

Министерство образования Российской Федерации

САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ

А.Я. Городецкий

ИНФОРМАЦИОННЫЕ СИСТЕМЫ

ВЕРОЯТНОСТНЫЕ МОДЕЛИ И
СТАТИСТИЧЕСКИЕ РЕШЕНИЯ

Учебное пособие

Санкт-Петербург

Издательство СПбГПУ

2003

УДК 681.518:519.2 (075.8)

ББК 32.973.202 я 73

Г 701

Городецкий А.Я. Информационные системы. Вероятностные модели и статистические решения. Учеб.пособие. СПб: Изд-во СПбГПУ, 2003. 326 с.

ISBN 5-7422-0381-X

Пособие соответствует дисциплинам государственного образовательного стандарта ЕН.Ф.01.7 «Методы оптимизации» и ЕН.Ф.01.8 «Теория принятия решений» по направлениям «Системный анализ и управление», «Информатика и вычислительная техника» при подготовке бакалавров, магистров и дипломированных специалистов по компьютерным сетям и основанным на них системам управления, а также по специальностям «Информационные системы и технологии» и «Вычислительные машины, комплексы, системы и сети».

Пособие также может быть полезно для инженеров и научных работников, интересующихся проблемами статистического оценивания и управления.

В учебном пособии в рамках вероятностно-статистических подходов представлены формируемые в информационных системах алгоритмы анализа и синтеза, которые используются в задачах измерения, моделирования и управления. Рассмотрены методы обнаружения, линейной и нелинейной фильтрации, интерполирования и прогнозирования, идентификации и статистического синтеза управления динамических систем с использованием понятий пространства состояний и марковских моделей процесса.

Ил.8. Библиогр.: 79 назв.

Печатается по решению редакционно-издательского совета Санкт-Петербургского государственного политехнического университета.

ISBN 5-7422-0381-X

© Санкт-Петербургский государственный
политехнический университет, 2003
© Городецкий А.Я., 2003

ПРЕДИСЛОВИЕ

В инфраструктуре современного информационно-индустриального общества информационные системы занимают одно из ключевых мест. Это вызвано возрастающей ролью информации в наукоемком промышленном производстве. Его эффективность из-за требований резкого снижения материало- и энергозатрат, а также экологической безопасности определяется не только количеством переработанного сырья, но и объемом овестьствовленной информации. Информация в современных условиях выступает, как ресурс, позволяющий минимизировать расходы других ресурсов (сырьевых, материальных, энергетических, трудовых, финансовых и т.д.). Требования к качеству информации в современных условиях настолько возросли, что, вообще говоря, трудно представить нормальное функционирование общества без соответствующего информационного обеспечения. Например, в условиях рыночной экономики особое значение приобретают достоверность и полнота информации, без которых невозможна маркетинговая, финансовая и инвестиционная деятельность. Кроме того, современные информационные технологии открыли новые возможности в организации научной деятельности, в создании современных форм обучения на базе технических средств информационных систем. Таким образом, решаемые информационными системами задачи разнообразны, зависят от конкретных сфер человеческой деятельности и практическая их направленность определяется поставленными на инженерном уровне целями управления, требованиями информационного обеспечения и т.д.

Имеющиеся затруднения в классификации и решении этих задач проистекают как по причине разнородности технической базы информационных технологий и разнообразия выполняемых этими системами функций, влияния структур управления с учетом реальных режимов и возмущений, в том числе и случайных, так и вследствие разных подходов при исследовании информационных систем и неоднозначности оценок эффективности их функционирования. Методы решения указанных задач, а также проведение наиболее перспективной технической политики в области создания новых

информационных технологий и аппаратно-программных средств передачи и обработки данных были бы наиболее плодотворны, если бы опирались на общую теорию информационных систем. Пока не представляется возможным рассматривать эту теорию как сложившееся научное направление и ожидать в ближайшем будущем подробное и всеобъемлющее ее изложение. С другой стороны, бурный научно-технический прогресс в информатизации общества диктует необходимость формирования и развития этой теории. Одним из возможных путей решения этой проблемы является самостоятельное развитие отдельных разделов теории с последующей их интеграцией. Следствием этого интеграционного процесса станет создание новых знаний и методик, что, в свою очередь, послужит толчком для следующего этапа развития теории информационных систем.

Предлагаемое учебное пособие посвящено одному из важных относительно самостоятельных разделов теории информационных систем, отражающему содержательный аспект информации - анализу вероятностных моделей процессов и формированию на их базе методов обнаружения и оценивания, а также статистического синтеза управления динамических систем.

Обучаемым по информационно-телекоммуникационным специальностям студентам вузов в настоящее время отводится много времени для изучения как общетеоретических, так и специальных учебных курсов по теориям вероятностей и оптимальных статистических решений. Это оправдано, так как теория динамических систем развивается в направлении более полных и глубоких подходов в анализе функционирования и эффективности, а также в методах синтеза этих систем при учете реальных режимов, возмущений и помех. Цель пособия – ознакомить читателя с современными вероятностно-статистическими методами анализа и синтеза динамических систем. Учебное пособие отличается от других аналогичных изданий по информационным системам. В нем круг рассматриваемых проблем, характер решаемых задач не выходит за рамки вероятностно-статистических подходов, базирующихся на методах теории марковских процессов. Такой подход позволяет единообразно и экономно анализировать и синтезировать динамические системы различного назначения.

Учебное пособие состоит из пяти разделов: «Методы описания случайных процессов», «Модели процессов в информационных системах», «Корреляционный анализ процессов», «Статистические решения и оптимальная обработка сигналов» и «Методы статистического синтеза оптимального управления динамических систем».

В первом и втором разделах даются краткие сведения о случайных величинах, процессах и методах их описания с помощью распределения вероятностей, плотностей вероятностей, характеристических и производящих функций и функционалов, а также других статистических характеристик. Обсуждаются математические модели и определяются статистические характеристики винеровского, гауссовских, марковских, точечных (случайных потоков) и фрактальных процессов, аппроксимирующих реальные сигналы в информационных системах.

В третьем разделе сформулированы задачи корреляционного анализа, основным содержанием которых является изучение прохождения (преобразования) случайных сигналов через линейные и нелинейные устройства (системы) с известными характеристиками, определение статистических характеристик в рамках корреляционной теории и оценка качества работы устройств (систем). Корреляционный анализ линейных и нелинейных систем осуществляется с единых позиций, поскольку для нелинейных систем применяется приближенный метод статистической линеаризации. Подробно рассмотрены вопросы применения этого метода при анализе точности работы нелинейных систем.

Четвертый раздел посвящен методам оптимальной обработки сигналов: задачам обнаружения и различения гипотез, оценивания параметров, интерполирования, фильтрации, прогнозирования и индетификации. Изложение материала проводится по единой методологии, именуемой в пособии совместным обнаружением и фильтрацией случайных процессов. В вычислительных сетях информационных систем для предотвращения возможных перегрузок оптимизации подвергаются сетевые процессы.

В пятом разделе рассматриваются методы синтеза оптимального управления динамических систем. Исследуются задачи статистического синтеза оптимального закона управления. Для определения этого закона в соответствии с постановкой задачи

выбирается критерий экстремального значения апостериорного среднего функционала качества, зависящего от вероятностных характеристик фазовых координат системы и управления. В общем виде решение указанной задачи достаточно сложно. Наиболее конструктивные физически содержательные результаты могут быть получены методами теории условных марковских процессов на основании достаточных статистик (оценок фазовых координат системы). Из всего многообразия подходов оптимального управления в учебном пособии рассматривают аналитические задачи управления, зависящего от времени: терминальные двухточечные и локальные задачи, опирающиеся на принцип максимума, а также базирующиеся на стохастическом варианте динамического программирования. При этом управление осуществляется по неполной информации, когда измерения из-за наличия шумов и помех не позволяют точно определить состояние системы.

ВВЕДЕНИЕ

Окружающий мир представляет собой совокупность взаимосвязанных биологических, технических, социально-экономических и других систем, нормальное функционирование которых невозможно без определенной организации. Наиболее существенными функциями в организации этих систем является, во-первых, управление, обеспечивающее сохранение структуры, режим ее деятельности, реализацию целевых программ организации, а также адаптацию к изменяющимся условиям; во-вторых, связь (передача информации), так как управление системой сопряжено с привлечением и обработкой больших объемов сведений о состоянии и поведении указанных систем. «Связь и управление являются основными процессами, характеризующими любую организацию» [Винер]. Информация как ресурс наряду с энергетическими и материальными ресурсами занимает ключевое место в процессе существования экосистем. В сфере человеческой деятельности информация является стратегическим ресурсом, оказывающим огромное влияние на процессы в экономике, образовании и культуре. Эффективность управления любой динамической (не только технической или кредитно-финансовой, но и административно-

хозяйственной) системой во многом определяется тем, как организовано хранение, поиск, передача, обработка и пополнение информации. Используемые ЭВМ автоматизированные системы, осуществляющие эти процессы с целью предоставления пользователям информации в соответствии с их запросами, называются информационными системами. Объектами, образующими информационные системы, являются языки программирования, компьютеры, вычислительные сети и другие аппаратно-программные средства, предназначенные для выполнения заданных процедур передачи и обработки данных. При этом необходимо отметить, что обработка данных осуществляется без конкретного смысла, т.е. без обработки содержания, и трансформация данных в информацию предусматривает наличие некоторых механизмов интерпретации результатов. Оценить смысловое содержание полученных данных может находящийся за пределами системы обработки данных человек-оператор (пользователь) через систему знаний с помощью разработанных алгоритмов, логических выводов и эвристических моделей. Пользователь при рассматриваемом подходе выступает в роли субъекта системы, определяющего ценность исходной и обработанной информации, и в совокупности с системой обработки данных формирует и организует эффективное функционирование информационной системы. Существенным в информации при этом оказывается содержательный аспект данных (сведений, фактов), а задачей информационной системы становится преобразование исходной («сырой») в результатную (обработанную) информацию, необходимую для принятия решения. Информация как бы становится одновременно сырьем и продуктом, что кардинально ее отличает как ресурс от материальных, энергетических и других видов ресурсов в сфере производства, для которой главным является производство и потребление материальных благ. В сферу влияния информационной системы в дополнение к ранее перечисленным объектам оказываются вовлеченными также система восприятия (наблюдения, результаты измерений), модели и методы анализа процессов, алгоритмы обработки сигналов, в том числе использующие теоретико-вероятностные подходы при исследовании динамических систем, а также прикладные программы, реализующие эти алгоритмы на языке компьютеров. В задачах управления на основании анализа

результатной информации, оценки вариантов синтеза осуществляется выбор рационального управления системой. В других задачах информационно-справочного и расчетного характера указанный анализ заканчивается логическим выводом, экспертными оценками или переходом к следующему этапу научно-исследовательской работы.

Обобщенная схема преобразования информации, взаимосвязей объектов и субъектов информационной системы представлена на рисунке В.1.

В связи с появлением больших распределенных систем актуальной становится проблема управления этими системами. Огромный поток данных, необходимость их передачи и обработки в интересах многочисленных расположенных в разных местах пользователей для выработки соответствующих управленческих решений делает не эффективным и практически не реальным хранение и обработку этих данных в одном месте. Решение этой проблемы взаимодействия пользователей лежит на путях создания локальных и глобальных сетевых инфраструктур (распределенных вычислительных или компьютерных сетей). Указанные сети являются важной компонентной информационной системы и структурно отображены на обобщенной схеме.

Обслуживающая информационную систему и пользователей система знаний занимает на схеме промежуточное положение. Это объясняется тем, что эта система в виде методик, сведений из книг, описания моделей и алгоритмов может быть придана непосредственно пользователям или в том или ином объеме в качестве набора программных модулей хранится в памяти информационной системы и по команде пользователей или внутренним командам взаимодействует с объектами системы обработки данных (вычислительной сети).

Представленная схема при всей своей условности достаточно универсальна и охватывает широкий спектр задач, выполняемых информационными системами. При использовании этих систем в качестве управляющей подсистемы – обязательного элемента системы управления, выход блока принятия решения, в котором производится оценка вариантов и выбор наиболее рационального управления, замыкают на вход динамической системы – объекта управления (охватывает эту систему обратной связью).

Если информационная система выполняет самостоятельные функции научно-исследовательского или информационно-справочного характера, то из схемы исключается динамическая система и к пользователям с блока принятия решений поступают данные в виде экспертных оценок, логических выводов, проектных решений и т.д. В случае обмена информацией между пользователями приоритетным становится содержательный аспект этих данных – исходная и результатная информация. В задачах идентификации и имитационного моделирования осуществляется соответственно формирование информационной модели системы и проверка этой модели на влияние неблагоприятных факторов и чувствительность к вариациям параметров. В качестве еще одного примера можно указать на интегрирование информационных систем в современные системы связи, что расширило возможности информационных технологий.

Современные методы описания процессов в информационных системах разнообразны и требуют различного смыслового содержания и представления для применения в инженерной практике. Среди них все большее значение приобретают теоретико-вероятностные методы исследований, основанные на вероятностной трактовке протекающих в информационных системах процессов. Вероятностный (статистический) подход позволяет более полно учесть состояние динамической системы, характер управляющих и возмущающих воздействий, результирующее поведение информационных потоков в больших вычислительных сетях и во многих случаях более адекватен практическим задачам.

Круг вытекающих из указанного подхода проблем, охватываемых пособием, достаточно широк: описание математических моделей случайных процессов в информационных системах, формирование на их базе статистических методов проверки гипотез и обнаружения, оценивания и фильтрации, интерполяции (сглаживания) и экстраполяции (прогнозирования), а также разработка алгоритмов оптимального управления стохастическими системами.

Значительное место уделено практическим аспектам применения методов статистического анализа и синтеза, использующих математический аппарат теории марковских процессов. Для успешного овладения этими методами изложение

материала базируется на математическом описании системы (его фазовых координат) в пространстве состояний, а при решении задач фильтрации – на способах совместного обнаружения и оценивания сигналов. Для получения реализуемого в инженерной практике единообразного алгоритма при исследовании линейных и нелинейных систем в данном пособии широко применяется приближенный метод статистической линеаризации нелинейностей.

Учитывая, что развитие систем связи, управления и т.д. характеризуется в настоящее время широким использованием цифровой вычислительной техники, в учебном пособии, в основном, уделяется внимание дискретным алгоритмам обработки информации и управлению дискретными стохастическими системами.

1. МЕТОДЫ ОПИСАНИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

1.1. ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ

При решении задач анализа и синтеза необходимыми для расчета данными служат полученные из опыта некоторые экспериментальные данные. Опытom называются наблюдения какого-либо явления при выполнении некоторого комплекса условий и действий, который каждый раз при повторении указанного опыта строго выполняется. Количественная характеристика опыта состоит в определении получаемой из опыта некоторой величины. Из-за влияния различных трудно учитываемых факторов результаты экспериментов в серии опытов имеют случайный непредсказуемый характер, а сами величины оказываются случайными. Несмотря на это в длинной серии опытов можно установить общие статистические закономерности, присущие реальным явлениям.

Эти закономерности отражаются в вероятности значений, полученных из опыта случайных величин. Так, например, если разбить на действительной числовой оси интервал возможных значений физической величины на конечное число непересекающихся подынтервалов и подсчитать в серии опытов число события попадания случайной величины в каждый из

подынтервалов, то отношение числа событий в одном из подынтервалов к общему числу опытов называется частотой появления событий в этом подынтервале. При достаточно длинной серии опытов эта частота мало изменяется и может служить количественной мерой вероятности появления или не появления рассматриваемых событий.

Остановимся на одном из способов определения случайного процесса на основе введенного понятия случайной величины. Под случайной величиной $\xi(t_i)$ понимается одно из значений случайной функции $\xi(t)$ при фиксированном аргументе t_i из произвольного множества $T, t_i \in T$. В зависимости от возможных значений случайная величина подразделяется на дискретную (из конечного или счетного множества X) или непрерывную (принимает непрерывные значения на всей действительной оси или ее интервале). Если аргумент t интерпретируется как время, то совокупность случайных величин называется случайным процессом^{*)}. Время может принимать дискретные или непрерывные значения. В соответствии с этим случайные процессы подразделяются на процессы с дискретным и непрерывным временем. Конкретный вид случайного процесса в результате отдельных экспериментов называется реализацией (траекторией или выборочной функцией).

Вероятностные характеристики случайного процесса могут быть определены на основе понятия совокупности (последовательности) случайных величин $\{\xi(t_i), i = \overline{1, n}\}$. Наиболее полной такой характеристикой является n -мерная функция распределения вероятностей

$$F(x_n, t_n; \dots; x_1, t_1) = P\{\xi(t_n) \leq x_n; \dots; \xi(t_1) \leq x_1\} \quad (1.1)$$

- вероятность того, что случайный процесс в любые возможные моменты времени t_1, \dots, t_n примет соответственно значения не выше уровней x_1, \dots, x_n из множества X . Эта функция удовлетворяет условиям неотрицательности $F(\cdot) \geq 0$ и согласованности

^{*)} Наряду с временным аргументом в теории случайных процессов в число аргументов могут входить координаты пространства. В этих задачах случайные факторы приводят к случайным полям, обладающим пространственно-временными характеристиками функций распределений. В данном пособии рассматриваются случайные функции временного аргумента.

$$\lim_{x_n \rightarrow \infty} F(x_n, t_n; \dots; x_1, t_1) = F(x_{n-1}, t_{n-1}; \dots; x_1, t_1).$$

Она является неубывающей функцией своих аргументов

$$F(-\infty, \dots, -\infty; t_n, \dots, t_1) = 0, \quad F(\infty, \dots, \infty; t_n, \dots, t_1) = 1.$$

Если функция дифференцируема по x_1, \dots, x_n , то можно определить n -мерную плотность вероятностей

$$p(x_n, t_n; \dots; x_1, t_1) = \frac{\partial^n F(x_n, t_n; \dots; x_1, t_1)}{\partial x_1 \dots \partial x_n}. \quad (1.2)$$

Плотность вероятностей (1.2) является неотрицательной функцией $p(\cdot) \geq 0$ и удовлетворяет условию нормировки

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_n, t_n; \dots; x_1, t_1) dx_1 \dots dx_n = 1.$$

Важным классом случайных процессов являются стационарные процессы. Случайный процесс называется стационарным, если функции распределения вероятностей инвариантны относительно сдвига времени для любых n и τ .

$$F(x_n, t_n; \dots; x_1, t_1) = F(x_n, t_n + \tau; \dots; x_1, t_1 + \tau).$$

Для стационарных процессов выражения для функции распределения не зависят от положения начала отсчета времени. Аналогичные соотношения выполняются и для плотностей вероятностей

$$p(x_n, t_n; \dots; x_1, t_1) = p(x_n, t_n + \tau; \dots; x_1, t_1 + \tau). \quad (1.3)$$

Если вероятностные характеристики случайных процессов не инвариантны к произвольному смещению начала времени, то процесс является нестационарным. Для стационарных случайных процессов одномерная функция плотности не зависит от времени; двумерная плотность зависит лишь от разности $t_2 - t_1$:

$$p(x_2, t_2; x_1, t_1) = p(x_2, x_1; t_2 - t_1),$$

n -мерная плотность вероятностей будет функцией $n-1$ разностей $t_i - t_1$, $i = \overline{2, n}$.

Перейдем к рассмотрению условных функций распределений. Вероятность совместного выполнения неравенств

$$\xi(t_n) \leq x_n, \xi(t_{n-1}) \leq x_{n-1}, \dots, \xi(t_{n-m}) \leq x_{n-m}$$

при условии, что

$$\xi(t_{n-m-1}) = x_{n-m-1}, \xi(t_{n-m-2}) = x_{n-m-2}, \dots, \xi(t_1) = x_1$$

где $x_i \in X, i = \overline{1, n}$, описывается условной функцией распределения

$$F(x_n, t_n; \dots; x_{n-m}, t_{n-m} \mid x_{n-m-1}, t_{n-m-1}; \dots; x_1, t_1)$$

Определим условную плотность вероятностей как производную по x_n, \dots, x_{n-m} функции распределения. С учетом формулы полной вероятности

$$p(x_{n-m-1}, t_{n-m-1}; \dots; x_1, t_1) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_n, t_n; \dots; x_1, t_1) dx_1 \dots dx_{n-m}$$

имеем

$$\begin{aligned} p(x_n, t_n; \dots; x_{n-m}, t_{n-m} \mid x_{n-m-1}, t_{n-m-1}; \dots; x_1, t_1) = \\ = \frac{p(x_n, t_n; \dots; x_1, t_1)}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_n, t_n; \dots; x_1, t_1) dx_n \dots dx_{n-m}} \end{aligned} \quad (1.4)$$

Соотношение (1.4) называется формулой Байеса для условных вероятностей. Как и безусловные условные плотности вероятностей удовлетворяют условиям неотрицательности

$$p(x_n, t_n; \dots; x_{n-m}, t_{n-m} \mid x_{n-m-1}, t_{n-m-1}; \dots; x_1, t_1) \geq 0$$

и нормировки

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_n, t_n; \dots; x_{n-m}, t_{n-m} | x_{n-m-1}, t_{n-m-1}; \dots; x_1, t_1) dx_n \dots dx_{n-m} = 1$$

В простейшем варианте двумерной условной плотности ($n=2, m=0$) формула Байеса принимает вид

$$p(x_2, t_2 | x_1, t_1) = \frac{p(x_2, t_2; x_1, t_1)}{p(x_1, t_1)} = \frac{p(x_1, t_1 | x_2, t_2) p(x_2, t_2)}{\int_{-\infty}^{\infty} p(x_1, t_1 | x_2, t_2) p(x_2, t_2) dx_2}.$$

В задачах теории случайных процессов довольно часто необходимо найти по известной плотности вероятностей $p_{\xi}(x)$ плотность вероятностей функции случайной величины $\eta = \varphi(\xi)$, т.е. $p_{\eta}(y)$.

Предположим, что функция $\varphi(x)$ имеет первые кусочно-непрерывные производные по x и не постоянна ни на каком множестве значений аргумента x , имеющем отличную от нуля вероятность. Кроме того, будем полагать что случайные величины связаны однозначной детерминированной зависимостью. В силу последнего предположения из того факта, что величина ξ заключена в интервале $(x, x+dx)$ достоверно следует, что η находится в интервале $(y, y+dy)$. И вероятности этих событий должны быть одинаковы $p_{\xi}(x)dx = p_{\eta}(y)dy$ или

$$p_{\eta}(y) = p_{\xi}(x) \left| \frac{dx}{dy} \right| = p_{\xi}(\varphi^{-1}) \left| \frac{d\varphi^{-1}}{dy} \right|. \quad (1.5)$$

Поскольку плотность вероятностей не может быть отрицательной, то в формулу (1.5) необходимо подставить модуль производной.

В случае многомерной плотности вероятностей имеем

$$\begin{aligned} p_{\eta}(y_n, t_n; \dots; y_1, t_1) &= p_{\xi}(x_n, t_n; \dots; x_1, t_1) |J(y)| = \\ &= p_{\xi}(\varphi_n^{-1}, t_n; \dots; \varphi_1^{-1}, t_1) |J(y)|, \end{aligned}$$

где якобиан преобразования вектора $x = \varphi^{-1}(y)$ имеет вид определителя

$$J(y) = \begin{vmatrix} \frac{\partial \varphi_1^{-1}}{\partial y_1} & \dots & \frac{\partial \varphi_1^{-1}}{\partial y_n} \\ \dots & \dots & \dots \\ \frac{\partial \varphi_n^{-1}}{\partial y_1} & \dots & \frac{\partial \varphi_n^{-1}}{\partial y_n} \end{vmatrix} = \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)}.$$

Если функция $\eta = \varphi(\xi)$ такова, что обратная ей функция неоднозначна, то одному значению y соответствует несколько ветвей $x = \varphi^{-1}(y)$. Обозначим через $x_{ik}(y_n, \dots, y_1)$ k -ую ветвь обратного преобразования, $i = \overline{1, n}$. В этом случае многомерная плотность вероятностей равна

$$p_\eta(y_n, t_n; \dots; y_1, t_1) = \sum_k p_\xi(x_{nk}, t_n; \dots; x_{1k}, t_1) \frac{\partial(x_{1k}, \dots, x_{nk})}{\partial(y_1, \dots, y_n)}.$$

1.2. МОМЕНТНЫЕ И КОРРЕЛЯЦИОННЫЕ ФУНКЦИИ

Полное вероятностное описание случайного процесса с помощью конечномерных распределений возможно тогда, когда процесс представляет собой конечную последовательность случайных величин $\{\xi(t_i), i = \overline{1, n}\}$, т.е. для процессов с дискретным временем. Если случайный процесс непрерывен во времени для его полного описания необходимо перейти к континуальным (непрерывным) характеристикам. Прежде чем перейти к вероятностному описанию непрерывного случайного процесса остановимся на имеющем большое значение для технических приложениях функциональном преобразовании – характеристической функции (хф) [1]. Хф первого порядка $\Theta_1(v)$ называется среднее значение случайной функции $\exp\{jv\xi(t)\}$

$$\Theta_1(v) = M[\exp\{jv\xi(t)\}] = \int_{-\infty}^{\infty} \exp\{jvx\} p(x, t) dx,$$

где v - произвольный действительный параметр. Из вида хф следует, что она является Фурье-преобразованием плотности вероятностей $p(x, t)$. Более полную информацию о случайном процессе можно получить из многомерной хф

$$\begin{aligned}\Theta_n(v_1, \dots, v_n) &= M \left[\exp \left\{ j \sum_{i=1}^n v_i \xi(t_i) \right\} \right] = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ j \sum_{i=1}^n v_i x_i \right\} p(x_n, t_n; \dots x_1, t_1) dx_1 \dots dx_n.\end{aligned}$$

По характеристической функции, применяя обратное преобразование Фурье, можно определить плотность вероятностей

$$p(x_n, t_n; \dots x_1, t_1) = (2\pi)^{-n} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left\{ -j \sum_{i=1}^n v_i x_i \right\} \Theta_n(v_1, \dots, v_n) dv_1 \dots dv_n.$$

Для исчерпывающего описания случайного процесса необходимо устремить $n \rightarrow \infty$, что приводит к так называемому континуальному характеристическому функционалу (ХФ)

$$\Theta[v, T] = M \left[\exp \left\{ j \int_0^T v(t) \xi(t) dt \right\} \right], \quad (1.6)$$

где $v(t)$ - вспомогательная действительная функция.

Таким образом, при описании случайных процессов можно с одинаковым правом использовать как плотности вероятностей, так и характеристическую функцию или ХФ. На практике для описания случайных процессов чаще используются система функции, получаемые из хф и ХФ. Причем эти системы функции обладают тем свойством, что функции более низкого порядка не несут информации о последующих функциях. К числу таких функций относятся неслучайные статистические характеристики: моментные и корреляционные функции. Причем наибольший интерес представляют корреляционные функции первых двух порядков, так как значимость многократных корреляций с увеличением порядка уменьшается. В дальнейшем при описании конкретных случайных процессов корреляционную функцию первого порядка будем называть также математическим ожиданием, а для корреляционной функции второго порядка название порядка опускать. По моментным и корреляционным функциям можно определить хф и, следовательно, плотность вероятностей. И наоборот, по хф - рассчитать указанные системы функции. Выражение ХФ в виде многомерного разложения

в ряд относительно моментных $m_n(\cdot)$ и корреляционных $k_n(\cdot)$ функций имеют следующий вид [2]

$$\begin{aligned}\Theta_n(v_1, \dots, v_n) &= 1 + \sum_{r=1}^{\infty} \frac{j^r}{r!} \sum_{\alpha, \dots, \omega=1}^n m_r(t_\alpha, \dots, t_\omega) v_\alpha \dots v_\omega, \\ \Theta_n(v_1, \dots, v_n) &= \exp \left\{ \sum_{r=1}^{\infty} \frac{j^r}{r!} \sum_{\alpha, \dots, \omega=1}^n k_r(t_\alpha, \dots, t_\omega) v_\alpha \dots v_\omega \right\}.\end{aligned}\quad (1.7)$$

Для ХФ разложение в ряд записывается в форме

$$\Theta[v, T] = 1 + \sum_{n=1}^{\infty} \frac{j^n}{n!} \int_0^T \dots \int_0^T m_n(t_1, \dots, t_n) \prod_{r=1}^n v(t_r) dt_1 \dots dt_n, \quad (1.8)$$

$$\Theta[v, T] = \exp \left\{ \sum_{n=1}^{\infty} \frac{j^n}{n!} \int_0^T \dots \int_0^T k_n(t_1, \dots, t_n) \prod_{r=1}^n v(t_r) dt_1 \dots dt_n \right\}. \quad (1.9)$$

Моментные и корреляционные функции определяются через ХФ путем n -кратного функционального дифференцирования

$$m_n(t_1, \dots, t_n) = M \{ \xi(t_1), \dots, \xi(t_n) \} = \frac{1}{j^n} \frac{\delta^n \Theta[v, T]}{\delta v(t_1) \dots \delta v(t_i) \dots \delta v(t_n)} \Big|_{v(t_i) = 0},$$

$$k_n(t_1, \dots, t_n) = \frac{1}{j^n} \frac{\delta^n \ln \Theta[v, T]}{\delta v(t_1) \dots \delta v(t_i) \dots \delta v(t_n)} \Big|_{v(t_i) = 0}.$$

Сравнивая выражения (1.8) и (1.9) можно получить следующие соотношения, связывающие моментные и корреляционные функции:

$$\begin{aligned}m_1(t) &= k_1(t); \\ m_2(t_1, t_2) &= k_2(t_1, t_2) + k_1(t_1)k_1(t_2); \\ m_3(t_1, t_2, t_3) &= k_3(t_1, t_2, t_3) + k_1(t_1)k_2(t_2, t_3) + \\ &\quad + k_1(t_2)k_2(t_1, t_3) + k_1(t_3)k_2(t_1, t_2) + \\ &\quad + k_1(t_1)k_1(t_2)k_1(t_3); \\ &\dots\end{aligned}\quad (1.10)$$

- корреляционная функция произведения $\eta(t)=\varphi(t)\xi(t)$, где $\varphi(t)$ - неслучайная, а $\xi(t)$ - случайная функция, определяется соотношением

$$k_{\eta}(t_1, t_2) = \varphi(t_1)\varphi(t_2)k_{\xi}(t_1, t_2);$$

- если $\eta(t) = \sum_{i=1}^k \varphi_i(t)\xi_i(t)$, где $\varphi_i(t)$ - неслучайная, а $\xi_i(t)$ - случайная функция, то

$$k_{\eta}(t_1, t_2) = \sum_{i=1}^k \sum_{j=1}^k \varphi_i(t_1)\varphi_j(t_2)k_{\xi_{ij}}(t_1, t_2).$$

Взаимная корреляционная функция $k_{\xi\eta}(t_1, t_2)$ случайных функций $\xi(t)$ и $\eta(t)$ имеет свойства в принципе аналогичные свойствам обычной корреляционной функции. Однако, некоторые отличия в свойствах этих функций имеются. Взаимная корреляционная функция не является симметричной функцией относительно аргументов. Но можно установить соотношение $k_{\xi\eta}(t_1, t_2) = k_{\eta\xi}(t_2, t_1)$, означающее, что при одновременной перестановке аргументов и порядка следования случайных функций значение взаимной корреляционной функции не изменяется. Если процессы $\xi(t)$ и $\eta(t)$ статистически независимы, то взаимная корреляционная функция равна нулю. В отличие от условия стационарности (1.3), характеризующего случайные процессы стационарные в строгом смысле, случайные процессы в рамках корреляционной теории оказываются стационарными в широком смысле. Для этих процессов математическое ожидание и дисперсия не зависят от времени, а корреляционная функция зависит от разности аргументов $t_2 - t_1 = \tau$.

$$k(t_1, t_2) = k(t_1, t_1 + \tau) = k(\tau), \quad D = k(0).$$

Случайные процессы стационарные строго (в узком смысле) всегда стационарны в широком смысле, но не наоборот. Очевидно, что в рамках корреляционной теории достаточно знать плотности или функции распределения вероятностей не выше второго порядка. Отметим, что корреляционная теория дает полное описание важного класса случайных процессов – гауссовских процессов.

В технических приложениях корреляционной теории для стационарных процессов нередко вместо корреляционной функции ограничиваются указанием интервала корреляции τ_k . При $\tau > \tau_k$

значения $\xi(t)$ и $\xi(t+\tau)$ для любого момента времени t можно считать практически некоррелированными. Обычно интервал корреляции определяют как половину ширины основания прямоугольника с высотой равной дисперсии $k(0)=D$, площадь которого равняется площади под кривой модуля корреляционной функции

$$\tau_k = \frac{1}{D} \int_0^{\infty} |k(\tau)| d\tau. \quad (1.12)$$

На практике обычно не представляется возможным использовать формулы (1.9), (1.10) и (1.11) для точного определения статистических характеристик, так как функции распределения или плотностей вероятностей часто оказываются неизвестными. Поэтому прибегают к определению оценок (статистик) этих характеристик путем обработки экспериментально полученных n реализаций случайного процесса $\xi(t)$ на основании следующих формул

$$m^*(t) = \frac{1}{n} \sum_{i=1}^n \xi_i(t) \quad (1.13)$$

- оценка математического ожидания случайного процесса;

$$D^*(t) = \frac{1}{n} \sum_{i=1}^n [\xi_i(t) - m^*(t)]^2 \quad (1.14)$$

- оценка дисперсии случайного процесса;

$$k^*(t_1, t_2) = \frac{1}{n} \sum_{i=1}^n [\xi_i(t_1) - m^*(t_1)][\xi_i(t_2) - m^*(t_2)] \quad (1.15)$$

- оценка корреляционной функции случайного процесса;

$$k_{\xi\eta}^*(t_1, t_2) = \frac{1}{n} \sum_{i=1}^n [\xi_i(t_1) - m_{\xi}^*(t_1)][\eta_i(t_2) - m_{\eta}^*(t_2)] \quad (1.16)$$

- оценка взаимной корреляционной функции случайного процесса.

Согласно этим формулам оценки представляют собой приближенные значения статистических характеристик. При $n \rightarrow \infty$

указанные оценки с вероятностью единицы сходятся к соответствующим статистическим характеристикам.

1.3. ЭРГОДИЧЕСКИЕ ПРОЦЕССЫ

Определение статистических характеристик на основании формул (1.13), (1.14), (1.15) и (1.16) предполагает достаточно большое число реализации случайного процесса. Если случайный процесс стационарен в строгом смысле, то представляется допустимым получения этих реализаций путем разделения результата одного опыта на необходимое число частей. Однако эта возможность в действительности имеет место не во всех случаях существования стационарного процесса, а лишь тогда, когда этот процесс эргодичен. Случайный стационарный процесс считается эргодическим, если любая статистическая характеристика, полученная усреднением по множеству возможных реализаций равна временному среднему, полученному усреднением за достаточно большой промежуток времени T по одной единственной реализации случайного процесса. В работе [3] показано, что необходимым и достаточным условием эргодичности стационарного случайного процесса является

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(1 - \frac{\tau}{T}\right) k_{\xi}(\tau) d\tau = 0,$$

где $k_{\xi}(\tau)$ - корреляционная функция процесса.

В результате в случае эргодичности стационарного процесса при вычислении оценок статистических характеристик вместо формул (1.13), (1.14), (1.15) и (1.16) могут использоваться более простые формулы

$$m^* = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi_1(t) dt \approx \frac{1}{T} \int_0^T \xi_1(t) dt,$$

$$D^* = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\xi_1(t) - m^*]^2 dt \approx \frac{1}{T} \int_0^T [\xi_1(t) - m^*]^2 dt$$

$$k^*(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T - \tau} \int_0^{T-\tau} [\xi_1(t) - m^*][\xi_1(t + \tau) - m^*] dt \approx$$

$$\approx \frac{1}{T} \int_0^T [\xi_1(t) - m^*][\xi_1(t + \tau) - m^*] dt,$$

$$k_{\xi\eta}^*(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T - \tau} \int_0^{T-\tau} [\xi_1(t) - m_{\xi}^*][\eta_1(t) - m_{\eta}^*] dt \approx$$

$$\approx \frac{1}{T} \int_0^T [\xi_1(t) - m_{\xi}^*][\eta_1(t) - m_{\eta}^*] dt,$$

при $T \gg \tau$,

где $\xi_1(t)$ и $\eta_1(t)$ - одна реализация соответственно случайных процессов $\xi(t)$ и $\eta(t)$.

1.4. СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ

В теории случайных процессов наряду с описанием сигналов во временной области используют частотные представления сигналов. Пусть для любой реализации случайного процесса на интервале $(0, T)$ выполняется условие

$$\frac{1}{T} \int_0^T |\xi(t)|^2 dt < \infty.$$

Это условие означает ограничение мощности сигнала. В этом случае спектральная функция конкретной реализации $\xi(t)$ может быть определена с помощью прямого преобразования Фурье

$$\Phi(j\omega) = \int_{-\infty}^{\infty} \xi(t) \exp\{-j\omega t\} dt.$$

В спектральной функции содержится вся информация о реализации $\xi(t)$. Последняя может быть восстановлена по спектральной функции путем применения обратного преобразования Фурье

$$\xi(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(j\omega) \exp\{j\omega t\} d\omega.$$

Из-за случайного характера непосредственное использование спектральной функции вызывает неудобства. В связи с этим целесообразно перейти к другому виду спектрального представления - спектральной плотности, полученной в результате усреднения квадрата модуля спектральной функции по всем реализациям. Для стационарного в широком смысле с нулевым математическим ожиданием процесса эта функция равняется

$$F(\omega) = \lim_{T \rightarrow \infty} M \left\{ \frac{1}{T} |\Phi_T(j\omega)|^2 \right\},$$

где $\Phi_T(j\omega) = \int_0^T \xi_T(t) \exp\{-j\omega t\} dt$, $\xi_T(t)$ - усеченная реализация,

равная нулю вне интервала $(0, T)$.

Спектральная плотность или энергетический спектр дает только усредненную картину распределения энергии по частотам гармонических составляющих, не учитывает их фазовую структуру. Поэтому по ней нельзя восстановить реализации случайного процесса. Спектральная плотность и корреляционная функция для стационарного процесса связаны между собой парой взаимного преобразования Фурье, называемых формулами Хинчина-Винера.

$$F(\omega) = \int_{-\infty}^{\infty} k(\tau) \exp\{-j\omega\tau\} d\tau, \quad (1.17)$$

$$k(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp\{j\omega\tau\} d\omega. \quad (1.18)$$

Учитывая четность спектральной плотности, формулы Винера-Хинчина можно также записать в виде

$$F(\omega) = 2 \int_0^{\infty} k(\tau) \cos \omega \tau d\tau,$$

$$k(\tau) = \frac{1}{\pi} \int_0^{\infty} F(\omega) \cos \omega \tau d\omega.$$

Если случайный процесс имеет отличное от нуля математическое ожидание $M\{\xi(t)\}=m_\xi$, то в выражении спектральной плотности появляется дополнительное дельтаобразное слагаемое $2\pi m_\xi^2 \delta(\omega)$, что соответствует появлению в спектре дискретной линии на нулевой частоте. Спектральная плотность является неотрицательной $F(\omega) \geq 0$ и четной $F(\omega) = F(-\omega)$ функцией своего аргумента.

На практике протяженность спектральной плотности стационарного случайного процесса характеризуется шириной полосы спектра – Δ_Π . Ее определяют как величину площади под кривой спектральной плотности, отнесенной к спектральной плотности на характерной частоте ω_0 :

$$\Delta_\Pi = \frac{1}{2\pi F(\omega_0)} \int_0^{\infty} F(\omega) d\omega. \quad (1.19)$$

В качестве $F(\omega_0)$ принимают максимум спектральной плотности или координату, соответствующую точке симметрии. Обычно при широкополосном спектре ею является спектральная плотность на нулевой частоте $F(0)$. Отметим также, что произведение интервала корреляции на ширину полосы $\tau_k \Delta_\Pi$ – величина постоянная для семейства спектральных плотностей заданной формы. Таким образом, можно сформулировать следующее утверждение: чем уже корреляционная функция, тем протяженнее спектр, и наоборот, чем протяженнее корреляционная функция, тем уже спектр.

2. МОДЕЛИ ПРОЦЕССОВ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

При аналитическом исследовании поведения реальных динамических систем важным является обоснование математических моделей процессов, протекающих в этих системах. В ходе формирования методов анализа и синтеза эти модели, с одной стороны, должны наиболее адекватно отражать свойства реальных сигналов, с другой – допускать исследования этих процессов известными теоретическими методами. Выход, видимо, следует искать на основе компромиссных решений в части выбора наиболее простых и в то же время содержательных (продуктивных) моделей при решении интересующих задач. Рассматриваемые в данном разделе модели случайных процессов являются наиболее типичными в практике исследований динамических систем с учетом реальных режимов, возмущений и помех и во многих случаях могут служить базовыми при формировании более сложных процессов.

2.1. ГАУССОВСКИЙ ПРОЦЕСС

В инженерной практике особую роль играют гауссовские случайные процессы благодаря возможности аппроксимации ими многих реальных процессов. К числу таких процессов, например, относятся собственные шумы электронных устройств, тепловые флуктуации, космические излучения и т.д. Они представляют собой суммарный эффект большого числа сравнительно малых элементарных импульсов, возникающих в случайные моменты времени. На основании центральной предельной теоремы теории вероятностей [4] плотность вероятности этой суммы неограниченно приближается к гауссовской с увеличением числа слагаемых, независимо от того, какие плотности вероятности имеют отдельные слагаемые.

Случайный процесс, у которой корреляционные функции третьего и более высоких порядков для любых t равняются нулю, называются гауссовским $k_n(t_1, \dots, t_n) = 0, n \geq 3$:

Характеристическая функция на основании (1.7)

$$\Theta_n(v_1, \dots, v_n) = \exp \left\{ j \sum_{i=1}^n k_1(t_i) v_i - \frac{1}{2} \sum_{i,k=1}^n k_2(t_i, t_k) v_i v_k \right\}.$$

В результате обратного Фурье-преобразования найдем n -мерную плотность вероятностей гауссовского процесса

$$\begin{aligned} p(x_n, t_n; \dots; x_1, t_1) &= \\ &= (2\pi)^{-n/2} |k_2|^{-1/2} \exp \left\{ -\frac{1}{2} (x - m_\xi)^T k_2^{-1} (x - m_\xi) \right\} = \quad (2.1) \\ &= (2\pi)^{-n/2} |k_2|^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{i,k=1}^n h_{ik} [x_i - m_\xi(t_i)] [x_k - m_\xi(t_k)] \right\} \end{aligned}$$

где $|k_2|$ - определитель корреляционной матрицы $k_2 = \|k_2(t_i, t_k)\|$ размера $(n \times n)$; h_{ik} - компоненты матрицы $\|h(t_i, t_k)\|$, обратной корреляционной матрице $\|k_2(t_i, t_k)\|$, x и $k_1 = m_\xi$ - векторы-столбцы размера $(n \times 1)$.

Для стационарного гауссовского процесса математическое ожидание $m_\xi(t_i) = m_\xi, i = \overline{1, n}$ постоянно, а корреляционные функции зависят от разности моментов времени $t_i - t_k$. Для гауссовских процессов стационарность в строгом и широком смысле совпадают.

В случае одномерного гауссовского процесса плотность вероятности имеет вид

$$p(x, t) = (2\pi)^{-1} D_t^{-1/2} \exp \left\{ -\frac{1}{2} D_t^{-1} [x - m_\xi(t)]^2 \right\}, \quad (2.2)$$

где $k_2(t, t) = D_t$.

Гауссовский процесс, у которого взаимные корреляционные функции равны нулю $k_2(t_i, t_k) = 0, i \neq k$, является процессом с независимыми значениями с плотностью вероятностей равной произведению n одномерных гауссовских плотностей

$$p(x_n, t_n; \dots; x_1, t_1) = (2\pi)^{-n/2} (D_{t_1} \dots D_{t_n})^{1/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n D_{t_i} [x_i - m_\xi(t_i)]^2 \right\},$$

где $k_2(t_i, t_i) = D_{t_i}$.

Условные плотности вероятностей совместно гауссовских процессов являются гауссовскими. Этот результат следует из формулы условной вероятности (1.4)

$$p(x_2, t_2 | x_1, t_1) = \frac{p(x_2, t_2; x_1, t_1)}{p(x_1, t_1)}. \quad (2.3)$$

Действительно, на основании (2.1) двумерная плотность вероятностей нестационарного гауссовского процесса

$$\begin{aligned} p(x_2, t_2; x_1, t_1) &= \\ &= (2\pi)^{-1} \left\{ D_{t_1} D_{t_2} [1 - r^2(t_1, t_2)] \right\}^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{i,k=1}^2 h(t_i, t_k) [x_i - m_\xi(t_i)] [x_k - m_\xi(t_k)] \right\} = \\ &= (2\pi)^{-1} \left\{ D_{t_1} D_{t_2} [1 - r^2(t_1, t_2)] \right\}^{-1/2} \exp \left\{ -\frac{1}{2[1 - r^2(t_1, t_2)]} \left(\frac{[x_1 - m_\xi(t_1)]^2}{D_{t_1}} - \right. \right. \\ &\quad \left. \left. - 2r(t_1, t_2) \frac{[x_1 - m_\xi(t_1)] [x_2 - m_\xi(t_2)]}{\sqrt{D_{t_1}} \sqrt{D_{t_2}}} + \frac{[x_2 - m_\xi(t_2)]^2}{D_{t_2}} \right) \right\}, \end{aligned}$$

где коэффициент корреляции $r(t_1, t_2) = \frac{k_2(t_1, t_2)}{\sqrt{D_{t_1} D_{t_2}}}$, $D_{t_1} = k_2(t_1, t_1)$,

$$D_{t_2} = k_2(t_2, t_2).$$

Используя формулу (2.3), с учетом (2.2) получим условную плотность вероятностей

$$\begin{aligned} p(x_2, t_2 | x_1, t_1) &= \left\{ 2\pi D_{t_2} [1 - r^2(t_1, t_2)] \right\}^{-1/2} \times \\ &\times \exp \left\{ -\frac{[x_2 \sqrt{D_{t_1}} - x_1 \sqrt{D_{t_2}} r(t_1, t_2)]^2}{2D_{t_1} D_{t_2} [1 - r(t_1, t_2)]} \right\}. \end{aligned} \quad (2.4)$$

При выводе формулы (2.4) математические ожидания приравнялись нулю $m_\xi(t_1) = m_\xi(t_2) = 0$.

Сумма гауссовских случайных процессов также имеет гауссовское распределение. Отсюда следует, что линейная

комбинация гауссовских случайных процессов

$\eta(t) = \sum_{i=1}^n C_i(t) \xi_i(t) + S_i(t)$, где $C_i(t)$ и $S_i(t)$ - заданные функции времени,

также представляет гауссовский случайный процесс. Например, сумма двух гауссовских процессов, имеющих математическое ожидание и корреляционные функции равные соответственно m_{ξ_1} , m_{ξ_2} и k_{ξ_1} , k_{ξ_2} представляет собой гауссовский процесс с математическим ожиданием $m_\eta = m_{\xi_1} + m_{\xi_2}$ и корреляционной функцией $k_\eta = k_{\xi_1} + k_{\xi_1 \xi_2} + k_{\xi_2 \xi_1} + k_{\xi_2}$, где $k_{\xi_1 \xi_2}$ и $k_{\xi_2 \xi_1}$ - взаимные корреляционные функции.

2.2. ВИНЕРОВСКИЙ ПРОЦЕСС. БЕЛЫЙ ШУМ

2.2.1. ОБЩИЕ ОПРЕДЕЛЕНИЯ

Винеровский процесс играет важную роль при формировании более сложных сигналов, в том числе марковских и фрактальных процессов. К понятию винеровского процесса можно прийти при рассмотрении следующей физической задачи. Частицы жидкости или газа в отсутствии внешних воздействий из-за столкновений с молекулами находятся в постоянном хаотическом движении (броуновском движении), интенсивность которого зависит от температуры и плотности среды. При этих столкновениях частицы изменяют свою скорость и направление движения. Если масса частицы равна m , то, пренебрегая силами трения, скорость движения частицы $B(t)$ по какой-либо координате на основании закона Ньютона определяется из соотношения

$$m \frac{dB(t)}{dt} = \eta(t),$$

где функция $\eta(t)$ является составляющей по этой координате случайной последовательности силовых толчков. Из условия симметрии направления этих толчков равновероятны и поэтому математическое ожидание этой функции равно нулю: $M\{\eta(t)\}=0$.

При определении статистических характеристик броуновского движения необходимо исходить из того, что в реальном физическом

эксперименте время корреляции процесса $\eta(t)$ конечно и, грубо говоря, не превосходит среднего времени между столкновениями τ_0 . Далее необходимо иметь ввиду, что реальные физические приборы, осуществляющие наблюдения и измерения, имеют конечное время разрешения Δt . В течение этого времени при большой концентрации молекул частица испытывает большое число столкновений, вследствие чего интервал измерения оказывается много больше интервала корреляции: $\Delta t \gg \tau_0$. В связи с изложенным на основании центральной предельной теоремы процесс $\eta(t)$ приближенно можно представить гауссовским процессом с математическим ожиданием равным нулю

$$m_\eta = M\{\eta(t)\} = 0$$

и дельтаобразной корреляционной функцией

$$k_{2\eta}(t_1, t_2) = M\{\eta(t_1)\eta(t_2)\} = \mathcal{N}_\eta \delta(t_2 - t_1),$$

где $F_\eta(\omega) = \mathcal{N}_\eta$ - спектральная плотность этого процесса, определяется из формулы (1.17), $\delta(t_2 - t_1)$ - дельта-функция (Приложение 1).

Спектральная плотность его постоянна во всем диапазоне изменения частот. Такой процесс получил название гауссовский белый шум. Он имеет бесконечную дисперсию (мощность) и служит математической моделью реальных широкополосных воздействий. В случае стационарного гауссовского белого шума его статистические характеристики имеют вид

$$m_\eta = 0, \\ k_{2\eta}(\tau) = M\{\eta(t + \tau)\eta(t)\} = \mathcal{N}_\eta \delta(\tau).$$

Винеровский процесс $B(t)$ по определению находится через белый шум $\eta(t)$ из стохастического дифференциального уравнения

$$\frac{dB(t)}{dt} = \eta(t), \quad B(t_0) = B_0. \quad (2.5)$$

Винеровский процесс после линейного интегрального преобразования (2.5) остается гауссовским процессом и с учетом

фильтрующих свойств дельта-функции и $B(t_0)=0$ имеет соответственно математическое ожидание, дисперсию и корреляционную функцию

$$m_B = M\{B(t)\} = \int_0^t M\{\eta(\tau)\}d\tau = 0, \quad (2.6)$$

$$D_B = M\{B^2(t)\} = \int_0^t \int_0^t M\{\eta(\tau_1)\eta(\tau_2)\}d\tau_1d\tau_2 = \mathcal{N}_\eta t, \quad (2.7)$$

$$\begin{aligned} k_{2B}(t_1, t_2) &= \int_0^{t_1} \int_0^{t_2} M\{\eta(\tau_1)\eta(\tau_2)\}d\tau_1d\tau_2 = \begin{cases} \mathcal{N}_\eta t_1, & t_1 \leq t_2 \\ \mathcal{N}_\eta t_2, & t_2 < t_1 \end{cases} = \\ &= \mathcal{N}_\eta \min(t_1, t_2), \quad t_1 > 0, t_2 > 0. \end{aligned} \quad (2.8)$$

Очевидно, корреляционную функцию винеровского процесса можно также представить для положительных t_1 и t_2 в виде

$$k_{2B}(t_1, t_2) = \frac{\mathcal{N}_\eta}{2} (t_1 + t_2 - |t_2 - t_1|). \quad (2.9)$$

Плотность распределения винеровского процесса имеет вид

$$p(B(t)) = \frac{1}{\sqrt{2\pi\mathcal{N}_\eta t}} \exp\left\{-\frac{B^2(t)}{2\mathcal{N}_\eta t}\right\}, \quad t > 0. \quad (2.10)$$

Рассмотрим некоторые свойства приращений винеровского процесса – их некоррелированность на неперекрывающихся интервалах времени и масштабируемость (самоподобие). Для моментов времени $t_2 > t_1 > t_0 > 0$ имеем

$$B(t_1) - B(t_0) = \int_{t_0}^{t_1} \eta(\tau)d\tau. \quad (2.11)$$

Отсюда на основании (2.6) и (2.7) математическое ожидание и дисперсия приращения винеровского процесса соответственно равны

$$M\{B(t_1) - B(t_0)\} = 0$$

$$M\{(B(t_1)-B(t_0))^2\}=\mathcal{N}_\eta(t_1-t_0)\sim(t_1-t_0). \quad (2.12)$$

Взаимная корреляционная функция приращений при выполнении условия (2.11) на основании (2.8) равна

$$\begin{aligned} M\{[B(t_2)-B(t_1)][B(t_1)-B(t_0)]\} &= \\ &= k_2(t_1, t_2) - k_2(t_1, t_1) - k_2(t_2, t_0) + k_2(t_1, t_0) = \\ &= \mathcal{N}_\eta t_1 - \mathcal{N}_\eta t_1 - \mathcal{N}_\eta t_0 + \mathcal{N}_\eta t_0 = 0. \end{aligned}$$

Таким образом, приращения процесса $B(t)$ некоррелированы и, поскольку имеют гауссовское распределение, они также независимы.

Кроме того, эти приращения стационарны, так как их математические ожидания равны постоянной величине (нулю), а дисперсии пропорциональны постоянным временным интервалам: $t_1-t_0=\dots=t_{k+1}-t_k$, $k=0,1,\dots$. Реализации винеровского процесса непрерывны ввиду того, что модуль его приращения $|\Delta B|=|B(t+\Delta t)-B(t)|$ характеризуется среднеквадратическим отклонением, порядок которого на основании (2.7) $\sim \Delta t^{1/2}$, и при $\Delta t \rightarrow 0$ стремится к нулю. Реализации винеровского процесса не дифференцируемы, так как

$$\lim_{\Delta t \rightarrow 0} |\Delta B| / \Delta t \sim \lim_{\Delta t \rightarrow 0} \Delta t^{-1/2} \rightarrow \infty.$$

Согласно свойству масштабируемости, вытекающему из соотношения (2.12), изменение масштаба времени в b раз сопровождается приращением координаты винеровского процесса в $b^{1/2}$ раз.

Это свойство можно представить в виде

$$B(bt)-B(bt_0)=b^{1/2}[B(t)-B(t_0)].$$

Особый характер поведения винеровского процесса является следствием его свойства, которое символически можно записать в виде $(\Delta B)^2 \sim \Delta t$. В то же время для дифференцируемой функции квадрат ее приращений имеет порядок малости квадрата приращения аргумента: $(\Delta f)^2 \sim \Delta t^2$.

Важным результатом этих свойств является соотношение для винеровского процесса

$$\begin{aligned}
J_m &= \lim_{\Delta t \rightarrow 0} \sum_{k=0}^{m-1} [B(t_{k+1}) - B(t_k)]^2 \\
&= \lim_{\Delta t \rightarrow 0} \mathcal{N}_\eta \sum_{k=0}^{m-1} (t_{k+1} - t_k) = \mathcal{N}_\eta (t_m - t_0),
\end{aligned} \tag{2.13}$$

где произвольное разбиение отрезка (t_0, t_m) :
 $t_0 < t_1 < \dots < t_k < \dots < t_m$, $\Delta t = \max_k (t_{k+1} - t_k)$.

Это соотношение означает, что последовательность случайных сумм слева при $\Delta t \rightarrow 0$ сходится по вероятности к неслучайной величине справа. Полученный результат является следствием того, что дисперсия суммы (2.13) при $\Delta t \rightarrow 0$ стремится к нулю: $\lim_{\Delta t \rightarrow 0} D(J_m) = \lim_{\Delta t \rightarrow 0} M\{J_m^2\} = 0$. При доказательстве соотношения (2.13) сначала рассмотрим дисперсию квадрата приращения отдельного слагаемого $\Delta B^2(t_k) = [B(t_{k+1}) - B(t_k)]^2$.

Согласно данным работы [5] эта дисперсия для гауссовского процесса при математическом ожидании $m_B = 0$ имеет вид

$$M\{\Delta B^4(t_k)\} - [M\{\Delta B^2(t_k)\}]^2 = 3D^2(t_k) - D^2(t_k) = 2\mathcal{N}_\eta^2 (t_{k+1} - t_k)^2, \tag{2.14}$$

где D_{t_k} - дисперсия приращения $\Delta B(t_k)$.

На основании (2.14) получаем

$$D(J_m) = M\{J_m^2\} = 2\mathcal{N}_\eta^2 \sum_{k=1}^{m-1} (t_{k+1} - t_k)^2 \leq 2\mathcal{N}_\eta^2 (t_m - t_0) \max(t_{k+1} - t_k).$$

Следовательно, при $\Delta t = \max(t_{k+1} - t_k) \rightarrow 0$, $D(J_m) \rightarrow 0$ и согласно неравенству Чебышева [4] для вероятностей $P\{|J_m - M\{J_m\}| \geq \varepsilon\} \leq D(J_m) / \varepsilon^2$ приходим к соотношению (2.13).

2.2.2. СТОХАСТИЧЕСКИЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ И ИНТЕГРАЛЫ

Следствием особых свойств винеровского процесса являются разные формы определения стохастических интегралов и соответственно решений описывающих поведение динамических

систем стохастических дифференциальных уравнений, в правой части которых присутствует винеровский процесс

$$dx(t)=f_t(x)dt+\gamma_t(x)dB(t), \quad x(t_0)=x(0) \quad (2.15)$$

где $f_t(\cdot)$ и $\gamma_t(\cdot)$ - детерминированные дифференцируемые функции, удовлетворяющих условию регулярности $|f_t(x)|^2+|\gamma_t(x)|^2 \leq k^2(1+|x|^2)$, k - постоянная величина.

Эти ограничения достаточны для существования и единственности решения уравнения (2.15). Используя дифференциальное соотношение винеровского процесса (2.5), можно уравнение (2.15) записать в эквивалентной форме, зависящей от белого шума

$$\frac{dx(t)}{dt} = f_t(x) + \gamma_t(x)\eta(t), \quad x(t_0) = x(0). \quad (2.16)$$

Случайный процесс $x(t)$, $t \geq t_0$ является решением стохастических дифференциальных уравнений (2.15) или (2.16).

Рассмотрим решение уравнения (2.15) в интегральной форме

$$x(t) = x(t_0) + \int_{t_0}^t f_\tau(x) d\tau + \int_{t_0}^t \gamma_\tau(x) dB(\tau). \quad (2.17)$$

Если вычисление первого интеграла в выражении (2.17) не имеет особенностей и проводится по обычным правилам математического анализа, то стохастические интегралы вида $\int \gamma_\tau(x) dB(\tau)$, представляющие собой случайные функции, требуют особого рассмотрения из-за присутствия в них дифференциала винеровского процесса. При условии интегрируемости функции $\gamma_t(x)$ с квадратом: $M \left\{ \int_{t_0}^t \gamma_\tau^2(x) d\tau \right\}$, представим стохастический интеграл, как предел в среднеквадратическом смысле соответствующих интегральных сумм

$$\int_{t_0}^t \gamma_\tau(x) dB(\tau) = \lim_{\substack{\Delta t \rightarrow 0 \\ m \rightarrow \infty}} \sum_{k=0}^{m-1} \gamma_{\tau_k}(x(\tau_k)) [B(t_{k+1}) - B(t_k)], \quad (2.18)$$

где $t_0 < \dots < t_m = t$, подынтервал $\Delta t = \max_k (t_{k+1} - t_k)$.

Рассмотрим два случая положения временного отсчета t_k : в середине подынтервала $\tau_k = (t_{k+1} + t_k)/2$ и в начале подынтервала $\tau_k = t_k$.

В силу дифференцируемости функции $\gamma_t(x)$ ее можно представить в виде разложения в ряд Тейлора в окрестности точки с координатой $x(\tau_k)$. Для первого случая $x(\tau_k) = \frac{x(t_{k+1}) + x(t_k)}{2} + 0(\Delta t)$ и это разложение принимает вид

$$\begin{aligned} \gamma_{\tau_k}(x(\tau_k)) &= \gamma_{\tau_k}\left(\frac{x(t_{k+1}) + x(t_k)}{2}\right) = \\ &= \gamma_{\tau_k}(x(t_k)) + \frac{1}{2} \frac{\partial \gamma_{t_k}(x_{t_k})}{\partial x_{t_k}} \times [x(t_{k+1}) - x(t_k)] + 0[x(t_{k+1}) - x(t_k)]. \end{aligned}$$

В результате интеграл (2.18) преобразуется к виду

$$\begin{aligned} J^* &= \int_{t_0}^t \gamma_{\tau}(x) dB(\tau) = \\ &= \lim_{\substack{\Delta t \rightarrow 0 \\ m \rightarrow \infty}} \sum_{k=0}^{m-1} \gamma_{\tau_k}\left(\frac{x(t_{k+1}) + x(t_k)}{2}\right) [B(t_{k+1}) - B(t_k)] = \\ &= \lim_{\substack{\Delta t \rightarrow 0 \\ m \rightarrow \infty}} \sum_{k=0}^{m-1} \left\{ \gamma_{t_k}(x(t_k)) + \frac{1}{2} \frac{\partial \gamma_{t_k}}{\partial x_{t_k}} [x(t_{k+1}) - x(t_k)] \right\} [B(t_{k+1}) - B(t_k)]. \end{aligned} \quad (2.19)$$

Выразим дифференциал в уравнении (2.15) в виде малого приращения

$$\begin{aligned} x(t_{k+1}) - x(t_k) &= \\ &= f_{\tau_k}(x(\tau_k))(t_{k+1} - t_k) + \gamma_{\tau_k}(x(\tau_k)) [B(t_{k+1}) - B(t_k)] + 0(\Delta t). \end{aligned} \quad (2.20)$$

Подставим выражение (2.20) в соотношение (2.19), получаем

$$J^* = J + \lim_{\substack{\Delta t \rightarrow 0 \\ m \rightarrow \infty}} \frac{1}{2} \sum_{k=0}^{m-1} \gamma_{t_k}(x(t_k)) \frac{\partial \gamma_{t_k}(x(t_k))}{\partial x_{t_k}} [B(t_{k+1}) - B(t_k)]^2 + 0(\Delta t).$$

Или с учетом (2.13) при $\Delta t \rightarrow 0$ и, пренебрегая членами порядка $0(\Delta t)$, имеем

$$J^* = J + \frac{\mathcal{N}_\eta}{2} \int_{t_0}^t \gamma_\tau(x) \frac{\partial \gamma_\tau(x)}{\partial x_\tau} d\tau, \quad (2.21)$$

где J обозначает интеграл (2.18) при $\tau_k = t_k$, т.е. когда временной отсчет берется в начале подынтервала.

Таким образом, получены различные значения интегралов J^* и J . Формальное отличие состоит в том, что при формировании интегралов значение функции $\gamma_t(x)$ в первом случае берется в середине подынтервала, во втором – на левом его конце.

Отметим, во-первых, что если вместо винеровского использовать обычную дифференцируемую функцию, то квадрат ее приращения имеет порядок Δt^2 и при $\Delta t \rightarrow 0$ второе слагаемое (2.21) пропадает. Во-вторых, в случае независимости функции $\gamma(\cdot)$ от x это слагаемое также пропадает. В результате разница между интегралами J^* и J исчезает. Интеграл J^* был введен Стратоновичем и назван симметризованным, а дифференциальное уравнение (2.15) при $\tau_k = (t_k + t_{k+1})/2$ – стохастическим симметризованным уравнением. Интеграл J и дифференциальное уравнение (2.15) при $\tau_k = t_k$ названы соответственно стохастическим интегралом и стохастическим уравнением в форме Ито.

Стохастические симметризованные интеграл и дифференциальные уравнения обладают важным свойством: для них справедливы обычные правила интегрирования и дифференцирования. Стохастические интеграл и уравнения в форме Ито этими свойствами не обладают. В результате использования симметризованной формы записи недифференцируемые случайные процессы приобретают свойства гладких дифференцируемых моделей. Однако, форма записи Ито также обладает рядом достоинств: записи стохастических выражений оказываются короче. Кроме того, они обладают рядом свойств, которые облегчают анализ алгоритмов обработки сигналов. В дальнейшем в пособии используются представления стохастических интегралов и уравнений в форме Ито.

2.3. МАРКОВСКИЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

Имеется класс математических моделей, аппроксимирующих широкий спектр случайных процессов и позволяющих решать многие практически важные задачи благодаря хорошо разработанному математическому аппарату. Такой класс составляют марковские процессы.

В этот класс включают гауссовские и негауссовские, непрерывные и импульсные сигналы и помехи. В качестве вырожденных случаев в него входят детерминированные и квазидетерминированные, а также пуассоновские процессы.

2.3.1. ОПРЕДЕЛЕНИЕ МАРКОВСКИХ ПРОЦЕССОВ

Основные виды марковских процессов классифицируются в соответствии с задаваемыми значениями на временных и числовых множествах. В зависимости от этого, непрерывные или дискретные значения принимает случайная величина на множествах X и ее аргумента T , рассматриваются следующие виды марковских процессов: 1) марковские цепи (дискретный процесс с дискретным временем); 2) непрерывный процесс с дискретным временем (непрерывнозначные случайные последовательности); 3) дискретный случайный процесс (дискретный процесс с непрерывным временем); 4) непрерывный процесс. Кроме перечисленных случайных процессов возможны более сложные смешанные процессы [6]. Наряду со скалярным (одномерным) процессом в технических приложениях находят применение векторный (многомерный) процесс $\xi(t)$, состоящий из компонент $\xi_1(t), \dots, \xi_r(t)$. Виды реализаций перечисленных случайных процессов представлены на Рис. 2.1.

Приведем общее определение марковского процесса, которое будет конкретизировано при рассмотрении упомянутых выше процессов.

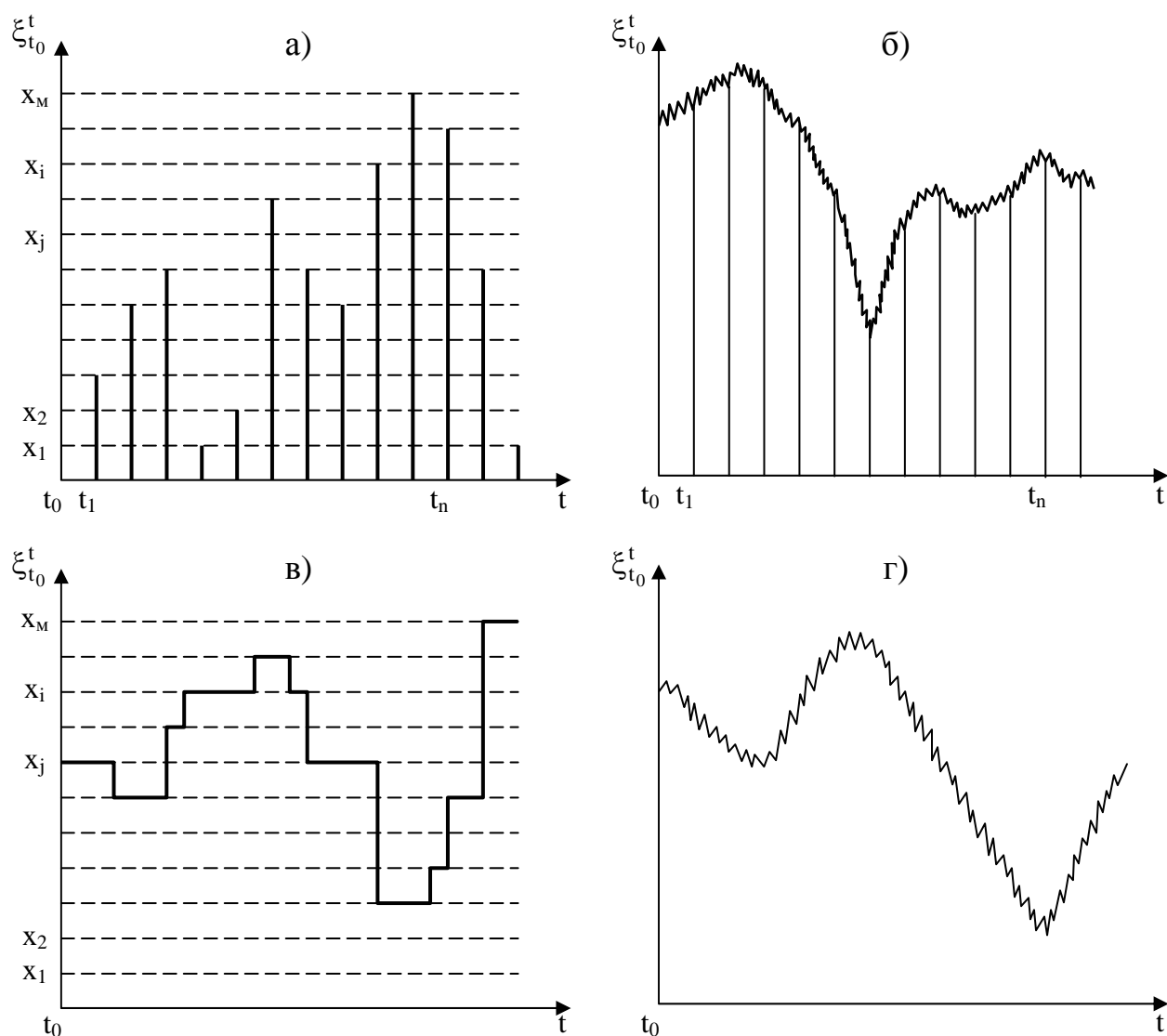


Рис. 2.1. Реализации марковских случайных процессов:
а – дискретный процесс с дискретным временем (марковская цепь);
б – непрерывный процесс с дискретным временем;
в – дискретный процесс с непрерывным временем;
г – непрерывный процесс.

Случайный процесс $\xi(t)$, $t \in (0, T)$ называется марковским, если для любых n моментов времени из интервала $(0, T)$ $t_1 < t_2 < \dots < t_n$ для условных распределений вероятностей справедливо соотношение

$$\begin{aligned} P\{\xi(t_n) \leq x_n \mid \xi(t_{n-1}) = x_{n-1}, \dots, \xi(t_1) = x_1\} = \\ = P\{\xi(t_n) \leq x_n \mid \xi(t_{n-1}) = x_{n-1}\} \end{aligned} \quad (2.22)$$

Такая запись означает, что при фиксированном состоянии процесса в настоящий момент времени t_{n-1} будущее для момента времени t_n не зависит от прошлого в моменты t_{n-2}, \dots, t_1 . Общим для

рассматриваемых процессов является свойство факторизации многомерных вероятностных характеристик, а именно представление этих характеристик (в форме вероятностей или плотностей вероятностей) в виде произведения одномерной начальной вероятности $P(x_1, t_1)$ на $n-1$ условных вероятностей перехода $\pi(x_i, t_i | x_{i-1}, t_{i-1})$, где индекс n соответствует моменту t_n ,

$$P\{\xi(t_n) = x_n, \dots, \xi(t_1) = x_1\} = P(x_1, t_1) \prod_{i=2}^n \pi(x_i, t_i | x_{i-1}, t_{i-1}).$$

С более подробным изложением теории, а также вопросов практического применения марковских процессов можно ознакомиться в работах [6-8].

2.3.2. ДИСКРЕТНЫЙ ПРОЦЕСС С ДИСКРЕТНЫМ ВРЕМЕНЕМ

Допустим, что случайный процесс $\xi(t)$ принимает конечное число значения из дискретного набора $x_i, \dots, x_j, \dots, x_i, \dots, x_k, \dots, x_M$ множества X . Для марковского дискретного процесса с дискретным временем t_n , $n=1, 2, \dots$ условие (2.22) имеет вид

$$\begin{aligned} P\{\xi(t_n) = x_{jn} | \xi(t_{n-1}) = x_{i, n-1}, \dots, \xi(t_1) = x_{k1}\} = \\ = P\{\xi(t_n) = x_{jn} | \xi(t_{n-1}) = x_{i, n-1}\} = \pi_{ji}(t_n | t_{n-1}). \end{aligned}$$

Условные вероятности перехода или переходные вероятности образуют квадратную матрицу переходных вероятностей $\Pi_n = \|\pi_{ji}(t_n | t_{n-1})\|$, $i, j=1, M$. Переходные вероятности удовлетворяют условиям неотрицательности $\pi_{ji}(t_n | t_{n-1}) \geq 0$ и нормировки $\sum_{i=1}^M \pi_{ji}(t_n | t_{n-1}) = 1$. Для полного определения цепи Маркова необходимо также задать вектор начальных вероятностей

$$P\{\xi(t_0) = x_1\} = p(x_1), \dots, P\{\xi(t_0) = x_j\} = p(x_j).$$

Если переходные вероятности не зависят от времени (индекса n), то цепь Маркова называется однородной. Для однородной цепи Маркова вероятность перехода за l шагов $P\{\xi(t_{n+l}) = x_j | \xi(t_n) = x_i\} = \pi_{ji}(l)$

определяется рекуррентным соотношением $\pi_{ji}(l) = \sum_{k=1}^M \pi_{jk}(1)\pi_{ki}(l-1)$.

Используя свойство условных вероятностей можно получить

$$\pi_{ji}(n+l) = \sum_{k=1}^M \pi_{jk}(n)\pi_{ki}(l). \quad (2.23)$$

Соотношение (2.23) является уравнением Колмогорова-Чепмена для переходных вероятностей. Зная вектор начальных вероятностей и вероятность перехода за n шагов можно найти безусловную вероятность

$$P\{\xi(t_n) = x_i\} = \sum_{j=1}^M p(x_j)\pi_{ji}(n).$$

2.3.3. НЕПРЕРЫВНЫЙ ПРОЦЕСС С ДИСКРЕТНЫМ ВРЕМЕНЕМ

В этом случае марковский процесс представляет собой последовательность непрерывных случайных величин

Пусть существует для этого процесса производная условной функции распределения

$$\frac{\partial P\{\xi(t_n) \leq x_n \mid \xi(t_{n-1}) = x_{n-1}\}}{\partial x_n} = \pi(x_n \mid x_{n-1}),$$

называемая переходной плотностью вероятности или плотностью вероятности перехода. В случае однородного процесса эта плотность не зависит от времени. Переходная плотность удовлетворяет условиям неотрицательности $\pi(x_n \mid x_{n-1}) \geq 0$ и нормировки $\int_{-\infty}^{\infty} \pi(x_n \mid x_{n-1}) dx_n = 1$. Уравнение Колмогорова-Чепмена

для рассматриваемого случая принимает вид

$$\pi(x_n \mid x_l) = \int_{-\infty}^{\infty} \pi(x_n \mid x_s)\pi(x_s \mid x_l) dx_s, \quad n > s > l.$$

По начальной плотности $p(x_1)$ и переходным вероятностям $\pi(x_i \mid x_{i-1})$ можно определить многомерную плотность вероятности

вектора $\{x_1, \dots, x_n\} : p(x_1, \dots, x_n) = p(x_1) \prod_{i=2}^n \pi(x_i | x_{i-1})$, выражающую свойства факторизации плотности вероятностей марковского непрерывного процесса с дискретным временем.

2.3.4. ДИСКРЕТНЫЙ ПРОЦЕСС С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ

Допустим, что процесс принимает конечные значения из дискретного набора $x_1, \dots, x_i, \dots, x_j, \dots, x_k, \dots, x_M$ множества X , но смена этих состояний происходит не в фиксированные, как в марковской цепи, а в произвольные моменты времени. Вероятности состояний $P\{\xi(t)=x_j\}=p_j(t)$ удовлетворяют условиям неотрицательности $p_j(t) \geq 0$ и нормировки $\sum_{j=1}^M p_j(t) = 1$ для любых t ($0 \leq t < \infty$). Марковское свойство для рассматриваемого процесса записывается в виде

$$P\{\xi(t)=x_i | \xi(s)=x_j, \dots, \xi(v)=x_k\} = P\{\xi(t)=x_i | \xi(s)=x_j\} \\ \text{при любых } t > s > \dots > v.$$

Переходные вероятности $P\{\xi(t)=x_i | \xi(s)=x_j\} = \pi_{ji}(s, t)$, $i, j = \overline{1, M}$ удовлетворяют условию неотрицательности $\pi_{ji}(s, t) \geq 0$ и нормировки $\sum_{j=1}^M \pi_{ji}(s, t) = 1$. Уравнение Колмогорова-Чепмена в этом случае принимает вид

$$\pi_{ji}(l, t) = \sum_{k=1}^M \pi_{jk}(l, s) \pi_{ki}(s, t). \quad (2.24)$$

Переходные вероятности (2.24) и вектор начальных условий $p_i(0) = P\{\xi(0)=x_i\}$, $i, j = \overline{1, M}$ полностью определяют рассматриваемый марковский процесс $p_i(t) = \sum_{j=1}^M p_j(0) \pi_{ji}(0, t)$.

Допустим далее, что для переходных вероятностей можно записать

$$\pi_{ji}(s, t) = \begin{cases} 1 - q_{jj}(t - s) + 0(t - s), & j = i, \\ q_{ji}(t - s) + 0(t - s), & j \neq i, \end{cases} \quad (2.25)$$

где $(t-s)$ - малый временной интервал, а $0(t-s)$ обозначает члены выше первого порядка малости относительно $(t-s)$. Как следует из соотношений (2.25), переход из одного состояния в любое другое возможное состояние зависит от рассматриваемого момента времени и пропорционален малому интервалу времени $(t-s)$. Кроме того, при $t=s$ система достоверно (с вероятностью равной единице) находится в состоянии j .

Функции $q_{ji}(t)$, $q_{ij}(t)$ неотрицательны. После подстановки переходных вероятностей (2.25) в уравнение (2.24), полагая $(t-s)=\Delta t \rightarrow 0$, получаем систему линейных дифференциальных уравнений

$$\frac{d\pi_{ji}(s,t)}{dt} = -q_{ii}(t)\pi_{ji}(s,t) + \sum_{\substack{k=1 \\ k \neq i}}^M \pi_{jk}(s,t)q_{ki}(t), \quad i,j = \overline{1,M}. \quad (2.26)$$

Умножив обе части уравнения (2.26) на $p(s)$, а затем просуммировав по всем значениям j , получим систему дифференциальных уравнений

$$\frac{dp_i(t)}{dt} = q_{ii}(t)p_i(t) + \sum_{\substack{k=1 \\ k \neq i}}^M p_k(t)q_{ki}(t), \quad i = \overline{1,M}, \quad (2.26)$$

определяющих эволюцию безусловных вероятностей состояний дискретного случайного процесса.

Дискретные марковские процессы широко применяются в практических задачах, в частности, при описании моделей случайных потоков, например, пуассоновского процесса. Рассмотрим однородный дискретный процесс, для которого q_{ji} являются постоянными величинами. Пуассоновский процесс описывается неубывающей целочисленной функцией $N_0^t = \{N_\tau, 0 \leq \tau \leq t\}$ при условии $j=i-1$ с постоянной интенсивностью изменения состояний $q_{ii}=q_{ji}=v>0$. Соотношения (2.25) для рассматриваемого случая при $s-t=\Delta t$ принимает вид

$$\pi_{ji}(\Delta t) \begin{cases} P\{N_o^{t+\Delta t} = i \mid N_o^t = i\} = 1 - \nu\Delta t + 0(\Delta t), \\ P\{N_o^{t+\Delta t} = i \mid N_o^t = i-1\} = \nu\Delta t + 0(\Delta t), \\ P\{N_o^{t+\Delta t} > i \mid N_o^t = i\} = 0(\Delta t). \end{cases}$$

Последнее равенство означает, что вероятность смены состояния более одного раза является малой величиной порядка $0(\Delta t)$. Оно выражает свойство ординарности. Для указанного процесса отличны от нуля только два параметра q_{ii} и q_{ji} . Поэтому уравнения (2.26) записываются в виде

$$\begin{aligned} dp_i(t) &= -\nu p_i(t)dt + \nu p_{i-1}(t)dt, \\ dp_0(t) &= -\nu p_0(t)dt. \end{aligned}$$

Решение этих дифференциальных уравнений при начальных условиях $p_i(0) = \begin{cases} 1, & i=0, \\ 0, & i=1,2,\dots \end{cases}$ приводит к системе рекуррентных соотношений

$$p_i(t) = \nu \exp\{-\nu t\} \int_0^t \exp\{\nu z\} p_{i-1}(z) dz, \quad i \geq 1, \quad (2.27)$$

$$p_0(t) = \exp\{-\nu t\}, \quad (2.28)$$

позволяющих последовательно для $i=1,2,\dots$ находить вероятности $p_i(t)$. В результате приходим к распределению Пуассона

$$p_i(t) = P\{N_o^t = i\} = (\nu t)^i \exp\{-\nu t\} / i!.$$

Отметим, что $p_0(t)$ является вероятностью отсутствия смены состояния (появления точки) на интервале $(0,t)$. При малом Δt вероятность (2.28) можно представить в виде разложения в ряд:

$$p_0(\Delta t) = 1 - \nu\Delta t + \frac{1}{2} \nu^2 \Delta t^2 - \dots = 1 - \nu\Delta t + 0(\Delta t). \quad (2.29)$$

Подставим (2.29) в (2.27) при $i=1$ и после замены t на Δt получим $p_1(\Delta t) = \nu\Delta t \exp\{-\nu\Delta t\}$. При малом Δt эту вероятность появления точки в интервале Δt также представим в виде разложения в ряд

$$p_1(\Delta t) = \nu \Delta t - \nu^2 \Delta t^2 - \dots = \nu \Delta t + O(\Delta t).$$

Отсюда следует $\nu = \lim_{\Delta t \rightarrow 0} p_1(\Delta t) / \Delta t$.

