\left\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\right\}$ - матрица ковариаций начального вектора состояния.

Не все компоненты вектора состояний могут быть измерены. Некоторые их них измеряются с некоторой погрешностью, поэтому оценим значения вектора состояний фильтром Калмана. Для этого воспользуемся каналом измерений).

H- матрица вида:

$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, если 2-я компонента вектора состояния - неизмеряемая величина;

$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, если 1-я компонента вектора состояния - неизмеряемая величина;

величина;

$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, если 1-я и 2-я компоненты вектора состояния – неизмеряемы.

В результате оценки вектора состояний фильтром Калмана получим, что оценка практически полностью совпадает со значениями вектора состояний, полученными в результате моделирования.

См. также дополнительный материал по к этой лекции в приложении В

Основная литература

1. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.
2. Бенькович Е.С. и др. Практическое моделирование динамических систем. -СПб: БХВ-Петербург, 2002

Дополнительная литература

3. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
4. Советов Б.Я., Яковлев С.А. Моделирование систем. – М.: Высшая школа. 2001

Тема 10 Методы идентификации с настраиваемыми адаптивными моделями

Лекция 14 Методы идентификации с настраиваемыми адаптивными моделями

Рассматриваем вопросы: методы идентификации с настраиваемыми адаптивными моделями; виды адаптивных моделей динамических объектов; модели линейные по параметрам, по сигналам; структурные схемы идентификации с применением адаптивных моделей; связь процесса настройки с характеристиками входного сигнала, условия независимости настроек параметров; виды критериев приближения моделей к объекту; упрощение синтезированных алгоритмов настройки адаптивных моделей.

Адаптивное управление — совокупность методов теории управления, позволяющих синтезировать системы управления, которые имеют возможность изменять параметры регулятора или структуру регулятора в зависимости от изменения параметров объекта управления или внешних возмущений, действующих на объект управления. Подобные системы управления называются адаптивными. Адаптивное управление широко используется во многих приложениях теории управления.

Классификация адаптивных систем. По характеру изменений в управляющем устройстве адаптивные системы делят на две большие группы:

- самонастраивающиеся (изменяются только значения параметров регулятора);
- самоорганизующиеся (изменяется структура самого регулятора).

По способу изучения объекта системы делятся **на поисковые и беспоисковые.**

В первой группе особенно известны экстремальные системы, целью управления которых является поддержание системы в точке экстремума статических характеристик объекта. В таких системах для определения управляющих воздействий, обеспечивающих движение к экстремуму, к управляющему сигналу добавляется поисковый сигнал. Беспоисковые адаптивные системы управления по способу получения информации для подстройки параметров регулятора делятся на:

- системы с эталонной моделью (ЭМ);
- системы с идентификатором, в литературе иногда называют, как системы с настраиваемой моделью (НМ).

Адаптивные системы с ЭМ содержат динамическую модель системы, обладающую требуемым качеством. Адаптивные системы с **идентификатором** делятся по способу управления на:

- прямой;
- косвенный(непрямой).

При косвенном адаптивном управлении сначала делается оценка параметров объекта, после чего на основании полученных оценок определяются требуемые значения параметров регулятора и производится их подстройка. При прямом адаптивном управлении благодаря учёту взаимосвязи параметров объекта и регулятора производится непосредственная оценка и подстройка пара-

метров регулятора, чем исключается этап идентификации параметров объекта. По способу достижения эффекта самонастройки системы с моделью делятся на

- системы с **сигнальной** (пассивной);
- системы с **параметрической** (активной) адаптацией.

В системах с сигнальной адаптацией эффект самонастройки достигается без изменения параметров управляющего устройства с помощью компенсирующих сигналов. Системы, сочетающие в себе оба вида адаптации называют комбинированными.

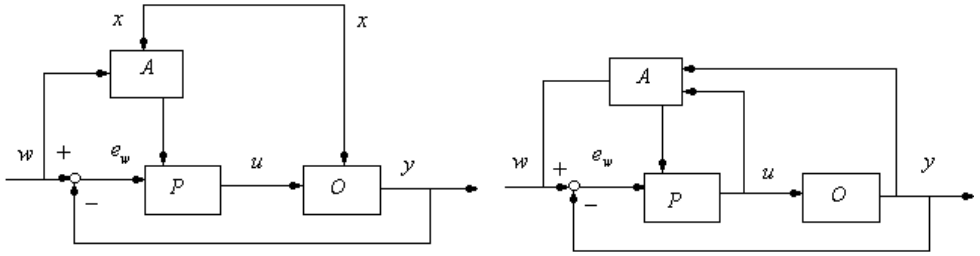
Применяется для управления нелинейной системой, и или системой с переменными параметрами. К примерам таких систем относят, например, асинхронные машины, транспортные средства на магнитной подушке, магнитные подшипники и т.п. Среди механических систем можно назвать инверсный маятник, подъемно транспортные машины, роботы, шагающие машины, подводные аппараты, самолеты, ракеты многие виды управляемого высокоточного оружия и т.п.

Таким образом, с целью повышения эффективности систем управления разрабатывается идея адаптивных систем управления. Главное отличие **адаптивных** систем управления от систем с фиксированными параметрами состоит в том, что они могут приспосабливаться (подстраиваться) к изменяющимся характеристикам объектов и протекающих в них процессов (см. рисунок 14.1). Существует два основных способа настройки регуляторов:

- настройка с прямой связью (адаптация по разомкнутому циклу);
- настройка с обратной связью (адаптация по замкнутому циклу).

Если известно, как должен настраиваться регулятор в зависимости от внешних входных факторов (доступных прямому измерению), то можно применять прямой метод настройки.

В условиях, когда невозможно оценить динамические свойства объекта непосредственно, приходится использовать настройку с обратной связью или адаптацию по замкнутому контуру. При этом необходимый минимум информации об объекте получают путем обработки измерений входных и выходных сигналов. Использование адаптации структурно равносильно введению второй обратной связи и соответственно второго замкнутого контура.



A – алгоритм настройки, P – регулятор, O – объект управления, w - вектор заданной переменной,
ew - ошибка управления ew =w – y.

Рисунок 14.1 - Структурная идентификация при адаптации

Все адаптивные регуляторы можно разделить на два класса: самооптимизирующиеся регуляторы и регуляторы с эталонной моделью.

Процесс адаптации в системах управления с регуляторами или регуляторами с эталонной моделью проходит в три этапа: идентификация объекта (эталонной модели) или системы управления в целом; расчет регулятора; настройка регулятора.

Идентификация в замкнутом контуре. Особенности оперативной идентификации дискретных моделей в замкнутом контуре (см. рисунок 14.2).

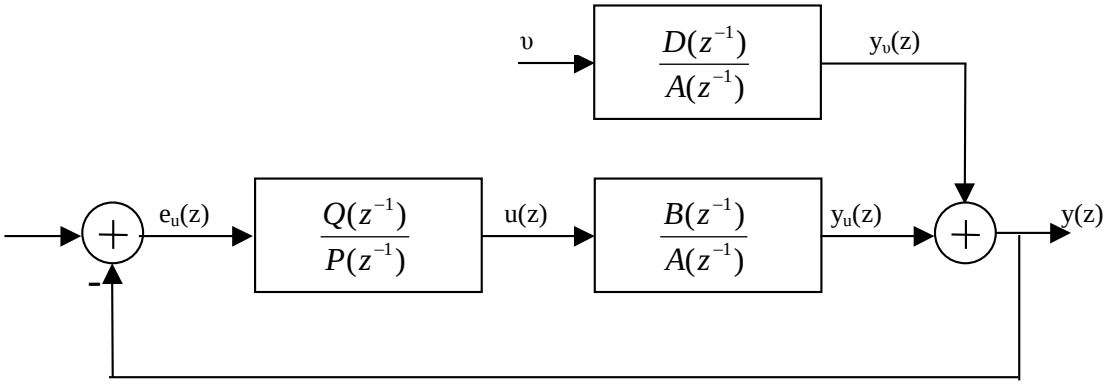


Рисунок 14.2 - Идентификация в замкнутом контуре

При такой постановке задачи возможно несколько вариантов подходов к идентификации модели.

1. Косвенная идентификация ОУ. Оценивают параметры замкнутого контура в целом и если параметры регулятора известны, характеристики объекта определяют путем аналитических преобразований.
2. Параметры модели объекта определяют непосредственно, минуя промежуток идентификации замкнутого объекта.
3. Измерению доступны только выходные системы.
4. Измеряют y и u.
5. Отсутствует внешнее возмущение на объект, тогда yu=y
6. На объект действуют внешние возмущения, но при идентификации их не учитывают.
7. Внешнее возмущение таково, что оно учитывается при идентификации параметров моделей.