2.3.5. НЕПРЕРЫВНЫЙ МАРКОВСКИЙ ПРОЦЕСС

Процессы этого типа протекают во времени непрерывно и имеют непрерывные реализации. Марковское свойство (2.22) для этих процессов означает, что для последовательных моментов времени $t_1 < t_2 < \dots < t_n$ плотность вероятностей перехода в момент t_n зависит от значений x_{n-1} в момент t_{n-1} и не зависит от остальных более ранних значений

$$\pi(x_n, t_n | x_{n-1}, t_{n-1}; \dots; x_1, t_1) = \pi(x_n, t_n | x_{n-1}, t_{n-1}).$$

Безусловная плотность вероятностей равна

$$p(x_n, t_n; \dots; x_1, t_1) = \pi(x_n, t_n | x_{n-1}, t_{n-1}) p(x_{n-1}, t_{n-1}; \dots; x_1, t_1) \quad (2.30)$$

Используя последовательно соотношение (2.30), можно получить условие факторизации многомерной плотности вероятностей

$$p(x_n, t_n; \dots; x_1, t_1) = p(x_1, t_1) \prod_{i=2}^n \pi(x_i, t_i | x_{i-1}, t_{i-1}). \quad (2.31)$$

Полученное выражение определяет многомерную плотность вероятностей марковского процесса через одномерную начальную плотность и произведения плотностей вероятностей перехода. Для простоты последующих расчетов в выражении для плотности перехода вместо произвольных моментов времени t_n, t_{n-1}, \dots используем моменты $t > t' > t_0$. Указанные плотности удовлетворяют условию неотрицательности $\pi(x, t | x_0, t_0) \geq 0$ и нормировки

$\int_{-\infty}^{\infty} \pi(x, t | x_0, t_0) dx = 1$. Уравнение Колмогорова-Чепмена имеет вид

$$\pi(x, t | x_0, t_0) = \int_{-\infty}^{\infty} \pi(x, t | x', t') \pi(x', t' | x_0, t_0) dx'. \quad (2.32)$$

Важное место среди марковских процессов занимают диффузионные процессы, имеющие непрерывные с вероятностью единица реализации. Статистические свойства диффузионного марковского процесса характеризуются коэффициентами сноса $K_1(x,t)$ и диффузии $K_2(x,t)$, которые соответственно равны

$$K_1(x,t) = \lim_{\Delta t \rightarrow 0} M \left\{ \frac{x(t + \Delta t) - x(t)}{\Delta t} \middle| x(t) \right\}, \quad (2.33)$$

$$K_2(x,t) = \lim_{\Delta t \rightarrow 0} M \left\{ \frac{[x(t + \Delta t) - x(t)]^2}{\Delta t} \middle| x(t) \right\}. \quad (2.34)$$

Коэффициенты более высокого порядка

$$K_n(x,t) = \lim_{\Delta t \rightarrow 0} M \left\{ \frac{[x(t + \Delta t) - x(t)]^n}{\Delta t} \middle| x(t) \right\} \quad (2.35)$$

полагаем равными нулю. Символ $M\{\cdot\}$ обозначает условные моменты приращения $x(t)$ за время Δt . Коэффициенты K_1 и K_2 характеризуют в первом случае среднее значение скорости изменения, во втором – скорость изменения дисперсии приращения марковского процесса в момент времени t .

Предположим, что плотность вероятностей перехода можно представить в виде разложения в ряд Тейлора по степеням $(x' - x_0)$:

$$\pi(x,t | x',t') = \pi(x,t | x_0,t_0) + \sum_{k=1}^{\infty} \frac{(x' - x_0)^k}{k!} \frac{\partial^k \pi(x,t | x_0,t_0)}{\partial x_0^k}. \quad (2.36)$$

После подстановки соотношения (2.36) в (2.32), интегрирования, а затем предельного перехода $(t' - t_0) \rightarrow 0$ с учетом (2.33) и (2.34), а также $K_n = 0, n \geq 3$ получаем прямое уравнение Колмогорова

$$\begin{aligned} \frac{\partial \pi(x,t | x_0,t_0)}{\partial t} = & - \frac{\partial}{\partial x} \{ K_1(x,t) \pi(x,t | x_0,t_0) \} + \\ & + \frac{1}{2} \frac{\partial^2}{\partial x^2} \{ K_2(x,t) \pi(x,t | x_0,t_0) \} \end{aligned}$$

ИЛИ

$$\frac{\partial \pi(x, t | x_0, t_0)}{\partial t} = L\{\pi(x, t | x_0, t_0)\}, \quad (2.37)$$

где $L\{\cdot\} = -\frac{\partial}{\partial x} K_1(x, t)\{\cdot\} + \frac{1}{2} \frac{\partial^2}{\partial x^2} K_2(x, t)\{\cdot\}$ - дифференциальный оператор.

Получено дифференциальное уравнение в частных производных второго порядка параболического типа марковского диффузионного процесса. Это уравнение также называют уравнением Фоккера-Планка-Колмогорова. Условие (2.35) характеризует быстроту уменьшения вероятности больших отклонений с уменьшением Δt . Поэтому среднее приращение диффузионного процесса за время Δt имеет порядок $\sqrt{\Delta t}$, хотя допускаются более быстрые изменения самого процесса $x(t)$ в противоположных направлениях.

Умножив обе части уравнения (2.37) на плотность вероятностей $p(x_0, t_0)$ и проинтегрировав затем по x_0 , получаем уравнение ФПК для безусловной плотности вероятностей

$$\frac{\partial p(x, t)}{\partial t} = L\{p(x, t)\}. \quad (2.38)$$

Для решения уравнения (2.38) необходимо задать начальную плотность вероятностей $p(x_0, t_0)$. В частности, если все реализации процесса начинаются из точки x_0 в момент времени t_0 , то $p(x_0, t_0) = \delta(x - x_0)$. Для однородного диффузионного процесса, когда переходная плотность вероятностей зависит от разностей координат $(x - x_0)$ и временных параметров $(t - t_0)$ коэффициенты K_1 и K_2 оказываются постоянными величинами.

На примере ранее рассмотренного винеровского процесса покажем, что он относится к марковским процессам. Для трех моментов времени t_1 , t_2 и t_3 на основании статической независимости приращения винеровского процесса на неперекрывающихся интервалах с учетом уравнения (2.5) можно записать

$$B(t_3) = \int_0^{t_2} \eta(\tau) d\tau + \int_{t_2}^{t_3} \eta(\tau) d\tau = B(t_2) + \int_{t_2}^{t_3} \eta(\tau) d\tau, \quad t_3 > t_2 > t_1 > 0.$$

$$\text{Но } B(t_2) = \int_0^{t_1} \eta(\tau) d\tau + \int_{t_1}^{t_2} \eta(\tau) d\tau = B(t_1) + \int_{t_1}^{t_2} \eta(\tau) d\tau.$$

Таким образом, при фиксированном $B(t_2)$ будущее значение $B(t_3)$ не зависит от прошлого значения $B(t_1)$, что подтверждает марковское свойство винеровского процесса.

Нетрудно получить коэффициенты сноса и диффузии для винеровского процесса

$$\begin{aligned} K_1(x, t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{t-\Delta t}^t M\{\eta(\tau)\} d\tau = 0, \\ K_2(x, t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{t-\Delta t}^t \int_{t-\Delta t}^t M\{\eta(\tau_1)\eta(\tau_2)\} d\tau_1 d\tau_2 = \\ &= \lim_{\Delta t \rightarrow 0} \frac{\mathcal{N}_\eta}{\Delta t} \int_{t-\Delta t}^t d\tau = \mathcal{N}_\eta. \end{aligned}$$

и уравнение ФПК принимает вид
$$\frac{\partial p(x, t)}{\partial t} = \frac{\mathcal{N}_\eta}{2} \frac{\partial^2 p(x, t)}{\partial x^2}.$$

Ввиду того, что коэффициент сноса равен нулю, такой процесс называется чисто диффузионным процессом.

Рассмотрим необходимое в дальнейшем приращение плотности вероятности перехода на малом интервале времени $(t, t+\Delta t)$. Обозначим процесс для двух близкоотстоящих моментов времени t и $t+\Delta t$ через x_t и $x_{t+\Delta t}$. При таком разбиении будем считать процесс x_t постоянным на этом интервале и равным своему значению в начале интервала. Это практически выполняется, если принять $\Delta t \ll \tau_x$, где τ_x - интервал корреляции x_t . Для рассматриваемого случая уравнение (2.37) записывается в виде

$$\frac{\partial \pi(x_{t+\Delta t} | x_t)}{\partial t} = L\{\pi(x_{t+\Delta t} | x_t)\} \quad (2.39)$$

при начальном условии $\lim_{\Delta t \rightarrow 0} \pi(x_{t+\Delta t} | x_t) = \delta(x_{t+\Delta t} - x_t)$, где $\delta(\cdot)$ - дельта-функция.

Для малых интервалов Δt имеет место приближенное равенство $\frac{\partial \pi(x_{t+\Delta t} | x_t)}{\partial t} = \frac{\Delta \pi_t}{\Delta t}$, где $\Delta \pi_t$ - приращение плотности за время Δt .

Поэтому выражение (2.39) при пренебрежении членами второго порядка малости и выше $O(\Delta t)$ переписывается в форме $\Delta \pi_t = \Delta t L \{ \delta(x_{t+\Delta t} - x_t) \}$. При малом Δt плотность вероятностей перехода для момента времени t равна начальной плотности $\delta(x_{t+\Delta t} - x_t)$ в момент времени t и приращению за время Δt

$$\pi(x_{t+\Delta t} | x_t) \approx \delta(x_{t+\Delta t} - x_t) + \Delta t L \{ \delta(x_{t+\Delta t} - x_t) \} \quad (2.40)$$

2.3.6. МНОГОМЕРНЫЙ МАРКОВСКИЙ ПРОЦЕСС

Обобщим полученные результаты на многомерный марковский диффузионный процесс $x(t) = \|x_1(t) \dots x_r(t)\|^T$, где $\{x_i, i = \overline{1, r}\}$ - компоненты вектора $x(t)$.

Марковское свойство для рассматриваемого случая записывается в форме

$$\begin{aligned} P\{\xi_1(t_n) \leq x_1, \dots, \xi_r(t_n) \leq x_r \mid \xi_1(t_{n-1}), \dots, \xi_1(t_1); \dots; \xi_r(t_{n-1}), \dots, \xi_r(t_1)\} = \\ = P\{\xi_1(t_n) \leq x_1, \dots, \xi_r(t_n) \leq x_r \mid \xi_1(t_{n-1}), \dots, \xi_r(t_{n-1})\}. \end{aligned}$$

Многомерное уравнение ФПК имеет вид

$$\frac{\partial p(x, t)}{\partial t} = L^r \{ p(x, t) \}, \quad (2.41)$$

где многомерный дифференциальный оператор

$$\begin{aligned} L^r \{ \cdot \} &= - \sum_{i=1}^r \frac{\partial}{\partial x_i} K_{1i}(x, t) \{ \cdot \} + \frac{1}{2} \sum_{i,j=1}^r \frac{\partial^2}{\partial x_i \partial x_j} K_{2ij}(x, t) \{ \cdot \}, \\ K_{1i}(x, t) &= \lim_{\Delta t \rightarrow 0} M \left\{ \frac{x_i(t + \Delta t) - x_i(t)}{\Delta t} \middle| x(t) \right\}, \\ K_{2ij}(x, t) &= \lim_{\Delta t \rightarrow 0} M \left\{ \frac{[x_i(t + \Delta t) - x_i(t)][x_j(t + \Delta t) - x_j(t)]}{\Delta t} \middle| x(t) \right\}, \quad i, j = \overline{1, r}. \end{aligned}$$

Отметим, что каждая из компонент многомерного процесса не является одномерным марковским процессом, поскольку зависит от

всех остальных или части компонент и поэтому не удовлетворяет марковскому свойству. Только в одном частном случае векторного марковского процесса, представляющего собой набор независимых процессов, рассматриваемые компоненты являются одномерными марковскими процессами.

2.3.7. АППРОКСИМАЦИЯ РЕАЛЬНЫХ СИГНАЛОВ МАРКОВСКИМИ ПРОЦЕССАМИ

Как уже отмечалось, марковские модели процессов широко используются при анализе и синтезе систем управления и связи, в различных задачах информационного обеспечения, статистической обработки результатов наблюдения и т.д. Свойство марковости является определенным ограничением на используемые реальные сигналы, но вполне достаточным для разработки физически содержательных методов анализа и оптимального синтеза. С одной стороны, это свойство не сужает существенно область применения оптимальных процедур обработки информации, с другой - не усложняет экспоненциально с ростом объема выборки (отсчетов) вычислительный процесс. Несомненным достоинством марковских процессов является возможность использования в расчетной практике рекуррентных (многошаговых) процедур. Применение этих процедур в решаемых на ЭВМ задачах анализа и синтеза существенно упрощает реализацию алгоритмов функционирования устройств информационных систем, так как используются последние результаты наблюдений и некоторые статистики от всех предшествующих наблюдений. Это обеспечивается упомянутым ранее свойством факторизации многомерной вероятности марковского процесса: представлением указанной вероятности (или плотности вероятностей) в виде произведения одномерной начальной вероятности на последовательность условных вероятностей перехода.

При составлении модели динамических систем используется информационный подход, основанный на преобразовании и оценке сигналов в соответствии с требуемыми алгоритмами обработки и управления. Динамика поведения этих моделей описывается с помощью переменных состояния (фазовых координат системы) и дифференциальных уравнений первого порядка относительно этих переменных. Очевидно, любая сложная динамическая система является многомерной и ее состояние характеризуется

совокупностью переменных состояния системы, объединенных в вектор состояния. Учитывая, что, в основном, реальная динамическая система функционирует под воздействием случайных управляющих сигналов, помех и шумов, состояние системы в любой момент времени является случайным.

Рассмотрение начнем с находящейся под воздействием входного непрерывного случайного сигнала системы, описываемой нелинейным дифференциальным уравнением первого порядка

$$\frac{dx(t)}{dt} = f_t(x) + \gamma_t(x)\xi(t), \quad x_{t_0} = x(0), \quad (2.42)$$

где $f_t(\cdot)$ и $\gamma_t(\cdot)$ - детерминированные дифференцируемые функции своих аргументов, удовлетворяющие условию регулярности $|f_t(x)|^2 + |\gamma_t(x)|^2 \leq k^2(1 + |x|^2)$, k - постоянная величина, $\xi(t)$ - входной стационарный центрированный случайный процесс с конечным временем корреляции τ_ξ . Приведем некоторые результаты, вытекающие из анализа уравнения (2.42). Если для случайной функции $\xi(t)$ ее интервал корреляции удовлетворяет условию

$$T_0 \approx \left| \frac{\partial f_t(x)}{\partial x} \right|^{-1} \gg \tau_\xi, \quad (2.43)$$

где T_0 - характерная постоянная времени системы, то вместо уравнения (2.42) можно использовать стохастическое дифференциальное уравнение (2.16)

$$\frac{dx(t)}{dt} = f_t(x) + \gamma_t(x)\eta(t), \quad x_{t_0} = x(0) \quad (2.44)$$

или в эквивалентной форме (2.15) уравнение

$$dx(t) = f_t(x)dt + \gamma_t(x)dB(t), \quad x_{t_0} = x(0), \quad (2.45)$$

решением которых является диффузионный марковский процесс, а $\eta(t)$ и $B(t)$ введенные в разд. 2.2. гауссовский белый шум и винеровский процесс. Интервал корреляции марковского процесса определяется характерной постоянной времени системы – T_0 . В силу ранее рассмотренных свойств винеровского процесса этот процесс не

дифференцируем, но сохраняет свойства непрерывности. Покажем, что случайный процесс, определяемый уравнением (2.45) является марковским процессом. Действительно, интеграл уравнения (2.17) из-за статистической независимости приращений винеровского процесса при $t < t_0$ и $t > t_0$ в момент времени $t > t_0$ зависит от x_0 и от значений правой части этого уравнения и не зависит от значений $x(t)$ до момента времени t_0 . Таким образом, правая часть этого уравнения при $t > t_0$ статистически независима от ее значений при $t < t_0$. Отсюда следует, что любое условное распределение плотности вероятностей интеграла уравнения относительно любой совокупности случайных величин $x(t_1), \dots, x(t_k), x_0$ при $t_1, \dots, t_k < t_0$ не зависит от $x(t_1), \dots, x(t_k)$ и определяется условной плотностью относительно x_0 . Поэтому рассматриваемый процесс по определению является марковским процессом. Итак, использование белого шума в качестве модели входного воздействия приводит к одномерному марковскому процессу. Причем замена процесса $\xi(t)$ на $\eta(t)$ оправдана, если не принимать во внимание характер процессов в системе на интервалах времени, соизмеримых с τ_ξ . Таким образом, для сравнительно инерционных систем, встречающихся, как правило, в технических приложениях, процессы будут определяться характерной постоянной времени системы. Это позволяет утверждать, что на интервалах времени значительно больших τ_ξ процессы, описываемые уравнениями (2.42) и (2.44), будут практически одинаковыми.

На основании (2.33) и (2.34) с учетом записи стохастического дифференциального уравнения (2.44) в форме Ито коэффициенты сноса и диффузии соответственно равны

$$K_1(x, t) = f_1(x),$$

$$K_2(x, t) = \mathcal{N}_\eta \gamma_t^2(x).$$

Если коэффициент сноса является линейной функцией $x(t)$, а коэффициент диффузии постоянной величиной, не зависящей от $x(t)$, то одномерная плотность вероятностей становится гауссовской. Как уже отмечалось, аналитический аппарат марковских процессов позволяет разработать общие методы определения статистических характеристик случайных процессов. Для гауссовских марковских процессов указанные методы достаточны просты и приводят к решению линейных дифференциальных уравнений относительно

математического ожидания и корреляционной функции (дисперсии) процесса. Отметим, что из-за необходимости выполнения условия уравнения (2.32), не каждый гауссовский процесс обладает марковскими свойствами. Действительно, после подстановки условной плотности (2.4) для моментов времени t_1, t_2 и t_2, t_3 в это уравнение, последующего интегрирования, получаем соотношение, из которого следует равенство для введенных в разделе 2.1. коэффициентов корреляции $r(t_1, t_3) = r(t_1, t_2)r(t_2, t_3)$.

Для стационарного процесса при $\tau_1 = t_2 - t_1$ и $\tau_2 = t_3 - t_2$ это выражение принимает форму $r(\tau_1 + \tau_2) = r(\tau_1)r(\tau_2)$. Этому соотношению соответствует корреляционная функция экспоненциального вида.

В качестве примера приведем уравнение напряжения на выходе апериодического звена (RC -цепочки) с постоянной времени $\alpha^{-1} = RC$ при воздействии на вход звена белого шума $\eta(t)$ со спектральной плотностью \mathcal{N}_η

$$\frac{dx(t)}{dt} = -\alpha x(t) + \gamma \eta(t), \quad \gamma = \alpha, \quad x_0 = x(0). \quad (2.46)$$

Для рассматриваемого случая коэффициенты сноса и диффузии соответственно равны $K_1 = -\alpha x(t)$, $K_2 = \mathcal{N}_\eta \alpha^2$.

Общее решение этого уравнения имеет вид

$$x(t) = x_0 \exp\{-\alpha t\} + \alpha \int_0^t \exp\{-\alpha(t - \tau)\} \eta(\tau) d\tau.$$

Математическое ожидание $x(t)$ с учетом $M\{\eta(t)\} = 0$ равно $m_x(t) = M\{x(t)\} = x_0 \exp\{-\alpha t\}$.

Корреляционная функция, принимая во внимание $M\{\eta(t_1)\eta(t_2)\} = \mathcal{N}_\eta \delta(t_2 - t_1)$,

$$\begin{aligned} K_{2x}(t_1, t_2) &= M\{[x(t_1) - m_x(t_1)][x(t_2) - m_x(t_2)]\} = \\ &= \mathcal{N}_\eta \int_0^{t_1} \int_0^{t_2} \exp\{-\alpha[(t_1 - \tau_1) + (t_2 - \tau_2)]\} \delta(t_2 - t_1) d\tau_1 d\tau_2. \end{aligned}$$

На основании фильтрующих свойств дельта-функции окончательно получаем

$$k_{2x}(t_1, t_1 + \tau) = \frac{\alpha N_\eta}{2} \exp\{-\alpha |\tau|\} (1 - \exp\{-2\alpha t_1\}), \quad \tau = t_2 - t_1.$$

Отсюда при $\tau=0$ дисперсия принимает вид

$$D_x(t) = \frac{\alpha N_\eta}{2} (1 - \exp\{-2\alpha t\}).$$

Для более точного описания реальных сигналов прибегают к многомерному марковскому процессу. Образуя векторный процесс $x(t) = \|x_1 \dots x_r\|^T$, можно заменить реальный исходный процесс компонентой r -мерного марковского процесса. С одной стороны, чем выше размерность марковского процесса, тем точнее аппроксимируется исходный процесс, с другой стороны, это усложняет модели сигналов, что приводит к более сложным синтезируемым алгоритмам обработки информации.

Рассмотрим механизм формирования гауссовского многомерного марковского процесса с использованием представлений о формирующем фильтре. Ранее было показано, при каких условиях реальный сигнал можно аппроксимировать одномерным марковским процессом. Рассмотрим случай, когда условие (2.43) не выполняется и интервал корреляции процесса $\xi(t)$ соизмерим с характерной постоянной времени системы. Можно представить, что воздействие $\xi(t)$ является реакцией некоторой фиктивной системы первого порядка на воздействие типа белого шума. Таким образом, приходим к двумерному марковскому процессу. В общем случае в зависимости от характера процесса фиктивная система (или формирующий фильтр) может иметь более высокую размерность и описываться стохастическими линейными дифференциальными уравнениями первого порядка (уравнениями состояния системы) относительно компонент (переменных состояний) векторного марковского процесса

Путем замены $s=d/dt$ непосредственно из уравнения (2.49) получают дробно-рациональную относительно s передаточную функцию формирующего фильтра

$$w(s) = \frac{\gamma_1 s^{r-1} + \gamma_2 s^{r-2} + \dots + \gamma_r}{s^r + f_1 s^{r-1} + \dots + f_r}.$$

По передаточной функции этого фильтра (после замены $s=j\omega$) и спектральной плотности белого шума \mathcal{N}_η на основании известной методики [9] определяют спектральную плотность компоненты $x_1(t)$ многомерного диффузионного марковского процесса

$$F_{x_1}(\omega) = |w(j\omega)|^2 \mathcal{N}_\eta.$$

Она также является дробно-рациональной относительно ω^2 функцией. Таким образом, гауссовский марковский случайный процесс имеет дробно-рациональную спектральную плотность. Следовательно, любой гауссовский случайный процесс с дробно-рациональным спектром может быть представлен марковским процессом. Приведенный результат имеет важное значение, так как спектральную плотность реального процесса всегда можно аппроксимировать с определенной степенью точности рациональной функцией частоты. И поэтому реальный даже, может быть, гауссовский, но не марковский процесс приближенно можно представить в виде марковского процесса.

Многомерный оператор уравнения ФПК для уравнения (2.48) принимает вид

$$L^r \{ \cdot \} = - \sum_{i=1}^r \frac{\partial}{\partial x_i} [F x(t)]_i \{ \cdot \} + \frac{1}{2} \sum_{i,j=1}^r [\Gamma \mathcal{N}_\eta \Gamma^T]_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \{ \cdot \},$$

где $[\Gamma \mathcal{N}_\eta \Gamma^T]_{ij}$ - (ij) компонента матрицы $\Gamma \mathcal{N}_\eta \Gamma^T$;

$[F x(t)]_i$ - i компонента вектора $F x(t)$.

Для общего уравнения многомерного диффузионного марковского процесса

$$\frac{dx(t)}{dt} = f_t(x) + \Gamma_t(x) \eta(t), \quad (2.50)$$

каждая из компонент которого задана стохастическим дифференциальным уравнением первого порядка в форме Ито

$$\frac{dx_i(t)}{dt} = f_{it}(x) + \sum_{j=1}^r \gamma_{ijt}(x) \eta_j(t),$$

многомерный оператор ФПК равен (2.41), где

$$K_{1i}(x, t) = f_{it}(x),$$

$$K_{2ij}(x, t) = \sum_{k=1}^r \mathcal{N}_{\eta_{kt}} \gamma_{ikt}(x) \gamma_{jkt}(x).$$

В технических приложениях также находят широкое применение сигналы, математическая модель которых описывается детерминированной функцией времени t и вектора случайных величин a : $S_t(a)$, где вектор a зависит от неизвестных параметров (a_1, \dots, a_k) . Эти параметры неизменны на интервале наблюдения, но имеют начальное распределение, т.е. являются случайными величинами. Процессы, зависящие от этих параметров, называются квазидетерминированными сообщениями. Такими процессами могут являться сигналы вида $x(t) = \sum_{j=1}^k a_j \varphi_j(t)$, где $\{\varphi_i(t)\}$ -

детерминированные функции времени. В частности, этим сигналом описывается равноускоренное (замедленное) движение объектов по координате x при $t_0=0$

$$x(t) = \sum_{j=1}^3 a_j t^{j-1}, \quad a = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}^T \quad (2.51)$$

$$\text{или } x(t) = x_0 + vt + a_y t^2/2,$$

$$\text{где } a_1 = x_0, \quad a_2 = v, \quad a_3 = a_y/2.$$

Очевидно, что квазидетерминированный процесс относится к вырожденному частному случаю марковского диффузионного процесса. От этой модели, в которой сигнал меняется случайным образом, всегда можно перейти к более простой с постоянными параметрами. Так как в конкретной реализации случайного процесса вектор параметров постоянен, необходимо считать $F=\Gamma=0$ и уравнение (2.48) принимает вырожденную форму $da_t=0$, решение

которого совпадает с заданным начальным условием $a_0 = a = //a_1 \quad a_2 \quad a_3//^T$.

Переходная плотность вероятностей для моментов времени t_{i-1} и t_i принимает дельтаобразный вид $\pi(a_i|a_{i-1})=\delta(a_i-a_{i-1})$, а многомерная плотность вероятностей $p(a_0, a_1, ..., a_n) = p(a_0) \prod_{i=1}^n \delta(a_i - a_{i-1})$, где $p(a_0)$ - начальная плотность вероятностей. При дальнейшем упрощении марковских моделей можно прийти к детерминированному процессу – известному сигналу $f(t)$. В этом случае многомерная плотность вероятностей имеет вид

$$p(x_0, x_1, ..., x_n) = \prod_{i=1}^n \delta(x_i - f(t_i)).$$

2.3.8. ВРЕМЕННАЯ ДИСКРЕТИЗАЦИЯ МАРКОВСКОГО ПРОЦЕССА

Применение марковских моделей процессов, как уже ранее отмечалось, существенно упрощает реализацию алгоритмов обработки информации, так как используются последние результаты наблюдения (измерения) и некоторые статистики от всех предшествующих наблюдений. В этом случае указанные алгоритмы оказываются рекуррентными, функционирующие в режиме дискретного времени. Они описываются системой уравнений первого порядка в конечных разностях (разностными уравнениями). Кроме того, указанный подход соответствует характеру вычислительного процесса цифровых ЭВМ: чередованию временных моментов приема отсчетов входного сигнала и выдачи результатов расчета, отделенных друг от друга интервалом обработки (расчета) входного сигнала.

Пусть поведение системы характеризуется одномерным стохастическим нелинейным дифференциальным уравнением (2.45). Решением этого уравнения является рассмотренное в разделе (2.22) интегральное представление в форме Ито

$$x(t) = x_0 + \int_{t_0}^t f_{\tau}(x) d\tau + \int_{t_0}^t \gamma_{\tau}(x) dB(\tau),$$

где дифференциал винеровского процесса $dB(\tau) = \eta(\tau) d\tau$.

Представим интервал (t_0, t) в виде непрерывной последовательности неперекрывающихся, примыкающих друг к другу одинаковых подынтервалов $\Delta t = t_{n+1} - t_n$, называемых шагом дискретизации, $t - t_0 = [k]\Delta t$, где $[k]$ - ближайшее целое число, удовлетворяющее соотношению $t - t_0 \geq k\Delta t$, $[k] = \overline{1, n}$.

Запишем интегральное представление для малого шага дискретизации. Учитывая непрерывность функций $f_\tau(\cdot)$ и $\gamma_\tau(\cdot)$, получим реализующую рекуррентную процедуру уравнение в конечной разности первого порядка

$$\hat{x}(t_{n+1, \varepsilon}) = \hat{x}(t_{n, \varepsilon}) + f(\hat{x}, t_{n, \varepsilon})\Delta t + \gamma(\hat{x}, t_{n, \varepsilon})\Delta B(t_{n, \varepsilon}) \quad (2.52)$$

где $0 \leq \varepsilon \leq 1$, $t_{n+1, \varepsilon} = \varepsilon t_{n+1} + (1 - \varepsilon)t_n$, $\Delta B(t_{n, \varepsilon}) = \eta(t_{n, \varepsilon})\Delta t$.

С помощью параметра ε отсчет приближенного значения \hat{x} фиксируется внутри интервала (t_n, t_{n+1}) . Действительно, это уравнение является рекуррентным, так как позволяет при увеличении n шаг за шагом, используя независимость приращений винеровского процесса на неперекрывающихся интервалах времени, получить последовательность дискретных отсчетов оценок процесса \hat{x} . Неточность полученных решений связана с тем, что уравнению (2.52) можно сопоставить бесконечное число уравнений в конечных разностях (через точки отсчетов можно провести бесконечное число кривых). Вследствие этого точное решение уравнения (2.52) представляет чрезвычайно сложную задачу. Кроме того, в правой части уравнения (2.52) присутствует белый шум, дисперсия которого бесконечна. В то же время в реальных условиях использования рекуррентных процедур в цифровых ЭВМ дискретные отсчеты случайных процессов должны соответствовать сигналам с ограниченной дисперсией. Конструктивное решение этих проблем может быть осуществлено сглаживанием сигналов на шаге дискретизации и приданию параметру ε определенного значения с последующей оценкой погрешности полученного решения.

Путем сглаживания на шаге дискретизации получают дискретный белый шум

$$\eta_n = \frac{1}{\Delta t} \int_{t_n}^{t_n + \Delta t} \eta(t) dt.$$

Действительно, дисперсия в этом случае на основании фильтрующих свойств δ -функции принимает конечное значение, зависящее от Δt

$$\begin{aligned} D_{\eta n} &= M\{\eta^2(t_n)\} = \frac{1}{\Delta t^2} \int_{t_n}^{t_n+\Delta t} \int_{t_n}^{t_n+\Delta t} M\{\eta(t_1)\eta(t_2)\} dt_1 dt_2 \\ &= \frac{1}{\Delta t^2} \int_{t_n}^{t_n+\Delta t} \int_{t_n}^{t_n+\Delta t} \mathcal{N}_{\eta n} \delta(t_2 - t_1) dt_1 dt_2 = \frac{\mathcal{N}_{\eta n}}{\Delta t}, \end{aligned}$$

и при $\Delta t \rightarrow 0$ бесконечно возрастает, что соответствует переходу к непрерывному времени. Математическое ожидание дискретного белого шума равно нулю

$$m_{\eta n} = M\{\eta(t_n)\} = \frac{1}{\Delta t} \int_{t_n}^{t_n+\Delta t} M\{\eta(t)\} dt = 0.$$

Таким образом, дискретный белый шум представляет собой последовательность независимых гауссовских случайных величин $\{\eta_n\}$ с нулевым математическим ожиданием и дисперсией $D_{\eta n} = \mathcal{N}_{\eta n} / \Delta t$.

Аналогичным образом в качестве временных дискретных отсчетов марковских процессов также необходимо использовать осредненные значения

$$x_n = \frac{1}{\Delta t} \int_{t_n}^{t_n+\Delta t} x(t) dt.$$

На практике, не внося сколько-нибудь существенную погрешность, учитывая малое значение шага дискретизации и непрерывность $x(t)$, ограничиваются временными отсчетами марковского процесса.

Достаточно просто приближенные решения разностных уравнений в дискретные моменты времени могут быть получены при $\varepsilon=1$, $\hat{x}(t_n) = x_n$. В этом случае нелинейное уравнение (2.52) принимает форму уравнения в конечных разностях первого порядка

$$x_{n+1} = x_n + f_n(x_n) \Delta t + \gamma_n(x_n) \eta_n, \quad x_0 = x(0)$$

где η_n - дискретный белый шум, $\gamma_n(x_n)=\gamma(x,t_n)\Delta t$, $f_n(x_n)=f(x,t_n)$.

Для наиболее распространенного в практических расчетах класса марковских процессов полагают $\gamma_n(x_n)=\gamma(t_n)\Delta t=\gamma_n$. В этом случае используемые далее в задачах анализа и синтеза нелинейное или линейное уравнение одномерного марковского процесса в дискретном времени могут быть представлены соответственно в форме

$$x_{n+1}=x_n+\Delta t f_n(x_n)+\gamma_n \eta_n, \quad x_0=x(0) \quad (2.53)$$

или

$$x_{n+1}=x_n+\Delta t f_n x_n+\gamma_n \eta_n, \quad x_0=x(0) \quad (2.54)$$

Для нелинейной дискретной системы многомерный марковский процесс на основании (2.50) записывается в векторной форме

$$x_{n+1}=x_n+\Delta t f_n(x_n)+\Gamma_n \eta_n, \quad x_0=x(0)$$

где x_n и $f_n(\cdot)$ – r -мерные векторы соответственно марковского процесса и вектор-функции; Γ_n - диагональная переходная матрица воздействия размера $(r \times r)$ с компонентами γ_{nii} ; η_n - r -мерный вектор дискретного белого шума, корреляционная матрица которого равна $M\{\eta_n \eta_l\}=Q_{\eta n} \delta_{nl}$, δ_{nl} - символ Кронекера: $\delta_{nl}=1$ при $n=l$ и $\delta_{nl}=0$ при $n \neq l$; матрица размера $(r \times r)$ $Q_{\eta n}$ равна: $Q_{\eta n}=\mathcal{N}_{\eta, t=n\Delta t}/\Delta t$, где $\mathcal{N}_{\eta, t}$ - неотрицательная диагональная матрица спектральных плотностей вектора белых шумов η_t .

В случае линейных дискретных систем, описываемых многомерным гауссовским марковским процессом, рекуррентное уравнение записывается в векторной форме

$$x_{n+1}=\Phi_n x_n+\Gamma_n \eta_n, \quad x_0=x(0) \quad (2.55)$$

где $\Phi_n=I+\Delta t F_n$ - переходная матрица состояния размера $(r \times r)$, I - единичная матрица; F_n - нестационарная матрица дискретных значений компонент f_{jn} уравнений состояния (2.48).

Полученные при $\varepsilon=1$ разностные уравнения, как уже было отмечено ранее, являются приближенными. В связи с этим необходимо оценить погрешность при переходе от стохастических дифференциальных уравнений к рекуррентным уравнениям с

дискретным временем. Среднеквадратическая погрешность в результате этого перехода в интервале $(0, T)$ на основании данных работы [10] определяется формулой

$$\sigma \leq \Delta t \left[\int_0^T M \{ [f'_t(x) \gamma_t(x)]^2 + [L_t(\gamma_t(x))]^2 \mid x(0) \} dt \right]^{1/2}, \quad (2.56)$$

где оператор $L_t = \frac{\partial}{\partial t}(\cdot) + f_t(x) \frac{\partial}{\partial x}(\cdot) + \frac{1}{2} \gamma_t^2(x) \frac{\partial^2}{\partial x^2}(\cdot)$.

С помощью неравенства (2.56) можно выбрать шаг дискретизации Δt . Например, уравнению (2.46) соответствует разностное уравнение одномерного гауссовского марковского процесса

$$x_{n+1} = \Phi_n x_n + \gamma_n \eta_n, \quad x_0 = x(0), \quad (2.57)$$

где $\Phi_n = 1 - \alpha \Delta t$, $\gamma_n = \alpha \Delta t$.

Имеем $f_t(x) = -\alpha x$, $\gamma_t(x) = \alpha \mathcal{N}_\eta^{1/2}$, откуда $\sigma \leq \alpha^2 \Delta t (\mathcal{N}_\eta T)^{1/2}$.

2.4. СЛУЧАЙНЫЕ ТОЧЕЧНЫЕ ПРОЦЕССЫ

Исследования динамики систем, моделями сигналов в которых служат точечные процессы (случайные потоки), получили в последнее время развитие в связи с освоением для передачи информации наиболее коротковолнового диапазона электромагнитного излучения (оптического, рентгеновского), внедрением информационно-управляющих систем, функционирующих на основе использования современных компьютерных технологий, при анализе отказов в сложных динамических системах, изучении транспортных потоков и т.д.

2.4.1. ИНТЕГРАЛЬНЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ПОТОКОВ

Случайный поток $N(\tau) = \{N_{t_0}^\tau, t_0 \leq \tau \leq t\}$ образуют неразличимые (одинаковые) точки (события), выпавшие по случайным законам на временном интервале (t_0, t) . Реализацию случайного потока можно представить в виде неубывающей ступенчатой функции $N_{t_0}^\tau$,

принимаящей целые неотрицательные значения, моменты роста (смены состояния) которой являются случайными (Рис. 2.2). Величина ступенек также является, вообще говоря, случайной. В дальнейшем будут рассматриваться ординарные потоки, т.е. потоки исключающие взаимное наложение точек. Поэтому величина ступеньки реализации случайного потока будет всегда равняться единице. Еще одним видом реализации случайного потока является последовательность дельта-импульсов (случайный импульсный поток), положение которых совпадает с координатами появления точек.

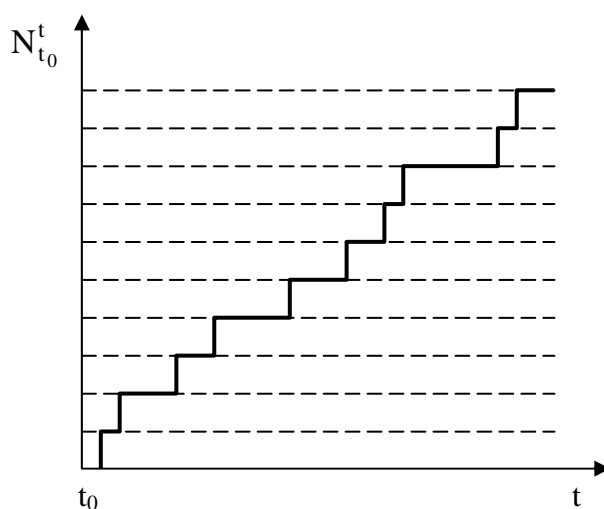


Рис. 2.2. Реализация случайного точечного процесса.

Для описания свойств случайных потоков удобно пользоваться производящей функцией [11]. Зададимся вероятностью появлений n точек на интервале $(0, T) - P_n(T)$. Для этой вероятности справедливы условия неотрицательности $P_n(T) \geq 0$ и нормировки $\sum_{n=0}^{\infty} P_n(T) = 1$. Производящая функция (пф) определяется как разложение в ряд по степеням вспомогательного параметра z

$$\mathcal{G}(z, T) = \sum_{n=0}^{\infty} P_n(T) z^n. \quad (2.58)$$

Соотношение (2.58) соответствует среднему значению величины z^n , поэтому формально его можно представить в виде $\mathcal{G}(z, T) = M\{z^n\}$. Коэффициенты ряда (2.58) равняются вероятности появления на интервале $(0, T)$ точек и определяются из соотношения

$$P_n(T) = \frac{1}{n!} \left. \frac{d^n \mathcal{G}(z, T)}{dz^n} \right|_{z=0}. \quad (2.59)$$

Формальными операциями над пф, например, дифференцированием по z и приравниванием оснований различных степеней параметра z нулю или единицы можно получить другие полезные интегральные характеристики. Например, систему коэффициентов, называемых факториальными моментами

$$F_n(T) = \left. \frac{d^n \mathcal{G}(z, T)}{dz^n} \right|_{z=1}. \quad (2.60)$$

Разложение пф в ряд относительно этих коэффициентов имеет вид

$$\mathcal{G}(z, T) = 1 + \sum_{n=1}^{\infty} \frac{F_n(T)}{n!} (z-1)^n.$$

С помощью факториальных моментов вычислим наиболее существенные интегральные характеристики (статистики отсчетов) – математическое ожидание $m_N(T)$ и дисперсию $D_N(T)$ числа появления точек на интервале $(0, T)$. Указанные статистики по определению равны

$$m_N(T) = M\{n\} = \sum_{n=0}^{\infty} n P_n(T), \quad (2.61)$$

$$D_N(T) = M\{[n - m_N(T)]^2\} = \sum_{n=0}^{\infty} n^2 P_n(T) - \left[\sum_{n=0}^{\infty} n P_n(T) \right]^2. \quad (2.62)$$

Получим ряд промежуточных результатов из формулы (2.58)

$$\left. \frac{d \mathcal{G}(z, T)}{dz} \right|_{z=1} = \sum_{n=0}^{\infty} n P_n(T), \quad (2.63)$$

$$\left. \frac{d^2 \mathcal{G}(z, T)}{dz^2} \right|_{z=1} = \sum_{n=0}^{\infty} n^2 P_n(T) - \sum_{n=0}^{\infty} n P_n(T). \quad (2.64)$$

С другой стороны, используя (2.60), имеем

$$\left. \frac{d\mathcal{G}(z, T)}{dz} \right|_{z=1} = F_1(T), \quad \left. \frac{d^2 \mathcal{G}(z, T)}{dz^2} \right|_{z=1} = F_2(T)$$

С учетом (2.63) и (2.64) на основании соотношений для рассматриваемых статистик (2.61) и (2.62) окончательно получаем

$$m_N(T) = F_1(T), \quad (2.65)$$

$$D_N(T) = F_1(T) + F_2(T) - F_1^2(T) \quad (2.66)$$

Для описания модулированных потоков можно получить еще одну полезную систему коэффициентов $X_n(T)$:

$$X_n(T) = \left. \frac{d^n \ln \mathcal{G}(z, T)}{dz^n} \right|_{z=0}, \quad (2.67)$$

$$\mathcal{G}(z, T) = \exp \left\{ \sum_{n=1}^{\infty} \frac{X_n(T)}{n!} (z^n - 1) \right\}.$$

Математическое ожидание и дисперсия числа появления точек на интервале $(0, T)$ соответственно равны

$$m_N(T) = X_1(T),$$

$$D_N(t) = X_1(T) + X_2(T).$$

Коэффициент $X_2(T)$ определяет дополнительную по сравнению с дисперсией пуассоновского процесса – $X_1(T)$ составляющую дисперсии, обусловленную модулирующим сигналом. Такой процесс называется двойным стохастическим пуассоновским процессом или процессом с двойной случайностью (одна случайность порождена пуассоновским процессом, другая – модулирующим сигналом).

2.4.2. ФУНКЦИИ ПЛОТНОСТИ И КОРРЕЛЯЦИИ ПЛОТНОСТИ

Наиболее полное описание случайных потоков можно получить с помощью производящего функционала (ПФ), который имеет вид [11]

$$L[u, T] = M \left\{ \prod_{i=1}^n [1 + u(t_i)] \right\}, \quad (2.68)$$

где $u(t)$ - вспомогательная действительная функция. Усреднение производится по числу n и моментам появления точек t_i на рассматриваемом интервале $(0, T)$.

Как и в случае характеристического функционала ПФ выражается через локальные характеристики – свои системы моментных $f_n(\cdot)$ и корреляционных $g_n(\cdot)$ функций, которые назовем соответственно функциями плотности и корреляции плотности n -ого порядка. Функция плотности n -ого порядка $f_n(t_1, \dots, t_n)$ характеризует совместную вероятность появления n точек в каждом из неперекрывающихся подынтервалов Δt_i безоотносительно к появлению дополнительного числа точек на остальных Δt - подынтервалах интервала $(0, T)$:

$$p_n = f_n(t_1, \dots, t_n) \Delta t_1 \dots \Delta t_i \dots \Delta t_n + O(\Delta t),$$

$$\text{где } \Delta t = \max \Delta t_i, i = \overline{1, n}, \lim_{\Delta t \rightarrow 0} O(\Delta t) / \Delta t = 0.$$

Функция первого порядка $f_1(t)$ имеет особое значение и называется интенсивностью потока (ИП). Наряду с функцией плотности для описания потока вводят другую систему функций – корреляционных или, как их будем по аналогии называть, функции корреляции плотности $g_n(t_1, \dots, t_n)$. Естественно допустить, например, что функции корреляции плотности второго порядка выражаются через функции плотности следующим образом:
 $g_2(t_1, t_2) = f_2(t_1, t_2) - f_1(t_1)f_1(t_2).$

Функция $f_2(t_1, t_2)$ характеризует совместную вероятность появлений точек вблизи моментов времени t_1 и t_2 и при разнесении аргументов стремится к произведению $f_1(t_1)f_1(t_2)$, каждый из сомножителей которого характеризует вероятность независимых событий. Следовательно, функция $g_2(t_1, t_2)$ при разнесении аргументов стремится к нулю, что означает ослабление корреляционных связей.

Соотношения, связывающие функции плотности и функции корреляции плотности любого порядка, можно получить на основании ПФ. Последний выражается через указанные функции в форме функциональных рядов

$$L[u, T] = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^T \dots \int_0^T f_n(t_1, \dots, t_n) \prod_{i=1}^n u(t_i) dt_1 \dots dt_n, \quad (2.69)$$

$$L[u, T] = \exp \left\{ \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^T \dots \int_0^T g_n(t_1, \dots, t_n) \prod_{i=1}^n u(t_i) dt_1 \dots dt_n \right\}, \quad (2.70)$$

где соответствующие функции определяются путем функционального дифференцирования

$$f_n(t_1, \dots, t_n) = \frac{\delta^n L[u, T]}{\delta u(t_1) \dots \delta u(t_i) \dots \delta u(t_n)} \Big|_{u(t_i) = 0},$$

$$g_n(t_1, \dots, t_n) = \frac{\delta^n \ln L[u, T]}{\delta u(t_1) \dots \delta u(t_i) \dots \delta u(t_n)} \Big|_{u(t_i) = 0}. \quad (2.71)$$

Сравнивая выражения (2.69) и (2.70), можно получить аналогичные по форме, что для моментных и корреляционных функций, соотношения, связывающие $f(\cdot)$ и $g(\cdot)$,

$$\begin{aligned} f_1(t) &= g_1(t); \\ f_2(t_1, t_2) &= g_2(t_1, t_2) + g_1(t_1)g_1(t_2); \\ f_3(t_1, t_2, t_3) &= g_3(t_1, t_2, t_3) + g_1(t_1)g_2(t_2, t_3) + g_1(t_2)g_2(t_1, t_3) + \\ &\quad + g_1(t_3)g_2(t_1, t_2) + g_1(t_1)g_1(t_2)g_1(t_3); \\ &\dots \end{aligned} \quad (2.72)$$

$$\begin{aligned} g_1(t) &= f_1(t); \\ g_2(t_1, t_2) &= f_2(t_1, t_2) - f_1(t_1)f_1(t_2); \\ g_3(t_1, t_2, t_3) &= f_3(t_1, t_2, t_3) - f_1(t_1)f_2(t_2, t_3) - f_1(t_2)f_2(t_1, t_3) - \\ &\quad - f_1(t_3)f_2(t_1, t_2) + 2f_1(t_1)f_1(t_2)f_1(t_3); \\ &\dots \end{aligned} \quad (2.73)$$

Обратим внимание на глубокую аналогию между полученными соотношениями и выражениями (1.9) и (1.10). Эта аналогия прослеживается в форме записи ХФ и ПФ, что указывает на внутреннее единство объединенной теории случайных процессов и потоков.

Наряду с безусловными функциями важную роль играют условные функции плотности, в частности, для случайных потоков со статистически зависимыми моментами появления точек. Обозначим координаты появления точек для первого подмножества T_n множества T через t_1, \dots, t_n для второго T_k через t'_1, \dots, t'_k , $T_k \in T / T_n$. Условные функции определяются через безусловные из следующих соотношений

$$f(t'_k, \dots, t'_1 | t_n, \dots, t_1) = \frac{f(t'_k, \dots, t'_1, t_n, \dots, t_1)}{f(t_n, \dots, t_1)}. \quad (2.74)$$

Очевидно, что эти условные функции характеризуют вероятность появления точек вблизи моментов времени t'_1, \dots, t'_k при условии появления точек в предшествующие моменты времени t_1, \dots, t_n .

В практических задачах, как и в теории случайных процессов, важную роль играют стационарные потоки. Стационарными называются потоки, функции плотности и корреляции плотности которых не меняются при произвольном в пределах области задания потока сдвиге аргумента. В этом случае ИП оказывается постоянной величиной, а остальные функции плотности зависят только от разностей аргументов

$$\begin{aligned} f_1(t) &= f_1 = \text{const}; \\ f_2(t_1, t_2) &= f_2(t_2 - t_1); \\ &\dots\dots\dots \\ f_n(t_1, \dots, t_n) &= f_n(t_2 - t_1, \dots, t_n - t_1); \end{aligned} \quad (2.75)$$

$$\begin{aligned} g_1(t) &= g_1 = \text{const}; \\ g_2(t_1, t_2) &= g_2(t_2 - t_1); \\ &\dots\dots\dots \\ g_n(t_1, \dots, t_n) &= g_n(t_2 - t_1, \dots, t_n - t_1). \end{aligned} \quad (2.76)$$

На интервале задания стационарного потока $(0, T)$ выделим подынтервал $(a, b) \in (0, T)$. Сдвинем в пределах этого интервала подынтервал (a, b) на произвольную величину Δ . Тогда для ранее

рассмотренных интегральных характеристик имеют место следующие соотношения

$$P_n(a,b)=P_n(a+\Delta,b+\Delta)=P_n(a-b),$$

$$F_n(a,b)=F_n(a-b),$$

$$X_n(a,b)=X_n(a-b).$$

В частности, для характеристики $P_n(a,b)$ это означает, что в случае стационарного точечного процесса вероятность появления точек в течение заданного отрезка времени $a-b$ зависит от величины этого отрезка и не зависит от его расположения на оси времени.

По аналогии со случайными процессами условия (2.75) и (2.76) трактуются как стационарность в строгом смысле. Под стационарностью в широком смысле понимается выполнение только первых двух равенств (2.75) и (2.76). Таким образом, из стационарности в строгом смысле всегда следует стационарность в широком смысле, но не наоборот.

Следующим свойством случайных потоков является ординарность. Поток называется ординарным, если вероятность появления более одной точки на любом малом интервале времени $(t, t+\Delta t)$ есть величина более высокого порядка малости, чем Δt . Для такого типа потоков выполняются соотношения

$$P\{(N_0^{t+\Delta t} - N_0^t) \geq 2\} = 0(\Delta t), \quad (2.77)$$

$$P\{(N_0^{t+\Delta t} - N_0^t) = 1\} = f_1(t)\Delta t + 0(\Delta t), \quad (2.78)$$

$$P\{(N_0^{t+\Delta t} - N_0^t) = 0\} = 1 - f_1(t)\Delta t + 0(\Delta t).^{*)} \quad (2.79)$$

Потоки, для которых это условие не выполняется, называются потоками кратных точек. Потоки, у которых вероятность появления n точек выражается через произведения вероятностей появления n_i точек, появившихся в непересекающихся подынтервалах (a_i, b_i) ,

$$T = \sum_{i=1}^M (a_i b_i), i = \overline{1, M} \quad \text{называются потоками с независимыми}$$

приращениями или потоками без последствия

*) Ввиду зависимости функции плотности первого порядка от времени ее локальные свойства отражают мгновенное поведение интенсивности, которое следует из формулы (2.78)

$$f_1(t) = \lim_{\Delta t \rightarrow 0} P\{(N_0^{t+\Delta t} - N_0^t) = 1\} / \Delta t$$

$$P_{n_1+\dots+n_i+\dots+n_M}(T) = \prod_{n=1}^M P_{n_i}(a_i, b_i). \quad (2.80)$$

Соотношение (2.80) выражает тот факт, что вероятность появления n_i точек в подынтервале (a_i, b_i) не зависит от того, сколько точек появилось вне этого подынтервала.

2.4.3. ПУАССОНОВСКИЙ ПРОЦЕСС

Пуассоновский точечный процесс играет фундаментальную роль в теории случайных потоков и является основополагающим для формирования других более сложных точечных процессов. Он часто используется в качестве модели физических процессов, в различных областях естественных наук и техники. Важная роль его заключается также и в том, что суперпозиция большого числа независимых разных потоков малой интенсивности образует поток, который близок к пуассоновскому. Пуассоновским процессом называется случайный поток, у которого функции корреляции плотности второго и более высокого порядка для любых t равны нулю

$$g_n(t_1, \dots, t_n) = 0, \quad n \geq 2.$$

Обозначим единственную характеристику отличную от нуля - интенсивность потока (ИП) $g_1(t) = f_1(t) = \nu(t)$. На основании (2.67) и (2.70) выражения для производящих функций и функционала принимают вид

$$\mathcal{G}(z, T) = \exp \left\{ (z-1) \int_0^T \nu(\tau) d\tau \right\}, \quad X_1(T) = \int_0^T \nu(\tau) d\tau, \quad (2.81)$$

$$L[u, T] = \exp \left\{ \int_0^T \nu(\tau) u(\tau) d\tau \right\}. \quad (2.82)$$

Используя (2.59) и (2.81) после n -кратного дифференцирования $\mathcal{G}(z, T)$ по z получим известную формулу Пуассона

$$P_n(T) = \frac{\left[\int_0^T \nu(\tau) d\tau \right]^n}{n!} \exp \left\{ - \int_0^T \nu(\tau) d\tau \right\}.$$

Функцию плотности $f_n(t_1, \dots, t_n)$ можно получить на основании (2.71) и (2.82) путем n -кратного функционального дифференцирования ПФ по $u(\tau)$. Она равна произведению ИП с разными аргументами

$$f_n(t_1, \dots, t_n) = \prod_{i=1}^n f_1(t_i) = \prod_{i=1}^n \nu(t_i).$$

Производящий функционал, определенный на сумме всех непересекающихся подынтервалов $\tau_i = a_i - b_i$, $T = \sum_{i=1}^M \tau_i$ равен произведению отдельных ПФ для каждого из подынтервалов:

$$L[u, T] = \exp \left\{ \sum_{i=1}^M \int_{\tau_i} \nu(\tau) u(\tau) d\tau \right\} = \prod_{i=1}^M L[u, \tau_i].$$

Таким образом, для пуассоновского потока статистические характеристики можно определить в каждом подынтервале независимо, т.е. он является потоком с независимыми приращениями. Если пуассоновский ординарный поток с независимыми приращениями является также и стационарным, то он становится простейшим.

Факториальные моменты можно получить из (2.60), используя выражение пф (2.81): $F_1(T) = \int_0^T \nu(\tau) d\tau$, $F_2(T) = \left[\int_0^T \nu(\tau) d\tau \right]^2$.

Математическое ожидание и дисперсия числа точек на интервале $(0, T)$ на основании (2.65) и (2.66) равны

$$m_N(T) = D_N(T) = \int_0^T \nu(\tau) d\tau. \quad (2.83)$$

Можно показать, что кумулянты третьего и более высоких порядков также равны (2.83). Это простая связь между статистическими характеристиками является отличительной чертой только пуассоновского процесса.

В заключение приведем известный из работы [12] полезный результат. Если сумма k независимых случайных величин распределена по закону Пуассона, то каждая из величин удовлетворяет распределению Пуассона.

С более подробным изложением свойств случайных потоков и, в частности пуассоновских, можно ознакомиться в работах [6, 11].

2.4.4. МОДУЛИРОВАННЫЕ ПОТОКИ

2.4.4.1. СЛУЧАЙНАЯ ИНТЕНСИВНОСТЬ

Сформируем одну из распространенных в технических приложениях модель импульсного процесса – случайный импульсный поток. Представим интервал наблюдения (t_0, t) в виде непрерывной последовательности неперекрывающихся одинаковых подынтервалов $\Delta_k = t_k - t_{k-1}$, $t - t_0 = \sum_{k=1}^n \Delta_k$. Обозначим на каждом

подынтервале скорость появления точек через $\Delta N_k / \Delta_k$, где ΔN_k – число появившихся на k -м подынтервале точек. Используя эти данные, образуем функцию

$$\eta^N(\Delta_k) = \lim_{\Delta_k \rightarrow 0} \frac{\Delta N_k}{\Delta_k} 1(\Delta_k),$$

где $1(\Delta_k)$ – единичная на подынтервале Δ_k функция, введена для описания протекающего во времени кусочно-постоянного процесса. При $\Delta_k \rightarrow 0$ ввиду ординарности потока в подынтервал может попасть не более одной точки. Обозначив подынтервалы, в которых появляются точки, индексом i , а также принимая во внимание, что $\lim_{\Delta_i \rightarrow 0} 1(\Delta_k) / \Delta_k = \delta(t - t_i)$, где $\delta(\cdot)$ – дельта-функция, получим при увеличении интервала наблюдения функцию

$$\eta^N(t) = \sum_i \delta(t - t_i), \quad (2.84)$$

описывающую реализацию скорости счета случайного потока событий. Другое название этой функции, вытекающее непосредственно из ее вида – случайный импульсный поток, или, как ее еще называют, случайная интенсивность. Ее математическим ожиданием является средняя скорость счета или интенсивность случайного импульсного потока. Оценка интенсивности на интервале (t_0, t) определяется выражением

$$\hat{f}_{1t} = \frac{1}{t - t_0} \int_{t_0}^t \eta^N(t) dt.$$

Для стационарного импульсного потока интенсивность вычисляется из формулы: $f_1 = \lim_{t \rightarrow \infty} \hat{f}_{1t}$.

Отметим, что реализация случайной интенсивности может быть получена путем дифференцирования функции случайного потока. Действительно, изображенную на рис.2.2 реализацию случайного процесса аналитически можно представить в виде зависимости $N_{t_0}^\tau = \sum_i 1(\tau - \tau_i)$, где единичная функция

$$1(\tau - \tau_i) = \begin{cases} 0, & \tau < \tau_i, \\ 1, & \tau \geq \tau_i, \end{cases}$$

τ_i - координата случайного события единичного роста. Дифференцируя функцию $N_{t_0}^\tau$ получаем (2.84).

Статистические характеристики случайной интенсивности - корреляционные функции разного порядка могут быть получены из сравнения характеристического (ХФ) и производящего (ПФ) функционалов, рассмотренных в разделах 1.2 и 2.4.

Подставим (2.84) в выражение (1.6). Используя фильтрующие свойства дельта-функции, получаем

$$\Theta[v, T] = M \left\{ \exp j \int_0^T \eta^N(t) v(t) dt \right\} = M \left\{ \exp j \sum_i^n v(t_i) \right\} = M \left\{ \prod_i^n \exp j v(t_i) \right\},$$

где операция усреднения производится по числу n и моментам t_i появления дельта-импульсов.

Произведя замену $\exp jv(t_i) = 1 + u(t_i)$, имеем

$$\Theta[v, T] = M \left\{ \prod_i^n [1 + u(t_i)] \right\}. \quad (2.85)$$

Но выражение справа от знака равенства формулы (2.85) является ПФ (2.68).

На основании изложенного можно получить соотношение, связывающее ХФ и ПФ

$$\Theta[v, T] = L\{\exp[jv(t)] - 1\}.$$

Используя эту формулу, а также выражения ХФ (1.8) и ПФ (2.70), после логарифмирования приходим к следующему соотношению

$$\begin{aligned} & j \int_0^T k_1(t) v(t) dt + j^2 \int_0^T \int_0^T k_2(t_1, t_2) v(t_1) v(t_2) dt_1 dt_2 + \dots = \\ & = \int_0^T g_1(t) \{\exp[jv(t)] - 1\} dt + \frac{1}{2} \int_0^T \int_0^T g_2(t_1, t_2) \{\exp[jv(t_1)] - 1\} \times \\ & \times \{\exp[jv(t_2)] - 1\} dt_1 dt_2 + \dots \end{aligned}$$

Разложим экспоненциальные члены в ряд по степеням $v(t)$ и, приравнявая члены с одинаковыми степенями, получаем

$$\begin{aligned} k_1(t) &= g_1(t), \\ k_2(t_1, t_2) &= g_1(t_1) \delta(t_1 - t_2) + g_2(t_1, t_2), \\ k_3(t_1, t_2, t_3) &= g_1(t_1) \delta(t_1 - t_2) \delta(t_1 - t_3) + \\ &+ g_2(t_1, t_3) \delta(t_1 - t_2) + g_2(t_2, t_3) \delta(t_2 - t_1) + \\ &+ g_2(t_1, t_2) \delta(t_1 - t_3) + g_3(t_1, t_2, t_3), \\ &\dots \end{aligned} \quad (2.86)$$

При выводе этих соотношений члены вида $\int_0^T g_1(t)v^2(t)dt$, $\int_0^T g_1(t)v^3(t)dt$, $\int_0^T \int_0^T g_2(t_1, t_2)v^2(t_1)v(t_2)dt_1dt_2$ и т.д. на основании фильтрующих свойств дельта-функции были заменены на тождественно равные им соотношения

$$\begin{aligned}\int_0^T g_1(t)v^2(t)dt &= \int_0^T \int_0^T g_1(t_1)\delta(t_1 - t_2)v(t_1)v(t_2)dt_1dt_2, \\ \int_0^T g_1(t)v^3(t)dt &= \int_0^T \int_0^T \int_0^T g_1(t_1)\delta(t_1 - t_2)\delta(t_1 - t_3)v(t_1)v(t_2)v(t_3)dt_1dt_2dt_3, \\ \int_0^T \int_0^T g_2(t_1, t_2)v^2(t_1)v(t_2)dt_1dt_2 &= \\ &= \int_0^T \int_0^T \int_0^T g_2(t_1, t_2)\delta(t_1 - t_3)v(t_1)v(t_2)v(t_3)dt_1dt_2dt_3.\end{aligned}$$

Дальнейшее изложение будет проводиться в рамках корреляционной теории, т.е. рассматриваются статистические характеристики не выше второго порядка (математическое ожидание, корреляционная функция, дисперсия, спектральная плотность). Как было отмечено ранее, это связано, во-первых, с тем, что учет корреляций более высокого порядка приводит к неоправданным для инженерных расчетов трудностям в вычислениях, во-вторых, для большинства практически важных задач роль высших корреляций с увеличением порядка быстро уменьшается. Учитывая формулы (2.86), для корреляционных функций первых двух порядков можно записать

$$k_1(t)=f_1(t), \quad k_2(t_1, t_2)=f_1(t_1)\delta(t_1 - t_2)+g_2(t_1, t_2).$$

Функцию корреляции плотности $g_2(t_1, t_2)$ необходимо учитывать, если существует статистические зависимости между моментами появления импульсов, например, для модулированных потоков, потоков восстановления и т.д. Приведем также используемые далее статистические характеристики стационарной случайной

интенсивности. Согласно сформулированным в разделе 1.1 требованиям стационарности в широком смысле математическое ожидание имеет постоянное значение, а корреляционная функция зависит от разности аргументов

$$k_1 = g_1 = f_1, \quad (2.87)$$

$$k_2(\tau) = f_1 \delta(\tau) + g_2(\tau). \quad (2.88)$$

Применяя к корреляционной функции (2.88) Фурье-преобразование (формулу Винера-Хинчина (1.17)), получаем спектральную плотность центрированной составляющей случайной интенсивности

$$F(\omega) = \int_{-\infty}^{\infty} k_2(\tau) \exp\{-j\omega\tau\} d\tau = f_1 + \int_{-\infty}^{\infty} g_2(\tau) \exp\{-j\omega\tau\} d\tau. \quad (2.89)$$

Отметим, что для пуассоновского импульсного процесса из-за статистической независимости моментов появления импульсов $g_2(t_1, t_2) = 0$, $f_1(t) = \nu(t)$ и статистические характеристики случайной интенсивности принимают вид

$$k_1(t) = \nu(t),$$

$$k_2(t_1, t_2) = \nu(t_1) \delta(t_2 - t_1),$$

$$F_t(\omega) = \nu(t).$$

Для стационарного процесса имеем

$$k_1(t) = \nu, \quad (2.89)$$

$$k_2(\tau) = \nu \delta(\tau), \quad (2.90)$$

$$F(\omega) = \nu. \quad (2.91)$$

2.4.4.2. УРАВНЕНИЕ ВОССТАНОВЛЕНИЯ

Один из методов определения корреляционной функции случайной интенсивности (2.88) основан на использовании уравнения восстановления. Входящую в корреляционную функцию (2.88) функцию корреляции плотности для стационарной в широком смысле

случайной интенсивности на основании (2.73) можно представить в виде

$$g_2(\tau) = f_2(t_1, t_2) - f_1(t_1)f_1(t_2) = f_1[f_1(t_2|t_1) - f_1] = f_1[f(\tau) - f_1], \quad (2.92)$$

так как условная функция плотности для стационарного процесса равна $f(t_2|t_1) = f(t_2 - t_1) = f(\tau)$, $\tau = t_2 - t_1$.

Условная функция плотности $f(\tau)$ характеризует вероятность появления дельта-импульса в окрестности момента времени t_2 при условии существования дельта-импульса в момент t_1 , $t_2 > t_1$.

Ее можно определить из интегрального уравнения восстановления, которое для стационарного процесса имеет вид [11,13]

$$f(\tau) = \psi(\tau) + \int_0^{\tau} \psi(\tau - t) f(t) dt. \quad (2.92)$$

Здесь $\psi(\tau)$ – плотность распределения вероятностей временных интервалов между соседними импульсами (точками). Таким образом, задаваясь этой функцией, можно из уравнения (2.92) определить условную функцию плотности $f(\tau)$, а на основании ее – функцию $g_2(\tau)$ и соответственно корреляционную функцию случайной интенсивности $k_2(\tau)^*$.

Если для функции $\psi(\tau)$ существует преобразование Лапласа - $\Psi(s)$, то, применяя к обеим частям уравнения (2.92) это преобразование, после упрощений получаем

$$F(s) = \frac{\Psi(s)}{1 - \Psi(s)}. \quad (2.93)$$

Осуществляя обратное преобразование, определяют по $F(s)$ условную плотность $f(\tau)$. Можно предположить более общий путь определения этой функции. Учитывая $|\Psi(s)| < 1$, соотношение (2.93) представим как сумму бесконечно убывающей геометрической прогрессии с первым членом и знаменателем равными $\Psi(s)$:

*) В дальнейшем рассматриваются случайные потоки восстановления, у которых интервалы между соседними точками – независимые случайные величины, имеющие одинаковую плотность распределения $\psi(\tau)$.

$$F(s) = \sum_{k=1}^{\infty} \Psi^k(s).$$

Осуществляя обратное преобразование Лапласа, получаем $f(\tau) = \sum_{k=1}^{\infty} \psi_k(\tau)$, где $\psi_k(\tau)$ определяется через интеграл свертки

$$\psi_k(\tau) = \int_0^t \psi_{k-1}(\tau') \psi(\tau - \tau') d\tau', \quad k \geq 2, \quad \psi_1(\tau) = \psi(\tau).$$

Используя изложенную методику, определим функцию $f(\tau)$ для пуассоновского импульсного потока. Этот процесс ввиду того, что интервалы между соседними импульсами независимы и имеют одинаковое распределение, может быть отнесен к частному случаю потоков восстановления. Имеем $\psi(\tau) = \nu \exp\{-\tau\nu\}$, $\Psi(s) = \frac{\nu}{s + \nu}$,

$F(s) = \frac{\nu}{s}$. Осуществляя обратное преобразование, получаем $f(\tau) = f(t_2|t_1) = f_1(t_2) = \nu$, т.е. условная функция плотности равняется безусловной. В результате функция $g_2(\tau)$ на основании (2.92) оказывается равной нулю и корреляционная функция случайной интенсивности согласно (2.88) равна (2.90).

2.4.5. МАРКОВСКИЙ ПРОЦЕСС, ПОРОЖДЕННЫЙ ПУАССОНОВСКИМ ПРОЦЕССОМ

Наряду с диффузионными марковскими процессами в технических приложениях находят применение марковские процессы, порожденные пуассоновским точечным процессом. Эти процессы могут служить моделью дробовых шумов в электронных и полупроводниковых устройствах, импульсных процессов на выходе детекторов фотонных излучений и т.д. Процессы указанного типа моделируются нелинейными дифференциальными уравнениями, в правых частях которых в качестве воздействующего сигнала присутствует пуассоновский точечный процесс

$$dx(t) = f_t(x)dt + \gamma_t(x)dN_t, \quad x(t_0) = x(0) \quad (2.94)$$

Покажем, что определяемый уравнением (2.94) процесс является марковским. Представим это уравнение в интегральной форме

$$x(t) = x(t_0) + \int_{t_0}^t f_{\tau}(x) d\tau + \int_{t_0}^t \gamma_{\tau}(x) dN(\tau).$$

Из-за независимости приращений пуассоновского процесса решение $x(t)$, равное $x(t_0)$ в момент времени t_0 , зависит от $x(t_0)$ и от значений правой части уравнений при $t > t_0$ и не зависит от $x(t)$ до момента времени t_0 . Таким образом, правая часть этого уравнения при $t > t_0$ статистически независима от ее значений при $t < t_0$. Отсюда следует, что любое условное распределение решения при $t > t_0$ относительно любой совокупности случайных величин $x(t_1), \dots, x(t_k), x(t_0)$ при $t_1 < \dots < t_k < t_0$ не зависит от $x(t_1), \dots, x(t_k)$ и определяется условной плотностью относительно только значения $x(t_0)$. Поэтому процесс, являющийся решением уравнения (2.94), по определению оказывается марковским.

Приведем также уравнение с использованием производной точечного процесса

$$\frac{dx(t)}{dt} = f_t(x) + \gamma_t(x) \eta^N(t), \quad x(t_0) = x(0),$$

где $\eta^N(t)$ - пуассоновский импульсный поток (2.84) или случайная интенсивность, статистические характеристики которых обсуждались в разд. 2.4.4.

В случае независимости коэффициента γ_t от $x(t)$ приходим к используемому в дальнейшем уравнению

$$\frac{dx(t)}{dt} = f_t(x) + \gamma_t \eta^N(t), \quad x(t_0) = x(0).$$

Для многомерного марковского процесса нелинейная система в пространстве состояний описывается дифференциальным уравнением в векторно-матричной форме

$$\frac{dx(t)}{dt} = f_t(x) + \Gamma_t \eta^N(t), \quad x(t_0) = x(0), \quad (2.95)$$

где $f_t(x)$ - r -мерный вектор-столбец нелинейных функций;

$\eta^N(t)$ – r - мерный вектор-столбец случайных интенсивностей;

Γ_t – диагональная матрица воздействий размера $r \times r$.

Для линейной системы дифференциальное уравнение многомерного марковского процесса имеет вид

$$\frac{dx(t)}{dt} = F_t x(t) + \Gamma_t \eta^N(t), \quad x(t_0) = x(0) \quad . \quad (2.96)$$

Используя аналогичный подход, что и в разд. 2.3.8, можно показать, что для дискретного времени многомерный марковский процесс характеризуется уравнением и в векторно-матричной форме для

линейной системы $x_{n+1} = \Phi_n x_n + \Gamma_n \eta_n^N$, $x_0 = x(0)$,

нелинейной системы $x_{n+1} = x_n + f_n(x_n) \Delta t + \Gamma_n \eta_n^N$,

где η_n^N – вектор дискретной случайной интенсивности, математическое ожидание которого представляет собой вектор $m_n = \nu_n$, а корреляционная матрица имеет вид $M\{\eta_n^N \eta_l^N\} = Q_{nl} \delta_{nl}$, δ_{nl} – символ Кронекера; $Q_{nl} = \nu_{l=n\Delta t} / \Delta t$, ν_t – диагональная матрица интенсивностей импульсного пуассоновского потока. Остальные обозначения уравнений аналогичны обозначениям для уравнения (2.55).

2.5. ФРАКТАЛЬНЫЕ ПРОЦЕССЫ

Анализ процессов передачи и преобразования информации в вычислительных сетях указывает на своеобразный характер их поведения, не поддающийся наглядной интерпретации в рамках известных детерминированных и случайных моделей. Переход к технологии пакетной коммутации и создание нового класса интегрированных сетевых приложений (виртуальная реальность, распределенные вычисления, видеоконференции, интернет-телефония и т.д.) сопровождаются появлением сложных явлений, называемых фрактальными процессами, исследование которых может быть проведено в рамках теоретико-вероятностных подходов. Процессы в сетях имеют случайный характер и обладают важными свойствами – масштабируемостью и тесно связанным с ней самоподобием.

Следует отметить, что с обладающими указанными свойствами явлениями исследователи сталкиваются уже давно при наблюдении объектов природного происхождения. Для многих этих объектов характерны определенные уровни регулярности и фрагментации. Эти свойства проявляются, например, в том, что профиль горы имеет сходство с контурами образующих ее холмов, контуры берегов рек и морей – с отдельными составляющую береговую линию фрагментами, профиль дерева имеет сходство со структурой ветвей и т.д.

Для анализа геометрических свойств рассматриваемых структур были введены математические объекты – фракталы, которые обнаруживают некоторую форму самоподобия: части целого могут характеризовать все целое путем масштабирования своей структуры. «Фракталами называются структуры, состоящие из частей, которые в каком смысле подобны целому» [14].

Связь между масштабируемостью и фрактальностью обнаруживается не только для природных объектов, но и в различных физических явлениях, при химических превращениях, а также во многих других наблюдаемых явлениях и объектах, в том числе – в случайных процессах [15,16,17,18,19].

Аналитически свойства масштабируемости и самоподобия, учитывая высказанные выше соображения о соотношениях части и целого, можно отобразить функциональным уравнением

$$U(r_L t) = \frac{1}{N} U(t), \quad (2.97)$$

где t - пространственно-временной аргумент, r_L и N - масштабирующие параметры.

Для описания характеристик фрактальных объектов или процессов вводится степенная с дробным показателем функция вида

$$U(t) = A t^{-\beta}, \quad 0 < \beta < 1. \quad (2.98)$$

После подстановки выражения (2.98) в уравнение (2.97) получаем соотношения, связывающие масштабирующие параметры с дробным показателем $\beta = \ln N / \ln r_L$. Нетрудно заметить, что при $\beta \rightarrow 0$ функция (2.98) приобретает характер медленно убывающего процесса или, как говорят, протяженной зависимости, кардинально

отличающей ее от обычных моделей с короткопротяженными зависимостями.

Описывающая масштабно-инвариантные свойства фракталов, убывающая с дробным показателем степенная функция в последнее время широко используется при анализе объектов природного и искусственного происхождения. В дальнейшем при изучении фрактальных процессов в вычислительных сетях рассматриваются зависящие от временного аргумента t и обладающие вышеуказанными свойствами следующие характеристики: импульсная переходная функция, корреляционная функция, дисперсия и спектральная плотность. Базовыми для формирования этих характеристик служат ранее рассмотренные точечные и винеровские процессы. С более подробным изложением свойств характеристик и параметров моделей фрактальных процессов можно ознакомиться в работе [20].

2.5.1. ФРАКТАЛЬНЫЙ ТОЧЕЧНЫЙ ПРОЦЕСС

При анализе моделей процессов в вычислительных сетях используются статистики числа отсчетов (приращений) точечного процесса на интервалах заданной длительности T (счетные статистики). Обозначим число выпавших на интервале $(t_n - T, t_n)$ точек через X_n . Сместим этот интервал на kT ($k \geq 1$) и обозначим число выпавших на интервале $(t_{n+k} - T, t_{n+k})$ точек через X_{n+k} . Корреляционная функция числа отсчетов в разнесенных на время, равное kT указанных интервалах определяется соотношением

$$C(k; T) = M\{X_n X_{n+k}\} - (\nu T)^2, \quad (2.99)$$

где ν - интенсивность пуассоновского точечного процесса.

Дисперсия числа отсчетов равна при $k=0$

$$D(T) = C(0; T). \quad (2.100)$$

Процедуры определения статистических характеристик (2.99) и (2.100) опираются на интегральные соотношения, связывающие искомые функции и процессы с известными статистическими характеристиками. Предварительно получим выражение статистик для непрерывного времени. Пусть $\xi(t)$ стационарный в широком смысле случайный процесс с известными математическим

ожиданием m_1 и корреляционной функцией $k_2(u)$. Математическое ожидание и корреляционная функция интеграла от этого процесса на заданном интервале $x_T(t) = \int_{t-T}^t \xi(t)dt$, $(t-T, t)$ соответственно равны

$$m_x = M\{x_T(t)\} = \int_{t-T}^t M\{\xi(t)\}dt = m_1 T,$$

$$k_x(\vartheta) = M\{[x_T(t) - m_x][x_T(t + \vartheta) - m_x]\} = \int_{t-T}^t \int_{t-T+\vartheta}^{t+\vartheta} k_2(u_1 - u_2) du_1 du_2.$$

Соотношение, связывающее на основании формулы Хинчина-Винера (1.18) корреляционную функцию и спектральную плотность, имеет вид

$$k_2(u_1 - u_2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp\{j\omega(u_1 - u_2)\} d\omega.$$

После подстановки полученного выражения в $k_x(\vartheta)$ и интегрирования, имеем

$$\begin{aligned} k_x(\vartheta) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) d\omega \int_{t-T}^t \exp\{-j\omega u_2\} du_2 \int_{t-T+\vartheta}^{t-\vartheta} \exp\{j\omega u_1\} du_1 = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \frac{2[1 - \cos \omega T]}{\omega^2} \exp\{j\omega \vartheta\} d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \left[\frac{\sin \omega T / 2}{\omega / 2} \right]^2 \exp\{j\omega \vartheta\} d\omega. \end{aligned}$$

Как следует из полученного выражения, интегралу от процесса $\xi(t)$ с известной корреляционной функцией $k_2(\tau)$ соответствует процесс с корреляционной функцией $k_x(\vartheta)$ и спектральной плотностью

$$F_x(\omega) = F(\omega) \left[\frac{\sin \omega T / 2}{\omega / 2} \right]^2.$$

Подставив в соотношение для $k_x(\vartheta)$ выражение $F(\omega) = \int_{-\infty}^{\infty} k_2(u) \exp\{-j\omega u\} du$, получим

$$k_x(\vartheta) = \int_{-\infty}^{\infty} k_2(u) du \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-j\omega(\vartheta - u)\} \left(\frac{\sin \omega T / 2}{\omega / 2} \right)^2 d\omega \right].$$

После замены $\vartheta - u = \tau$ ($\vartheta > u$) интеграл в квадратных скобках оказывается табличным и равным $T - \tau$ при $\tau < T$ и нулю при $\tau > T$. Присоединяя к полученному результату значение этого интеграла для области $u > \vartheta$, получаем окончательное выражение для корреляционной функции и дисперсии

$$k_x(\vartheta) = \int_{-T}^T (T - |\tau|) k_2(\vartheta - \tau) d\tau,$$

$$D_x = k_x(0) = \int_{-T}^T (T - |\tau|) k_2(\vartheta - \tau) d\tau = 2 \int_0^T (T - \tau) k_2(\tau) d\tau.$$

При определении статистических характеристик числа отсчетов на интервалах заданной длительности T исходное интегральное соотношение для дискретного времени $t_1, t_2, \dots, t_n, \dots$ имеет вид

$$X_n = \int_{t_n - T}^{t_n} \eta^N(t) dt, \text{ где } \eta^N(t) - \text{случайная интенсивность или случайный}$$

импульсный поток (2.84) с известными математическим ожиданием ν , корреляционной функцией $k_2(\tau)$ и спектральной плотностью $F(\omega)$. Интервал между отдельными отсчетами ϑ оказывается кратным длительности T и равным kT , где k - параметр смещения.

На основании изложенного, учитывая обозначения для счетных статистик, имеем

$$m_x = \nu T, \quad (2.101)$$

$$C(k; T) = \int_{-T}^T (T - |\tau|) k_2(kT - \tau) d\tau, \quad (2.102)$$

$$D(T) = 2 \int_0^T (T - \tau) k_2(\tau) d\tau. \quad (2.103)$$

При анализе процессов в вычислительных сетях используются следующие статистические характеристики стационарного точечного процесса

Характеристики первого порядка:

- интенсивность точечного процесса (средняя скорость счета точечного процесса) ν .

Характеристики второго порядка:

- моментная функция второго порядка случайной интенсивности $G_N(\tau)$;

- спектральная плотность, соответствующая этой функции $F_N(\omega)$;

- корреляционная функция числа отсчетов $C(k; T)$;

- нормированная корреляционная функция числа отсчетов (коэффициент корреляции) $r(k; T)$;

- нормированная дисперсия числа отсчетов (фактор Фано) $F(T)$.

Моментная функция второго порядка случайной интенсивности точечного процесса по определению равна

$$G_N(\tau) = \lim_{\Delta t \rightarrow 0} \frac{M\{\Delta N_t \Delta N_{t+\tau}\}}{\Delta t^2},$$

где ΔN_t характеризует появление по крайней мере одной точки в бесконечно малом интервале $(t - \Delta t, t)$, τ - интервал времени между событиями появления точек.

На основании соотношений (1.9) и (2.88) эта функция через рассмотренную в разд. 2.4.4.1 корреляционную функцию случайной интенсивности $k_2(\tau)$ выражается следующим образом

$$G_N(\tau) = m_2(\tau) = k_2(\tau) + \nu^2 = \nu \delta(\tau) + g_2(\tau) + \nu^2 = \nu \delta(\tau) + R_I(\tau), \quad (2.104)$$

где составляющую $R_I(\tau) = g_2(\tau) + \nu^2$ можно интерпретировать как моментную функцию модулирующего точечный процесс сигнала $I(t)$.

Особенностью этого сигнала является то, что он порождает корреляционные функции с протяженной зависимостью, приводящие к большому числу комбинаций фрактальных процессов со свойствами самоподобия. В силу указанной интерпретации такие

процессы также называют двойным стохастическим пуассоновским процессом или точечным процессом с двойной случайностью (одна случайность порождена пуассоновским процессом, другая – сигналом $I(t)$). Отметим, что модуляция точечного процесса другими сигналами, например, марковскими с экспоненциальной корреляционной функцией, имеющей короткопротяженную зависимость, порождает модели процессов, не обладающие фрактальными свойствами и поэтому не адекватные процессам в вычислительных сетях.

Форму записи функции корреляции плотности $g_2(\tau)$ или, что тоже самое, корреляционной функции модулируемой сигналом $I(t)$ составляющей случайной интенсивности, можно получить разными методами. Согласно одному из них эту функцию определяют через интеграл свертки

$$I(t) = \int_0^{\infty} g(t - \tau) \eta^N(\tau) d\tau, \quad (2.105)$$

где $g(t) = Kt^{(\alpha/2)-1}$ - импульсная переходная функция степенного вида; $\eta^N(t)$ - воздействующий стационарный импульсный пуассоновский процесс с интенсивностью ν ; α - фрактальный параметр ($0 < \alpha < 1$)^{*)}.

На основании теоремы Кембелла о суперпозиции независимых случайных воздействий (разд.3.3.3) для области $t > 0$ ($t - \tau > 0$) эта функция с учетом (2.105) определяется из выражения

$$g_2(\tau) = M\{I(t)I(t + \tau)\} - \nu^2 = \nu \int g(t)g(t + \tau)dt = \nu K^2 \int_0^{\infty} (t^2 + \tau^2)^{(\alpha/2)-1} dt.$$

После замены $z = t/\tau$ приходим к табличному интегралу. В результате получаем

$$g_2(\tau) = \nu K^2 \tau^{\alpha-1} \int_0^{\infty} z^{(\alpha/2)-1} dz = \nu K^2 \tau^{\alpha-1} \frac{\Gamma(\alpha/2)\Gamma(1-\alpha)}{\Gamma(1-\alpha/2)}, \quad (2.106)$$

^{*)} Наряду с параметром α для описания фрактальных процессов используют параметр Херста: $H = (1 + \alpha)/2$ ($1/2 < H < 1$).

где $\Gamma(\cdot)$ - гамма-функция.

Принимая во внимание, что корреляционная функция является четной функцией своего аргумента и присоединяя к (2.106) результат интегрирования по области $t < 0$ ($t - \tau < 0$), получаем окончательно

$$g_2(\tau) = \nu^2 \left(\frac{|\tau|}{\tau_0} \right)^{\alpha-1}, \quad (2.107)$$

$$\text{где } \tau_0^{\alpha-1} = \frac{\nu \Gamma(1 - \alpha/2)}{K^2 \Gamma(\alpha/2) \Gamma(1 - \alpha)}.$$

Статистические характеристики второго порядка числа отсчетов определяются следующим образом. Фактор Фано вычисляется на основании формул (2.103), (2.104) и (2.107)

$$\begin{aligned} F(T) &= D(T)(\nu T)^{-1} = 2(\nu T)^{-1} \int_0^T (T - \tau) k_2(\tau) d\tau = \\ &= 2(\nu T)^{-1} \int_0^T (T - \tau) (G_N(\tau) - \nu^2) d\tau = \\ &= 2(\nu T)^{-1} \left[\int_0^T (T - \tau) \nu \delta(\tau) d\tau + \frac{\nu^2}{\tau_0^{\alpha-1}} \int_0^T (T - \tau) \tau^{\alpha-1} d\tau \right]. \end{aligned} \quad (2.108)$$

Первый интеграл в (2.108) на основании фильтрующих свойств дельта-функции равен $\nu T/2$. После вычисления второго интеграла получаем $\frac{T^{\alpha+1}}{\alpha(1+\alpha)}$. Выражение для фактора Фано принимает окончательный вид

$$F(T) = 1 + \left(\frac{T}{T_0} \right)^\alpha, \quad (2.109)$$

где

$$T_0 = \frac{\alpha(1+\alpha)}{2\nu\tau_0^{1-\alpha}}. \quad (2.110)$$

Приведем также выражение для дисперсии числа отсчетов

$$D(T) = \nu TF(T) = \nu T \left[1 + \left(\frac{T}{T_0} \right)^\alpha \right]. \quad (2.111)$$

Выражение для корреляционной функции числа отсчетов вычисляется при $k \geq 1$ с учетом (2.104) и (2.107) из соотношения (2.102)

$$\begin{aligned} C(k; T) &= \int_{-T}^T (T - |\tau|) k_2(kT - \tau) d\tau = \int_{-T}^T (T - |\tau|) [G_N(kT - \tau) - \nu^2] d\tau = \\ &= \nu \left[\int_{-T}^T (T - \tau) \delta(kT - \tau) d\tau + \int_{-T}^T (T - \tau_1) k_2(kT + \tau_1) d\tau_1 \right] + \\ &+ \frac{\nu^2}{\tau_0^{\alpha-1}} \left[\int_{-T}^T (T - \tau) (kT - \tau)^{\alpha-1} d\tau + \int_{-T}^T (T - \tau_1) (kT + \tau_1)^{\alpha-1} d\tau_1 \right], \end{aligned}$$

где $\tau_1 = -\tau$.

Первые два интеграла J_1 и J_2 на основании фильтрующих свойств дельта-функции равны нулю. Третий J_3 и четвертый J_4 интегралы соответственно равны

$$\begin{aligned} J_3 &= \frac{\nu^2}{\tau_0^{\alpha-1}} \int_{-T}^T (T - \tau) (kT - \tau)^{\alpha-1} d\tau = \\ &= \frac{\nu^2 T^{\alpha+1}}{\tau_0^{\alpha-1}} \left\{ \frac{1}{2} [k^\alpha - (k-1)^\alpha] + \frac{1}{\alpha} (k-1)^\alpha + \frac{[(k-1)^{\alpha+1} - k^{\alpha+1}]}{\alpha(1+\alpha)} \right\}, \end{aligned} \quad (2.112)$$

$$\begin{aligned} J_4 &= \frac{\nu^2}{\tau_0^{\alpha-1}} \int_{-T}^T (T - \tau_1) (kT - \tau_1)^{\alpha-1} d\tau = \\ &= \frac{\nu^2 T^{\alpha+1}}{\tau_0^{\alpha-1}} \left\{ \frac{1}{2} [(k+1)^\alpha - k^\alpha] + \frac{1}{\alpha} (k+1)^\alpha + \frac{[(k+1)^{\alpha+1} - k^{\alpha+1}]}{\alpha(1+\alpha)} \right\}. \end{aligned} \quad (2.113)$$

Корреляционная функция числа отсчетов на основании (2.110), (2.112) и (2.113) принимает окончательный вид

$$C(k;T) = J_1 + J_2 + J_3 + J_4 =$$

$$= \frac{1}{2} \nu T \left(\frac{T}{T_0} \right)^\alpha [(k+1)^{\alpha+1} - 2k^{\alpha+1} + (k-1)^{\alpha+1}]. \quad (2.114)$$

Отметим, фактор Фано является удобной для подтверждения фрактальных свойств сетевого трафика при обработке экспериментальных данных характеристикой. Действительно, при $T \gg T_0$ зависимость $F(T)$ в двойном логарифмическом масштабе представляет собой приблизительно прямую с положительным наклоном, равным фрактальному параметру α (для пуассоновского процесса наклон равен нулю). Таким образом, определяя выборочные значения $D(T)$ и νT как функции текущего интервала T , можно оценить фрактальный параметр α .

Спектральная плотность случайной интенсивности, соответствующая моментной функции $G_N(\tau)$ (2.104) с учетом (2.107), определяется через Фурье-преобразование (формула Хинчина-Винера) и после вычислений равна

$$F_N(\omega) = \int_{-\infty}^{\infty} G_N(\tau) \exp\{-j\omega\tau\} d\tau = 2\pi\nu^2 \delta(\omega) + \nu \left(\frac{|\omega|}{\omega_0} \right)^{-\alpha} + \nu, \quad (2.115)$$

где.

$$\omega_0^\alpha = 2\nu \cos \frac{\pi\alpha}{2} \Gamma(\alpha) \tau_0^{1-\alpha} \quad (2.116)$$

$\delta(\cdot)$ - дельта-функция, $\Gamma(\cdot)$ - гамма-функция.

Рассматриваемую спектральную плотность можно представить также в виде

$$F_N(\omega) = F_I(\omega) + \nu. \quad (2.117)$$

Здесь $F_I(\omega) = \int_{-\infty}^{\infty} R_I(\tau) \exp\{-j\omega\tau\} d\tau = 2\pi\nu^2 \delta(\omega) + \nu \left(\frac{|\omega|}{\omega_0} \right)^{-\alpha}$ -

спектральная плотность модулирующего сигнала $I(t)$.

Приведем еще одно выражение спектральной плотности

$$F_N(\omega) = 2\pi \nu^2 \delta(\omega) + F(\omega), \quad (2.118)$$

где $F(\omega) = \nu \left(\frac{|\omega|}{\omega_0} \right)^{-\alpha} + \nu$ - спектральная плотность
центрированной составляющей случайной интенсивности.

Обратим внимание на своеобразный характер поведения спектральной плотности $F(\omega)$ для фрактальных процессов: ее неограниченное увеличение при $\omega \rightarrow 0$: $F(\omega) \sim \omega^{-\alpha}$.

Используя формулы (2.110) и (2.116), получаем соотношение, связывающее параметры T_0 и ω_0 :

$$\omega_0^\alpha T_0^\alpha = \cos(\pi\nu/2) \Gamma(\alpha + 2). \quad (2.119)$$

Введем нормированную корреляционную функцию (коэффициент корреляции) приращений точечного процесса (числа отсчетов)

$$r(k; T) = \frac{C(k; T)}{D(T)} = \frac{T^\alpha}{2(T^\alpha + T_o^\alpha)} [(k+1)^{\alpha+1} - 2k^{\alpha+1} + (k-1)^{\alpha+1}], \quad (2.120)$$

$$|r(k; T)| \leq 1.$$

Представим выражения $(k+1)^{\alpha+1}$ и $(k-1)^{\alpha+1}$ при $k \gg 1$ в виде трех членов разложения

$$(k+1)^{\alpha+1} \approx k^{\alpha+1} + (\alpha+1)k^\alpha + \frac{1}{2}\alpha(1+\alpha)k^{\alpha-1},$$

$$(k-1)^{\alpha+1} \approx k^{\alpha+1} - (\alpha+1)k^\alpha + \frac{1}{2}\alpha(1+\alpha)k^{\alpha-1}.$$

После подстановки полученных выражений в соотношение (2.120) и упрощений приходим к коэффициенту корреляции, поведение которого (точнее поведение его «хвоста») аппроксимируется выражением

$$r(k; T) \sim \frac{\alpha(\alpha+1)}{2 \left[1 + \left(\frac{T_0}{T} \right)^\alpha \right]} k^{\alpha-1}. \quad (2.121)$$

Полученное соотношение ясно указывает на атрибуты фрактального процесса – протяженную зависимость и самоподобие (масштабируемость). Причем эти свойства проявляются тем интенсивнее, чем больше интервал отсчета T по отношению к фрактальному времени установки T_0 . В этом случае принято говорить об асимптотическом самоподобии в том смысле, что коэффициент корреляции сохраняет свою структуру при $T \gg T_0$ и зависит от убывающего с дробным показателем степени масштабируемого параметра k

$$r(k; T) \sim \frac{1}{2} \alpha (\alpha + 1) k^{\alpha-1}. \quad (2.122)$$

Нетрудно заметить, что если в уравнении (2.97) для этого случая время заменить на параметр смещения k , то соответствующие параметры функционального уравнения равны

$$r_L = N^{-1/\beta}, \quad \beta = 1 - \alpha = \ln N / \ln r_L.$$

Ввиду того, что спектральная плотность случайной интенсивности является степенной функцией $|\omega|$ с показателем, равным фрактальному параметру α , взятым с обратным знаком, спектральная плотность числа отсчетов (приращения точечного процесса) на основании (2.118) также оказывается степенной функцией $|\omega|$ и при $\omega \rightarrow 0$ неограниченно увеличивается как

$$F_T(\omega) \sim |\omega|^{-\alpha}. \quad (2.123)$$

Свойство самоподобия обнаруживается у агрегированного счетного процесса. Указанный процесс формируется как последовательность средневзвешенных величин из отсчетов X_n на m одинаковых неперекрывающихся интервалах длительностью T

$$X^{(m)} = \{X_k, k = 0, 1, \dots\} = \left\{ \frac{X_1 + \dots + X_m}{m}, \dots, \frac{X_{km+1} + \dots + X_{(k+1)m}}{m} \right\}, \quad (2.124)$$

где m и k соответственно параметры агрегирования и смещения,

$$X_k^{(m)} = \frac{1}{m} \sum_{n=km+1}^{(k+1)m} X_n, \quad m=1, 2, \dots$$

Для агрегированного процесса статистики второго порядка имеют вид

$$C^{(m)}(k;T) = m^{-2} \int_{-mT}^{mT} (mT - |\tau|)(G(kTm - \tau) - \nu^2) d\tau = , \quad (2.125)$$

$$= m^{-2} C(k, mT)$$

$$r^{(m)}(k;T) = \frac{1}{2 \left[1 + \left(\frac{T_0}{mT} \right)^\alpha \right]} [(k+1)^{\alpha+1} - 2k^{\alpha+1} + (k-1)^{\alpha+1}]. \quad (2.126)$$

При $m \rightarrow \infty$ коэффициент корреляции практически инвариантен к параметру агрегирования и сохраняет в асимптотическом смысле свою структуру. Последнее означает, что исходный и агрегированный процессы имеют одинаковую форму коэффициента корреляции

$$r^{(m)}(k;T) \sim \frac{1}{2} [(k+1)^{\alpha+1} - 2k^{\alpha+1} + (k-1)^{\alpha+1}].$$

Дисперсия агрегированного счетного процесса при больших m изменяется по медленно затухающему закону $m^{\alpha-1}$. Действительно, при $k=0$ с учетом (2.111) и (2.125) имеем

$$D^{(m)}(T) = \frac{\nu m T}{m^2} \left[1 + \left(\frac{mT}{T_0} \right)^\alpha \right] = \nu T \left[m^{-1} + \left(\frac{T}{T_0} \right)^\alpha m^{\alpha-1} \right] \sim m^{\alpha-1}. \quad (2.127)$$

Для сравнения обратим внимание на то, что для обычных короткопротяженных зависимостей дисперсия агрегированного процесса

$$D^{(m)}(T) \sim m^{-1}. \quad (2.128)$$

2.5.2. ФРАКТАЛЬНЫЙ ВИНЕРОВСКИЙ ПРОЦЕСС

Наряду с точечными процессами при анализе моделей сигналов в вычислительных сетях используют непрерывный с вероятностью единица фрактальный винеровский процесс. Для описания непрерывного обладающего фрактальными свойствами процесса в

работе [21] было введено обобщенное броуновское движение или, как теперь его будем называть, фрактальный винеровский процесс $B_H(t)$, который записывается в форме дробного интеграла

$$B_H(t) = \frac{1}{\Gamma(H + \frac{1}{2})} \int_{-\infty}^t g(t - \tau) dB(\tau), \quad (2.129)$$

где $B(\tau)$ - винеровский процесс, $\Gamma(\cdot)$ - гамма функция, H - параметр Херста, $H=(\alpha+1)/2$.

Импульсная переходная функция равна

$$g(t - \tau) = \begin{cases} (t - \tau)^{H-1/2}, & 0 \leq \tau < t, \\ (t - \tau)^{H-1/2} - (-\tau)^{H-1/2}, & \tau < 0. \end{cases} \quad (2.130)$$

Использование в формуле (2.129) импульсной переходной функции степенного вида приводит к протяженной коррелированной зависимости процесса $B_H(t)$, а также указывает на самоподобный характер фрактального винеровского процесса. На основании очевидного соотношения $g(bt - b\tau) = b^{H-1/2} g(t - \tau)$, где b - масштабный параметр, а также свойства самоподобия винеровского процесса $dB(b\tau) = b^{1/2} dB(\tau)$ из формулы (2.129) получаем

$$b^{-H} B_H(bt) = B_H(t), \quad (2.131)$$

что подтверждает самоподобный характер фрактального винеровского процесса.

Для приращения этого процесса математическое ожидание и дисперсия с учетом рассмотренных в разделе 2.2.1 свойств винеровского процесса

$$M\{dB(\tau)\} = 0,$$

$$M\{dB(\tau_1)dB(\tau_2)\} = M\{\eta(\tau_1)\eta(\tau_2)\}d\tau_1d\tau_2 = \mathcal{N}_\eta\delta(\tau_2 - \tau_1)d\tau_1d\tau_2$$

на основании (2.129) соответственно равны

$$M\{B_H(t) - B_H(t_0)\} = 0,$$

$$M\{[B_H(t) - B_H(t_0)]^2\} \sim (t - t_0)^{2H}.$$

Заметим, что при $H=1/2$ процесс $B_H(t)$ становится винеровским с дисперсией и математическим ожиданием равным соответственно (2.12) и нулю.

Используя аналогичный подход, можно от характеристик винеровского процесса перейти к характеристикам фрактального винеровского процесса. Например, знание математического ожидания (2.6) и корреляционной функции (2.9) позволяет их записать для фрактального винеровского процесса в форме

$$\begin{aligned} M\{B_H(t)\} &= 0, \\ k_{2H}(t_1, t_2) &\sim \frac{1}{2}[t_1^{2H} + t_2^{2H} - |t_2 - t_1|^{2H}]. \end{aligned} \quad (2.132)$$

Коэффициент корреляции для стационарных приращений фрактального винеровского процесса на интервалах $(t_n - T, t_n)$ и $(t_{n+k} - T, t_{n+k})$ заданной длительности T , разнесенных на время kT , где k - параметр смещения, можно представить, как и для приращений точечного процесса, выражением

$$r_H(k; T) \sim \frac{1}{2}[(k+1)^{\alpha+1} - 2k^{\alpha+1} + (k-1)^{\alpha+1}].$$

При $k=1$, что соответствует корреляционной зависимости приращений процесса на соседних интервалах времени, получаем

$$r_H(1; T) \sim 2^\alpha - 1 = 2^{2H-1} - 1.$$

При больших k коэффициент корреляции аппроксимируется выражением (2.122) или

$$r_H(k; T) \sim H(2H-1)k^{2H-2}.$$

Из этого выражения следует, что чем больше параметр Херста, тем более протяженной зависимостью обладает $r_H(k; T)$

Как и для счетных статистик можно показать, что поведение спектральной плотности приращений фрактального броуновского движения при $\omega \rightarrow 0$ описывается зависимостью (2.123).

Если обозначить приращение фрактального винеровского процесса на интервале через X_n , то у агрегированного процесса

$$X^{(m)} = \{X_k^{(m)}, k = 0, 1, \dots\}, \quad X_k^{(m)} = \frac{1}{m} \sum_{n=kmH}^{(k+1)m} X_n \quad \text{коэффициент корреляции}$$

сохраняет свою структуру, а дисперсия изменяется согласно соотношению (2.127). Напомним, что для процессов с короткопротяженными зависимостями статистик дисперсия приращений агрегированного процесса изменяется в соответствии с выражением (2.128).

3. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ПРОЦЕССОВ

3.1. ДИНАМИЧЕСКИЕ СИСТЕМЫ И ИХ ХАРАКТЕРИСТИКИ

Динамическая система представляет собой устройство, преобразующее (перерабатывающее) входной воздействующий сигнал в выходной для достижения определенных практических целей. С математической точки зрения, преобразование входного процесса $\xi(t)$ в выходной – $x(t)$ можно представить формализованным соотношением, которое символически записывается в виде $x(t)=w[\xi(t)]$. Здесь $w[\cdot]$ является оператором, так как выходной процесс можно рассматривать, как результат выполнения некоторых операций над входным воздействием $\xi(t)$. В общем случае система может быть многомерной (многоканальный), т.е. иметь несколько входов и выходов. Поэтому входные и выходные процессы необходимо рассматривать как состоящие из набора компонент векторов определенной размерности: $\xi(t)=\|\xi_1(t), \dots, \xi_p(t)\|^T$, $x(t)=\|x_1(t), \dots, x_r(t)\|^T$.

Динамические системы и соответствующие им операторы могут быть детерминированными или случайными. В первом случае каждой конкретной реализации входного процесса соответствует определенная реализация выходного процесса. Во втором - одной и той же реализации входного процесса соответствуют различные реализации выходного процесса. В детерминированных системах, которые в дальнейшем рассматриваются, их поведение описывается дифференциальными уравнениями с детерминированными (постоянными или зависящими от времени) коэффициентами, а сама

случайность выходного процесса обусловлена случайным характером входного процесса.

По характеру зависимости выходного процесса от входного системы классифицируются следующим образом: безынерционные (без памяти) и инерционные, стационарные и нестационарные, линейные и нелинейные. В безынерционных системах значения выходного процесса в любой момент времени зависят от значения входного процесса в тот же момент времени. В инерционных системах значения выходного процесса в некоторый момент времени зависят не только от значений входного процесса в этот же момент времени, но и от значений этого процесса в предшествующие моменты времени. Любую сложную систему можно представить в виде комбинаций отдельных линейных и нелинейных устройств (звеньев). Поэтому анализ системы может быть сведен к анализу преобразования сигналов через отдельные устройства. В дальнейшем в качестве таких устройств рассматриваются детерминированные линейные инерционные и нелинейные безынерционные устройства. Реальные устройства, как правило, оказываются одновременно нелинейными и инерционными. Решение задач анализа для этого общего случая сопряжено с большими трудностями. Поэтому в практических задачах прибегают к вышеуказанным более простым моделям устройств, которые в большинстве инженерных задач приводят к удовлетворительным результатам.

Важным классом динамических систем являются стационарные системы. Для этих систем реакция на входное воздействие зависит только от времени действия этого воздействия и не зависит от его временного момента приложения. Другими словами, сдвиг входного воздействия приводит к такому же сдвигу выходного процесса: $x(t-\tau)=w[\xi(t-\tau)]$. Если это условие не выполняется, система оказывается нестационарной.

В современной теории систем при описании динамических процессов используют переменные состояния и уравнения первого порядка относительно этих переменных. Состояние системы в любой текущий момент времени характеризуется совокупностью фазовых координат - переменных состояний, объединенных в вектор состояния. Метод исследования динамических систем, как и ранее рассмотренный в разд.2.3.7 формирующих фильтров, использующих этот способ описания, принято называть методом пространства

состояний. Таким образом, состояние системы может быть охарактеризовано r -мерным вектором выходных сигналов $x(t)=\|x_1(t),\dots,x_r(t)\|^T$. Динамика изменения состояния непрерывной нелинейной системы с детерминированными параметрами описывается системой дифференциальных уравнений первого порядка в векторно-матричной форме

$$\frac{dx(t)}{dt} = f_t(x) + \Gamma_t \xi(t), \quad x(t_0) = x(0), \quad (3.1)$$

где $\xi(t)$ - вектор произвольных входных сигналов. В частном случае, когда входной сигнал является гауссовским белым шумом, уравнение (3.1) становится стохастическим дифференциальным уравнением.

Для линейных систем уравнение принимает вид

$$\frac{dx(t)}{dt} = F_t x(t) + \Gamma_t \xi(t), \quad x(t_0) = x(0), \quad (3.2)$$

где F_t и Γ_t аналогичны векторам и матрицам уравнений (2.48) .

В случае линейных систем справедлив принцип суперпозиции, согласно которому линейная комбинация входных воздействий связана с выходным процессом линейным оператором на основании соотношения

$$x(t) = w \left[\sum_{i=1}^k C_i \xi_i(t) \right] = \sum_{i=1}^k C_i w[\xi_i(t)], \quad (3.3)$$

где C_i являются постоянными или случайными величинами, не зависящими от времени. Оператор, для которого принцип суперпозиции не применим, называется нелинейным.

Последовательно дифференцируя уравнения первого порядка системы (3.2), можно получить дифференциальное уравнение r -ого порядка относительно какой-либо компоненты вектора выходного процесса, например, $x(t)=x_1(t)$

$$\begin{aligned}
& a_r(t) \frac{d^r x_t}{dt^r} + a_{r-1}(t) \frac{d^{r-1} x_t}{dt^{r-1}} + \dots + a_0(t) x_t = \\
& = b_{r-1}(t) \frac{d^{r-1} \xi_t}{dt^{r-1}} + b_{r-2}(t) \frac{d^{r-2} \xi_t}{dt^{r-2}} + \dots + b_0(t) \xi_t, \quad (3.4) \\
& x(t_0) = x(0), \quad x'(t_0) = x_1(0), \dots, x^{(r-1)}(t_0) = x_{r-1}(0).
\end{aligned}$$

Оператор системы, описываемой уравнением (3.4) имеет вид $w(p,t)=B(p,t)/A(p,t)$, где собственный оператор $A(p,t)=\sum_{i=0}^r a_i p^i$, входной оператор $B(p,t)=\sum_{i=0}^{r-1} b_i p^i$, $p=d/dt$ - оператор дифференцирования^{*)} Сокращенная запись уравнения, принимает форму

$$A(p,t)x(t)=B(p,t)\xi(t). \quad (3.5)$$

Из (3.4) формально следует равенство, определяющее выходной процесс в явном виде $x(t)=w(p,t)\xi(t)$. При заданном операторе линейной системы принцип суперпозиции позволяет свести исследование реакции системы (поведение выходного процесса системы) на произвольное воздействие к исследованию реакции системы на типовое воздействие. В теории линейных систем в качестве одного из таких воздействий используется дельта-функция. Используя фильтрующие свойства дельта-функции, произвольное воздействие можно представить в виде бесконечной последовательности дельта-импульсов в следующие друг за другом моменты времени с интенсивностями, равными значениям воздействующего сигнала в эти моменты времени

$$\xi(t) = \int_0^t \xi(u) \delta(t-u) du. \quad (3.6)$$

Реакцией предварительно невозбужденной системы на воздействие дельта-импульса является импульсная переходная

^{*)} Название «собственный оператор» обусловлено тем, что многочлен $A(p,t)$ характеризует собственное движение системы, т.е. движение при отсутствии внешних воздействий

функция системы $g(t,u)$. Для системы (3.4) эта реакция может быть установлена дифференциальным уравнением

$$\begin{aligned} a_r(t) \frac{d^r}{dt^r} g(t,u) + a_{r-1}(t) \frac{d^{r-1}}{dt^{r-1}} g(t,u) + \dots + a_0 g(t,u) = \\ = b_{r-1}(t) \frac{d^{r-1}}{dt^{r-1}} \delta(t-u) + b_{r-2}(t) \frac{d^{r-2}}{dt^{r-2}} \delta(t-u) + b_0 \delta(t-u) \end{aligned}$$

или в операторной форме

$$g(t,u) = w(p,u) \delta(t-u). \quad (3.7)$$

Момент времени i является параметром, определяющим момент приложения дельта-импульса. Анализируя действия этого импульса, отметим, что до момента его приложения система остается не возбужденной. Это означает, что функция $g(t,u)$ тождественно равна нулю при $t \leq u$, а в момент времени $t=u$ она сама и ее производные имеют разрыв. Условие $g(t,u)=0$ при $t < u$ соответствует физической реализуемости линейной динамической системы. На основании соотношений (3.6) и (3.7), используя принцип суперпозиции (3.3), получаем при нулевых начальных условиях выражение для выходного сигнала системы

$$\begin{aligned} x(t) = w(p,t) \xi(t) = w(p,t) \left[\int_0^t \xi(u) \delta(t-u) du \right] = \\ = \int_0^t \xi(u) w(p,t) \delta(t-u) du = \int_0^t g(t,u) \xi(u) du. \end{aligned} \quad (3.8)$$

Соотношения (3.6) и (3.8) раскрывают смысл функции $g(t,u)$. Выходной сигнал системы $x(t)$ в момент времени t представляет собой бесконечную сумму для разных моментов времени u ($0 < u \leq t$) значений входного сигнала $\xi(t)$, умноженных на весовой коэффициент $g(t,u)$, учитывающий вклад сигнала $\xi(u)$ в выходной сигнал в момент времени t .

Рассмотрим импульсную переходную функцию, у которой моменты ее измерения t и подачи на вход системы дельта-импульса u сдвинуты на t_0 : $g(t',u') = g(t+t_0, u+t_0)$.

Обозначим интервал времени между моментами измерения и приложения дельта-импульса $\tau=t'-u'=t-u$, получим $g(\tau, u)$. Если $g(\tau, u)=g(\tau)$, т.е. реакция системы зависит от интервала времени τ и не зависит от момента приложения дельта-импульса, то динамическая система становится стационарной. Выходной сигнал на основании (3.8) при нулевых начальных условиях определяется формулой

$$x(t) = \int_0^t g(t-u)\xi(u)du = \int_0^t g(u)\xi(t-u)du. \quad (3.9)$$

Понятия стационарности и линейности не связаны между собой. Стационарная система может быть как линейной, так и нелинейной, и соответственно линейная система может быть как стационарной, так и нестационарной. Рассматриваемая ранее линейная система ввиду зависимости оператора системы $w(p, t)$ явно от текущего времени является нестационарной. Покажем, что система с постоянными коэффициентами дифференциальных уравнений (с оператором, независящим от времени – $w(p)$) является стационарной. Для этого случая дифференциальное уравнение (3.5) в операторной форме имеет вид

$$A(p)x(t)=B(p)\xi(t). \quad (3.10)$$

Заменим параметр дифференцирования p на комплексный параметр s преобразования Лапласа. В этом случае изображение Лапласа входного и выходного сигналов принимают форму $\mathcal{L}\{\xi(t)\}=\bar{\xi}(s)$ и $\mathcal{L}\{x(t)\}=\bar{x}(s)$ и при нулевых начальных условиях связаны соотношением $\bar{x}(s)=w(s)\bar{\xi}(s)$.

Обозначив сигнал в пространстве оригиналов, соответствующий изображению $w(s)$ через $h(t)$ имеем на основании теоремы свертывания

$$x(t) = \int_0^t h(t-u)\xi(u)du. \quad (3.11)$$

Сравнивая (3.9) с выражением (3.11), приходим к равенству $h(\tau)=g(\tau)$, означающее зависимость функции $g(\cdot)$ только от одного параметра $\tau=t-u$. Формула (3.9), как ранее было показано, выражает

свойство стационарности. Следовательно, линейная система с постоянными коэффициентами стационарна.

Из соотношения (3.11) следует, что характеристика $w(s)$ является изображением Лапласа импульсной переходной функции линейной стационарной системы: $w(s)=\mathcal{L}\{g(t)\}$. В теории систем эта характеристика имеет особое название - передаточная функция системы. Путем замены $s=j\omega$ может быть получена частотная характеристика системы $w(j\omega)=|w(j\omega)|\exp\{j\varphi(\omega)\}$, которая связана с импульсной переходной функцией парой преобразования Фурье

$$w(j\omega) = \int_{-\infty}^{\infty} g(\tau) \exp\{-j\omega\tau\} d\tau,$$

$$g(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w(j\omega) \exp\{j\omega\tau\} d\omega,$$

где $|w(j\omega)|$ - амплитудно-частотная характеристика, $\varphi(\omega)$ - фазо-частотная характеристика.

Представим динамическую систему, состоящую из k последовательно соединенных линейных устройств с передаточными функциями $w_i(s)$. Передаточная функция системы имеет вид $w(s) = \prod_i^k w_i(s)$. Если обозначить через $g_{10}(\tau_1), g_{20}(\tau_2), \dots, g_{k0}(\tau_k) = g(t)$ импульсные переходные функции на выходе подсистем $w_1(s), w_1(s)w_2(s), \dots, \sum_i^k w_i(s)$, то импульсная переходная функция системы имеет вид

$$g(t) = g_{k0}(t) = \int_0^t g_k(t - \tau_{k-1}) d\tau_{k-1} \int_0^{\tau_{k-1}} g_{k-1}(t - \tau_{k-2}) d\tau_{k-2} \dots$$

$$\dots \int_0^{\tau_3} g_3(t - \tau_2) d\tau_2 \int_0^{\tau_2} g_2(t - \tau_1) g_{10}(\tau_1) d\tau_1, \quad g_{10}(\tau_1) = g_1(\tau).$$

На основании свойства ассоциативности свертка функций дает одинаковый результат независимо от того, в каком порядке выполняются операции свертывания.

При ненулевых начальных условиях выходной процесс определяется соотношением $x(t) = \varphi(t) + \int_0^t g(t-u)\xi(u)du$, где $\varphi(t)$ является решением однородного уравнения $A(p)x(t)=0$ при начальных условиях $x(0)=x(0), x'(0)=x_1(0), \dots, x^{(r-1)}(0)=x_{r-1}(0)$.

Для дискретных линейных систем модели случайных процессов в дискретном времени можно получить путем перехода от непрерывной модели с помощью следующей процедуры. Обозначим значение процесса x в равностоящие моменты времени $t=k\Delta t, k=0,1,\dots$ через $x(k\Delta t)=x_k$. Введем разностный оператор ∇ :

$$\nabla x_k = x_{k+1} - x_k, \nabla^2 x_k = \nabla x_{k+1} - \nabla x_k, \dots, \nabla^m x_k = \nabla^{m-1} x_{k+1} - \nabla^{m-1} x_k.$$

В этом случае формальным аналогом уравнения (3.4) с постоянными коэффициентами является линейное разностное уравнение порядка r

$$(\alpha_r \nabla^r + \alpha_{r-1} \nabla^{r-1} + \dots + \alpha_0)x_k = (\beta_{r-1} \nabla^{r-1} + \beta_{r-2} \nabla^{r-2} + \dots + \beta_0)\xi_k.$$

Необходимо отметить, что если от непрерывной модели можно всегда осуществить переход к дискретной, то обратный переход в общем случае невозможен. Во многих задачах дискретные модели представляют самостоятельный интерес и соответствующие ей непрерывные модели не существуют.

Аналогом формулы (3.9) для линейных стационарных систем при нулевых начальных условиях является соотношение $x_k = \sum_{i=1}^{k-1} g_{k-i}\xi_i$, где g_i - некоторые весовые коэффициенты. Физической реализуемости указанной системы соответствует условие $g_{k-i}=0$ при $k < i$.

Вернемся к рассмотрению системы уравнений в пространстве состояний (3.2). Решением ее является вектор состояний, определяемый при нулевых начальных условиях из соотношения

$$x(t) = \int_{t_0}^t g(t,u)\xi(u)du, \quad (3.12)$$

где $g(t,u)$ - матрица импульсных переходных функций. Для уравнения с постоянными коэффициентами компоненты этой матрицы зависят от разности моментов времени $\tau=t-u$.

Представим эту матрицу в следующем виде

$$g(t,u)=\Phi_0(t,u)\Gamma_u, \quad (3.13)$$

где $\Phi_0(t,u)$ - матрица весовых коэффициентов, являющаяся решением дифференциального уравнения

$$\frac{\partial \Phi_0(t,u)}{\partial t} = F_t \Phi_0(t,u) + I_\delta(t-u), \quad \Phi_0(u,u) = 0. \quad (3.14)$$

при входном воздействии в виде диагональной матрицы

$$I_\delta(t-u) = \begin{Bmatrix} \delta(t-u) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \delta(t-u) \end{Bmatrix},$$

$\delta(t-u)$ - дельта-функция^{*)}

Решением уравнения (3.14) является соотношение

$$\Phi_0(t,u) = \exp \left\{ \int_u^t F_\tau d\tau \right\}.$$

Аналогичный результат можно получить после решения однородного уравнения с отличными от нуля начальными условиями

$$\frac{\partial \Phi_0(t,u)}{\partial t} = F_t \Phi_0(t,u), \quad \Phi_0(u,u) = I(u), \quad (3.15)$$

где единичная матрица $I(u) = \begin{Bmatrix} 1(u) & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1(u) \end{Bmatrix}.$

В матрице $\Phi_0(t,u)$ ее компоненты непрерывны по t и дискретны по u . В некоторых приложениях удобно иметь матрицу $\Phi_0^T(u,t)$, в

^{*)} В теории линейных дифференциальных уравнений матрицу $\Phi_0(t,u)$ также называют фундаментальной матрицей.

которой компоненты непрерывны по u и дискретны по t . Эта матрица является решением сопряженного к (3.15) уравнения

$$\frac{\partial \Phi_0^T(u, t)}{\partial u} = \Phi_0^T(u, t) F_t^T, \quad \Phi_0^T(t, t) = I(t). \quad (3.16)$$

Общее решение уравнения (3.2) с учетом начальных условий на основании принципа суперпозиции запишется в виде

$$x(t) = \Phi_0(t, t_0)x(t_0) + \int_{t_0}^t \Phi_0(t, u) \Gamma_u \xi(u) du, \quad (3.17)$$

где $\Phi_0(t, t_0)$ - матрица весовых коэффициентов, являющаяся решением однородного уравнения

$$\frac{\partial \Phi(t, t_0)}{\partial t} = F_t \Phi(t, t_0)$$

с начальными условиями - единичной матрицей $I(t_0)$.

3.2. УСТОЙЧИВОСТЬ ДИНАМИЧЕСКИХ СИСТЕМ

Одним из обязательных требований, обеспечивающим работоспособность динамических систем, является устойчивость. Под устойчивостью обычно понимается свойство системы возвращаться к первоначальному состоянию после прекращения действия внешних возмущений. Дальнейшее рассмотрение материала в пособии не связано с применением критериев устойчивости и динамическую систему полагают всегда устойчивой. Поэтому ограничимся кратким изложением известных в теории систем понятий и определений устойчивости.

Допустим, что поведение системы описывается уравнением в векторной форме $\frac{dx}{dt} = f_t(x)$, характеризующим собственное (свободное) движение системы. Компоненты вектора x имеют единственное решение $x_i = \lambda_i(x_{i0}, t)$, удовлетворяющее начальным условиям $x_{i0} = \lambda_i(x_{i0}, t_0)$ для момента времени t_0 , $t > t_0$. Рассмотрим некоторое решение $x_i = \lambda_i$ на интервале (t_0, ∞) . Это решение называется устойчивым по Ляпунову при $t \rightarrow \infty$, если для любого $\varepsilon > 0$ существует

такое $\delta > 0$, зависящее от ε и t_0 , что любое решение $x_i = \varphi_i(t)$ для которого при $t = t_0$ выполняются неравенство $|\varphi_i(t_0) - \lambda_i(t_0)| < \delta$, удовлетворяет неравенству $|\varphi_i(t) - \lambda_i(t)| < \varepsilon$, для всех $i = \overline{1, r}$. Геометрически это означает, что все решения, которые при $t = t_0$ начинаются в δ -окрестности, никогда не покинут ε -трубку решения $\lambda(t)$.

Решение $\lambda_i(t)$ называется неустойчивым, если существует $\varepsilon > 0$ такое, что для любого $\delta > 0$ найдется такой момент времени $t = t_1$ что для некоторого $i = k$ и $t = t_1$ будет выполняться неравенство $|\varphi_k(t_1) - \lambda_k(t_1)| \geq \varepsilon$, несмотря на то, что $|\varphi_i(t_0) - \lambda_i(t_0)| < \delta$ для всех $i = \overline{1, r}$.

Обратим внимание на одно обстоятельство. Приведенное определение устойчивости предусматривает сравнение свободных движений при двух близких начальных состояниях. Движение системы свободно, если в процессе ее работы возмущения не действуют. В реальных условиях всегда имеется возмущение. Поэтому определение устойчивости целесообразно изменить, используя положение о возможности замены начальных состояний эквивалентными воздействиями. Как с физической, так и с математической точек зрения нет никакой необходимости различать воздействия, эквивалентные начальным условиям, и обычные воздействия. Поэтому определение устойчивости движения системы следует основывать на оценке суммарного воздействия, включающего в себя как эквивалентное воздействие, так и сигнал, поступающий извне на вход системы. В этом случае движение системы будет устойчиво, если при малом изменении суммарного входного воздействия $\delta \xi(t)$ изменения выходного сигнала системы $\delta x(t)$ малы по абсолютной величине в диапазоне достаточно больших значений аргумента. Система, устойчивая в смысле уточненного определения называется системой, устойчивой по воздействиям.

В настоящее время наиболее изучены вопросы устойчивости линейных динамических систем с постоянными коэффициентами. При исследовании устойчивости таких систем широко применяются критерии, позволяющие по виду передаточной функции $w(s)$ системы установить устойчива система или нет. В этом случае движение системы устойчиво, если все полюса передаточной функции (корни характеристического уравнения $A(p) = 0$) имеют отрицательные вещественные части. Граничные случаи, когда один или несколько

полюсов функции $w(s)$ находятся в точке $s=0$ или на мнимой оси относятся, как правило, к неработоспособным системам.

Благодаря теоремам Ляпунова об устойчивости по первому приближению теория устойчивости линейных систем служит основным инструментом при исследовании устойчивости реальных систем при малых возмущениях. Однако при исследовании устойчивости при больших возмущениях теория линейных систем в большинстве случаев либо вообще не позволяет обнаружить важные свойства системы, либо приводит к результатам мало пригодным для количественных оценок. Наиболее существенное развитие при исследовании устойчивости нелинейных систем и больших возмущениях получил прямой метод Ляпунова, для которого сформулирован ряд теорем [22,23,24]. Это метод позволяет с помощью вспомогательных знакопостоянных и знакоопределенных функций исследовать устойчивость решений нелинейных дифференциальных уравнений, не производя решения самих уравнений.

3.3. СЛУЧАЙНЫЕ ПРОЦЕССЫ В ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМАХ

3.3.1. ОБЩИЕ ОПРЕДЕЛЕНИЯ

Соотношение (3.17) определяет реакцию динамической системы с известными параметрами на входное воздействие или возмущение $\xi(t)$. Если повторять опыты по исследованию реакции системы на эти возмущения, создаваемые реальными устройствами, и осуществлять запись выходных сигналов при каждом опыте, то можно обнаружить, что записи не совпадают. С какой бы тщательностью не ставился последующий опыт, на вход системы не удастся подать сигналы аналогичные по характеру изменения, что и в предыдущем опыте. На закон изменения входного сигнала находящейся в условиях эксплуатации динамической системы оказывают влияния большое количество различных не поддающихся всестороннему теоретическому учету факторы: помехи и шумы, климатические и метеорологические условия, источники электромагнитных излучений окружающей среды. Следовательно, выходные сигналы системы также будут меняться от опыта к опыту. Процессы, конкретное протекание или реализации которых изменяются от опыта к опыту,

как было отмечено ранее, являются случайными процессами. В ряде практических задач случайностью процессов в системах можно пренебречь, считать, что функционирование системы осуществляется при некоторых осредненных характеристиках внешних условий, а сами входные и выходные сигналы полагать практически детерминированными. Наряду с ними имеется большой класс задач, где учет осредненных условий недостаточен и необходимо знать возможные отклонения от средних значений и влияние этих отклонений на характеристики качества работы системы. Наиболее типичными примерами таких задач являются задачи, связанные с разработкой систем связи в условиях воздействия шумов и внешних помех, с созданием систем управления летательных и других подвижных объектов, приводящих объект управления в заданную точку с наименьшей погрешностью, с проектированием адаптивных систем, определяющих режим своей дальнейшей работы по результатам статистической обработки информации о внешних условиях до данного момента времени и т.д.

В разд.1 и 2 рассматривались типы случайных процессов и методы их описания. Было отмечено, что наиболее полной вероятностной характеристикой случайного процесса являются законы распределения и, в частности, функции распределения вероятностей (плотностей вероятностей) переменных состояния системы. Также было отмечено, что задачей исследования динамических систем является преобразование входного воздействующего процесса в выходной. Для случайных процессов эта задача формулируется следующим образом. Предполагая известными параметры математической модели динамической системы (или известным оператор этой системы), необходимо определить вероятностные характеристики выходного процесса по вероятностным характеристикам входного воздействия. В наиболее общей постановке задача анализа сводится к определению плотности вероятностей $p(x,t)$ или параметров ее заменяющих (моментных или корреляционных функций разных порядков) случайных процессов на выходе системы с известными параметрами (описываемой известными дифференциальными уравнениями) по плотности вероятностей $p(\xi,t)$ случайных процессов на входе системы.

За исключением гауссовских и марковских процессов и небольшого числа частных примеров, не существует прямых

методов, которые позволили бы непосредственно по плотности $p(\xi, t)$ находить плотность вероятностей $p(x, t)$. Существуют приближенные достаточно сложные и трудоемкие методы пересчета вероятностных характеристик, суть которых заключается в следующем. По известной плотности вероятностей $p(\xi, t)$ на основании прямого преобразования Фурье определяют характеристическую функцию. Представляя характеристическую функцию в виде многомерного разложения в ряд (1.6), определяют моментные или корреляционные функции процесса $\xi(t)$. На основании характеристик динамической системы, например, импульсной переходной функции, пересчитывают их и находят некоторое число моментных или корреляционных функций выходного процесса $x(t)$. Определив их, записывают в виде конечного ряда разложение характеристической функции выходного процесса. А затем путем обратного преобразования Фурье - приближенное выражение плотности вероятностей $p(x, t)$.

В случае гауссовского процесса задача упрощается, поскольку он полностью определяется двумя статистическими характеристиками: вектором математического ожидания и матрицей корреляционных функций. После линейного преобразования выходной процесс системы остается гауссовским с пересчитанными математическими ожиданиями и корреляционными функциями. Для марковских процессов плотность вероятностей выходного процесса определяется из уравнения ФПК (2.38), коэффициенты сноса и диффузии которого зависят от параметров дифференциальных уравнений динамической системы.

Рассмотренные выше подходы по статистическому исследованию динамических систем являются содержанием задач вероятностного анализа. Определяющим при этом оказывается знание функции распределения (плотности вероятностей) или моментных (корреляционных функций первого, второго и более высоких порядков). Наиболее полно и в достаточно общем виде эти задачи решаются для линейных систем и гауссовских процессов с использованием моментных (корреляционных) функций первых двух порядков. Во многих случаях для исследования не только линейных, но и нелинейных систем также практически достаточными являются оценки статистических характеристик первых двух порядков. Область вероятностного анализа, ограничивающаяся изучением

указанных статистических характеристик при преобразовании случайных сигналов, получила название корреляционной теории (корреляционного анализа).

3.3.2. ПРЕОБРАЗОВАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ В ЛИНЕЙНЫХ СИСТЕМАХ

В рамках корреляционной теории задачей вероятностного анализа является установление правил пересчета статистических характеристик первых двух порядков (математического ожидания, дисперсий, корреляционных функций, спектральных плотностей). Воспользуемся приведенными в разделе 3.1 данными по определению статистических характеристик через операции взятия математических ожиданий. Выходной процесс $x(t)$ одномерной линейной динамической системы при воздействии входного сигнала при нулевых начальных условиях определяется, начиная с момента времени t_0 , соотношением (3.8). Учитывая, что для линейных систем операции взятия математического ожидания и интегрирования можно менять местами, получаем выражения математического ожидания, корреляционной функции и дисперсии выходного процесса системы

$$\begin{aligned}
 m_x(t) = M\{x(t)\} &= M\left\{\int_{t_0}^t g(t,u)\xi(u)du\right\} = \int_{t_0}^t g(t,u)M\{\xi(u)\}du = \\
 &= \int_{t_0}^t g(t,u)m_\xi(u)du;
 \end{aligned}
 \tag{3.18}$$

$$\begin{aligned}
k_x(t_1, t_2) &= M \{ [x(t_1) - m_x(t_1)][x(t_2) - m_x(t_2)] \} \\
&= M \left\{ \int_{t_0}^{t_1} \int_{t_0}^{t_2} g(t_1, u_1) g(t_2, u_2) \times \right. \\
&\quad \times [\xi(u_1) - m_\xi(u_1)][\xi(u_2) - m_\xi(u_2)] du_1 du_2 \Big\} = \\
&= \int_{t_0}^{t_1} \int_{t_0}^{t_2} g(t_1, u_1) g(t_2, u_2) \times \\
&\quad \times M \{ [\xi(u_1) - m_\xi(u_1)][\xi(u_2) - m_\xi(u_2)] \} du_1 du_2 = \\
&= \int_{t_0}^{t_1} \int_{t_0}^{t_2} g(t_1, u_1) g(t_2, u_2) k_\xi(u_1, u_2) du_1 du_2;
\end{aligned} \tag{3.19}$$

$$D_x(t) = k_x(t, t) = \int_{t_0}^t \int_{t_0}^t g(t_1, u_1) g(t, u_2) k_\xi(u_1, u_2) du_1 du_2, \tag{3.20}$$

где $m_\xi(t)$ и $k_\xi(t_1, t_2)$ - математическое ожидание и корреляционная функция входного сигнала.

Для стационарных систем в формулах (3.18), (3.19) и (3.20) необходимо подставить импульсную переходную функцию, зависящую от разности аргументов: $g(t, u) = g(t - u)$.

Допустим, что на вход стационарной системы подается стационарный в широком смысле процесс, математическое ожидание которого постоянно, а корреляционная функция зависит от разности аргументов. Установившемся режиму соответствует достаточно большое время работы системы, что обеспечивается, если нижние пределы интегрирования в формулах (3.18), (3.19) и (3.20) принять равными $-\infty$. Соответствующие выражения статистических характеристик принимают вид

$$m_x = \int_{-\infty}^{\infty} g(t - u) m_\xi du = m_\xi \int_0^{\infty} g(\lambda) d\lambda, \tag{3.21}$$

$$\begin{aligned}
k_x(\tau) &= \int_{-\infty}^t \int_{-\infty}^t g(t_1 - u_1) g(t_2 - u_2) k_\xi(u_1 - u_2) du_1 du_2 = \\
&= \int_0^\infty \int_0^\infty g(\lambda_1) g(\lambda_2) k_\xi(\tau + \lambda_1 - \lambda_2) d\lambda_1 d\lambda_2,
\end{aligned} \tag{3.22}$$

$$D_x = k_x(0) = \int_0^\infty \int_0^\infty g(\lambda_1) g(\lambda_2) k_\xi(\lambda_1 - \lambda_2) d\lambda_1 d\lambda_2, \tag{3.23}$$

где $\tau=t_2-t_1$, $\lambda=t-u$, $\lambda_1=t_1-u_1$, $\lambda_2=t_2-u_2$.

Если входной процесс нестационарный, а преобразующая система по-прежнему стационарна, то статистические характеристики при нулевых начальных условиях, соответственно равны

$$m_x(t) = \int_0^t g(t-u) m_\xi(u) du, \tag{3.24}$$

$$k_x(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} g(t_1 - u_1) g(t_2 - u_2) k_\xi(u_1, u_2) du_1 du_2, \tag{3.25}$$

$$D_x = k_x(t, t) = \int_0^t \int_0^t g(t - u_1) g(t - u_2) k_\xi(u_1, u_2) du_1 du_2. \tag{3.26}$$

В частном случае, когда входной сигнал - стационарный белый шум $\xi(t)=\eta(t)$ с $m_\eta=0$ и корреляционной функцией $k_\eta(\tau)=\mathcal{N}_\eta \delta(\tau)$, статистические характеристики выходного сигнала на основании (3.21), (3.22) и (3.23) с учетом фильтрующих свойств свойств дельта-функции соответственно равны

$$m_x = 0;$$

$$k_x(\tau) = \mathcal{N}_\eta \int_0^\infty g(\lambda) g(\lambda + \tau) d\lambda;$$

$$D_x = \mathcal{N}_\eta \int_0^\infty g^2(\lambda) d\lambda.$$

Спектральная плотность выходного процесса при стационарных воздействиях на основании формул (1.17) и (3.22) принимает вид

$$F_x(\omega) = \int_{-\infty}^{\infty} k_x(\tau) \exp\{-j\omega\tau\} d\tau = \\ = \int_0^{\infty} g(\lambda_1) d\lambda_1 \times \int_0^{\infty} g(\lambda_2) d\lambda_2 \int_{-\infty}^{\infty} k_{\xi}(\tau + \lambda_1 - \lambda_2) \exp\{-j\omega\tau\} d\tau.$$

Введем новую переменную $\tau + \lambda_1 - \lambda_2 = u$, получаем

$$F_x(\omega) = \int_0^{\infty} g(\lambda_1) \exp\{-j\omega\lambda_1\} d\lambda_1 \int_0^{\infty} g(\lambda_2) \exp\{j\omega\lambda_2\} d\lambda_2 \times \\ \times \int_{-\infty}^{\infty} k_{\xi}(u) \exp\{-j\omega u\} du.$$

Но согласно свойствам преобразования Лапласа и известному соответствию импульсной переходной функции передаточной функции имеем

$$\int_0^{\infty} g(\lambda_1) \exp\{j\omega\lambda_1\} d\lambda_1 = w(-j\omega), \quad \int_0^{\infty} g(\lambda_2) \exp\{-j\omega\lambda_2\} d\lambda_2 = w(j\omega).$$

Кроме того, $\int_{-\infty}^{\infty} k_{\xi}(u) \exp\{-j\omega u\} du = F_{\xi}(\omega)$, поэтому

$$F_x(\omega) = w(j\omega)w(-j\omega)F_{\xi}(\omega) = |w(j\omega)|^2 F_{\xi}(\omega).$$

Соотношение (3.27) позволяет определить спектральную плотность выходного процесса системы с амплитудно-частотной характеристикой $|w(j\omega)|$ по известной спектральной плотности входного воздействия. При входном сигнале - белом шуме, имеем $F_{\xi}(\omega) = \mathcal{N}_{\eta}$, и спектральная плотность выходного сигнала равна $F_x(\omega) = |w(j\omega)|^2 \mathcal{N}_{\eta}$. Если выражение $|w(j\omega)|^2$ является дробно-рациональной относительно ω^2 функцией, то выходной сигнал представляет собой, как было отмечено в разделе 2.3.7, гауссовский марковский процесс.

Для многомерной системы, описываемой линейным дифференциальным уравнением в векторно-матричной форме (3.2), статистические характеристики первых двух порядков могут быть получены следующим образом. Учитывая, что для линейных систем операции взятия математического ожидания и интегрирования можно менять местами, аналогичным образом, что для одномерного случая, на основании (3.12) при нулевых начальных условиях после усреднения можно записать вектор математического ожидания выходного процесса $m_x(t) = M\{x(t)\} = \int_{t_0}^t g(t,u)m_\xi(u)du$ или с учетом (3.13)

$$m_x(t) = \int_{t_0}^t \Phi_0(t,u)\Gamma_u m_\xi(u)du.$$

Матрицы корреляционных функций и дисперсий соответственно определяются из выражений

$$\begin{aligned} k_x(t_1, t_2) &= M\{[x(t_1) - m_x(t_1)][x(t_2) - m_x(t_2)]^T\} = \\ &= \int_{t_0}^{t_1} \int_{t_0}^{t_2} \Phi_0(t_1, u_1)\Gamma_{u_1} k_\xi(u_1, u_2)\Gamma_{u_2}^T \Phi_0^T(t_2, u_2)du_1 du_2, \end{aligned}$$

$$D_x(t) = k_x(t, t) = \int_{t_0}^t \int_{t_0}^t \Phi_0(t, u_1)\Gamma_{u_1} k_\xi(u_1, u_2)\Gamma_{u_2}^T \Phi_0^T(t, u_2)du_1 du_2,$$

где $k_\xi(u_1, u_2)$ - матрица корреляционных функций входного сигнала.

Для стационарной системы с постоянными коэффициентами дифференциальных уравнений имеем

$$\begin{aligned} m_x(t) &= \int_{t_0}^t \Phi_0(t-u)\Gamma m_\xi(u)du, \\ k_x(t_1, t_2) &= \int_{t_0}^{t_1} \int_{t_0}^{t_2} \Phi_0(t_1-u_1)\Gamma k_\xi(u_1, u_2)\Gamma^T \Phi_0^T(t_2-u_2)du_1 du_2, \end{aligned}$$

$$D_x(t) = \int_{t_0}^t \int_{t_0}^t \Phi_0(t-u_1) \Gamma k_\xi(u_1, u_2) \Gamma^T \Phi_0^T(t-u_2) du_1 du_2.$$

В установившемся режиме при воздействии стационарного процесса $\xi(t)$ соответствующие статистические характеристики равны

$$m_x = \int_{-\infty}^t \Phi_0(t-u) \Gamma m_\xi du = \int_0^\infty \Phi_0(\lambda) \Gamma m_\xi d\lambda,$$

$$\begin{aligned} k_x(\tau) &= \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} \Phi_0(t_1-u_1) \Gamma k_\xi(u_1, u_2) \Gamma^T \Phi_0^T(t_2-u_2) du_1 du_2 = \\ &= \int_0^\infty \int_0^\infty \Phi_0(\lambda_1) \Gamma k_\xi(\tau + \lambda_1 - \lambda_2) \Gamma^T \Phi_0^T(\lambda_2) d\lambda_1 d\lambda_2, \end{aligned}$$

$$\begin{aligned} D_x(t) &= \int_{-\infty}^t \int_{-\infty}^t \Phi_0(t-u_1) \Gamma k_\xi(u_1, u_2) \Gamma^T \Phi_0^T(t-u_2) du_1 du_2 = \\ &= \int_0^\infty \int_0^\infty \Phi_0(\lambda_1) \Gamma k_\xi(\lambda_1 - \lambda_2) \Gamma^T \Phi_0^T(\lambda_2) d\lambda_1 d\lambda_2, \end{aligned}$$

где $\tau=t_2-t_1$, $\lambda=t-u$, $\lambda_1=t_1-u_1$, $\lambda_2=t_2-u_2$.

В частном случае, когда воздействующий сигнал представляет собой многокомпонентный белый шум $\xi(t)=\eta(t)$ с $m_\eta=0$ и корреляционной матрицей $\mathcal{N}_\eta \delta(\tau)$, имеем

$$m_x = 0, \quad (3.28)$$

$$k_x(\tau) = \int_0^\infty \Phi_0(\lambda) \Gamma \mathcal{N}_\eta \Gamma^T \Phi_0^T(\lambda + \tau) d\lambda, \quad (3.29)$$

$$D_x = \int_0^\infty \Phi_0(\lambda) \Gamma \mathcal{N}_\eta \Gamma^T \Phi_0^T(\lambda) d\lambda. \quad (3.30)$$

Напомним, что при воздействии на линейную систему белого шума уравнения состояния системы описывают многомерный

гауссовский марковский процесса, а формулы (3.28), (3.29) и (3.30) выражают статистические характеристики этого процесса.

3.3.3. СЛУЧАЙНЫЕ ИМПУЛЬСНЫЕ ПРОЦЕССЫ

Полученная выше импульсная переходная функция является решением дифференциального уравнения (3.7) при входном сигнале - дельта-импульсе и нулевых начальных условиях. Очевидно, что при воздействии случайной последовательности дельта-импульсов на выходе системы, характеризуемой этим уравнением, возникает последовательность импульсных переходных функций, начальные временные координаты которых совпадают с координатами дельта-импульсов. Для формирования импульсных процессов разного типа наиболее общим уравнением служит (2.95) или линейное уравнение с постоянными коэффициентами, решением которого является выражение (3.9) при $\xi(t)=\eta^N(t)$. Рассматриваемый подход достаточно конструктивен, так как во-первых, расширяет арсенал используемых в информационных системах моделей марковских процессов, во-вторых, позволяет смоделировать реальные импульсные сигналы и определить статистические характеристики импульсных и «гладких» (из-за перекрытия импульсных переходных функций) выходных процессов системы. Для наблюдаемых в устройствах обработки информации импульсных процессов довольно часто в качестве случайного параметра рассматривается только положение импульса по временной оси. Остальные параметры (форма, амплитуда, длительность импульса и т.д.) полагают постоянными. Сформированный таким образом импульсный процесс называется случайной импульсной последовательностью. Статистические характеристики случайных импульсных последовательностей (математическое ожидание, корреляционная функция, спектральная плотность) могут быть получены на основании рассмотренных в разделе 3.3.2 методов преобразования случайных процессов.

Перейдем к рассмотрению основных характеристик случайных импульсных последовательностей. Допустим, что на линейную стационарную систему в момент времени t_i воздействует дельта-импульс. Выходным сигналом системы при нулевых начальных условиях является импульс, описываемый импульсной переходной функцией $g(t-t_i)$. «Пропуская» последовательность дельта-импульсов через эту систему, которую назовем формирующим фильтром,

получим на выходе фильтра последовательность импульсов, форма которых определяется функцией формы $\Psi(t)=g(t)$.

Используя выражение (2.84) для случайного импульсного потока и фильтрующее свойство дельта-функции, получим на основании формулы (3.9) выражение, описывающее случайную импульсную последовательность

$$\begin{aligned} x(t) &= \int_0^t g(t-u) \eta^N(u) du = \int_0^t g(t-u) \sum_i \delta(u-t_i) du = \\ &= \int_0^t \Psi(t-u) \sum_i \delta(u-t_i) du = \sum_i \Psi(t-t_i). \end{aligned}$$

Таким образом, случайная импульсная последовательность является суперпозицией импульсных переходных функций (фильтрованным пуассоновским процессом).

Определение характеристик приведем для двух типов стационарных импульсных последовательностей, построенных на основе пуассоновского импульсного потока: с экспоненциальными и прямоугольными импульсами единичной амплитуды.

Воздействующий стационарный пуассоновский импульсный поток на основании формул (2.89), (2.90) и (2.91) имеет математическое ожидание, корреляционную функцию и спектральную плотность, соответственно равные

$$m_\eta = \nu, \quad k_\eta(\tau) = \nu \delta(\tau) \quad \text{и} \quad F_\eta(\omega) = \nu. \quad (3.31)$$

Математическое ожидание, корреляционная функция и спектральная плотность случайной импульсной последовательности на основании соотношений (3.21), (3.22) и (3.27) равны

$$m_x = \nu \int_0^\infty \Psi(\lambda) d\lambda, \quad (3.32)$$

$$\begin{aligned}
k_x(\tau) &= \int_0^\infty \int_0^\infty \Psi(\lambda_1) \Psi(\lambda_2) \delta(\tau + \lambda_1 - \lambda_2) d\lambda_1 d\lambda_2 = \\
&= \nu \int_0^\infty \Psi(\lambda) \Psi(\tau + \lambda) d\lambda,
\end{aligned}
\tag{3.33}$$

$$F_x(\omega) = |w_\Psi(j\omega)|^2 \nu, \tag{3.34}$$

где $w_\Psi(j\omega)$ - Фурье-преобразование функции формы.

Для импульсной последовательности с экспоненциальными импульсами имеем

$$\Psi(t) = \alpha \exp\{-\alpha t\}, \quad w_\Psi(j\omega) = \frac{\alpha}{\alpha + j\omega}.$$

Математическое ожидание на основании (3.32) равно

$$m_x = \alpha \nu \int_0^\infty \exp\{-\alpha \nu\} d\lambda = \nu.$$

Корреляционную функцию определяют из следующих соображений. На основании формулы (3.33)

$$k'_x(\tau) = \alpha^2 \nu \int_0^\infty \exp\{-\alpha(2\lambda + \tau)\} d\lambda.$$

В этом интеграле величина $\tau > 0$, что соответствует «прямому» течению времени ($\lambda > 0$). Для величины $\tau_1 = -\tau < 0$ и «обратному» течению времени ($\lambda_1 = -\lambda < 0$) имеем

$$k''_x(\tau) = \alpha^2 \nu \int_{-\infty}^0 \exp\{\alpha(2\lambda_1 + \tau_1)\} d\lambda_1.$$

Объединив результаты интегрирования по обеим областям, получим

$$k_x(\tau) = \frac{\alpha \nu}{2} \exp\{-\alpha |\tau|\}.$$

Спектральная плотность на основании (3.34) равна

$$F_x(\omega) = \frac{\alpha^2}{\alpha^2 + \omega^2}.$$

В случае импульсной последовательности с прямоугольными импульсами имеем

$$\Psi(t) = 1(t) - 1(t - \tau_u), \quad w_\Psi = \frac{1 - \exp\{-j\omega\tau_u\}}{j\omega},$$

где τ_u - длительность импульса,

$$1(t - \tau_u) = \begin{cases} 0, & t < \tau_u, \\ 1, & t \geq \tau_u. \end{cases}$$

Статистические характеристики для рассматриваемого случая имеют вид

$$m_x = \nu \int_0^\infty [1(\lambda) - 1(\lambda - \tau_u)] d\lambda = \nu\tau_u,$$

$$k_x(\tau) = \nu \int_0^\infty [1(\lambda) - 1(\lambda - \tau_u)][1(\lambda + \tau) - 1(\lambda + \tau - \tau_u)] d\lambda,$$

Объединив результаты интегрирования по областям $\tau > 0$, $\lambda > 0$ и $\tau < 0$, $\lambda < 0$, получим

$$k_x(\tau) = \begin{cases} 0, & |\tau| > \tau_u, \\ \nu\tau_u \left(1 - \frac{|\tau|}{\tau_u}\right), & |\tau| < \tau_u. \end{cases}$$

Спектральная плотность, вычисленная из формулы (3.34), равна

$$F_x(\omega) = \frac{4\nu}{(\omega\tau_u)^2} \sin^2\left(\frac{\omega\tau_u}{2}\right).$$

Обобщим полученные результаты на задачи преобразования случайных импульсных потоков линейными стационарными

устройствами с известной импульсной переходной функцией $g(t)$. Для стационарных пуассоновских импульсных потоков после подстановки в формулы (3.21), (3.22) и (3.23) соответствующих исходных характеристик (3.31), используя фильтрующие свойства дельта-функции, получим математическое ожидание, корреляционную функцию, дисперсию и спектральную плотность случайного процесса на выходе преобразующего устройства

$$m_x = \nu \int_0^{\infty} g(\lambda) d\lambda, \quad (3.35)$$

$$k_x(\tau) = \nu \int_0^{\infty} g(\lambda) g(\lambda + \tau) d\lambda, \quad (3.36)$$

$$D_x = k_x(0) = \nu \int_0^{\infty} g^2(\lambda) d\lambda, \quad (3.37)$$

$$F_x(\omega) = |w(j\omega)|^2 \nu,$$

где $|w(j\omega)|$ - амплитудно-частотная характеристика преобразующего устройства^{*)}.

Если пуассоновский импульсный поток модулируется стационарным случайным сигналом $z(t)$ с математическим ожиданием равным нулю и корреляционной функцией $k_z(\tau)$, то соответствующие характеристики преобразованного процесса $x(t)$ на основании формул (3.21), (3.22) и (3.23), а также (2.87) и (2.88), в которых необходимо положить $f_1 = \nu$, равны

$$m_x = M\{\nu\} \int_0^{\infty} g(\lambda) d\lambda$$

$$k_x(\tau) = M\{\nu\} \int_0^{\infty} g(\lambda) g(\lambda + \tau) d\lambda + m_z \int_0^{\infty} \int_0^{\infty} g(\lambda_1) g(\lambda_2) k_z(\tau + \lambda_1 - \lambda_2) d\lambda_1 d\lambda_2,$$

^{*)} Формулы (3.35), (3.36) и (3.37) являются следствием теоремы Кембелла о суперпозиции независимых случайных возмущений (импульсов).

$$D_x = k_x(0) = M\{v\} \int_0^{\infty} g^2(\lambda) d\lambda + m_z \int_0^{\infty} \int_0^{\infty} g(\lambda_1) g(\lambda_2) k_z(\lambda_1 - \lambda_2) d\lambda_1 d\lambda_2,$$

$$F_x(\omega) = |w(j\omega)|^2 [M\{v\} + m_z F_z(\omega)],$$

где m_z - масштабный коэффициент, $M\{\cdot\}$ означает операцию усреднения относительно плотности вероятностей сигнала z .

В случае нестационарного пуассоновского импульсного потока, когда интенсивность является неслучайной функцией времени $v(t)$, а корреляционная функция - $k_z(t_1, t_2) = v(t_1) \delta(t_1 - t_2)$, соответствующие статистические характеристики выходного процесса определяются из формул (3.24), (3.25) и (3.26)

$$m_x(t) = \int_0^t g(t-u) v(u) du, \quad (3.38)$$

$$k_x(t_1, t_2) = \int_0^{\min(t_1, t_2)} g(t_1-u) g(t_2-u) v(u) du, \quad (3.39)$$

$$D_x(t) = k_x(t, t) = \int_0^t g^2(t-u) v(u) du. \quad (3.40)$$

В заключение этого раздела на основании полученных результатов определим математическое ожидание и дисперсию случайного процесса на выходе сглаживающего фильтра (линейного измерителя скорости счета) с импульсной переходной функцией $g(t) = \alpha_f \exp\{-\alpha t\}$ при подаче на его вход нестационарного импульсного потока с ИП, аппроксимируемого выражением $v(t) = v_0 + A \exp\{-\beta H\}$, $H = H_0 - Vt$. С такими задачами сталкиваются при измерении высоты положения летательного аппарата с помощью радиоизотопных высотомеров по обратно-рассеянному излучению. Здесь H - высота, v_0 - интенсивность фоновой составляющей потока излучения, A и β - параметры высотомера. Рассмотрим установившийся режим, которому соответствует постоянная скорость спуска V . На основании формулы (3.38), математическое ожидание случайного процесса на выходе фильтра в установившемся режиме

$$\begin{aligned}
m_x &= \alpha_\Phi \int_{-\infty}^t [\nu_0 + A \exp\{-\beta(H_0 - Vu)\}] \exp\{-\alpha_\Phi(t - u)\} du = \\
&= \nu_0 + \frac{A \exp\{-\beta H\}}{1 - \alpha_\Phi^{-1} \beta V}.
\end{aligned}$$

Используя формулу (3.40), получаем дисперсию выходного процесса

$$\begin{aligned}
D_x(t) &= \alpha_\Phi^2 \int_{-\infty}^t [\nu_0 + A \exp\{-\beta(H_0 - Vu)\}] \exp\{-2\alpha_\Phi(t - u)\} du = \\
&= \frac{\alpha_\Phi \nu_0}{2} + \frac{A \alpha_\Phi \exp\{-\beta H\}}{2 + \alpha_\Phi^{-1} \beta V}.
\end{aligned}$$

3.3.4. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК

3.3.4.1. ВОЗДЕЙСТВУЮЩИЙ ПРОЦЕСС – БЕЛЫЙ ШУМ

Одной из важных задач корреляционного анализа является составление дифференциальных уравнений статистических характеристик вектора состояния системы и интегрирования их при заданных начальных условиях. В практических расчетах корреляционного анализа, как уже было отмечено ранее, ограничиваются наиболее доступными для исследований статистическими характеристиками первого и второго порядка. Отметим, что для линейных систем и гауссовских процессов знание этих характеристик полностью определяют плотности вероятностей случайных процессов.

Дальнейшее изложение материала базируется на использовании методов теории марковских процессов. Вопросы аппроксимации реальных сигналов в динамических системах марковскими процессами рассматривались в разделе 2.3. Было показано, что при воздействии белого шума (или винеровского процесса) характер динамики линейных и нелинейных систем в пространстве состояний определяется стохастическими дифференциальными уравнениями (2.48) и (2.50), описывающими многомерный марковский процесс.

Рассмотрение начнем с линейных динамических систем, описываемых стохастическими дифференциальными уравнениями в векторно-матричной форме (2.48)

$$\frac{dx(t)}{dt} = F_t x(t) + \Gamma_t \eta(t), \quad x(t_0) = x(0). \quad (3.41)$$

Применяя операцию математического ожидания к уравнению (3.41) с учетом характеристик гауссовского белого шума (раздел 2.2) приходим к дифференциальному уравнению вектора математического ожидания

$$\frac{dm_x(t)}{dt} = F_t m_x(t), \quad m_x(t_0) = m_x(0). \quad (3.42)$$

Почленно вычитая уравнение (3.42) из уравнения (3.41), получаем линейное уравнение относительно центрированной составляющей вектора состояния $x^0(t) = x(t) - m_x(t)$

$$\frac{dx^0(t)}{dt} = F_t x^0(t) + \Gamma_t \eta(t). \quad (3.43)$$

Определим дифференциальное уравнение для матрицы дисперсии вектора состояния системы из выражения

$$D(t) = k_x(t, t) = M\{x^0(t)x^{0T}(t)\}, \quad (3.44)$$

где $x^{0T}(t)$ - транспонированный вектор, удовлетворяющий сопряженному уравнению

$$\frac{dx^{0T}(t)}{dt} = x^{0T}(t)F_t^T + \eta^T(t)\Gamma_t^T. \quad (3.45)$$

Дифференцируя выражение (3.44) по времени, а также учитывая, что операции определения математического ожидания и дифференцирования линейны и независимы, получаем

$$\frac{dD(t)}{dt} = M\left\{\frac{dx^0(t)}{dt}x^{0T}(t)\right\} + M\left\{x^0(t)\frac{dx^{0T}(t)}{dt}\right\}. \quad (3.46)$$

Подставляя в соотношение (3.46) выражения для производных (3.43) и (3.45), имеем

$$\frac{dD(t)}{dt} = F_t D(t) + D(t) F_t^T + M\{\Gamma_t \eta(t) x^{0T}(t)\} + M\{x^0(t) \eta^T(t) \Gamma_t^T\}. \quad (3.47)$$

Математическое ожидание выражений в фигурных скобках соотношения (3.47) определяется на основании формул (3.12) и (3.13) при $\xi(t) = \eta(t)$ с учетом значения корреляционной функции $k_\eta(t, u) = \mathcal{N}_\eta(t) \delta(t - u)$. Принимая во внимание фильтрующие свойства дельта-функции, имеем

$$\begin{aligned} M\{\Gamma_t \eta(t) x^{0T}(t)\} &= M\left\{\int_{t_0}^t \Gamma_t \eta(t) \eta^T(u) \Gamma_u^T \Phi_0^T(t, u) du\right\} = \\ &= \int_{t_0}^t \Gamma_t \mathcal{N}_{\eta t} \delta(t - u) \Gamma_u^T \Phi_0^T(t, u) du = \frac{1}{2} \Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T \Phi_0^T(t, t) = \\ &= \frac{1}{2} \Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T. \end{aligned} \quad (3.48)$$

Здесь учтено, что матрица $\Phi_0(t, t)$ является единичной, а половинный множитель появился вследствие того, что интеграл от дельта-функции при аргументе равном верхнему значению интеграла, равен 0,5.

Аналогичным путем было определено соотношение

$$\begin{aligned} M\{x^0(t) \eta^T(t) \Gamma_t^T\} &= M\left\{\int_{t_0}^t \Phi_0(t, u) \Gamma_u \eta(u) \eta^T(t) \Gamma_t^T du\right\} = \\ &= \frac{1}{2} \Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T. \end{aligned} \quad (3.49)$$

Подставляя полученные выражения (3.48) и (3.49) в (3.47), приходим к уравнению для матрицы дисперсии

$$\frac{dD(t)}{dt} = F_t D(t) + D(t) F_t^T + \Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T. \quad (3.50)$$

Уравнения (3.42) и (3.50) описывают эволюцию во времени вектора математического ожидания и матрицы дисперсии вектора состояния системы. Обратим внимание на то, что для линейных систем, во-первых, уравнения определенного порядка содержат соответствующие статистические характеристики данного порядка, во-вторых, уравнения (3.42) и (3.50) не связаны между собой и каждое решается независимо друг от друга.

В координатно- скалярном представлении для уравнения (3.42) и (3.50) можно записать

$$\begin{aligned}\frac{dm_{x_i}(t)}{dt} &= \sum_{j=1}^r f_{ij} m_{x_j}(t), \quad m_{x_i}(t_0) = m_{x_i}(0), \\ \frac{dD_{ij}(t)}{dt} &= \sum_{k=1}^r [f_{ik} D_{kj}(t) + D_{ik}(k) f_{jk}] + \sum_{l=1}^r \gamma_{il} \gamma_{jl} \mathcal{N}_{\eta ll}, \quad D_{ij}(t_0) = D(0).\end{aligned}$$

Для стационарной системы соответствующие статистические характеристики определяются из уравнений

$$\begin{aligned}\frac{dm_x(t)}{dt} &= F m_x(t), & m_x(t_0) &= m_x(0), \\ \frac{dD(t)}{dt} &= F D(t) + D(t) F^T + \Gamma \mathcal{N}_{\eta} \Gamma^T, & D(t_0) &= D(0).\end{aligned}$$

Одним из распространенных режимов стационарных устойчивых систем является установившийся режим при воздействии стационарного сигнала $\eta(t)$. Соответствующие статистические характеристики на основании выражений (3.42) и (3.50) определяются из соотношений

$$\begin{aligned}m_x &= 0, \\ F D + D F^T + \Gamma \mathcal{N}_{\eta} \Gamma &= 0.\end{aligned}$$

Для дискретных линейных систем, динамика которых описывается уравнением (2.55) после выполнения соответствующих преобразований уравнения для вектора математического ожидания и матрицы дисперсии принимают вид

$$m_{x,n+1} = \Phi_n m_{x,n}, \quad m_{x0} = m_x(0), \quad (3.51)$$

$$D_{n+1} = \Phi_n D_n \Phi_n^T + \Gamma_n Q_{\eta n} \Gamma_n^T, \quad D_0 = D(0), \quad (3.52)$$

где $\Phi_n = 1 + \Delta t F_n$, $Q_{\eta n} = \mathcal{N}_{\eta, t=\Delta t} / \Delta t$.

В некоторых практических задачах корреляционного анализа помимо значений вектора математического ожидания и матрицы дисперсий необходимы более полные вероятностные характеристики, которые даются матрицей корреляционных функций. Для ее определения воспользуемся уравнением состояния линейной динамической системы (3.41), в правой части которого в качестве воздействующего сигнала используется стационарный гауссовский белый шум. Для центрированной составляющей это уравнение имеет вид (3.43). По определению матрица корреляционных функций равняется

$$\begin{aligned} k_x(t_1, t_2) &= M\{x^0(t_2)x^{0T}(t_1)\} = M\{M\{x^0(t_2)|x^0(t_1)\}x^{0T}(t_1)\} = \\ &= M\{\hat{x}^0(t_2)x^{0T}(t_1)\}, \end{aligned} \quad (3.53)$$

где $\hat{x}^0(t_2) = M\{x^0(t_2)|x^0(t_1)\}$ - условное математическое ожидание. Оно определяется в результате усреднения уравнения (3.43) относительно условной плотности вероятностей $p(x, t_2|x, t_1)$ из соотношения

$$\frac{d\hat{x}^0(t_2)}{dt_2} = F_{t_2} \hat{x}^0(t_2), \quad \hat{x}^0(t_1) = x^0(t_1). \quad (3.54)$$

Решение его запишем в форме

$$\hat{x}^0(t_2) = \Phi_0(t_2, t_1)x^0(t_1), \quad (3.55)$$

где $\Phi_0(t_2, t_1)$ - матрица весовых коэффициентов, являющаяся решением уравнения (3.14).

Подставляя (3.55) в выражение (3.53), а также учитывая (3.44), имеем

$$\begin{aligned} k_x(t_1, t_2) &= M\{\hat{x}^0(t_2)x^{0T}(t_1)\} = M\{\Phi_0(t_2, t_1)x^0(t_1)x^{0T}(t_1)\} = \\ &= \Phi_0(t_2, t_1)D(t_1). \end{aligned}$$

Аналогичным образом после соответствующих преобразований можно записать для случая $t_1 \geq t_2$ уравнение

$$\hat{x}^{0T}(t_1) = x^{0T}(t_2)\Phi_0^T(t_1, t_2)$$

и

$$k_x(t_2, t_1) = D(t_2)\Phi_0^T(t_1, t_2).$$

Таким образом, для всей области изменений t_1 и t_2 матрица корреляционных функций записывается в форме

$$k_x(t_1, t_2) = \begin{cases} \Phi_0(t_2, t_1)D(t_1), & t_2 \geq t_1, \\ D(t_2)\Phi_0^T(t_1, t_2), & t_1 \geq t_2. \end{cases} \quad (3.56)$$

При $t_1 = t_2 = t$ имеем $k_x(t, t) = D(t)$, поскольку $\Phi(t, t) = \Phi^T(t, t) = I(t)$.

Как следует из полученных выражений, для определения матрицы корреляционных функций вектора состояния системы необходимо сначала определить матрицу дисперсий, а затем умножить ее соответственно при $t_2 \geq t_1$ слева, а при $t_1 \geq t_2$ справа на матрицу весовых коэффициентов.

3.3.4.2. ВОЗДЕЙСТВУЮЩИЙ ПРОЦЕСС – ПУАССОНОВСКИЙ ИМПУЛЬСНЫЙ ПОТОК

Исходным при определении соответствующих статистических характеристик является дифференциальное уравнение вектора состояния системы, описывающего многомерный марковский процесс в форме (2.96)

$$\frac{dx(t)}{dt} = F_t x(t) + \Gamma_t \eta^N(t), \quad x(t_0) = x(0). \quad (3.57)$$

Применяя операцию математического ожидания, а также учитывая рассмотренные в разделе 2.4.3 характеристики пуассоновского процесса, определяем дифференциальное уравнение вектора математического ожидания размера $(r \times 1)$

$$\frac{dm_x(t)}{dt} = F_t m_x(t) + \Gamma_t \nu(t), \quad m_x(t_0) = m_x(0), \quad (3.58)$$

где $\nu(t)$ - вектор интенсивностей импульсного потока размера $(r \times 1)$.