При рассмотрении задачи будем предполагать, что объект управления в большом, а регулятор линейный, стационарен и не подвергается воздействию помех.

Передаточные функции объекта:

1. $Gp(z) = \dfrac{b_1 z^{-1} + ... + b_{mb} z^{-mb}}{1 + a_1 z^{-1} + ... + a_{ma} Z^{-ma}} \cdot Z^{-d}$
2. $Gpv(z) = \dfrac{1 + d_1 z^{-1} + ... + d_{md} z^{-md}}{1 + a_1 z^{-1} + ... + a_{ma} Z^{-ma}}$
3. $Gr(z) = \dfrac{g_0 + g_0 z^{-1} + ... + g_0 z^{-v}}{1 + p_1 z^{-1} + ... + p_M Z^{-M}}$

Если применить за вход u(z), а за выход y(z) определим передаточные функции:

$$W(z) = \frac{y(z)}{V(Z)} = \frac{Gpv}{1 + Gr + Gp} = \frac{D(z^{-1})P(z^{-1})}{A(z^{-1})P(z^{-1}) + B(z^{-1})Q(z^{-1})} = \frac{B(z^{-1})}{\alpha(z^{-1})} = \frac{(r)}{(l)}$$

l=max[ma+μ, mb+υ+d]

θTα,β=[α1... αl β1...βr]

θa,b=F/ θα,β /

Для однозначного решения задачи необходимо выполнение определенных условий, которые называются условия идентификации объекта в замкнутом контуре.

Объект выходящий в состав замкнутого контура управления называется параметрически идентифицируемым, если с помощью некоторого метода идентификации можно получить самостоятельные оценки его параметров.

$$\lim_{N \to \infty} E\{\theta(N)\} = \theta$$

Для объекта у которого выполняется это условие можно сформулировать 2 условия, которое называется условие идентифицируемости объекта в замкнутом контуре. Порядок моделей идентифицируемого объекта ma и mb и порядок числителя модели md должны быть известны заранее.

Если оценки параметров моделей Ai и Bi, число которых равно (ma + mb) определяется из уравнения по l–параметрам αi и если у полиномов D(z-1) и A(z-1) нет общих корней, то для получения однозначного решения надо выполнение неравенства:

$l \geq m_a + m_b$

Тогда, $\max\left|m_a + \mu, m_b + \upsilon + d\right| \geq m_a + m_b$ или $\max\left|\mu - m_b, \upsilon + d - m_a\right| \geq \tau$

$\upsilon > \mu - d - m_a - m_b$

$\upsilon \geq m_a - d$; $\mu \geq m_b$; $\upsilon > m_a$; $\mu \geq m_b$.

Если объект удовлетворяет этим условиям, то задача косвенной идентификации параметров объекта может быть решена однозначно.

Прямая идентификация параметров объекта в замкнутом контуре (см. рисунок 14.3).

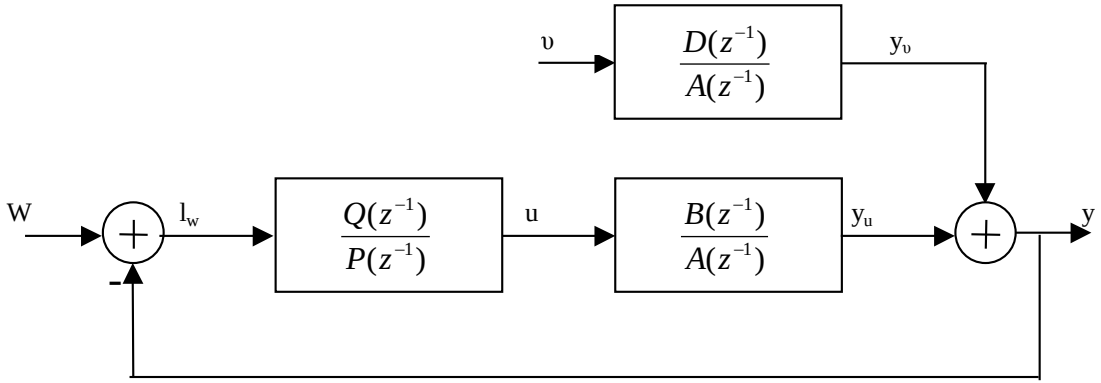


Рисунок 14.3 - Прямая идентификация параметров объекта в замкнутом контуре

W=0; lw=-y

$$\frac{U(z)}{V(z)} = \frac{-Gr \cdot Gpv}{1 + Gr \cdot Gp}$$

$$\frac{y(z)}{V(z)} = \frac{Gpv}{1 + Gr \cdot Gp}$$

$$\frac{y(z)}{U(z)} = \frac{y(z)V(z)}{U(z)V(z)} = -\frac{1}{Gr}$$

Представим полученные значения в виде разностных уравнений:

$$A(z^{-1})y(z) = B(z^{-1}) \cdot z^{-d}U(z) - D(z^{-1}) \cdot V(z)$$

$$[A(z^{-1}) \cdot P(z^{-1}) + B(z^{-1}) \cdot z^{-d}Q(z^{-1})]y(z) = D(z^{-1}) \cdot U(z)$$

$$Q(z^{-1})y(z) = -P(z^{-1})U(z)$$

$$A(z^{-1})P(z^{-1})y(z) - B(z^{-1})z^{-d}P(z^{-1})U(z) = D(z^{-1})P(z^{-1})V(z)$$

$$A(z^{-1})y(z) - B(z^{-1})z^{-d}U(z) = D(z^{-1})V(z)$$

Отличие этого подхода – U(z) вход объекта в замкнутом контуре непроизводная функция и является функцией y(z).

Идентификация в замкнутом контуре при внешних возмущениях на выходе регулятора (см. рисунок 14.4)

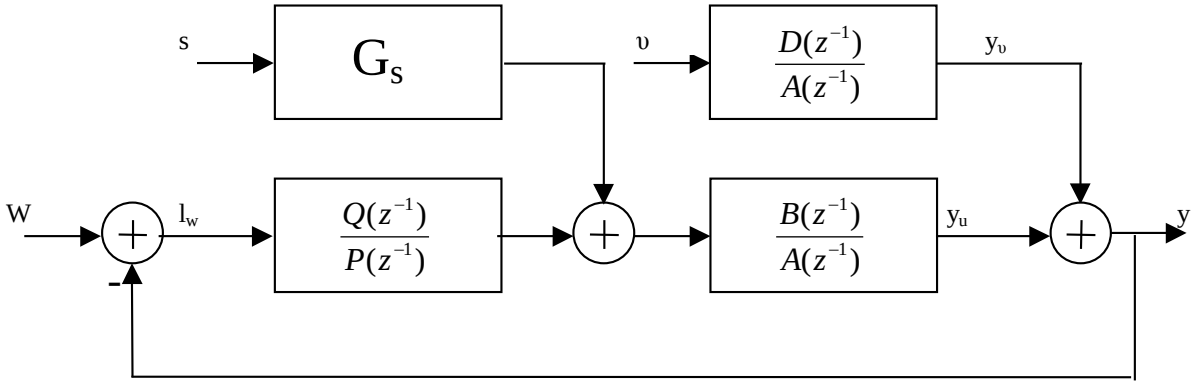


Рисунок 14.4 - Идентификация в замкнутом контуре при внешних возмущениях

Задача сводится к идентификации на основании разносного уравнения, причем форма их будет такой же, как и без начала помехи.

$$A(z^{-1})y(z) - B(z^{-1})z^{-d}U(z) = D(z^{-1})V(z)$$

$$U(z) = Ur(z) + Us(z)$$

При любых значениях полином ν и μ не является линейной комбинацией вектора данных, т.е. всегда возможна идентификация объекта в независимости от порядка ν и μ.

Рационально можно использовать схему, при которой на начальном этапе идентификации используют МНК, а при достижении некоторой стабильности используют более точные алгоритмы получения оценок.