Почленно вычитая из уравнения (3.57) уравнение (3.58), получаем линейное уравнение относительно центрированных составляющих $x^0(t) = x(t) - m_x(t)$, $\eta^{N0}(t) = \eta^N(t) - \nu(t)$:

$$\frac{dx^0(t)}{dt} = F_t x^0(t) + \Gamma_t \eta^{N0}(t), \quad (3.59)$$

Определим дифференциальное уравнение для матрицы дисперсии вектора состояния $x(t)$ из выражения

$$D(t) = k_x(t, t) = M\{x^0(t)x^{0T}(t)\}, \quad (3.60)$$

где $x^{0T}(t)$ - транспонированный вектор $x^0(t)$, удовлетворяющий сопряженному уравнению

$$\frac{dx^{0T}(t)}{dt} = x^{0T}(t)F_t^T + (\eta^{N0}(t))^T \Gamma_t^T. \quad (3.61)$$

Дифференцируя выражение (3.60) по времени, а также учитывая, что операции определения математического ожидания и дифференцирования линейны и независимы, получаем

$$\frac{dD(t)}{dt} = M\left\{\frac{dx^0(t)}{dt}x^{0T}(t)\right\} + M\left\{x^0(t)\frac{dx^{0T}(t)}{dt}\right\}. \quad (3.62)$$

Подставляя в соотношение (3.62) выражения для производных (3.59) и (3.61), имеем

$$\frac{dD(t)}{dt} = F_t D(t) + D(t)F_t^T + M\{\Gamma_t \eta^{N0}(t)x^{0T}(t)\} + M\{x^0(t)(\eta^{N0}(t))^T \Gamma_t^T\}.$$

Математические ожидания выражений в фигурных скобках определяются на основании формул (3.12), (3.13) и (3.60) при воздействии пуассоновского импульсного потока с известной корреляционной диагональной матрицей

$$k_\eta(t, u) = M\{\eta^{N0}(t)(\eta^{N0}(u))^T\} = \nu(t)\delta(t-u).$$

Дальнейшие преобразования, аналогичные ранее рассмотренным в разделе 3.3.4.1, приводят к дифференциальному уравнению матрицы дисперсии

$$\frac{dD(t)}{dt} = F_t D(t) + D(t) F_t^T + \Gamma_t \nu(t) \Gamma_t^T, \quad D(t_0) = D(0). \quad (3.63)$$

или в координатно-скалярном представлении

$$\frac{dm_{x_i}(t)}{dt} = \sum_{j=1}^r f_{ij} m_{x_j}(t) + \sum_{l=1}^r \gamma_{il} \nu_l, \quad m_{x_i}(t_0) = m_{x_i}(0),$$

$$\frac{dD_{ij}(t)}{dt} = \sum_k^r [f_{ik} D_{kj}(t) + D_{ik}(t) f_{jk}] + \sum_{l=1}^r \gamma_{il} \gamma_{jl} \nu_l, \quad D_{ij}(t_0) = D_{ij}(0).$$

Уравнения (3.58) и (3.63) описывают эволюцию во времени вектора математического ожидания и матрицы дисперсии линейной динамической системы при воздействии пуассоновского импульсного потока. Как следует из вида этих уравнений, они не связаны между собой и решаются отдельно.

Обратим внимание на структурную идентичность уравнений (3.50) и (3.63) для рассматриваемых типов воздействий: белого шума и случайного импульсного потока. Хотя физическая картина поведения реализаций этих процессов различна: в первом случае непрерывный, во втором - импульсный процесс. Неполнота описания векторного процесса системы только первыми двумя статистическими характеристиками (в рамках корреляционной теории) нивелирует эти процессы. Для более точной их идентификации необходимо учитывать статистические характеристики третьего и более высоких порядков.

Для стационарных линейных систем соответствующие статистические характеристики определяются из уравнений

$$\begin{aligned} \frac{dm_x(t)}{dt} &= F m_x(t) + \Gamma \nu, & m_x(t_0) &= m_x(0), \\ \frac{dD(t)}{dt} &= F D(t) + D(t) F^T + \Gamma \mathbf{V} \Gamma^T, & D(t_0) &= D(0). \end{aligned}$$

В случае дискретных линейных систем, динамика которых описывается уравнением (2.96), после выполнения соответствующих преобразований, рекуррентные уравнения вектора математического ожидания и матрицы дисперсий принимают вид

$$\begin{aligned} m_{x,n+1} &= \Phi_n m_{xn} + \Gamma_n v_n, & m_{x0} &= m_x(0), \\ D_{n+1} &= \Phi_n D_n \Phi_n^T + \Gamma_n Q_n \Gamma_n^T, & D_0 &= D(0), \end{aligned}$$

где $\Phi_n = 1 + \Delta t F_n$, $Q_n = v_{t=n\Delta t} / \Delta t$.

Приведем без вывода формулу матрицы корреляционных функций вектора состояния системы

$$k_x(t_1, t_2) = \begin{cases} \Phi_0(t_2, t_1) D(t_1), & t_2 \geq t_1, \\ D(t_2) \Phi_0^T(t_1, t_2), & t_1 \geq t_2. \end{cases}$$

где $\Phi_0(t_2, t_1)$ - матрица весовых коэффициентов, $D(t)$ - матрица дисперсий.

3.4. СЛУЧАЙНЫЕ ПРОЦЕССЫ В НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМАХ

3.4.1. ОБЗОР МЕТОДОВ

Исследование нелинейных систем, находящихся под воздействием случайных сигналов, обычно является достаточно сложной задачей. Решение описывающих эти системы нелинейных дифференциальных уравнений зависит от вида внешних воздействий, начальных условий, типа нелинейностей и в отличие от линейной системы не может быть записано в общем виде. Остановимся на некоторых наиболее важных в практическом отношении задачах анализа таких систем. Методы исследования нелинейных систем могут быть разделены на точные и приближенные. К точным можно отнести методы, позволяющие получить характеристики, полностью описывающие случайный процесс (n -мерные плотности вероятностей, моментные, корреляционные функции и т.д.). К ним относятся, например, методы с использованием функциональных рядов Вольтерра [25], канонических разложений [26], основанный на интегрировании уравнений Фоккера-Планка-Колмогорова.

Метод с использованием рядов Вольтерра основан на представлении решения нелинейного дифференциального уравнения (2.42) в виде функционального ряда

$$x(t) = C_0 + \sum_{i=1}^n C_i \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g_i(t, \tau_1, \dots, \tau_i) \xi(\tau_1) \dots \xi(\tau_i) d\tau_1 \dots d\tau_i,$$

где функции $g_i(\cdot)$ - ядра нелинейной системы.

При этом предполагается, что функция $f(\cdot)$ аппроксимируется рядом Тейлора, а функции $g_i(\cdot)$ не зависят от процесса $x(t)$.

Первый член ряда (не считая постоянной составляющей) характеризует линейную часть системы. Второй - операцию возведения в квадрат входного воздействия, третий - возведение в куб и т.д. Вычисления корреляционных функций первого (математического ожидания) и более высоких порядков производятся по известным правилам и является весьма трудной задачей. Облегчить задачу может ортогональное разложение ядер высокого порядка нелинейной системы.

В методе канонических разложений используется представление двумерной плотности вероятностей в виде двойного ряда по ортогональным полиномам:

$$p(x_1, t_1; x_2, t_2) = p(x_1, t_1) p(x_2, t_2) \sum_{n,m=1} a_{nm}(t_1, t_2) Q_n(x_1, t_1) Q_m(x_2, t_2),$$

где $p(x, t)$ - одномерная плотность вероятностей $Q_n(x, t)$, $Q_m(x, t)$ - полиномы, удовлетворяющие условию ортогональности

$$\int_{-\infty}^{\infty} p(x, t) Q_n(x, t) Q_m(x, t) dx = \begin{cases} 1, & m = n, \\ 0, & m \neq n. \end{cases}$$

Коэффициенты $a_{nm}(t_1, t_2)$ можно заранее вычислить из формул

$$a_{nm} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, t_1; x_2, t_2) Q_n(x_1, t_1) Q_m(x_2, t_2) dx_1 dx_2.$$

При использовании уравнения ФПК выходной случайный процесс полагают марковским. Решение этого уравнения проводится численными методами и, как правило, ограничивается нахождением

одномерной, в некоторых случаях - двумерной плотности вероятностей. Точное аналитическое решение можно получить для ряда частных задач. Так, в методе, рассмотренном в [27], вводятся новые переменные, которые позволяют свести уравнение ФПК к уравнению, решение которого известно. Основная трудность этого метода заключается в выборе функции, связывающей старые и новые переменные. С наиболее простым решением сталкиваются в стационарном режиме, когда при $t \rightarrow \infty$ $p(x,t)$ стремится к стационарной плотности $p(x)$, не зависящей от начальных условий. Уравнение (2.38) в этом случае преобразуется к виду

$$d\{K_2(x)p(x)\}/dx - 2K_1(x)p(x) = 0. \quad (3.64)$$

Решением уравнения (3.64) при нулевых граничных условиях является выражение

$$p(x) = \frac{C}{K_2(x)} \exp \left\{ 2 \int_0^x \frac{K_1(\lambda)}{K_2(\lambda)} d\lambda \right\},$$

где постоянная C определяется из условия нормировки

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

Например, для линейной устойчивой системы, описываемой уравнением (2.46), стационарная плотность определяется из уравнения (3.64) и имеет гауссовский вид

$$p(x) = (2\pi D)^{-1/2} \exp \left\{ -\frac{x^2}{2D_x} \right\},$$

где $D_x = \alpha N_r / 2$.

Ввиду того, что применение точных методов исследований нелинейных систем связано с большими трудностями, большое распространение получили приближенные методы, в частности, основанные на линеаризации нелинейной функции. Простейшим видом линеаризации при условии малой величины дисперсии центрированной составляющей случайного сигнала является линеаризация нелинейной дифференцируемой функции путем замены ее линейной зависимостью (так называемая локальная аппроксимация

или линеаризация). Пусть нелинейность описываемая однозначной безынерционной нечетной функцией $f_t(x)$, где $x(t)=m_x(t)+x^0(t)$ - случайный сигнал, центрированная составляющая которого $x^0(t)$ имеет малую дисперсию $M\{[x^0(t)]^2\}$. Это означает, что случайный сигнал, в основном, незначительно отклоняется от своего математического ожидания $m_x(t)$. Локальная линеаризация состоит в замене нелинейности приближенной линеаризованной зависимостью, определяемой первыми двумя членами разложения функции $f_t(x)$ в ряд Тейлора в точке $m_x(t)$ относительно центрированного сигнала $x^0(t)$

$$y(t) = f_t(x) \approx f_t(m_x) + \frac{\partial f_t(m_x)}{\partial m_x} x^0(t). \quad (3.65)$$

Приближенная зависимость линейна относительно $x^0(t)$ и нелинейна относительно $m_x(t)$ и заменяет кривую $y(t)=f_t(x)$ касательной в окрестности подвижной точки $m_x(t)$, т.е. непосредственно (локально) вблизи этой точки. Погрешность зависимости (3.65) оценивается по статистическим характеристикам (математическому ожиданию и дисперсии) отброшенных членов разложения в ряд Тейлора.

В случае многомерной нелинейной функции $y(t)=f_t(x)$ линеаризованная зависимость записывается в виде

$$y(t) \approx f_t(m_x) + \frac{\partial f_t(m_x)}{\partial m_x} x^0(t).$$

где $y(t)$, x^0 и $f_t(m_x)$ - r -мерные векторы-столбцы соответственно выходного, входного сигналов и нелинейного преобразования вектора математического ожидания $m_x(t)$; $\frac{\partial f_t(m_x)}{\partial m_x}$ - матрица Якоби с компонентами $\frac{\partial f_{ti}(m_x)}{\partial m_{x_j}}$, $i, j = \overline{1, r}$, i - номер строки; j - номер столбца.

Для определения статистических характеристик процесса на выходе нелинейности при локальной линеаризации может быть использован рассмотренный в предыдущем разделе математический аппарат линейного корреляционного анализа.

Если дисперсия центрированной составляющей случайного сигнала на входе нелинейности большая или в системе содержатся

элементы с существенно нелинейными характеристиками (релейные элементы, элементы с ограниченными зонами линейности, нечувствительности и т.д.), то при исследовании процессов применяют методы статистической линеаризации.

3.4.2. МЕТОДЫ СТАТИСТИЧЕСКОЙ ЛИНЕАРИЗАЦИИ

Локальная линеаризация неприменима к недифференцируемым нелинейностям, а также тогда, когда изменение входного сигнала превышает зону, в которой нелинейность может быть заменена линейной зависимостью. Для указанных случаев не существует достаточно простой связи между входными и выходными сигналами и соответственно между входными и выходными статистическими характеристиками. Однако при статистическом анализе системы можно построить относительно простую приближенную линеаризованную зависимость, статистически эквивалентную исходной нелинейности. Указанная линеаризация нелинейности получила название статистической линеаризации. Метод статистической линеаризации состоит в замене безынерционной нелинейной функции

$$y=f(x) \quad (3.66)$$

эквивалентной в вероятностном смысле линеаризованной зависимостью между входным и выходными сигналами. Указанная задача сводится к нахождению наилучшего в рамках корреляционной теории приближения к нелинейному преобразованию. Поэтому метод статистической линеаризации позволяет учесть основные закономерности нелинейного преобразования в рамках первых двух вероятностных моментов случайного процесса

$$m_y = \int_{-\infty}^{\infty} f(x)p(x)dx, \quad (3.67)$$

$$D_y = \int_{-\infty}^{\infty} f^2(x)p(x)dx - m_y^2. \quad (3.68)$$

Представим однозначную нелинейную функцию (3.66) приближенным соотношением

$$z=f_0+Kx^0. \quad (3.69)$$

где $x^0=x-m_x$ - входной центрированный случайный сигнал.

Входящими в это соотношение параметрами являются статистическая характеристика нелинейности $-f_0$ и статистический коэффициент усиления $-K$. Указанные параметры - неслучайные нелинейные функции статических характеристик входного сигнала: m_x и D_x .

В качестве критерия эквивалентности принимается выполнение равенства $m_y=m_z$. Отсюда имеем

$$f_0 = f_0(m_x, D_x) = \int_{-\infty}^{\infty} f(x) p(x) dx.$$

При рассмотрении критериев для определения статистических коэффициентов усиления остановимся на двух наиболее известных способах. Согласно первому способу требуется выполнение равенства дисперсий $D_z=D_y$. Так как из эквивалентной зависимости (3.69) следует $D_z = K_1^2 D_x$, то коэффициент K_1 с учетом (3.69) определяется из соотношения

$$K_1 = K_1(m_x, D_x) = \left[\frac{1}{D_x} \left(\int_{-\infty}^{\infty} f^2(x) p(x) dx - m_y^2 \right) \right]^{1/2}. \quad (3.70)$$

При втором способе коэффициент K_2 определяется из условия минимума среднеквадратической ошибки аппроксимации

$$M\{\varepsilon^2\} = \min_{K_2} M\{(y-z)^2\}.$$

Принимая во внимание $m_z=m_y$, а также $z^0=K_2x^0$ получаем

$$y-z=m_y+y^0-m_z-z^0=y^0-K_2x^0 \text{ и} \\ M\{\varepsilon^2\} = M\{(y-z)^2\} = D_y - 2K_2D_{xy} + K_2^2D_x,$$

где $D_{xy}=M\{x^0y^0\}$.

Необходимым условием экстремума $M\{\varepsilon^2\}$ по K_2 является $\partial M\{\varepsilon^2\} / \partial K_2 = -2D_{xy} + 2K_2D_x = 0$, отсюда $K_2=D_{xy}/D_x$.

Поскольку

$$D_{xy} = M\{(f(x) - m_y)x^0\} = \int_{-\infty}^{\infty} f(x)(x - m_x)p(x)dx,$$

то статистический коэффициент усиления по второму способу принимает вид

$$K_2 = K_2(m_x, D_x) = \frac{M\{f(x)(x - m_x)\}}{D_x} = \frac{\int_{-\infty}^{\infty} f(x)(x - m_x)p(x)dx}{D_x}. \quad (3.71)$$

Соотношения (3.69) и (3.70) или (3.71) позволяют представить нелинейный элемент приближенно в виде линейного элемента

$$f(x) \approx f_0 + K(m_x, D_x)(x - m_x) = K_0(m_x, D_x)m_x + K(m_x, D_x)(x - m_x),$$

если известна одномерная плотность вероятностей $p(x)$. Здесь $K_0(m_x, D_x)$ - статистический коэффициент усиления по математическому ожиданию. Значение коэффициентов K_0 и K (K_1 или K_2) для различных типов нелинейностей приведена в соответствующих руководствах [28,29].

В замкнутых динамических системах, содержащих нелинейный элемент, истинное значение плотности вероятностей сигнала на входе этого элемента неизвестно. В методе статистической линеаризации эта плотность приближенно принимается гауссовской. В этом и заключается приближенность указанного метода. Рассматриваемое допущение базируется на свойстве сложных систем, содержащих линейные инерционные элементы, нормализовать законы распределения случайных сигналов, т.е. приближать эти сигналы к гауссовским процессам. Это означает, что плотность вероятностей входного сигнала нелинейного элемента, находящегося между линейными инерционными элементами в замкнутой системе, можно приближенно считать гауссовской. Если в формулах (3.67), (3.70) и (3.71) плотность вероятностей $p(x)$ предположить гауссовской, то входящие в них интегралы могут быть сведены к интегралам вероятностей, которые вычисляются или определяются из таблиц, как функции двух переменных m_x и D_x . В случае гауссовской плотности

$$p(x) = (2\pi D_x)^{-1/2} \exp\left\{-\frac{(x-m_x)^2}{2D_x}\right\}$$

определение статистического коэффициента усиления по второму способу упрощается и он вычисляется из формулы

$$\begin{aligned} K(m_x, D_x) &= \frac{\partial f_0(m_x, D_x)}{\partial m_x} = \frac{\partial}{\partial m_x} \left\{ \int_{-\infty}^{\infty} f(x) p(x) dx \right\} = \\ &= \frac{1}{D_x} \int_{-\infty}^{\infty} f(x)(x-m_x) p(x) dx \end{aligned} \quad (3.72)$$

Обобщим полученные результаты на многомерные процессы. При статистической линеаризации по второму способу многомерная безынерционная однозначная нелинейная функция $f(x)$ аппроксимируется выражением $f_0 + K[x-m_x]$, где $x = \|x_1 \dots x_r\|^T$ - вектор-столбец на входах нелинейностей, $f_0(m_x, D_x) = M\{f(x)\}$ - векторная статистическая характеристика; $m_x = M\{\|x_1 \dots x_r\|^T\}$ - вектор-столбец математических ожиданий входных процессов; $D_x = M\{(x-m_x)(x-m_x)^T\}$ - матрица дисперсий (m_x и D_x - параметры многомерной гауссовской плотности); $K = K(m_x, D_x)$ - матрица статистических коэффициентов усиления, компоненты которой K_{ij} определяется на основании критерия минимума среднеквадратической ошибки аппроксимации для несвязанных переменных вектора x из уравнений

$$\sum_{j=1}^r K_{ij} D_{jl} = D_{fl}, \quad i, l = \overline{1, r},$$

где $D_{jl} = M\{(x_j - m_{x_j})(x_l - m_{x_l})\}$, $D_{ff} = M\{f_i(x)(x_l - m_{x_l})\}$, $f_i(x)$ - компонента вектор-функции $f(x)$, которая аппроксимируется выражением

$$f_{0i} + \sum_{j=1}^r K_{ij}(x_j - m_{x_j}), \quad K_{ij} = \frac{\partial f_{0i}(m_x, D_x)}{\partial m_{x_j}}.$$

Оба рассмотренных способа статистической линеаризации по методической погрешности практически равноценны. Значения

статистических характеристик нелинейностей по этим способам одинаковы, и различия имеют место при определении статистических коэффициентов усиления. При втором способе процедура вычисления проще. С другой стороны, эти значения для ряда конкретных нелинейностей оказывается заниженным. В этом случае целесообразно использовать среднеарифметическое из значений этих коэффициентов по каждому из способов.

3.4.3. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК

3.4.3.1. ВОЗДЕЙСТВУЮЩИЙ ПРОЦЕСС – БЕЛЫЙ ШУМ

Динамическая система описывается в пространстве состояний системой дифференциальных уравнений в векторной форме (2.50). В результате статистической линеаризации нелинейности $f_t(x)$ по известной методике раздела 3.4.2 приходим к линеаризованному уравнению

$$\frac{dx(t)}{dt} = f_{0t}(m_x, D_x) + K_t(m_x, D_x)x^0(t) + \Gamma_t\eta(t), \quad x(t_0) = x(0), \quad (3.73)$$

где $f_{0t}(\cdot)$ - векторная статистическая характеристика, $K_t(\cdot)$ - матрица статистических коэффициентов усиления, зависящих от вектора математического ожидания - $m_x(t)$ и матрицы дисперсий $D_x(t)$.

Для дискретного времени на основании уравнения (3.73) можно получить линеаризованное уравнение в конечных разностях

$$x_{n+1} = x_n + f_{0n}(m_x, D_x)\Delta t + K_n(m_x, D_x)x_n^0\Delta t + \Gamma_n\eta_n, \quad x_0 = x(0), \quad (3.74)$$

где Γ_n - матрица воздействия размера $(r \times r)$; η_n - дискретный белый шум с математическим ожиданием равным нулю и корреляционной матрицей $M\{\eta_n, \eta_l\} = Q_{\eta n} \delta_{nl}$, δ_{nl} - символ Кронекера, $Q_{\eta n} = N_{\eta t=n\Delta t}/\Delta t$, $N_{\eta t}$ - диагональная положительная матрица спектральных плотностей белого шума $\eta(t)$.

В результате статистического осреднения уравнения (3.73) получаем дифференциальное уравнение для вектора математического ожидания

$$\frac{dm_x(t)}{dt} = f_{0t}(m_x, D_x), \quad m_x(t_0) = m_x(0). \quad (3.75)$$

Используя рассмотренную в разделе 3.4.2 методику, сначала определяют дифференциальное уравнение относительно центрированной составляющей $x^0(t)$

$$\frac{dx^0(t)}{dt} = K_t(m_x, D_x)x^0(t) + \Gamma_t \eta(t).$$

На основании выражения для матрицы дисперсий (3.44), где транспонированный вектор $x^{0T}(t)$ определяется из сопряженного уравнения

$$\frac{dx^{0T}(t)}{dt} = x^{0T}(t)K_t^T(m_x, D_x) + \eta^T \Gamma_t^T,$$

после дифференцирования приходят к соотношению

$$\begin{aligned} \frac{dD_x(t)}{dt} = & K_t(m_x, D_x)D_{xt} + D_{xt}K_t^T(m_x, D_x) + \\ & + M\{\Gamma_t \eta(t)x^0(t)\} + M\{x^0(t)\eta^T(t)\Gamma_t^T\}, \end{aligned}$$

$$\text{где } M\{\Gamma_t \eta(t)x^{0T}\} = M\left\{\int_{t_0}^t \Gamma_t \eta(t)\eta^T(t)\Gamma_u^T \Phi_0^T(t, u)du\right\} = \frac{1}{2}\Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T,$$

$$M\{x^0(t)\eta^T(t)\Gamma_t^T\} = M\left\{\int_{t_0}^t \Phi_0(t, u)\Gamma_u \eta(u)\eta^T(t)\Gamma_t^T du\right\} = \frac{1}{2}\Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T,$$

$$K_\eta(t, u) = \mathcal{N}_{\eta t} \delta(t - u), \quad k_\eta(t, u) = \mathcal{N}_{\eta t} \delta(u - t).$$

Здесь $\Phi_0(t, u)$ и $\Phi_0^T(u, t)$ матрицы весовых коэффициентов, являющиеся решением соответствующих линеаризованных уравнений

$$\frac{d\Phi_0(t, u)}{dt} = K_t(m_x, D_x)\Phi_0(t, u) + I_\delta(t - u), \quad (3.76)$$

$$\frac{d\Phi_0^T(u, t)}{du} = \Phi_0^T(u, t) K_u^T(m_x, D_x) + I_\delta(u - t), \quad (3.77)$$

где $I_\delta(t-u)$ и $I_\delta(u-t)$ - диагональные матрицы дельта-функций.

В результате линеаризованное дифференциальное уравнение матрицы дисперсии вектора состояния нелинейной системы принимает вид

$$\begin{aligned} \frac{dD_x(t)}{dt} = & K_t(m_x, D_x)D_x(t) + \\ & + D_x(t)K_t^T(m_x, D_x) + \Gamma_t \mathcal{N}_{\eta\eta} \Gamma_t^T, \quad D_x(t_0) = D_x(0). \end{aligned} \quad (3.78)$$

В координатно-скалярном представлении уравнения (3.75) и (3.78) имеют вид

$$\begin{aligned} \frac{dm_{x_i}(t)}{dt} = & f_{0i}(m_x, D_x), \quad m_{x_i}(t_0) = m_{x_i}(0), \\ \frac{dD_{xij}(t)}{dt} = & \sum_{k=1}^r [K_{ik}(m_x, D_x)D_{kj}(t) + D_{ik}(t)K_{jk}(m_x, D_x)] + \sum_{l=1}^r \gamma_{il}\gamma_{jl}\mathcal{N}_{\eta ll}, \\ D_{ij}(t_0) = & D_{ij}(0), \quad i, j = \overline{1, r}. \end{aligned}$$

В отличие от линейной системы уравнения (3.75) и (3.78) связаны между собой и их необходимо решать совместно. Обратим внимание на то, что эти уравнения линейные по форме являются по существу нелинейными, так как параметры f_{0t} и K_t зависят от математического ожидания и дисперсии вектора состояния системы нелинейно. В этом проявляется сущность метода статистической линеаризации, сохраняющая специфические особенности исходной нелинейной системы, в том числе и несоблюдение принципа суперпозиции. Как следует из структуры полученных уравнений метод статистической линеаризации обеспечивает единообразный алгоритм анализа линейных и нелинейных систем.

Для дискретных нелинейных систем, динамика которых описывается линеаризованным уравнением (3.74) после выполнения

соответствующих преобразований можно определить рекуррентные уравнения вектора математических ожиданий и матрицы дисперсий

$$m_{x,n+1} = m_{xn} + f_{0n}(m_x, D_x)\Delta t, \quad m_{x0} = m_x(0), \quad (3.79)$$

$$D_{x,n+1} = D_{xn} + K_n(m_x, D_x)D_{xn}\Delta t + D_{xn}K_n^T(m_x, D_x)\Delta t + \Gamma_n Q_{\eta n} \Gamma_n^T, \quad (3.80)$$

$$D_{x0} = D_x(0).$$

Корреляционная функция по форме аналогична соответствующим выражениям, полученным для линейной системы, и имеет вид

$$k_x(t_1, t_2) = \begin{cases} \Phi_0(t_2, t_1) D_x(t_1), & t_2 \geq t_1, \\ D_x(t_2) \Phi_0^T(t_1, t_2), & t_1 \geq t_2. \end{cases} \quad (3.81)$$

Особенность их определения состоит в том, что дисперсия $D_x(t)$ и матрицы $\Phi_0(t_2, t_1)$ и $\Phi_0^T(t_1, t_2)$ являются решениями линеаризованных уравнений (3.78) и (3.76) и (3.77). Корреляционная функция симметрична относительно своих аргументов. Поэтому достаточно рассчитать верхнее или нижнее произведение выражений (3.81).

3.4.3.2. Воздействующий процесс – пуассоновский импульсный поток

Динамическая система описывается в пространстве состояний системой дифференциальных уравнений в векторной форме (2.95). В результате статистической линеаризации нелинейности приходят к линеаризованному уравнению

$$\frac{dx(t)}{dt} = f_{0t}(m_x, D_x) + K_t(m_x, D_x)x^0(t) + \Gamma_t \eta_t^N, \quad x(t_0) = x(0). \quad (3.82)$$

После статистического осреднения уравнения (3.82) дифференциальное уравнение вектора математического ожидания принимает вид

$$\frac{dm_x(t)}{dt} = f_{0t}(m_x, D_x) + \Gamma_t \nu(t), \quad m_x(t_0) = m_x(0) \quad (3.83)$$

где $\mathbf{v}(t)$ - вектор интенсивностей импульсного потока размера $(r \times 1)$.

Исходным при определении уравнения для матрицы дисперсии является выражение (3.46), где уравнения для центрированных составляющих векторов состояния системы имеют вид

$$\begin{aligned}\frac{dx^0(t)}{dt} &= K_t(m_x, D_x)x^0(t) + \Gamma_t \eta^{N0}(t), \\ \frac{dx^{0T}(t)}{dt} &= x^{0T}(t)K_t^T(m_x, D_x) + [\eta^{N0}(t)]^T \Gamma_t^T,\end{aligned}$$

Здесь $\eta^{N0}(t) = \eta^N(t) - \mathbf{v}(t)$ - вектор центрированного импульсного потока.

После дифференцирования (3.44) по времени и проведения известных из раздела 3.4.2 преобразований получаем дифференциальное уравнение матрицы дисперсии

$$\begin{aligned}\frac{dD_x(t)}{dt} &= K_t(m_x, D_x)D_x(t) + \\ &+ D_x(t)K_t^T(m_x, D_x) + \Gamma_t \mathbf{v}(t)\Gamma_t^T, \quad D_x(t_0) = D_x(0).\end{aligned}\tag{3.84}$$

где $\mathbf{v}(t)$ - положительная диагональная матрица интенсивностей импульсного потока размера $(r \times r)$.

В координатно-скалярном представлении уравнения (3.83) и (3.84) принимают вид

$$\begin{aligned}\frac{dm_{x_i}(t)}{dt} &= f_{0i}(m_x, D_x) + \sum_{l=1}^r \gamma_{il} v_l(t), \quad m_{x_i}(t_0) = m_{x_i}(0), \\ \frac{dD_{x,ij}(t)}{dt} &= \sum_{k=1}^r [K_{ik}(m_x, D_x)D_{kj}(t) + D_{ik}(t)K_{jk}(m_x, D_x)] + \sum_{l=1}^r \gamma_{il}\gamma_{jl}v_{ll}(t), \\ D_{x,ij}(t_0) &= D_{x,ij}(0).\end{aligned}$$

Для дискретных нелинейных систем после выполнения статистической линеаризации и соответствующих преобразований

рекуррентные уравнения вектора математического ожидания и матрицы дисперсий

$$m_{x,n+1} = m_{xn} + f_{0n}(m_x, D_x)\Delta t + \Gamma_n v_n, \quad m_x(t_0) = m_x(0), \quad (3.85)$$

$$D_{x,n+1} = D_{xn} + K_n(m_x, D_x)D_{xn}\Delta t + D_{xn}K_n^T(m_x, D_x)\Delta t + \Gamma_n Q_{vn}\Gamma_n^T, \quad (3.86)$$

$$D_{x0} = D_x(0).$$

где $Q_{vn} = v_{t=n\Delta t}/\Delta t$.

Выражение корреляционной функции по форме аналогично соотношениям (3.81). При этом необходимо обратить внимание на то, что матрицы $\Phi_0(t_2, t_1)$, $\Phi^T(t_1, t_2)$ и матрица дисперсии $D(t)$ являются решением линеаризованных уравнений (3.76), (3.77) и (3.84).

3.4.4. ПОГРЕШНОСТЬ СТАТИСТИЧЕСКОЙ ЛИНЕАРИЗАЦИИ

В практических задачах применения статистической линеаризации необходимо оценить погрешность получаемых решений. Известны задачи, где этот способ вносит существенные ошибки, например в безынерционных системах, в системах без обратной связи и с выраженными резонансными свойствами. В этих случаях на вход нелинейного элемента подается сигнал отличный от гауссовского процесса. По существу гауссовский или близкий к нему закон распределения входного процесса являются единственным ограничением приближенного метода статистической линеаризации.

Используя методику работы [30], оценим эту погрешность. Нелинейная система в пространстве состояний описывается уравнением в векторной форме

$$\frac{dx(t)}{dt} = f_t(x) + \Gamma_t \mu(t), \quad x(t_0) = x(0), \quad (3.87)$$

где $\mu(t)$ - обобщенный воздействующий процесс: пуассоновский импульсный поток или белый шум. В последнем случае $m_\mu = 0$.

После применения метода статистической линеаризации уравнение (3.87) заменяется на приближенное линеаризованное уравнение

$$\frac{dx(t)}{dt} = f_{0t}(m_x, D_x) + K_t(m_x, D_x)x^0(t) + \Gamma_t \mu(t), \quad x(t_0) = x(0). \quad (3.88)$$

Этому уравнению соответствуют уравнения вектора математического ожидания и матрицы дисперсии

$$\begin{aligned} \frac{dm_x(t)}{dt} &= f_{0t}(m_x, D_x) + m_\mu(t), \quad m_x(t_0) = m_x(0), \\ \frac{dD_x(t)}{dt} &= K_t(m_x, D_x)D_x(t) + \\ &+ D_x(t)K_t^T(m_x, D_x) + \Gamma_t \mathcal{N}_{\mu t} \Gamma_t^T, \quad D_x(t_0) = D_x(0). \end{aligned}$$

Используя решения этих уравнений, запишем точное выражение нелинейной функции

$$f_t(x) = f_{0t}(m_x, D_x) + K_t(m_x, D_x)x^0(t) + \Psi_t(x), \quad (3.89)$$

где $\Psi_t(x)$ - вектор-столбец нелинейных функций, характеризующих отклонение истинного значения функции $f_t(x)$ от приближенной линеаризованной зависимости.

Обозначим приближенное решение уравнения (3.88) через $x(t)$, а точное решение уравнения (3.87) - через $x(t) + \Delta(t)$. Подставим выражение (3.89) в уравнение (3.87), получаем уравнение для точного решения

$$\begin{aligned} \frac{dx(t)}{dt} + \frac{d\Delta(t)}{dt} &= f_{0t}(m_x, D_x) + K_t(m_x, D_x)(x^0(t) + \\ &+ \Delta^0(t)) + \Psi_t(x + \Delta(t)) + \Gamma_t \mu(t), \quad x(t_0) + \Delta(t_0) = x(0). \end{aligned}$$

Вычитая из этого соотношения почленно уравнение (3.88), получаем для вектора ошибок $\Delta(t)$ уравнение

$$\frac{d\Delta(t)}{dt} = K_t(m_x, D_x)\Delta^0(t) + \Psi_t(x + \Delta(t)), \quad (3.90)$$

где $\Psi(x + \Delta(t)) = f_t(x + \Delta(t)) - f_{0t}(m_x, D_x) - K_t(m_x, D_x)(x^0(t) + \Delta^0(t))$.

Ввиду того, что вектор ошибок $\Delta(t)$ имеет вероятностный характер, для него в рамках корреляционной теории можно определить вектор математического ожидания $m_\Delta(t)$ и матрицу дисперсии $D_\Delta(t)$. Указанные параметры вектора ошибки характеризуют погрешность статистической линеаризации. После осреднения векторного уравнения (3.90) последовательно относительно x и Δ уравнение для вектора математического ожидания ошибки принимает вид

$$\frac{dm_\Delta(t)}{dt} = M_\Delta \{M_x \{\Psi_t(x + \Delta(t))\}\}$$

или

$$\frac{dm_\Delta(t)}{dt} = M_\Delta \{\hat{\Psi}(\Delta(t))\}, \quad (3.91)$$

где $\hat{\Psi}(\Delta(t)) = M_x \{\Psi_t(x + \Delta(t))\}$.

Представим функцию $\hat{\Psi}(\Delta(t))$ первыми тремя членами разложения в ряд Тейлора $\hat{\Psi}(\Delta(t)) = \hat{\Psi}_{0t} + B\Delta(t) + R$, где функция $\hat{\Psi}_{0t}$ равна $\hat{\Psi}(\Delta(t))$ при $\Delta(t)=0$, B - матрица с компонентами

$b_{kj} = \frac{\partial \hat{\Psi}_k(\Delta(t))}{\partial \Delta_j \Delta_i}$, R - вектор с компонентами $\Delta^T C_k \Delta$, C_k - матрица с

компонентами $C_{kij} = \frac{1}{2} \frac{\partial^2 \hat{\Psi}_k(\Delta(t))}{\partial \Delta_i \partial \Delta_j} \Big|_{\Delta=0}$.

С учетом полученных результатов приходим на основании (3.91) к уравнению относительно вектора математического ожидания ошибки

$$\frac{dm_\Delta(t)}{dt} = Bm_\Delta(t) + M_\Delta \{R\}, \quad m_\Delta(t_0) = 0.$$

Уравнение матрицы дисперсий ошибок получают из выражения $D_\Delta(t) = \{\Delta_t^0 (\Delta_t^0)^T\}$, используя уравнение (3.90), а также уравнение ему сопряженное.

В скалярно-координатном представлении компоненты матрицы имеют вид

$$D_{\Delta ij}(t) = \sum_{k,l=1}^r \int_{t_0}^t \int_{t_0}^t g_{ik}(t, \tau) g_{li}(t, \tau') K_{\Psi_{kl}}(\tau, \tau') d\tau d\tau',$$

где $g_{ik}(t, \tau)$ - импульсная переходная функция с i -ого входа на j -ый выход системы, $K_{\Psi_{kl}} = M\{\Psi_{\tau,k}(x + \Delta)\Psi_{\tau,l}(x + \Delta)\}$.

В работах [28,31,32] приведена оценка погрешности для частных задач с разными нелинейностями, а также для узкополосных и широкополосных систем.

4. СТАТИСТИЧЕСКИЕ РЕШЕНИЯ И ОПТИМАЛЬНАЯ ОБРАБОТКА СИГНАЛОВ

4.1. СВЕДЕНИЯ ИЗ ТЕОРИИ СТАТИСТИЧЕСКИХ РЕШЕНИЙ

4.1.1. ОСНОВНЫЕ ПОНЯТИЯ

Основой для статистических решений и оптимальной обработки сигналов служат аналитические методы математической статистики. В математической статистике исходными являются имеющие случайный характер результаты наблюдения или экспериментальные данные из опытов. Задачей математической статистики является, во-первых, определения по результатам наблюдения путем надлежащей (оптимальной) обработки сигналов оценок параметров или характеристик (называемых статистическими) явлений или событий, связанных с наблюдениями, во-вторых, исследование точности их приближения к оцениваемым статистическим характеристикам^{*)}. Любая функция экспериментальных данных из опыта, которая не зависит от неизвестных статистических характеристик, называется статистикой. Оценкой статистической характеристики называется статистика, реализация которой принимается за неизвестное истинное значение параметра.

Задачи статистических решений возникают при наблюдении реализаций $\{y_{t_0}^t\}$ на текущем интервале времени (t_0, t) случайного

^{*)} В дальнейшем все подлежащие оценке величины - функции распределения, плотности вероятностей, их числовые параметры, называются статистическими характеристиками.

процесса $y(t)$, протекающих в непрерывном или дискретном времени. Пусть некоторый неизвестный наблюдателю параметр, принадлежащий заданному интервалу или всей действительной числовой оси, принимает одно или несколько значений. Предполагается, что распределение вероятностей наблюдаемого процесса зависит от этого параметра. Применительно к задачам проверки гипотез параметр принимает два или несколько значений, которым соответствует отсутствие или наличие одного или несколько полезных сигналов. Задача статистической проверки гипотез в этом случае формулируется следующим образом. Пусть наблюдаемая реализация $y_{t_0}^t$ является либо шумом, либо некоторой смесью шума и полезного сигнала, порождающего два или несколько различных событий (гипотез). Необходимо за фиксированное время на основании наблюдения оптимальным способом (с наименьшими потерями) определить, какая из возможных ситуаций имеет место. Для типичной задачи оценивания параметров наблюдаемый процесс также задается в виде смеси полезного сигнала и шума. Требуется определить какое значение принял сигнал (или его параметр). Оптимальному правилу, как и в случае проверки гипотез, отвечает процедура с наименьшими потерями. Если оцениваемый параметр принимает дискретные значения (случайные величины с априорным распределением), то задачу оценивания параметров можно свести к многоальтернативной задаче проверки гипотез. Когда параметр является случайным процессом, она становится задачей фильтрации. Нетрудно заметить, что задача оценки параметров является частным случаем задачи фильтрации, если за время наблюдения оцениваемый случайный процесс остается неизменным, либо существенно не изменяется. Процедура получения оценок этих параметров сводится к отысканию решающего правила $\gamma(y_{t_0}^t)$, как функции от реализации наблюдаемого сигнала. В результате того или иного принятого решения возможны ошибки. Потери, которые в связи с этим несет наблюдатель, можно охарактеризовать функцией потерь (штрафов) $\Pi[x, \gamma(y_{t_0}^t)]$, учитывающей величину потерь, возникающих вследствие принятия решения γ при условии, что истинное значение параметра x . Вид этой функции выбирается из практических соображений с учетом получения несложных решений. Функции

потерь используется для сравнения решающих правил и выбора из них более предпочтительно.

4.1.2. БАЙЕСОВСКИЕ ПРАВИЛА РЕШЕНИЯ

Наилучшее решающее правило можно определить при так называемом байесовском подходе. Поскольку решение $\gamma(y_{t_0}^t)$ зависит от реализации случайного процесса, т.е. является случайной величиной, функция потерь является случайной. Поэтому выбор решающего правила целесообразно проводить путем операции усреднения. Так, беря математическое ожидание от этой функции, получаем условный риск

$$r(x, \gamma) = M \{ \Pi[x, \gamma(y_{t_0}^t)] | x \} = \int_Y \Pi(x, \gamma) p(y_{t_0}^t | x) dy_{t_0}^t, \quad (4.1)$$

где $p(y_{t_0}^t | x)$ - функция правдоподобия (условная плотность вероятностей наблюдаемого сигнала при фиксированном значении x).

Согласно байесовскому подходу наиболее предпочтительным решающим правилом получения оценок следует считать таким, которым минимизируется условный риск (4.1) для всех значений x . Для решения этой задачи возьмем повторное математическое ожидание от условного риска относительно априорной плотности вероятностей $p(x)$. В результате получим выражение среднего риска

$$R(\gamma) = \int_X r(x, \gamma) p(x) dx = \int_X \int_Y \Pi(x, \gamma) p(x, y_{t_0}^t) dy_{t_0}^t dx. \quad (4.2)$$

Для решающего правила $\hat{\gamma}$ минимизирующего средний риск относительно всех других решающих правил γ , выполняется неравенство

$$R(\hat{\gamma}) \leq R(\gamma).$$

Решение $\hat{x} = \hat{\gamma}$ называется оптимальным байесовским решением относительно априорного распределения $p(x)$ и выбранной функции потерь. Минимальное значение среднего риска является мерой качества работы байесовских систем обнаружения, оценивания параметров и фильтрации случайных сообщений. На основании

формулы для условных плотностей вероятностей (1.4), которая для рассматриваемого случая имеет вид

$$p(x | y_{t_0}^t) = \frac{p(x, y_{t_0}^t)}{p(y_{t_0}^t)} = \frac{p(y_{t_0}^t | x)p(x)}{\int_{-\infty}^{\infty} p(y_{t_0}^t | x)p(x)dx}, \quad (4.3)$$

перепишем средний риск

$$R(\gamma) = \int_Y R(\gamma | y_{t_0}^t) p(y_{t_0}^t) dy_{t_0}^t,$$

где

$$R(\gamma | y_{t_0}^t) = M\{\Pi(x, \gamma) | y_{t_0}^t\} = \int_X \Pi(x, \gamma) p(x | y_{t_0}^t) dx \quad (4.4)$$

- апостериорный риск; $p(x | y_{t_0}^t)$ - апостериорная плотность вероятностей^{*)}.

Так как $p(y_{t_0}^t) \geq 0$, то отсюда следует, что минимумы среднего и апостериорных рисков достигаются при одном и том же значении $\hat{\gamma}$. Таким образом, синтез байесовских систем можно проводить, используя средний или апостериорный риски. Полученные результаты обобщаются и на случай дискретных наборов наблюдаемых сигналов $y = \|y_1, \dots, y_n\|^T$ и оцениваемых параметров $x = \|x_1, \dots, x_n\|^T$. Средний и апостериорный риски при достаточно длинной серии экспериментов приближенно равняются в пространстве выборок математическим ожиданиям функции потерь

$$R(\gamma) = \sum_{j=1}^n \sum_{k=1}^n p_j \Pi_{jk} P(y_k | x_j),$$

$$R(\gamma | y) = \sum_{j=1}^n \Pi_{jk} P(x_j | y).$$

^{*)} Апостериорную плотность вероятностей, как и другие апостериорные статистические характеристики, получают после опыта по экспериментальным данным на основании реализаций случайного процесса.

Для этой задачи совокупность наблюдений (y_1, \dots, y_n) параметров (x_1, \dots, x_n) принадлежит одному и тому же пространству выборок; n - число непересекающихся подобластей пространства выборок; p_j - априорная вероятность состояния x_j ; $P(y_k|x_j)$ - условная вероятность попадания выборки в область y_k , если действительно имеет место состояние x_j ; $P(x_j|y)$ - апостериорная вероятность состояния x_j . Каждому ошибочному решению $k \neq j$ приписывается штраф Π_{jk} .

Оптимальному выбору решения соответствуют следующие неравенства

$$R(\hat{\gamma}) \leq R(\gamma),$$

$$R(\hat{\gamma} | y) \leq R(\gamma | y).$$

Таким образом, при синтезе байесовских систем необходимо задать функцию потерь и определить по известным реализациям наблюдаемого сигнала апостериорные функции распределения или плотности вероятностей. Нахождение этих характеристик является одним из существенных моментов в теории синтеза алгоритмов информационных систем. Выбор вида функции потерь в известной степени субъективен и зависит от конкретной ситуации. В технических приложениях наиболее распространенными функциями потерь, зависящими от величины ошибки $\varepsilon_x = x - \hat{x}$, являются

- простая $\Pi(x, \hat{x}) = 1 - \delta(\varepsilon_x)$, $\delta(\cdot)$ - дельта- функция;
- модульная $\Pi(x, \hat{x}) = |\varepsilon_x|$;
- квадратичная $\Pi(x, \hat{x}) = \varepsilon_x^2$.

Если апостериорный риск дифференцируем, то после подстановки вышеприведенных функций потерь в (4.4), определения экстремума, получаем следующие результаты.

Для простой функции потерь на основании фильтрующих свойств дельта-функции из выражения (4.4) получаем

$$R(\gamma | y_{t_0}^t) = 1 - p(x | y_{t_0}^t) \Big|_{x = \hat{x}}.$$

После определения экстремума для простой функции потерь оптимальная оценка определяется по критерию максимума АПВ

$$\left. \frac{\partial R(\gamma | y_{t_0}^t)}{\partial x} \right|_{x=\hat{x}} = \left. \frac{\partial p(x | y_{t_0}^t)}{\partial x} \right|_{x=\hat{x}} = 0 \quad (4.5)$$

Так как логарифм есть монотонная функция своего положительного аргумента, то вместо (4.5) используют соотношение

$$\left. \frac{\partial \ln p(x | y_{t_0}^t)}{\partial x} \right|_{x=\hat{x}} = 0.$$

Для модульной функции потерь апостериорный риск принимает вид

$$R(\gamma | y_{t_0}^t) = \int_X |\varepsilon_x| p(x | y_{t_0}^t) dx = \int_{-\infty}^{\hat{x}} \varepsilon_x p(x | y_{t_0}^t) dx - \int_{\hat{x}}^{-\infty} \varepsilon_x p(x | y_{t_0}^t) dx,$$

и условию экстремума соответствует выражение

$$\frac{\partial R(\gamma | y_{t_0}^t)}{\partial x} = \int_{-\infty}^{\hat{x}} p(x | y_{t_0}^t) dx - \int_{\hat{x}}^{-\infty} p(x | y_{t_0}^t) dx = 0, \quad (4.7)$$

где область определения параметра \hat{x} принадлежит действительной оси.

Из (4.7) следует, что байесовская оценка в этом случае совпадает с условной медианой АПВ.

Байесовская оценка при квадратичной функции потерь оптимальна по критерию минимума дисперсии ошибки оценки или, что то же самое, апостериорного риска

$$D(t) = R(\gamma | y_{t_0}^t) = \int_{-\infty}^{\infty} \varepsilon_x^2 p(x | y_{t_0}^t) dx = \int_{-\infty}^{\infty} (x - \hat{x})^2 p(x | y_{t_0}^t) dx. \quad (4.8)$$

Взяв производную по $x = \hat{x}$ от апостериорного риска и приравняв ее нулю, получаем, что эта оценка представляет собой апостериорное среднее

$$\hat{x} = M\{x | y_{t_0}^t\} = \int_{-\infty}^{\infty} x p(x | y_{t_0}^t) dx. \quad (4.9)$$

Покажем, что апостериорной оценке также соответствует минимальное значение безусловной дисперсии ошибки

$$M\{(x - \hat{x})^2\} = M_Y\{M\{(x - \hat{x})^2 | y_{t_0}^t\}\},$$

где $M\{(x - \hat{x})^2 | y_{t_0}^t\}$ - апостериорное математическое ожидание при фиксированной реализации наблюдаемого процесса $y_{t_0}^t$; $M_Y\{\cdot\}$ - математическое ожидание по всем возможным реализациям процесса $y(t)$.

Предположим, что известна совместная плотность вероятностей процессов x и $y_{t_0}^t$ - $p(x, y_{t_0}^t)$. Безусловная дисперсия ошибки для рассматриваемого случая равна

$$\begin{aligned} M\{(x - \hat{x})^2\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \hat{x})^2 p(x, y_{t_0}^t) dx dy_{t_0}^t = \\ &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} (x - \hat{x})^2 p(x, y_{t_0}^t) dx \right\} p(y_{t_0}^t) dy_{t_0}^t. \end{aligned} \quad (4.10)$$

Выражение в фигурных скобках соотношения (4.10) является апостериорной дисперсией ошибки. Представим ее с учетом двух последних слагаемых в сумме равных нулю в развернутом виде

$$\begin{aligned} &\int_{-\infty}^{\infty} (x - \hat{x})^2 p(x, y_{t_0}^t) dx = \\ &= M\{x^2 | y_{t_0}^t\} - 2\hat{x}M\{x | y_{t_0}^t\} + \hat{x}^2 + \\ &+ [M\{x | y_{t_0}^t\}]^2 - [M\{x | y_{t_0}^t\}]^2 = \\ &= [M\{x | y_{t_0}^t\} - \hat{x}]^2 + M\{x^2 | y_{t_0}^t\} - [M\{x | y_{t_0}^t\}]^2. \end{aligned} \quad (4.11)$$

По определению безусловная дисперсия ошибки является положительной величиной и как следует из соотношений (4.10) и (4.11), принимает минимальное значение при оценке $\hat{x} = M\{x | y_{t_0}^t\}$, равной апостериорному среднему.

Байесовские оценки для рассматриваемых видов потерь одинаковы, если АПВ имеет один максимум (униmodalна) и

симметрична относительно своей медианы (среднего значения). Очевидно, этим требованиям отвечает гауссовская апостериорная плотность вероятностей.

Если оцениваемый параметр является вектором, то аналогом квадратичной функции потерь является квадратичная форма

$$\Pi(x, \hat{x}) = \sum_{i,j=1}^r A_{ij} (x_i - \hat{x}_i)(x_j - \hat{x}_j) = (x - \hat{x})^T A (x - \hat{x}),$$

где A - положительно определенная матрица с весовыми компонентами A_{ij} . В частности, эта матрица может быть единичной и в этом случае $\Pi(x, \hat{x}) = (x - \hat{x})^T (x - \hat{x})$.

Минимизация апостериорного риска $R(\gamma | y_{t_0}^t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x - \hat{x})^T (x - \hat{x}) p(x | y_{t_0}^t) dx$ приводит к формуле, определяющей апостериорное среднее вектора оцениваемого параметра

$$\hat{x} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x p(x | y_{t_0}^t) dx \quad (4.12)$$

$$\text{с компонентами } \hat{x}_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i p(x | y_{t_0}^t) dx.$$

Можно показать, как и для случая скалярных процессов, что минимальные значения матриц апостериорных и безусловных дисперсий ошибок достигается для апостериорного вектора оценок \hat{x} .

Важными характеристиками, по которым отбирают оценки, являются состоятельность, несмещенность и эффективность. Оценка \hat{x}_n называется состоятельной, если она сходится по вероятности к среднему значению оцениваемого случайного параметра $M\{x\}$ при неограниченном увеличении выборки n , т.е. при произвольном $\varepsilon > 0$ имеет место

$$\lim_{n \rightarrow \infty} P\{|\hat{x}_n - M\{x\}| \geq \varepsilon\} = 0.$$

Для несмещенной оценки среднее значение этой оценки при любом n равно математическому ожиданию параметра

$$M\{\hat{x}_n\} = \int_Y x_n p(y_1^n) dy_1^n.$$

В частности, байесовская оценка при квадратичной функции потерь является несмещенной.

Эффективной в классе несмещенных оценок называют оценку с наименьшей дисперсией

$$M\{(\hat{x}_n - x)^2\} \geq M\{(\hat{x}_{n,эфф} - x)^2\} = D_{эфф}.$$

Часто эффективность определяется из условия достижения нижней границы в неравенстве Крамера-Рао.

Необходимо пользоваться такими оценками, для которых дисперсия ошибки равна или близка $D_{эфф}$. Для нахождения этих оценок в многих случаях оказывается полезным понятие достаточной оценки (статистики). Определяют достаточную оценку z , как некоторую функцию наблюдаемой реализации $y_{t_0}^t$ так, чтобы соблюдалось равенство для апостериорных рисков или АПВ

$$M\{\Pi(x, \gamma) | y_{t_0}^t\} = M\{\Pi(x, \gamma) | z\},$$

$$p(x | y_{t_0}^t) = p(x | z).$$

Достаточными статистиками могут быть оценки параметров, неизвестных процессов, сами наблюдаемые реализации. В последнем случае существует более компактное разбиение пространства наблюдений за счет сокращения объема выборки наблюдаемых отсчетов или длины наблюдаемой реализации без изменения величины апостериорного риска или АПВ.

Достаточная статистика содержит всю информацию о неизвестном параметре, которую можно получить при наблюдениях. В случае марковских сигналов для оценки параметра в момент времени t_{n+1} нет необходимости определять каждый раз все отсчеты наблюдаемого процесса из интервала (t_0, t_{n+1}) достаточно получить последний отсчет для t_{n+1} и иметь лишь одну функцию от реализации на интервале (t_0, t_n) - достаточную статистику. Если $z_n = \hat{x}_n$

достаточная статистика, то зависящая от выборочных значений реализации y_1^{n+1} апостериорная плотность вероятностей $p(x | y_1^{n+1})$ преобразуется в АПВ вида $p(x | z_n, y_{n+1})$.

4.2. ОБСУЖДЕНИЕ ПОДХОДОВ РЕШЕНИЯ ЗАДАЧИ

На основании полученных результатов изложим в общих чертах возможные подходы определения оценок неизвестных параметров или случайных процессов. Представим рассматриваемые процессы, протекающими в дискретном времени $(t_n, n=0,1,2,\dots)$.

Наблюдаемый сигнал и оцениваемый процесс являются для этого случая последовательностями отсчетов соответственно $\{y_n\}$ и $\{x_n\}$. Апостериорную плотность вероятностей на основании формулы условных вероятностей (1.4) представим в виде

$$\begin{aligned} p(x_{n+1} | y_1^{n+1}) &= \\ &= c_{n+1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(y_1^{n+1} | x_{n+1}, \dots, x_1, x_0) p(x_{n+1}, \dots, x_1, x_0) dx_0 \dots dx_n, \end{aligned}$$

где x_0 - начальное значение параметра, y_1^{n+1} - последовательность наблюдений (y_1, \dots, y_{n+1}) ; плотность вероятностей $p(x_{n+1}, \dots, x_1, x_0) = p(x_0)p(x_1|x_0)\dots p(x_{n+1}|x_0, \dots, x_n)$

Сомножитель $[p(y_1^{n+1})]^{-1}$ не зависит от $\{x_{n+1}\}$, его учитывают в виде нормировочной постоянной c_{n+1} . Интегрирование осуществляется по всей области определения вектора $x_0^n = (x_0, x_1, \dots, x_n)$.

Используя, например, критерий минимума среднеквадратической ошибки можно, вообще говоря, получить оптимальную оценку \hat{x}_{n+1} и дисперсию ошибки оценки D_{n+1} . Но такой способ решения задачи не эффективен, поскольку представляет собой трудно разрешимую проблему из необходимости определения $n+1$ -кратного интеграла. Существенными преимуществами в отношении вычислений и моделирования обладают методы, использующие марковские модели процессов и связанные с ними рекуррентные процедуры. Пусть x является непрерывным марковским процессом. В дискретном времени его также можно представить в виде

последовательности отсчетов $\{x_{n+1}\}$. Плотность вероятностей марковского процесса (2.31) позволяет, используя двумерные переходные вероятности $\pi(x_{n+1}|x_n)$ и известную функцию правдоподобия $p(y_{n+1}|x_{n+1})$, по отсчетам наблюдаемого процесса y_{n+1} рекуррентно вычислять апостериорные плотности вероятностей $\{p(x_{n+1}|y_1^{n+1})\}$. А по ней на основании формул (4.9) и (4.8) - оценки и дисперсии ошибок оценок. И хотя представленная картина определения оценок носит качественный характер (подробное и строгое изложение рекуррентных алгоритмов будет проведено в разделе 4.6.2) можно сформулировать важный результат: для получения оценки в момент времени t_{n+1} используется последний результат (отсчет) наблюдения y_{n+1} и достаточная статистика, как функция всех предшествующих наблюдений. Полагая в данном случае достаточной статистикой оценку параметров $z = \hat{x}$, можно записать операторное рекуррентное уравнение

$$\hat{x}_{n+1} = J_{n+1}(y_{n+1}, \hat{x}_n),$$

где оператор $J_{n+1}(\cdot)$ зависит от статистических свойств оцениваемого сигнала x и наблюдаемого процесса y . Общая методика данного подхода к рекуррентным процедурам совместного обнаружения и фильтрации применительно к радиотехническим задачам впервые была рассмотрена в работах [33,34,35]. Аналогичные задачи, в которых моделями наблюдаемых сигналов являются случайные потоки, рассматривались в [36,37,38].

Точное решение уравнений оценивания (фильтрации) вызывает, за исключением ряда частных случаев, большие технические трудности. Обычно ищут приближенные решения. Одним из таких решений является представление АПВ в виде разложения его логарифма в ряд по заданной системе функций $\varphi_i(x_t)$ относительно некоторой оценки x_t^0

$$\ln p(x_t | y_{t_0}^t) \approx \ln p(x_t^0 | y_{t_0}^t) - \sum_{i=1}^K h_{it} \varphi_i(x_t), \quad (4.13)$$

где h_{it} - неизвестные апостериорные параметры i -го порядка, K - порядок аппроксимации.

Очевидно, чем больше K , тем точнее может быть описана АПВ. С другой стороны, с увеличением порядка аппроксимации сложнее решить задачу синтеза. Для упрощения урезают число аппроксимирующих АПВ членов ряда. Случай $K=2$ соответствует квазиоптимальному в гауссовском приближении методу фильтрации. В качестве системы функций выбирают степенную функцию $\varphi_i(x_t) = (x_t - x_t^0)^i / i!$. В этом случае логарифм АПВ является усеченным (до K -го члена) разложением в ряд Тейлора, в котором апостериорные параметры определяются из соотношения

$$h_{it} = - \frac{\partial^i \ln p(x_t | y_{t_0}^t)}{\partial x^i} \Big|_{x_t = x_t^0}.$$

Если обеспечить $h_{1t} = \frac{\partial \ln p(x_t | y_{t_0}^t)}{\partial x_t} \Big|_{x_t = x_t^0} = 0$, то x_t^0 является оценкой \hat{x}_t , соответствующей критерию максимума апостериорной плотности вероятностей. Заметим, что определение системы неизвестных параметров h_{2t}, \dots, h_{kt} эквивалентно заданию приближенного значения АПВ (до K члена разложения ряда).

Использование при синтезе приближенного значения АПВ неизбежно приводит к некоторым потерям оптимальности. Отказ от полной оптимальности - вынужденная мера и является своеобразной платой за упрощение технических решений.

4.3. ОБНАРУЖЕНИЕ И РАЗЛИЧЕНИЕ СИГНАЛОВ

В простейшей задаче обнаружения некоторый источник порождает сигналы в виде двух значений параметра. По результатам наблюдения $y_{t_0}^t$, представляющего собой смесь полезного сигнала и шума, необходимо с наименьшими потерями решить, какая из возможных ситуаций имеет место. Предлагаемая задача обнаружения в теории статистических решений эквивалентна проверки простой гипотезы - утверждения, что $\Theta=1$ есть сигнал против простой альтернативы - утверждения, что $\Theta=0$ нет сигнала. При байесовском походе решающее правило должно минимизировать средний риск (4.2). При выводе выражения среднего риска полагаем заданными априорные вероятности каждой из гипотез $p_0=P(\Theta=0)$ и $p_1=P(\Theta=1)$, $p_0+p_1=1$. Функция потерь здесь переходит в матрицу

$$\Pi = \begin{vmatrix} \Pi(0, \gamma_0) & \Pi(0, \gamma_1) \\ \Pi(1, \gamma_0) & \Pi(1, \gamma_1) \end{vmatrix},$$

где γ_0 - решение, что $\Theta=0$, γ_1 - решение, что $\Theta=1$. Положим без существенных ограничений общности подхода

$$\Pi(0, \gamma_0) = \Pi(1, \gamma_1) = 0, \quad \Pi(0, \gamma_1) > 0, \quad \Pi(1, \gamma_0) > 0.$$

Разобьем пространство наблюдений Γ на две непересекающиеся подобласти Γ_1 и Γ_0 , $\Gamma = \Gamma_0 \cup \Gamma_1$. Если результат наблюдения оказался в Γ_1 , то принимается гипотеза $\Theta=1$, если в Γ_0 - то ее альтернатива $\Theta=0$. В двухальтернативной задаче обнаружения принятие решения сопровождается ошибками первого и второго рода. Ошибка первого рода α - альтернатива $\Theta=0$ отвергается тогда, когда в действительности она верна; ошибка второго рода β - отвергается гипотеза $\Theta=1$ в то время, как она верна. В этом случае вероятности ошибочных решений

$$\alpha = P\{y_{t_0}^t \in \Gamma_1 | 0\} = \int_{\Gamma_1} p(y_{t_0}^t | 0) dy_{t_0}^t,$$

$$\beta = P\{y_{t_0}^t \in \Gamma_0 | 0\} = \int_{\Gamma_0} p(y_{t_0}^t | 1) dy_{t_0}^t.$$

где $p(y_{t_0}^t | 0)$ и $p(y_{t_0}^t | 1)$ - функции правдоподобия соответственно при гипотезах $\Theta=0$ и $\Theta=1$. Для указанного разбиения пространства наблюдения на две подобласти средний риск (4.2) можно представить в виде суммы двух слагаемых

$$R(\gamma) = \alpha \Pi(0, \gamma_1) p_0 + \beta \Pi(1, \gamma_0) p_1. \quad (4.14)$$

Поскольку $\Gamma_1 = \Gamma / \Gamma_0$, то на основании условия нормировки можно записать $\int_{\Gamma_1} p(y_{t_0}^t | 0) dy_{t_0}^t = 1 - \int_{\Gamma_0} p(y_{t_0}^t | 0) dy_{t_0}^t$ и средний риск (4.14) становится равным

$$R(\gamma) = \Pi(0, \gamma_1) p_0 - \int_{\Gamma_0} \{p(y_{t_0}^t | 0) \Pi(0, \gamma_1) p_0 - p(y_{t_0}^t | 1) \Pi(1, \gamma_0) p_1\} dy_{t_0}^t.$$

Ввиду того, что $\Pi(0, \gamma_1)p_0$ - постоянная неотрицательная величина, минимальное значение среднего риска будет получено, если подинтегральная функция будет неотрицательной, т.е. в подобласть Γ_0 включены все точки пространства наблюдения, для которых выполняется

$$p(y_{t_0}^t | 0)\Pi(0, \gamma_1)p_0 > p(y_{t_0}^t | 1)\Pi(1, \gamma_0)p_1. \quad (4.15)$$

Точки, для которых выполняется обратное неравенство, следует отнести к подобласти Γ_1 . На основании (4.15) приходим к байесовскому оптимальному решению

$$\Lambda(y_{t_0}^t) = \frac{p(y_{t_0}^t | 1)}{p(y_{t_0}^t | 0)} \underset{\gamma_0}{\overset{\gamma_1}{>}} h = \frac{\Pi(0, \gamma_1)p_0}{\Pi(1, \gamma_0)p_1}. \quad (4.16)$$

Таким образом, оптимальное (согласно байесовскому подходу) правило решения сводится к формированию статистики, называемой отношением правдоподобия $\Lambda(y_{t_0}^t)$ (ОП), и к сравнению с его порогом обнаружения h . Величина порога определяется априорными вероятностями наличия и отсутствия сигнала и задаваемой матрицей функций потерь. Отметим, что решающему правилу (4.16) эквивалентно неравенство

$$\ln \Lambda(y_{t_0}^t) \underset{\gamma_0}{\overset{\gamma_1}{>}} h_1,$$

так как обе части неравенства (4.16) являются положительными величинами. Для простых функций потерь $\Pi(0, \gamma_0)=\Pi(1, \gamma_1)=0$, $\Pi(0, \gamma_1)=\Pi(1, \gamma_0)=1$ средний риск равен $R(\gamma)=\alpha p_0+\beta p_1$, который совпадает с полной вероятностью ошибки. Таким образом, при простых функциях потерь минимум среднего риска приводит к минимуму полной вероятности ошибки. Байесовский критерий обнаружения можно записать на основании (1.4) в терминах апостериорной вероятности. Оптимальному решению, называемому в этом случае правилом решения идеального наблюдателя, соответствует неравенство

$$l(y_{t_0}^t) = \frac{p(1 | y_{t_0}^t)}{p(0 | y_{t_0}^t)} \underset{\gamma_0}{\overset{\gamma_1}{>}} 1, \text{ где } l(y_{t_0}^t) = \Lambda(y_{t_0}^t) \frac{p_1}{p_0}.$$

Если представить непрерывную реализацию $y_{t_0}^t$ в дискретном времени в виде n -мерного вектора отсчетов y_1^n , то отношение правдоподобия $\Lambda(y_1^n)$ и соответственно $l(y_1^n)$ оказываются достаточными статистиками обнаружения. Достаточная статистика позволяет свести n -мерный вектор наблюдения к скалярным величинам $\Lambda(y_1^n)$ или $l(y_1^n)$ при соблюдении равенства апостериорных рисков

$$M\{\Pi(\Theta, \gamma(y_1^n) | y_1^n)\} = M\{\Pi(\Theta, \gamma(y_1^n) | \Lambda(y_1^n))\}.$$

Полученные результаты можно распространить на задачи различения многих гипотез или многоальтернативного обнаружения. С указанными проблемами сталкиваются при обработке большого числа сигналов в радиосвязи и радиолокации, в задачах распознавания и разрешения многих целей и т.д. Необходимо отметить, что задачи многоальтернативного обнаружения, по существу, адекватны статистической теории распознавания образов, имеющей широкое применение в теории статистических решений. Применительно к задачам оценивания при многоальтернативном обнаружении Θ ассоциируется с самим сигналом x (или параметрами этого сигнала) из непрерывного множества действительной оси.

Пусть наблюдаемая реализация $y_{t_0}^t$ представляет собой шум или смесь шума с одним из возможных сигналов, который отождествляется с параметром Θ_i , $i = \overline{0, M}$. Обозначим через Γ_i подобласть пространства наблюдения Γ , в которой принимается решение о верности гипотезы - утверждения, что Θ_i есть сигнал. Оптимальное правило состоит в принятии решения минимизирующего средний риск

$$R(\gamma) = \sum_{i=0}^M \sum_{j=0}^M \Pi(\Theta_j, \gamma_i) p_i \int_{\Gamma_i} p(y_{t_0}^t | \Theta_j) dy_{t_0}^t.$$

Минимизация достигается определением подобласти Γ_i , в которой средний риск минимален. При $\Pi(\Theta_j, \gamma_i)=1$, $i \neq j$ и $\Pi(\Theta_j, \gamma_i)=0$ для всех остальных $i, j = \overline{0, M}$ указанный критерий приводит к принятию решения о верности той из гипотез Θ_i , для которой апостериорные вероятности имеют максимальное значение $P(\Theta_i | y_{t_0}^t) > P(\Theta_j | y_{t_0}^t)$ для $i \neq j$.

Это правило можно записать по-другому

$$p_i \Lambda_i(y_{t_0}^t) > p_j \Lambda_j(y_{t_0}^t),$$

$$\text{где } \Lambda_{i(j)} = \frac{p(y_{t_0}^t | \Theta_{i(j)})}{p(y_{t_0}^t | \Theta_0)}.$$

Если к тому же априорные вероятности p_i всех состояний оказываются одинаковыми, то это правило сводится к выбору наибольшего ОП

$$\Lambda_i(y_{t_0}^t) > \Lambda_j(y_{t_0}^t), \quad i \neq j.$$

Многоальтернативное обнаружение можно использовать в задачах совместного обнаружения сигнала и оценки его параметров, которые принимают одно из значений μ_i , $i = \overline{0, M}$. В этом случае также приходим к многоканальной системе обработки сигнала, в которой оценка параметра сигнала определяется по номеру канала с максимальным выходным эффектом. Структура канала задается условным ОП:

$$\Lambda(y_{t_0}^t | \mu_i) = \frac{p(y_{t_0}^t | \mu_i, \Theta_i)}{p(y_{t_0}^t | \Theta_0)}.$$

4.4. НЕБАЙЕСОВСКИЕ ПРАВИЛА РЕШЕНИЯ

4.4.1. КРИТЕРИЙ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Байесовский подход связан с предположением, что априорные распределения состояний случайных величин и функции потерь известны. Если этих априорных сведений нет, то используются

небайесовские методы принятия решений. К их числу относят критерий максимального правдоподобия.

Представим на основании формулы Байеса (1.4) соотношение (4.6) в виде

$$\left. \frac{\partial \ln p(y_{t_0}^t | x)}{\partial x} \right|_{x = \hat{x}} + \left. \frac{\partial \ln p(x)}{\partial x} \right|_{x = \hat{x}} = 0.$$

В практических задачах сталкиваются с ситуацией, когда априорную плотность вероятностей $p(x)$ затруднительно заранее определить. Тогда ее аппроксимируют достаточно широким равномерным распределением и принимают $\partial \ln p(x) / \partial x = 0$. В этом случае оценка по критерию максимума АПВ совпадает с небайесовской оценкой по критерию максимума функции правдоподобия

$$\left. \frac{\partial \ln p(y_{t_0}^t | x)}{\partial x} \right|_{x = \hat{x}} = 0. \quad (4.17)$$

Условию экстремума (4.17) соответствует такая оценка параметра, для которой при заданном объеме наблюдаемых отсчетов $y_1^n = (y_1, \dots, y_n)$ функция $p(y_1^n | x)$ достигает максимума. Оценка, удовлетворяющая указанному критерию, называется максимально правдоподобной.

Для независимых отсчетов из уравнения (4.17) следует

$$\sum_{i=1}^n \frac{\partial}{\partial x} \ln p(y_i | x) = 0.$$

Указанный критерий получил широкое распространение в задачах оценивания постоянных параметров (случайных величин) $x = \Theta$. Эти оценки при $t \rightarrow \infty$ асимптотически состоятельны, эффективны и нормально распределены [39].

4.4.2. КРИТЕРИЙ НЕЙМАНА-ПИРСОНА

Решающее правило в этом случае разбивает пространство наблюдения Γ также на две непересекающиеся подобласти Γ_1 и Γ_0 ,

$\Gamma = \Gamma_1 \cup \Gamma_0$. Но выбор этих подобластей производится из других соображений. Используя ранее рассмотренные вероятности ошибок первого рода - α (вероятность ложной тревоги) и вероятность ошибки второго рода - β (вероятность пропуска сигнала). Вероятность $1-\beta$ есть вероятность правильного обнаружения. Согласно этому критерию оптимальным решением считается такое, которое обеспечивает минимум вероятности пропуска сигнала β (максимум $1-\beta$) при условии, что вероятность ложной тревоги α не превышает заданного числа. Этому решающему правилу эквивалентно неравенство (4.16)

$$\Lambda(y_{t_0}^t) = \frac{p(y_{t_0}^t | 1)}{p(y_{t_0}^t | 0)} \underset{\gamma_0}{\overset{\gamma_1}{>}} h, \quad (4.18)$$

где порог выбирается из условия

$$P\{\Lambda(y_{t_0}^t) \geq h | 0\} = P\{y_{t_0}^t \in \Gamma_1 | 0\} = \int_h^\infty P(\Lambda | 0) d\Lambda = \alpha. \quad (4.19)$$

Для этого требуется доказать, что для любой подобласти $\tilde{\Gamma} \in \Gamma_1$, удовлетворяющей условию $P\{y_{t_0}^t \in \tilde{\Gamma} | 0\} \leq \alpha$ справедливо неравенство

$$1 - \beta = P\{y_{t_0}^t \in \Gamma_1 | 1\} \geq P\{y_{t_0}^t \in \tilde{\Gamma} | 1\} = 1 - \beta_{\tilde{\Gamma}}$$

или $\beta_{\tilde{\Gamma}} \geq \beta$.

Обозначим через G пересечение подобласти Γ_1 и $\tilde{\Gamma}$, а также разности подобластей $\Gamma_1 / G = A$, $\tilde{\Gamma} / G = B$. Так как $A \in \Gamma_1$, то с учетом (4.18) имеем

$$\int_A p(y_{t_0}^t | 1) dy_{t_0}^t \geq h \int_A p(y_{t_0}^t | 0) dy_{t_0}^t. \quad (4.20)$$

На основании (4.20) можно записать

$$\begin{aligned}
\int_{\Gamma_1} p(y_{t_0}^t | 1) dy_{t_0}^t &= \int_A p(y_{t_0}^t | 1) dy_{t_0}^t + \int_G p(y_{t_0}^t | 1) dy_{t_0}^t \geq \\
&\geq h \int_A p(y_{t_0}^t | 0) dy_{t_0}^t + \int_G p(y_{t_0}^t | 1) dy_{t_0}^t = \\
&= h \int_{\Gamma_1} p(y_{t_0}^t | 0) dy_{t_0}^t - h \int_G p(y_{t_0}^t | 0) dy_{t_0}^t + \int_G p(y_{t_0}^t | 1) dy_{t_0}^t,
\end{aligned} \tag{4.21}$$

Так как $\alpha = \int_{\Gamma_1} p(y_{t_0}^t | 0) dy_{t_0}^t \geq \int_{\tilde{\Gamma}} p(y_{t_0}^t | 0) dy_{t_0}^t$, то выражение (4.21)

можно переписать

$$\begin{aligned}
\int_{\Gamma_1} p(y_{t_0}^t | 1) dy_{t_0}^t &\geq h \int_{\tilde{\Gamma}} p(y_{t_0}^t | 0) dy_{t_0}^t - h \int_G p(y_{t_0}^t | 0) dy_{t_0}^t + \\
&+ \int_G p(y_{t_0}^t | 1) dy_{t_0}^t = h \int_B p(y_{t_0}^t | 0) dy_{t_0}^t + \int_G p(y_{t_0}^t | 1) dy_{t_0}^t.
\end{aligned} \tag{4.22}$$

Ввиду того, что B - подобласть наблюдения, где гипотеза $\Theta=1$ отвергается, для этой области имеет место $\frac{p(y_{t_0}^t | 1)}{p(y_{t_0}^t | 0)} \leq h$ и поэтому

$$\int_B p(y_{t_0}^t | 1) dy_{t_0}^t \leq h \int_B p(y_{t_0}^t | 0) dy_{t_0}^t.$$

Отсюда на основании (4.22) имеем

$$\int_{\Gamma_1} p(y_{t_0}^t | 1) dy_{t_0}^t \geq \int_B p(y_{t_0}^t | 1) dy_{t_0}^t + \int_G p(y_{t_0}^t | 1) dy_{t_0}^t = \int_{\tilde{\Gamma}} p(y_{t_0}^t | 1) dy_{t_0}^t$$

или

$$P\{y_{t_0}^t \in \Gamma_1\} \geq P\{y_{t_0}^t \in \tilde{\Gamma} | 1\}, \quad \beta_{\tilde{\Gamma}} \geq \beta,$$

что и требовалось доказать.

Согласно рассмотренному решающему правилу оптимальный обнаружитель формирует ОП и подает его на пороговое устройство, где осуществляется процедура сравнения с порогом h , который определяется при заданном α и известном распределении $p(\Lambda|0)$ из формулы (4.19).

4.4.3. МИНИМАКСНОЕ РЕШАЮЩЕЕ ПРАВИЛО

Это правило в теории статистических решений рассматривается как специальный случай байесовского решения для наименее благоприятного распределения состояний случайной величины $p(x)$. Минимаксное решение минимизирует максимальный по всем значениям параметра x средний риск $R_{MM}(\gamma) = \min_{\gamma} \max_{p(x)} R(\gamma)$. Это

решение является наилучшим решением в наихудшей относительно x ситуации. В общем случае отыскание минимаксного решения - довольно сложная задача. При более детальном анализе можно показать, что при некоторых слабых ограничениях указанное решение можно получить, если априорное распределение не зависит от x . Таким неблагоприятным распределением часто оказывается равномерное распределение.

В случае проверки простой гипотезы параметр принимает два значения $\Theta=0$ или $\Theta=1$ и указанное априорное распределение определяется лишь одной вероятностью $p(\Theta=1)=p_1$, $p(\Theta=0)=p_0=1-p_1$. Обратим внимание на то, что ошибка первого и второго рода зависят от априорных вероятностей согласно формулам

$$\alpha(p_1)=P(\Lambda>h|0), \beta(p_1)=P(\Lambda<h|1),$$

где порог обнаружения h определяется из формулы (4.16). Для получения наибольшего значения среднего риска продифференцируем (4.14) по p_1 и приравняем результат дифференцирования нулю, получаем

$$\Pi(0, \gamma_1) \alpha(p_{1MM}) = \Pi(1, \gamma_0) \beta(p_{1MM}).$$

Решая это уравнение, находим то значение p_{1MM} , которому соответствует максимальное значение среднего риска по априорному распределению. В результате приходим к следующему минимаксному правилу решения

$$\Lambda(y_{t_0}^t) \begin{matrix} > \\ < \end{matrix} \begin{matrix} \gamma_1 \\ \gamma_0 \end{matrix} h,$$

где на основании (4.16) $h = \frac{\Pi(0, \gamma_1)}{\Pi(1, \gamma_0)} \cdot \frac{1 - p_{1MM}}{p_{1MM}}$.

Разность $R_{MM}(\gamma) - R(\gamma) > 0$ между минимаксным и байесовским средними рисками является своеобразной платой за отсутствие информации об априорном распределении.

Как следует из предыдущего, байесовский и минимаксный критерии, критерий Неймана-Пирсона приводят к решающему правилу обнаружения, основанному на сравнении ОП с некоторым порогом. Различие между правилами, оптимальными по указанным критериям, состоит в различном выборе порога h . Таким образом, рассмотренные выше критерии приводят к единообразной процедуре принятия решения по наблюдаемой выборке $y_{t_0}^t$ в интервале (t_0, t)

$$\Lambda(y_{t_0}^t) \underset{\gamma_0}{\overset{\gamma_1}{>}} h.$$

4.4.4. ПОСЛЕДОВАТЕЛЬНЫЙ КРИТЕРИЙ ОЦЕНИВАНИЯ

Предыдущие критерии основываются на предположении, что объем наблюдаемой выборки $y_{t_0}^t (y_1^n)$ фиксирован. Однако время, затрачиваемое на принятие решения для этих случаев, может оказаться столь значительным, что вопрос о числе измерений становится открытым. Последовательный критерий или критерий Вальда позволяет объем измерений ставить в зависимости от того, когда наблюдатель убедится в правильности одной из гипотез. Если ранее использовались две подобласти в пространстве наблюдения, то при последовательных испытаниях вводится еще одна промежуточная область, в которой окончательное решение не принимается. При последовательных испытаниях принимается одно из трех решений: принять гипотезу γ_1 , принять альтернативу γ_0 , произвести следующее измерение. Как и в критерии Неймана - Пирсона выбираются приемлемые значения ошибок α и β . По результатам n измерений формируется ОП - $\Lambda(y_1^n)$. Полученное значение ОП сравнивается с двумя порогами h и h' . Если $\Lambda(y_1^n) < h$, то гипотеза γ_0 принимается. Если $\Lambda(y_1^n) \geq h'$ принимается решение в

пользу γ_1 . Если $h < \Lambda(y_1^n) < h'$, то делается следующее $n+1$ измерение и вычисляется новое значение ОП. Так продолжается до тех пор, пока не будет принята одна из гипотез. Вычисление этих порогов достаточно сложная задача. Вальд показал, что они определяются следующими неравенствами $h' \leq (1-\beta)/\alpha$, $h \geq \beta/(1-\alpha)$. Основное достоинство последовательного критерия состоит в том, что в среднем число измерений существенно уменьшается (минимизируется средний размер выборки), хотя в отдельных случаях время принятия решения может быть значительным.

4.5. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ

Ранее рассматривались точечные оценки параметра $\hat{x}_n = \hat{\gamma}(y_1^n)$ из непрерывного или конечного множеств для момента времени t_n . Поскольку оценки параметра при каждом опыте будут отличаться друг от друга, целесообразно также ввести интервальную оценку.

Под этой оценкой понимается интервал с нижней $\hat{\Theta}_n^H$ и верхней $\hat{\Theta}_n^B$ границами являющимися функциями выборки объема n , содержащий оцениваемый параметр $x = \Theta$ с заданной вероятностью α

$$P\{\hat{\Theta}_n^H < \Theta < \hat{\Theta}_n^B\} = \alpha. \quad (4.23)$$

Указанный интервал называется доверительным, α - коэффициентом доверия. Задача определения доверительного интервала может быть решена, если известно условное распределение точечной оценки параметра $\Theta - P(\hat{\Theta}_n | y_1^n)$.

Выразим указанные границы через положительный параметр ε

$$\hat{\Theta}_n^H = \hat{\Theta}_n - \varepsilon\Theta, \quad \hat{\Theta}_n^B = \hat{\Theta}_n + \varepsilon\Theta.$$

На основании полученных выражений соотношение (4.23) записывается в форме

$$P\{\Theta(1 - \varepsilon) < \hat{\Theta}_n < \Theta(1 + \varepsilon)\} = \alpha.$$

В результате можно получить формулу, связывающую относительную величину доверительного интервала

$2\varepsilon = (\hat{\Theta}_n^B - \hat{\Theta}_n^H) / \Theta$ с коэффициентом доверия α и объемом выборки n .

$$\int_{\Theta(1-\varepsilon)}^{\Theta(1+\varepsilon)} P(\hat{\Theta}_n | y_1^n) d\hat{\Theta} = \alpha.$$

Таким образом, при заданном значении величины доверительного интервала для достижения выбранного коэффициента доверия можно определить объем выборки. Аналогичным образом можно отыскать другие характеристики - доверительный интервал или коэффициент доверия по двум другим задаваемым характеристикам.