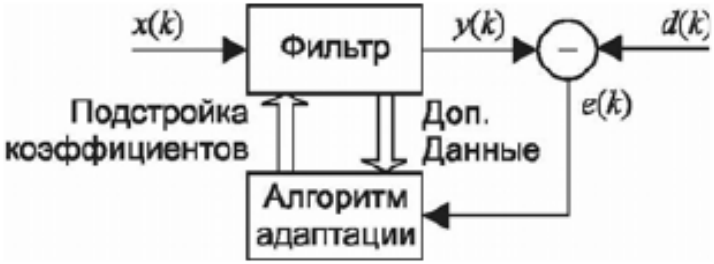


Рисунок 14.5 – Общая структура адаптивного фильтра

Основная литература

1. Фомин В Н., Фрадков О.Л., Якубович В.А. Адаптивное управление динамическими объектами. М : Наука, 1981
2. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.
3. Гроп Д. Методы идентификации систем. - М.: Мир, 1979

Дополнительная литература

4. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970. –312 с.
5. Бокс Дж, Дженкинс. Анализ временных рядов. Прогноз и управление. Выпуск 1. –М.: Мир, 1970, 406с

Тема 11 Методы идентификации нелинейных динамических характеристик

Лекция 15 Методы идентификации нелинейных динамических характеристик

В лекции рассмотрены методы идентификация нелинейных динамических объектов включая вопросы: применение гармонической линеаризации при идентификации нелинейных объектов; использование метода статистической линеаризации для идентификации нелинейных объектов; идентификация нелинейных объектов с использованием функциональных степенных рядов

Существует несколько методов идентификации нелинейных систем, некоторые из них:

1. метод прямого поиска;
2. аппроксимация нелинейности;
3. модель Гаммерштейна;
4. метод Винера;
5. двухэтапная процедура; п
6. применение гармонической линеаризации при идентификации объектов;
7. использование метода статистической линеаризации для идентификации нелинейных объектов; и
8. идентификация нелинейных объектов с использованием функциональных степенных рядов
9. градиентные методы идентификации нелинейных систем и др.

Сущность метода прямого поиска:

- нелинейную функцию f(x) преобразуют в линейную функцию /л(x);
- далее применяют любой метод идентификации линейных систем.

Допустим, что нелинейная модель объекта имеет вид:

$$y = a_0 x_1^{в1} x_2^{в2},$$

где x_1, x_2 - входные параметры, y - выходной параметр, а0, в1, в2 - искомые параметры. Выполняем логарифмирование и получаем:

$$\ln y = \ln a_0 + в_1 \ln x_1 + в_2 \ln x_2;$$

$$Z = \ln y, \quad \varepsilon_1 = \ln x_1, \quad \varepsilon_2 = \ln x_2, \quad A = \ln a_0;$$

$$Z = A + в_1 \varepsilon_1 + в_2 \varepsilon_2$$

Рассматриваем только положительные значения y.

Аппроксимация нелинейности. Таблично заданная функция (явно нелинейная) аппроксимируется с помощью полинома произвольным методом. Полученный полином и есть модель нашего объекта. Ограничения: функция должна быть непрерывна. Существует теорема Вейерштрасса, которая доказывает, что все нелинейности можно описать полиномом:

$$y = a_0 + a_1 x_1 + a_2 x_1^2 + в_1 x_2 + в_2 x_2^2$$

а) Замена линейной переменной и сведение к регрессии;
б) Применение интегральных формул.

Модель Гаммерштейна

Входной сигнал u(t) известен.

Если известна функциональная зависимость f(u(t)) - вид нелинейности, то вводим Z=f(u(t)). Идентификация сводится к определению параметров линейной части:

$$y(t) = A \cdot Z(t).$$

Если функциональная зависимость f(u(t)) не известна, то строится таблица этой нелинейной зависимости. По этой таблице любой интерпретируемой формулой получаем аппроксимирующий полином нелинейности f*(u(t)). Зная параметры аппроксимирующего полинома, вводим Z(t) = f(u(t)) и, снимая соответствующие ему y(t), решаем задачу идентификации:

y(t)= A · Z (t).

xi — a0 ^+ ai xi, функция является нелинейной.

x2 = J (xi)

Пример: Система приводится к следующему виду (см. рисунок 15.1):

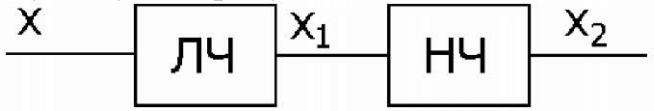


Рисунок 15.1 – Разбиение системы на линейную и нелинейную части

Схема нелинейной системы. Используя метод интерполяции, аппроксимируем полином

$$y = в_0 + в_1 x_1 + в_1 x_1^2 + \ldots + в_m x_1^m$$

Составляем обобщенный вектор:

$$V = \begin{bmatrix} 1 \\ x \\ 1 \\ x_1 \\ x_1^2 \\ \vdots \\ x_1^m \end{bmatrix} \begin{matrix} \rightarrow a_0 \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ \vdots \\ v_{m+3} \end{matrix}$$

Тогда искомая матрица:

$$A = \begin{bmatrix} a_0 & a_1 & в_0 & в_1 & \ldots в_m \end{bmatrix}$$

может быть получена по выражению:

где

$$A = YV^T \left[VV^T \right]^{-1},$$

$$Y = \begin{bmatrix} x_2(1) & x_2(2) \ldots & x_2(k) \end{bmatrix}$$

$$V = \begin{bmatrix} v_1(1) & \cdots & v_1(k) \\ v_2(1) & \cdots & v_2(k) \\ \vdots & & \vdots \\ v_{m+3}(1) & \cdots & v_{m+3}(k) \end{bmatrix}$$

В инженерной практике большое применение находят приближенные методы, основанные на замене действительных зависимости между входной и выходной переменными приближенными линейными. При этом линеаризацию необходимо производить так, чтобы учесть хотя бы приближенно нелинейные свойства звеньев, т.е. чтобы для линеаризованных элементов не выполняется принцип суперпозиции.

Линеаризация нелинейных характеристик путем разложения в ряд состоит в замене характеристики y = f(x) приближенной линейной зависимостью, определяемой двумя первыми членами разложения характеристики в ряд Тейлора . Пусть характеристика y = f(x) дифференцируема и входной сигнал x (f) мало отличается от некоторого среднего значения x0, тогда зависимость y = f(x) можно заменить приближенной:

$$y = f(x0) + f'(x-0)(x - x0). \qquad (15.1)$$

Замена нелинейной зависимости y =f(x) линейной геометрически представляет собой замену кривой y= f(x), касательной к ней в точке x0.

Действующие внешние возмущения можно представить как стационарные случайные функции x (г) с математическим ожиданием тх и центрированной случайной составляющей л(г):

$$x(t)\hat{}mx + x(t). \qquad (15.2)$$

В этом случае практически линеаризацию нелинейной характеристики целесообразно производить относительно центрированного входного случайного сигнала x(t), т.е. за центр разложения .x0 в (15.1) взять математическое ожидание тх входного сигнала x(1). В результате получается:

$$y (г) *./>У + ./' (тх) x (г). \qquad (15.3)$$

Таким образом, приближенная зависимость (15.3) линейна только относительно случайной составляющей x(t) входного сигнала и нелинейно относительно математического ожидания тх, поэтому принцип суперпозиции здесь неприменим.

Гармоническая линеаризация. В целом ряде практических задач приходится рассматривать воздействие на линейное звено гармонических колебаний

$$X(t)= A \sin \omega t = A \sin \psi,\ \psi=\omega t. \qquad (15.4)$$

Выходной сигнал нелинейного звена также будем периодическим, но не гармоническим.

Идея гармонической линеаризации состоит в том, что выходные периодические колебания y(t) разлагают в ряд Фурье и для дальнейших исследований ограничиваются рассмотрением лишь первых гармоник этого ряда. В этом случае нелинейная зависимость

y= y(t)=f(Asinψ),

заменяется приближенной

Y(t)=a0+asin ωt+bcosωt=a0+ q1x+q2x/ω,

где

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(A\sin\phi)d\phi$$

$$q_1 = \frac{1}{\pi A}\int_0^{2\pi} f(A\sin\phi)\sin\phi d\phi$$

$$q_2 = \frac{1}{\pi A}\int_0^{2\pi} f(A\sin\phi)\cos\phi d\phi$$

Статистическая линеаризация. Метод приближенной замены нелинейной характеристики эквивалентными в вероятностном смысле линейными зависимостями называется методом статистической линеаризации. В результате такой линеаризации нелинейная зависимость y=f(t) заменяется приближенной:

y(t)=kamx + k,x0(t).

где mx = const — математическое ожидание стационарного случайного сигнала на входе нелинейного элемента; x0(t) — центрированная случайная составляющая входного сигнала x (t).

Предполагается, что выходной стационарный случайный сигнал может быть представлен в виде:

y(t) = ту. + y/(t)

где ту — математическое ожидание y(t); y/(t) -центрированная случайная составляющая y(t).

Коэффициент к0 = ту/тх называется статистическим коэффициентом усиления нелинейного звена по математическому ожиданию. Коэффициент k1=±σy/ σ x .

Идентификация нелинейных объектов с использованием функциональных степенных рядов Производная функции определяется разностным отношением:

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z+\Delta z) - f(z)}{\Delta z}.$$

Если существует предел и не зависит от того, как стремится к нулю, то функция является аналитической. Следовательно в окрестности некоторой точки а можно ее разложить в ряд Тейлора:

$$f(z) = f(a) + \frac{z-a}{1!} f'(a) + \ldots + \frac{z-a}{n!} f^n(a) + \ldots$$

Если а=0 , то получаем ряд Маклорена:

$$f(z) = f(0) + \frac{z}{1!} f'(0) + \ldots + \frac{z^n}{n!} f^n(0) + \ldots$$

Учитывая , что последующие слагаемые выше второго достаточно малые величины, можно заменить функцию линейной частью разложения

Градиентные методы идентификации нелинейных систем. Для общей задачи минимизации функционала:

$$J = \frac{1}{2}\int_0^t [z(t) - y(t)]^2 dt \qquad \text{(15.5), (15.6)}$$

$$x(t_0) = x_0$$

при ограничениях: $\dot{x} = f[x(t), p(t), t]$, где f – нелинейная вектор – функция.

Случай p_i - const $\dot{p} = 0$ будем задаваться начальным значением p^i , i- номер итерации, и решив систему дифференциальных уравнений оценим величину функции штрафа J^i. Слегка изменяя p^i , для нового значения $p_j^i + h_j$ найдем штраф:

$J^i + \eta_j$, j = 1,n, n - число неизвестных коэффициентов

j-ю компоненту вектора-градиента функции штрафа можно приближенно оценить как

$$\frac{dJ}{dp_j^i} \approx \frac{(J^i + \eta_j) - J^i}{(p_j^i + h_j) - p_j^i} = \frac{\eta_j}{h_j} \qquad (15.7)$$

Повторяя эту процедуру для возмущений различных компонент вектора параметров, определим приближенное значение вектора-градиента dJ/dp^i Первое приращение вектора параметров в направлении наискорейшего спуска к минимуму функции штрафа составит

$$\Delta p^i = K^i \frac{dI}{dp^i}, \qquad (15.8)$$

и K^i – выбирается из условия

$$\min_{K^i} I(p^i - K^i \frac{dJ}{dp^i}) \qquad (15.9)$$

а новое приближение для вектора параметров определится как:

$p^{i+1} = pi + \Delta p^i$.

Простота приближенного метода позволила положить его в основу нескольких итерационных схем, однако трудно оценить ошибку, связанную с приближенным вычислением dJ/dp^i (процедура точного вычисления свелась бы к уже известным алгоритмам решения динамических задач). Приближенная процедура приводит к существенным ошибкам, особенно тогда, когда функция штрафа не очень чувствительна к изменению вектора параметров. Последнее, к сожалению, довольно часто имеет место, если измерения или наблюдения искажены помехой и имеются неизвестные входные сигналы.

Алгоритм вычислений следующий

1. Задаемся начальными значениями вектора параметров **p**.
2. Решаем дифференциальные уравнения (2)
3. Вычисляем значения функционала (1)
4. Вычисляем компоненты вектора-градиента функционала (1) по ф.(3).
5. Определяем новые значения p по ф.4 из условия (5)
6. Переходим к п.2 алгоритма, если компоненты вектора-градиента больше некоторой величины ε.

Если коэффициенты системы есть функции времени, т.е. p = p(t), то можно применить способы аппроксимации функции p(t).

1. Кусочно-постоянными функциями:

$$p^i = (p_i^1, p_i^2, ..., p_i^m)$$

2. Кусочно-линейными функциями вида, показанными на рисунке 15.2.

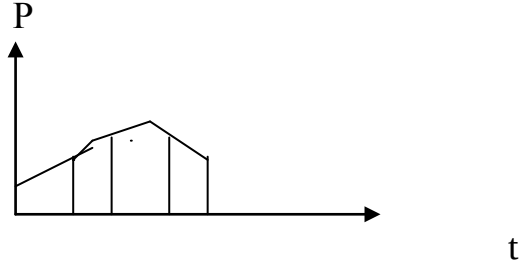


Рисунок 15.2 – Кусочно-линейная аппроксимация нелинейной функции

3. Полиномиальная гипроксимация.

$$p_i = \sum_{k=0}^{m_i} p_{i,k} t^k, \quad m_i = t_f / h_p,$$

4.Сплайн-аппроксимация (см. рисунок 15.3).

$$p_i = \sum_{k=0}^{m_i} p_{i,k} t^k, \quad m_i = t_f / h_{sp}, \text{ где } h_{2p} << h_p$$

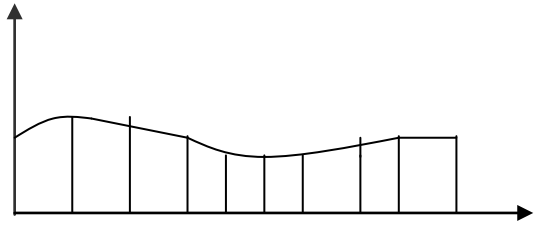


Рисунок 15.3 – Сплайн - аппроксимация нелинейной функции

Основная литература

1. Гроп Д. Методы идентификации систем. - М.: Мир, 1979
2. Иванов А. И. Быстрая идентификация нелинейных динамических объектов // CD-ROM “Бизнес-игры”. − М.: CD-ROM изд-во “Compact Book Publishing”, 1996. − 226 с.

Дополнительная литература

3. Бокс Дж, Дженкинс. Анализ временных рядов. Прогноз и управление. Выпуск 1. –М.: Мир, 1970, 406с
4. Бокс Дж, Дженкинс. Анализ временных рядов. Прогноз и управление. Выпуск 2. –М.: Мир, 1974, 193с

Приложение А. Условные обозначения

Объекты и системы в целом

U- вход (или вектор входов) объекта или системы, параметры управления;

X или иногдаY- выход (или вектор выходов) объекта или системы, параметры состояния;

$x^{(1)} = \frac{dx}{dt}$, $u^{(1)} = \frac{du}{dt}$ - производные по времени X и U;

V- возмущения (или вектор возмущений);W, W(p) - передаточная функция) объекта;

W^*, $W^*(p)$ - ПФ объекта без учета запаздывания;

W_1, $W_1(p)$ и т.д. - ПФ объекта в многоконтурных АСР;

$W_{зам}$, $W_{зам}(p)$ - ПФ замкнутой АСР;

i- порядковый номер W, R, элементов векторов U, X, V;

m- порядок полинома в знаменателе W(p);

n- порядок полинома в числителе W(p) вида, (n>m);

a_i - коэффициенты полинома в числителе W(p), i=1,2,…n;

b_i - коэффициенты полинома в знаменателе W(p), i=1,2,…m;

Частотные характеристики

ω; ω*; $ω_{раб}$ - частота; критическая частота; рабочая частота;

$ω_1$; $ω_к$; $Δ_ω$ - начальная, конечная частота; шаг по частоте;

N - количество точек по частоте;

φ ; $φ_{об}$; $φ_{рег}$ - фаза (ФЧХ); ФЧХ объекта; ФЧХ регулятора;

A ; $A_{об}$; $A_{рег}$ - амплитуда (АЧХ); АЧХ объекта; АЧХ регулятора;

j - мнимая единица ($\sqrt{-1}$); p, s - оператор Лапласа;

W(jω) ; R(jω) - АФЧХ объекта; АФЧХ регулятора;

ImW ; ReW - мнимая ; действительная часть АФЧХ объекта;

$Im_{Зн}$; $Re_{Зн}$ - мнимая ; действительная часть АФЧХ знаменателя ПФ;

$Im_{Чис}$; $Re_{Чис}$ - мнимая ; действительная часть АФЧХ числителя ПФ;

$φ_{зн}$; $A_{зн}$ - ФЧХ ; АЧХ знаменателя ПФ; $φ_{чис}$; $A_{чис}$ - ФЧХ; АЧХ числителя ПФ;

$φ_{зс}$; $A_{зс}$ - ФЧХ; АЧХ замкнутой АСР;

Цифровые модели и системы

j - номер точки по времени; N - количество точек по времени;

$X_j; \tilde{X}_j; U_j; E_j$ -выход, отфильтрованный выход, вход, ошибка в момент времени j;

T_o- период опроса датчика; шаг (такт) квантирования по времени ($T=2\pi/\omega$);

ω_o - частота опроса датчика ($\omega=2\pi/T$);

D - запаздывание, равное целому числу тактов квантирования по времени ($d=\tau/T_o$);

$Q_П$, $Q_И$, $Q_Д$ - П, И, Д параметры настройки цифрового регулятора;

Переходные процессы и их качество

T, t - текущее время; T_1 ; $T_к$; ΔT - начальное время; конечное время; шаг по времени;

N - количество точек по времени; ΔU - амплитуда ступенчатого возмущения на входе;

x_{o1}, x_{o1}, x_{o1} - динамические ошибки; $x_{ст}$ - статическая (установившаяся) ошибка.