Наиболее разработаны методы получения доверительных оценок для гауссовского распределения оценки исследуемого параметра [39].

4.6. ОБЩИЕ МЕТОДЫ СИНТЕЗА СИСТЕМ ОБНАРУЖЕНИЯ И ФИЛЬТРАЦИИ

4.6.1. ФОРМУЛИРОВАНИЕ ЗАДАЧИ СИНТЕЗА

Методы синтеза базируются на алгоритмах последовательного (совместного) обнаружения и фильтрации случайных процессов. Указанный подход соответствует характеру работы реальной системы: сначала поиску (обнаружению), а затем оцениванию принимаемых сигналов. Аналитические способы решения этих задач опираются на один из важных научных разделов теории статистических решений - методах оптимальной фильтрации. Основная цель этих методов состоит в определении структур и соответствующих параметров оптимальных фильтров, обеспечивающих получение с наибольшей точностью оценок передаваемых сигналов при наличии случайных помех. Оценка является точечной, т.е. определяется для момента времени $t=t_1$, на основании всей доступной для наблюдения информации за время (t_0, t) . Причем последующие наблюдения не используются для улучшения ранее полученных оценок. Наряду с задачей фильтрации в технических приложениях существуют задачи экстраполяции (прогнозирования) и интерполяции (сглаживания), для которых

оценки определяются соответственно в моменты времени $t_1 > t$ и $t_0 < t_1 < t$.

Методы оптимальной фильтрации опираются на свойства стохастической наблюдаемости. Благодаря этим свойствам обеспечивается определение части ненаблюдаемых (оцениваемых) сигналов по другой связанной с ней доступной для наблюдения группе сигналов. Стохастическая наблюдаемость определяется асимптотическим поведением апостериорной плотности вероятностей или дисперсий ошибок оцениваемых процессов при увеличении числа наблюдений (или наблюдаемой реализации). Следствием этого поведения является сходимость по вероятности к конечному значению указанных ошибок. При этом возможны два типа задачи фильтрации. Для первого типа отыскиваются оценки квазидетерминированных процессов. Ошибки оценки в этом случае при неограниченном времени наблюдения стремятся к нулю. Указанному типу задач соответствует поведение систем с неслучайными выходными сигналами, измерения которых засорены помехами. Для другого типа задач оцениваемый на выходе системы процесс случаен и измерения также засорены помехами. Ошибки оценки здесь отличны от нуля и не превосходят заданной величины при неограниченном времени наблюдения.

Наиболее полно получила развитие теория линейной оптимальной фильтрации. В работе [40] показано, что при критерии минимума среднеквадратической ошибки и гауссовском распределении рассматриваемых случайных процессов оптимальная система фильтрации является линейной. Это означает также, что оптимальный алгоритм обработки результатов является линейным, а сам наблюдаемый сигнал связан с оцениваемым процессом линейной зависимостью. Важное значение для теории линейной фильтрации имели результаты работ Колмогорова [41] и Винера [42], получившие дальнейшее развитие в трудах отечественных и зарубежных исследователей. Согласно этим результатам импульсная переходная характеристика оптимального линейного фильтра на выходе которого отыскивается оценка процесса x , а на входе существует наблюдаемый сигнал $y = x + \lambda$, где λ - независимая от оцениваемого процесса x помеха, точно определяемая с помощью интегрального уравнения Винера-Хопфа через известные корреляционную функцию процесса x и взаимную корреляционную функцию x и наблюдаемого сигнала y .

Более простой алгоритм фильтрации можно получить, если отказаться от вышеуказанных корреляционных функций и представить оцениваемый процесс, как было показано в разделе 2.3.7, в виде решений системы стохастических дифференциальных уравнений. При этом оптимальный фильтр, называемый фильтром Калмана-Бьюси [43], полностью описывается дифференциальными или разностными уравнениями. Наличие уравнений указанного вида вместо интегральных существенно облегчает определение структуры оптимального фильтра. Необходимо также отметить, что с помощью фильтра Калмана-Бьюси могут быть получены оптимальные оценки для более широкого класса многомерных случайных процессов: с учетом ненулевых начальных условий и детерминированной составляющей сигнала, при наличии корреляционных связей между «небелой» помехой и оцениваемым процессом, при нестационарном режиме работы фильтра и т.д.

Для многих практических задач, где процессы не являются гауссовскими и, кроме того, подвергаются нелинейному преобразованию, оптимальная оценка может быть получена в классе нелинейных фильтров, т.е. оптимальный алгоритм обработки результатов наблюдения является нелинейным. Нахождение структур нелинейных фильтров, которым занимается теория нелинейной фильтрации, является достаточно сложной задачей. Наиболее существенные результаты здесь получены при рассмотрении оценки как условного марковского процесса и аддитивной белой гауссовской помехе. Результаты разработанной Р.Л. Стратоновичем [44] теории нелинейной фильтрации не противоречат данным теории линейной фильтрации и являются более общими по сравнению с ними. Это означает, что при гауссовском характере процессов, независимой аддитивной помехе и линейном преобразовании оцениваемого процесса нелинейная фильтрация вырождается в линейную, осуществляемую по критерию минимума среднеквадратической погрешности.

Задача синтеза формулируется следующим образом. Имеются априорные сведения о сигналах, помехах и способах их функционального взаимодействия. Для наблюдения (измерения) доступна векторная реализация случайного процесса $y_{t_0}^t$ на текущем интервале времени (t_0, t) , являющаяся детерминированной (неслучайной) функцией полезного сигнала $S_t(x)$, некоторой

детерминированной функцией управления $u(t)$, квантового шума $\chi(t)$ и помехи $\lambda(t)$

$$y(t) = \Psi_t \{ S_t(x_t), u(t), \chi(t), \lambda(t), \Theta \}. \quad (4.24)$$

Соотношение (4.24) получило название уравнение наблюдения. Оно описывает статистические процессы передачи информации по каналу связи. Удовлетворяющий ранее рассмотренному в разделе 2.2.2. условию регулярности полезный сигнал $S_t(x)$, представляет собой в общем случае нелинейную вектор-функцию времени t и многомерного марковского процесса $x(t)$. Марковский процесс может быть гауссовским и негауссовским, с дискретным и непрерывным временем, из конечного или непрерывного множества на действительной оси. Также известны вероятностные характеристики квантового шума $\chi(t)$ и помехи $\lambda(t)$. Параметр Θ - случайная величина с вероятностями $p_1(\Theta=1)$ и $p_0(\Theta=0)$, $p_1+p_0=1$; $\Theta=1$ или $\Theta=0$ характеризует соответственно наличие или отсутствие полезного сигнала.

На основании приведенных априорных сведений необходимо в каждый текущий момент времени t сформировать апостериорную плотность вероятностей (АПВ), позволяющую определить оценку вектора $\hat{x}(t)$ согласно одному из рассмотренных в разделе 4.1.2 критериев качества.

4.6.2. РЕКУРРЕНТНЫЕ УРАВНЕНИЯ ФИЛЬТРАЦИИ И ОБНАРУЖЕНИЯ

В данном разделе излагается метод определения рекуррентных алгоритмов для уравнения наблюдения^{*)}

$$y(t) = \Psi_t \{ S_t(x_t), \lambda(t), \Theta \}.$$

Решается задача синтеза для дискретного времени и одномерного марковского процесса, принимающего значения из непрерывного множества на действительной числовой оси.

Уравнение наблюдения для этого случая имеет вид

$$y_{n+1} = \Psi_{n+1} \{ S_{n+1}(x_{n+1}), \lambda_{n+1}, \Theta \}. \quad (4.25)$$

^{*)} Решения задачи синтеза с учетом вектора управления $u(t)$, а также квантового шума $\chi(t)$, обусловленного статистическим характером квантовых измерений, представлены соответственно в разделах 4.12 и 4.10.

Рассматривается на временном интервале (t_0, t_{n+1}) условная вероятность оцениваемого процесса в виде последовательности отсчетов $x_1^{n+1} = \{x_1, \dots, x_j, \dots, x_{n+1}\}$ в дискретные моменты времени t_j , $j = \overline{1, n+1}$ и случайной величины Θ при условии осуществления события $y_1^{n+1} = \{y_1, \dots, y_j, \dots, y_{n+1}\}$ - последовательности из $n+1$ отсчетов наблюдаемого сигнала в те же моменты времени - $P(x_1^{n+1}, \Theta | y_1^{n+1})$. Начальное значение марковского процесса x_0 , соответствующее моменту времени t_0 , удовлетворяет априорной плотности вероятностей $p(x_0)$.

На основании соотношения для условных вероятностей (1.4) можно записать

$$P(x_1^{n+1}, \Theta | y_1^{n+1}) = P(x_1^{n+1}, \Theta | y_{n+1}, y_1^n) = \frac{P(x_1^{n+1}, \Theta, y_{n+1} | y_1^n)}{P(y_{n+1} | y_1^n)}, \quad (4.26)$$

где условная вероятность в числителе (4.26) с учетом условия факторизации и независимости x_{n+1} от y_1^n имеет вид

$$p(x_1^{n+1}, \Theta, y_{n+1} | y_1^n) = \begin{cases} P(x_1^n, \Theta = 1 | y_1^n) \pi(x_{n+1} | x_n) p(y_{n+1} | x_1^{n+1}, y_1^n), \\ P(x_1^n, \Theta = 0 | y_1^n) \delta(x_{n+1}) p(y_{n+1} | y_1^n). \end{cases} \quad (4.27)$$

Введенная во второй строчке формулы (4.27) взамен $\pi(x_{n+1} | x_n)$ дельта-функция $\delta(x_{n+1})$ соответствует состоянию $\Theta=0 (S_{n+1}(x_{n+1})=0)$. В условной плотности вероятности $p(y_{n+1} | x_1^{n+1}, y_1^n)$ опустим y_1^n , так как при $\Theta=1$ сигнал y_{n+1} зависит только от x_1^{n+1} , а последовательность y_1^n не дает никакой дополнительной информации при наличии x_1^{n+1} . Кроме того, при $\Theta=0$ $p(y_{n+1} | y_1^n) = p(y_{n+1})$ в виду независимости отсчетов наблюдаемого сигнала в отсутствие оцениваемого процесса.

Запишем рассматриваемую условную плотность вероятностей с учетом высказанных замечаний

$$\begin{aligned}
p(y_{n+1} | x_1^{n+1}, y_1^n) &= p(y_{n+1} | x_1^{n+1}) = \frac{p(y_{n+1}, x_1^{n+1})}{p(x_1^{n+1})} = \\
&= \frac{p(x_1^n | y_{n+1}, x_{n+1}) p(y_{n+1} | x_{n+1}) p(x_{n+1})}{p(x_1^{n+1})} = p(y_{n+1} | x_{n+1}),
\end{aligned} \tag{4.28}$$

так как последовательность x_1^n не зависит от y_{n+1} , а $p(x_1^{n+1}) = p(x_1^n | x_{n+1}) p(x_{n+1})$. Таким образом, плотность вероятностей отсчета наблюдаемого сигнала y_{n+1} зависит от значения марковского процесса для того же момента времени x_{n+1} . Поскольку отсчет x_{n+1} в силу марковости процесса зависит только от предшествующего значения x_n , то наблюдаемый сигнал также является марковским процессом. Условная плотность вероятностей $p(y_{n+1} | x_{n+1})$ по определению является функцией правдоподобия (ФП) и характеризует статистические процессы передачи информации по каналу связи. После подстановки (4.28) в выражение (4.27), а затем – в (4.26), интегрирование обеих частей этого выражения по x_1, \dots, x_{n-1} получим

$$\begin{aligned}
P(x_{n+1}, x_n, \Theta | y_1^{n+1}) &= \frac{1}{p(y_{n+1})} \times \\
&\times \begin{cases} P(x_n, \Theta = 1 | y_1^n) \pi(x_{n+1} | x_n) p(y_{n+1} | x_{n+1}), \\ P(x_n, \Theta = 0 | y_1^n) \delta(x_{n+1}) p(y_{n+1}). \end{cases}
\end{aligned} \tag{4.29}$$

Представим соответствующие апостериорные вероятности в виде

$$P(x_n, \Theta | y_1^n) = \begin{cases} p(x_n | \Theta = 1, y_1^n) P(\Theta = 1 | y_1^n), & (4.30) \\ p(x_n | \Theta = 0, y_1^n) p(\Theta = 0 | y_1^n). & (4.31) \end{cases}$$

Обозначим $p(x_n | \Theta = 0, y_1^n) = \delta(x_n)$, а отношение апостериорных вероятностей в момент времени t_n

$$l_n = \frac{P(\Theta = 1 | y_1^n)}{P(\Theta = 0 | y_1^n)}.$$

Из равенства $P(\Theta=1|y_1^n)+P(\Theta=0|y_1^n)=1$ следует

$$P(\Theta = 1 | y_1^n) = \frac{l_n}{1 + l_n}, \quad (4.32)$$

$$P(\Theta = 0 | y_1^n) = \frac{1}{1 + l_n}. \quad (4.33)$$

После подстановки (4.32) и (4.33) в (4.30) и (4.31), а затем в (4.29) и последующего интегрирования по x_n , с учетом фильтрующих свойств дельта-функции $\delta(x_n)$ получим

$$\begin{aligned} P(x_{n+1}, \Theta | y_1^{n+1}) &= \\ &= \frac{1}{p(y_{n+1})} \begin{cases} \frac{l_n}{1 + l_n} \int_{-\infty}^{\infty} w_n(x_n) p(y_{n+1} | x_{n+1}) \pi(x_{n+1} | x_n) dx_n, \\ \frac{1}{1 + l_n} \delta(x_{n+1}) p(y_{n+1}). \end{cases} \end{aligned} \quad (4.34)$$

где $p(x_n | \Theta=1, y_1^n) = w_n(x_n)$ - апостериорная (условная) плотность вероятностей (АПВ) в момент времени t_n , описывает поведение условного марковского процесса.

Учитывая, что отношение апостериорных вероятностей в момент времени t_{n+1}

$$l_{n+1} = \frac{\int_{-\infty}^{\infty} P(x_{n+1}, \Theta = 1 | y_1^{n+1}) dx_{n+1}}{\int_{-\infty}^{\infty} P(x_{n+1}, \Theta = 0 | y_1^{n+1}) dx_{n+1}},$$

на основании выражения (4.34) получаем рекуррентный алгоритм отношения апостериорных вероятностей

$$l_{n+1} = \frac{l_n}{p(y_{n+1})} \int_{-\infty}^{\infty} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1},$$

где

$$\tilde{w}_n(x_{n+1}) = \int_{-\infty}^{\infty} \pi(x_{n+1} | x_n) w_n(x_n) dx_n \quad (4.35)$$

- экстраполированная плотность вероятностей в момент времени t_{n+1} . Так как отношения правдоподобия (ОП) и апостериорных вероятностей связаны между собой соотношением $\Lambda_{n+1} = \frac{p_0}{p_1} l_{n+1}$, приходим к рекуррентному алгоритму ОП

$$\Lambda_{n+1} = \frac{\Lambda_n}{p(y_{n+1})} \int_{-\infty}^{\infty} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1}. \quad (4.36)$$

Плотность вероятностей $p(y_{n+1})$ соответствует состоянию при отсутствии полезного сигнала ($\Theta=0$). ОП является положительной величиной и поэтому в расчетах вместо (4.36) используют логарифм ОП.

Для получения рекуррентного алгоритма фильтрации воспользуемся соотношением для условных плотностей вероятностей (1.4)

$$w_{n+1}(x_{n+1}) = p(x_{n+1} | \Theta = 1, y_1^{n+1}) = \frac{P(x_{n+1}, \Theta = 1 | y_1^{n+1})}{\int_{-\infty}^{\infty} P(x_{n+1}, \Theta = 1 | y_1^{n+1}) dx_{n+1}}. \quad (4.37)$$

После подстановки в (4.37) выражения (4.34) с учетом (4.35) получим рекуррентное уравнение АПВ

$$w_{n+1}(x_{n+1}) = c_{n+1} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}). \quad (4.38)$$

Входящая в уравнение (4.38) экстраполированная плотность вероятностей $\tilde{w}_n(x_{n+1})$ для момента времени t_{n+1} зависит от значения АПВ $w_n(x_n)$, взятого на предыдущем моменте времени t_n .

Нормировочная постоянная $c_{n+1} = \left[\int_{-\infty}^{\infty} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1} \right]^{-1}$ не зависит от состояния x_{n+1} . Значение ОП и АПВ на первом шаге обнаружения и фильтрации зависят от начальных условий

$$\Lambda_1 = \frac{\Lambda_0}{p(y_1)} \int_{-\infty}^{\infty} p(y_1 | x_1) \tilde{w}_0(x_1) dx_1,$$

$$w_1(x_1) = c_1 p(y_1 | x_1) \tilde{w}_0(x_1),$$

$$\text{где } \tilde{w}_0(x_1) = \int_{-\infty}^{\infty} \pi(x_1 | x_0) p(x_0) dx_0.$$

Полученные рекуррентные алгоритмы позволяют поэтапно шаг за шагом формировать, начиная с априорных данных (начальных условий), оценки марковского процесса x_n и статистики оптимального обнаружения $\ln \Lambda_n$. При этом, если алгоритм оценивания оптимальной фильтрации (4.38) можно применять независимо от алгоритма оптимального обнаружения (4.36), то последний невозможно использовать без алгоритма (4.38). Таким образом, система, реализующая рекуррентные алгоритмы (4.36) и (4.38) в принципе решает задачу совместного обнаружения и фильтрации случайных процессов. Суть метода решения этой задачи заключается в разбиении области возможных значений оцениваемого сигнала x_n на действительной оси на некоторое количество интервалов или каналов, вычисления значений АПВ для каждого из них и выбора такого канала (и соответственно значения x_n), в котором АПВ имеет максимальное значение. Реализация этой процедуры для задаваемых в обобщенном виде наблюдаемого сигнала (4.25) и функции правдоподобия $p(y_n | x_n)$ вызывает значительные трудности. В дальнейшем при конкретизации измерительной процедуры, уточнении функциональных зависимостей входящих в наблюдаемый процесс сигналов и помех и их статистических характеристик удастся сложную задачу свести к ряду более простых задач. В частности, если сигналы и помехи можно аппроксимировать гауссовскими распределениями, наблюдаемый процесс и оцениваемый сигнал связаны линейной зависимостью, а помеха входит в уравнение наблюдения аддитивно, то можно достаточно просто получить точные решения задач обнаружения и фильтрации. Кроме того, в отличие от рассматриваемого в разд. 4.3 и 4.4 задач обнаружения здесь отсутствует требование некоррелированности (независимости) отсчетов наблюдаемого сигнала: статистика ОП определяется рекуррентным алгоритмом (4.36).

Обратим внимание на то, что, во-первых, наблюдаемый сигнал в рекуррентные выражения входит через ФП $p(y_n|x_n)$. При конкретизации задач синтеза задание типа ФП определяется измерительной процедурой, которая в свою очередь, задается уравнением наблюдения. Во-вторых, ввиду того, что вероятность правильного обнаружения меньше единицы возникает возможность получать так называемые псевдооценки в отсутствие полезного сигнала.

Структурная схема оптимальной системы совместного обнаружения и фильтрации, описываемая алгоритмами (4.36) и (4.38), представлена на Рис. 4.1. В связи с рекуррентным характером этих алгоритмов ряд блоков схемы охвачен обратной связью с задержкой на интервал дискретизации (БЗ). Устройство включения (УВ) в конечный момент времени t_m , определяемый алгоритмом обнаружения, пропускает вычисленную в блоке формирования отношения правдоподобия (БФОП) статистику $\ln\Lambda_m$ на пороговое устройство (ПУ). В нем в результате сравнения $\ln\Lambda_m$ с порогом в соответствии с алгоритмом (4.16) выносится решение о наличии или отсутствии полезного сигнала. Величина порога, как отмечалось в разделе 4.4.3, зависит от типа оптимального обнаружителя. После превышения порога в ПУ (решение γ_1) с выхода ключевого устройства (КУ) поступает из блока фильтрации (БФ) оценка (если γ_1 верно) или псевдооценка (если γ_1 ложно).

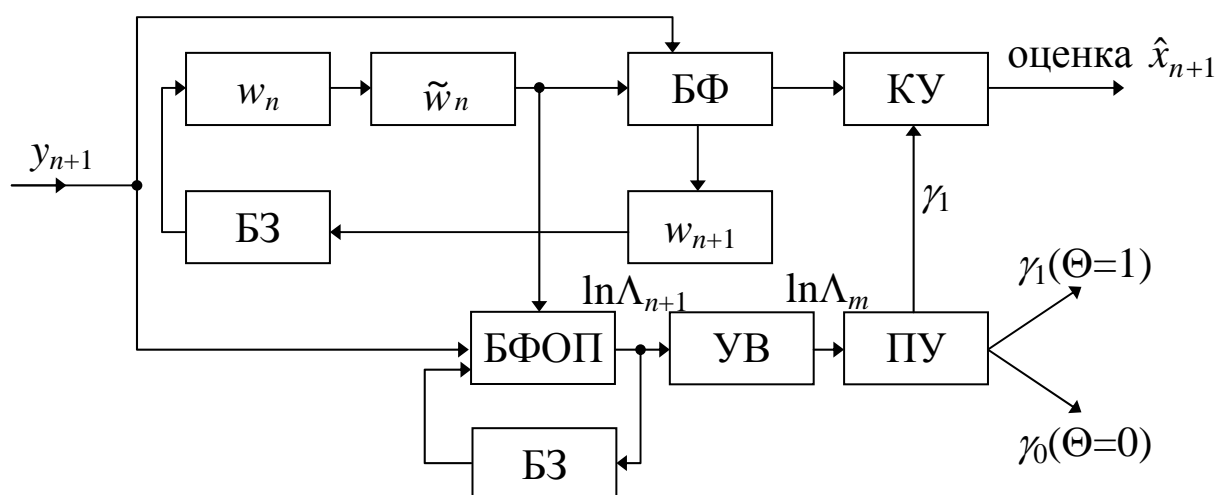


Рис. 4.1. Структурная схема системы совместного обнаружения и фильтрации.

4.6.3. УРАВНЕНИЕ ДЛЯ АПОСТЕРИОРНЫХ ПАРАМЕТРОВ

В рамках дальнейшей конкретизации методов фильтрации одномерного марковского сигнала в дискретном времени перейдем к отысканию рекуррентных уравнений для апостериорных параметров. Эти параметры представляют, как было отмечено в разделе 4.2, логарифм АПВ с определенной степенью точности (до K -го члена аппроксимирующего ряда). Отыскиваются следующие апостериорные параметры

$$\hat{x}_{n+1} = \int_{-\infty}^{\infty} x_{n+1} w_{n+1}(x_{n+1}) dx_{n+1} \quad (4.39)$$

- оценка марковского процесса (апостериорное среднее),

$$D_{n+1} = (h_{2,n+1})^{-1} = \int_{-\infty}^{\infty} (x_{n+1} - \hat{x}_{n+1})^2 w_{n+1}(x_{n+1}) dx_{n+1} \quad (4.40)$$

- апостериорная дисперсия ошибки оценки,

$$(h_{i,n+1})^{-1} = \int_{-\infty}^{\infty} (x_{n+1} - \hat{x}_{n+1})^i w_{n+1}(x_{n+1}) dx_{n+1}, \quad i \geq 3 \quad (4.41)$$

- апостериорные центральные моменты третьего и более высокого порядка.

Полагаем, что апостериорная плотность вероятностей АПВ (4.38) и экстраполированная плотность вероятностей ЭПВ (4.35) соответственно аппроксимируется выражениями

$$w_{n+1}(x_{n+1}) = A_1 \exp \left\{ - \sum_{i=2}^K \frac{h_{i,n+1}}{i!} (x_{n+1} - \hat{x}_{n+1})^i \right\}, \quad (4.42)$$

$$\tilde{w}_n(x_{n+1}) = A_2 \exp \left\{ - \sum_{i=2}^K \frac{g_{i,n+1}}{i!} (x_{n+1} - \tilde{x}_{n+1})^i \right\}, \quad (4.43)$$

где A_1 и A_2 - постоянные; K - заданный порядок аппроксимации; \hat{x}_{n+1} и \tilde{x}_{n+1} - оптимальные оценки одномерного марковского процесса, соответствующие максимумам АПВ и ЭВП на шаге $n+1$.

Коэффициенты разложений (4.42) и (4.43) вычисляются из соотношений:

$$h_{i,n+1} = - \frac{\partial^i \ln w_{n+1}(x_{n+1})}{\partial x_{n+1}^i} \Big|_{x_{n+1} = \hat{x}_{n+1}},$$

$$g_{i,n+1} = - \frac{\partial^i \ln \tilde{w}_{n+1}(x_{n+1})}{\partial x_{n+1}^i} \Big|_{x_{n+1} = \tilde{x}_{n+1}}.$$

Прологарифмируем, а затем продифференцируем выражения (4.38) по x_{n+1} K раз. После замены $x_{n+1} = \tilde{x}_{n+1}$ получаем

$$\begin{aligned} & \frac{\partial^i \ln w_{n+1}(x_{n+1})}{\partial x_{n+1}^i} \Big|_{x_{n+1} = \tilde{x}_{n+1}} = \\ & = \frac{\partial^i \ln \tilde{w}_n(x_{n+1})}{\partial x_{n+1}^i} \Big|_{x_{n+1} = \tilde{x}_{n+1}} + \frac{\partial^i \Pi_{n+1}(x_{n+1})}{\partial x_{n+1}^i} \Big|_{x_{n+1} = \tilde{x}_{n+1}}, \quad (4.44) \\ & i = \overline{1, K} \end{aligned}$$

где $\Pi_{n+1}(x_{n+1}) = \ln p(y_{n+1} | x_{n+1})$ - логарифм ФП.

В результате подстановки (4.42) и (4.43) в выражение (4.44) приходим к усеченной до K порядка включительно системе уравнений относительно апостериорных параметров (обратных апостериорных центральных моментов) разного порядка

$$- \sum_{i=2}^K \frac{h_{i,n+1}}{(i-1)!} (\tilde{x}_{n+1} - \hat{x}_{n+1})^{i-1} = \Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) \quad (4.45)$$

после первого дифференцирования,

$$\begin{aligned} h_{i,n+1} &= g_{i,n+1} - \Pi_{n+1}^{(i)}(\tilde{x}_{n+1}) - \\ & - \sum_{j=1}^{K-i} \frac{(\tilde{x}_{n+1} - \hat{x}_{n+1})^j}{j!} [g_{i+j,n+1} - \Pi_{n+1}^{(i+j)}(\tilde{x}_{n+1})], \quad i = \overline{2, K} \end{aligned} \quad (4.46)$$

после второго и всех последующих дифференцирований, где введено обозначение

$$\left. \frac{\partial^i \Pi_{n+1}(x_{n+1})}{\partial x_{n+1}^i} \right|_{x_{n+1} = \tilde{x}_{n+1}} = \Pi_{n+1}^{(i)}(\tilde{x}_{n+1}).$$

После K -ого дифференцирования параметр $h_{K,n+1}$ принимает вид

$$h_{K,n+1} = g_{K,n+1} - \Pi_{n+1}^{(K)}(\tilde{x}_{n+1}). \quad (4.47)$$

По определению параметр $h_{1,n+1}$ равен нулю и полученная оценка (апостериорное среднее), как показано в разделе 4.2, соответствует максимуму апостериорной плотности вероятности.

Система уравнений (4.45) и (4.46) является рекуррентной, так как позволяет определить $\hat{x}_{n+1}, h_{2,n+1}, \dots, h_{K,n+1}$ по значениям $\tilde{x}_{n+1}, g_{2,n+1}, \dots, g_{K,n+1}$ зависящим от апостериорных центральных моментов на предыдущем шаге (момента времени t_n). Обратим также внимание на то, что в уравнение (4.46) для параметра i -го порядка входят экстраполированные параметры более высоких $(K-i)$ порядков.

В представленных моделирующих задачу фильтрации уравнениях присутствуют экстраполированные оценка \tilde{x} и параметры g_i . Для решения уравнений (4.45) и (4.46) необходимо получить явную зависимость апостериорных значений оценки \hat{x} и центральных моментов разного порядка h_i^{-1} от указанных экстраполированных параметров. Для решения этой задачи в качестве исходных соотношений используются

$$\tilde{x}_{n+1} = \int_{-\infty}^{\infty} x_{n+1} \tilde{w}_n(x_{n+1}) dx_{n+1} \quad (4.48)$$

- экстраполированная оценка,

$$\tilde{D}_{n+1} = (g_{2,n+1})^{-1} = \int_{-\infty}^{\infty} (x_{n+1} - \tilde{x}_{n+1})^2 \tilde{w}_n(x_{n+1}) dx_{n+1} \quad (4.49)$$

- экстраполированная дисперсия,

$$(g_{i,n+1})^{-1} = \int_{-\infty}^{\infty} (x_{n+1} - \tilde{x}_{n+1})^i \tilde{w}_n(x_{n+1}) dx_{n+1}, \quad i \geq 3 \quad (4.50)$$

- экстраполированные центральные моменты третьего и более высокого порядка.

Ввиду того, что далее задачи синтеза решаются в рамках корреляционной теории, ограничимся определением экстраполированных значений оценки и дисперсии ошибки оценки.

Входящая в выражения (4.48), (4.49) и (4.50) экстраполированная плотность $\tilde{w}_n(x_{n+1})$ была получена (4.35) и зависит от переходных вероятностей $\pi(x_{n+1}|x_n)$ и АПВ на шаге n . Воспользуемся соотношением (2.40), которое для дискретного времени и при пренебрежении членами порядка $0(\Delta t)$ имеет вид

$$\pi(x_{n+1}|x_n) = \delta(x_{n+1} - x_n) + \Delta t L\{\delta(x_{n+1} - x_n)\}. \quad (4.51)$$

Рассматриваемый подлежащий оценке одномерный марковский процесс описывается в дискретном времени уравнением в конечных разностях (2.53)

$$x_{n+1} = x_n + f_n(x_n)\Delta t + \gamma_n \eta_n, \quad x_0 = x(0).$$

С учетом выражения для оператора ФПК (2.37) запишем соотношение (4.51) в развернутом виде

$$\begin{aligned} \pi(x_{n+1}|x_n) &= \delta(x_{n+1} - x_n) - \Delta t \frac{\partial}{\partial x_{n+1}} [f_n(x_n) \delta(x_{n+1} - x_n)] + \\ &+ \frac{\gamma_n^2}{2} Q_{\eta} \frac{\partial^2}{\partial x_{n+1}^2} \delta(x_{n+1} - x_n). \end{aligned}$$

После подстановки этого соотношения в выражение (4.35), последующих вычислений с учетом фильтрующих свойств дельта-функции получаем

$$\begin{aligned} \tilde{w}_n(x_{n+1}) &= w_n(x_{n+1}) - \Delta t \frac{\partial}{\partial x_{n+1}} [f_n(x_{n+1}) w_n(x_{n+1})] + \\ &+ \frac{\gamma_n^2}{2} Q_{\eta} \frac{\partial^2 w_n(x_{n+1})}{dx_{n+1}^2}. \end{aligned} \quad (4.52)$$

Экстраполированная оценка вычисляется в результате подстановки выражения (4.52) в формулу (4.48) и интегрирования по частям

$$\tilde{x}_{n+1} = \int_{-\infty}^{\infty} x_{n+1} \tilde{w}_n(x_{n+1}) dx_{n+1} = \hat{x}_n + \Delta t M\{f_n(x_{n+1}) | y_1^n\}, \quad (4.53)$$

где введено обозначение

$$\int_{-\infty}^{\infty} f_n(x_{n+1}) w_n(x_{n+1}) dx_{n+1} = M\{f_n(x_{n+1}) | y_1^n\}, \quad \text{а промежуточный}$$

член $x_{n+1} f_n(x_{n+1}) w_n(x_{n+1}) \Big|_{-\infty}^{\infty}$ приравнен нулю при неограниченном изменении аргумента.

Для определения экстраполированной дисперсии используется формула (4.49). После подстановки в нее соотношения (4.52), интегрирования по частям и приравнивания нулю промежуточных членов вида $x_{n+1} f_n(x_{n+1}) \frac{\partial w_n(x_{n+1})}{\partial x_{n+1}} \Big|_{-\infty}^{\infty}$, $x_{n+1}^2 f_n(x_{n+1}) w_n(x_{n+1}) \Big|_{-\infty}^{\infty}$, с учетом выражения (4.53) имеем

$$\begin{aligned} \tilde{D}_{n+1} &= \int_{-\infty}^{\infty} (x_{n+1} - \tilde{x}_{n+1})^2 \tilde{w}_n(x_{n+1}) dx_{n+1} = \\ &= \int_{-\infty}^{\infty} [(x_{n+1} - \hat{x}_n) - \Delta t M\{f_n(x_{n+1}) | y_1^n\}]^2 \tilde{w}_n(x_{n+1}) dx_{n+1} = \\ &= D_n + 2\Delta t M\{(x_{n+1} - \hat{x}_n) f_n(x_{n+1}) | y_1^n\} + \gamma_n^2 Q_m, \end{aligned} \quad (4.54)$$

где $D_n = M\{x_n^2 | y_1^n\} - [M\{x_n | y_1^n\}]^2 = \overset{\wedge}{x_n^2} - (\hat{x}_n)^2$ - дисперсия ошибки оценки на шаге n .

Выражения (4.53) и (4.54) можно рассматривать как результат экстраполяции (эволюции априорных данных согласно уравнению ФПК) на один шаг полученных на предыдущем шаге оценки марковского процесса и дисперсии ошибки оценки в отсутствие наблюдения.

Задача обнаружения одномерного марковского процесса решается после подстановки плотности вероятностей $\tilde{w}_n(x_{n+1})$ в выражение (4.36) и конкретизации вида ФП.

Решение в общем виде уравнений фильтрации и обнаружения представляют собой, как правило, чрезвычайно сложную задачу и определяется видом ФП и входящими в эти уравнения нелинейными функциями. В дальнейшем за счет некоторой потери оптимальности в результате упрощений полученных алгоритмов приходят к более простым приближенным методам решения задач синтеза.

4.6.4. ОБЩИЕ АЛГОРИТМЫ СИНТЕЗА ДЛЯ МНОГОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

При выводе уравнений фильтрации и обнаружения для многомерного марковского процесса используются векторно-матричные уравнения оцениваемого процесса

$$x_{n+1} = x_n + \Delta t f_n(x_n) + \Gamma_n \eta_n \quad (4.55)$$

и наблюдаемого сигнала

$$y_{n+1} = \Psi_{n+1}(S_{n+1}(x_{n+1}), \lambda_{n+1}, \Theta), \quad (4.56)$$

где x_n и $f_n(\cdot)$ соответственно вектор и вектор-функция оцениваемого процесса размера $(r \times 1)$; Γ_n - диагональная матрица размера $(r \times r)$, y_{n+1} - вектор-столбец наблюдаемых сигналов размера $(q \times 1)$; $S_n(\cdot)$ - детерминированная вектор-функция размера $(q \times 1)$, описывающая нелинейное преобразование вектора x ; η_n и λ_n - вектор-столбцы дискретных белых шумов размеров соответственно $(r \times 1)$ и $(p \times 1)$; Θ - вектор размера $(r \times 1)$ с компонентами $\Theta_i = 1$ или $\Theta_i = 0$, характеризующими наличие или отсутствие соответствующих компонент полезного сигнала. Кроме того, имеют место следующие соотношения $q \leq r$ и $p \leq q$. Первое неравенство отражает тот факт, что не все фазовые координаты доступны для наблюдений (измерений), второе учитывает, что не каждое измерение засорено помехами.

Рассматривается на временном интервале (t_0, t) последовательность отсчетов векторов оцениваемых процессов и наблюдаемых сигналов

$$x_1^{n+1} = \{x_1, \dots, x_j, \dots, x_{n+1}\},$$

$$y_1^{n+1} = \{y_1, \dots, y_j, \dots, y_{n+1}\},$$

где $x_j = \|x_{1j} \dots x_{rj}\|^T$, $y_j = \|y_{1j} \dots y_{qj}\|^T$, $j = \overline{1, n+1}$

Используя методику, аналогичную при выводе уравнений АПВ и ОП для одномерного процесса, можно показать, что указанные статистические характеристики определяются рекуррентными соотношениями

$$w_{n+1}(x_{n+1}) = c_{n+1} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) \quad (4.57)$$

$$\Lambda_{n+1} = \frac{\Lambda_n}{p(y_{n+1})} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1}, \quad (4.58)$$

$$\text{где } c_{n+1} = \left[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(y_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1} \right]^{-1},$$

$$\tilde{w}_n(x_{n+1}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \pi(x_{n+1} | x_n) w_n(x_n) dx_n.$$

Как и в случае одномерных процессов, апостериорные параметры вычисляют на основании представления логарифмов многомерных АПВ и ЭПВ в виде конечной суммы разложения в ряд Тейлора относительно параметров h и g . В рамках корреляционной теории рассматривается разложение в ряд с порядком аппроксимации $K=2$, что соответствует получению приближенных оценок x^* . Имеем

$$\begin{aligned} w_{n+1}(x_{n+1}) &= A_1 \exp \left\{ -\frac{1}{2} (x_{n+1} - x_{n+1}^*)^T H_{n+1} (x_{n+1} - x_{n+1}^*) \right\} = \\ &= A_1 \exp \left\{ -\frac{1}{2} \sum_{i=1}^r \sum_{j=1}^r h_{ij,n+1} (x_{i,n+1} - x_{i,n+1}^*) (x_{j,n+1} - x_{j,n+1}^*) \right\}, \\ \tilde{w}_{n+1}(x_{n+1}) &= A_2 \exp \left\{ -\frac{1}{2} (x_{n+1} - \tilde{x}_{n+1})^T G_{n+1} (x_{n+1} - \tilde{x}_{n+1}) \right\} = \\ &= A_2 \exp \left\{ -\frac{1}{2} \sum_{i=1}^r \sum_{j=1}^r g_{ij,n+1} (x_{i,n+1} - \tilde{x}_{i,n+1}) (x_{j,n+1} - \tilde{x}_{j,n+1}) \right\}, \end{aligned}$$

$$\text{где } h_{ij,n+1} = -\frac{\partial^2 \ln w_{n+1}(x_{n+1})}{\partial x_{i,n+1} \partial x_{j,n+1}} \left| \begin{array}{l} x_{i,n+1} = x_{i,n+1}^* \\ x_{j,n+1} = x_{j,n+1}^* \end{array} \right.,$$

$$g_{ij,n+1} = -\frac{\partial^2 \ln \tilde{w}_{n+1}(x_{n+1})}{\partial x_{i,n+1} \partial x_{j,n+1}} \left| \begin{array}{l} x_{i,n+1} = \tilde{x}_{i,n+1} \\ x_{j,n+1} = \tilde{x}_{j,n+1} \end{array} \right.$$

H_{n+1} и G_{n+1} - положительно определенные матрицы размера $(r \times r)$, компоненты которых соответственно равны $h_{ij,n+1}$ и $g_{ij,n+1}$; x_i^* и \tilde{x}_i - компоненты векторов апостериорных и экстраполированных оценок, соответствующих максимумам АПВ и ЭПВ.

Прологарифмируем соотношение (4.57), а затем дважды продифференцируем по x_{n+1} . В результате первого дифференцирования после подстановки $x_{n+1} = \tilde{x}_{n+1}$ получаем

$$H_{n+1}(x_{n+1}^* - \tilde{x}_{n+1}) = \Pi_{n+1}^{(1)}(\tilde{x}_{n+1}). \quad (4.59)$$

После второго дифференцирования и подстановки $x_{n+1} = \tilde{x}_{n+1}$ имеем

$$H_{n+1} = G_{n+1} - \Pi_{n+1}^{(2)}(\tilde{x}_{n+1}), \quad (4.60)$$

где $\Pi_{n+1}^{(1)}(\tilde{x}_{n+1})$ - r -мерный вектор - столбец, i - компонента которого $\partial \Pi_{n+1}(\tilde{x}_{n+1}) / d\tilde{x}_{i,n+1}$; $\Pi_{n+1}^{(2)}(\tilde{x}_{n+1})$ - матрица размера $(r \times r)$, (i,j) -компонента которой $\partial^2 \Pi_{n+1}(\tilde{x}_{n+1}) / d\tilde{x}_{i,n+1} d\tilde{x}_{j,n+1}$.

Преобразуем полученные уравнения в удобную для дальнейшего использования форму. Умножим соотношение (4.59) слева на обратную матрицу H_{n+1}^{-1} . Учитывая, что $H_{n+1}^{-1} H_{n+1} = I$ (I - единичная матрица), получим рекуррентное уравнение оценок в векторной форме.

$$x_{n+1}^* = \tilde{x}_{n+1} + D_{n+1} \Pi_{n+1}^{(1)}(\tilde{x}_{n+1}), \quad x_0^* = x(0). \quad (4.61)$$

где $D_{n+1}=H_{n+1}^{-1}$ - матрица дисперсии ошибок, (i,j) -компонента которой $D_{ij,n+1}=H_{ij,n+1}/|H_{n+1}|$, H_{ij} - алгебраическое дополнение компоненты h_{ij} определителя $|H_{n+1}|$.

Для определения матрицы дисперсий ошибок умножим справа (4.60) на $[G_{n+1}-\Pi_{n+1}^{(2)}(\tilde{x}_{n+1})]^{-1}$, а затем полученное выражение слева на D_{n+1} . Учитываем $D_{n+1}=H_{n+1}^{-1}$, а также $\tilde{D}_{n+1}^{-1}=G_{n+1}$, где \tilde{D}_{n+1}^{-1} - матрица экстраполированных дисперсий ошибок, (ij) -компонента которой $\tilde{D}_{ij,n+1}=G_{ij,n+1}/|G_{n+1}|$, $G_{ji,n+1}$ - алгебраическое дополнение элемента g_{ij} определителя $|G_{n+1}|$. В результате получим рекуррентное уравнение дисперсий ошибок в векторной форме

$$D_{n+1}=[\tilde{D}_{n+1}^{-1}-\Pi_{n+1}^{(2)}(\tilde{x}_{n+1})]^{-1}, \quad D_0=D(0). \quad (4.62)$$

Необходимые для решения уравнений фильтрации и обнаружения экстраполированные вектор \tilde{x}_{n+1} и матрица дисперсий определяются из выражений

$$\tilde{x}_{n+1}=\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_{n+1} \tilde{w}_n(x_{n+1}) dx_{n+1}, \quad (4.63)$$

$$\tilde{D}_{n+1}=\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_{n+1}-\tilde{x}_{n+1})(x_{n+1}-\tilde{x}_{n+1})^T \tilde{w}_n(x_{n+1}) dx_{n+1}. \quad (4.64)$$

Используя методику получения этих статистических характеристик для одномерных процессов, приходят к следующим соотношениям в векторной форме

$$\tilde{x}_{n+1}=x_n^*+\Delta t M\{f_n(x_{n+1})|y_1^n\}, \quad (4.65)$$

$$\begin{aligned} \tilde{D}_{n+1}= & D_n+\Delta t M\{f_n(x_{n+1})(x_{n+1}-x_{n+1}^*)^T|y_1^n\}+ \\ & +\Delta t M\{(x_{n+1}-x_{n+1}^*)f_n^T(x_{n+1})|y_1^n\}+\Gamma_n Q_{\eta n} \Gamma_n^T. \end{aligned} \quad (4.66)$$

4.7. ЛИНЕЙНЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ И ФИЛЬТРАЦИИ

4.7.1. АЛГОРИТМЫ СИНТЕЗА ДЛЯ ОДНОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

Точное решение задачи синтеза для одномерного (скалярного) марковского процесса можно получить если наблюдаемый y_n и полезный $S_n(x)$ сигналы линейны относительно оцениваемого гауссовского марковского процесса x_n , начальное значение этого процесса x_0 также имеет гауссовское распределение. Кроме того, дискретный белый шум λ_n входит в уравнение наблюдения аддитивно (в виде слагаемой составляющей) и является гауссовским процессом с независимыми значениями.

При выводе уравнений фильтрации и обнаружения, исходными являются уравнения марковского процесса (2.54)

$$x_{n+1} = \Phi_n x_n + \gamma_n \eta_n, \quad x_0 = x(0) \quad (4.67)$$

и наблюдаемого сигнала (4.25)

$$y_{n+1} = \Theta C_{n+1} x_{n+1} + \lambda_{n+1}, \quad S_{n+1}(x_{n+1}) = C_{n+1} x_{n+1} \quad (4.68)$$

в дискретном времени, где $\Phi_n = 1 + \Delta t f_n$, дискретные белые шумы η_{n+1} и λ_{n+1} некоррелированы^{*)}.

Функция правдоподобия на основании (4.68) и плотность вероятностей наблюдаемого сигнала в отсутствии полезного сигнала соответственно равны

$$p(y_{n+1} | x_{n+1}) = (2\pi Q_{\lambda, n+1})^{-1/2} \exp \left\{ -\frac{(y_{n+1} - C_{n+1} x_{n+1})^2}{2Q_{\lambda, n+1}} \right\} \quad (4.69)$$

$$p(y_{n+1}) = (2\pi Q_{\lambda, n+1})^{-1/2} \exp \left\{ -\frac{y_{n+1}^2}{2Q_{\lambda, n+1}} \right\}, \quad (4.70)$$

где $Q_{\lambda, n+1} = \mathcal{N}_{\lambda, n+1} / \Delta t$ - дисперсия дискретного белого шума, \mathcal{N}_{λ} - спектральная плотность белого шума λ .

^{*)} В дальнейшем обозначение Θ не используется, так как предполагается, что полезный сигнал присутствует в наблюдении.

Ввиду того, что все процессы в рассматриваемой задаче синтеза являются гауссовскими, АПВ (4.42) также описывается гауссовским распределением с параметрами \hat{x}_{n+1} и $h_{2,n+1}$. Таким образом, указанная задача сводится к вычислению этих параметров на основании дискретных отсчетов наблюдаемого сигнала $\{y_{n+1}\}$ с помощью рекуррентных алгоритмов (4.45) и (4.46), которые при $K=2$ принимают форму

$$\hat{x}_{n+1} = \tilde{x}_{n+1} + D_{n+1} \Pi_{n+1}^{(1)}(\tilde{x}_{n+1}), \quad \hat{x}_0 = x(0), \quad (4.71)$$

$$\begin{aligned} D_{n+1} &= (h_{2,n+1})^{-1} = \\ &= [g_{2,n+1} + \Pi_{n+1}^{(2)}(\tilde{x}_{n+1})]^{-1} = [\tilde{D}_{n+1}^{-1} + \Pi_{n+1}^{(2)}(\tilde{x}_{n+1})]^{-1}, \quad D_0 = D(0) \end{aligned} \quad (4.72)$$

Так как $f_n(x_n) = f_n x_n$, то на основании формул (4.54) и (4.55) получим экстраполированные параметры

$$\begin{aligned} \tilde{x}_{n+1} &= \Phi_n \hat{x}_n, \\ \tilde{D}_{n+1} &= D_n + 2\Delta t f_n D_n + \gamma_n^2 Q_{\eta n}. \end{aligned}$$

В результате вычислений $\Pi_{n+1}^{(1)}(\tilde{x}_{n+1})$ и $\Pi_{n+1}^{(2)}(\tilde{x}_{n+1})$ на основании (4.69) имеем

$$\begin{aligned} \Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) &= \left. \frac{\partial \ln p(y_{n+1} | x_{n+1})}{\partial x_{n+1}} \right|_{x_{n+1} = \tilde{x}_{n+1}} = \\ &= C_{n+1} Q_{\lambda, n+1}^{-1} (y_{n+1} - C_{n+1} \tilde{x}_{n+1}), \end{aligned} \quad (4.73)$$

$$\begin{aligned} \Pi_{n+1}^{(2)}(\tilde{x}_{n+1}) &= \left. \frac{\partial^2 \ln p(y_{n+1} | x_{n+1})}{\partial x_{n+1}^2} \right|_{x_{n+1} = \tilde{x}_{n+1}} = \\ &= -C_{n+1}^2 Q_{\lambda, n+1}^{-1} \end{aligned} \quad (4.74)$$

Подставляя соотношения (4.73) и (4.74) соответственно в выражения (4.71) и (4.72), получаем рекуррентные уравнения фильтрации одномерного гауссовского марковского процесса в дискретном времени

$$\hat{x}_{n+1} = \tilde{x}_{n+1} + D_{n+1} C_{n+1} Q_{\lambda, n+1}^{-1} (y_{n+1} - C_{n+1} \tilde{x}_{n+1}), \quad \hat{x}_0 = x(0), \quad (4.75)$$

$$\begin{aligned} D_{n+1} &= [\tilde{D}_{n+1}^{-1} + C_{n+1}^2 Q_{\lambda, n+1}^{-1}]^{-1} = \\ &= \tilde{D}_{n+1} - \frac{\tilde{D}_{n+1}^2 C_{n+1}^2}{\tilde{D}_{n+1} C_{n+1}^2 + Q_{\lambda, n+1}}, \quad D_0 = D(0). \end{aligned} \quad (4.76)$$

Линейное уравнение (4.75) называется уравнением оценки, а нелинейное уравнение (4.76) - уравнением дисперсии ошибки оценки. Как следует из приведенных уравнений фильтрации, при линейном преобразовании гауссовского марковского процесса и аддитивной независимой помехе уравнение оценки линейно относительно самой оценки, а нелинейное уравнение дисперсии не зависит от наблюдаемого сигнала, оценки и решается независимо от уравнения оценки.

Перейдем к определению рекуррентного алгоритма обнаружения для рассматриваемой задачи. После подстановки в выражение (4.36) плотностей вероятностей (4.69) и (4.70) и сокращения общих сомножителей в числителе и знаменателе ОП получаем

$$\begin{aligned} \Lambda_{n+1} &= \\ &= \Lambda_n \int_{-\infty}^{\infty} \exp \left\{ \frac{\Delta t}{2\mathcal{N}_{\lambda, n+1}} [2C_{n+1} y_{n+1} x_{n+1} - C_{n+1}^2 x_{n+1}^2] \right\} \tilde{w}_n(x_{n+1}) dx_{n+1}. \end{aligned} \quad (4.77)$$

Разложим экспоненциальный член выражения (4.77) в ряд по степеням Δt , имеем

$$\begin{aligned} \Lambda_{n+1} &= \Lambda_n + \Lambda_n \int_{-\infty}^{\infty} \left\{ \frac{\Delta t}{2\mathcal{N}_{\lambda, n+1}} [2C_{n+1} y_{n+1} x_{n+1} - C_{n+1}^2 x_{n+1}^2] + \right. \\ &\quad \left. + \frac{(\Delta t y_{n+1})^2}{2\mathcal{N}_{\lambda, n+1}^2} C_{n+1}^2 x_{n+1}^2 \right\} \tilde{w}_n(x_{n+1}) dx_{n+1} + o(\Delta t), \end{aligned} \quad (4.78)$$

где $o(\Delta t)$ - члены порядка малости $(\Delta t)^2$ и выше.

В фигурных скобках выражения (4.78) содержится слагаемое, зависящее от $(\Delta t)^2$. Это связано с тем, что множитель $(\Delta t y_{n+1})^2 \rightarrow (\Delta t \lambda_{n+1})^2$, а $(\Delta t \lambda_{n+1})^2 \sim (\Delta B)^2$, где ΔB - приращение винеровского процесса. При малом Δt этот множитель на основании данных раздела 2.2.1 может быть заменен на математическое ожидание квадрата этого приращения $\mathcal{N}_\lambda \Delta t$.

Таким образом в выражение (4.78) указанное слагаемое входит зависящим от первой степени приращения Δt : $\Delta t C_{n+1}^2 x_{n+1}^2 / 2 \mathcal{N}_{\lambda, n+1}$.

В результате получаем искомый рекуррентный алгоритм ОП в форме

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n C_{n+1} y_{n+1} \tilde{x}_{n+1} Q_{\lambda, n+1}^{-1}, \quad \Lambda_0 = \Lambda(0). \quad (4.79)$$

Полученные в данном разделе значения \hat{x}_{n+1} и Λ_{n+1} являются достаточными статистиками линейной задачи обнаружения и фильтрации.

4.7.2. АЛГОРИТМЫ СИНТЕЗА ДЛЯ МНОГОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

Векторно-матричное уравнение марковского процесса (4.55) и наблюдаемого сигнала (4.56) записывается в форме

$$x_{n+1} = \Phi_n x_n + \Gamma_n \eta_n, \quad x_0 = x(0), \quad (4.80)$$

$$y_{n+1} = C_{n+1} x_{n+1} + \lambda_{n+1}, \quad S_{n+1}(x_{n+1}) = C_{n+1} x_{n+1}, \quad (4.81)$$

где C_{n+1} - матрица коэффициентов размера $(q \times r)$. Остальные вектора и матрицы рассматривались в разделе 4.6.4. Вектора дискретных шумов η_n и λ_n полагают независимыми, а начальные условия x_0 принадлежащими гауссовскому распределению. Многомерная функция правдоподобия с учетом вида наблюдаемого сигнала (4.81) описывается гауссовским распределением

$$p(y_{n+1} | x_{n+1}) = (2\pi)^{-q/2} |Q_{\lambda, n+1}|^{-1/2} \times \\ \times \exp \left\{ -\frac{1}{2} (y_{n+1} - C_{n+1} x_{n+1})^T Q_{\lambda, n+1}^{-1} (y_{n+1} - C_{n+1} x_{n+1}) \right\}, \quad (4.82)$$

где $Q_{\lambda,n+1} = \mathcal{N}_{\lambda,n+1}/\Delta t$ - положительная диагональная матрица дисперсий вектора дискретных белых шумов λ_{n+1} .

АПВ и ЭПВ также описывается гауссовскими распределениями (4.59) и (4.60). Воспользуемся полученным ранее выражениями (4.61), (4.62) и (4.58) для определения векторно-матричных уравнений оценки, дисперсий ошибки оценки, а также уравнения отношения правдоподобия.

Вектор $\Pi_{n+1}^{(1)}(\tilde{x}_{n+1})$ и матрица $\Pi_{n+1}^{(2)}(\tilde{x}_{n+1})$ определяются на основании ФП (4.82) и принимают форму

$$\Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) = C_{n+1}^T Q_{\lambda,n+1}^{-1} (y_{n+1} - C_{n+1} \tilde{x}_{n+1}),$$

$$\Pi_{n+1}^{(2)}(\tilde{x}_{n+1}) = -C_{n+1}^T Q_{\lambda,n+1}^{-1} C_{n+1}.$$

В результате рекуррентные уравнения оценок и дисперсией ошибок оценок с учетом $f_n(x_n) = F_n x_n$ принимают вид

$$\hat{x}_{n+1} = \tilde{x}_{n+1} + D_{n+1} C_{n+1}^T Q_{\lambda,n+1}^{-1} (y_{n+1} - C_{n+1} \tilde{x}_{n+1}), \quad x_0 = x(0), \quad (4.83)$$

$$D_{n+1} = [\tilde{D}_{n+1}^{-1} + C_{n+1}^T Q_{\lambda,n+1} C_{n+1}]^{-1}, \quad D_0 = D(0), \quad (4.84)$$

где $\tilde{x}_{n+1} = \Phi_n \hat{x}_n$, $\tilde{D}_{n+1} = D_n + \Delta t F_n D_n + \Delta t D_n F_n^T + \Gamma_n Q_{\eta n} \Gamma_n^T$ или $\tilde{D}_{n+1} = \Phi_n D_n \Phi_n^T + \Gamma_n Q_{\eta n} \Gamma_n^T$, $\Phi_n = I + \Delta t F_n$.

Решение уравнения (4.84) численными методами при высокой размерности вызывает большие трудности из-за необходимости обращения матрицы в квадратных скобках. Если число наблюдаемых сигналов q меньше размера вектора марковского процесса r то решение задачи можно упростить с помощью матричного тождества

$$\begin{aligned} & [\tilde{D}_{n+1}^{-1} + C_{n+1}^T Q_{\lambda,n+1} C_{n+1}]^{-1} = \\ & = \tilde{D}_{n+1} - \tilde{D}_{n+1} C_{n+1}^T (C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + Q_{\lambda,n+1})^{-1} C_{n+1} \tilde{D}_{n+1}. \end{aligned} \quad (4.85)$$

Указанная формула справедлива, если матрицы \tilde{D}_{n+1} , $Q_{\lambda,n+1}$ и C_{n+1} имеют размерность соответственно $(r \times r)$, $(q \times q)$ и $(q \times r)$. Отметим, что при скалярном наблюдаемом сигнале (4.81) необходимость в обращении матрицы вообще отпадает, так как выражение в круглых скобках (4.85) оказывается скалярной величиной.

После подстановки выражения (4.85) в уравнении (4.84) получаем уравнение дисперсии ошибки оценки

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{D}_{n+1} C_{n+1}^T [C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + Q_{\lambda, n+1}]^{-1} C_{n+1} \tilde{D}_{n+1}. \quad (4.86)$$

В некоторых задачах для упрощения расчетов уравнение оценки можно представить в иной форме за счет преобразования одного из сомножителей (4.83) с помощью выражения (4.86)

$$\begin{aligned} D_{n+1} C_{n+1}^T Q_{\lambda, n+1}^{-1} &= \tilde{D}_{n+1} C_{n+1}^T Q_{\lambda, n+1}^{-1} - \\ &- \tilde{D}_{n+1} C_{n+1}^T [C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + Q_{\lambda, n+1}]^{-1} C_{n+1} \tilde{D}_{n+1} C_{n+1}^T Q_{\lambda, n+1}^{-1} = \\ &= \tilde{D}_{n+1} C_{n+1}^T (C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + Q_{\lambda, n+1})^{-1} [(C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + \\ &+ Q_{\lambda, n+1}) Q_{\lambda, n+1}^{-1} - C_{n+1} \tilde{D}_{n+1} C_{n+1}^T Q_{\lambda, n+1}^{-1}]. \end{aligned} \quad (4.87)$$

Принимая во внимание, что выражение в квадратных скобках ввиду $Q_{\lambda, n+1} Q_{\lambda, n+1}^{-1} = I$ равно единичной матрице, рекуррентное уравнение оценки при фильтрации многомерного марковского процесса принимает окончательный вид

$$\begin{aligned} \hat{x}_{n+1} &= \hat{x}_{n+1} + \tilde{D}_{n+1} C_{n+1}^T [C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + Q_{\lambda, n+1}]^{-1} (y_{n+1} - C_{n+1} \tilde{x}_{n+1}), \\ \hat{x}_0 &= x(0). \end{aligned}$$

Как и для ранее рассмотренных одномерных процессов, можно сделать вывод, что для линейной задачи, гауссовских процессов и аддитивной независимой помехи уравнение оценки (4.88) линейно, а нелинейное уравнение дисперсии (4.86) не зависит от векторов оценки и наблюдаемых сигналов и может быть решено предварительно до решения уравнения оценки, т.е. независимо от процесса обработки наблюдаемого сигнала.

Приведем также уравнение фильтрации для распространенной в практических приложениях задачи оценивания одной из компонент многомерного марковского процесса и скалярном наблюдаемом сигнале, зависящим от этой компоненты. Обозначим указанную компоненту x_{1n} . В матрице C_{n+1} все компоненты за исключением $C_{11, n+1}$ равны нулю и наблюдаемый сигнал принимает вид $y_{n+1} = C_{11, n+1} x_{1, n+1} + \lambda_{n+1}$

$$\hat{x}_{n+1} = \tilde{x}_{n+1} + \tilde{R}_{n+1} C_{11,n+1} [\tilde{D}_{11,n+1} C_{11,n+1}^2 + \mathcal{N}_{\lambda,n+1} / \Delta t]^{-1} (y_{n+1} - C_{11,n+1} \tilde{x}_{1,n+1}), \quad \hat{x}_0 = x(0),$$

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{T}_{n+1} C_{11,n+1}^2 [\tilde{D}_{11,n+1} C_{11,n+1}^2 + \mathcal{N}_{\lambda,n+1} / \Delta t]^{-1}, \quad D_0 = D(0),$$

где $\tilde{R}_{n+1} = \| D_{11,n+1} \dots \tilde{D}_{r1,n+1} \|^T$, \tilde{T}_{n+1} - матрица размера $(r \times r)$, компонента которой $\tilde{D}_{i1,n+1} \times \tilde{D}_{j1,n+1}$, $i, j = \overline{1, r}$.

Рекуррентный алгоритм обнаружения для рассматриваемой линейной задачи может быть получен на основании выражения для многомерной ФП (4.82). Используя рассмотренный в разделе 4.7.1 подход для одномерных процессов, приходим к рекуррентному уравнению обнаружения

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n \sum_{i=1}^q \sum_{j=1}^r C_{ij,n+1} y_{i,n+1} \hat{x}_{j,n+1} Q_{\lambda i,n+1}, \quad \Lambda(t_0) = \Lambda(0).$$

В заключение рассмотрим линейные рекуррентные алгоритмы фильтрации и обнаружения рассмотренного в разделе 2.3.7 квазидетерминированного процесса вида $x(t) = \sum_{i=1}^k a_i t^{i-1}$, где подлежащие оценке параметры a_i - случайные величины, компоненты вектора $a = \|a_1 \dots a_k\|^T$. В дискретном времени квазидетерминированный процесс и скалярный наблюдаемый сигнал описываются соответственно соотношениями

$$\begin{aligned} x_{n+1} &= \sum_{i=1}^k a_i t_{n+1}^{i-1}, \\ y_{n+1} &= C_{n+1} T_{n+1} a + \lambda_{n+1}, \end{aligned} \quad (4.89)$$

где $T_{n+1} = \|1 t_{n+1} \dots t_{n+1}^{k-1}\|$, C_{n+1} и λ_{n+1} скалярные величины.

Функция правдоподобия имеет вид

$$\begin{aligned} p(y_{n+1} | a) &= \\ &= (2\pi Q_{\lambda,n+1})^{-1/2} \exp \left\{ - \frac{(y_{n+1} - C_{n+1} a^T T_{n+1}^T)(y_{n+1} - C_{n+1} T_{n+1} a)}{2Q_{\lambda,n+1}} \right\} \end{aligned} \quad (4.90)$$

где $Q_{\lambda,n+1}=N_{\lambda,n+1}/\Delta t$.

Переходная плотность вероятностей, как было показано ранее, принимает дельтаобразный вид и в параметрическом представлении оказывается равной $\pi(a_{n+1}|a_n)=\delta(a_{n+1}-a_n)$. Вследствие этого оператор ФПК приравнивается к нулю, экстраполированная плотность вероятностей заменяется согласно соотношению (4.35) апостериорной плотностью вероятностей

$$w_n(a_{n+1}) = \int_{-\infty}^{\infty} \delta(a_{n+1} - a_n) w_n(a_n) da_n.$$

С учетом высказанных замечаний рекуррентные уравнения АПВ и ОП принимают вид

$$w_{n+1}(a_{n+1}) = c_{n+1} p(y_{n+1} | a_{n+1}) w_n(a_{n+1}), \quad (4.91)$$

$$\Lambda_{n+1} = \frac{\Lambda_n}{p(y_{n+1})} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(y_{n+1} | a_{n+1}) w_n(a_{n+1}) da_{n+1}, \quad (4.92)$$

$$\text{где } c_{n+1} = \left[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(y_{n+1} | a_{n+1}) w_n(a_{n+1}) da_{n+1} \right]^{-1}.$$

Первая и вторая производные логарифмы ФП (4.90) по $a=\hat{a}_n$ принимает соответственно форму вектора и матрицы

$$\Pi_{n+1}^{(1)}(\hat{a}_n) = C_{n+1} T_{n+1}^T Q_{\lambda,n+1}^{-1} [y_{n+1} - C_{n+1} T_{n+1} \hat{a}_n], \quad (4.93)$$

$$\Pi_{n+1}^{(2)}(\hat{a}_n) = -C_{n+1}^2 T_{n+1}^T T_{n+1} Q_{\lambda,n+1}^{-1}. \quad (4.94)$$

Рекуррентные уравнения оценки и дисперсии ошибки оценки на основании формул (4.61) и (4.62)

$$\begin{aligned} \hat{a}_{n+1} &= \hat{a}_n + D_{n+1} \Pi_{n+1}^{(1)}(\hat{a}_n), \quad \hat{a}_0 = a(0), \\ D_{n+1} &= [D_n^{-1} + \Pi_{n+1}^{(2)}(\hat{a}_n)]^{-1}, \quad D_0 = D(0). \end{aligned}$$

или, учитывая соотношения (4.93) и (4.94), а также матричное тождество (4.85),

$$\begin{aligned} \hat{a}_{n+1} &= \hat{a}_n + C_{n+1} D_n T_{n+1}^T [C_{n+1}^2 T_{n+1} D_{n+1} T_{n+1}^T + Q_{\lambda,n+1}]^{-1} \times \\ &\times (y_{n+1} - C_{n+1} T_{n+1} \hat{a}_n), \quad \hat{a}_0 = a(0), \end{aligned} \quad (4.95)$$

$$D_{n+1} = D_n - C_{n+1}^2 D_n T_{n+1}^T [C_{n+1}^2 T_{n+1} D_{n+1} T_{n+1}^T + Q_{\lambda, n+1}]^{-1} T_{n+1} D_n, \quad (4.96)$$

$$D_0 = D(0).$$

Рекуррентный алгоритм отношения правдоподобия квазидетерминированного процесса может быть получен из выражения (4.79) путем замены \tilde{x}_{n+1} на $T_{n+1} \hat{a}_n$

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n C_{n+1} y_{n+1} T_{n+1} \hat{a}_n Q_{\lambda, n+1}^{-1}, \quad \Lambda_0 = \Lambda(0).$$

В скалярно-координатном виде уравнения (4.95) и (4.96) принимают форму

$$\hat{a}_{i, n+1} = \hat{a}_{in} + \frac{C_{n+1} \sum_{j=1}^k D_{ij, n} t_{n+1}^{j-1} \left(y_{n+1} - C_{n+1} \sum_{j=1}^k \hat{a}_{jn} t_{n+1}^{j-1} \right)}{C_{n+1}^2 \sum_{l, j=1}^k D_{lj, n} t_{n+1}^{j+l-2} + Q_{\lambda, n+1}}, \quad a_{i0} = a_i(0),$$

$$D_{iq, n+1} = D_{iq, n} - \frac{C_{n+1}^2 \sum_{l, j=1}^k D_{ij, n} D_{lq, n} t_{n+1}^{j+l-2}}{C_{n+1}^2 \sum_{l, j=1}^k D_{lj, n} t_{n+1}^{j+l-2} + Q_{\lambda, n+1}}, \quad D_{iq, 0} = D_{iq}(0), \quad i, q = \overline{1, k}.$$

В случае фильтрации не зависящего от времени одного параметра a_1 соответствующие уравнения

$$\hat{a}_{1, n+1} = \hat{a}_{1n} + C_{n+1} D_n [C_{n+1}^2 D_n + Q_{\lambda, n+1}]^{-1} (y_n - C_{n+1} \hat{a}_{1n}), \quad a_{10} = a_1(0),$$

$$D_{n+1} = D_n - C_{n+1}^2 D_n^2 [C_{n+1}^2 D_n + Q_{\lambda, n+1}]^{-1}, \quad D_0 = D(0),$$

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n C_{n+1} y_{n+1} \hat{a}_{1n} Q_{\lambda, n+1}^{-1}, \quad \Lambda_0 = \Lambda(0).$$

4.7.3. АЛГОРИТМЫ СИНТЕЗА ДЛЯ НЕПРЕРЫВНОГО ВРЕМЕНИ

Перейдем к рассмотрению уравнений линейной фильтрации для непрерывного времени, совершив предельный переход $\Delta t \rightarrow 0$ в соответствующих уравнениях с дискретным временем. Рассмотрим начнем с одномерного марковского процесса. Уравнение оценки и дисперсии ошибки оценки при наблюдаемом сигнале (4.68)

соответственно имеют вид (4.75) и (4.76). Подставив в эти уравнения функции $\tilde{x}_{n+1} = \Phi_n \hat{x}_n$, $\tilde{D}_{n+1} = D_n + 2\Delta t f_n D_n + \gamma_n^2 Q_{\eta n}$, где $\Phi_n = 1 + \Delta t f_n$, $\gamma_n = \gamma(t_n)\Delta t$, $Q_{\eta n} = \mathcal{N}_{\eta n}/\Delta t$, $Q_{\lambda, n+1} = \mathcal{N}_{\lambda, n+1}/\Delta t$, после предельного перехода получим дифференциальные уравнения оценки и дисперсии ошибки оценки

$$\frac{d\hat{x}(t)}{dt} = f_t \hat{x}(t) + D_t C_t N_{\lambda t}^{-1} [y(t) - C_t \hat{x}(t)], \quad x(t_0) = x(0), \quad (4.97)$$

$$\frac{dD_t}{dt} = 2f_t D_t + \gamma_t^2 \mathcal{N}_{\eta t} - C_t^2 D_t^2 \mathcal{N}_{\lambda t}^{-1}, \quad D_{t_0} = D(0). \quad (4.98)$$

Нелинейное уравнение дисперсии (4.98) является частным случаем уравнения Риккати $dD_t/dt = a(t) + b(t)D_t + c(t)D_t^2$. Решение этого уравнения получают численными методами с помощью ЭВМ. Уравнение Риккати можно свести к двум линейным нестационарным уравнениям путем подстановки $D_t = R(t)/P(t)$

$$\frac{dR(t)}{dt} = \frac{1}{2}b(t)R(t) + a(t)P(t), \quad P(t_0) = P(0), \quad (4.99)$$

$$\frac{dP(t)}{dt} = -\frac{1}{2}b(t)P(t) + c(t)R(t), \quad R(t_0) = R(0). \quad (4.100)$$

Если коэффициенты уравнений (4.99) и (4.100) имеют постоянные значения, то эти уравнения позволяют получить аналитическое решение задачи [45].

В случае фильтрации многомерного марковского процесса (4.80) и уравнения наблюдения (4.81) соответствующие уравнения оценки и дисперсии ошибки оценки в дискретном времени имеют вид (4.87) и (4.86). Учитывая что $\tilde{x}_{n+1} = \Phi_n \hat{x}_n$, $\tilde{D}_{n+1} = \Phi_n D_n \Phi_n^T + \Gamma_n Q_{\eta n} \Gamma_n^T$, где $\Phi_n = I + \Delta t F_n$, $\Gamma_n = \Delta t \Gamma(t_n)$, $Q_{\lambda, n+1} = \mathcal{N}_{\lambda, n+1}/\Delta t$, $Q_{\eta n} = \mathcal{N}_{\eta n}/\Delta t$, F_n - матрица коэффициентов уравнения состояния, $\Gamma(t_n)$ - диагональная матрица, после предельного перехода $\Delta t \rightarrow 0$ получаем уравнения линейной многомерной фильтрации

$$\frac{d\hat{x}(t)}{dt} = F_t \hat{x}(t) + D_t C_t^T \mathcal{N}_{\lambda t}^{-1} [y(t) - C_t \hat{x}(t)], \quad \hat{x}(t_0) = x(0), \quad (4.101)$$

$$\frac{dD_t}{dt} = F_t D_t + D_t F_t^T + \Gamma_t \mathcal{N}_{\eta t} \Gamma_t^T - D_t C_t^T \mathcal{N}_{\lambda t}^{-1} C_t D_t, \quad D_{t_0} = D(0). \quad (4.102)$$

Нелинейное уравнение дисперсии (4.103) является матричным уравнением Риккати. Как и для одномерного случая, это уравнение решается численными методами с помощью ЭВМ. Это уравнение путем замены $D_t = R(t)/P(t)$ может быть сведено к решению системы $2q$ линейных дифференциальных уравнений

$$\begin{aligned} \left\| \frac{dP(t)}{dt} \right\| &= \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} P(t) \\ R(t) \end{bmatrix}, \quad P(t_0) = P(0), \\ \left\| \frac{dR(t)}{dt} \right\| &= \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} P(t) \\ R(t) \end{bmatrix}, \quad R(t_0) = R(0), \end{aligned}$$

где функции F_{ij} удовлетворяют уравнению (4.102) в форме

$$\frac{dD_t}{dt} = F_{22} D_t - D_t F_{11} + F_{21} - D_t F_{12} D_t.$$

Обратим внимание на то, что, как и для дискретного времени, уравнение оценки линейно относительно самой оценки, а уравнение дисперсии не зависит от наблюдаемого сигнала и оценки, решается независимо от уравнения оценки. Кроме того, в правых частях уравнения оценки первое слагаемое, а в уравнении дисперсии первых три слагаемых определяют эволюцию априорных составляющих соответствующих статистик согласно уравнению ФПК. В этом нетрудно убедиться, сравнив дифференциальные уравнения (3.42) и (3.50) с уравнениями (4.101) и (4.102)

При фильтрации квазидетерминированных процессов на основании уравнений (4.95) и (4.96) после предельного перехода можно получить

$$\frac{d\hat{a}(t)}{dt} = C_t \mathcal{N}_{\lambda t}^{-1} D_t T^T [y(t) - C_t T \hat{a}(t)], \quad \hat{a}(t_0) = a_0,$$

$$\frac{dD_t}{dt} = -C_t^2 \mathcal{N}_{\lambda t}^{-1} D_t T^T T D_t, \quad D(t_0) = D_0, \quad T = \begin{bmatrix} 1 & t & \dots & t^{k-1} \end{bmatrix}.$$

4.8. КВАЗИОПТИМАЛЬНЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ И ФИЛЬТРАЦИИ

4.8.1. ПОСТАНОВКА ЗАДАЧИ

Общие методы статистического синтеза для марковских процессов были сформулированы в разд. 4.6.2 в виде алгоритмов фильтрации (4.38) и обнаружения (4.36). Было отмечено, что решения в общем виде нелинейных с негауссовскими процессами задач синтеза представляют собой, как правило, трудноразрешимую проблему. При конкретизации методов синтеза были заданы семейства плотностей вероятностей АПВ и ЭПВ. И уже в рамках применения этих плотностей получены выражения апостериорных параметров (условных центральных моментов), число которых определяется порядком аппроксимации указанных плотностей вероятностей. При практическом моделировании прибегают к приближенным методам и ограничиваются усеченной определяющей эти параметры системой уравнений, которая, вообще говоря, уже не дает оптимальное решение. Система уравнений при этом оказывается конечной и замкнутой относительно ограниченного числа апостериорных параметров. Во многих задачах прикладного характера это оправдано, так как апостериорные параметры высших порядков оказываются малосущественными и усеченная система дает решения близкие к оптимальным.

К числу приближенных, получивших название квазиоптимальных методов синтеза, относятся методы, позволяющие решить задачу синтеза в рамках корреляционной теории, т.е. на основе использования апостериорных параметров не выше второго порядка ($K=2$). Достаточным условием рассматриваемого подхода является гауссовская аппроксимация АПВ. О степени близости АПВ к гауссовскому распределению зависит эффективность квазиоптимальных методов обнаружения и фильтрации. Очевидно, требования к указанной аппроксимации можно сформулировать с учетом ограничений на величины апостериорных параметров. Степень близости к гауссовскому распределению можно охарактеризовать коэффициентами [46]

$$\gamma_{in} = \frac{h_{in}}{h_n^{i/2}}, \quad h_{2n} = h_n, \quad i = \overline{3, K}. \quad (4.103)$$

где апостериорные параметры h_{in} получены из соотношений (4.46) и (4.47). Если выполняется условие $|\gamma_{in}| \ll 1$, то аппроксимация АПВ гауссовским распределением оказывается справедливой и решение задачи синтеза базируется на квазиоптимальных алгоритмах в гауссовском приближении. Сформулированные условия можно раскрыть на основании выражений для h_{in} (4.103) с помощью соотношений

$$[g_n - \Pi_n^{(2)}(\tilde{x}_n)]^{i/2} \gg |g_{in} - \Pi_n^{(i)}(\tilde{x}_n)|.$$

Ввиду того, что $g_n - \Pi_n^{(2)}(\tilde{x}_n) = h_n$ эти неравенства принимают следующий вид

$$[g_n - \Pi_n^{(2)}(\tilde{x}_n)]^{i/2} \gg |g_n - \Pi_n^{(2)}(\tilde{x}_n)| / L_n^{i-2} \quad (4.104)$$

или $h_n^{(i-2)/2} \gg L_n^{-(i-2)}$, где L_n характеризует относительный порядок параметров

$$L_n \approx \min(l_n, \bar{l}_n), \quad l_n^{i-2} = |\Pi_n^{(2)}(\tilde{x}_n)| / |\Pi_n^{(i)}(\tilde{x}_n)|, \quad \bar{l}_n^{i-2} = |g_n| / |g_{in}|.$$

Анализ неравенств (4.104) показывает, что их выполнение обеспечивается уже при $i=3$, т.е. $h_n^{1/2} \gg L_n^{-1}$ или, принимая во внимание, что апостериорная дисперсия $D_n = h_n^{-1}$, получаем

$$\sqrt{D_n} \ll L_n. \quad (4.105)$$

Условие (4.105) означает не только близость АПВ к гауссовскому распределению, но и высокую апостериорную точность (малую апостериорную дисперсию).

В рассматриваемых ниже приближенных подходах уделяется внимание наиболее простым с точки зрения технической реализации квазиоптимальным методам: локальной гауссовской аппроксимации и статистической линеаризации.

4.8.2. МЕТОД ЛОКАЛЬНОЙ ГАУССОВСКОЙ АППРОКСИМАЦИИ

4.8.2.1. АЛГОРИТМЫ СИНТЕЗА ДЛЯ ОДНОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

При локальной гауссовской аппроксимации текущая оценка x^* фильтруемого процесса x определяется в области малых отклонений или ошибок $|x - x^*|$, что объясняет использование термина «локальный». По-существу в основе этого метода лежит представление нелинейных дифференцируемых функций первыми членами разложения этих функций в ряд Тейлора. При квадратичной функции потерь это означает малое значение апостериорной дисперсии ошибки $M\{(x - x^*)^2\}$ (большую апостериорную точность).

При выводе уравнений фильтрации и обнаружения исходными являются уравнения одномерного марковского процесса (2.53)

$$x_{n+1} = x_n + \Delta t f_n(x_n) + \gamma_n \eta_n, \quad x_0 = x(0) \quad (4.106)$$

и скалярного наблюдаемого сигнала

$$y_{n+1} = S_{n+1}(x_{n+1}) + \lambda_{n+1} \quad (4.107)$$

в дискретном времени.

Функции правдоподобия и плотность вероятностей в отсутствии полезного сигнала на основании (4.107) соответственно равны

$$p(y_{n+1} | x_{n+1}) = (2\pi Q_{\lambda, n+1})^{-1/2} \exp \left\{ -\frac{[y_{n+1} - S_{n+1}(x_{n+1})]^2}{2Q_{\lambda, n+1}} \right\}, \quad (4.108)$$

$$p(y_{n+1}) = (2\pi Q_{\lambda, n+1})^{-1/2} \exp \left\{ -\frac{y_{n+1}^2}{2Q_{\lambda, n+1}} \right\}, \quad (4.109)$$

где $Q_{\lambda, n+1}$ - дисперсия дискретного белого шума λ_{n+1} .

Первая и вторая производные по $x_{n+1} = \tilde{x}_{n+1}$ логарифма ФП (4.108) соответственно принимают форму

$$\Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) = \frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} Q_{\lambda, n+1}^{-1} [y_{n+1} - S_{n+1}(\tilde{x}_{n+1})], \quad (4.110)$$

$$\begin{aligned} \Pi_{n+1}^{(2)}(\tilde{x}_{n+1}) = \\ = Q_{\lambda, n+1}^{-1} \left\{ [y_{n+1} - S_{n+1}(\tilde{x}_{n+1})] \frac{\partial^2 S_{n+1}(\tilde{x}_{n+1})}{\partial x_{n+1}^2} - \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right]^2 \right\}, \end{aligned} \quad (4.111)$$

где экстраполированные параметры после подстановки первых двух членов разложения в ряд нелинейной функции

$$f_n(x_n) \approx f_n(x_n^*) + (x_n - x_n^*) \frac{\partial f_n(x_n^*)}{\partial x_n^*}$$

в соотношения (4.53) и (4.54) и последующих упрощений принимают вид

$$\tilde{x}_{n+1} = x^* + \Delta t f_n(x_n^*), \quad (4.112)$$

$$\tilde{D}_{n+1} = D_n + 2\Delta t \frac{\partial f_n(x_n^*)}{\partial x_n^*} D_n + \gamma_n^2 Q_{\eta n}. \quad (4.113)$$

Уравнение (4.113) можно переписать в иной форме

$$\tilde{D}_{n+1} = \Phi_n^2 D_n + \gamma_n^2 Q_{\eta n}, \quad (4.114)$$

где $\Phi_n = 1 + \Delta t \frac{\partial f_n(x^*)}{\partial x^*}$, $Q_{\eta n} = \mathcal{N}_{\eta n} / \Delta t$, $\gamma_n = \chi(t_n) \Delta t$.

Если подставить выражения (4.110) и (4.111) соответственно в уравнения (4.45) и (4.47) при $K=2$, то приходим к уравнениям оценки и дисперсии ошибки оценки в дискретном времени при локальной гауссовской аппроксимации. Обратим внимание, что в отличие от алгоритмов линейной фильтрации, описываемых уравнениями (4.75) и (4.76), в рассматриваемом случае уравнение оценки нелинейно относительно самой оценки, а в уравнении дисперсии присутствует наблюдаемый сигнал и оценка. Это уравнение необходимо решать совместно с уравнением оценки.