Приложение Б. Глоссарий. Основная терминология

Термин	Толкование
DCS	Сокр. от англ. Distributed Control System — распределённая система управления)
HMI	Сокр. от англ. Human Machine Interface — * человеко-машинный интерфейс
SCADA.	Сокр. от англ. Supervisory Control And Data Acquisition. В АСУТП это диспетчерское управление и сбор данных. Под термином SCADA понимают инструментальную программу для разработки программного обеспечения систем управления технологическими процессами в реальном времени (АСУ ТП) и удаленного сбора данных (телемеханика). Реже термин SCADA-система используют для обозначения программно-аппаратного комплекса сбора данных (телемеханического комплекса)
Автоматизированная информационная система (АИС)	Человеко-машинная система (взаимосвязанная совокупность средств, методов и персонала), обеспечивающая сбор, хранение, обработку, поиск и выдачу необходимой потребителю информации.
Автоматизированная система управления (АСУ)	Человеко-машинная система, реализующая автоматизированный сбор и переработку информации, необходимой для принятия решений по управлению объектом. АСУ создают для оптимального управления в различных сферах деятельности.
Авторегрессионная модель	Модель, которая статистически описывает связи значений одного и того же показателя в различные моменты времени Y (t) = f (y(t-1)). Авторегресся часто используется в качестве линейной модели для прогнозирования
Адаптивная система	Самоприспосабливающаяся система - система, автоматически изменяющая алгоритмы своего функционирования и (иногда) свою структуру с целью сохранения или достижения оптимального состояния при изменении внешних условий.
Адаптивная система в кибернетике	Система, которая в процессе эволюции и функционирования демонстрирует способность системы к целенаправленному приспосабливающемуся поведению в сложных средах. Адаптивная система может приспосабливаться к изменениям как внутренних, так и внешних условий

Адаптивное управление	Совокупность методов теории управления, позволяющих синтезировать системы управления, которые имеют возможность изменять параметры регулятора или структуру регулятора в зависимости от изменения параметров объекта управления или внешних возмущений, действующих на объект управления. Подобные системы управления называются адаптивными. Адаптивное управление широко используется во многих приложениях теории управления.
Адекватность модели	Совпадение свойств (функций/параметров/характеристик и т. п.) модели и соответствующих свойств моделируемого объекта. Адекватностью называется совпадение модели моделируемой системы в отношении цели моделирования
Анализ временных рядов	Применим к одиночным или связанным временным рядам и позволяет выделять различные формы периодичности и взаимовлияния временных процессов, а также осуществлять прогнозирование будущего поведения временного ряда.
Аппроксимация	Математический метод, в основе которого лежит замена одних математических объектов другими, близкими к исходным в том или ином смысле, но более простыми.
Априори	Знание (информация), полученное до опыта и независимо от него
Винеровское оценивание	Задача нахождения импульсной характеристики линейной стационарной системы, которая минимизирует среднюю квадратическую ошибку между реальным y(t) и желаемым d(t) выходными сигналами при бесконечном времени наблюдения.
Временной ряд	Последовательность наблюдений за изменениями во времени значений параметров (атрибутов, признаков, показателей) некоторого объекта или процесса.
Дисперсионный анализ	(от латинского Dispersio – рассеивание / на английском Analysis Of Variance - ANOVA) применяется для исследования влияния одной или нескольких качественных переменных (факторов) на одну зависимую количественную переменную
Дисперсионный анализ	применяется для исследования влияния одной или нескольких качественных переменных (факторов) на одну зависимую количественную переменную (отклик).
Дисперсия	Статистическая характеристика, которая описывает меру разброса значений случайной величины относительно ее среднего значения (математического ожидания). В русскоязычной литературе дисперсия обозначается D[X], а в англоязычной var(X) (от англ. variance - дисперсия).
Доверительный интервал	Наиболее вероятный диапазон изменения наблюдений случайной величины
Идентификация	(в теории управления) — определение структуры системы и ее параметров путем анализа входных и выходных данных данной системы.

Идентификация систем	Математический аппарат для построения математических моделей динамической системы по измеренным данным. Математическая модель в данном контексте означает математическое описание динамики поведения какой-либо системы или процесса в частотной или временной области, к примеру, физических процессов (движение механической системы под действием силы тяжести), экономического процесса (реакция биржевых котировок на внешние возмущения) и т.п. В настоящее время эта область теории управления хорошо изучена и находит широкое применение на практике.
Идентифицируемость	Система идентифицируема, если по изменениям координат состояния системы можно определить ее параметры
Интеллектуальная система (ИС)	Техническая или программная система, способная решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти такой системы. Структура интеллектуальной системы включает три основных блока — базу знаний, решатель и интеллектуальный интерфейс]. Интеллектуальные системы изучаются наукой искусственный интеллект.
Интерполяция	Метод нахождения неизвестных промежуточных значений некоторой функции по имеющемуся дискретному набору ее известных значений.
Искусственный интеллект	Наука, которая входит в комплекс компьютерных наук, а создаваемые на её основе технологии к информационным технологиям. Задачей этой науки является воссоздание с помощью вычислительных систем и иных искусственных устройств разумных рассуждений и действий]
Кибернетика	Наука об общих закономерностях процессов управления и передачи информации в машинах, живых организмах и обществе впервые был предложен Норбертом Винером в 1948 году
Кластерный анализ	осуществляет разбиение объектов на заданное число удаленных друг от друга классов, а также строит дерево классификаций объектов посредством иерархического объединения их в группы (кластеры).
Корреляционный анализ	это группа статистических методов, направленная на выявление и математическое представление структурных зависимостей между выборками.
Коэффициент детерминации (R^2)	Показатель являющийся статистической мерой согласия, с помощью которой можно определить, насколько уравнение регрессии соответствует реальным данным. Чем ближе значение к 1 тем ближе модель к эмпирическим наблюдениям.

Марковский процесс	Случайный процесс, эволюция которого после любого заданного значения времянно́го параметра t не зависит от эволюции, предшествовавшей t, при условии, что значение процесса в этот момент фиксировано («будущее» процесса не зависит от «прошлого» при известном «настоящем»; другая трактовка«будущее» процесса зависит от «прошлого» лишь через «настоящее»).
Математическое ожидание	Наиболее вероятное (ожидаемое) значение случайной величины. Если для каждого значения, принимаемого случайной величиной X известна его вероятность, то ее математическое ожидание будет равно сумме произведений каждого значения на его вероятность.
Математическое программирование	Математическая дисциплина, посвященная теории и методам нахождения экстремумов (максимумов или минимумов) функций многих переменных при наличии дополнительных ограничений на эти переменные
Метод Байеса	Аналитический метод, который эффективно используется при сравнении гипотез. В этом методе вероятности всех возможных исходов эксперимента объединяются с гипотезами, известными до проведения эксперимента, и затем вычисляется вероятность того, что данные гипотезы подтвердятся в ходе эксперимента. В основе метода лежит теорема Байеса
Метод Монте-Карло	Синоним - метод статистических испытаний. Один из методов статистического моделирования, основанный на кибернетической идее “черного ящика”. Применяется в тех случаях, когда построение аналитической модели явления очень трудоемко или вообще невозможно (например, при решении сложных задач теории массового обслуживания и ряда других задач, связанных с изучением случайных процессов).
Метод наименьших квадратов (МНК)	Метод оценки параметров модели (например, регрессионной) на основании экспериментальных данных, содержащих случайные ошибки. Поэтому параметры модели должны быть рассчитаны так, чтобы минимизировать разницу между экспериментальными данными и теоретическими (вычисленными при помощи предложенной модели). Для этого и используется метод наименьших квадратов (МНК).
Моделирование	Процесс построения и исследования модели
Модель	Объект-заместитель объекта-оригинала, инструмент для познания, который исследователь ставит между собой и объектом и с помощью которого изучает некоторые свойства оригинала. В качестве модели выступает другой материальный или мысленно представляемый объект, замещающий в процессе исследования объект-оригинал. Соответствие свойств модели исходному объекту характеризуется адекватностью..
Модель имитационная (Симулятор)	Модель изучаемой системы (объекта, процесса), предназначенная для использования в процессе компьютерного моделирования

Мониторинг	Наблюдение за состоянием некоторого объекта или процесса, дающее возможность отслеживать их развитие, оценивать, оперативно выявлять результаты воздействия различных внешних факторов.
Наблюдаемость	Система наблюдаема, если все ее состояние можно непосредственно или косвенно определить по выходному вектору системы
Отрицательная обратная связь	Тип обратной связи, при которой выходной сигнал передается обратно на вход для погашения части входного сигнала. Отрицательная обратная связь делает систему более устойчивой к случайному изменению параметров. Методы математического анализа систем, в том числе и охваченных отрицательной обратной связью, подробно рассматриваются теорией автоматического управления.
Проверка статистических гипотез	Проверки статистических гипотез - класс базовых задач в математической статистике.
Программное обеспечение АСУ	Иерархический целевой комплекс программ для реализации целей и задач АСУ, а также обеспечения работы ее технических средств.
Регрессионные процедуры	Позволяют рассчитать модель, описываемую некоторым уравнением и отражающую функциональную зависимость между экспериментальными количественными переменными, а также проверяют гипотезу об адекватности модели экспериментальным данным. По полученным результатам можно оценить природу и степень зависимости переменных и предсказать новые значения зависимой переменной.
Серый ящик	Объект исследований, о внутреннем устройстве которого либо известно частично, либо существуют некоторые гипотезы. В отличие от черного, модели серого ящика учитывают помимо связей между реакциями и внешними воздействиями и те частичные сведения, которые известны о его внутреннем строении.
Система автоматического управления (САУ, АСР)	Система, в которой управление полностью выполняют технические средства, без непосредственного участия человека в этом процессе.
Случайный процесс	Множество (ансамбль) случайных функций xk(t), называемых возможными реализациями данного процесса. Если время t дискретно, то случайный процесс называется случайной последовательностью. Типичной случайной последовательностью является временной ряд. Наблюдаемая последовательность значений временного ряда – это случайная функция или одна из возможных реализаций случайного процесса.

Среднеквадратическое отклонение	Синонимы - стандартное отклонение. Статистическая характеристика распределения случайной величины, показывающая среднюю степень разброса значений величины относительно математического ожидания. Обозначается греческой буквой "сигма"
Стационарная система	Класс систем, для которых реакция на любое заданное входное воздействие одинакова для каждого момента его приложения (с точностью до времени запаздывания). Т.е. параметры такой системы не меняются во времени. более узком смысле стационарность означает, что при запаздывании входного сигнала по времени на некоторую величину выходной сигнал будет задерживаться на ту же самую величину. Стационарные системы выделяют в отдельный класс, поскольку их проще моделировать и анализировать, чем не стационарные. В частности, их динамика может быть описана единственной простой функцией (передаточной функцией).
Теорема Байеса	Теорема, которая позволяет определять вероятность наступления события (например, принадлежности наблюдения к одному из классов по значению признаков данного объекта) на основе некоторой частичной информации о нем.
Теория вероятностей	Раздел математики, изучающий закономерности случайных явлений, событий и величин, их свойства и операции над ними. Методы теории вероятности являются основой для математической статистики.
Теория игр	Раздел математики, изучающий математические модели принятия решений в конфликтных ситуациях, когда интересы участников либо противоположны, и тогда эти модели называются антагонистическими играми, либо не совпадают, хотя и неполярные, и тогда речь идет об играх с непротивоположными интересами.
Теория информации	Отрасль теории вероятности и математической статистики, определяющая понятие информации и методы ее представления, преобразования и передачи. Ее основателем является Клод Шеннон (1916-2001)
Теория массового обслуживания	Синоним - теория очередей. Область прикладной математики, использующая методы теории вероятностей и математической статистики. Система массового обслуживания состоит из некоторого числа единиц или каналов, работа которых заключается в выполнении поступающих по этим каналам заявок. Примерами таковых являются телефонные станции, ремонтные мастерские, супермаркеты и т.д. Вид и количество поступающих на эти системы заявок различны и, вообще говоря, случайны. Сложность описания процессов, протекающих в системе массового обслуживания, заключается в том, что система в случайные моменты времени переходит из одного состояния в другое, постоянно меняется число заявок, число занятых каналов, число заявок в очереди и т.д. Теория очередей описывает закономерности функционирования таких систем с целью повышения их

	эффективности (пропускной способности, издержек и т.д.) и является очень быстро развивающимся разделом теории вероятностей, т.к. ее применение на практике чрезвычайно широко.
Теория нечетких множеств	Раздел прикладной математики, посвященный методам анализа неопределенных данных, например качественного характера, в которых описание неопределенностей реальных явлений и процессов проводится с помощью понятия о множествах, не имеющих четких границ.
Теория принятия решений (ТПР)	Область исследования, вовлекающая понятия и методы математики, статистики, экономики, менеджмента и психологии; изучает закономерности выбора людьми путей решения разного рода задач, а также исследует способы поиска наиболее выгодных из возможных решений.
Теория случайных процессов	Является важнейшим разделом теории вероятностей и может использоваться в аналитических технологиях для анализа временных рядов.
Управляемость	Система управляема, если для любого момента времени при любых состояниях существует такое управление u, которое переводит начальное состояние системы в конечное
Уравнение Винера-Хопфа	Используется в кибернетике, при решении задачи выделения, фильтрации полезного сигнала из его смеси с шумом
Факторный анализ	Основной задачей является нахождение в многомерном пространстве первичных переменных (значения которых регистрируются в эксперименте), сокращенной системы вторичных переменных (факторов). Метод факторного анализа первоначально был разработан в психологии с целью выделения отдельных компонентов человеческого интеллекта из многомерных данных по измерению различных проявлений умственных способностей.
Фильтр Винера	Конечное или параллельное устройство оценки, использующее решение уравнения Винера–Хопфа, которое дает оценку в виде взвешенной суммы конечного числа наблюдаемых выборок.
Фильтр Калмана	Эффективный (имеющий способ гарантированно достигать результат за конечное число действий) рекурсивный фильтр, оценивающий вектор состояния динамической системы, используя ряд неполных и зашумленных измерений. Назван в честь Рудольфа Калмана.
Функционал	Переменная величина, заданная на множестве функций, т. е. зависящая от одной или нескольких функций. Есть и другое определение: функция, аргументы которой также представляют собой функции некоторых переменных. Наконец, по третьему и весьма распространенному определению – то же, что и функция.
Функция распределения	В математической статистике это функция, определяющая вероятность того, что случайная величина примет некоторое значение или будет принадлежать заданному диапазону значений
Целевая функция	В широком смысле целевая функция есть математическое выражение некоторого критерия качества одного объекта (решения, процесса и т.д.) в сравнении с другим. Синоним: критерий оптимальности

Цифровой фильтр	Любой фильтр, обрабатывающий цифровой сигнал с целью выделения и/или подавления определённых частот этого сигнала. В отличие от цифрового аналоговый фильтр имеет дело с аналоговым сигналом, его свойства недискретны, соответственно передаточная функция зависит от внутренних свойств составляющих его элементов.
Черный ящик	Термин, используемый в точных науках (в частности, системотехнике, кибернетике и физике) для обозначения системы, механизм работы которой очень сложен, неизвестен или неважен в рамках данной задачи. Такие системы обычно имеют некий «вход» для ввода информации и «выход» для отображения результатов работы. Состояние выходов обычно функционально зависит от состояния входов
Шум	Аддитивная или мультипликативная составляющая сигнала, имеющая случайный характер и не несущая полезной информации. Различают белый шум (имеющий непрерывный равномерный спектр) и цветной шум
Экспертная система	Система искусственного интеллекта (программно-аппаратный комплекс), разработанная на основе специальных знаний экспертов об определенной предметной области (информатике, математике, медицине, геологии и т. д.).
Экстраполяция	Приближённая оценка значений функции f (x) в точках x , лежащих вне отрезка [x 0 , x n], по её значениям, расположенным внутри данного отрезка. Иными словами, это разновидность
Эргодический процесс	Случайный процесс эргодичен, если с вероятностью, равной единице, все его статистические характеристики можно предсказать по одной реализации из ансамбля процесса с помощью усреднения по времени

Приложение В. Матрица наблюдаемости и фундаментальная матрица

(материалы с форума к теме лекции №13)

Студентам каких-либо радиотехнических специальностей часто читают и даже пытаются объяснять, а иногда и вдолбить, такой алгоритм фильтрации, как рекуррентный байесовский фильтр, типа фильтр Калмана. Сам знаю, что на первых порах такой алгоритм крайне тяжело воспринимаются. И, как правило, такой алгоритм студентам преподносят на матером техническом языке. И пока ты сидишь и вспоминаешь значение того или иного словечка, понимание сути этого алгоритма до тебя так и не доходит. Единственное, что в итоге запоминается, что этот фильтр придуманный Калманом очень хорошая штука, которая почти нигде не используется, потому что он трудно реализуем в аппаратуре.