Можно предложить другие алгоритмы гауссовского приближения, отличающиеся дополнительными членами вследствие расширения набора нелинейных функций, изменение формы их представления при разложении в ряд и т.д. Целесообразность учета

этих членов определяется при конкретизации и анализе исходных данных задач. В связи с этим необходимо заметить, что учет дополнительных членов (в нашем случае слагаемое в фигурных скобках (4.111), зависящее от наблюдаемого сигнала y_{n+1}) при моделировании не только усложняет задачу синтеза, но и ухудшает точностные характеристики алгоритма.

Хотя нет основания придавать этим результатам общий характер, тем не менее они ставят под сомнение необходимость учитывать дополнительные члены. Рассматриваемый ниже алгоритм фильтрации с уравнением дисперсии, не зависящим от наблюдаемого сигнала, получил название алгоритма расширенного фильтра Калмана.

Усредним вторую производную логарифма ФП относительно плотности вероятностей $p(y_{n+1}|x_{n+1})$, сделав ее тем самым инвариантной к изменениям наблюдаемого сигнала y_{n+1} . В результате получим

$$\begin{aligned}\bar{\Pi}_{n+1}^{(2)}(x_{n+1}) &= M\{\Pi_{n+1}^{(2)}(x_{n+1}) | x_{n+1}\} = \\ &= \int_{-\infty}^{\infty} \Pi_{n+1}^{(2)}(x_{n+1}) p(y_{n+1} | x_{n+1}) dy_{n+1} = \\ &= -Q_{\lambda, n+1}^{-1} \left[\frac{\partial S_{n+1}(x_{n+1})}{\partial x_{n+1}} \right]^2,\end{aligned}\tag{4.115}$$

так как из-за равенства нулю математического ожидания дискретного белого шума имеет место $M\{(y_{n+1} - S_{n+1}(x_{n+1})) | x_{n+1}\} = 0$.

После подстановки $x_{n+1} = \tilde{x}_{n+1}$, а также соотношений (4.110) и (4.115) в уравнения (4.45) и (4.47) получаем рекуррентный алгоритм фильтрации одномерного марковского процесса в дискретном времени

$$\begin{aligned}x_{n+1}^* &= \tilde{x}_{n+1} + D_{n+1} \frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} Q_{\lambda, n+1}^{-1} [y_{n+1} - S_{n+1}(\tilde{x}_{n+1})], \\ D_{n+1} &= \left[\tilde{D}_{n+1}^{-1} + Q_{\lambda, n+1}^{-1} \left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right)^2 \right]^{-1}\end{aligned}$$

или с учетом соотношений (4.85) и (4.87)

$$x_{n+1}^* = \tilde{x}_{n+1} + \tilde{D}_{n+1} \frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \times$$

$$\times \left[\tilde{D}_{n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right)^2 + Q_{\lambda, n+1} \right]^{-1} [y_{n+1} - S_{n+1}(\tilde{x}_{n+1})], \quad x_0^* = x(0), \quad (4.116)$$

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{D}_{n+1}^2 \left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right)^2 \times$$

$$\times \left[\tilde{D}_{n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right)^2 + Q_{\lambda, n+1} \right]^{-1}, \quad D_0 = D(0). \quad (4.117)$$

При определении рекуррентного алгоритма обнаружения воспользуемся выражениями (4.36) и (4.108). После вычислений, аналогичных при выводе формулы (4.79), приходим к рекуррентному соотношению

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n y_{n+1} \tilde{S}_{n+1}(x_{n+1}) Q_{\lambda, n+1}^{-1}, \quad (4.118)$$

где

$$\tilde{S}_{n+1}(x_{n+1}) = \int_{-\infty}^{\infty} S_{n+1}(x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1}. \quad (4.119)$$

Определение экстраполированной функции $\tilde{S}_{n+1}(x_{n+1})$ является достаточно сложной задачей. Упрощение достигается при использовании первых трех членов разложения функции $\tilde{S}_{n+1}(x_{n+1})$ в ряд в окрестности точки \tilde{x}_{n+1} . После вычислений получаем алгоритм отношения правдоподобия

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n y_{n+1} Q_{\lambda, n+1}^{-1} \left[S_{n+1}(\tilde{x}_{n+1}) + \frac{1}{2} \tilde{D}_{n+1} \frac{\partial^2 S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}^2} \right], \quad \Lambda_0 = \Lambda(0).$$

4.8.2.2. АЛГОРИТМЫ СИНТЕЗА ДЛЯ МНОГОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

При определении рекуррентных уравнений фильтрации и обнаружения для многомерного марковского процесса используются соотношения (4.61), (4.62), (4.58). Исходными являются в дискретном времени r -мерное уравнение марковских процессов

$$x_{n+1} = x_n + \Delta t f_n(x_n) + \Gamma_n \eta_n, \quad x_0 = x(0) \quad (4.120)$$

и q -мерное уравнение наблюдаемого сигнала

$$y_{n+1} = S_{n+1}(x_{n+1}) + \lambda_{n+1} \quad (4.121)$$

в векторной форме, $q \leq r$.

Многомерная функция правдоподобия имеет вид

$$p(y_{n+1} | x_{n+1}) = (2\pi)^{-q/2} |Q_{\lambda, n+1}|^{-1/2} \times \\ \times \exp \left\{ -\frac{1}{2} [y_{n+1} - S_{n+1}(x_{n+1})]^T Q_{\lambda, n+1}^{-1} [y_{n+1} - S_{n+1}(x_{n+1})] \right\}, \quad (4.122)$$

Первая и вторая производные логарифма ФП для расширенного фильтра Калмана принимает соответственно форму вектора и матрицы

$$\Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) = \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right]^T Q_{\lambda, n+1}^{-1} [y_{n+1} - S_{n+1}(\tilde{x}_{n+1})], \quad (4.123)$$

$$\bar{\Pi}_{n+1}^{(2)}(\tilde{x}_{n+1}) = - \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right]^T Q_{\lambda, n+1}^{-1} \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right], \quad (4.124)$$

где $\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}}$ - матрица Якоби размера $(q \times r)$, (i, j) компонента которой $\partial S_{i, n+1}(\tilde{x}_{n+1}) / \partial x_{j, n+1}$ (i - номер строки, j - номер столбца).

Разложим нелинейную вектор-функцию $f_n(x_n)$ в многомерный ряд Тейлора в окрестности вектора x_n^* . После подстановки первых двух членов разложения в формулы (4.65) и (4.66) получаем выражения для экстраполированных вектора оценок и матрицы. С

учетом сказанного после подстановки выражений (4.123) и (4.124) в уравнения (4.61) и (4.62), использования тождества (4.85) и соотношения (4.87), получим рекуррентные уравнения оценки и дисперсии ошибки оценки для локальной гауссовской аппроксимации

$$x_{n+1}^* = \tilde{x}_{n+1} + \tilde{D}_{n+1} \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right]^T \times \\ \times \left[\left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda, n+1} \right]^{-1} \times \quad (4.125) \\ \times [y_{n+1} - S_{n+1}(\tilde{x}_{n+1})], \quad x_0^* = x(0),$$

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{D}_{n+1} \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right]^T \times \\ \times \left[\left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda, n+1} \right]^{-1} \times \quad (4.126) \\ \times \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1})}{\partial \tilde{x}_{n+1}} \right] \tilde{D}_{n+1}, \quad D_0 = D(0),$$

где

$$\tilde{x}_{n+1} = x_n^* + \Delta t f_n(x_n^*), \quad (4.127)$$

$$\tilde{D}_{n+1} = D_n + \Delta t \left(\frac{\partial f_n(x_n^*)}{\partial x_n^*} \right) D_n + \Delta t D_n \left(\frac{\partial f_n(x_n^*)}{\partial x_n^*} \right)^T + \Gamma_n Q_{\eta n} \Gamma_n^T, \quad (4.128)$$

$\partial f_n(x_n^*) / \partial x_n^*$ - матрица Якоби размера $(r \times r)$, (i, j) компонента которой $\partial f_{in}(x_n^*) / \partial x_{jn}^*$ (i - номер строки, j - номер столбца).

На основании полученных результатов можно прийти к заключению, что как и для скалярных процессов, в рассматриваемой задаче уравнение оценки нелинейно относительно самого вектора оценки, а уравнение дисперсии зависит от вектора оценки и его необходимо решать совместно с уравнением оценки.

Рекуррентный алгоритм отношения правдоподобия определяется из выражения (4.58) на основании многомерной ФП (4.122). Используя подход для скалярных процессов, после вычислений приходим к рекуррентному уравнению отношения правдоподобия

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n \sum_{i=1}^q y_{i,n+1} \tilde{S}_{i,n+1}(x) Q_{\lambda i, n+1}^{-1}.$$

Перейдем к формированию рекуррентных уравнений фильтрации и обнаружения квазидетерминированного процесса вида (4.89) Скалярный наблюдаемый сигнал описывается соотношением

$$y_{n+1} = S_{n+1}(a) + \lambda_{n+1}, \quad (4.129)$$

где вектор подлежащих оценке параметров $a = \|a_1 \dots a_k\|^T$.
Функция правдоподобия имеет вид

$$p(y_{n+1} | a) = (2\pi Q_{\lambda, n+1})^{-1/2} \exp \left\{ -\frac{[y_{n+1} - S_{n+1}(a)]^2}{2Q_{\lambda, n+1}} \right\}, \quad (4.130)$$

где $Q_{\lambda, n+1} = \mathcal{N}_{\lambda, n+1} / \Delta t$.

Как и для линейной задачи, воспользуемся тем обстоятельством, что оператор ФПК равен нулю и уравнения АПВ и ОП соответственно описываются соотношениями (4.91) и (4.92). Первая и вторая производные логарифма ФП (4.130) по $a = a_n^*$ принимают соответственно форму вектора и матрицы

$$\Pi_{n+1}^{(1)}(a_n^*) = \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T Q_{\lambda, n+1}^{-1} [y_{n+1} - S_{n+1}(a_n^*)], \quad (4.131)$$

$$\bar{\Pi}^{(2)}(a_n^*) = -Q_{\lambda, n+1}^{-1} \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right), \quad (4.132)$$

где $\partial S_{n+1}(a_n^*) / \partial a_n^*$ - вектор-строка из компонент $\partial S_{n+1}(a_n^*) / \partial a_{jn}^*$, $j = \overline{1, k}$.

После подстановки выражений (4.131) и (4.132) в уравнения (4.61) и (4.62), используя тождество (4.85), а также соотношение (4.87), приходим к рекуррентному алгоритму фильтрации параметров квазидетерминированного процесса расширенного фильтра Калмана

$$\begin{aligned}
 a_{n+1}^* &= a_n^* + D_n \left[\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right]^T \times \\
 &\times \left[\left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T + Q_{\lambda, n+1} \right]^{-1} \times \\
 &\times [y_{n+1} - S_{n+1}(a_n^*)], \quad a_0^* = a(0),
 \end{aligned} \tag{4.133}$$

$$\begin{aligned}
 D_{n+1} &= D_n - D_n \left[\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right]^T \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right) D_n \times \\
 &\times \left[\left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T + Q_{\lambda, n+1} \right]^{-1},
 \end{aligned} \tag{4.134}$$

$$D_0 = D(0),$$

где выражения в квадратных скобках являются скалярными величинами.

Рекуррентный алгоритм отношения правдоподобия получают из соотношения (4.118) после замены $\tilde{S}_{n+1}(x_{n+1})$ на функцию $S_{n+1}(a_n^*)$. Указанную функцию под знаком интеграла представляют первыми тремя членами разложения в ряд Тейлора в окрестности вектора a_n^* . После подстановки результата интегрирования в формулу (4.118) имеем

$$\Lambda_{n+1} = \Lambda_n y_{n+1} Q_{\lambda, n+1}^{-1} \left[S_{n+1}(a_n^*) + \frac{1}{2} \sum_{i,j=1}^k D_{ij,n} \frac{\partial^2 S_{n+1}(a_n^*)}{\partial a_{in}^* \partial a_{jn}^*} \right].$$

В скалярно-координатной форме уравнения (4.133) и (4.134) принимают вид

$$a_{i,n+1}^* = a_{in}^* + \frac{\sum_{l=1}^k D_{il,n} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{ln}^*} [y_{n+1} - S_{n+1}(a_n^*)]}{\sum_{l,m=1}^k D_{lm,n} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{ln}^*} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{mn}^*} + Q_{\lambda,n+1}}, \quad a_{i0}^* = a_i(0),$$

$$D_{ij,n+1} = D_{ij,n} - \frac{\sum_{l,m=1}^k D_{il,n} D_{mj,n} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{ln}^*} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{mn}^*}}{\sum_{l,m=1}^k D_{lm,n} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{ln}^*} \frac{\partial S_{n+1}(a_n^*)}{\partial a_{mn}^*} + Q_{\lambda,n+1}},$$

$$D_{ij,0} = D_{ij}(0), \quad i, j = \overline{1, k}.$$

4.8.2.3. АЛГОРИТМЫ СИНТЕЗА ДЛЯ НЕПРЕРЫВНОГО ВРЕМЕНИ

Для многомерных процессов уравнения фильтрации в непрерывном времени при локальной гауссовской аппроксимации получают путем предельного перехода в уравнениях (4.125) и (4.126). Подставим в эти уравнения значения экстраполированных параметров (4.127) и (4.128), а также учтем $\Gamma_n = \Delta t \Gamma(t_n)$, $Q_{\eta n} = \mathcal{N}_{\eta n} / \Delta t$, $Q_{\lambda, n+1} = \mathcal{N}_{\lambda, n+1} / \Delta t$. После предельного перехода $\Delta t \rightarrow 0$ получим дифференциальные уравнения оценок и дисперсии ошибок оценок для расширенного фильтра Калмана

$$\frac{dx^*(t)}{dt} = f_t(x_t^*) + D(t) \left(\frac{\partial S_t(x_t^*)}{\partial x^*} \right)^T \mathcal{N}_{\lambda t}^{-1} [y(t) - S_t(x_t^*)], \quad x^*(t_0) = x(0),$$

$$\frac{dD(t)}{dt} = \frac{\partial f_t(x_t^*)}{\partial x^*} D(t) + D(t) \left(\frac{\partial f_t(x_t^*)}{\partial x^*} \right)^T + \Gamma_t \mathcal{N}_{\lambda t} \Gamma_t^T -$$

$$- D(t) \left(\frac{\partial S_t(x_t^*)}{\partial x^*} \right)^T \mathcal{N}_{\lambda t}^{-1} \left(\frac{\partial S_t(x_t^*)}{\partial x^*} \right) D(t), \quad D(t_0) = D(0).$$

Приведем также полученные путем предельного перехода дифференциальные уравнения фильтрации в непрерывном времени квазидетерминированных процессов

$$\frac{da^*(t)}{dt} = D(t) \left(\frac{\partial S_t(a_t^*)}{\partial a^*} \right)^T \mathcal{N}_{\lambda t}^{-1} [y(t) - S_t(a_t^*)], \quad a^*(t_0) = a(0),$$

$$\frac{dD(t)}{dt} = -\mathcal{N}_{\lambda t}^{-1} D(t) \left(\frac{\partial S_t(a_t^*)}{\partial a^*} \right)^T \left(\frac{\partial S_t(a_t^*)}{\partial a^*} \right) D(t), \quad D(t_0) = D(0).$$

4.8.3. МЕТОД СТАТИСТИЧЕСКОЙ ЛИНЕАРИЗАЦИИ

В отличие от метода локальной гауссовской аппроксимации, пригодного для малых ошибок фильтрации в окрестности оценки x^* , в методе статистической линеаризации оценка определяется в значительно более широкой области изменения случайных функций. Как было показано в разд. 3.4.2, для многих нелинейных функций, в том числе и для существенно нелинейных, можно построить в вероятностном смысле приближенную относительно простую зависимость между входным и выходным сигналами. Это достигается заменой безынерционной нелинейной функции статистически эквивалентным линеаризованным преобразованием, учитывающим наиболее существенные свойства исходной нелинейной функции. Метод статистической линеаризации, в основе которого также лежит гауссовское приближение апостериорной плотности вероятности, получил широкое распространение, так как позволяет построить сравнительно несложные алгоритмы синтеза для нелинейных систем с инерционными звеньями при гауссовском распределении ошибок измерения.

4.8.3.1. АЛГОРИТМЫ СИНТЕЗА ДЛЯ ОДНОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

Согласно общей идеи статистической линеаризации безынерционные нелинейные функции $f(x)$ и $S(x)$ описываются линеаризованными зависимостями (3.69)

$$f(x) \approx f_0 + K_f(x - x^*), \quad (4.135)$$

$$S(x) \approx S_0 + K_S(x - \tilde{x}), \quad (4.136)$$

где f_0 и S_0 - моментные функции, называемые статистическими характеристиками нелинейных функций (НФ), K_f и K_S -

статистические коэффициенты усиления, $(x-x^*)$ и $(x-\tilde{x})$ - центрированные составляющие случайных процессов.

Запишем линеаризованные выражения (4.135) и (4.136) с учетом зависимостей между статистическими характеристиками и коэффициентами усиления НФ (3.72) для дискретного времени

$$\begin{aligned} f_n(x_n) &\approx f_{0n}(x_n^*, D_n) + \frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*} (x_n - x_n^*) = \\ &= f_{0n} + \frac{\partial f_{0n}}{\partial x^*} (x_n - x^*), \end{aligned} \quad (4.137)$$

$$\begin{aligned} S_{n+1}(x_{n+1}) &\approx S_{0,n+1}(\tilde{x}_{n+1}, \tilde{D}_{n+1}) + \\ &+ \frac{\partial S_{0,n+1}(\tilde{x}_{n+1}, \tilde{D}_{n+1})}{\partial \tilde{x}_{n+1}} (x_{n+1} - \tilde{x}_{n+1}) = \\ &= S_{0,n+1} + \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} (x_{n+1} - \tilde{x}_{n+1}). \end{aligned} \quad (4.138)$$

Уравнения одномерного марковского процесса и скалярного наблюдаемого сигнала соответственно имеют вид (4.106) и (4.107). Функция правдоподобия и плотность вероятностей в отсутствии полезного сигнала равны (4.108) и (4.109). После подстановки выражения (4.138) в ФП первая и вторая производные по $x_{n+1}=\tilde{x}_{n+1}$ логарифма ФП соответственно равны

$$\Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) = Q_{\lambda,n+1}^{-1} \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} [y_{n+1} - S_{0,n+1}], \quad (4.139)$$

$$\Pi_{n+1}^{(2)}(\tilde{x}_{n+1}) = -Q_{\lambda,n+1}^{-1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^2. \quad (4.140)$$

Подставляя выражения (4.139) и (4.140) соответственно в уравнения (4.45) и (4.47) при $K=2$, получим с учетом соотношений (4.85) и (4.87) рекуррентный алгоритм фильтрации одномерного марковского процесса в дискретном времени

$$x_{n+1}^* = \tilde{x}_{n+1} + \tilde{D}_{n+1} \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \left[\tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^2 + Q_{\lambda,n+1} \right]^{-1} \times \quad (4.141)$$

$$\times [y_{n+1} - S_{0,n+1}], \quad x_0^* = x(0),$$

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{D}_{n+1}^2 \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \left[\tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^2 + Q_{\lambda,n+1} \right]^{-1}, \quad (4.142)$$

$$D_0 = D(0).$$

Используя в выражениях (4.53) и (4.54) представления нелинейной функции (4.135), после вычислений получим экстраполированные оценки и дисперсии ошибок оценок

$$\tilde{x}_{n+1} = x_n^* + \Delta t f_{0n}, \quad (4.143)$$

$$\tilde{D}_{n+1} = D_n + 2\Delta t \frac{\partial f_{0n}}{\partial x_n^*} D_n + \gamma_n^2 Q_{\eta n}. \quad (4.144)$$

Как следует из полученных выражений, в виду зависимости $S_{0,n+1}$ и $\partial S_{0,n+1}/\partial \tilde{x}_{n+1}$ от экстраполированных параметров \tilde{x}_{n+1} и \tilde{D}_{n+1} линеаризированные уравнения (4.141) и (4.142) следует решать совместно.

В связи с появлением дополнительной операции определения статистических характеристик НФ, рассмотрим порядок решений этих уравнений. С помощью вычисленных на шаге n оценки x_n^* и дисперсии D_n определяют по приведенной ниже методике статистические характеристику $f_{n0}(x_n^*, D_n)$ и коэффициент усиления $\partial f_{0n}(x_n^*, D_n)/\partial x_n^*$ линеаризируемой НФ $f_n(x_n)$, которые далее используются в выражениях (4.143) и (4.144) при вычислениях экстраполированных параметров \tilde{x}_{n+1} и \tilde{D}_{n+1} .

Применяя уже упомянутую методику на шаге $n+1$ определяют статистические характеристику $S_{0,n+1}(\tilde{x}_{n+1}, \tilde{D}_{n+1})$ и коэффициент усиления $\partial S_{0,n+1}(\tilde{x}_{n+1}, \tilde{D}_{n+1})/\partial \tilde{x}_{n+1}$ линеаризируемой НФ $S_{n+1}(x_{n+1})$. Подставляя полученные значения \tilde{x}_{n+1} , \tilde{D}_{n+1} , $S_{0,n+1}$ и $\partial S_{0,n+1}/\partial \tilde{x}_{n+1}$, а

также результаты измерения y_{n+1} в уравнения (4.141) и (4.142), рассчитываем новые значения оценки x_{n+1}^* и D_{n+1} . Затем цикл вычислений повторяется с определения на шаге $n+2$ значений $f_{0,n+1}$ и $\partial f_{0,n+1}/\partial x_{n+1}^*$, и соответственно, экстраполированных параметров \tilde{x}_{n+2} и \tilde{D}_{n+2} .

Рассмотрим методы получения статистических характеристик НФ экспоненциального типа $S_{n+1}(x_{n+1})=A_{n+1}\exp\{-\alpha x_{n+1}\}$, а также НФ, аппроксимируемых кусочно-линейной функцией.

Статистическую характеристику НФ экспоненциального типа можно определить в результате усреднения НФ относительно гауссовской экстраполированной плотности вероятностей

$$\begin{aligned} S_{0,n+1} &= M\{S_{n+1}(x_{n+1}) | y_1^n\} = \\ &= A_{n+1}(2\pi\tilde{D}_{n+1})^{-1/2} \int_{-\infty}^{\infty} \exp\left\{-\alpha x_{n+1} - \frac{(x_{n+1} - \tilde{x}_{n+1})^2}{2\tilde{D}_{n+1}}\right\} dx_{n+1}. \end{aligned} \quad (4.145)$$

Введем замену $z = (x_{n+1} - \tilde{x}_{n+1})\tilde{D}_{n+1}^{-1/2} + \alpha\tilde{D}_{n+1}^{1/2}$. После подстановки z^2 в выражение (4.145) и интегрирования получим

$$S_{0,n+1} = A_{n+1} \exp\left\{-\alpha\tilde{x}_{n+1} + \frac{\alpha^2}{2}\tilde{D}_{n+1}\right\}.$$

Статистический коэффициент усиления вычисляется из соотношения

$$K_{S,n+1} = \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} = -\alpha A_{n+1} \exp\left\{-\alpha\tilde{x}_{n+1} + \frac{\alpha^2}{2}\tilde{D}_{n+1}\right\}.$$

При кусочно-линейной аппроксимации НФ воспользуемся изложенным в работе [29] методом. НФ $y=S(x)$ (Рис. 4.2) по оси x разбивается на q участков, в пределах каждого из которых она полагается линейной. При $x_1=\infty$ и $x_q=-\infty$ можно записать значения кусочно-линейной функцией для первого участка $S_1(x)\approx y_2+\alpha_1(x-x_2)$ при $\infty>x>x_2$, для остальных участков $S_k(x)\approx y_k+\alpha_k(x-x_k)$ при $x_k>x>x_{k+1}$, $k=2,...,q-1$, где α_k - значение крутизны k -го участка.

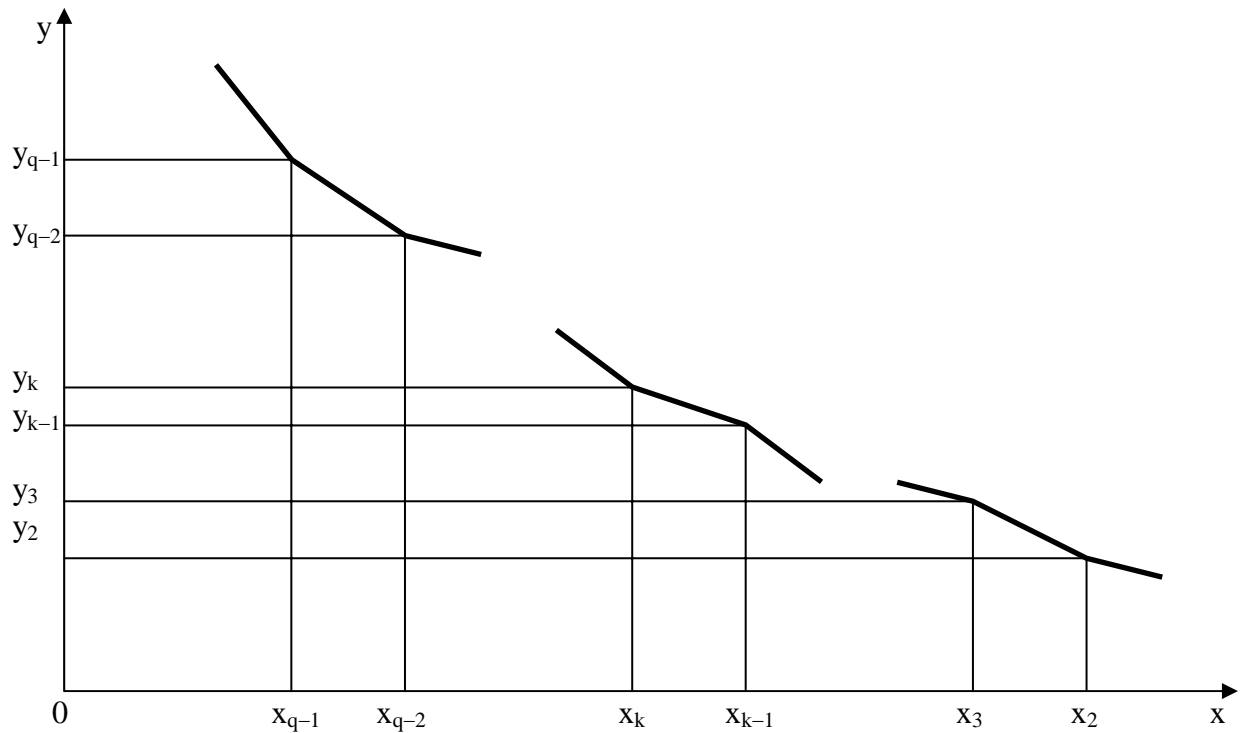


Рис. 4.2. Аппроксимация нелинейной характеристики
кусочно-линейной функцией

Заменив переменную x на $\xi = (x - \tilde{x}) / \sqrt{\tilde{D}}$, получим для первого и остальных участков

$$S_1(\xi \sqrt{\tilde{D}} + \tilde{x}) = y_2 + \alpha_1(\tilde{x} - x_2) + \alpha_1 \xi \sqrt{\tilde{D}} \quad \text{при } \infty > \xi > \xi_2,$$

$$S_k(\xi \sqrt{\tilde{D}} + \tilde{x}) = y_k + \alpha_k(\tilde{x} - x_k) + \alpha_k \xi \sqrt{\tilde{D}} \quad \text{при } \xi_k > \xi > \xi_{k+1}.$$

На основании принципа суперпозиции запишем для статических характеристик и коэффициентов усиления

$$S = \sum_{k=1}^q S_k, \quad K_0 = \sum_{k=1}^q K_{0k},$$

$$\text{где } S_k = \int_{\xi_k}^{\xi_{k+1}} S_k(\xi \sqrt{\tilde{D}} + \tilde{x}) \varphi(\xi) d\xi, \quad K_{0k} = \int_{\xi_k}^{\xi_{k+1}} \frac{\xi S_k(\xi \sqrt{\tilde{D}} + \tilde{x})}{\sqrt{\tilde{D}}} \varphi(\xi) d\xi,$$

$\varphi(\xi) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}\xi^2\right\}$ - функция плотности вероятностей.

После интегрирования получаем для первого участка (результаты приводятся для дискретного времени $t_n, n=1,2,\dots$)

$$S_{1n} = \frac{1}{2} \{y_2 + \alpha_1 [\tilde{x}_n - x_2]\} [\Phi(\xi_2) + 1] - \alpha_1 \varphi(\xi_2) \sqrt{\tilde{D}_n},$$

$$K_{01,n} = \frac{1}{\sqrt{\tilde{D}_n}} \{y_2 + \alpha_1 [\tilde{x}_n - x_2]\} [-\varphi(\xi_2)] + \\ + \alpha_1 \sqrt{\tilde{D}_n} \left\{ \frac{1}{2} [\Phi(\xi_2) + 1] - \xi_2 \varphi(\xi_2) \right\};$$

для k -го участка ($k=2,3,\dots,q-1$)

$$S_{kn} = \frac{1}{2} \{y_k + \alpha_k [\tilde{x}_n - x_k]\} [\Phi(\xi_{k+1}) - \Phi(\xi_k)] + \\ + \alpha_k \sqrt{\tilde{D}_n} [\varphi(\xi_k) - \varphi(\xi_{k+1})],$$

$$K_{0k,n} = \frac{1}{\sqrt{\tilde{D}_n}} \{y_k + \alpha_k [\tilde{x}_n - x_k]\} \left[\varphi(\xi_k) - \varphi(\xi_{k+1}) + \alpha_k \sqrt{\tilde{D}_n} \times \right. \\ \left. \times \left\{ \xi_k \varphi(\xi_k) - \xi_{k+1} \varphi(\xi_{k+1}) + \frac{1}{2} [\Phi(\xi_{k+1}) - \Phi(\xi_k)] \right\} \right];$$

для q -го участка

$$S_{qn} = \frac{1}{2} \{y_q + \alpha_q [\tilde{x}_n - x_q]\} [1 - \Phi(\xi_q)] + \alpha_q \sqrt{\tilde{D}_n} \varphi(\xi_q),$$

$$K_{0q,n} = \\ = \frac{1}{\sqrt{\tilde{D}_n}} \left[\{y_q + \alpha_q (\tilde{x}_n - x_q)\} \varphi(\xi_q) + \alpha_q \sqrt{\tilde{D}_n} \left\{ \xi_q \varphi(\xi_q) + \frac{1}{2} [1 - \Phi(\xi_q)] \right\} \right].$$

Значения интеграла вероятностей $\Phi(\xi_k) = 2 \int_0^{\xi_k} \varphi(\xi) d\xi$ и функции плотности вероятностей $\varphi(\xi_k)$ приведены в [29].

Рекуррентный алгоритм отношения правдоподобия можно получить если подставить представление НФ (4.38) в соотношение (4.119), а результат интегрирования в уравнение (4.118)

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n y_{n+1} S_{0,n+1} Q_{\lambda,n+1}^{-1}. \quad (4.146)$$

4.8.3.2. АЛГОРИТМЫ СИНТЕЗА ДЛЯ МНОГОМЕРНОГО МАРКОВСКОГО ПРОЦЕССА

При определении рекуррентных уравнений фильтрации и обнаружения используется многомерные уравнения марковских процессов (4.120), наблюдаемого сигнала (4.121) и функция правдоподобия (4.122).

После статистической линеаризации многомерных НФ последние принимают в дискретном времени форму

$$\begin{aligned} f_n(x_n) &\approx f_{0n}(x_n^*, D_n) + \frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*} (x_n - x_n^*) = \\ &= f_{0n} + \frac{\partial f_{0n}}{\partial x_n^*} (x_n - x_n^*), \end{aligned} \quad (4.147)$$

$$\begin{aligned} S_{n+1}(x_{n+1}) &\approx S_{0,n+1}(\tilde{x}_{n+1}, \tilde{D}_{n+1}) + \\ &+ \frac{\partial S_{0,n+1}(\tilde{x}_{n+1}, \tilde{D}_{n+1})}{\partial \tilde{x}_{n+1}} (x_{n+1} - \tilde{x}_{n+1}) = \\ &= S_{0,n+1} + \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} (x_{n+1} - \tilde{x}_{n+1}), \end{aligned} \quad (4.148)$$

где $f_n(x_n)$, f_{0n} , $(x_n - x_n^*)$, $(x_{n+1} - \tilde{x}_{n+1})$ - вектора размера $(r \times 1)$, $\partial f_{0n} / \partial x_n^*$ - матрица Якоби размера $(r \times r)$, (ij) – компонента которой $\partial f_{0i,n} / \partial x_{jn}^*$ (i - номер строки, j - номер столбца); $S_{n+1}(x_{n+1})$ и $S_{0,n+1}$ - вектора размера $(q \times 1)$, $\partial S_{0,n+1} / \partial \tilde{x}_{n+1}$ - матрица Якоби размера $(q \times r)$, (ij) компонента которой $\partial S_{0i,n+1} / \partial \tilde{x}_{j,n+1}$ (i - номер строки, j - номер столбца).

В результате подстановки соотношений (4.174) и (4.148) в выражения (4.120), (4.122) и (4.58), определения производных логарифма ФП-вектора $\Pi_{n+1}^{(1)}(\tilde{x}_{n+1})$ и матрицы $\Pi_{n+1}^{(2)}(\tilde{x}_{n+1})$, использования их в уравнениях (4.61) и (4.62) получаем рекуррентные уравнения оценки и дисперсии ошибки оценки при статистической линеаризации нелинейных функций

$$\begin{aligned}
x_{n+1}^* &= \tilde{x}_{n+1} + \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T \times \\
&\times \left[\left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \\
&\times [y_{n+1} - S_{0,n+1}], \quad x_0^* = x(0),
\end{aligned} \tag{4.149}$$

$$\begin{aligned}
D_{n+1} &= \tilde{D}_{n+1} - \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T \times \\
&\times \left[\left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \\
&\times \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1}, \quad D_0 = D(0),
\end{aligned} \tag{4.150}$$

где

$$\tilde{x}_{n+1} = x_n^* + \Delta t f_{0n}, \tag{4.151}$$

$$\tilde{D}_{n+1} = D_n + \Delta t \left(\frac{\partial f_{0n}}{\partial x_n^*} \right) D_n + \Delta t D_n \left(\frac{\partial f_{0n}}{\partial x_n^*} \right)^T + \Gamma_n Q_{\eta n} \Gamma_n^T. \tag{4.152}$$

Как следует из полученных результатов, уравнение оценки нелинейно относительно самого вектора оценки, а уравнение дисперсии зависит от вектора оценки и его необходимо решать совместно с уравнением оценки.

Используя уже известный подход для скалярных процессов, приходим к рекуррентному уравнению отношения правдоподобия

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n \sum_{i=1}^q y_{i,n+1} S_{0i,n+1} Q_{\lambda i,n+1}^{-1}.$$

При формировании рекуррентных алгоритмов фильтрации и обнаружения квазидетерминированных процессов (4.89) используется скалярное уравнение наблюдаемого сигнала (4.129) и

функция правдоподобия (4.130). После статистической линеаризации скалярной функции $S_{n+1}(a)$, последняя принимает в дискретном времени вид

$$\begin{aligned} S_{n+1}(a) &\approx S_{0,n+1}(a_n^*, D_n) + \frac{\partial S_{0,n+1}(a_n^*, D_n)}{\partial a_n^*} (a - a_n^*) = \\ &= S_{0,n+1} + \frac{\partial S_{0,n+1}}{\partial a_n^*} (a - a_n^*), \end{aligned} \quad (4.153)$$

где $\partial S_{0,n+1}/\partial a_n^*$ - вектор-строка размера $(1 \times k)$ из компонент $\partial S_{n+1}(a_n^*)/\partial a_{nj}^*$, $j=1, k$; $(a - a_n^*)$ - вектор-столбец размера $(k \times 1)$.

В результате подстановки (4.153) в ФП (4.62), определения производных логарифма ФП-вектора $\Pi_{n+1}^{(1)}(a_n^*)$ и матрицы $\Pi_{n+1}^{(2)}(a_n^*)$, использования их в уравнениях (4.61) и (4.62) при $K=2$ получаем рекуррентные уравнения оценки и дисперсии ошибки сигнала при статистической линеаризации нелинейных функций

$$a_{n+1}^* = a_n^* + D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T \left[\left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \quad (4.154)$$

$$\times [y_{n+1} - S_{0,n+1}], \quad a_0^* = a(0),$$

$$\begin{aligned} D_{n+1} &= D_n - D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right) D_n \times \\ &\times \left[\left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T + Q_{\lambda,n+1} \right]^{-1}, \quad D_0 = D(0), \end{aligned} \quad (4.155)$$

где выражения в квадратных скобках являются скалярными величинами.

К рекуррентному алгоритму отношения правдоподобия можно прийти после подстановки выражения (4.153) в интегральное соотношение (4.119) и последующего вычисления по формуле (4.118)

$$\Lambda_{n+1} = \Lambda_n + \Lambda_n y_{n+1} S_{0,n+1} Q_{\lambda,n+1}^{-1}.$$

В скалярно-координатной форме уравнения (4.154) и (4.155) принимают вид

$$a_{i,n+1}^* = a_{in}^* + \frac{\sum_{l=1}^k D_{il,n} \frac{\partial S_{0,n+1}}{\partial a_{ln}^*} [y_{n+1} - S_{0,n+1}]}{\sum_{l,m=1}^k D_{lm,n} \frac{\partial S_{0,n+1}}{\partial a_{ln}^*} \frac{\partial S_{0,n+1}}{\partial a_{mn}^*} + Q_{\lambda,n+1}}, \quad a_{i0}^* = a_i(0),$$

$$D_{ij,n+1} = D_{ij,n} - \frac{\sum_{l,m=1}^k D_{il,n} D_{mj,n} \frac{\partial S_{0,n+1}}{\partial a_{ln}^*} \frac{\partial S_{0,n+1}}{\partial a_{mn}^*}}{\sum_{l,m=1}^k D_{lm,n} \frac{\partial S_{0,n+1}}{\partial a_{ln}^*} \frac{\partial S_{0,n+1}}{\partial a_{mn}^*} + Q_{\lambda,n+1}}, \quad D_{ij,0} = D_{ij}(0).$$

4.8.3.3. АЛГОРИТМЫ СИНТЕЗА ДЛЯ НЕПРЕРЫВНОГО ВРЕМЕНИ

Для многомерных процессов уравнения фильтрации в непрерывном времени при статистической линеаризации нелинейностей получают путем предельного перехода в уравнениях (4.149) и (4.150). Подставим в эти уравнения экстраполированные параметры (4.151) и (4.152), а также учтем $\Gamma_n = \Delta t \Gamma(t_n)$, $Q_{\eta n} = \mathcal{N}_{\eta n} / \Delta t$, $Q_{\lambda,n+1} = \mathcal{N}_{\lambda,n+1} / \Delta t$.

После предельного перехода $\Delta t \rightarrow 0$ получаем дифференциальные уравнения оценок и дисперсии ошибок оценок

$$\frac{dx^*(t)}{dt} = f_{0t}(x_t^*, D_t) + D(t) \left(\frac{\partial S_{0t}(x_t^*, D_t)}{\partial x_t^*} \right)^T \times$$

$$\times \mathcal{N}_{\lambda t}^{-1} [y(t) - S_{0t}(x_t^*, D_t)], \quad x^*(t_0) = x(0),$$

$$\frac{dD(t)}{dt} = \frac{\partial f_{0t}(x_t^*, D_t)}{\partial x_t^*} D(t) + D(t) \left(\frac{\partial f_{0t}(x_t^*, D_t)}{\partial x_t^*} \right)^T + \Gamma_t \mathcal{N}_{\lambda t} \Gamma_t^T -$$

$$- D(t) \left(\frac{\partial S_{0t}(x_t^*, D_t)}{\partial x_t^*} \right)^T \mathcal{N}_{\lambda t}^{-1} \left(\frac{\partial S_{0t}(x_t^*, D_t)}{\partial x_t^*} \right) D(t), \quad D(t_0) = D(0).$$

Для квазидетерминированных процессов после предельного перехода в уравнениях (4.154) и (4.155) соответствующие дифференциальные уравнения принимают вид

$$\begin{aligned} \frac{da^*(t)}{dt} &= D(t) \left(\frac{\partial S_{0t}(a_t^*, D_t)}{\partial a_t^*} \right)^T \times \\ &\times \mathcal{N}_{\lambda t}^{-1} [y(t) - S_{0t}(a_t^*, D_t)], \quad a^*(t_0) = a(0), \\ \frac{dD(t)}{dt} &= -\mathcal{N}_{\lambda t}^{-1} D(t) \left(\frac{\partial S_{0t}(a_t^*, D_t)}{\partial a_t^*} \right)^T \left(\frac{\partial S_{0t}(a_t^*, D_t)}{\partial a_t^*} \right) D(t), \quad D(t_0) = D(0). \end{aligned}$$

4.9. ЭКСТРАПОЛЯЦИЯ И ИНТЕРПОЛЯЦИЯ

4.9.1. ОБЩИЕ ОПРЕДЕЛЕНИЯ

Рассматриваемые задачи экстраполяции (прогнозирования) и интерполяции являются задачами оценивания, для которых оценки определяются соответственно в моменты времени $t_1 > t$ или $t_0 < t_1 < t$, где (t_0, t) - время измерения наблюдаемого процесса.

Качество работы алгоритмов экстраполяции и интерполяции непрерывных динамических систем, как и для ранее рассмотренных задач фильтрации, определяется решающим правилом $\gamma_{t_1}(y_{t_0}^t)$, $t_1 \geq t$. Оптимальному решающему правилу при байесовском подходе соответствует $\hat{\gamma}_{t_1}$, для которого относительно всех других решающих правил γ_{t_1} , обеспечивается минимум среднего риска $R(\hat{\gamma}_{t_1}) \leq R(\gamma_{t_1})$ (см. раздел 4.1.2). При квадратичной функции потерь и апостериорном риске $R(\gamma_{t_1} | y_{t_0}^t) = M\{[x_{t_1} - M\{x_{t_1} | y_{t_0}^t\}]^2 | y_{t_0}^t\}$ оптимальная оценка определяется как апостериорное среднее

$$\begin{aligned} \hat{\gamma}_{t_1}(y_{t_0}^t) &= \hat{x}(t_1) = M\{x_{t_1} | y_0^t\} = \\ &= \int_{-\infty}^{\infty} x(t_1) p(x_{t_1} | y_{t_0}^t) dx_{t_1} = \int_{-\infty}^{\infty} x(t_1) w_t(x_{t_1}) dx_{t_1}. \end{aligned} \quad (4.156)$$

Дисперсия ошибки оценки принимает вид

$$D(t_1) = \int_{-\infty}^{\infty} [x(t_1) - \hat{x}(t_1)]^2 w_t(x_{t_1}) dx_{t_1} . \quad (4.157)$$

Поэтому для рассматриваемых алгоритмов, как и для алгоритмов фильтрации, основополагающая роль в получении оценки принадлежит апостериорной плотности вероятностей $w_t(x_{t_1}), t_1 \geq t$.

Ограничимся обсуждением алгоритмов в дискретном времени. На основании формул (4.156) и (4.157) апостериорная оценка и дисперсия ошибки оценки определяется из соотношений

$$\hat{x}_m = \int_{-\infty}^{\infty} x_m w_{nm}(x_m) dx_m , \quad (4.158)$$

$$D_m = (h_{2m})^{-1} = \int_{-\infty}^{\infty} (x_m - \hat{x}_m)^2 w_{nm}(x_m) dx_m , \quad (4.159)$$

где апостериорная плотность вероятностей $w_{nm}(x_m) = p(x_m | y_1^n)$.

При $m=n$ АПВ соответствует задаче фильтрации, при $m>n$ - задаче экстраполяции, при $m<n$ - задаче интерполяции.

4.9.2. АЛГОРИТМ ЭКСТРАПОЛЯЦИИ

Поскольку оцениваемый процесс марковский, на первом шаге экстраполяции при отсутствии наблюдения имеем ЭПВ (4.35)^{*)}

$$w_{n,n+1}(x_{n+1}) = \int_{-\infty}^{\infty} \pi(x_{n+1} | x_n) w_n(x_n) dx_n ,$$

где $w_n(x_n)$ - апостериорная плотность вероятностей, соответствующая последнему измерению на шаге n , $\pi(x_{n+1}|x_n)$ - переходная плотность вероятностей, параметры которой определяются уравнением ФПК. Продолжая процедуру получения экстраполяционных оценок, приходим к рекуррентному алгоритму

^{*)} Для удобства дальнейших преобразований обозначение экстраполированной плотности изменено.

$$w_{n,m+1}(x_{m+1}) = \int_{-\infty}^{\infty} \pi(x_{m+1} | x_m) w_{nm}(x_m) dx_m \quad (4.160)$$

при начальном условии $w_{nn}(x_n) = w_n(x_n)$.

Подставляя выражение (4.160) в формулы (4.158) и (4.159), получим экстраполяционные оценку и дисперсию ошибки оценки.

Рассмотрим точное решение задачи экстраполяции одномерного гауссовского стационарного марковского процесса (2.46) с коэффициентом корреляции на интервале дискретизации

$$r(\Delta t) = k_x(\Delta t) / D_x = \exp\{-\alpha|\Delta t|\}.$$

Переходная плотность вероятностей для этого случая на основании формулы (2.4) равна

$$\pi(x_{m+1} | x_m) = [2\pi D_x (1 - r^2(\Delta t))]^{-1/2} \exp\left\{-\frac{[x_{m+1} - x_m r(\Delta t)]^2}{2D_x [1 - r^2(\Delta t)]}\right\}, \quad (4.161)$$

где $r(\Delta t) = r(t_1, t_2)$, $\Delta t = t_2 - t_1$, $D_x = \alpha \mathcal{N}_\eta / 2$ - априорная дисперсия марковского процесса.

Поскольку x_n и помеха – дискретный белый шум η_n являются гауссовскими, то АПВ также является гауссовской плотностью вероятностей

$$w_{nm}(x_m) = A_1 \exp\left\{-\frac{1}{2} h_{2nm} (x_m - \hat{x}_{nm})^2\right\}. \quad (4.162)$$

Подставляя формулы (4.161) и (4.162) в выражение (4.160) после вычисления интеграла и приравнивания соответствующих членов по разные стороны равенства этого выражения, приходим к рекуррентному алгоритму определения оценки и дисперсии ошибки оценки прогноза

$$\hat{x}_{n,m+1} = \hat{x}_{nm} r(\Delta t), \quad \hat{x}_{nn} = \hat{x}_n,$$

$$h_{2n,m+1} = (D_{n,m+1})^{-1} = h_{2n,m} [h_{2n,m} D_x (1 - r^2(\Delta t)) + r^2(\Delta t)]^{-1}, \quad h_{2nn} = h_{2n}$$

или $D_{n,m+1} = D_{nm} r^2(\Delta t) + D_x [1 - r^2(\Delta t)]$, $D_{nn} = D_n$, где начальные условия x_n и D_n являются оценкой и дисперсией ошибки оценки при последнем измерении процесса.

Приведем также рекуррентные выражения, в которых начальные условия входят непосредственно

$$\hat{x}_{n,m+1} = \hat{x}_n r^{m-n+1}(\Delta t), \quad (4.163)$$

$$D_{n,m+1} = D_n r^{2(m-n+1)}(\Delta t) + D_x [1 - r^{2(m-n+1)}(\Delta t)]. \quad (4.164)$$

Выражения в форме (4.163) и (4.164) пригодны для корреляционной функции экспоненциального вида. Для других процессов, как например, фрактальных корреляционная функция имеет степенной характер. Поэтому, введя параметр смещения $k=m-n+1$, эти выражения можно записать в более общем виде

$$\hat{x}_{n+k} = \hat{x}_n r(k; \Delta t),$$

$$D_{n+k} = D_n r^2(k; \Delta t) + D_x [1 - r^2(k; \Delta t)].$$

Полученные результаты можно обобщить на случай, когда измерения производятся точно. В этой частной задаче $D_n=0$ и x_n соответствует последнему измерению без помех. Рекуррентные алгоритмы экстраполяции в этом случае принимают вид

$$\hat{x}_{n+k} = x_n r(k; \Delta t), \quad (4.165)$$

$$D_{n+k} = D_x [1 - r^2(k; \Delta t)]. \quad (4.166)$$

4.9.3. АЛГОРИТМ ИНТЕРПОЛЯЦИИ

В инженерной практике в зависимости от роли, которые выполняют моменты времени оценивания t_1 и наблюдения t , рассматриваются три вида задач интерполяции. Если зафиксировать момент времени t , то оценка рассматривается как функция изменения t_1 , чему соответствует обратная интерполяция. Если зафиксировать момент времени t_1 , то оценка строится как функция изменения момента t - прямая интерполяция. И, наконец, можно зафиксировать разность $t-t_1$ и рассматривать задачу интерполяции с постоянным запаздыванием. Ниже рассматривается задача обратной интерполяции в дискретном времени. С другими задачами интерполяции можно ознакомиться в работах [5,47,48].

Как и в предыдущем случае, необходимо найти АПВ $w_{nm}(x_m) = p(x_m | y_1^n)$ при $m < n$. Она позволяет находить оценки и

дисперсии ошибки оценки интерполяции. При решении задачи введем условную плотность вероятностей, которая определяется формулой

$$p(x_m | x_{m+1}, y_1^m) = \frac{p(x_m, x_{m+1} | y_1^m)}{\int_{-\infty}^{\infty} p(x_m, x_{m+1} | y_1^m) dx_m}. \quad (4.167)$$

Учитывая равенство $p(x_{m+1}, x_m | y_1^m) = \pi(x_{m+1} | x_m) p(x_m | y_1^m)$, соотношение (4.167) можно представить в форме

$$p(x_m | x_{m+1}, y_1^m) = \frac{\pi(x_{m+1} | x_m) w_m(x_m)}{\int_{-\infty}^{\infty} \pi(x_{m+1} | x_m) w_m(x_m) dx_m}, \quad w_m(x_m) = p(x_m | y_1^m). \quad (4.168)$$

На основании марковского характера процесса m_m можно также записать

$$p(x_m | x_{m+1}, y_1^m) = p(x_m | x_{m+1}, y_1^n), \quad m < n. \quad (4.169)$$

Подставляя (4.168) с учетом соотношения (4.169) в выражение $w_{n,m+1}(x_{m+1}) p(x_m | x_{m+1}, y_1^n) = p(x_m, x_{m+1} | y_1^n)$ и интегрируя затем по x_{m+1} , получаем

$$\int_{-\infty}^{\infty} w_{n,m+1}(x_{m+1}) \frac{\pi(x_{m+1} | x_m) w_m(x_m)}{\int_{-\infty}^{\infty} \pi(x_{m+1} | x_m) w_m(x_m) dx_m} dx_{m+1} = w_{nm}(x_m) \quad (4.170)$$

где $m=n-1, n-2, \dots, 1$, начальное условие $w_{nn}(x_n) = w_n(x_n)$.

В отличие от задачи экстраполяции, когда требовалось знание конечной плотности вероятности $w_n(x_n)$, в задаче обратной интерполяции необходим набор АПВ. Поэтому в этом случае, кроме блока фильтрации, необходим блок запоминающих функций $w_m(x_m)$.

Этот блок подает соответствующие значения этих функций в последовательные моменты времени на схему с обратной связью,

структура которой определяется рекуррентным соотношением (4.170).

Как и в предыдущем случае, представим результаты точного решения задачи интерполяции одномерного гауссовского марковского процесса (2.46). Подставим (4.161), а также соотношения $w_{n,m+1}(x_{m+1})$, $w_{nm}(x_m)$, и $w_m(x_m)$ в форму (4.162) в выражение (4.170). В результате получаем

$$\hat{x}_{nm} = \frac{\hat{x}_{n,m+1}r(\Delta) + \hat{x}_m h_{2m} D_x [1 - r^2(\Delta t)]}{h_{2m} D_x [1 - r^2(\Delta t)] + r^2(\Delta t)},$$

$$h_{2nm} = D_{nm}^{-1} =$$

$$= \frac{h_{2n,m+1} \{h_{2m} D_x [1 - r^2(\Delta t)] + r^2(\Delta t)\}^2}{h_{2n,m+1} D_x^2 [1 - r^2(\Delta t)]^2 h_{2m} + h_{2n,m+1} D_x [1 - r^2(\Delta t)] r^2(\Delta t) + r^2(\Delta t)},$$

где начальные значения $\hat{x}_{nm} = \hat{x}_n$, $h_{2nm} = h_{2n}$, а характеристики \hat{x}_m и $h_{2m} = D_m^{-1}$ получены из уравнений фильтрации (4.75) и (4.76).

В заключение отметим, что дополнительное наблюдение (y_τ , $t_1 < \tau < t$) может улучшить в задачах интерполяции оптимальное оценивание или, по крайней мере, оставить его прежним. В последнем случае оцениваемый параметр статистически не зависит от дополнительного наблюдения

$$M\{x_{t_1} \mid y_1^t\} = M\{x_{t_1} \mid y_1^{t_1}\}.$$

4.10. АЛГОРИТМЫ СИНТЕЗА ДЛЯ СЛУЧАЙНЫХ ТОЧЕЧНЫХ ПРОЦЕССОВ

4.10.1. ФОРМУЛИРОВАНИЕ ЗАДАЧИ

Одной из важных областей приложения моделей случайных процессов – точечных процессов (случайных потоков), являются системы передачи информации фотонным излучением. В фотонных излучениях, к которым относятся оптическое, рентгеновское и гамма-излучения, заметно проявляется квантовая структура, обусловленная дискретной природой электромагнитного излучения. Возрастающий интерес к проблемам передачи информации фотонным излучением

вызван освоением для этих целей наиболее коротковолнового диапазона электромагнитного излучения и внедрением фотонных устройств в системы связи для управления динамическими системами, а также в коммуникационные системы передачи данных. Случайным потоком моделируются импульсные процессы на выходе детекторов излучений, которые далее поступают в устройства обработки фотонных систем связи для решения задач обнаружения и фильтрации передаваемым этим излучением информации. В этом отношении указанные системы по выполняемым задачам и предъявляемым к ним требованиям – обеспечению надежной передачи информации (в смысле получения оценок процесса с наименьшей погрешностью), не отличаются от рассмотренных ранее систем связи. Отличие состоит в характере наблюдаемого сигнала, модель реализации которого представляет собой модулированный импульсный поток с параметром – модулируемой интенсивностью $S_t(x_t)$, а также в наличии квантового шума $\chi(t)$.

Теория оценивания параметров фотонного излучения базируется на вероятностной концепции квантовых измерений [49,13]. Согласно этой теории при амплитудной модуляции электромагнитного поля посылаемое источником излучения сообщение связано со средним числом (отсчетов в условную единицу времени) зарегистрированных квантов (модулируемой интенсивностью $S_t(x_t)$) и может быть обнаружено с помощью детектора при наблюдении за скоростью счета импульсных сигналов – реализации случайного импульсного потока (2.84). При описании этих сигналов важно располагать сведениями не только о среднем числе отсчетов, но и о характере флуктуаций их числа около среднего значения, которые задаются функциями распределения или другими статистическими характеристиками. Эти флуктуации (квантовый шум $\chi(t)$), обусловленные статистическим характером квантовых измерений, расцениваются как мешающие составляющие сигналов и в дальнейшем квалифицируются как эквивалентный шум в задачах обнаружения и фильтрации. Если модуляция фотонного излучения отсутствует, что равнозначно статистической независимости регистрируемых детектором квантов, то их число описывается пуассоновским распределением.

При модуляции излучения случайным процессом вид функции распределения меняется. Для строгого решения задачи оценивания

статистику числа отсчетов целесообразно описать с помощью заменяющих эту функцию системами корреляционных или моментных функций первого, второго и более высоких порядков. Для их получения необходимо более полное описание модулированного поля излучения с помощью квантовых корреляционных функций. Как и для обычных случайных процессов, указанная система функций определяется путем функционального дифференцирования логарифма квантовомеханического функционала [50]. Можно показать [51], что сформированный в соответствии с условиями задачи указанный функционал совпадает с классическим функционалом (1.6), описывающим случайный процесс на выходе детектора. Им оказывается случайный импульсный поток (2.84). Таким образом, при взаимодействии поля излучения с идеальным детектором процесс на выходе последнего аппроксимируется случайным импульсным потоком $\eta^N(\tau) = \{\eta_{t_0}^{N\tau}, t_0 < \tau < t\}$, где $\eta_{t_0}^{N\tau}$ - отдельная реализация этого потока. Статистические характеристики случайного импульсного потока дают необходимую информацию о поле во временной области в месте расположения детектора. Причем квантовые корреляционные функции, обусловленные корреляциями сообщений и поэтому включающие в себя полную об оцениваемом процессе, однозначно связаны с корреляционными функциями случайного импульсного потока. Следовательно, если сообщение обладает корреляциями вплоть до n -го порядка, то те же по порядку корреляции существуют у модулируемой интенсивности. Таким образом, основным информационным параметром при измерении фотонного излучения является модулируемая интенсивность, которая однозначно связана с передаваемыми этим излучением сообщениями.

При непосредственном наблюдении доступной является не сама модулируемая интенсивность $S_t(x_t)$, зависящая от объединенных в вектор x_t самих сообщений и параметров среды распространения излучения, а некоторая связанная с ней реализация случайного импульсного потока. Причем случайность в наблюдении обусловлена статистическим характером квантовых измерений и, возможно, векторным процессом x_t , если его компоненты случайные. В общем виде модель вектора наблюдаемого сигнала (случайного импульсного процесса на выходе детектора) описывается случайной функцией от многокомпонентного сообщения и других параметров излучения

$$y(t)=\Psi_t[S_t(x_t),\chi(t),\lambda(t),\Theta].$$

Отметим, что при статистическом описании канала связи переходы входных сигналов x в выходные y задаются условной вероятностью – функцией правдоподобия (ФП). В математическом плане ФП полностью описывает статистические процессы передачи информации в канале связи.

4.10.2. РЕКУРРЕНТНЫЕ УРАВНЕНИЯ ФИЛЬТРАЦИИ И ОБНАРУЖЕНИЯ

Уравнения наблюдения в дискретном времени с учетом обозначения наблюдаемого сигнала (точечного процесса) $y_n=N_n$ имеет вид

$$N_{n+1}=\Psi_{n+1}[S_{n+1}(x_{n+1}),\chi_{n+1},\Theta], \quad (4.171)$$

где x_{n+1} - скалярный или многомерный марковский процесс.

Остальные обозначения уравнения (4.171) обсуждались в разд.4.6.

При формировании уравнения наблюдения необходимо обратить внимание на то, что ошибки в обнаружении и фильтрации вносят квантовый шум χ и аддитивная помеха λ . В фотонных системах в возникновение ошибочных решений наиболее существенный вклад вносит квантовый шум. Поэтому основное внимание при формировании алгоритмов обработки фотонного излучения уделяется учету этого дестабилизирующего фактора, что нашло отражение в записи уравнения (4.171). Учет помехи λ (собственный шум системы измерения и другие приведенные ко входу системы помехи), описываемой аддитивным гауссовским процессом, будет проведен в конце раздела.

Рассматривается модель наблюдаемого сигнала в виде последовательности целочисленных отсчетов случайного потока $\{N_{n+1} \gg 1\}$ на непересекающихся, примыкающих друг к другу временных подынтервалах $(t_n, t_{n+1})^{*)}$

*) При конкретизации задач синтеза могут использоваться и другие виды наблюдаемых сигналов: случайный ординарный поток точек, последовательность случайных подынтервалов между точками ординарного потока и профильтрованный случайный поток. С алгоритмами фильтрации и обнаружения для перечисленных наблюдаемых сигналов можно ознакомиться в работе [13].

Рассмотрим на временном интервале (t_1, t_{n+1}) условную вероятность оцениваемого процесса $x_1^{n+1} = \{x_1, x_2, \dots, x_{n+1}\}$ в дискретные моменты времени t_i , $i=1, 2, \dots, n+1$ и случайной величины Θ при условии осуществления события $N_1^{n+1} = \{N_1, N_2, \dots, N_{n+1}\}$ - последовательности из $n+1$ целочисленных отсчетов наблюдаемого сигнала в те же моменты времени - $P(x_1^{n+1}, \Theta | N_1^{n+1})$. Начальное значение x_0 , соответствующее моменту времени t_0 , удовлетворяет априорной плотности вероятностей $p(x_0)$.

На основании соотношения для условных вероятностей (1.4) можно записать

$$P(x_1^{n+1}, \Theta | N_1^{n+1}) = P(x_1^{n+1}, \Theta | N_{n+1}, N_1^n) = \frac{P(x_1^{n+1}, \Theta, N_{n+1} | N_1^n)}{P(N_{n+1} | N_1^n)}, \quad (4.172)$$

где условная вероятность в числителе (4.172) с учетом условия факторизации (2.31) и независимости x_{n+1} от N_1^n имеет вид

$$\begin{aligned} P(x_1^{n+1}, \Theta, N_{n+1} | N_1^n) = \\ = \begin{cases} P(x_1^n, \Theta = 1 | N_1^n) \pi(x_{n+1} | x_n) P(N_{n+1} | x_1^{n+1}, N_1^n), \\ P(x_1^n, \Theta = 0 | N_1^n) \delta(x_{n+1}) P(N_{n+1} | N_1^n). \end{cases} \quad (4.173) \end{aligned}$$

Введенная во второй строчке формулы (4.173) взамен $\pi(x_{n+1} | x_n)$ дельта- функция $\delta(x_{n+1})$ соответствует состоянию $\Theta=0$, т.е. отсутствию полезного сигнала.

В условной вероятности $P(N_{n+1} | x_1^{n+1}, N_1^n)$ опустим N_1^n , так как при $\Theta=1$ отсчет N_{n+1} зависит только от x_1^{n+1} , а последовательность отсчетов N_1^n не дает никакой дополнительной информации при наличии x_1^{n+1} . Так как отсчет x_{n+1} в силу марковости процесса зависит только от значения на предшествующем шаге, то наблюдаемый сигнал на шаге $n+1$ также является марковским процессом, зависящим от исходного марковского процесса для того же момента времени, и рассматриваемая условная вероятность принимает вид $P(N_{n+1} | x_{n+1})$. Эта вероятность называется функцией правдоподобия и характеризует статистические процессы передачи информации по

каналу связи. Кроме того, при $\Theta=0$ $P(N_{n+1}|N_1^n)=P(N_{n+1})$ в виду независимости отсчетов точечного процесса.

Используя далее методику, изложенную в разд.4.6.2, заменив y_{n+1} и y_n соответственно на N_{n+1} и N_n приходим к рекуррентным алгоритмам обнаружения

$$\Lambda_{n+1} = \frac{\Lambda_n}{P(N_{n+1})} \int_{-\infty}^{\infty} P(N_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1}, \quad \Lambda_0 = \Lambda(0) \quad (4.174)$$

и фильтрации

$$w_{n+1}(x_{n+1}) = c_{n+1} P(N_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}), \quad w_0(x_0) = p(x_0). \quad (4.175)$$

где $\tilde{w}_n(x_{n+1})$ и $w_{n+1}(x_{n+1})$ - соответственно экстраполированная (ЭПВ) и апостериорная (АПВ) плотности вероятности,

$$c_{n+1} = \left[\int_{-\infty}^{\infty} P(N_{n+1} | x_{n+1}) \tilde{w}_n(x_{n+1}) dx_{n+1} \right]^{-1}.$$

Полученные рекуррентные алгоритмы совместного обнаружения и фильтрации позволяют формировать в дискретном времени шаг за шагом, начиная с априорных данных (начальных условий), оценки марковского процесса и статистики отношения правдоподобия с помощью многоканального устройства, структурная схема которого представлена на Рис. 4.3.

Устройство включения (УВ) в конечный момент времени t_m пропускает сформированную в блоке формирования отношения правдоподобия (БФОП) статистику $\ln \Lambda_m$ на пороговое устройство (ПУ). В нем в результате сравнения $\ln \Lambda_m$ с порогом выносится решение о наличии ($\Theta=1$) или отсутствии ($\Theta=0$) полезного сигнала. После превышения порога в ПУ (решение γ_1) с выхода ключевого устройства (КУ) поступает полученная в блоке фильтрации (БФ) оценка (если γ_1 верно) или псевдооценка (если γ_1 ложно).

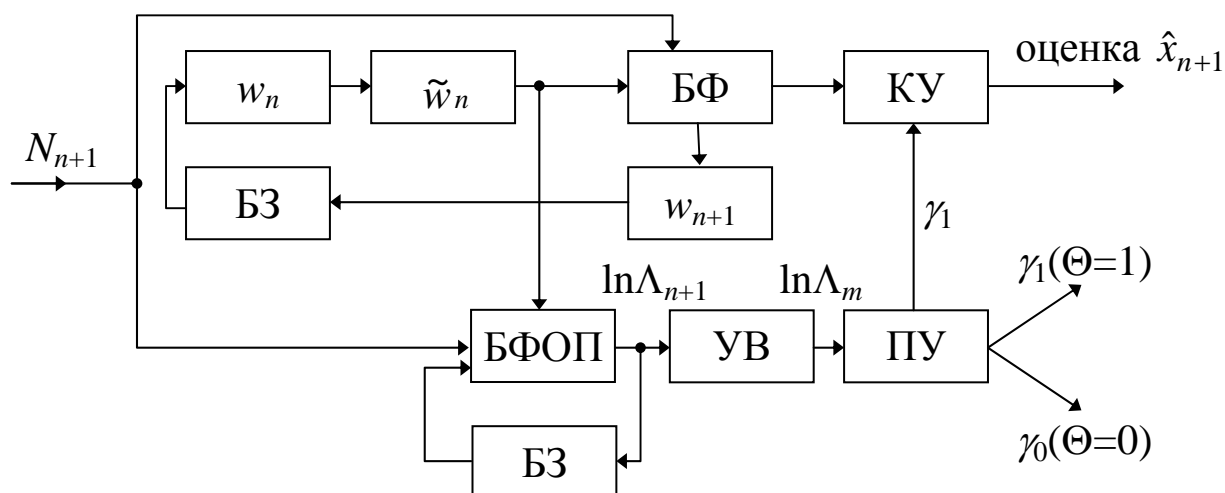


Рис. 4.3. Структурная схема системы совместного обнаружения и фильтрации для случайного точечного процесса.

4.10.3. УРАВНЕНИЯ ДЛЯ АПОСТЕРИОРНЫХ ПАРАМЕТРОВ

Решение задачи синтеза, как уже было отмечено ранее, также возможно путем определения заменяющих АПВ параметров: апостериорного среднего (4.39) и апостериорных центральных моментов разного порядка (4.40) и (4.41). В качестве исходных выражений рассматриваются рекуррентные уравнения обнаружения (4.174) и фильтрации (4.175). Используя методику разд.4.6.3, можно получить систему уравнений относительно апостериорных параметров разного порядка (4.45), (4.46) и (4.47), где $\Pi_{n+1}(x_{n+1}) = \ln P(N_{n+1}|x_{n+1})$ - логарифм ФП. В дальнейшем задача решается в рамках корреляционной теории, т.е. на основе использования апостериорных параметров первых двух порядков, что приводит к необходимости гауссовской аппроксимации, требования которой рассматривались в разд.4.8.1.

В задачах синтеза при передаче информации фотонным излучением могут быть получены только нелинейные уравнения оценки и дисперсии ошибки оценки. Квазиоптимальный подход для решения этих задач приводит к необходимости использования методов локальной гауссовской аппроксимации и статистической линеаризации нелинейных функций.

4.10.3.1. МЕТОД ЛОКАЛЬНОЙ ГАУССОВСКОЙ АППРОКСИМАЦИИ

Для скалярных процессов рекуррентные уравнения для апостериорного среднего оценки x_{n+1}^* и дисперсии ошибки оценки $D_{n+1}=(h_{2,n+1})^{-1}$ при $K=2$ на основании соотношений (4.45) и (4.47) имеют вид

$$x_{n+1}^* = \tilde{x}_{n+1} + D_{n+1} \Pi_{n+1}^{(1)}(\tilde{x}_{n+1}), \quad x_0^* = x(0), \quad (4.176)$$

$$D_{n+1} = [g_{2,n+1} - \Pi_{n+1}^{(2)}(\tilde{x}_{n+1})]^{-1} = \tilde{D}_{n+1} \{1 - \tilde{D}_{n+1} \Pi_{n+1}^{(2)}(\tilde{x}_{n+1})\}^{-1}, \quad (4.177)$$

$$D_0 = D(0),$$

где \tilde{x}_{n+1} и \tilde{D}_{n+1} - экстраполированные оценки и дисперсии ошибки оценки, для марковских процессов определяются из выражений (4.112) и (4.114).

С учетом замечания о зависимости ФП на шаге $n+1$ от марковского процесса для того же момента времени одношаговая функция правдоподобия записывается в форме распределения Пуассона

$$P(N_{n+1} | x_{n+1}) = \frac{\{[S_{n+1}(x_{n+1}) + \nu_{0,n+1}]\Delta t\}^{N_{n+1}}}{N_{n+1}!} \times \quad (4.178)$$

$$\times \exp\{-[S_{n+1}(x_{n+1}) + \nu_{0,n+1}]\Delta t\},$$

где $S_{n+1}(x_{n+1})$ - модулируемая интенсивность (полезный сигнал), $\nu_{0,n+1}$ - заданная детерминированная составляющая интенсивности случайного потока, включающая фоновую и другие известные неслучайные сигналы.

Первая и вторая производные логарифма ФП по $x_{n+1}=\tilde{x}_{n+1}$ соответственно равны

$$\Pi_{n+1}^{(1)}(\tilde{x}_{n+1}) = \frac{\partial S_{n+1}(\tilde{x}_{n+1}) / \partial \tilde{x}_{n+1}}{S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}} \{N_{n+1} - [S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}]\Delta t\}, \quad (4.179)$$

$$\begin{aligned} \Pi_{n+1}^{(2)}(\tilde{x}_{n+1}) = & \frac{N_{n+1} - [S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}]\Delta t}{S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}} \cdot \frac{\partial^2 S_{n+1}(\tilde{x}_{n+1})}{\partial x_{n+1}^2} - \\ & - N_{n+1} \left[\frac{\partial S_{n+1}(\tilde{x}_{n+1}) / \partial \tilde{x}_{n+1}}{S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}} \right]^2. \end{aligned} \quad (4.180)$$

В результате использования выражений (4.179) и (4.180) соответственно в (4.176) и (4.177) получим рекуррентные уравнения оценки и дисперсии ошибки оценки. К рекуррентному алгоритму ОП можно прийти после подстановки (4.178) в выражение (4.174) с

учетом $P(N_{n+1}) = \frac{(\nu_{0,n+1}\Delta t)^{N_{n+1}}}{N!} \exp\{-\nu_{0,n+1}\Delta t\}$. Имеем

$$\begin{aligned} \Lambda_{n+1} = \Lambda_n \int_{-\infty}^{\infty} \left\{ \frac{S_{n+1}(x_{n+1})}{\nu_{0,n+1}} + 1 \right\}^{N_{n+1}} \exp\{-S_{n+1}(x_{n+1})\Delta t\} \times \\ \times \tilde{w}_n(x_{n+1}) dx_{n+1}, \quad \Lambda_0 = \Lambda(0). \end{aligned} \quad (4.181)$$

Реализация рекуррентных алгоритмов (4.176), (4.177) и (4.181) после подстановки соотношений (4.179) и (4.180) в уравнения (4.176) и (4.177) наталкивается на серьезные технические трудности. Задача заключается в привлечении таких приближенных методов, которые при минимальных потерях качества обработки позволили бы получить пригодные для использования на практике расчетные процедуры. Одним из таких подходов является рассмотренный в разд. 4.8.2 расширенный фильтр Калмана. Следуя этому подходу, усредним функцию $\Pi_{n+1}^{(2)}(x_{n+1})$ относительно условной вероятности $P(N_{n+1}|x_{n+1})$. Учитывая, что $M\{N_{n+1}|x_{n+1}\} = [S_{n+1}(x_{n+1}) + \nu_{0,n+1}]\Delta t$, получаем

$$\bar{\Pi}_{n+1}^{(2)}(\tilde{x}_{n+1}) = - \frac{\Delta t [\partial S_{n+1}(\tilde{x}_{n+1}) / \partial \tilde{x}_{n+1}]^2}{S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}}. \quad (4.182)$$

В результате подстановки соотношений (4.179) и (4.182) соответственно в выражения (4.176) и (4.177) после преобразований приходим к рекуррентным уравнениям фильтрации

$$x_{n+1}^* = \tilde{x}_{n+1} + D_{n+1} \frac{\partial S_{n+1}(\tilde{x}_{n+1}) / \partial \tilde{x}_{n+1}}{S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}} \times$$

$$\times \{N_{n+1} - [S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}] \Delta t\}, \quad x_0^* = x(0),$$
(4.183)

$$D_{n+1} = \tilde{D}_{n+1} -$$

$$- \frac{\tilde{D}_{n+1}^2 [\partial S_{n+1}(\tilde{x}_{n+1}) / \partial \tilde{x}_{n+1}]^2}{\tilde{D}_{n+1} [\partial S_{n+1}(\tilde{x}_{n+1}) / \partial \tilde{x}_{n+1}]^2 + [S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}] / \Delta t},$$
(4.184)

$$D_0 = D(0).$$

При анализе уравнений (4.183) и (4.184), как и для дискретных систем, можно прийти к заключению, что уравнение оценки нелинейно относительно самой оценки, а уравнение дисперсии зависит от оценки и его необходимо решать совместно с уравнением оценки.

Кроме того, при сравнении (4.183) и (4.184) с уравнениями (4.116) и (4.117) обнаруживается, что вместо спектральной плотности белого шума \mathcal{N}_λ присутствует выражение $[S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}]$. Это указывает на наличие в задачах фильтрации фотонного излучения эквивалентного дискретного шума (квантового шума) с условной дисперсией $[S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}] / \Delta t$. В связи с этим обратим внимание на одно важное обстоятельство, отличающее синтез для точечных процессов от ранее рассмотренных алгоритмов обнаружения и фильтрации. В рассматриваемом случае ввиду зависимости параметров шума от оценок марковского процесса не удастся сформировать уравнение наблюдения, представляющего собой аддитивную смесь полезного сигнала и независимой от него помехи. Поэтому определение параметров эквивалентного шума, обусловленного квантовой природой носителя (фотонного излучения) информации, является одним из существенных моментов при определении алгоритмов совместного обнаружения и фильтрации. Более подробное обоснование введения эквивалентного белого шума и определение его параметров будет проведено в разд.4.11.

В задачах передачи информации фотонным излучением, как уже ранее отмечалось, наиболее существенным источником ошибки является квантовый шум. Вместе с тем в фотонных системах присутствуют шумы, которые в математической модели

наблюдаемого сигнала для многих практических ситуаций можно учесть в виде аддитивной независимой от оцениваемого процесса случайной составляющей с гауссовским распределением. К числу таких шумовых сигналов, которые назовем внешними шумами, можно отнести шумы электронных схем, темновой ток детекторов, квантованный шум цифровых фильтров и т.д. Рассмотрим один из возможных приближенных подходов, учитывающий этот внешний шум и позволяющий получить физически содержательные результаты в рамках корреляционной теории. Принимая во внимание, что параметры помехи в ранее рассмотренных рекуррентных уравнениях фильтрации учитывались в виде дисперсии, предлагается дополнить ее дисперсией эквивалентного белого шума. В результате дисперсия суммарного шума в дискретном времени принимает вид

$$D_{\Sigma,n+1} \approx [\mathcal{N}_{\lambda,n+1} + S_{n+1}(\tilde{x}_{n+1}) + \nu_{0,n+1}] / \Delta t,$$

где $\mathcal{N}_{\lambda,n+1} / \Delta t$ - дисперсия дискретного внешнего шума.

Необходимо также учесть, что модулирующий случайный процесс принимает с некоторой вероятностью отрицательное значение (например, гауссовский марковский процесс) и не может точно формировать модулируемую интенсивность, которая принципиально является положительной величиной. Для сведения к минимуму нежелательных искажений будем полагать, что дисперсия модулируемой интенсивности $M\{S_{n+1}^2(x_{n+1})\} - \{M\{S_{n+1}(x_{n+1})\}\}^2$ существенно меньше квадрата безусловной интенсивности $\{M\{S_{n+1}(x_{n+1})\}\}^2$. В этом случае вероятность появления отрицательного значения $S_{n+1}(x_{n+1})$ пренебрежимо мала (правило трех сигм).

Эффективность функционирования разработанных алгоритмов ввиду дискретного характера отсчетов наблюдаемого сигнала зависит от полосы пропускания частот (или обратной величины – интервала корреляции) модулирующего марковского процесса. Для минимизации ошибочных решений необходимо обеспечить $\tau_x \gg \Delta t$, где τ_x - интервал корреляции процесса, Δt - интервал отсчетов точечного процесса.

Рассмотрим одну частную задачу, в которой скалярный наблюдаемый сигнал в виде последовательности целочисленных отсчетов $\{N_n\}$ зависит от какой-либо компоненты многомерного

марковского процесса, например x_{1n} . Поскольку соответствующие алгоритмы получают таким же путем что и в скалярном варианте задачи разд.4.7.2, приведем окончательные выражения алгоритма фильтрации в векторной форме

$$\begin{aligned}
x_{n+1}^* &= \tilde{x}_{n+1} + \tilde{R}_{n+1} \frac{\partial S_{n+1}(\tilde{x}_{1,n+1})}{\partial \tilde{x}_{1,n+1}} \times \\
&\times \left[\Delta t \tilde{D}_{11,n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{1,n+1})}{\partial \tilde{x}_{1,n+1}} \right)^2 + S_{n+1}(\tilde{x}_{1,n+1}) + \nu_{0,n+1} \right]^{-1} \times \\
&\times [N_{n+1} - (S_{n+1}(\tilde{x}_{1,n+1}) + \nu_{0,n+1})\Delta t], \quad x_0^* = x(0), \\
D_{n+1} &= \tilde{D}_{n+1} - \tilde{T}_{n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{1,n+1})}{\partial \tilde{x}_{1,n+1}} \right)^2 \times \\
&\times \left[\tilde{D}_{11,n+1} \left(\frac{\partial S_{n+1}(\tilde{x}_{1,n+1})}{\partial \tilde{x}_{1,n+1}} \right)^2 + \frac{S_{n+1}(\tilde{x}_{1,n+1}) + \nu_{0,n+1}}{\Delta t} \right]^{-1}, \quad D_0 = D(0).
\end{aligned}$$

где $\tilde{R} = \|\tilde{D}_{11,n+1} \dots \tilde{D}_{r1,n+1}\|^T$, \tilde{T}_{n+1} - матрица размера $r \times r$, компонента которой $\tilde{D}_{i1,n+1} \tilde{D}_{j1,n+1}$, $i, j = \overline{1, r}$.

При формировании рекуррентных алгоритмов фильтрации квазидетерминированного процесса вида (4.89) учитывается, что оператор ФПК равен нулю, а уравнение АПВ имеет вид (4.91). Первая и вторая производные логарифма ФП для расширенного фильтра Калмана принимают векторно-матричные формы

$$\Pi_{n+1}^{(1)}(a_n^*) = \frac{[\partial S_{n+1}(a_n^*) / \partial a_n^*]^T}{S_{n+1}(a_n^*) + \nu_{0,n+1}} [N_{n+1} - (S_{n+1}(a_n^*) + \nu_{0,n+1})\Delta t] \quad (4.185)$$

$$\overline{\Pi}_{n+1}^{(2)}(a_n^*) = \frac{\Delta t}{S_{n+1}(a_n^*) + \nu_{0,n+1}} \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right), \quad (4.186)$$

где $\partial S_{n+1}(a_n^*) / \partial a_n^*$ - вектор-строка из компонент $\partial S_{n+1}(a_n^*) / \partial a_{jn}$, $j = \overline{1, k}$.

После подстановки выражений (4.185) и (4.186) в уравнения (4.61) и (4.62), используя тождество (4.85), а также соотношение (4.87), получим рекуррентный алгоритм фильтрации параметров квазидетерминированного процесса

$$\begin{aligned}
a_{n+1}^* &= a_n^* + D_n \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T \times \\
&\times \left[\Delta t \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T + S_{n+1}(a_n^*) + v_{0,n+1} \right]^{-1} \times \\
&\times [N_{n+1} - (S_{n+1}(a_n^*) + v_{0,n+1})\Delta t], \quad a_0^* = a(0), \\
D_{n+1} &= D_n - D_n \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right) D_n \times \\
&\times \left[\left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{n+1}(a_n^*)}{\partial a_n^*} \right)^T + \frac{S_{n+1}(a_n^*) + v_{0,n+1}}{\Delta t} \right]^{-1}, \\
D_0 &= D(0).
\end{aligned}$$

где выражения в квадратных скобках являются скалярными величинами.

4.10.3.2. МЕТОД СТАТИСТИЧЕСКОЙ ЛИНЕАРИЗАЦИИ

Нелинейные функции $f_n(x_n)$ и $S_{n+1}(x_{n+1})$ аппроксимируются выражениями (4.137) и (4.138) для скалярных процессов или (4.147) и (4.148) для многомерных процессов. Первая и вторая производные логарифма ФП (4.178) для определения соответствующих рекуррентных уравнений могут быть получены согласно методике разд.4.10.3.1. Ввиду единообразной процедуры получения этих рекуррентных уравнений приведем окончательные выражения уравнений оценки и дисперсий ошибки оценки для типовых задач при скалярном наблюдаемом сигнале в виде целочисленных отсчетов случайного потока на непересекающихся и примыкающих к друг другу подынтервалах (t_n, t_{n+1}) наблюдаемого интервала (t_1, t_{n+1}) .