Действительно, реализация фильтра Калмана на борту космического аппарата или в навигационной аппаратуре потребителя была затруднительной, так как требовался достаточно мощный процессор и приличный объем памяти, т.е. значительное увеличение габаритов аппаратуры и, соответственно, массы.

Как правило, использовался наиболее простой (и в реализации) метод наименьших квадратов (МНК), алгоритм фильтрации по полной выборке.

По сравнению с фильтром Калмана, МНК не использует априорной информации о движении объекта и оценка состояния выполняется на основе выборки измерений. Т.е. до получение первой качественной оценки состояния объекта помимо времени на получение информации от навигационных спутников (альманахи, эфемериды и т.д.) прибавляется время на накопление измерений. Навигаторы с МНК плохо определяют именно движущиеся объекты, но если, с уже включенным приемником, вы постоите минут 5 на одном месте, тогда навигатор может показать ваше местоположение с точностью до сантиметров.

Фильтр Калмана, в свою очередь, после получения первых измерений, в состоянии достаточно качественно сделать оценку движение объекта. Априорная информация о движении объекта в алгоритме выражается в матрицах наблюдаемости (Н) и фундаментальной матрице (Ф).

Итак. Вот, собственно, сам фильтра Калмана (точнее, одно из его представлений):

$$
\begin{aligned}
q^{*}_{i+1} &= \hat{q}_{i+1} + K_{i+1}[R_{И(i+1)} - R_{0(i+1)}(\hat{q}_{i+1}, Q_{i+1})], \\
K_{q(i+1)} &= [I - K_{i+1}H_{i+1}]K_{q0(i+1)}[I - K_{i+1}H_{i+1}]^{T} + K_{i+1}W_{i+1}K^{T}_{i+1}, \\
K_{i+1} &= K_{q0(i+1)}H^{T}_{i+1}[H_{i+1}K_{q0(i+1)}H^{T}_{i+1} + W_{i+1}]^{-1}, \\
K_{q0(i+1)} &= \Phi_i K_{qi}\Phi_i^{T} + V_i.
\end{aligned}
\qquad (1)
$$

q^*_{i+1} - апостериорный вектор состояния, размерностью 6x1 ;

$\hat{q}_{i+1}$ - априорный вектор состояния, размерностью 6x1 ;

K_{i+1} - матрица коэффициентов усиления, размерностью 6xn ;

$R_{И(i+1)}$ - "истинный" вектор измерений, размерностью nx1 ;

$R_{0(i+1)}(\hat{q}_{i+1}, Q_{i+1})$ - навигационный (расчетный) вектор измерений, размерностью nx1 ;

$K_{q(i+1)}$ - априорная матрица ковариаций среднеквадратических отклонений, характеризующих точность оценки вектора состояния, размерностью 6x6 ;

H_{i+1} - матрица наблюдаемости, размерностью nx6 ;

I - единичная матрица, размерностью 6x6 ;

W_{i+1} - матрица ошибок измерения, размерностью nxn ;

$K_{q0(i+1)}$ - апостериорная матрица ковариаций среднеквадратических отклонений, характеризующих точность оценки вектора состояния, размерностью 6x6 ;

Φ_i - фундаментальная матрица, размерностью 6x6 ;

V_i - матрица влияния внешних возмущений на вектор состояния.

n – количество измерений.

Пример взят из книги Сетевые спутниковые радионавигационные системы (http://isns.ru/download/SSRNS_GPS_.djvu) Для простоты рассмотрим фильтр для вектора состояния q из шести элементов: три проекции радиус-вектора (радиус-вектор с началом в центре Земли, направленный в центр масс объекта) и три проекции вектора скорости объекта.

Для студентов, изучающих такой фильтр, трудности доставляют, в основном, именно эти две матрицы: **матрица наблюдаемости и фундаментальная матрица**.

Попробуем объяснить построение этих матриц как можно проще (в пределах разумного).

Начнем с матрицы наблюдаемости.

Предположим, что мы находимся на борту космического аппарата. Помимо нас на борту есть такие навигационные средства, как ИНС (инерциальная навигационная система, пусть будет платформенная) и спутниковая навигационная система (какой-нибудь GPS/ГЛОНАСС-приемник). Предположим, что в каждый момент времени нам с помощью ИНС, известны все параметры нашего движения q={Rx, Ry, Rz, Vx, Vy, Vz,}. И вроде бы всё замечательно. Но интуиция, а может эхо давно забытых лекций, подсказывает нам, что эти платформенные ИНС имеют обычай «уходить», т.е. накапливать ошибку. И, значит, со временем станет не очень понятно, как мы вообще движемся. Чтобы скорректировать ИНС к нам на помощь приходит достаточно точный GPS/ГЛОНАСС – приемник. (Правда слегка недоделанный: Нам еще самим нужно будет высчитывать псевдодальности и псевдоскорости.) Но и приемник не может выдавать абсолютно точные оценки. А если вспомнить, что наш космический аппарат находится в агрессивной среде, где на него воздействуют нецентральные гравитационные силы Земли, Луны и Солнца.. Может быть там еще есть и верхняя атмосфера Земли. Короче говоря, сплошные ошибки и погрешности. Но чтобы отфильтровать эти ошибки

нам как раз пригодится фильтр Калмана. Но вот беда: в алгоритме фильтрации нужно реализовать некую матрицу наблюдаемости (H), которая связывает измерения GPS/ГЛОНАСС – приемника и вектор состояния, получаемый с помощью ИНС. В, общем-то, это некая матрица перехода связывающая наши реальные(по ИНС) местоположение и скорость с измерениями GPS/ГЛОНАСС – приемника.

Есть такая функция y=A*x, так вот, если сопоставить эту функцию с тем, что я написал выше, получится: [вектор состояния(навигационный (по ИНС))] = A* [измерения приемника].

Так вот A – это и есть H, матрица наблюдаемости.

Размерность этой матрицы nx6. 6 - это число элементов вектора состояния, n – число измерений. Предположим, что наш GPS/ГЛОНАСС – приемник наблюдает в данный момент только 3 навигационных спутника. Тогда от этих спутников мы можем получить максимум 6 измерений: псевдодальность и псевдоскорость от одного спутника, псевдодальность и псевдоскорость от другого и т.д. Следовательно, размерность матрицы 6x6.

Вот нужно теперь определить эту матрицу. Роясь в карманах, находим маленькую книжечку. Нам её, незаметно, засунули в центре управления полетами. В ней находим главу «Модель измерений». Как правило, такая модель описывается формулой(в данном случае двумя: для псевдодальности и псевдоскорости). Из чего эти формулы состоят, разбирать не будем. Предположим, что эти формулы есть, и все их компоненты нам понятны.

Заполним первую строку матрицы наблюдаемости. Если перевести текст по заполнению матрицы, из умных книг, и адаптировать для его для среднего студента получится приблизительно следующее: возьмем формулу модели измерения для псевдодальности и продифференцируем сначала по Rx, таким образом получим первый элемент матрицы наблюдаемости, потом по Ry – получим второй элемент матрицы наблюдаемости, первой строки и т.д. до Vz.

Для заполнения второй строки необходимо взять уравнение модели измерения для псевдоскорости и проделать с ней такие же операции.

Таким образом, вперемежку, аналитически заполнится вся матрица наблюдаемости. Останется только запрограммировать её.

Правильно ли вы составили матрицу наблюдаемости? Её строки, по сути, будут единичными векторами, направленными на навигационный спутник. Так что, если модуль строки матрицы наблюдаемости равен 1, то скорее всего вы правильно продифференцировали.

Теперь фундаментальная матрица. Эта матрица, описывающая динамику состояния объекта. На пальцах её суть можно объяснить следующим образом:

Ваш вестибулярный аппарат задаёт, связанную с вами, систему координат.