Фильтрация скалярных процессов

$$x_{n+1}^* = \tilde{x}_{n+1} + \tilde{D}_{n+1} \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \left[\tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^2 \Delta t + S_{0,n+1} + \nu_{0,n+1} \right]^{-1} \times$$

$$\times [N_{n+1} - (S_{0,n+1} + \nu_{0,n+1})\Delta t], \quad x_o^* = x_0,$$

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{D}_{n+1}^2 \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^2 \left[\tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^2 + \frac{S_{0,n+1} + \nu_{0,n+1}}{\Delta t} \right]^{-1},$$

$$D_o = D(0),$$

где экстраполированные параметры \tilde{x}_{n+1} и \tilde{D}_{n+1} определяются из выражений (4.143) и (4.144).

Фильтрация компоненты многомерного марковского процесса

$$x_{n+1}^* = \tilde{x}_{n+1} + \tilde{R}_{n+1} \frac{\partial S_{0,n+1}}{\partial \tilde{x}_{1,n+1}} \left[\Delta t \tilde{D}_{11,n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{1,n+1}} \right)^2 + S_{0,n+1} + \nu_{0,n+1} \right]^{-1} \times$$

$$\times [N_{n+1} - (S_{0,n+1} + \nu_{0,n+1})\Delta t], \quad x_o^* = x(0),$$

$$D_{n+1} = \tilde{D}_{n+1} - \tilde{T}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{1,n+1}} \right)^2 \left[\tilde{D}_{11,n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{1,n+1}} \right)^2 + \frac{S_{0,n+1} + \nu_{0,n+1}}{\Delta t} \right]^{-1},$$

$$D_o = D(0),$$

где вектор \tilde{R}_{n+1} и матрица \tilde{T}_{n+1} приведены в разд.4.10.3.1.

Фильтрация параметров квазидетерминированного процесса.

$$a_{n+1}^* = a_n^* + D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T \left[\Delta t \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T + S_{0,n+1} + \nu_{0,n+1} \right]^{-1} \times$$

$$\times [N_{n+1} - (S_{0,n+1} + \nu_{0,n+1})\Delta t], \quad a_o^* = a_0,$$

$$D_{n+1} = D_n - D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right) D_n \times \\ \times \left[\left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right) D_n \left(\frac{\partial S_{0,n+1}}{\partial a_n^*} \right)^T + \frac{S_{0,n+1} + \nu_{0,n+1}}{\Delta t} \right]^{-1}, \quad D_0 = D(0).$$

4.11. НИЖНЯЯ ГРАНИЦА ДИСПЕРСИИ ОШИБКИ

4.11.1. ОТНОШЕНИЕ СИГНАЛ/ШУМ

Введем параметр отношение сигнал/шум (отношение мощности полезного сигнала к мощности помехи) для дискретного времени

$$q_{x,n+1} = -\bar{\Pi}_{n+1}^{(2)}(x_{n+1}) D_0, \quad (4.187)$$

где D_0 - мощность (априорная дисперсия) оцениваемого процесса, $\bar{\Pi}_{n+1}^{(2)}(x_{n+1})$ - введенная в разд.4.8.2 усредненная с плотностью вероятностей $p(y_{n+1}|x_{n+1})$ вторая производная логарифма ФП. Операцией усреднения добиваются независимости параметра q_x от флуктуации наблюдаемого сигнала y . Физический смысл формулы (4.187) можно раскрыть на основании следующих соображений. Мощность помехи зависит от ширины ФП (ФП полагаем унимодальной функцией). Чем шире ФП, тем больше мощность помехи. Функция $\bar{\Pi}_{n+1}^{(2)}(x_{n+1})$ пропорциональна средней крутизне ФП для всех значений y в точке, соответствующей состоянию x_{n+1} . Чем «шире» ФП, тем меньше крутизна и, следовательно, меньше параметр q_x .

Для линейной задачи ФП имеет вид (4.69) и этот параметр не зависит от x_{n+1}

$$q_{x,n+1} = C_{n+1}^2 Q_{\lambda,n+1}^{-1} D_0, \quad Q_{\lambda,n+1} = \mathcal{N}_{\lambda,n+1} / \Delta t. \quad (4.188)$$

Для нелинейных задач в случае гауссовского приближения при локальной гауссовской аппроксимации на основании соотношения (4.115) этот параметр равен

$$q_{x,n+1} = \left(\frac{\partial S_{n+1}(x_{n+1})}{\partial x_{n+1}} \right)^2 Q_{\lambda,n+1}^{-1} D_0. \quad (4.189)$$

Для точечных процессов и локальной гауссовской аппроксимации на основании соотношения (4.182) этот параметр оказывается равным

$$q_{x,n+1} = \Delta t \left(\frac{\partial S_{n+1}(x_{n+1})}{\partial x_{n+1}} \right)^2 [S_{n+1}(x_{n+1}) + \nu_{0,n+1}]^{-1} D_0.$$

По аналогии с предыдущим введем усредненную дисперсию ошибки $\bar{D}_{n+1} = \bar{h}_{2,n+1}^{-1}$, $\bar{h}_{2,n+1}^{-1} = [\bar{g}_{2,n+1} - \bar{\Pi}_{n+1}^{(2)}(x_{n+1})]^{-1}$ или $\bar{D}_{n+1} = \left[\bar{g}_{2,n+1} + \frac{q_{x,n+1}}{D_0} \right]^{-1}$. О качестве фильтрации будем судить по относительной дисперсии ошибки

$$\delta_{n+1}^2 = \bar{D}_{n+1} / D_0 = [\bar{g}_{2,n+1} D_0 + q_{x,n+1}]^{-1}. \quad (4.190)$$

Как следует из (4.190), относительная дисперсия ошибки, соответствующая моменту времени t_{n+1} , с увеличением q_x уменьшается и в асимптотике при $q_{x,n+1}/g_{2,n+1}D_0 \gg 1$ стремится к $q_{x,n+1}^{-1}$, т.е. не зависит от дисперсий D_0 и $g_{2,n+1}$.

Из полученных результатов также следует, что для нелинейных систем отношение сигнал/шум зависит от оцениваемого процесса x и в этом смысле является условной характеристикой. Для получения используемой далее безусловной характеристики необходимо этот параметр усреднить относительно априорной плотности вероятностей $p(x_{n+1})$

$$q_{0,n+1} = M_x \{ q_{x,n+1} \}. \quad (4.191)$$

Для линейной задачи имеет место тождество $q_{x,n+1} = q_{0,n+1}$.

Для многомерных процессов отыскивается соответствующая матрица, диагональные элементы которой являются отношением сигнал/шум для каждой компоненты оцениваемых процессов. Например, при локальной гауссовской аппроксимации указанная матрица имеет вид

$$\left(\frac{\partial S_{n+1}(x_{n+1})}{\partial x_{n+1}}\right)^T Q_{\lambda,n+1}^{-1} \left(\frac{\partial S_{n+1}(x_{n+1})}{\partial x_{n+1}}\right) D_0,$$

где $Q_{\lambda,n+1}$ и D_0 - диагональные матрицы соответственно размеров $(q \times q)$ и $(r \times r)$; $\partial S_{n+1}(x_{n+1})/\partial x_n$ - матрица Якоби размера $(q \times r)$, компонента которой $\partial S_{i,n+1}(x_{n+1})/\partial x_{j,n+1}$ (i - номер строки, j - номер столбца).

4.11.2. НЕРАВЕНСТВО КРАМЕРА-РАО

4.11.2.1. НИЖНЯЯ ГРАНИЦА ДИСПЕРСИИ ОШИБКИ ОЦЕНИВАЕМОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

В практических задачах получение точных алгоритмов оптимальной оценки наталкивается на серьезные технические трудности. Поэтому прибегают к приближенным методам решений уравнений оптимизации, что приводит к замене оптимальных алгоритмов квазиоптимальными. При обосновании применения квазиоптимальных алгоритмов важным является оценка потерь, вызванных неоптимальностью используемых алгоритмов. Для ее получения необходимо определить нижнюю границу дисперсии ошибки. И по степени близости к ней можно судить об эффективности используемых квазиоптимальных алгоритмов. Указанная граница может быть получена в общем виде для широкого класса задач и разных типов наблюдаемых сигналов.

При оценке параметра - случайной величины x по результатам наблюдения y указанная граница формируется на основании неравенства Крамера-Рао: средний квадрат (дисперсия) ошибки для любой несмещенной оценки параметра удовлетворяет этому неравенству

$$D_\varepsilon = M\{(\hat{x} - x)^2\} \geq - \left[M \left\{ \frac{\partial^2 \ln p(x, y_{t_0}^t)}{\partial x^2} \right\} \right]^{-1}, \quad \varepsilon = |\hat{x} - x| \quad (4.192)$$

или

$$D_{\varepsilon} \geq - \left[M \left\{ \frac{\partial^2 \ln p(y_{t_0}^t | x)}{\partial x^2} \right\} + M \left\{ \frac{\partial^2 \ln p(x)}{\partial x^2} \right\} \right]^{-1},$$

где $p(x, y_{t_0}^t)$ и $p(x)$ - априорная совместная плотность вероятностей параметра и наблюдаемого сигнала и априорная плотность вероятностей параметра. Математическое ожидание в правой части неравенства (4.192) берется относительно $p(x, y_{t_0}^t)$. При выводе этого неравенства используют очевидное соотношение для точечной несмещенной оценки $M\{\hat{x}\}=x$, где усреднение проводится с плотностью вероятностей $p(y_{t_0}^t | x)$. На основании соотношения $M\{(\hat{x}-x)\}=0$ имеем

$$\int_{-\infty}^{\infty} (\hat{x} - x) p(y_{t_0}^t | x) dy_{t_0}^t = 0. \quad (4.193)$$

Умножим обе части выражения (4.193) на $p(x)$, а затем продифференцируем по x , получаем

$$\int_{-\infty}^{\infty} p(x, y_{t_0}^t) dy_{t_0}^t = \int_{-\infty}^{\infty} (\hat{x} - x) \frac{\partial p(x, y_{t_0}^t)}{\partial x} dy_{t_0}^t. \quad (4.194)$$

После интегрирования обеих частей выражения (4.194) по x приходим к следующему соотношению

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\hat{x} - x) \frac{\partial p(x, y_{t_0}^t)}{\partial x} dx dy_{t_0}^t = 1 \text{ или с учетом}$$

$$\frac{\partial p(x, y_{t_0}^t)}{\partial x} = p(x, y_{t_0}^t) \frac{\partial \ln p(x, y_{t_0}^t)}{\partial x} \quad (4.195)$$

имеем

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ (\hat{x} - x) \sqrt{p(x, y_{t_0}^t)} \right\} \left\{ \sqrt{p(x, y_{t_0}^t)} \frac{\partial \ln p(x, y_{t_0}^t)}{\partial x} \right\} dx dy_{t_0}^t = 1$$

при условии дифференцируемости плотностей вероятностей по x и существования приведенных интегралов.

Используем теперь неравенство Коши-Буняковского, которое приводит к соотношению

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\hat{x} - x)^2 p(x, y_{t_0}^t) dx dy_{t_0}^t \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\frac{\partial \ln p(x, y_{t_0}^t)}{\partial x} \right]^2 p(x, y_{t_0}^t) dx dy_{t_0}^t \geq 1$$

или

$$D_{\varepsilon} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\hat{x} - x)^2 p(x, y_{t_0}^t) dx dy_{t_0}^t \geq \left[M \left\{ \left[\frac{\partial \ln p(x, y_{t_0}^t)}{\partial x} \right]^2 \right\} \right]^{-1}.$$

Далее продифференцируем $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y_{t_0}^t) dx dy_{t_0}^t = 1$ дважды по x и с учетом (4.195) получаем

$$M \left\{ \left[\frac{\partial \ln p(x, y_{t_0}^t)}{\partial x} \right]^2 \right\} = -M \left\{ \frac{\partial^2 \ln p(x, y_{t_0}^t)}{\partial x^2} \right\},$$

что и доказывает неравенство (4.192).

Оценка, для которой дисперсия ошибки, достигает нижней границы называется эффективной. Анализ показывает, что для существования эффективной оценки АПВ – $p(x|y_{t_0}^t)$ должна быть гауссовской. При анализе квазиоптимальных алгоритмов нижняя граница определяет потенциальную точность систем фильтрации.

Неравенство Крамера-Рао обобщается на задачу оценивания вектора параметров $x = \|x_1 \dots x_r\|^T$. В этом случае ошибки оказываются коррелированными и их совокупность описывается матрицей дисперсий ошибок D_{ε} , (i, j) компонента которой равна $M\{(\hat{x}_i - x_i)(\hat{x}_j - x_j)\}$, $i, j = 1, r$. Неравенство Крамера-Рао для нижней границы имеет вид

$$D_{\varepsilon} \geq J^{-1}, \quad (4.196)$$

где J^{-1} - матрица, обратная матрице Фишера, компоненты которой

$$J_{ij} = M \left\{ \frac{\partial \ln p(x, y_{t_0}^t)}{\partial x_i} \cdot \frac{\partial \ln p(x, y_{t_0}^t)}{\partial x_j} \right\} = -M \left\{ \frac{\partial^2 \ln p(x, y_{t_0}^t)}{\partial x_i \partial x_j} \right\}.$$

Все оценки являются эффективными, если $D_\varepsilon = J^{-1}$, что имеет место при гауссовском характере многомерной АПВ. Компоненты главной диагонали матрицы J^{-1} являются нижними границами дисперсий ошибок соответствующих параметров.

4.11.2.2. НИЖНЯЯ ГРАНИЦА ДИСПЕРСИИ ОШИБКИ ОЦЕНИВАЕМОГО СЛУЧАЙНОГО ПРОЦЕССА

При фильтрации случайного процесса последний необходимо представить в виде последовательности случайных величин - временных отсчетов в дискретные моменты t_j , $j=1, n+1$, объединенных в вектор $x_1^{n+1} = \{x_1, \dots, x_j, \dots, x_{n+1}\}$.

Нахождение нижней границы для дисперсии ошибки сводится к определению матрицы Фишера в соответствии с выражением (4.196). Однако, при этом возникают трудности вычислительного характера из-за необходимости обращения матрицы Фишера высокого порядка, который возрастает с увеличением номера момента времени t_n . Существенно упрощает вычислительный процесс переход к диагональной матрице, для которой только компоненты, расположенные на главной диагонали, не равны нулю: $J_{ij}=0$, $i \neq j$; $J_{ij} \neq 0$, $i=j$.

Сначала рассмотрим задачу фильтрации дискретизированного во времени одномерного марковского процесса (4.67) при наличии независимой от полезного сигнала аддитивной помехи типа белого шума. Представим совместную плотность вероятностей векторов случайных величин x_1^j и наблюдаемых сигналов y_1^j для момента времени t_j в виде $p(x_1^j, y_1^j) = p(y_1^j) p(x_1^j | y_1^j)$.

В разд.4.6.2 показано, что для марковских процессов последовательность плотностей вероятностей $\{p(x_1^j, y_1^j)\}$ эквивалентна последовательности, полученных с помощью рекуррентного алгоритма выражений АПВ

$$\{p(x_j|y_1^j)\} = \{w_j(x_j)\}, j=1, \dots, n+1,$$

где АПВ для момента времени t_j определяется из соотношения

$$w_j(x_j) = c_j p(y_j|x_j) \tilde{w}_{j-1}(x_j).$$

Эквивалентность здесь означает переход к выражению АПВ, позволяющему определить оценку \hat{x}_j по результатам наблюдения для того же момента времени y_j и некоторой статистики от всех предшествующих наблюдений y_1^{j-1} .

На основании указанного условия эквивалентности можно записать

$$\begin{aligned} p(x_1^j, y_1^j) &= p(y_1^j) p(x_1^j | y_1^j) \rightarrow p(y_1^j) p(x_j | y_1^j) = \\ &= c'_j p(y_j | x_j) \tilde{w}_{j-1}(x_j) \end{aligned} \quad (4.197)$$

где $c' = c p(y_1^j)$.

Прологарифмируем в выражении (4.197) стоящее справа от стрелки соотношение, а затем определим вторые производные функции $\ln p(x_j|y_1^j)$ по параметрам x_k и x_l , $k, l = \overline{1, j}$. Поскольку функции $\ln p(y_j|x_j)$ и $\ln \tilde{w}_{j-1}(x_j)$ зависят только от отсчета процесса x - в момент времени t_j , не равной нулю оказывается вторая производная функции $\ln p(x_j|y_1^j)$ по x_j .

В случае линейной задачи и фильтрации одномерного гауссовского марковского процесса ФП имеет вид (4.69). Используя методику разд.4.7.1 после вычислений имеем

$$\frac{\partial^2 \ln p(x_1^j, y_1^j)}{\partial x_k \partial x_l} \rightarrow \frac{\partial^2 \ln p(x_j | y_1^j)}{\partial x_j^2} = -(g_{2j} + C_j^2 Q_{\lambda j}^{-1}), \quad j = \overline{1, n+1}. \quad (4.198)$$

Для получения нижней границы дисперсии ошибки необходимо сначала усреднить выражение (4.198) с вероятностью $p(x_1^j, y_1^j)$, а затем изменить его знак на обратный. Ввиду того, что это выражение не зависит от процессов x и y компоненты главной диагонали Фишера принимают вид $J_{jj} = g_{2j} + C_j^2 Q_{\lambda j}^{-1}$, $J_{ij} = 0$ при $i \neq j$.

С учетом параметра отношения сигнал/шум нижняя граница дисперсии ошибки оценки равна

$$D_{\varepsilon j} = (g_{2j} + q_{0j}/D_0)^{-1}.$$

Покажем, что диагональной матрице Фишера соответствуют наиболее эффективные оценки. Ограничимся рассмотрением матрицы второго порядка. Для двух отсчетов x_1 и x_2 дисперсия совместно эффективных оценок имеет вид

$$D_{11} = \frac{J_{22}}{|J|} = \frac{1}{J_{11} \left(1 - \frac{J_{12}^2}{J_{11}J_{22}} \right)}, \quad D_{22} = \frac{J_{11}}{|J|} = \frac{1}{J_{22} \left(1 - \frac{J_{12}^2}{J_{11}J_{22}} \right)},$$

где $|J| = J_{11}J_{22} - J_{12}J_{21}$, $J_{12} = J_{21}$.

Умножим вторые слагаемые в круглых скобках на сомножитель $|J|^2/|J|^2$, получим

$$D_{11} = - \frac{1}{M \left\{ \frac{\partial^2 \ln p(x_1^2, y_1^2)}{\partial x_1^2} \right\}} \cdot \frac{1}{\left(1 - \frac{k_{12}^2}{D_{11}D_{22}} \right)}, \quad (4.199)$$

$$D_{22} = - \frac{1}{M \left\{ \frac{\partial^2 \ln p(x_1^2, y_1^2)}{\partial x_2^2} \right\}} \cdot \frac{1}{\left(1 - \frac{k_{12}^2}{D_{11}D_{22}} \right)}, \quad (4.200)$$

где $k_{12} = k_{21} = J_{12}/|J|$.

Первые сомножители в выражениях (4.199) и (4.200) являются компонентами диагональной матрицы Фишера. Из этих выражений следует, что наличие конечных корреляций между ошибками оценок приводит к увеличению дисперсии ошибок оценок. И только при $k_{12} = k_{21} = 0$ дисперсии ошибок оказываются наименьшими по величине, а оценки, следовательно, наиболее эффективными.

Таким образом, для линейной задачи нижняя граница дисперсии ошибки для момента времени t_{n+1} равна ранее вычисленной дисперсии ошибки оптимальной оценки (4.76)

$$(g_{2,n+1} + C_{n+1}^2 Q_{\lambda,n+1}^{-1})^{-1} = \tilde{D}_{n+1} - \tilde{D}_{n+1}^2 C_{n+1}^2 [\tilde{D}_{n+1} C_{n+1}^2 + Q_{\lambda,n+1}]^{-1},$$

где $g_{2,n+1} = \tilde{D}_{n+1}^{-1}$.

Для нелинейной задачи при локальной гауссовской аппроксимации (порядок аппроксимации $K=2$) с наблюдаемым сигналом (4.107) и ФП (4.108) вторая производная логарифма совместной плотности оцениваемого процесса и наблюдаемого сигнала

$$\frac{\partial^2 \ln p(x_1^j, y_1^j)}{\partial x_k \partial x_l} \rightarrow \frac{\partial^2 \ln p(x_j, y_1^j)}{\partial x_j^2} = -\{g_{2j} - \Pi_j^{(2)}(x_j)\} \quad j = \overline{1, n+1}, (4.201)$$

где вторая производная логарифма ФП

$$\Pi_j^{(2)}(x_j) = Q_{\lambda j}^{-1} \left\{ [y_j - S_j(x_j)] \frac{\partial^2 S_j(x_j)}{\partial x_j^2} - \left(\frac{\partial S_j(x_j)}{\partial x_j} \right)^2 \right\}.$$

Усредним выражение (4.201) относительно плотности вероятностей $p(x_1^j, y_1^j)$. После перемены знака получим

$$-M \left\{ \frac{\partial^2 \ln p(x_1^j, y_1^j)}{\partial x_k \partial x_l} \right\} \rightarrow -M \left\{ \frac{\partial^2 \ln p(x_j, y_1^j)}{\partial^2 x_j} \right\} = g_{2j} - M \{ \Pi_j^{(2)}(x_j) \}.$$

Обратим внимание, что функция $\Pi_j^{(2)}(x_j)$ зависит от отсчетов процесса x в момент времени t_j . Поэтому математическое ожидание от этой функции следует брать с плотностью вероятностей $p(x_j, y_j)$ и в два этапа: сначала с условной плотностью вероятностей $p(y_j|x_j)$ затем с априорной плотностью вероятностей $p(x_j) \rightarrow M_x\{\cdot\}$. Учитывая, что $M\{y_j|x_j\} = S_j(x_j)$, получим компоненты диагональной матрицы Фишера

$$\begin{aligned} J_{jj} &= g_{2j} - M \{ \Pi_j^{(2)}(x_j) \} = g_{2j} + Q_{\lambda j}^{-1} M_x \left\{ \left(\frac{\partial S_j(x_j)}{\partial x_j} \right)^2 \right\} = \\ &= g_{2j} - M_x \{ \bar{\Pi}_j^2(x_j) \}. \end{aligned}$$

Или, используя параметр безусловного отношения сигнал/шум (4.191), имеем

$$J_{jj} = g_{2j} + q_{0j}/D_0, \quad (4.202)$$

где

$$q_{0j} = Q_{\lambda j}^{-1} D_0 M_x \left\{ \left(\frac{\partial S_j(x_j)}{\partial x_j} \right)^2 \right\}. \quad (4.203)$$

В этом случае дисперсия ошибки оценки удовлетворяет неравенству

$$D_{\varepsilon j} \geq J_{jj}^{-1}, \quad j = \overline{1, n+1}. \quad (4.204)$$

Для случайных точечных процессов, используя аналогичный подход, также можно определить для j -го момента времени на основании безусловного отношения сигнал/шум компоненты диагональной матрицы Фишера (4.202) и получить соответствующее неравенство для оценки снизу дисперсии ошибки (4.204), где

$$q_{0j} = \Delta t D_0 M_x \left\{ \left(\frac{\partial S_j(x_j)}{\partial x_j} \right)^2 [S_j(x_j) + \nu_{0j}]^{-1} \right\}. \quad (4.205)$$

Определение нижней границы J_{jj}^{-1} с учетом (4.203) и (4.205) из-за необходимости усреднения при произвольном характере $p(x)$ является технически сложной в вычислительном отношении задачей. Если плотность вероятностей $p(x)$ аппроксимируется гауссовским распределением, а нелинейные функции имеют аналитический вид, то подобно рассмотренному в разд.4.8.3 методу в ряде задач удастся определить статистические характеристик нелинейных функций. В более сложных случаях эти характеристики можно получить, осуществив кусочно-линейную аппроксимацию нелинейных функций.

Получение этой границы при некоторых допущениях не связаны с большими трудностями. Так, при большом отношении сигнал/шум ошибки оценок оказываются малыми, а уравнения фильтрации при гауссовском приближении практически точными. Компоненты матрицы Фишера в этом случае принимают вид

$$J_{jj} \approx g_{2j} + Q_{\lambda j}^{-1} \left(\frac{\partial S_j(\tilde{x}_j)}{\partial \tilde{x}_j} \right)^2$$

или для точечных процессов

$$J_{jj} \approx g_{2j} + \Delta t \left(\frac{\partial S_j(\tilde{x}_j)}{\partial \tilde{x}_j} \right)^2 [S_j(\tilde{x}_j) + \nu_{0j}]^{-1}.$$

Нижними границами после замены $g_{2j} = \tilde{D}^{-1}$ являются решения уравнений (4.126) или (4.184).

Для многомерных процессов и линейной задачи условие эквивалентности для многомерной плотности вероятностей имеет вид

$$\frac{\partial^2 \ln p(x_1^j, y_1^j)}{\partial x_k \partial x_l} \rightarrow \frac{\partial^2 \ln p(x_j, y_1^j)}{\partial x_j^2} = -\{\tilde{D}_j^{-1} + C_j^T Q_{\lambda j}^{-1} C_j\} = -\{\tilde{D}_j^{-1} + q_{0j} D_0^{-1}\},$$

где диагональная матрица размера $(r \times r)$ отношений сигнал/шум $q_{0j} = C_j^T Q_{\lambda j}^{-1} C_j D_0$; x_k, x_l, x_j - векторы оцениваемых процессов для различных моментов времени размера $(r \times 1)$, C_j - матрица размера $(q \times r)$, $Q_{\lambda j}$, D_0 и \tilde{D}_j^{-1} - диагональные матрицы соответственно размеров $(q \times q)$, $(r \times r)$ и $(r \times r)$.

В этом случае матрица нижних границ дисперсий ошибок равна

$$D_{\varepsilon j} = [\tilde{D}_j^{-1} + q_{0j} D_0^{-1}]^{-1}, \quad j = \overline{1, n+1}.$$

Для нелинейной задачи при гауссовской аппроксимации матрица нижних границ дисперсий ошибок определяется из неравенств

$$D_{\varepsilon j} \geq \left[\tilde{D}_j^{-1} + M_x \left\{ \left(\frac{\partial S_j(x_j)}{\partial x_j} \right)^T Q_{\lambda j}^{-1} \left(\frac{\partial S_j(x_j)}{\partial x_j} \right) \right\} \right]^{-1}$$

для дискретных систем;

$$D_{ej} \geq \left[\tilde{D}_j^{-1} + \Delta t M_x \left\{ \left(\frac{\partial S_j(x_j)}{\partial x_j} \right)^T [S_j(x_j) + v_{oj}]^{-1} \left(\frac{\partial S_j(x_j)}{\partial x_j} \right) \right\} \right]$$

для точечных процессов,

где $\partial S_j(x_j)/\partial x_j$ - матрица Якоби размера $(q \times r)$, $Q_{\lambda j}$ и $[S_j(x_j) + v_{oj}]^{-1}$ - диагональные матрицы размера $(q \times q)$.

4.12. УЧЕТ УПРАВЛЕНИЯ В ЗАДАЧАХ ОЦЕНИВАНИЯ СОСТОЯНИЯ ДИНАМИЧЕСКИХ СИСТЕМ

В задачах оптимального управления используются по сравнению с ранее рассмотренными (4.120) более полные учитывающие управление модели состояния системы. Управление входит в уравнение состояний системы в виде детерминированной аддитивной составляющей – последовательности сигналов управления $\{u_n\}$, зависящих от оценок состояния системы в дискретном времени

$$x_{n+1} = x_n + \Delta t f_n(x_n) + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0), \quad (4.206)$$

где V_n детерминированная матрица коэффициентов обратной связи размера $(r \times q)$, u_n - вектор управления размера $(q \times 1)$.

Наряду с (4.206) в задачах управления используется сокращенная запись уравнений

$$x_{n+1} = f_n(x_n, u_n, \eta_n) \quad (4.207)$$

или

$$x_{n+1} = f_n(x_n) + V_n u_n + \Gamma_n \eta_n. \quad (4.208)$$

С помощью сигналов $\{u_n\}$ осуществляется целенаправленное управление динамическими системами.

С учетом управления получим рекуррентные уравнения оценки и дисперсии ошибки оценки в линейной задаче для скалярных и многомерных процессах, а также для задач гауссовской локальной и статистической линеаризации.

Для скалярных процессов уравнение одномерного гауссовского марковского процесса имеет вид

$$x_{n+1} = \Phi_n x_n + v_n u_n + \gamma_n \eta_n, \quad x_0 = x(0), \quad (4.209)$$

где $\Phi_n = 1 + \Delta t f_n$, v_n - детерминированный зависящий от времени параметр обратной связи.

С учетом ФП (4.69) можно получить рекуррентные уравнения фильтрации: оценки (4.71) и дисперсии ошибки оценки (4.72). Остановимся подробнее на определении экстраполированных параметров. Экстраполированная оценка на основании формулы (4.53) равна

$$\tilde{x}_{n+1} = \int_{-\infty}^{\infty} x_{n+1} \tilde{w}_n(x_{n+1}) dx_{n+1} = \Phi_n \hat{x}_n + v_n u_n,$$

так как $M\{v_n u_n | y_1^n\} = v_n u_n$.

Для экстраполированной дисперсии уравнение (4.54) сохраняет свой вид ввиду детерминированного характера управления. Как следует из полученных выражений, введение сигналов управления отражается только на форме экстраполированной оценки, остальные входящие в рекуррентные алгоритмы выражения и уравнения сохраняют свой вид. Используя аналогичный подход, можно определить рекуррентные алгоритмы для других задач оценивания состояния системы при наличии управляющих сигналов.

Линейная задача фильтрации многомерных процессов.

Уравнение многомерного гауссовского марковского процесса имеет вид

$$x_{n+1} = \Phi_n x_n + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0), \quad (4.210)$$

где Φ_n и Γ_n ранее введенные при анализе уравнения (4.80) матрицы.

Экстраполированная оценка равна

$$\tilde{x}_{n+1} = \Phi_n \hat{x}_n + V_n u_n.$$

Уравнения (4.88), (4.86) и (4.66) сохраняют свой вид.

Задача фильтрации при локальной гауссовской линеаризации.

Уравнение многомерного марковского процесса имеет вид

$$x_{n+1} = x_n + \Delta t f_n(x_n) + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0), \quad (4.211)$$

где $f_n(x_n)$ - вектор размера $(r \times 1)$.
Экстраполированная оценка равна

$$\tilde{x}_{n+1} = x_n^* + \Delta t f_n(x_n^*) + V_n u_n.$$

Уравнения (4.125), (4.126) и (4.128) сохраняют свой вид.

Задача фильтрации при статистической линеаризации нелинейностей.

Уравнение многомерного марковского процесса имеет вид (4.211). Экстраполированная оценка равна

$$\tilde{x}_{n+1} = x_n^* + \Delta t f_{0n}(x_n^*, D_n) + V_n u_n.$$

Уравнения (4.149), (4.150) и (4.152) сохраняют свой вид.

Для точечных процессов также можно получить аналогичные результаты: наличие управляющих сигналов отражается только на форме записи экстраполированной оценки.

4.13. СТАТИСТИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ДИНАМИЧЕСКИХ СИСТЕМ

В задачах синтеза оптимального управления необходимо располагать математической (информационной) моделью динамической системы - объекта управления. Такой моделью, как следует из раздела 3.1, служит система уравнений (или оператор системы), описывающая поведение динамической системы. Математическая модель может быть получена на основании теоретического анализа известных законов механики и физики, а также в результате экспериментальных исследований входных и выходных сигналов (фазовых координат) системы, установления соотношений между ними. В последнем случае имеют дело с задачей идентификации динамических систем. В большинстве случаев эта задача должна формулироваться как статистическая, так как учитывает в процессе нормальной эксплуатации случайный характер входных сигналов, воздействующих шумов и помех. При статистическом подходе оператор или уравнения динамической системы определяются на основании вероятностных методов обработки наблюдаемой информации и получения статистических характеристик выходных сигналов системы.

При решении практических задач прибегают к пробным (зондирующим) воздействиям с целью получения апостериорных характеристик: оценок фазовых координат и их корреляционных функций. При наиболее общем подходе в задачах идентификации определяется структура детерминированных операторов, аппроксимирующих наилучшим образом с точки зрения принятого критерия качества оператор идентифицируемой динамической системы. В данном разделе рассматривается одна из частных задач идентификации, в которой получают оценки ряда параметров динамической системы известной структуры.^{*)}

Полагаем, что поведение динамической системы описывается многомерным марковским процессом в пространстве состояний: системой линейных разностных уравнений первого порядка в матричной форме

$$x_{n+1} = \Phi_n x_n + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0). \quad (4.212)$$

Здесь $\Phi_n = I + \Delta t F_n$, I - единичная матрица; η_n - вектор зондирующих сигналов, обычно аппроксимируемых дискретным белым шумом с известной диагональной матрицей дисперсии и нулевым математическим ожиданием. Другие типы зондирующих сигналов на основании методики раздела 2.3.7 путем расширения вектора состояния системы всегда можно свести к дискретному белому шуму. Вектор детерминированных зондирующих сигналов u_n является аналогом введенных в разделе 4.12 сигналов управления.

Матрица F_n размера $r \times r$ с неполностью известными компонентами (параметрами) имеет вид $F_n = \|f_{ij,n}\|$.

Требуется определить оценки неизвестных параметров всех или части фазовых координат системы при воздействии случайных зондирующих сигналов по вектору наблюдения размера q ($q \leq r$)

$$y_n = C_n x_n + \lambda_n,$$

где λ_n - случайная помеха при измерениях, представляющая собой дискретный белый шум с известной диагональной матрицей дисперсии и нулевым математическим ожиданием.

^{*)} С различными методами идентификации можно ознакомиться в работах [52, 53, 54, 30]

Номинальное значение матриц $F_n^0 = \|f_{ij,n}^0\|$ и $\Phi_n^0 = I + \Delta t F_n^0$ полагают известными, а отклонение от неизвестных параметров малыми по величине. При известных статистических характеристиках зондирующих сигналов и номинальных значениях параметров матрицы Φ_n^0 из уравнения (4.212) можно получить матричное уравнение для определения вектора m_x^0

$$m_{x,n+1}^0 = \Phi_n^0 m_{xn}^0 + V_n u_n. \quad (4.213)$$

Используя уравнения (4.212) и (4.213), после линеаризации относительно номинальных значений имеем

$$\begin{aligned} x_{n+1} &= \Phi_n^0 m_{xn}^0 + (\Phi_n - \Phi_n^0) m_{xn}^0 + \Phi_n^0 (x_n - m_{xn}^0) + V_n u_n + \Gamma_n \eta_n \text{ или} \\ x_{n+1} &= \Phi_n^0 x_n + \Phi_n m_{xn}^0 - \Phi_n^0 m_{xn}^0 + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0). \end{aligned} \quad (4.214)$$

Расширим вектор состояния системы за счет неизвестных параметров матрицы Φ_n . При этом полагаем, что за время наблюдения отклонения параметров от номинальных значений являются постоянными случайными величинами. Введем векторы $\Phi_{in} = \|\Phi_{1i,n} \dots \Phi_{ri,n}\|^T$, $i = \overline{1, r}$, образующие матрицу $\Phi_n \|\Phi_{1n} \dots \Phi_{in} \dots \Phi_{rn}\|$.

Если теперь пронумеровать все компоненты вектора x_n и матрицы Φ_n и образовать из них вектор z , то придем к матричному уравнению расширенного вектора состояния системы

$$z_{n+1} = \Psi_n z_n - \Phi_n^0 m_{xn}^0 + V_n u_n + \Gamma_n \eta_n, \quad z_0 = z(0),$$

где прямоугольная матрица

$$\Psi_n = \begin{vmatrix} \Phi_{11,n}^0 & \dots & \Phi_{1r,n}^0 & m_{x_1}^0 & \dots & m_{x_r}^0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Phi_{r1,n}^0 & \dots & \Phi_{rr,n}^0 & m_{x_1}^0 & \dots & m_{x_r}^0 \end{vmatrix}.$$

Далее для определения оценок компонент вектора z используются рассмотренные в разделе 4.7.2 методы теории линейной фильтрации. Уравнение оценки имеет вид

$$\hat{z}_{n+1} = \tilde{z}_{n+1} - D_{n+1} C_{n+1}^T Q_{n+1}^{-1} [y_{n+1} - C_{n+1} \tilde{z}_{n+1}], \quad z_0 = z(0),$$

где вектор $\tilde{z}_{n+1} = \Psi_n \hat{z} - \Phi_n^0 m_{xn}^0 + V_n u_n$.

Блочная матрица D_{n+1} состоит из матриц дисперсий ошибок $D_{x,n+1}$, $D_{\Phi,n+1}$ и взаимных дисперсий ошибок $D_{x\Phi,n+1}$ *)

$$D_{n+1} = \left\| \begin{array}{c|c|c|c} D_x & D_{x\Phi_1} & \cdots & D_{x\Phi_r} \\ \hline D_{x\Phi_1} & D_{\Phi_1} & \cdots & \cdots \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline D_{x\Phi_r} & \cdots & \cdots & D_{\Phi_r} \end{array} \right\|.$$

Компоненты блочной матрицы определяются из матричного уравнения дисперсии ошибок оценок

$$D_{n+1} = \tilde{D}_{n+1} + \tilde{D}_{n+1} C_{n+1}^T (C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + Q_{\lambda,n+1})^{-1} C_{n+1} \tilde{D}_{n+1}, \quad D_0 = D(0),$$

где $\tilde{D}_{n+1} = \Psi_n D_n \Psi_n^T + \Gamma_n Q_{\eta n} \Gamma_n^T$.

Погрешность получаемых оценок параметров определяется матрицами D_x и D_{Φ_i} . Предельная минимальная погрешность для устойчивой стационарной системы получается из решений уравнений дисперсии при $n \rightarrow \infty$ ($t_n \rightarrow \infty$).

Перейдем к идентификации динамической системы, описываемой нелинейным векторным уравнением в дискретном времени

$$x_{n+1} = x_n + \Delta t f_n(x_n, \Phi_n) + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0). \quad (4.215)$$

Рассмотрим общие подходы определения оценок не полностью известных параметров вектора Φ при использовании метода статистической линеаризации. Предварительно путем усреднения из выражения (4.215) получают уравнение относительно номинальных значений вектора математических ожиданий фазовых координат системы m_x^0 и компонент вектора Φ^0

$$m_{x,n+1}^0 = m_{xn}^0 + \Delta t f_{0n}(m_{xn}^0, D_n^0, \Phi_n) + V_n u_n, \quad (4.216)$$

*) В обозначениях блочной матрицы для сокращений записей опущен подстрочный индекс $n+1$

где D_n^0 - дисперсия, соответствующая номинальному значению m_{xn}^0 , вычисленная из совместного решения уравнений (4.215) и (4.216). Полагая отклонения неизвестных параметров от номинальных значений малыми, линеаризируем векторную нелинейность

$$f(x, \Phi) = f_0(m_x^0, D^0, \Phi^0) + \frac{\partial f_0}{\partial \Phi^0} (\Phi - \Phi^0) + \frac{\partial f_0}{\partial m_x^0} (x - m_x^0). \quad (4.217)$$

Используя выражение (4.217) и подставляя уравнение (4.216) в (4.215), получаем линеаризованное уравнение

$$x_{n+1} = \left(I + \Delta t \frac{\partial f_{0n}}{\partial m_{xn}^0} \right) x_n + \Delta t \frac{\partial f_{0n}}{\partial \Phi_n^0} \Phi_n + \Delta t \left(f_{0n} - \frac{\partial f_{0n}}{\partial \Phi_n^0} \Phi_n^0 - \frac{\partial f_{0n}}{\partial m_{xn}^0} m_{xn}^0 \right) + \\ + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0),$$

которое по структуре аналогично уравнению (4.214). Поэтому дальнейшее решение задачи идентификации проводится по алгоритму линейной задачи. Процедуры дополнительных операций, связанных с определением статистических характеристик и коэффициентов усиления, ранее рассматривались в разд.3.4.2 и 4.8.3.

4.14. ПРОГНОЗИРОВАНИЕ И УПРАВЛЕНИЕ В КОМПЬЮТЕРНЫХ СЕТЯХ

4.14.1. ВВЕДЕНИЕ

Одним из важных требований, предъявляемых к современным высокоскоростным средствам компьютерных телекоммуникаций и характеризующих эффективность их функционирования, является быстрая и надежная (в смысле наименьшей ошибки) передача информации. Количественная оценка эффективности работы сетей отражается в параметре - пропускной способности или верхней границе средней скорости передачи данных в цифровой форме. Задача оптимизации сетей заключается в уменьшении потерь пропускной способности и тем самым в обеспечении более высокого качества использования сетевых ресурсов. Проблема состоит в том, что переход к технологии пакетной коммутации и создание

интегрированных сетевых приложений сопровождается, как было отмечено в разд.2.5, появлением сложных явлений, исследование которых может быть проведено в рамках теоретико-вероятностных подходов. Анализ потоков информации в виде последовательности серий или пачек пакетов (сетевого трафика) показывает, что эти потоки сопровождаются существенными случайными флуктуациями - пачечностью трафика и наличием больших межпачечных интервалов (прерывистым потоком пачек пакетов)^{*)}. Поэтому при организации внутрисетевого взаимодействия, кроме среднего значения скорости передачи данных, необходимо учитывать ее пиковое значение. В результате в сети следует иметь значительные запасы по пропускной способности, что приводит к нерациональному использованию сетевых ресурсов. Объяснения этому следует искать в особенностях построения сетей. Сетевая конфигурация включает в себя узлы, в которых расположены сетевые устройства (буфера, маршрутизаторы, коммутаторы), обеспечивающие требуемые маршруты прохождения пакетов. Из-за ограниченного объема памяти этих устройств возникают очереди, часть пакетов может быть даже потеряно, что приводит к необходимости их повторной передаче. Все это вызывает дополнительные временные задержки. Из-за нерегулярного влияния при передаче информации этих факторов поведение сетевого трафика принимает случайный характер.

Содержанием задач прогнозирования (экстраполяции) и управления в компьютерных сетях является формирование алгоритма минимизации потерь в пропускной способности или, что тоже самое, предотвращения перегрузки сетевых устройств. Необходимо иметь ввиду, что координация взаимодействия всех информационных ресурсов сети обеспечивается комплексом программных средств или блоком протоколов. Наряду с многочисленными функциями по сбору информации, анализу и контролю за состоянием сетевых объектов, формированию соответствующих команд управления сетевым взаимодействием, на протоколы возложены функции управления трафиком для экономной и эффективной передачи пакетов данных,

^{*)} В современных компьютерных сетях техника передачи информации предусматривает, что поток байтов разбивается на отдельные пакеты (пакетизируется), в которых помимо передаваемой присутствует служебная информация с указанием адреса назначения пакета. Далее информация передается на пакетном уровне по дуплексному (двунаправленному) каналу связи

управления их очередями и предотвращения тупиковых ситуаций, например, переполнения буферов в промежуточных узлах сети.

Модели фрактальных процессов, аппроксимирующих поведение сетевого трафика, рассматривались в разделе 2.5. В рамках корреляционной теории были получены статистические характеристики этих процессов (математическое ожидание, корреляционные функции и коэффициенты корреляции, спектральные плотности). С другой стороны, общие подходы по прогнозированию процессов, опирающиеся на соответствующие алгоритмические приемы, обсуждались в разделе 4.9. Все это позволяет перейти к решению конкретных задач по прогнозированию и управлению сетевым трафиком [55, 56, 20].

4.14.2. АЛГОРИТМ ДЛЯ МОДЕЛЕЙ СЕТЕВОГО ТРАФИКА ТИПА ПРИРАЩЕНИЙ ТОЧЕЧНОГО ПРОЦЕССА

При формализации задачи упомянутые выше случайные флуктуации скорости передачи данных для стационарного процесса приращений (отсчетов) X_n случайного точечного процесса могут быть охарактеризованы математическим ожиданием $m_x = \nu T$ (2.101) и коэффициентом корреляции $r(k; T)$ (2.120), где ν - средняя скорость (интенсивность) точечного процесса; T - интервал заданной длительности, k - число интервалов T , на которые разнесены отсчеты^{*)}. Управление информационными потоками между узлами виртуального соединения осуществляется на основе прогноза приращений точечного процесса с помощью модифицированного варианта протокола *UDP*. Выберем два узла, из которых i -ый узел является источником, а j -ый узел - приемником. Допустим, что интенсивность потока (пропускная способность участка сети между i и j узлами) определяется очередью в узле j , возникшей, например, из-за ограниченного объема памяти буфера в этом узле, низкой интенсивностью разгрузки этого буфера, из-за прибывающих в этот узел пакетов с других соединений сети и т.д. В связи с этим интенсивность потока информации от узла i к узлу j понижается, а в случае переполнения буфера в узле j передача информации прекращается, что сопровождается потерей части пакетов. Для

^{*)} Для указанных моделей точечным процессом аппроксимируется последовательность пачек пакетов. В качестве случайных величин рассматриваются интервалы между точками (пачками пакетов) [57] .

предотвращения полной потери пропускной способности необходимо регулировать уровень загрузки этого буфера. Для этого воспользуемся оптимальными в средне-квадратическом смысле прогнозируемыми оценками приращений точечного процесса. Полагаем, что на рассматриваемом участке сети практически точно измеряется число отсчетов за время $(t_n - T, t_n)$.

Одновременно осуществляется прогноз на некотором интервале упреждения kT на основании формулы (4.165) при

$$\hat{X}_{n+k} = r(k; T)(X_n - \nu T) + \nu T, \quad k = 1, 2, \dots,$$

где $r(k; T)$ в зависимости от исходных данных и особенностей решения задач принимает одну из форм (2.120), (2.121) и (2.122); \hat{X}_{n+k} - оценка числа отсчетов за время $(t_{n+k} - T, t_{n+k})$, слагаемое νT учитывает известную постоянную составляющую приращений процесса.

Если величина прогноза оказывается больше порога зависящего от уровня загрузки буфера и определяемого некоторым адаптивным алгоритмом, то по сигналу обратной связи интенсивность генерации с узла i уменьшается на величину, зависящую от уровня загрузки буфера и значений оценок прогноза. Хотя скорость передачи данных из-за этого на участке сети понижается, но в связи с сохранением процесса передача информации и уменьшением числа потерянных пакетов удастся в среднем уменьшить потери в пропускной способности этого соединения. Если необходимо сохранить пропускную способность, то по сигналу обратной доступные сетевые ресурсы перераспределяются в пользу этого участка виртуального соединения сети.

Качество прогноза оценивается по величине дисперсии ошибки при заданном параметре смещения k (4.166)

$$D_\varepsilon = D(T)[1 - r^2(k; T)],$$

где $D(T)$ - априорная дисперсия числа отсчетов (2.111).

Как следует из этого выражения, с возрастанием параметра k , что соответствует увеличению глубины прогноза, качество прогноза ухудшается (увеличивается ошибка). В пределе при $k \rightarrow \infty$ дисперсия ошибки прогноза стремится к априорной дисперсии $D(T)$.

4.14.3. АЛГОРИТМ ДЛЯ МОДЕЛЕЙ СЕТЕВОГО ТРАФИКА ТИПА ФРАКТАЛЬНОГО БРОУНОВСКОГО ДВИЖЕНИЯ

В современных компьютерных сетях соединение между пользователями осуществляется ТСП/IP протоколом. В соответствии с алгоритмом функционирования этого протокола пропускная способность со стороны источника в фазе медленного старта определяется текущим окном перегрузки, равным числу разрешенных к передаче пакетов до прихода пакетов подтверждения. На Рис. 4.4. представлен процесс формирования и эволюции окна перегрузки размера W_i для выбранного виртуального соединения источник-приемник по одному из вариантов реализации ТСП/IP протокола, где i - номер шага фазы медленного старта.

На первом шаге после отправки пакета с номером 1 разрешается передать еще W_1-1 пакетов до получения пакета подтверждения на посланный пакет с номером 1. На втором и всех последующих шагах, если пакеты подтверждения поступают более или менее регулярно, величина окна каждый раз увеличивается в два раза. В результате достигается максимально возможная (но не более заявленного со стороны приемника окна перегрузка) для данного соединения и принятого протокола пропускная способность соединения. Алгоритм, формирующий окно перегрузки по сигналу обратной связи в виде пакетов подтверждения, называется алгоритмом скользящего окна, так как при очередном получении пакета подтверждения окно перемещается (и увеличивается на один пакет на каждый шаг алгоритма), захватывая очередные пакеты, которые разрешается передавать без подтверждения. Допустим, что из-за очередей в промежуточных узлах - маршрутизаторах, а также из-за переполнения буферов в этих узлах соединение сети не справляется с нагрузкой. Вследствие этого часть пакетов чрезмерно задерживается в пути и даже, может быть, теряется (на Рис. 4.4 задерживается один пакет на $i+1$ шаге). В этом случае пакеты подтверждения не отсылаются и протоколом ТСП на стороне источника на шаге $i+2$ формируется окно уменьшенного размера, равное своему значению W_i на i -м шаге, и достоверность передачи обеспечивается через механизм повторения посылки пакетов. На этом заканчивается фаза медленного старта и начинается фаза управления перегрузкой, сопровождающая для налаживания и поддержания соединения

локальными воздействиями и определением каждый раз текущего окна перегрузки.

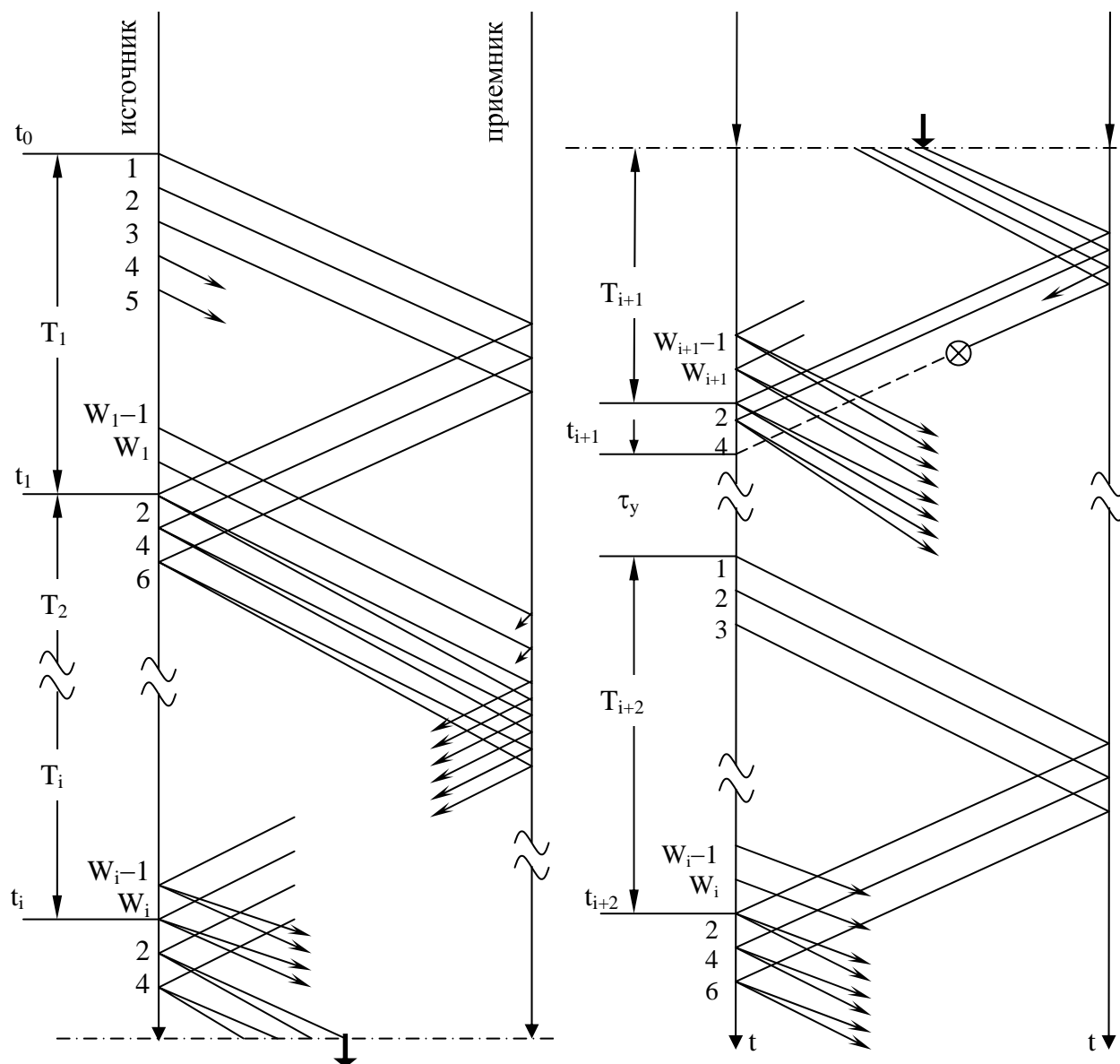


Рис. 4.4. Процесс формирования и эволюции окна перегрузки.

Интервал времени между моментами посылки пакета в направлении приемника получения пакета подтверждения определяется RTT-задержкой (round trip time)^{*)}. Чтобы избежать длительных простоев из-за ожидания потерянных или задержавшихся пакетов вводится пороговое значение RTT-задержки (тайм-аут - τ_y). При превышении RTT-задержкой порога пакеты считаются потерянными. Очевидно, в связи с уменьшением окна перегрузки

^{*)} На рисунке представлена упрощенная модель эволюции режима медленного старта, не учитывающая случайные флуктуации RTT-задержек

пропускная способность соединения снижается. Анализ работы существующего алгоритма работы протокола показывает, что имеются значительные резервы уменьшения потерь пропускной способности в режимах медленного старта и управления перегрузкой. Во-первых, перегрузка не прогнозируется, а обнаруживается по самому факту отсутствия пакетов подтверждения после очередного перемещения окна, во-вторых, по этому факту нельзя судить о величине перегрузки и следующее состояние окна устанавливается методом «проб и ошибок». Задача заключается в формировании заблаговременно, не заходя в критическую область обнаружения потерянных пакетов, оценок как месторасположения временного интервала возможного проявления перегрузки, так и величины самой перегрузки. Решение такой задачи можно получить в рамках методов идентификации процессов, использования при моделировании РТТ-задержек фрактального броуновского движения и формирования оценок прогноза. Временной интервал РТТ-задержки складывается из времени обработки и распространения, а также затрат времени на очереди вдоль маршрута прохождения пакетов. Для стационарного процесса РТТ-задержку, которую обозначим через T_i , можно записать в виде (Рис. 4.5.)

$$T_i = T_{1i} + T_{2i} + T_{np},$$

где $i=1,2,\dots$ номер задержки (шага), интервалы T_{1i} и T_{2i} соответствуют временам пересылки пакета от источника к приемнику и обратно, T_{np} - время обработки информации в приемнике. Совокупность интервалов $\{T_i\}$ образует дискретную последовательность РТТ-задержек. Интервалы T_{1i} и T_{2i} отражают задержки в обработке и передаче информации, а также наличие очередей в промежуточных узлах сети. Для известного маршрута движения пакета можно выделить постоянную минимальную обусловленную отсутствием очередей составляющую T_0 и случайную составляющую из-за задержек в очередях и связанную со случайным поведением сетевого трафика ΔT_i : $T_i = T_0 + \Delta T_i$.

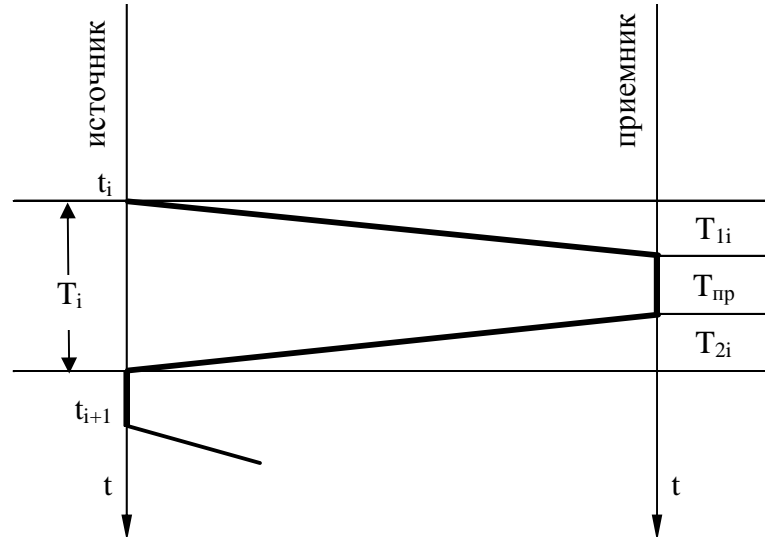


Рис. 4.5. Процесс РТТ-задержки.

Рассматривая выражение $T_i - M\{T_i\}$, где $M\{T_i\} = T_0 + \Delta T_{cp}$, ΔT_{cp} - среднее значение приращений РТТ-задержек, в качестве случайного приращения РТТ-задержки, сформируем для момента времени t_n фрактальный броуновский процесс [58].

$$B_H(t_n) = \sum_i^n [T_i - (T_0 + \Delta T_{cp})].$$

Полагаем, что моменты времени t_i ($i = \overline{1, n}$) образуют регулярную последовательность дискретных временных отсчетов. Для определенности интервалы между отсчетами принимаем равными средней величине РТТ-задержки $\Delta = T_0 + \Delta T_{cp}$. В соотношении (2.132) отождествим текущие времена t_2 и t_1 с моментами времени t_{n+k} и t_n ($t_{n+k} - t_n = k\Delta$), где k - параметр смещения. Используя это выражение, запишем соотношение для корреляционной функции процесса

$$k_{2H}(t_n, t_{n+k}) = M\{B_H(t_n)B_H(t_{n+k})\} \sim \frac{1}{2}[t_n^{2H} + t_{n+k}^{2H} - |t_{n+k} - t_n|^{2H}].$$

Из-за нестационарного характера фрактального броуновского движения коэффициент корреляции записывается в форме, зависящей от текущего времени t_n

$$r_n(k; \Delta) = \frac{k_{2n}(t_n, t_{n+k})}{D(t_n)} = \frac{t_n^{2H} + t_{n+k}^{2H} - |t_{n+k} - t_n|^{2H}}{2t_n^{2H}} =$$

$$= \frac{1}{2}[1 + S_{n,n+k}^{2H} - |S_{n,n+k} - 1|^{2H}], \quad (4.218)$$

где $S_{n,n+k} = t_{n+k}/t_n$.

Оптимальная в среднеквадратическом смысле оценка прогноза фрактального броуновского движения в момент времени t_{n+k} при известном (измеренном точно) указанном процессе в момент t_n на основании (4.165) равна

$$\hat{B}_n(t_{n+k}) = M\{B_H(t_{n+k}) | B_H(t_n)\} = r_n(k; \Delta) B_H(t_n)^* \quad (4.219)$$

Оптимальная оценка прогноза РТТ-задержки для момента времени t_{n+k} принимает вид

$$\hat{T}_{n+k} = \hat{B}_H(t_{n+k}) - \hat{B}_H(t_{n+k-1}) + T_0 + \Delta T_{cp} =$$

$$= [r_n(k; \Delta) - r_n(k-1; \Delta)] B_H(t_n) + T_0 + \Delta T_{cp}. \quad (4.220)$$

Приведем также используемую далее оценку прогноза РТТ-задержки на основании следующего за последним измеренным для момента времени t_n значения фрактального броуновского движения. На основании формулы (4.220) имеем

$$\hat{T}_{n+1} = \hat{B}_H(t_{n+1}) - B_H(t_n) + T_0 + \Delta T_{cp}, \quad (4.221)$$

где $\hat{B}_H(t_{n+1}) = r_n(1; \Delta) B_H(t_n)$.

Для случая на рис.4.4 наличие возможной перегрузки на $i+1$ шаге могло быть обнаружено и тем самым с большой вероятностью предотвращено по результатам измерения РТТ-задержек или фрактального броуновского движения к моменту времени t_i , формированию согласно формуле (4.221) соответствующей оценки прогноза и сравнения ее с тайм-аутом. Кроме того, по величине спрогнозированной РТТ-задержки на $i+1$ шаге можно судить об уровне перегрузки и по сигналу обратной связи соответственно уменьшить (перенастроить) величину окна.

*) Выражение (4.219) впервые было получено в работе [21].

Для рассматриваемого случая дисперсия ошибки прогноза задержки определяется выражением

$$D_{\varepsilon} = M\{(\hat{T}_{n+1} - T_{n+1})^2\}, \quad (4.222)$$

где $\hat{T}_{n+1} = [r_n(1; \Delta) - 1]B_H(t_n) + T_0 + \Delta T_{cp}$, $T_{n+1} = B_H(t_{n+1}) - B_H(t_n) + T_0 + \Delta T_{cp}$.

После подстановки выражений \hat{T}_{n+1} и T_{n+1} в соотношение (4.222), последующих вычислений, получаем

$$D_{\varepsilon} = M\{[r_n(1; \Delta)B_H(t_n) - B_H(t_{n+1})]^2\} = D_{n+1} - r_n^2(1; \Delta)D_n,$$

$$\text{где } D_n \sim t_n^{2H}, \quad D_{n+1} \sim t_{n+1}^{2H}, \quad r_n(1; \Delta) = \frac{1}{2}[1 + S_{n,n+1} - |S_{n,n+1} - 1|^{2H}],$$

$$S_{n,n+1} = t_{n+1} / t_n.$$

5. СТАТИСТИЧЕСКИЙ СИНТЕЗ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДИНАМИЧЕСКИХ СИСТЕМ

5.1. ВЕДЕНИЕ В ПРОБЛЕМУ СИНТЕЗА

В задачах управления информационная система выполняет функции управляющей подсистемы, играющей ключевую роль в оптимизации системы управления. Эта подсистема формирует результатную информацию или в контексте рассматриваемой проблемы осуществляет синтез оптимального управления динамической системы - объекта управления.

В теории оптимального управления в зависимости от цели и средств оптимизации получили наиболее широкое распространение подходы, связанные с определением оптимального линейного оператора [41, 53, 59, 60], осуществляющие параметрический синтез [61, 62, 63, 64], а также аналитические методы решения задач [65, 66, 67, 68, 69]. В первом случае при известном оптимальном операторе системы управления и операторе объекта в рамках корреляционной теории определяется оператор управляющей подсистемы. Использование этого метода не всегда приводит к физически реализуемым результатам. Кроме того, он довольно часто дает решения, критичные к небольшим вариациям параметров системы

[70, 71]. При параметрическом синтезе для заданной структуры управляющей подсистемы предлагаются вычислительные алгоритмы определения параметров оптимального закона управления. К числу этих методов относятся процедуры, основанные на стохастической аппроксимации, случайного поиска, на градиентных и неградиентных способах оптимизации. Аналитические методы, которые, в основном, рассматриваются в пособии, позволяют найти структуру и закон оптимального управления для функционалов качества, имеющих аналитический вид. К этому направлению относятся задачи синтеза управления детерминированных динамических систем при воздействующих неслучайных сигналах, основанные на классическом вариационном исчислении. Однако, оптимизационные задачи, возникающие в теории управления, не всегда могут быть сведены к известным вариационным методам. Причиной этому могут служить ограничения на управление и фазовые координаты систем, использование функционалов качества не аналитического вида, случайный характер протекающих в динамической системе процессов, что приводит к функционалам качества стохастического вида. Для их решения были разработаны специальные методы синтеза, базирующиеся на принципе максимума А.С.Понтрягина и динамическом программировании Р.Беллмана, которые в дальнейшем были обобщены на стохастические задачи. Аналитические решения указанных задач достигаются современными методами теории статистических решений, опирающимися на байесовский критерий условного риска и вариационные методы оптимизации. Для получения конструктивных технически реализуемых решений используется математический аппарат теории условных марковских процессов. Указанное направление успешно развивается на основе описания объектов управления в пространстве состояний и разработки алгоритмов оптимального управления с использованием цифровой вычислительной техники. Все это послужило мощным стимулом к дальнейшему развитию современной теории управления динамических систем и формированию теории статистического синтеза этих систем как самостоятельной дисциплины [67, 72, 73].

Задачи синтеза, использующие вероятностные модели процессов, являются более сложными по сравнению с детерминированными задачами, так как используются вероятностные критерии, что потребовало применять дополнительные операции

осреднения результатов измерения относительно плотностей вероятностей. Сложность решения указанной задачи заключается в том, что апостериорная плотность зависит от всех отсчетов вектора измерений, размерность которого увеличивается с возрастанием времени измерения. Существенное упрощение можно получить, если использовать вектор достаточных статистик вполне определенной размерности. Он является вектор-функцией последнего отсчета вектора измерений и некоторой статистики от всех предшествующих измерений. В качестве достаточных статистик рассматриваются условные марковские процессы - оценки фазовых координат динамической системы по результатам измерений. С введением достаточных координат задача стохастического синтеза оптимального управления упрощается и может быть условно разделена на две: задачу вычисления достаточных статистик на основании доступных для измерения координат объекта и задачу определения оптимального управления, как функции оптимизационных критериев и достаточных координат. В связи с этим управляющая подсистема распадается на два блока: блок обработки информации и блок оптимального управления. В общем случае обе задачи решаются совместно. Однако для линейных систем, гауссовских сигналов, аддитивной помехи и квадратичного функционала качества справедлива теорема разделения [74]. Согласно этой теореме задачу статистического синтеза можно свести к двум последовательно решаемым в аналитическом виде задачам: определению достаточных статистик (задаче фильтрации) и оптимального управления. Причем задача фильтрации решается независимо от задачи оптимизации управления. При указанном подходе можно определить требуемый закон управления и тем самым сформировать структуру системы управления, работающую по принципу обратной связи. Теорема разделения может быть использована для приближенного решения задачи управления нелинейными системами с применением ранее рассмотренных методов локальной и статистической линеаризации нелинейностей.

Задача оптимального управления тесно связана с одним из важных свойств динамических систем - стохастической управляемостью. Для динамических систем с детерминированными процессами свойство управляемости сводилось к возможности перевода динамической системы из заданного начального состояния

в требуемое конечное за ограниченный интервал времени с использованием кусочно-непрерывного управления. Для динамических систем, находящихся под воздействием случайных процессов, это свойство, которое взаимосвязано с рассмотренным в разделе 4.6.1 свойством наблюдаемости, определяется асимптотическим поведением апостериорной плотности вероятностей подверженных случайным вариациям фазовых координат системы при неограниченном времени наблюдения. Наиболее наглядно это может быть продемонстрировано на характере эволюции во времени компонент матрицы апостериорных дисперсий ошибок управления системы

$$D_{\varepsilon} = \lim_{t \rightarrow \infty} M \{ (x_t - x_t^{np})(x_t - x_t^{np})^T \mid y_{t_0}^t \}$$

при $\lim_{t \rightarrow \infty} M \{ (x_t - x_t^{np}) \} = 0$.

где x_t^{np} - требуемый программный вектор фазовых координат системы.

Если сумма диагональных компонент матрицы дисперсии ошибок (след матрицы дисперсии) по модулю ограничена при $t \rightarrow \infty$, то система стохастически управляема.

В наиболее общем случае расчет оптимального управления состоит из следующих этапов:

- формирования функционала качества;
- выбора критерия оптимизации;
- установление оптимального закона управления;
- определение структуры оптимально-управляющей подсистемы.

5.2. КРИТЕРИИ ОПТИМИЗАЦИИ

Понятие оптимальности определяется принятым критерием оптимизации и система, оптимальная по одному критерию, не является оптимальной по другому критерию. За характеристику оптимальности обычно принимают некоторый показатель (функционал качества).

Функционал качества является скалярной функцией, в общем случае зависящей от координат объекта $x(t)$, требуемых значений этих координат $x^{np}(t)$, управления $u(t)$ и времени t

$$J(x_t, x_t^{np}, u_t, t).$$

Задача оптимизации состоит в выборе критерия оптимизации и отыскании на основании этого критерия экстремального относительно вектора управления $u(t)$ значения функционала качества: минимума общей стоимости, потерь, штрафа или максимума другого функционала - суммарного результата, выигрыша.

При формировании задачи в общем виде полагаем, что поведение объекта управления описывается в пространстве состояний r -мерным векторно-матричным уравнением

$$\dot{x}_t = f_t(x_t, \eta_t, u_t), \quad x(t_0) = x(0), \quad (5.1)$$

в котором в силу формирования закона управления введен на основании данных раздела 4.12 детерминированный вектор управления $u(t)$. Управление объектом определяется видом функционала качества и организуется на основании измерений всех или части фазовых координат объекта общим числом q , $q \leq r$. Это равносильно предположению, что управление зависит от информации о координатах системы, физически доступных измерению. Обычно на вектор управления накладываются ограничения, связанные с физическими свойствами объекта или конечной мощностью органов управления. Эти ограничения в общем случае можно представить в символической форме

$$u \in U, \quad (5.2)$$

где под U понимается допустимое множество векторов управления.

Ограничения могут иметь аналитический вид, например, при управлении в среднем применяют условие

$$M\{u^T(t)W_t u(t)\} \leq \gamma_1(t),$$

где W_t - диагональная матрица с положительными компонентами.

Если ограничены общие ресурсы, то применяются интегральные (изопериметрические) условия в форме

$$\int_{t_0}^t M\{u^T(\tau)W_\tau u(\tau)\}d\tau \leq \gamma_2(t).$$

В инженерной практике наиболее широкое применение получил функционал качества в квадратичном виде и соответственно байесовский подход при определении оценки этого функционала - минимум условного риска.

В общем случае функционал качества в квадратичной форме записывается в виде

$$\begin{aligned} J(x_t, x_t^{np}, u_t, t) = & [x(t) - x^{np}(t)]^T K_t [x(t) - x^{np}(t)] + \\ & + \int_{t_0}^t \{ [x(\tau) - x^{np}(\tau)]^T L_\tau [x(\tau) - x^{np}(\tau)] + u^T(\tau) W_\tau u(\tau) \} d\tau, \end{aligned} \quad (5.3)$$

где K_t и L_t - заданные соответственно положительно и неотрицательно определенные матрицы размера $(q \times q)$. В зависимости от исходных данных и характера решений частных задач могут отсутствовать некоторые составляющие функционала (5.3). Кроме того, при решении ряда задач для упрощений расчетов принимают $x^{np}(t)=0$.

Критерии оптимизации выбираются, исходя из смысла решаемой задачи, например, минимума времени управления при переводе объекта из начального в требуемое конечное состояние или, если учитывается переходный процесс - экстремального значения интеграла от функционала состояния объекта. В других задачах ограничиваются заданием времени управления $t-t_0$, где t может обозначать текущее время или при $t=t_k$ - конечный момент времени управления. В первом случае критерий оптимизации называется локальным, так как требуется обеспечить экстремальное значение функционала качества в каждый текущий момент времени t ; во втором - терминальным, так как управление выбирают из условия экстремального значения функционала качества в конечный момент

времени t_k , который может быть фиксированным или подвижным. Так как в последнем случае часть условий задается в начальный момент времени, а часть - в конечный момент времени приходится решать двухточечную краевую задачу. При этом необходимо иметь ввиду, что из-за стохастического характера задачи эти начальные и конечные состояния объекта задаются в виде вероятностных характеристик: векторов математических ожиданий и матриц дисперсий векторов оценок или их ошибок. Поэтому двухточечные задачи решаются с учетом ограничений на конечное состояние объекта в виде $\hat{x}^T(t_k) \hat{x}(t_k) \leq \rho_1$ или $M\{[\hat{x}(t_k) - x(t_k)][\hat{x}(t_k) - x(t_k)]^T\} \leq \rho_2$

При формировании оптимального управления текущее значение функционала качества принимается за дополнительную координату, $x_{r+1}(t) = J(x_t, x_t^{np}, u_t, t)$, а его производная записывается в форме дифференциального уравнения

$$\dot{x}_{r+1}(t) = \frac{\partial J}{\partial t} + \sum_{i=1}^r \frac{\partial J}{\partial x_i} \frac{\partial x_i}{\partial t} + \sum_{i=1}^k \frac{\partial J}{\partial u_i} \frac{\partial u_i}{\partial t} + \sum_{i=1}^r \frac{\partial J}{\partial x_i^{np}} \frac{\partial x_i^{np}}{\partial t}, \quad x_{r+1}(t_0) = J_{t_0},$$

$$\text{где } \frac{\partial x_i}{\partial t} = f_{it}(x_t, \eta_t, u_t).$$

Функционал, характеризующий качество управления в задаче определения управления, переводящего объект из $x(t_0)$ в состояние $x(t_k)$ за минимальное время $t_k - t_0$ (задача быстрогодействия), имеет вид

$$J = \int_{t_0}^t d\tau.$$

Для дополнительной координаты состояния $x_{r+1}(t)$ дифференциальное уравнение запишется следующим образом

$$\dot{x}_{r+1}(t) = 1, \quad x_{r+1}(t_0) = t_0 \quad (5.4)$$

и оптимизируемый функционал принимает форму

$$J = x_{r+1}(t_k) - t_0. \quad (5.5)$$

Если отыскивается оптимальное управление, соответствующее минимуму интеграла от текущего состояния, то функционал качества принимает вид

$$J = x_{r+1}(t_k) = \int_{t_0}^{t_k} \varphi_\tau(x_\tau, x_\tau^{np}, u_\tau) d\tau,$$

а дифференциальное уравнение по дополнительной координате

$$\dot{x}_{r+1}(t) = \varphi_t(x_t, x_t^{np}, u_t), \quad x_{r+1}(t_0) = 0.$$

Общее выражение функционала качества для задачи управления конечным состоянием в момент t_k можно представить в форме

$$J = F_{t_k}(x_{t_k}, x_{t_k}^{np}) + \int_{t_0}^{t_k} \varphi_\tau(x_\tau, x_\tau^{np}, u_\tau) d\tau \quad (5.6)$$

а дифференциальное уравнение по дополнительной координате в виде

$$\begin{aligned} \dot{x}_{r+1}(t) = & \frac{\partial F_t}{\partial t} + \sum_{i=1}^r \frac{\partial F_t}{\partial x_i} f_{it}(x_t, \eta_t, u_t) + \varphi_t(x_t, x_t^{np}, u_t) + \\ & + \sum_{i=1}^r \frac{\partial F_t}{\partial x_i^{np}} f_{it}^{np}(x_t^{np}), \quad x_{r+1}(t_0) = J_{t_0}, \end{aligned}$$

где F_{t_k} характеризует конечное (терминальное) состояние объекта.

Задача определения оптимального управления конечным состоянием в этом случае также обращается в задачу оптимизации по дополнительной координате $x_{r+1}(t)$ в момент времени $t=t_k$ расширенного вектора состояния системы.

В ряде задач оптимизации управления заданы ограничения на вектор конечных координат объекта. В общем виде эти ограничения подчиняются уравнениям связи

$$Q_j[x_1(t_k), \dots, x_r(t_k)] = 0, \quad j = \overline{1, s}, \quad (5.7)$$

где Q_j - произвольные дифференцируемые функции.

В этом случае двухточечная задача оптимального управления перевода объекта из начального состояния в конечное состоит в отыскании условного экстремума функционала качества J при выполнении условий на координаты объекта (5.7) и на управление

(5.2). Задача нахождения условного экстремума относительно управления может быть сведена на основании известного метода Лагранжа к отысканию безусловного экстремума выражения

$$\Pi(t_k) = x_{r+1}(t_k) + \sum_{j=1}^s \lambda_j Q_j[x_1(t_k), \dots, x_r(t_k)], \quad (5.8)$$

где λ_j - неопределенные множители Лагранжа.

Если уравнения связи отсутствуют, то минимизируемый функционал (5.8) становится равным дополнительной координате $x_{r+1}(t_k)$.^{*)}

При локальном критерии вариационная задача вырождается, минимизируемый функционал (5.8) оказывается равным дополнительной координате $x_{r+1}(t)$ зависящей от текущего времени. Таким образом, функционал качества превращается в скалярную функцию времени и его оптимальность обеспечивается, как будет показано далее, только выбором локально-оптимального вектора управления.

Так как функционал качества зависит от векторов случайных координат объекта, он оказывается случайной скалярной функцией. При выборе критерия оптимизации ее необходимо преобразовать в неслучайную функцию, например, в вероятностную оценку. Ввиду того, что источником информации при формировании оптимального управления являются результаты измерений этих координат целесообразно выбрать в качестве оценки апостериорное среднее (условное математическое ожидание) функционала качества относительно реализации наблюдаемого процесса $y(t)$

$$\hat{J} = M\{J(x_\tau, x_\tau^{np}, u_\tau, \tau | y_{t_0}^\tau)\}, \quad t_0 \leq \tau \leq t_k. \quad (5.9)$$

Наилучшей оценке при байесовском подходе и квадратичной функции потерь соответствует минимальное математическое

^{*)} Параметры объекта под влиянием внешней среды, внутренних условий могут меняться. В связи с этим изменяется зависящий от этих параметров функционал качества. Для того, чтобы он снова достиг оптимального значения, с помощью дополнительных устройств и обслуживающих их алгоритмов адаптации изменяют параметры формируемых в информационных системах алгоритмов управления. С адаптивными системами и устройствами можно ознакомиться в работах [52] и [75].

ожидание квадрата ошибки - безусловной дисперсии ошибки оценки функционала качества

$$M_y\{M\{(J-\hat{J})^2|y_{t_0}^\tau\}\}, \quad (5.10)$$

где $M_y\{\cdot\}$ означает повторное математическое ожидание относительно всех возможных реализаций процесса $y(t)$.

Рассмотрим апостериорное среднее квадрата ошибки с учетом дополнительных двух последних членов в сумме равных нулю

$$\begin{aligned} & M\{(J-\hat{J})^2|y_{t_0}^\tau\} + [M\{J|y_{t_0}^\tau\}]^2 - [M\{J|y_{t_0}^\tau\}]^2 = \\ & = [M\{J|y_{t_0}^\tau\}]^2 - 2\hat{J}M\{J|y_{t_0}^\tau\} + \hat{J}^2 + M\{J^2|y_{t_0}^\tau\} - [M\{J|y_{t_0}^\tau\}]^2. \end{aligned}$$

После преобразования получаем

$$M\{(J-\hat{J})^2|y_{t_0}^\tau\} = [M\{J|y_{t_0}^\tau\} - \hat{J}]^2 + M\{J^2|y_{t_0}^\tau\} - [M\{J|y_{t_0}^\tau\}]^2. \quad (5.11)$$

Подставляя (5.11) в соотношение (5.10), имеем

$$\begin{aligned} M_y\{M\{(J-\hat{J})^2|y_{t_0}^\tau\}\} &= M_y\{[M\{J|y_{t_0}^\tau\} - \hat{J}]^2\} + \\ &+ M_y\{M\{J^2|y_{t_0}^\tau\}\} - M_y\{[M\{J|y_{t_0}^\tau\}]^2\}. \end{aligned} \quad (5.12)$$

По определению безусловная дисперсия ошибки оценки функционала качества является неотрицательной величиной и, как следует из (5.12), минимальное ее значение достигается, если первое слагаемое равно нулю. Таким образом минимум выражения (5.10) получают при оценке \hat{J} , равной апостериорному среднему функционала качества: $M\{J|y_{t_0}^\tau\} = \hat{J}$.

В общем виде задача статистического синтеза оптимального управления объектом, состояние которого описывается векторно-матричным уравнением (5.1), а измерение - вектором $y(t)$, зависящим от этого состояния, состоит в определении вектора управления $u(t)$, удовлетворяющего ограничениям (5.2) и доставляющего экстремум апостериорному среднему оптимизируемого функционала (5.8)

$$\text{exstr}_{u \in U} M\{\Pi(t_k)|y_{t_0}^\tau\}, \quad t \leq \tau \leq t_k,$$

при переходе объекта из начального состояния в некоторую определенную область пространства состояний. Заметим, что всегда можно отыскать минимум апостериорного среднего функционала $\hat{\Pi} = M\{\Pi|y\}$, умножив на -1 максимум этого функционала.

Рассматриваемые далее методы оптимизации: принцип максимума (минимума) и динамического программирования целесообразно применять с учетом особенностей решения конкретных задач. Принцип максимума рекомендуется применять при наличии ограничений в виде неравенств на фазовые координаты объекта и управление, причем объект описывается линейными дифференциальными уравнениями. Динамическое программирование целесообразно применять для объектов, описываемых конечно-разностными уравнениями, т.е. для дискретных динамических систем с учетом возможностей цифровой вычислительной техники.

5.3. МЕТОДЫ ОПТИМИЗАЦИИ

5.3.1. ПРИНЦИП МАКСИМУМА ДЛЯ НЕПРЕРЫВНОГО ВРЕМЕНИ

5.3.1.1. ОБЩИЕ ОПРЕДЕЛЕНИЯ

В соответствии с принципом максимума [67, 76] вводятся вспомогательные переменные $\psi(t)$, которые определяются из линейной системы дифференциальных уравнений

$$\dot{\psi}_i = - \sum_{j=1}^{r+1} \psi_j \frac{\partial f_j}{\partial x_i}, \quad i = \overline{1, r+1}, \quad (5.13)$$

где функции f_i ($j = \overline{1, r}$) являются правыми частями уравнений (5.1), а $f_{r+1} = \dot{x}_{r+1}(t)$ - производными дополнительной координаты в общем виде (5.3) или (5.6).