Вот вы стоите неподвижно в углу. Угол комнаты задает инерциальную систему координат. В какой-то момент вы решили развернуться на 180 градусов и делаете это(долго разворачиваетесь, секунд 5). Так вот фундаментальная матрица - это некое описание вашего движения, относительно угла комнаты, с момента времени начала вашего поворота до его окончания, т.е. движение связанной системы координат(в проекциях на оси инерциальной) относительно инерциальной в заданном интервале времени.

Более менее разобрались, что это за матрица такая. Теперь её нужно заполнить…

Примени метод Рунге-Кутта для формирования фундаментальной матрицы.

Заполнение фундаментальной матрицы.

1. Берется “истинный” вектор состояния на момент времени начала вашего поворота и численно интегрируется на интервале времени в 5 секунд (интегрируется формула, выражающая модель нашего движения).Полученный вектор состояния запоминается. Назовем его X.
2. Берется “истинный” вектор состояния до вашего поворота и к первой компоненте вектора прибавляется некоторое приращение(100 метров). Пример: Rx+100 метров (если Rx в метрах). Остальные компоненты вектора не трогаем. Полученный вектор состояния интегрируется на 5 секунд. В результате получаем вектор X*.
3. Покомпонентно из вектора состояния X вычитаем X* и делим на приращение(также покомпонентно). В результате получаем 1-й столбец фундаментальной матрицы.
4. Для получения остальных столбцов фундаментальной матрицы выполняются 2,3 пункты. (Номер элемента вектора состояния, к которому осуществляется приращение, соответствует номеру столбца фундаментальной матрицы).
5. Если модель движения выражается формулой для кругового движения и ваш космический аппарат представлен материальной точкой, тогда очевидно на протяжении всего вашего полета, пересчитывая фундаментальную матрицу её значения будут практически постоянными.

.

Методическое обеспечение дисциплины и ТСО.

Основная литература

1. Эйкхофф П. Основа идентификации систем управления. - М.: Мир, 1975.
2. Ахназарова С.Л., Кафаров В.В. Методы оптимизации эксперимента в химической технологии: Учебное пособие для вузов. - 2-е изд., перераб. и дополненное. -М.: Высшая школа, 1985. -327с.
3. Инков А.М. Моделирование и идентификация объектов управления. Методические указания к выполнению лабораторных работ для студентов спец. 050702. Шымкент, ЮКГУ, 2010 г., -78 с.
4. Цыпкин Я.З. Информационная теория систем. - М.: Наука, 1995.
5. Льюнг Я. Идентификация систем. Теория для пользователя. - М.: Наука, 1991.
6. Гроп Д. Методы идентификации систем. - М.: Мир, 1979.
7. Советов Б.Я., Яковлев С.А. Моделирование систем. Практикум. - М.: Высш. школа, 1999
8. Советов Б.Я., Яковлев С.А. Моделирование систем. – М.: Высшая школа. 2001

Дополнительная литература

9. Современные методы идентификации систем. Пер. с англ. /Под ред. П.Эйкхоффа. - М.: Мир, 1983.
10. Толчеев В.О., Ягодкина Т.В. Методы идентификации линейных одномерных динамических систем. -М.: Изд-во МЭИ, 1997
11. Практикум по автоматике и системам управления производственными процессами: учеб. пособие для вузов /под ред. И.М.Масленникова. -М.: Химия, 1986. -336с.
12. Построение математических моделей химико-технологических процессов. Под ред. Дудникова Е.Г. - Л.: Химия, 1970.
13. Райбман Н.С., Чадеев В.М. Построение моделей производства. - М.: Энергия, 1975.
14. Цирлин А.М., Балакирев В.С., Дудников Е.Г. Вариационные методы оптимизации управляемых объектов. М.: Наука, 1984.
15. Балакришнан А. Теория фильтрации Калмана: Пер. с англ. –М.:, Мир, 1988 -108с
16. Рузинов Л.П. Статистические методы оптимизации химических процессов. -М.: Химия, 1972.
17. Бенькович Е.С. и др. Практическое моделирование динамических систем. -СПб: БХВ-Петербург, 2002.
18. Боровиков В.П., Боровиков И.П. Statictica - статистический анализ и обработка данных в среде Windows,- М.: Инф. издат.дом «Филинъ», 2004
19. Фомин В Н., Фрадков О.Л., Якубович В.А. Адаптивное управление динамическими объектами. М : Наука, 1981.
20. Бокс Дж, Дженкинс. Анализ временных рядов. Прогноз и управление. Выпуск 1. –М.: Мир, 1970, 406с.
21. Бокс Дж, Дженкинс. Анализ временных рядов. Прогноз и управление. Выпуск 2. –М.: Мир, 1974, 193с.
22. Спиди Л., Браун Р, Гудвин Дж. Теория управления. Идентификация и оптимальное управление. –М.: Мир, 1973, 247с.
23. Сейдж Э.П., Уайт Ч. Оптимальное управление системами: Пер. с англ. /Под ред. Левина Б.Р.. –М.: Радио и связь, 1982. -398с.
24. В.Н. Фомин. Рекуррентное оценивание и адаптивная фильтрация. –М.: "Наука", 1984

25. Семенов А. Д., Артамонов Д. В., Брюхачев А. В. Идентификация объектов управления: Учебн. пособие. - Пенза: Изд-во Пенз. гос. ун-та, 2003.- 211 с
26. Иванов А. И. Быстрая идентификация нелинейных динамических объектов // CD-ROM “Бизнес-игры”. − М.: CD-ROM изд-во “Compact Book Publishing”, 1996. − 226 с.
27. Mohinder S. Grewal Kalman Filtering: Theory and Practice Using MATLAB, Second Edition, California State University at Fullerton, John Wiley & Sons, Inc.,2003
28. Bendat, Julius S. and Allan G. Piersol: Random Data - Analysis and Measurement Procedures (2nd edition); New York: Wiley, 2006
29. Bendat, Julius, and Allan Piersol: Engineering Application of Correlation and Spectral Analysis (2nd edition); New York: Wiley, 2003
30. Desrochers, Alan A. Modeling and Control of Automated Manufacturing Systems-, Washington DC: IEEE Computer Society Press, 1990.
31. Ljung, Lennart, and Torkel Glad: Modeling of Dynamic Systems; Englewood Cliffs, NJ: Prentice Hall, 1994
32. David G. Introduction to Dynamic Systems — Theory, Models and Applications-, New/ York: Wiley, 1979

Компьютерные программы

1. Инков А.М. Программа для ПК «ЛАБ1. Виртуальная лабораторная работа №1 по дисциплине МИОУ. Определение статической характеристики объекта»
2. Инков А.М. Программа для ПК «ЛАБ2. Виртуальная лабораторная работа №2 по дисциплине МИОУ. Однофакторный дисперсионный анализ объекта»
3. Инков А.М. Программа для ПК «ЛАБ3. Виртуальная лабораторная работа №3 по дисциплине МИОУ. Статистическая оценка параметров распределения случайных величин. Проверка гипотез»
4. Инков А.М. Программа для ПК «ЛАБ4. Виртуальная лабораторная работа №4 по дисциплине МИОУ. Планирование эксперимента»
5. Инков А.М. Программа для ПК «ЛАБ4а. Виртуальная лабораторная работа №4А по дисциплине МИОУ. Исследование методом ПФЭ»
6. Инков А.М. Программа для ПК «ЛАБ5. Виртуальная лабораторная работа №5 по дисциплине МИОУ. Исследование динамики объектов управления»
7. Инков А.М. Программа для ПК «Оценивание состояния системы дискретным фильтром Калмана»
8. MATLAB, Mahtcad, Statistica VisSim, LabView, SPSS и др.

Плакаты (имеется аналогичный раздаточный материал и слайды)

1 Плакат (слайд) по каждой из лекций

Пособия в электронном виде, имеющиеся на кафедре

1. На кафедре АТУ ЮКГУ имеются в электронном виде книги [1-8, 16, 19- 22, 24-25, 27]
2. Все учебно-методические материалы по дисциплине на электронных носителях.

Александр Михайлович Инков

ТЕХНОЛОГИИ ПРОГРАММИРОВАНИЯ

Конспект лекций
для студентов специальности
5В070200 (050702) «Автоматизация и управление»

Редактор _____

Подписано в печать «__» ____2010.

Формат бумаги X^xY 1/16
Бумага типографская. Печать офсетная. Объем 7 п.л.
Тираж ___ экз. Заказ № ___

© Издание Южно-Казахстанского государственного университета
им. М.Ауезова

Издательский центр ЮКГУ им. М.Ауезова, г. Шымкент, пр. Тауке-хана, 5