Сформируем скалярную функцию Гамильтона (стохастический гамильтониан)

$$H_\psi = H_\psi(x, \psi, \eta, u, t) = \psi^T f_t(x, \eta, u) = f_t^T(x, \eta, u) \psi = \sum_{i=1}^{r+1} \psi_i f_{it}. \quad (5.14)$$

Из выражения (5.14) следуют соотношения $\frac{\partial H_\psi}{\partial \psi_i} = f_{it}$ и

$$\frac{\partial H_\psi}{\partial x_i} = \sum_{j=1}^{r+1} \psi_j \frac{\partial f_{jt}}{\partial x_i}, \quad i = \overline{1, r+1}.$$

С учетом формул (5.1) и (5.13) получаем дифференциальные уравнения переменных в канонической форме Гамильтона

$$\dot{x}_i = \frac{\partial H_\psi}{\partial \psi_i}, \quad (5.15)$$

$$\dot{\psi}_i = -\frac{\partial H_\psi}{\partial x_i}, \quad i = \overline{1, r+1} \quad (5.16)$$

с граничными условиями при t_0 и t_k

$$x_i(t_0) = x_i(0), \quad \psi_i(t_k) = -\frac{\partial \Pi(t_k)}{\partial x_i(t_k)}. \quad (5.17)$$

Принимая во внимание выражение (5.7), соотношение (5.17) принимает форму

$$\psi_i(t_k) = -\sum_{j=1}^s \lambda_j \frac{\partial Q_j}{\partial x_i(t_k)}, \quad i = \overline{1, r}, \quad (5.18)$$

$$\psi_{r+1}(t_k) = -1. \quad (5.19)$$

Соотношение (5.18) выражает условие трансверсальности. Это условие позволяет, располагая в момент времени t_k частью свободных координат состояния объекта общим числом $r-q$ при $r>q$, получить систему $r-q$ соотношений для определения вектора вспомогательных переменных $\psi(t_k)$. Добавляя к ней заданные q ограничений на вектор координат из (5.7), приходим к полной системе r соотношений для момента времени t_k .

Особенностью применения принципа максимума является его тесная связь с принципом Гамильтона в теоретической механике. Используя этот принцип, приведенные соотношения могут быть получены достаточно простым способом и, что не менее важно, поддаются наглядной интерпретации (приложение 3).

Согласно принципу максимума в двухточечной задаче при детерминированных сигналах и оптимальных значениях векторов управления, удовлетворяющих ограничению (5.2), и состояния x при измерении вектора y , для которых функционал $\Pi(t_k)$ принимает минимальное значение, условие оптимальности состоит в достижении гамильтониана на интервалах $t_0 \leq \tau \leq t_k$ значений $\max_{u \in U} H_\psi(x, \psi, \eta, u, \tau)$ или $\sup_{u \in U} H_\psi(x, \psi, \eta, u, \tau)$, т.е. гамильтониан в первом случае достигает максимального, во втором - наибольшего значения. В работе [76] показано, что максимальное значение гамильтониана равно нулю.

Принцип максимума доказан как необходимый признак оптимальности для нелинейных объектов. В этом случае оптимальные траектории координат объекта являются экстремалиями. Проверкой необходимо убедиться, какая из них соответствует оптимальному управлению. Из необходимого условия $\max_{u \in U} H_\psi = 0$ следует, что вдоль оптимальной траектории включая начальные и конечные точки $(r+1)$ -мерные вектора ψ и \dot{x} взаимно ортогональны.

Отметим, что условия максимума гамильтониана является более общими по сравнению с уравнением Эйлера в классическом вариационном исчислении. В случае когда управление принадлежит открытой области (отсутствуют ограничения), из принципа максимума следует условие оптимальности классического вариационного исчисления.

В работах [77, 78] принцип максимума был распространен на стохастические задачи. В этом случае стохастический принцип максимума формулируется следующим образом. В двухточечной задаче для оптимальных значений векторов управления u , удовлетворяющих ограничению (5.2), и состояния объекта x (5.1) при измерении вектора y апостериорное среднее функционала $\Pi(t_k)$ принимает минимальное значение, а апостериорное среднее стохастического гамильтониана имеет максимальное или наибольшее значение относительно управления на интервале времени $t_0 \leq \tau \leq t_k$

$$\max_{u \in U} \hat{H}_\psi(\hat{x}_\tau, \hat{\psi}_\tau, u_\tau, \eta_\tau, \tau) = \max_{u \in U} M\{H(x_\tau, \psi_\tau, u_\tau, \eta_\tau, \tau) | y_{t_0}^\tau\},$$

$$\sup_{u \in U} \hat{H}_\psi(\hat{x}_\tau, \hat{\psi}_\tau, u_\tau, \eta_\tau, \tau) = \sup_{u \in U} M\{H(x_\tau, \psi_\tau, u_\tau, \eta_\tau, \tau) | y_{t_0}^\tau\},$$

где апостериорные оценки переменных

$$\hat{x} = M\{x | y_t^\tau\} \text{ и } \hat{\psi} = M\{\psi | y_t^\tau\}$$

определяются из уравнений

$$\hat{x}_i = M\left\{\frac{\partial H_\psi}{\partial \psi_i} \middle| y_{t_0}^\tau\right\}, \quad i = \overline{1, r+1}, \quad (5.20)$$

$$\hat{\psi}_i = -M\left\{\frac{\partial H_\psi}{\partial x_i} \middle| y_{t_0}^\tau\right\}, \quad i = \overline{1, r+1}, \quad (5.21)$$

при следующих граничных условиях

$$\begin{aligned} \hat{x}_i(t_0) &= x_i(0), \\ \hat{\psi}_i(t_k) &= -M\left\{\frac{\partial \Pi(t_k)}{\partial x_i(t_k)} \middle| y_{t_0}^\tau\right\}, \quad i = \overline{1, r+1}. \end{aligned} \quad (5.22)$$

Для решения задачи оптимального управления и получения конкретных результатов в качестве оценок состояния системы по результатам измерения необходимо воспользоваться результатами линейной и нелинейной фильтрации (разд. 4.7 и 4.8).

Если управление принадлежит открытой области, то вектор управления определяется единственным образом из векторно-матричного уравнения

$$M\left\{\frac{\partial H_\psi}{\partial u} \middle| y_{t_0}^\tau\right\} = 0, \quad t_0 \leq \tau \leq t_k \quad (5.23)$$

при условии, что апостериорная оценка гамильтониана имеет экстремум.

Следует отметить, что решение в общем виде задачи статистического синтеза оптимального управления достаточно сложно. Это обстоятельство приводит к необходимости решения сложных задач численными методами путем подбора соответствующих начальных и конечных условий. Однако для линейных объектов и измерителей, аддитивного управления и

квадратичного функционала качества необходимые и достаточные условия принципа максимума обоснованы [74]. На основании теоремы разделения задача может быть существенно упрощена и разделена на два этапа решения: первый - получение апостериорных оценок \hat{x} и $\hat{\psi}$, второй - определение оптимального управления u .

5.3.1.2. УПРАВЛЕНИЕ ПРИ ОПТИМАЛЬНОМ БЫСТРОДЕЙСТВИИ

Определение оптимального быстродействия при детерминированных воздействиях и при ограничении на управление явилось одной из первых задач оптимального управления при формировании программной траектории движения объекта и расчете программы управления. Она является актуальной и при разработке управляющей подсистемы, обеспечивающей быстродействие для объекта, находящегося под воздействием случайных сигналов, при вероятностном характере его начального и конечного состояний.

При решении задачи стохастического синтеза по оптимальному быстродействию поведения объекта с учетом аддитивного управления описывается линейным дифференциальным уравнением в векторно-матричной форме

$$\dot{x}(t) = F_t x(t) + V_t u(t) + \Gamma_t \eta(t), \quad x(t_0) = x(0) \quad (5.24)$$

с учетом ограничений $x(t_k) = x_k$.

Эти ограничения так же как и начальные условия, как правило, задаются векторами математических ожиданий и матрицами дисперсий: $M\{x(0)\}$, $M\{x_k\}$ и D_0 , D_k . В процессе управления осуществляется измерение вектора $y(t)$ в соответствии с уравнением наблюдения

$$y(t) = C_t x(t) + \lambda(t). \quad (5.25)$$

Это уравнение отражает тот факт, что в реальной задаче при неточных измерениях из-за помехи λ невозможно реализовать управление по полной информации о координатах состояния объекта.

Необходимо определить вектор оптимального управления u , переводящий объект из начального состояния x_0 в конечное x_k за минимальное время $t_k - t_0$ при учете того, что максимальное значение компонент вектора управления ограничено U_i . Для дополнительной координаты дифференциальное уравнение имеет вид (5.4)

$$\dot{x}_{r+1}=1, x_{r+1}(t_0)=t_0 \quad (5.26)$$

и минимизируемый функционал (5.8) принимает форму

$$\Pi(t_k) = x_{r+1}(t_k) + \sum_{i=1}^r \lambda_i [x_i(t_k) - x_{ik}],$$

где второе слагаемое учитывает ограничения в виде уравнений связи. Стохастический гамильтониан с учетом уравнений (5.24) и (5.26) принимает вид

$$\begin{aligned} \hat{H}_\psi &= M \{ [\psi^T (F_t x_t + V_t u_t + \Gamma_t \eta_t) + \psi_{r+1}] | y_{t_0}^t \} = \\ &= M \{ \psi^T F_t x_t | y_{t_0}^t \} + \sum_{i=1}^r \hat{\psi}_i V_{it} u_{it} + M \{ \psi^T \Gamma_t \eta_t | y_{t_0}^t \} + \hat{\psi}_{r+1}, \end{aligned} \quad (5.27)$$

где ψ^T - вектор-строка размера $(1 \times r)$.

Согласно стохастическому принципу максимума оптимальное управление доставляет максимум апостериорному среднему гамильтониана. Как следует из выражения (5.27) этот максимум по отношению к управлению достигается при $\sum_{i=1}^r \hat{\psi}_i u_{it} > 0, V_{it} > 0, |u_i| = U_i$.

Таким образом, оптимальное управление, обеспечивающее минимальное время (максимальное быстроедействие) движения объекта из начального состояния в конечный, является релейным и определяется знаком апостериорного среднего вспомогательной функции $\hat{\psi}$

$$u_i = U_i \text{sign} \hat{\psi}_i, \quad i = \overline{1, r}.$$

Функции $\hat{\psi}_i$ и x_i вычисляются из совместно решаемых уравнений (5.20) и (5.21) с учетом формулы для конечного состояния (5.22)

$$\begin{aligned} \hat{\psi}(t_k) &= -\lambda, \\ \hat{\psi} &= -F^T \hat{\psi}, \quad \hat{\psi}_{r+1} = 0, \quad \hat{\psi}_{r+1}(t_k) = -1 \end{aligned} \quad (5.28)$$

и уравнений линейной фильтрации (4.102) и (4.103) при наличии управления

$$\frac{dx(t)}{dt} = F_t \hat{x}(t) + D(t) C_t^T \mathcal{N}_{\lambda t}^{-1}(y_t - C_t \hat{x}_t) + V_t U_t \text{sign} \hat{\psi}, \quad \hat{x}(t_0) = x_0,$$

$$\frac{dD(t)}{dt} = F_t D(t) + D(t) F_t^T + \Gamma_t \mathcal{N}_x \Gamma_t^T - D(t) C_t^T \mathcal{N}_{\lambda t} C_t D(t), \quad D(t_0) = D(0).$$

При этом учитываются наложенные на вектор фазовых координат ограничения на конечное состояния, например $M\{x_k^T x_k\} \leq \rho$.

Решение этой двухточечной задачи достаточно сложно и достигается численными методами путем подбора переменных. Наиболее простое решение можно получить, подобрав для каждой составляющей вспомогательной функции стационарное решение уравнения (5.28)

$$\hat{\psi}_i = \sum_{j=1}^r b_{ij} \exp\{\alpha_j t\}, \quad (5.29)$$

где α_j - корни характеристического уравнения, а коэффициенты b_{ij} зависят от неопределенных множителей Лангража λ_i . Как следует из анализа этого выражения, сумма (5.29) переходит через ноль не более $r-1$ раз. Поэтому управление u_i имеет в общем случае $r-1$ переключений, т.е. r интервалов, на которых управление равно $\pm U_i$.

5.3.1.3. УПРАВЛЕНИЕ ПРИ ЛОКАЛЬНОМ КРИТЕРИИ ОПТИМИЗАЦИИ

Управление при локальном критерии оптимизации характерно для систем программного управления различными объектами промышленного назначения, а также полетом летательных и космических аппаратов при отсутствии ограничений на фазовые координаты объекта.

Выражения для рассматриваемого критерия можно получить, применяя принцип максимума. Введем в качестве дополнительной координаты зависящий от времени функционал качества

$$J(x, x^{\text{пп}}, u, t) = x_{r+1}(t),$$

являющейся решением дифференциального уравнения

$$\dot{x}_{r+1}(t) = \dot{J}(x, x^{\text{пп}}, u, t), \quad x_{r+1}(t_0) = J_{t_0}$$

где x^{np} - вектор теоретически требуемых (программируемых) фазовых координат, удовлетворяющих нелинейному уравнению $\dot{x}^{np}(t)=f_t(x^{np}(t))$, $x^{np}(t_0)=x^{np}(0)$, а в случае линейной задачи уравнению

$$\dot{x}^{np}(t)=F_t x^{np}(t), \quad x^{np}(t_0)=x^{np}(0),$$

Функция Гамильтона имеет вид

$$H_\psi(x, x^{np}, u, \psi, t) = \psi^T \dot{x} + \psi_{r+1} \dot{x}_{r+1}. \quad (5.30)$$

Ввиду отсутствия ограничений и требований на конечное состояние объекта, а также из-за необходимости минимизации функционала качества в каждый текущий момент времени минимизируемый функционал $\Pi(t)=x_{r+1}(t)$, $\psi_{r+1}(t)=-1$, а вектор вспомогательной переменной ψ размера $(r \times 1)$ равен нулю. Поэтому для второго слагаемого соотношения (5.30) можно записать

$$H_\psi = -\dot{J}(x_t, x_t^{np}, u_t).$$

С учетом изложенного стохастический принцип максимума при измерении вектора u и ограничениях на вектор управления (5.2) принимает форму

$$\begin{aligned} \max_{u \in U} \hat{H}_\psi(\hat{x}_t, x_t^{np}, \psi_t, u_t, t) &= \max_{u \in U} \{-M \{ \dot{J}(x_t, x_t^{np}, u_t) | y_{t_0}^t \} \} = \\ &= \max_{u \in U} \left\{ -M \left\{ \left[\frac{\partial J}{\partial t} + \left(\frac{\partial J}{\partial x_t} \right)^T \dot{x}_t + \left(\frac{\partial J}{\partial x_t^{np}} \right)^T \dot{x}_t^{np} + \left(\frac{\partial J}{\partial u_t} \right)^T u_t \right] y_{t_0}^t \right\} \right\}. \end{aligned}$$

Таким образом, при локальном критерии для обеспечения минимального значений апостериорного среднего функционала качества в каждый текущий момент времени за счет локально-оптимального вектора управления достаточно, чтобы апостериорное среднее производной этого функционала по времени, взятое с обратным знаком, имело бы максимальное значение.

Если управление принадлежит открытой области, то вектор оптимального управления удовлетворяет уравнению (5.23) и стохастический принцип максимума записывается в форме

$$M \left\{ \left[\frac{\partial^2 J}{\partial u_t \partial t} + \left(\frac{\partial^2 J}{\partial u_t \partial x_t} \right)^T \dot{x}_t + \left(\frac{\partial^2 J}{\partial u_t \partial x_t^{np}} \right)^T \dot{x}_t^{np} + \right. \right. \\ \left. \left. + \left(\frac{\partial^2 J}{\partial u_t \partial u_t^T} \right)^T \dot{u}_t \right] y_{t_0}^t \right\} = 0. \quad (5.31)$$

Важным достоинством локального критерия является то, что двухточечная задача при его применении не возникает. Это значительно упрощает алгоритм оптимального управления.

Рассмотрим на основе локального критерия стохастическую задачу синтеза управления линейной системы при воздействии случайных возмущений и неточных измерениях.

Уравнения объекта управления и наблюдения, как и в предыдущей задаче, определяются соответственно выражениями (5.24) и (5.25).

Минимизируемый функционал имеет вид положительно-определенной квадратичной формы (5.3), удовлетворяющей дифференциальному векторно-матричному уравнению

$$\dot{x}_{r+1}(t) = \dot{J}(x_t, x_t^{np}, u, t) = 2(\dot{x}_t - \dot{x}_t^{np})^T K_t (x_t - x_t^{np}) + \\ + (x_t - x_t^{np})^T \dot{K}_t (x_t - x_t^{np}) + (x_t - x_t^{np})^T L_t (x_t - x_t^{np}) + \\ + u_t^T W_t u_t, \quad x_{r+1}(t_0) = x_{r+1}(0), \quad (5.32)$$

где учтено, что $(\dot{x}_t - \dot{x}_t^{np})^T K_t (x_t - x_t^{np}) = (x_t - x_t^{np})^T K_t (\dot{x}_t - \dot{x}_t^{np})$.

Матрицы K_t , L_t и W_t задаются из физических требований, диктуемых смыслом задачи оптимизации, назначением и структурой объекта. Для устойчивого объекта матрицы K_t и L_t можно определить из дополнительного уравнения связи [30]

$$F_t^T K_t + K_t F_t = -L_t.$$

В предположении, что управление принадлежит открытой области с учетом уравнений (5.24) и (5.32) стохастический принцип максимума на основании (5.31) приводит к уравнению

$$\frac{\partial M\{\dot{J} | y_{t_0}^t\}}{\partial u_t} = V_t^T K_t(\hat{x}_t - x_t^{np}) + W_t u_t^{onm} = 0. \quad \text{Отсюда} \quad \text{вектор}$$

оптимального управления

$$u_t^{onm} = -W_t V_t^T K_t(\hat{x}_t - x_t^{np}), \quad (5.33)$$

где \hat{x}_t - апостериорное среднее (достаточная статистика) вектора состояния объекта, определяемого из уравнений (4.102) и (4.103).

Соотношение (5.33) подтверждает справедливость теоремы разделения, так как вектор управления u_t^{onm} является детерминированной линейной вектор-функции оценок фазовых координат для момента времени t . По этой причине управление в данный момент времени определяется не всей траекторией наблюдаемого процесса $y_{t_0}^t$ и его достаточной статистикой для того же момента времени, в которой в силу марковского свойства уже учтены все предыдущие значения наблюдаемого процесса.

5.3.2. ПРИНЦИП МАКСИМУМА (МИНИМУМА) ДЛЯ ДИСКРЕТНОГО ВРЕМЕНИ

Необходимость в управлении дискретной динамической системой возникает в связи с дискретностью самого процесса управления из-за применения цифровой вычислительной техники, а также для получения приближенных (численных) решений непрерывных задач. Строго говоря, принцип максимума не распространяется на дискретные динамические системы, если описывающие их поведение уравнения в конечных разностях не связаны с конечномерной аппроксимацией непрерывного процесса. При обосновании применения принципа максимума для дискретных систем обычно исходят из того, что, чем меньше шаг дискретизации, тем точнее выполняется этот принцип.

В общем виде дискретная динамическая система с учетом вектора управления u описывается векторным нелинейным конечно-разностным уравнением (4.207)

$$x_{n+1} = f_n(x_n, u_n, \eta_n), \quad (5.34)$$

где $n=0,1,2,\dots$ - моменты дискретизации текущего дискретного времени, $t_n=n\Delta t$, Δt - шаг дискретизации. В случае терминальной задачи $n=0,k$, $k+1$ - количество временных шагов управления, t_{k+1} - конечный момент состояния системы. Функционал качества при переходе к дискретному процессу записывается в виде $J(x_n, x_n^{np}, u_n, n)$ и принимает одну из форм (5.3) и (5.6) после замены непрерывного аргумента t на дискретный, а интегралов на суммы.

5.3.2.1. УПРАВЛЕНИЕ ПРИ ТЕРМИНАЛЬНОМ КРИТЕРИИ ОПТИМИЗАЦИИ

Рассматривается статистический синтез оптимального дискретного управления конечным состоянием объекта. Необходимое для формирования этого управления апостериорное среднее функционала на основании (5.6) имеет вид

$$\hat{J} = M \left\{ \left[F(x_{k+1}, x_{k+1}^{np}) + \sum_{n=0}^k \varphi_n(x_n, x_n^{np}, u_n) \right] y_1^k \right\}, \quad (5.35)$$

где первое слагаемое является заданной функцией конечного состояния, а второе учитывает интегральную составляющую функционала общего вида, x_n^{np} - вектор теоретически требуемых (программируемых) фазовых координат объекта в дискретном времени.

Представим функцию конечного состояния в виде разложения в ряд^{*)}

$$\begin{aligned} F(x_{k+1}) &= F(x_k) + \left(\frac{\partial F(x_k)}{\partial x_k} \right) (x_{k+1} - x_k) + 0(\Delta x) = \\ &= F(x_0) + \sum_{n=0}^k \left(\frac{\partial F(x_n)}{\partial (x_n)} \right) (x_{n+1} - x_n) + 0(\Delta x), \end{aligned}$$

где $\Delta x = \max \Delta x_n = \max (x_{n+1} - x_n)$, $0 \leq n \leq k$.

Выбираем интервал дискретизации Δt таким малым, чтобы членами разложения второго и более высоких порядков малости

^{*)} Для упрощения записи переменную x_n^{np} опускаем.

относительно $\Delta x \rightarrow 0$ (Δx) можно было бы пренебречь. В результате получаем функционал качества в форме

$$\hat{J} = M \left\{ \sum_{n=0}^k f_{r+1,n}(x_n, u_n, \eta_n) \mid y_1^k \right\}, \quad (5.36)$$

где функция $f_{r+1,n}(x_n, u_n, \eta_n)$ входит в рекуррентное уравнение для дополнительной координаты

$$x_{r+1,n+1} = x_{r+1,n} + f_{r+1,n}(x_n, u_n, \eta_n), \quad x_{r+1,0} = 0, \quad 0 \leq n \leq k \quad (5.37)$$

и имеет вид

$$f_{r+1,n}(x_n, u_n, \eta_n) = \varphi_n(x_n, u_n, \eta_n) + \left(\frac{\partial F(x_n)}{\partial x_n} \right) (x_{n+1} - x_n) + F(x_0). \quad (5.38)$$

Таким образом, за счет расширения вектора состояния системы еще на одну координату x_{r+1} функционал качества (5.35) приобретает вид

$$\hat{J} = M \{ x_{r+1,k+1} \mid y_1^k \}. \quad (5.39)$$

В результате задача оптимального управления конечным состоянием, как и задача с непрерывным временем, свелась к оптимизации апостериорного среднего по дополнительной координате $x_{r+1,k+1}$ расширенного вектора состояния объекта и заключается в отыскании среди допустимых такого вектора управления, который минимизировал бы функционал (5.39). Скалярная функция Гамильтона (стохастический гамильтониан) в дискретном времени с учетом уравнения состояния объекта (5.34) принимает вид

$$\begin{aligned} H_{\psi}(x_n, \psi_{n+1}, \eta_n, u_n, n) &= \psi_{n+1}^T x_{n+1} = \\ &= \psi_{n+1}^T f_n(x_n, u_n, \eta_n) = \sum_{i=1}^{r+1} \psi_{i,n+1} f_{in}. \end{aligned} \quad (5.40)$$

Эта функция удовлетворяет уравнениям

$$x_{i,n+1} = \frac{\partial H_\psi}{\partial \psi_{i,n+1}} = f_{in}, \quad i = \overline{1, r+1}, \quad (5.41)$$

$$\psi_{in} = \frac{\partial H_\psi}{\partial x_{in}} = \sum_{j=1}^{r+1} \psi_{j,n+1} \frac{\partial f_{jn}}{\partial x_{in}}, \quad i = \overline{1, r+1}. \quad (5.42)$$

с граничными условиями

$$x_{i0} = x_i(0), \quad \psi_{r+1,k+1} = 1,$$

$$\psi_{i,k+1} = \frac{\partial \Pi_{k+1}}{\partial x_{i,k+1}}, \quad i = \overline{1, r+1}.$$

Покажем, что уравнения (5.41) и (5.42) после предельного перехода $\Delta t \rightarrow 0$ принимают форму дифференциальных уравнений (5.15) и (5.16). Конечно-разностные уравнения (5.34) и (5.42) в случае малых значений Δt и пренебрежения членами порядка малости $O(\Delta t)$ можно заменить их аналогами

$$x_{i,n+1} = x_{in} + \Delta t f_{in}(x_n, u_n, \eta_n),$$

$$\psi_{in} = \psi_{i,n+1} + \Delta t \sum_{j=1}^{r+1} \psi_{j,n+1} \frac{\partial f_{jn}(x_n, u_n, \eta_n)}{\partial x_i}.$$

При выводе последнего выражения согласно формуле (5.42) использовался гамильтониан

$$H_\psi = \psi_{n+1}^T [x_n + \Delta t f_n(x_n, u_n, \eta_n)]. \quad (5.43)$$

С учетом соотношений (5.41), (5.42) и (5.43) конечно-разностные уравнения при $\Delta t \rightarrow 0$ принимают форму исходных дифференциальных уравнений (5.15) и (5.16).

Обратим внимание на то, что в рассматриваемой задаче с дискретным временем по сравнению с задачей для непрерывных процессов знак вспомогательной переменной ψ изменен на противоположный, т.е. положительный. Следовательно, изменился знак гамильтониана H_ψ на положительный и принцип оптимальности трансформировался в принцип минимума.

В задачах статистического синтеза оптимального управления необходимые выражения апостериорных средних стохастического гамильтониана и соответствующих переменных принимают вид

$$\begin{aligned}\hat{H}_\psi &= M\{H_\psi(x_n, \psi_{n+1}, u_n, n) | y_1^n\}, \\ \hat{x}_{i,n+1} &= M\left\{\frac{\partial H_\psi}{\partial \psi_{i,n+1}} \middle| y_1^n\right\},\end{aligned}\tag{5.44}$$

$$\hat{\psi}_{in} = M\left\{\frac{\partial H_\psi}{\partial x_{in}} \middle| y_1^n\right\}.\tag{5.45}$$

В общем случае двухточечной задачи дискретных систем принцип минимума формулируется следующим образом. Для оптимальных значений векторов состояния x и управления u , удовлетворяющего ограничениям (5.2) и обеспечивающего перевод объекта из начального состояния в конечное при измерении вектора y на интервале (t_1, t_k) , апостериорное среднее функционала Π_{k+1} принимает минимальное значение, а необходимое условие оптимальности состоит в достижении апостериорного среднего стохастического гамильтониана минимального или наименьшего значения

$$\begin{aligned}\min_{u \in U} \hat{H}_\psi(\hat{x}_n, \hat{\psi}_{n+1}, u_n, \eta_n, n) &= \min_{u \in U} M\{H_\psi(x_n, \psi_{n+1}, u_n, \eta_n, n) | y_1^n\}, \\ \inf_{u \in U} \hat{H}_\psi(\hat{x}_n, \hat{\psi}_{n+1}, u_n, \eta_n, n) &= \inf_{u \in U} M\{H_\psi(x_n, \psi_{n+1}, u_n, \eta_n, n) | y_1^n\}, \\ 0 &\leq n \leq k.\end{aligned}$$

Если управление принадлежит открытой области, то вектор управления в дискретном времени определяется из векторно-матричного уравнения

$$M\left\{\frac{\partial H_\psi}{\partial u_n} \middle| y_1^n\right\} = 0, \quad 0 \leq n \leq k.\tag{5.46}$$

Задачи по оптимизации управления, переводящего объект из начального состояния в конечное при минимизации функционала конечного состояния, имеют большое практическое значение, когда

важен конечный результат. К их числу относится, например, рассматриваемая задача определения корректирующих импульсов $\{u_n\}$ траектории движения космического аппарата, обеспечивающих вывод этого аппарата на орбиту с требуемой конечной точностью при ограничениях на энергозатраты (топливо).

Указанный вид задач относится к терминальным с фиксированным временем управления $t_0-t_k=k\Delta t$, $t_0=0$ и свободным правым концом фазовых координат. Поэтому ввиду отсутствия ограничений $\Pi_{k+1}=x_{r+1,k+1}$; $\psi_{i,k+1}=0$ для $i=\overline{1,r}$; $\psi_{r+1,n+1}=1$ для $n=\overline{0,k}$.

Рассмотрим один из возможных вариантов решения задачи указанного типа для линейного объекта с аддитивным управлением, описываемого векторно-матричным уравнением (4.210)

$$x_{n+1}=\Phi_n x_n+V_n u_n+\Gamma_n \eta_n, \quad x_0=x(0), \quad (5.47)$$

где $\Phi_n=1+\Delta t F_n$, $V_n=\Delta t V(t_n)$, $\Gamma_n=\Delta t \Gamma(t_n)$.

Вектор измерения определяется из уравнения (4.81)

$$y_{n+1}=C_{n+1}x_{n+1}+\lambda_{n+1}, \quad n=\overline{0,k}.$$

На вектор управления никаких ограничений не накладывается. Отличие этой задачи от классической вариационной задачи заключается в ее статистическом характере: действии случайных возмущений на объект, в случайных начальных условиях и неточных измерениях. Входящие в функционал качества (5.35) функции имеют следующий вид

$$F(x_n, x_n^{np})=(x_n-x_n^{np})^T K_n (x_n-x_n^{np}), \quad (5.48)$$

$$\varphi_n(x_n, x_n^{np}, u_n)=(x_n-x_n^{np})^T L_n (x_n-x_n^{np})+u_n^T W_n u_n,$$

где второе слагаемое функции φ_n учитывает в функционале качества суммарные энергозатраты.

Диагональную матрицу K_n , как для задачи с непрерывным временем, полагаем заданной. В связи с заменой в функционале интеграла на сумму дискретные отчеты соответствующих матриц связаны с отсчетами их непрерывных аналогов зависимостями $L_n=\Delta t L(t_n)$, $W_n=\Delta t W(t_n)$.

Стохастический гамильтониан с учетом уравнения программируемого вектора состояния

$$x_{n+1}^{np} = \Phi_n x_n^{np}, \quad x_0^{np} = x^{np}(0), \quad (5.49)$$

соотношений (5.38), (5.40), (5.47) и (5.48) равен

$$\begin{aligned} H_\psi(x_n, \psi_{n+1}, \eta_n, u_n, n) &= (x_n - x_n^{np})^T L_n (x_n - x_n^{np}) + u_n^T W_n u_n + \\ &+ 2[(x_n - x_n^{np})^T K_n (x_{n+1} - x_n^{np} - x_n + x_n^{np})] + \psi_{n+1}^T x_{n+1} + F(x_0) = \\ &= (x_n - x_n^{np})^T L_n (x_n - x_n^{np}) + u_n^T W_n u_n + 2(x_n - x_n^{np})^T K_n \times \\ &[\Phi_n x_n + V_n u_n + \Gamma_n \eta_n - \Phi_n x_n^{np} - x_n + x_n^{np}] + \\ &+ \psi_{n+1}^T (\Phi_n x_n + V_n u_n + \Gamma_n \eta_n) + F(x_0). \end{aligned} \quad (5.50)$$

Оптимальное управление отыскивается в открытой области из условия (5.46). Учитывая выражение (5.50), получаем

$$M \left\{ \frac{\partial H_\psi}{\partial u_n} \middle| y_1^n \right\} = M \{ [2W_n u_n^{onm} + 2V_n^T K_n (x_n - x_n^{np}) + V_n^T \psi_{n+1}] | y_1^n \} = 0.$$

После операции усреднения относительно вектора измерения y (получения апостериорного среднего) приходим к векторно-матричному уравнению

$$2V_n^T K_n (\hat{x}_n - x_n^{np}) + 2W_n u_n^{onm} + V_n^T \hat{\psi}_{n+1} = 0.$$

Отсюда оптимальный вектор управления

$$u_n^{onm} = -W_n^{-1} V_n^T \left[K_n (\hat{x}_n - x_n^{np}) + \frac{1}{2} \hat{\psi}_{n+1} \right], \quad n = \overline{0, k}, \quad (5.51)$$

где $\hat{x}_n = M\{x_n | y_1^n\}$, $\hat{\psi}_{n+1} = M\{\psi_{n+1} | y_1^n\}$.

Оценка вектора вспомогательной переменной $\hat{\psi}$ определяется из выражения (5.45), которое на основании (5.50) с учетом соотношения $\Phi_n = I + \Delta t F_n$ принимает рекуррентную форму

$$\begin{aligned}\hat{\psi}_n = & 2L_n(\hat{x}_n - x_n^{np}) + 2K_n(\hat{x}_{n+1} - x_n^{np}) - 2K_n(\hat{x}_n - x_n^{np}) + \\ & + \hat{\psi}_{n+1} + \Delta t F_n^T \hat{\psi}_{n+1} + 2\Delta t F_n^T K_n(\hat{x}_n - x_n^{np})\end{aligned}\quad (5.52)$$

или

$$\begin{aligned}\hat{\psi}_n = & 2L_n(x_n - x_n^{np}) + 2K_n(\hat{x}_{n+1} - x_{n+1}^{np}) - 2K_n(\hat{x}_n - x_n^{np}) + \\ & + \hat{\psi}_{n+1} + \Delta t F_n^T \hat{\psi}_{n+1} + 2\Delta t F_n^T K_n(\hat{x}_{n+1} - x_{n+1}^{np}) + 0_1(\Delta t),\end{aligned}\quad (5.53)$$

где член второго порядка малости относительно Δt : $0_1(\Delta t) = -2\Delta t^2 F_n^T K_n(\varphi_n^\Delta - F_n x_n^{np})$, появился вследствие замены в последнем слагаемом соотношения (5.52) вектора ошибки $\delta\hat{x}_n = \hat{x}_n - x_n^{np}$ на $\delta\hat{x}_{n+1} = \hat{x}_{n+1} - x_{n+1}^{np}$,

$$\begin{aligned}\varphi_n^\Delta = & F_n \hat{x}_n + V(t_n)u_n + \\ & + \tilde{D}_{n+1} C_{n+1}^T [\Delta t C_{n+1} \tilde{D}_{n+1} C_{n+1}^T + \mathcal{N}_{\lambda, n+1}]^{-1} [y_{n+1} - C_{n+1} \tilde{x}_{n+1}].^{*})\end{aligned}$$

Введем обозначения для нового вектора

$$z_n = K_n(x_n - x_n^{np}) + \frac{1}{2}\psi_n. \quad (5.54)$$

В результате рекуррентное уравнение (5.53) относительно этого вектора принимает вид

$$\hat{z}_n = \hat{z}_{n+1} + \Delta t F_n^T \hat{z}_{n+1} + L_n \delta\hat{x}_n + 0_1(\Delta t) = \Phi_n^T \hat{z}_{n+1} + L_n \delta\hat{x}_n + 0_1(\Delta t), \quad (5.55)$$

где $\hat{z}_n = M\{z_n | y_1^n\}$, $\delta(\hat{x}_n) = M\{\delta(x) | y_1^n\}$.

Граничное состояние вектора \hat{z} ввиду $\psi_{k+1}=0$ ($i=\overline{1, r}$) согласно формуле (5.54)

$$z_{k+1} = K_{k+1} \delta\hat{x}_{k+1}. \quad (5.56)$$

Вектор оптимального управления (5.51) связан с вектором \hat{z} зависимостью

*) Процедура замены сводится к введению в сомножитель $\delta\hat{x}_n$ этого слагаемого выражения $(\hat{x}_{n+1} - x_{n+1}^{np}) - (\hat{x}_{n+1} - x_{n+1}^{np})$ и определения на основании уравнений (4.88) и (5.49) значений $(\hat{x}_{n+1} - \hat{x}_n)$ и $(x_{n+1}^{np} - x_n^{np})$.

$$u_n^{onm} = -W_n^{-1}V_n^T \hat{z}_{n+1} + \Delta t W_n^{-1}V_n^T (\varphi_n^\Delta - F_n x_n^{np}), \quad (5.57)$$

где второе слагаемое в правой части (5.57) появилось вследствие замены $\delta \hat{x}_n$ на $\delta \hat{x}_{n+1}$.

На основании уравнений (5.47), (5.49) и (4.88) с учетом (5.57) представим вектор ошибок и его апостериорное среднее в виде

$$\begin{aligned} \delta x_{n+1} &= x_{n+1} - x_{n+1}^{np} = \Phi_n \delta x_n + V_n u_n + \Gamma_n \eta_n, \\ \delta \hat{x}_{n+1} &= \Phi_n \delta \hat{x}_n - V_n W_n^{-1} V_n^T z_{n+1} + \Gamma_n \eta_n + 0_2(\Delta t), \end{aligned} \quad (5.58)$$

где слагаемое $0_2(\Delta t) = \Delta t^2 V_n(t_n) W_n^{-1}(t_n) V_n^T(t_n) (\varphi_n^\Delta - F_n x_n^{np})$ ввиду $V_n = \Delta t_n(t_n)$, $W_n = \Delta t W(t_n)$ является членом более высокого порядка малости относительно Δt .

Вектор \hat{z} отыскивается в результате совместного решения уравнений (5.55) и (5.58) в виде линейной формы

$$\hat{z}_n = P_n \delta \hat{x}_n, \quad (5.59)$$

где P_n подлежащая определению неизвестная матрица. Подставив выражение (5.59) в уравнение (5.55) имеем

$$P_n \delta \hat{x}_n = \Phi_n^T P_{n+1} \delta \hat{x}_{n+1} + L_n \delta \hat{x}_n + 0_1(\Delta t).$$

А после подстановки $\delta \hat{x}_{n+1}$ из уравнения (5.58) с учетом (5.59) получаем уравнение

$$\begin{aligned} P_n \delta \hat{x}_n &= \Phi_n^T P_{n+1} [I + V_n W_n^{-1} V_n^T P_{n+1}]^{-1} [\Phi_n \delta \hat{x}_n + \Gamma_n \hat{\eta}_n] + \\ &+ L_n \delta \hat{x}_n + 0(\Delta t), \end{aligned} \quad (5.60)$$

где I - единичная матрица.

В дальнейшем из-за существенно малых значений интервала дискретизации Δt слагаемыми вида $0_1(\Delta t)$, $0_2(\Delta t)$ и $0(\Delta t)$ пренебрегаем.

Для того, чтобы матрица P не зависела от результатов измерений фазовых координат, усредним (5.60) относительно всех возможных реализаций вектора y . Обозначив $M\{\delta \hat{x}\} = m_\delta$, с учетом $M\{\hat{\eta}\} = 0$ имеем

$$\{P_n - \Phi_n^T P_{n+1} [I + V_n W_n^{-1} V_n^T P_{n+1}]^{-1} \Phi_n - L_n\} m_\delta = 0. \quad (5.61)$$

Ввиду того, что условие (5.61) должно выполняться при любом m_δ , получаем не зависящее от фазовых координат объекта конечно-разностное (рекуррентное) нелинейное матричное уравнение

$$P_n = \Phi_n^T P_{n+1} [I + V_n W_n^{-1} V_n^T P_{n+1}]^{-1} \Phi_n + L_n. \quad (5.62)$$

Используем формулу приложения 2 для представления выражения в квадратных скобках уравнения (5.62) в виде ряда по степеням матрицы $V_n W_n^{-1} V_n^T P_{n+1}$. После преобразования и приведения подобных, пренебрегая членами второго и более высоких порядков относительно Δt , приходим к нелинейному матричному уравнению Риккати в дискретном времени

$$P_n = P_{n+1} + \Delta t F_n^T P_{n+1} + \Delta t P_{n+1} F_n - P_{n+1} V_n W_n^{-1} V_n^T P_{n+1} + L_n. \quad (5.63)$$

Уравнение (5.63) решается численными методами при граничном (конечном) условии P_{k+1} . Граничное условие определяется из выражений (5.59) и (5.56): $\hat{z}_{k+1} = P_{k+1} \delta \hat{x}_{k+1}$, $\hat{z}_{k+1} = K_{n+1} \delta \hat{x}_{k+1}$. В результате имеем $P_{k+1} = K_{k+1}$.

Физически реализуемое управление в момент времени t_n в случае марковских процессов зависит от оценок фазовых координат объекта для того же момента времени \hat{x}_n . Однако, согласно формуле (5.57) оптимальное управление зависит от вектора $\delta \hat{x}_{n+1}$ в момент времени t_{n+1} ввиду присутствия вектора $\hat{z}_{n+1} = P_{n+1} \delta \hat{x}_{n+1}$. Для приведения в соответствие с реальной процедурой измерения, как и в предыдущих случаях, заменим вектор $\delta \hat{x}_{n+1}$ на $\delta \hat{x}_n$. В результате вновь полученное слагаемое $-\Delta t W_n^{-1} V_n^T P_{n+1} (\varphi_n^\Delta - F_n x_n^{np})$ вместе со вторым слагаемым выражения (5.57) после подстановки выражения оптимального управления в уравнение фильтрации (5.65) оказываются членами второго порядка малости относительно Δt ввиду $V_n = \Delta t V(t_n)$. Это позволяет представить вектор оптимального управления в окончательном виде

$$u_n^{onm} = -W_n^{-1} V_n^T P_{n+1} (\hat{x}_n - x_n^{np}), \quad (5.64)$$

так как при $\Delta t \rightarrow 0$ этими слагаемыми можно пренебречь.

Оценка вектора \hat{x}_{n+1} определяется на основании дискретного алгоритма фильтрации, описываемого рекуррентными уравнениями (4.88) и (4.86) с учетом рекомендаций разд.4.12

$$\begin{aligned}\hat{x}_{n+1} = \tilde{x}_{n+1} + \tilde{D}_{n+1}C_{n+1}^T[C_{n+1}\tilde{D}_{n+1}C_{n+1}^T + \\ + Q_{\lambda,n+1}]^{-1}[y_{n+1}C_{n+1}\tilde{x}_{n+1}], \quad \hat{x}_0 = x(0),\end{aligned}\quad (5.65)$$

$$\begin{aligned}D_{n+1} = \tilde{D}_{n+1} - \tilde{D}_{n+1}C_{n+1}^T[C_{n+1}\tilde{D}_{n+1}C_{n+1}^T + \\ + Q_{\lambda,n+1}]^{-1}C_{n+1}\tilde{D}_{n+1}, \quad D_0 = D(0),\end{aligned}\quad (5.66)$$

$$\begin{aligned}\text{где } \tilde{x}_{n+1} = \Phi_n \hat{x}_n - V_n W_n^{-1} V_n^T P_{n+1} (\hat{x}_n - x_n^{np}), \\ \tilde{D}_{n+1} = \Phi_n D_n \Phi_n^T + \Gamma_n Q_{\eta n} \Gamma_n^T\end{aligned}$$

Из соотношения(5.64) следует, что при линейном объекте и функционале качества в квадратичной форме управление является детерминированной линейной функцией вектора оценки фазовых координат объекта. В этом определении содержится подтверждение сформулированной ранее теоремы разделения.

5.3.2.2. УПРАВЛЕНИЕ ПРИ ЛОКАЛЬНОМ КРИТЕРИИ ОПТИМИЗАЦИИ

К стохастическому принципу минимума для локального критерия оптимизации можно прийти, определив стохастический гамильтониан с учетом высказанных в разделе 5.3.1.2 замечаний об отсутствии ограничений и требований на конечное состояние объекта, а также о необходимости минимизации функционала качества на каждом шаге дискретного времени. В результате стохастический гамильтониан при $\psi_{n+1}=0$ ($i=\overline{1,r}$) и $\psi_{r+1,n}=1$ оказывается равным $H_\psi(f_{r+1,n}(x_n, u_n, \eta_n))$ и стохастический принцип минимума состоит в минимизации его апостериорного среднего в каждый момент времени t_n

$$\min_{u_n \in U} M\{f_{r+1,n}(x_n, u_n, \eta_n) | y_1^n\}, \quad n = 0, 1, \dots,$$

где на основании данных раздела 5.2.2.1

$$f_{r+1,n}(x_n, u_n, \eta_n) = \varphi_n(x_n, u_n, \eta_n) + \left(\frac{\partial F(x_n)}{\partial x_n} \right) (x_{n+1} - x_n) + F_0(x_0) \quad (5.67)$$

Если управление принадлежит открытой области, то вектор оптимального управления определяется из условия

$$M \left\{ \frac{\partial H_\psi}{\partial u_n} \bigg| y_1^n \right\} = M \left\{ \frac{\partial f_{r+1,n}(x_n, u_n, \eta_n)}{\partial u_n} \bigg| y_1^n \right\} = 0. \quad (5.68)$$

При функционале качества квадратичного типа функции, входящие в выражение (5.67), имеют вид (5.48) и стохастический гамильтониан принимает форму (5.50) без слагаемого $\psi_{n+1}^T (\Phi_n x_n + V_n u_n + \Gamma_n \eta_n)$. После операции усреднения условие (5.68) приводит к векторно-матричному уравнению $V_n^T K_n (\hat{x}_n - x_n^{np}) + W_n u_n^{onm} = 0$.

Отсюда оптимальный вектор управления

$$u_n^{onm} = -W_n^{-1} V_n^T K_n (\hat{x}_n - x_n^{np}). \quad (5.69)$$

Оценка вектора \hat{x}_{n+1} определяется из уравнений фильтрации (5.65) и (5.66), где $\tilde{x}_{n+1} = \Phi_n \hat{x}_n - V_n W_n^{-1} V_n^T K_n (\hat{x}_n - x_n^{np})$.

Матрицу K_n , как и для задачи с непрерывным временем (раздел 5.2.1.2), определяем из уравнения связи $F_n^T K_n + K_n F_n = -L_n$. Полученные соотношения для локального критерия оптимизации также подтверждают справедливость теоремы разделения, так как оптимальное управление является детерминированной линейной функцией оценки вектора состояния объекта.

Сравним полученное решение с решением задачи для терминального критерия оптимальности на основании минимизации конечного состояния объекта на заданном интервале времени управления. Из формул (5.64) и (5.69) следует, что структура выражений оптимального управления одинакова. Различие состоит в различных значениях компонент матриц P и K . При $t=t_{k+1}$ ввиду $P_{k+1}=K_{k+1}$ (раздел 5.2.2.1) эти формулы дают одинаковый результат. Из этого сравнения следует, что управление по локальному критерию более интенсивное, так как требует выполнения критерия в каждый

отсчет времени. В тоже время по этой причине оно менее экономичное.

5.3.3. ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

5.3.3.1. ВВЕДЕНИЕ

Динамическое программирование основано на принципе оптимальности, сформулированном Р. Беллманом как необходимое условие оптимального процесса для дискретных систем. Этот принцип утверждает, что оптимальная стратегия не зависит от предистории системы и определяется только начальными условиями и конечной целью. В этом случае вариационная задача рассматривается как многошаговый процесс решения более простых задач и оптимальное управление отыскивается последовательно шаг за шагом. Отличительная особенность рассматриваемого здесь подхода состоит в том, что участки оптимальной траектории определяются в обратной последовательности, начиная с конечного (целевого) состояния объекта. В соответствии с принципом оптимальности при делении траектории на несколько участков можно, двигаясь в обратном направлении, убедиться в оптимальности участка примыкающего к последнему, а затем в оптимальности всех предшествующих участков. Оптимальность отдельных участков зависит от оптимальности всей траектории.

5.3.3.2. СТАТИСТИЧЕСКИЙ СИНТЕЗ ДИСКРЕТНОГО УПРАВЛЕНИЯ

Дискретная динамическая система описывается векторно-матричным в конечных разностях уравнением (5.34)

$$x_{n+1}=f_n(x_n,u_n,\eta_n), \quad x_0=x(0).$$

Критерий оптимальности состоит в минимизации апостериорного среднего функционала качества типа (5.35)

$$\begin{aligned} \hat{J} &= M \left\{ \left[F(x_{k+1}) + \sum_{n=0}^k \varphi_n(x_n, u_n) \right] y_1^k \right\} = \\ &= M_y \left\{ F(x_{k+1}) + \sum_{n=0}^k \varphi_n(x_n, u_n) \right\}, \end{aligned} \quad (5.70)$$

где функция конечного состояния $F(x_{k+1})$ задана.

Рассмотрим задачу управления, в которой определяются векторы управления u_0, u_1, \dots, u_k , обеспечивающие перевод объекта из состояния x_0 в состояние x_{k+1} при измерении вектора y_n за $k+1$ шагов управления с минимальным значением функционала (5.70).

Обозначим через \hat{J}_k апостериорное среднее последней частичной суммы на интервале $k\Delta t \leq t \leq (k+1)\Delta t$ полной суммы функционала качества, которое принимает вид

$$\begin{aligned}\hat{J}_k(\hat{x}_k, u_k) &= M_y\{F(x_{k+1}) + \varphi_k(x_k, u_k)\} = \\ &= M_y\{F(x_k, u_k, \eta_k) + \varphi_k(x_k, u_k)\}.\end{aligned}\quad (5.71)$$

В этом выражении подчеркнута зависимость \hat{J}_k от вектора состояния x_k и управления u_k . На основании принципа оптимальности управление u_k зависит от вектора оценки \hat{x}_k и должно минимизировать частичную сумму (5.71).

Обозначим это минимальное значение через $R_k(\hat{x}_k)$

$$R_k(\hat{x}_k) = \min_{u_k} \hat{J}_k(\hat{x}_k, u_k). \quad (5.72)$$

Эта функция, а также найденный из выражения (5.71) оптимальный вектор u_k^{onm} должны быть заданы в момент времени $t_k = k\Delta t$ как функции вектора оценки \hat{x}_k , полученного из уравнений дискретного алгоритма фильтрации, $u_k^{onm}(\hat{x}_k)$. После определения оптимального управления u_k^{onm} рассмотрим интервал $(k-1)\Delta t \leq t \leq (k+1)\Delta t$. С учетом выражения (5.70) апостериорное среднее частичной суммы полной суммы зависит от управлений u_k и u_{k-1}

$$\begin{aligned}\hat{J}_{k-1}(\hat{x}_{k-1}, u_{k-1}, u_k) &= \hat{J}_k(\hat{x}_k, u_k) + M_y\{\varphi_{k-1}(x_{k-1}, u_{k-1})\} = \\ &= \hat{J}_k(\hat{x}_{k-1}, u_{k-1}, u_k) + M_y\{\varphi_{k-1}(x_{k-1}, u_{k-1})\},\end{aligned}\quad (5.73)$$

где оценка \hat{x}_{k-1} получена из уравнений дискретного алгоритма фильтрации.

После проведения минимизации выражения (5.73) получим следующую функцию, которую будем называть также функцией будущих потерь

$$R_{k-1}(\hat{x}_{k-1}) = \min_{u_k, u_{k-1}} \hat{J}_{k-1}(\hat{x}_{k-1}, u_{k-1}, u_k).$$

Учитывая, что первое слагаемое выражения (5.73) зависит от u_{k-1} и u_k , а второе только от u_{k-1} , получим

$$R_{k-1}(\hat{x}_{k-1}) = \min_{u_k, u_{k-1}} J_k(\hat{x}_{k-1}, u_{k-1}, u_k) + \min_{u_{k-1}} M_y\{\varphi_{k-1}(x_{k-1}, u_{k-1})\}$$

или

$$R_{k-1}(\hat{x}_{k-1}) = \min_{u_{k-1}} \{R_k(\hat{x}_{k-1}, u_{k-1}) + M_y\{\varphi_{k-1}(x_{k-1}, u_{k-1})\}\}, \quad (5.74)$$

так как R_k является результатом минимизации \hat{J}_k по управлению u_k и, следовательно, не зависит от него. Выражение (5.74) должно быть минимизировано по вектору u_{k-1} при заданной оценке вектора координат \hat{x}_{k-1} . В результате определяют оптимальное управление $u_{k-1}^{onm} = u_{k-1}(\hat{x}_{k-1})$ и значение функции будущих потерь $R_{k-1}(\hat{x}_{k-1})$ в момент времени $t_{k-1} = (k-1)\Delta t$. Продолжая эту процедуру для интервала времени $(k-l+1)\Delta t \leq t \leq (k+1)\Delta t$, получим рекуррентную формулу динамического программирования

$$R_{k-l+1}(\hat{x}_{k-l+1}) = \min_{u_{k-l+1}} \{R_{k-l+2}(x_{k-l+1}, u_{k-l+1}) + M_y\{\varphi_{k-l+1}(x_{k-l+1}, u_{k-l+1})\}\}, \quad l = \overline{1, k+1}. \quad (5.75)$$

В результате минимизации выражения (5.75) определяются оптимальное управление $u_{k-l+1}^{onm} = u_{k-l+1}(\hat{x}_{k-l+1})$ и значение функции будущих потерь $R_{k-l+1}(\hat{x}_{k-l+1})$, как функции вектора оценок (достаточных статистик) \hat{x}_{k-l+1} . Как следует из характера самой процедуры оптимальность достигается не только на избранном интервале $(k-l+1)\Delta t \leq t \leq (k-l+2)\Delta t$, но и на всем будущем интервале $(k-l+1)\Delta t \leq t \leq (k+1)\Delta t$. Таким образом, функция будущих потерь

$$R_{k-l+1}(\hat{x}_{k-l+1}) = \min_{\{u_i\}} M \left\{ \left[F(x_{k+1}) + \sum_{i=k-l+1}^k \varphi(x_i, u_i) \right] y_1^{k-l+1} \right\}, \quad i = \overline{k-l+1, k}$$

определяет минимальное значение функционала (5.71), которое достигается на временном интервале $(k-l+1)\Delta t \leq t \leq (k+1)\Delta t$ при условии существования вектора измерения y_1^{k-l+1} .

В результате последовательного применения этой формулы можно найти вектор оптимального управления u_0^{onm} для первого участка траектории и таким образом функцию будущих потерь $R_0(x_0)$ равную минимуму функционала (5.71). В результате полностью определяют векторную последовательность оптимального управления u_k, u_{k-1}, \dots, u_0 .

Далее приступают к формированию оптимального управления траектории движения объекта. По результатам первого измерения y_1 с учетом u_0^{onm} определяют по известному вектору начального состояния x_0 оценку \hat{x}_1 . Повторяя измерения и каждый раз используя ранее полученные значения u_n^{onm} , определяют последовательность векторов оценок $\hat{x}_n, n=k-l+1, l=k, k-1, \dots, 0$.

В общем случае метод динамического программирования, к сожалению, приводит к очень громоздким процедурам отыскания решений. Причем на каждом шаге работы алгоритма необходимо вычислить и запомнить скалярные функции $R_{k-l+1}(\hat{x}_{k-l+1})$ и $R_{k-l+2}(\hat{x}_{k-l+1})$ многих переменных состояния объекта. Зависимости $u_n^{onm} = u_n(\hat{x}_n)$ довольно часто имеют не аналитический вид. Только в простых случаях, один из которых рассмотрен ниже, можно получить результаты в аналитической форме. Основные методы расчета здесь являются численными. Несмотря на это метод динамического программирования позволяет свести задачу оптимизации сложной скалярной функции (функционала качества) многих векторных аргументов к минимизации последовательности скалярных функции от одного векторного аргумента.

Используем метод динамического программирования для определения в явном виде оптимального управляющего воздействия, как функции оценок координат (достаточных статистик) объекта. Движение объекта характеризуется линейным векторно-матричным уравнением в дискретном времени (4.210)

$$x_{n+1} = \Phi_n x_n + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0). \quad (5.76)$$

Линейный измеритель описывается уравнением

$$y_{n+1} = C_{n+1}x_{n+1} + \lambda_{n+1}.$$

Апостериорное среднее показателя качества записывается в квадратичной форме

$$\hat{J} = M_y \left\{ F(x_{k+1}) + \sum_{n=0}^k \varphi_n(x_n, u_n) \right\}, \quad (5.77)$$

где $F(x_{k+1}) = x_{k+1}^T K_{k+1} x_{k+1}$, $\varphi_n(x_n, u_n) = x_n^T L_n x_n + u_n^T W_n u_n$.

Как и ранее полагаем $L_n = \Delta t L(t_n)$, $W_n = \Delta t W(t_n)$, $V_n = \Delta t V(t_n)$. Оценки вектора фазовых координат объекта определяются дискретным линейным алгоритмом фильтрации с помощью уравнений оценок (4.88) и дисперсии ошибок оценок (4.86).

Рассмотрение начнем с определения функции будущих потерь на последнем интервале управления $k\Delta t \leq t \leq (k+1)\Delta t$. Последнюю частичную сумму функционала качества (5.77) с учетом уравнения (5.76) запишем в виде

$$\begin{aligned} \hat{J}_k(\hat{x}_k, u_k) &= M_y \{ [\Phi_k(\hat{x}_k + (x_k - \hat{x}_k)) + V_k u_k + \Gamma_k \eta_k]^T \times \\ &\quad \times K_{k+1} [\Phi_k(\hat{x}_k + (x_k - \hat{x}_k)) + V_k u_k + \Gamma_k \eta_k] + \\ &\quad + (\hat{x}_k + (x_k - \hat{x}_k))^T L_k (\hat{x}_k + (x_k - \hat{x}_k)) + u_k^T W_k u_k \} = \quad (5.78) \\ &= \hat{x}_k^T \Phi_k^T K_{k+1} \Phi_k \hat{x}_k + 2\hat{x}_k^T \Phi_k^T K_{k+1} V_k u_k + u_k^T P_k u_k + \hat{x}_k^T L_k \hat{x}_k + \\ &\quad + Sp(L_k D_k) + Sp(\Phi_k^T K_{k+1} \Phi_k D_k) + Sp(\Gamma_k^T K_{k+1} \Gamma_k D_{\eta k}). \end{aligned}$$

где $Sp(\cdot)$ обозначает взятие следа матрицы, $P_k = W_k + V_k^T K_{k+1} V_k$, $D_k = M_y \{ (x_k - \hat{x}_k)(x_k - \hat{x}_k)^T \}$.

На основании принципа оптимальности необходимо получить минимальное значение частичной суммы $R_k(\hat{x}_k) = \min_{u_k} \hat{J}_k(\hat{x}_k, u_k)$.

Полагая матрицу P_k положительно определенной, а область управления открытой, приходим к оптимальному управлению на шаге k из уравнения $\partial \hat{J}_k(\hat{x}_k, u_k) / \partial u_k = 0$. Отсюда $u_k^{onm} = -A_k \hat{x}_k$, где $A_k = P_k^{-1} V_k^T K_{k+1} \Phi_k$. Подставляя значение u_k^{onm} в выражение (5.78), получаем будущие потери на шаге k

$$R_k(\hat{x}_k) = \hat{x}_k^T \Lambda_k x_k + c_k, \quad (5.79)$$

$$\text{где } \Lambda_k = L_k + \Phi_k^T K_{k+1} \Phi_k - A_k^T P_k A_k, \\ c_k = Sp(L_k D_k) + Sp(\Phi_k^T K_{k+1} \Phi_k D_k) + Sp(\Gamma_k^T K_{k+1} \Gamma_k D_{\eta k})$$

Как следует из полученного результата, оптимальное управление u_k на интервале $t_k \leq t \leq t_{k+1}$ зависит от вектора оценки \hat{x}_k и решает целевую задачу - минимизирует частичную сумму \hat{J}_k .

На следующем интервале $(k-1)\Delta t \leq t \leq (k+1)\Delta t$ частичная сумма апостериорного среднего функционала качества имеет вид

$$\hat{J}_{k-1}(\hat{x}_{k-1}, u_{k-1}, u_k) = \hat{J}_k(\hat{x}_k, u_k) + M_y \{x_{k-1}^T L_{k-1} x_{k-1} + u_{k-1}^T W_{k-1} u_{k-1}\} = \\ = \hat{J}_k(\hat{x}_{k-1}, u_{k-1}, u_k) M_y \{x_{k-1}^T L_{k-1} x_{k-1} + u_{k-1}^T W_{k-1} u_{k-1}\}.$$

Минимизация его по управлению u_{k-1} приводит к функции будущих потерь на шаге $k-1$

$$R_{k-1}(\hat{x}_{k-1}) = \min_{u_k, u_{k-1}} \hat{J}_{k-1}(\hat{x}_{k-1}, u_{k-1}, u_k) = \\ = \min_{u_{k-1}} \{R_k(\hat{x}_{k-1}, u_{k-1}) + \hat{x}_{k-1}^T L_{k-1} \hat{x}_{k-1} + u_{k-1}^T W_{k-1} u_{k-1}\} + \\ + Sp(l_{k-1} D_{k-1}). \quad (5.80)$$

Для определения $R_k(\cdot)$ в выражении (5.80) в функцию будущих потерь (5.79) необходимо подставить уравнение линейной фильтрации, которое для рассматриваемой задачи имеет вид

$$\hat{x}_k = \tilde{x}_k + D_k C_k^T Q_{\lambda k}^{-1} (y_k - C_k \tilde{x}_k), \\ D_k = \tilde{D}_k - \tilde{D}_k C_k^T [C_k \tilde{D}_k C_k^T + Q_{\lambda k}]^{-1} C_k \tilde{D}_k,$$

$$\text{где } \tilde{x}_k = \Phi_{k-1} \hat{x}_{k-1} + V_{k-1} u_{k-1}, \quad \tilde{D}_k = \Phi_{k-1} D_{k-1} \Phi_{k-1}^T + \Gamma_{k-1} Q_{\eta, k-1} \Gamma_{k-1}^T.$$

После минимизации, определения оптимального управления $u_{k-1}^{\text{опт}}$ и последующей его подстановки в (5.80) получаем следующие соотношения

$$u_{k-1}^{\text{онм}} = -A_{k-1} \hat{x}_{k-1}, \\ R_{k-1}(\hat{x}_{k-1}) = \hat{x}_{k-1}^T \Lambda_{k-1} \hat{x}_{k-1} + c_{k-1}, \quad (5.81)$$

где

$$\begin{aligned}
A_{k-1} &= P_{k-1}^{-1} V_{k-1}^T \Lambda_k \Phi_{k-1}, \\
P_{k-1} &= W_{k-1} + V_{k-1}^T \Lambda_k V_{k-1}, \\
\Lambda_{k-1} &= L_{k-1} + \Phi_{k-1}^T \Lambda_k \Phi_{k-1} - A_{k-1}^T P_{k-1} A_{k-1}.
\end{aligned} \tag{5.82}$$

Продолжая эту процедуру с помощью рекуррентной формулы (5.82), доходим до начального интервала $0 \leq t \leq t_{k+1}$, для которого находим оптимальное управление u_0^{opt} . В результате определяем последовательность оптимальных управлений u_0, u_1, \dots, u_k для всей траектории движения объекта. Полученный алгоритм оптимального управления подтверждает справедливость упомянутой ранее теоремы разделения, согласно которой в линейных дискретных системах с квадратичным функционалом качества при аддитивных гауссовских воздействиях оптимальное управление распадается на две последовательно решаемых задачи: определения оценок фазовых координат объекта и закона оптимального управления.

Наряду с рассмотренным разработаны методы динамического программирования для прямой последовательности (прямого хода), при произвольном задании интервала времени $t_{k+1} - t_0$, а также начальных и конечных состояний. Очевидно, задачи со свободным конечным состоянием и фиксированным начальным целесообразно решать прямым ходом, а с заданным конечным и свободным начальным состоянием - обратным ходом. Отметим, что рассмотренные алгоритмы определения оптимального управления являются рекуррентными и поэтому удобными для реализации на вычислительных машинах.

5.3.3.3. СТАТИСТИЧЕСКИЙ СИНТЕЗ УПРАВЛЕНИЯ НЕПРЕРЫВНЫХ СИСТЕМ. УРАВНЕНИЕ БЕЛЛМАНА.

Динамическое программирование при некоторых допущениях может использоваться для непрерывных систем. Пусть движение объекта описывается матрично-векторным уравнением (5.1)

$$\dot{x}_t = f_t(x_t, \eta_t, u_t), \quad x(t_0) = x(0).$$

Управление объектом производится на заданном интервале времени (t_0, t) . Требуется выбрать вектор управления u_t ,

принадлежащий области U и минимизирующий апостериорное среднее функционала качества,

$$\hat{J} = M \left\{ \left[F(x_{t_k}) + \int_{t_0}^{t_k} \varphi_\tau(x_\tau, u_\tau) d\tau \right] y_{t_0}^t \right\} = M_y \left\{ F(x_{t_k}) + \int_{t_0}^{t_k} \varphi_\tau(x_\tau, u_\tau) d\tau \right\},$$

где функция конечного состояния $F(x_{t_k})$ задана.

В процессе управления измеряется вектор $y(t)$, который связан с вектором состояния объекта управления (5.25).

Осуществим предельный переход в рекуррентной формуле (5.75), приняв момент времени t_{k-l+1} за t , а t_{k-l+2} за $t+\Delta t$. Учитывая, что оценка \hat{x}_{k-l+1} , управление u_{k-l+1} , а также измерение в момент времени t_{k-l+2} с точностью до величин высших порядков относительно Δt определяют оценку \hat{x}_{k-l+2} , приходим к следующему выражению будущих потерь на интервале (t, t_k)

$$R_t(\hat{x}_t) = \min_{u_t} \{ R_{t+\Delta t}(\hat{x}_{t+\Delta t}) + M_y \{ \varphi_t(x_t, u_t) \Delta t \} \}, \quad (5.83)$$

где $R_{t+\Delta t}(x_{t+\Delta t})$ - будущие потери на интервале $(t+\Delta t, t_k)$.

Первый член выражения (5.83) разложим в ряд Тейлора по степеням Δt с точностью до линейного члена^{*)}

$$\begin{aligned} R_{t+\Delta t}(\hat{x}_{t+\Delta t}) &= R_t(\hat{x}_t) + \frac{\partial R_t(\hat{x}_t)}{\partial t} \Delta t + \\ &+ \sum_{i=1}^r \frac{\partial R_t(\hat{x}_t)}{\partial \hat{x}_{it}} M_y \{ f_{it}(x_t, \eta_t, u_t) \} \Delta t. \end{aligned} \quad (5.84)$$

Учтем, что $R_t(\hat{x}_t)$ и $\partial R_t(\hat{x}_t)/\partial t$ не содержат u_t , так как по управлению произведена минимизация функции $R_t(\cdot)$. Поэтому $R_t(\hat{x}_t)$ и $\partial R_t(\hat{x}_t)/\partial t$ можно вынести за знак \min_{u_t} . После подстановки соотношения (5.84) в формулу (5.83) приведения подобных, сокращения на Δt , устремляя $\Delta t \rightarrow 0$, получаем нелинейное уравнение Беллмана

^{*)} Предположение, что функция $R_t(\cdot)$ непрерывна и имеет непрерывные частные производные по всем компонентам вектора x_t , является основным допущением при применении метода динамического программирования к непрерывным системам.

$$\begin{aligned}
-\frac{\partial R_t(\hat{x}_t)}{\partial t} = \\
= \min_{u_t} \left\{ \sum_{i=1}^r \left[\frac{\partial R_t(\hat{x}_t)}{\partial \hat{x}_i} M_y \{ f_{it}(x_t, \eta_t, u_t) \} \right] + M_y \{ \varphi_t(x_t, u_t) \} \right\}. \quad (5.85)
\end{aligned}$$

Если первое слагаемое в правой части уравнения (5.85) представить в виде скалярного произведения векторов, то это уравнение принимает форму

$$-\frac{\partial R_t(\hat{x}_t)}{\partial t} = \min_{u_t} \left\{ \left(\frac{\partial R_t(\hat{x}_t)}{\partial \hat{x}_t^T} \cdot M_y \{ f_t(x_t, \eta_t, u_t) \} \right) + M_y \{ \varphi_t(x_t, u_t) \} \right\} \quad (5.86)$$

при граничном условии $R_{t_k}(x_{t_k}) = M_y \{ F(x(t_k)) \}$.

Уравнение Беллмана дает необходимые условия оптимальности и служит для определения оптимального вектора управления путем минимизации правой части уравнения (5.86) в каждый текущий момент времени. Решение этого уравнения связано с техническими трудностями и, как правило, осуществляется численными методами.

Можно показать эквивалентность методов динамического программирования и принципа максимума. Введем непрерывный аналог функционала качества

$$J_t(x_t, u_t) = F(x(t_k)) + \int_t^{t_k} \varphi_\tau(x_\tau, u_\tau) d\tau$$

и функцию будущих потерь

$$R_t(\hat{x}_t) = \min_{u_t} \{ M_y \{ J_t(x_t, u_t) \} \}, \quad t > t_0. \quad (5.87)$$

Так как

$$R_{t+\Delta t}(\hat{x}_{t+\Delta t}) = \min_{u_t} \{ M_y \{ J_{t+\Delta t}(x_{t+\Delta t}, u_{t+\Delta t}) \} \}, \quad (5.88)$$

то соотношение (5.83) с учетом (5.87) и (5.88) можно представить в виде

$$\min_{u_t} \{ M_y \{ J_{t+\Delta t}(x_{t+\Delta t}, u_{t+\Delta t}) + \varphi_t(x_t, u_t) \Delta t - J_t(x_t, u_t) \} \} = 0. \quad (5.89)$$

При условии существования производной функционала J_t по x и t запишем разложение по малому параметру Δt в линейном приближении

$$J_{t+\Delta t}(x_{t+\Delta t}, u_{t+\Delta t}) = J_t(x_t, u_t) + \frac{\partial J_t(x_t, u_t)}{\partial t} \Delta t + \sum_{i=1}^r \frac{\partial J_t(x_t, u_t)}{\partial x_{it}} f_{it}(x_t, \eta_t, u_t) \Delta t. \quad (5.90)$$

Подставляя выражение (5.90) в соотношение (5.89) и сокращая на Δt , получаем

$$\min_{u_t} \left\{ M_y \left\{ \frac{\partial J_t(x_t, u_t)}{\partial t} + \sum_{i=1}^r \frac{\partial J_t(x_t, u_t)}{\partial x_{it}} f_{it}(x_t, \eta_t, u_t) + \varphi_t(x_t, u_t) \right\} \right\} = 0. \quad (5.91)$$

Введем вектор вспомогательных переменных с компонентами

$$\psi_0 = -\frac{\partial J_t(x_t, u_t)}{\partial t}, \quad \psi_i = -\frac{\partial J_t(x_t, u_t)}{\partial x_{it}}, \quad \psi_{r+1} = -1, \quad i = \overline{1, r}.$$

После подстановки этих компонент в соотношение (5.91) с учетом того, что его минимальное значение при изменении знака на обратный равно максимальному значению, получаем

$$\begin{aligned} \max_{u_t} \left\{ M_y \left\{ \psi_0 + \sum_{i=1}^r \psi_i f_i + \psi_{r+1} \varphi_t(x_t, u_t) \right\} \right\} &= \max_{u_t} \left\{ M_y \left\{ \sum_{i=0}^{r+1} \varphi_i f_i \right\} \right\} = \\ &= \max_{u_t} \{ M_y \{ H_\psi(x_t, \psi, u_t, t) \} \} = 0 \end{aligned} \quad (5.92)$$

где $f_0=1$, $\varphi_t(x_t, u_t)=f_{r+1}$, H_ψ - функция Гамильтона.

На основании полученного выражения (5.92) можно сделать вывод, что при оптимальном управлении $u(t)$ максимизируется апостериорное среднее гамильтониана. При этом в любой точке оптимальной траектории объекта это максимальное значение равно нулю.

5.3.4. МЕТОДЫ ПРИБЛИЖЕННОГО СИНТЕЗА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

5.3.4.1. ПОСТАНОВКА ЗАДАЧИ

Результаты теории оптимального синтеза управления приближенно распространяются на нелинейные уравнения состояния динамической системы (4.55) и наблюдаемого процесса (4.56). Рассматриваемые в этом разделе приближенные решения задач оптимального синтеза достигаются применением статистической линеаризации векторных нелинейностей $f_n(x_n)$ и $S_n(x_n)$ в дискретном времени (4.147) и (4.148). Представим многомерную нф и векторно-матричное линеаризованное уравнение в конечных разностях в следующем виде

$$f_n(x_n) = \frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*} x_n + f'_{0n}, \quad (5.93)$$

$$x_{n+1} = \Phi_n x_n + \Delta t f'_{0n} + V_n u_n + \Gamma_n \eta_n, \quad x_0 = x(0),$$

где $f'_{0n} = f_{0n}(x_n^*, D_n) - \frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*} x_n^*$; $\Phi_n = I + \Delta t \frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*}$; I - единичная матрица; $\frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*}$ - матрица Якоби, (ij) компонента которой $\partial f_{0in} / \partial x_{jn}$ (i - номер строки, j - номер столбца); $V_n = \Delta t V(t_n)$, $\Gamma_n = \Delta t \Gamma(t_n)$.

Вектор измерения определяется из уравнения наблюдения

$$y_n = S_n(x_n) + \lambda_n.$$

Соответствующие функции, входящие в функционал качества квадратичной формы имеют вид (5.38) и (5.48). На вектор управления ограничения не накладываются.^{*)}

^{*)} Приближенные методы синтеза управления также могут быть распространены на рассмотренный в разделе 4.10 наблюдаемый сигнал типа случайный точечный процесс, зависящий от фазовых координат объекта [79].

5.3.4.2. УПРАВЛЕНИЕ ПРИ ТЕРМИНАЛЬНОМ КРИТЕРИИ ОПТИМИЗАЦИИ

Для рассматриваемой задачи апостериорное среднее функционала качества равно (5.39) $\hat{J} = M\{x_{r+1,k+1} | y_1^k\}$, где дополнительная координата x_{r+1} определяется из рекуррентного уравнения (5.37). Стохастический гамильтониан в дискретном виде имеет вид

$$H_\psi = \psi_{n+1}^T f_n(x_n, u_n, \eta_n) = \sum_{i=1}^{r+1} \psi_{i,n+1} f_{in}, \quad (5.94)$$

где вектор-функция $f_n(x_n, u_n, \eta_n)$ состоит из компонент - правых частей уравнения состояния объекта (4.93) и функции (5.38).

На основании соотношений (5.93), (5.48), (5.94) стохастический гамильтониан с учетом программируемого вектора состояния $x_{n+1}^{np} = \Phi_n x_n^{np} + \Delta t f'_{0n}$ принимает форму

$$\begin{aligned} H_\psi(x_n, \psi_{n+1}, u_n, \eta_n t_n) &= (x_n - x_n^{np})^T L_n (x_n - x_n^{np}) + u_n^T W_n u_n + \\ &+ 2[(x_n - x_n^{np})^T K_n (x_{n+1} - x_{n+1}^{np} - x_n + x_n^{np})] + \psi_{n+1}^T x_{n+1} + F(x_0) = \\ &= (x_n - x_n^{np})^T L_n (x_n - x_n^{np}) + u_n^T W_n u_n + 2(x_n - x_n^{np})^T K_n [\Phi_n x_n + \\ &+ V_n u_n + \Gamma_n \eta_n - \Phi_n x_n^{np} - x_n + x_n^{np}] + \psi_{n+1}^T (\Phi_n x_n + \Delta t f'_{0n} + \\ &+ V_n u_n + \Gamma_n \eta_n) + F(x_0). \end{aligned} \quad (5.95)$$

Оптимальное управление отыскивается в открытой области из условия (5.46)

$$M \left\{ \frac{\partial H_\psi}{\partial u_n} \middle| y_1^n \right\} = M_y \{ 2W_n u_n + 2V_{n+1}^T K_n (x_n - x_n^{np}) + \psi_{n+1}^T V_n \} = 0.$$

После операции усреднения относительно вектора измерения y получаем векторно-матричное уравнение, из которого определяют вектор оптимального управления

$$u_n^{onm} = -W_n^{-1} V_n^T \left[K_n (\hat{x}_n - x_n^{np}) + \frac{1}{2} \hat{\psi}_{n+1} \right]. \quad (5.96)$$

Оценка вектора вспомогательной переменной $\hat{\psi}$ определяется из выражения (5.45) с учетом $\Phi_n = I + \Delta t \partial f_{0n}(x_n^*, D_n) / \partial x_n^*$ и вида стохастического гамильтониана (5.95)

$$\begin{aligned} \hat{\psi}_n = & 2L_n(\hat{x}_n - x_n^{np}) + 2K_n(\hat{x}_{n+1} - x_{n+1}^{np}) - 2K_n(\hat{x}_n - x_n^{np}) + \\ & + \hat{\psi}_{n+1} + \Delta t \left(\frac{\partial f_{0n}}{\partial x^*} \right)^T \hat{\psi}_{n+1} + 2\Delta t \left(\frac{\partial f_{0n}}{\partial x^*} \right)^T K_n(\hat{x}_n - x_n^{np}) + O_1(\Delta t). \end{aligned} \quad (5.97)$$

Сравнивая выражения (5.97) и (5.53) убеждаемся, что структурно они одинаковы и отличие их состоит в присутствии матрицы $\partial f_{0n}(x_n^*, D_n) / \partial x_n^*$ вместо F , а также в наборе параметров слагаемых с более высоким порядком малости относительно Δt . Соотношения для оптимального вектора управления (5.51) и (5.96) точно совпадают. Поэтому процедура дальнейших преобразований аналогична ранее проведенной для линейной задачи. Она состоит в введении вектора \hat{z} , отыскании его в виде линейной формы $\hat{z}_n = P_n \delta \hat{x}_n$, где $\delta \hat{x}_n = \hat{x}_n - x_n^{np}$ - вектор ошибки, P_n - подлежащая определению неизвестная матрица. В результате дальнейших преобразований приходим к нелинейному матричному уравнению в дискретном времени относительно этой матрицы

$$\begin{aligned} P_n = & P_{n+1} + \Delta t \left(\frac{\partial f_{0n}}{\partial x^*} \right)^T P_{n+1} + \Delta t P_{n+1} \left(\frac{\partial f_{0n}}{\partial x^*} \right) - \\ & - P_{n+1} V_n W_n^{-1} V_n^T P_{n+1} + L_n, \end{aligned} \quad (5.98)$$

которое позволяет определить вектор управления

$$u_n^{opt} = W_n^{-1} V_n^T P_{n+1} (x_n^* - x_n^{np}).$$

На основании полученных данных, а также уравнений фильтрации (4.149) и (4.150) можно прийти к алгоритму квазиоптимального синтеза терминального управления динамической системы в дискретном времени, описываемому системой уравнений

$$\begin{aligned}
x_{n+1}^* &= \tilde{x}_{n+1} + \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T \left[\left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \\
&\quad \times [y_{n+1} - S_{0,n+1}], \quad x_0^* = x(0), \\
D_{n+1} &= \tilde{D}_{n+1} - \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T \left[\left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \\
&\quad \times \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1}, \quad D_0 = D(0), \\
P_n &= P_{n+1} + \Delta t \left(\frac{\partial f_{0n}}{\partial x_n^*} \right)^T P_{n+1} + \Delta t P_{n+1} \left(\frac{\partial f_{0n}}{\partial x_n^*} \right) - P_{n+1} V_n W_n^{-1} V_n^T P_{n+1} + \\
&\quad + L_n, \quad P_{k+1} = K_{k+1}, \\
\tilde{x}_{n+1} &= x_n^* + \Delta t f_{0n} - V_n W_n^{-1} V_n^T P_{n+1} (x_n^* - x_n^{np}), \\
\tilde{D}_{n+1} &= D_n + \Delta t \left(\frac{\partial f_{0n}}{\partial x_n^*} \right) D_n + \Delta t D_n \left(\frac{\partial f_{0n}}{\partial x_n^*} \right)^T + \Gamma_n Q_{\eta n} \Gamma_n^T,
\end{aligned}$$

где $f_{0n} = f_{0n}(x_n^*, D_n)$, $S_{0,n+1} = S_{0,n+1}(\tilde{x}, \tilde{D})$.

Хотя формально задача решена, необходимо обратить внимание на сложность ее решения, вытекающую из-за невозможности применения теоремы разделения. Действительно, определяющая управление матрица P из уравнения (5.98) зависит от матрицы статистических коэффициентов усиления $\partial f_0(x^*, D)/\partial x^*$, которая, в свою очередь, определяется после статистической линеаризации нелинейных функций от оценок фазовых координат. В результате для вычисления квазиоптимального управления необходимо совместно решать приведенную выше систему уравнений.

5.3.4.3. УПРАВЛЕНИЕ ПРИ ЛОКАЛЬНОМ КРИТЕРИИ ОПТИМИЗАЦИИ

Необходимые исходные данные для определения квазиоптимального управления при локальном критерии в дискретном времени сформулированы в разделе 5.3.2.2. Отличие от линейной задачи состоит в том, что при формировании стохастического гамильтониана в функционале качества

квадратичной формы функция $f_{r+1,n}$ зависит от компонент матрицы статистических коэффициентов усиления

$$H_\psi = f_{r+1,n}(x_n, u_n, \eta_n) = (x_n - x_n^{np})^T L_n (x_n - x_n^{np}) + u_n^T W_n u_n + \\ + 2(x_n - x_n^{np})^T K_n [\Phi_n x_n + V_n u_n + \Gamma_n \eta_n - \Phi_n x_n^{np} - \\ - x_n + x_n^{np}] + F(x_0), \quad (5.99)$$

$$\text{где } \Phi_n = I + \Delta t \frac{\partial f_{0n}(x_n^*, D_n)}{\partial x_n^*}.$$

В предположении, что управление принадлежит открытой области, после усреднения относительно вектора наблюдения y на основании соотношения (5.99) получаем

$$M \left\{ \frac{\partial H_\psi}{\partial u_n} \middle| y_1^n \right\} = V_n^T K_n (x_n^* - x_n^{np}) + W_n u_n^{onm} = 0.$$

Отсюда вектор оптимального управления

$$u_n^{onm} = -W_n^{-1} V_n^T K_n (x_n^* - x_n^{np}). \quad (5.100)$$

Для устойчивого объекта аналогично линейной задаче вводится дополнительное уравнение связи для матриц K_n и L_n

$$\left(\frac{\partial f_{0n}}{\partial x_n^*} \right)^T K_n + K_n \left(\frac{\partial f_{0n}}{\partial x_n^*} \right) = -L_n. \quad (5.101)$$

Определенная из этого уравнения матрица K_n после подстановки в выражение (5.100) позволяет получить оптимальный вектор управления и тем самым решить задачу квазиоптимального управления для нелинейных объекта и измерителя. Как следует из выражения (5.100) управление u формально является детерминированной линейной функцией вектора оценки x^* состояния объекта. Однако, компоненты матриц, входящих в уравнение (5.101), зависят от вектора этих оценок. Вследствие этого указанное уравнение связано с вектором оценки, и следовательно, теорема разделения несправедлива. В итоге для определения оптимального управления необходимо совместно решать систему уравнений

$$\begin{aligned}
x_{n+1}^* &= \tilde{x}_{n+1} + \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T \left[\left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \\
&\quad \times [y_{n+1} - S_{0,n+1}], \quad x_0^* = x(0), \\
D_{n+1} &= \tilde{D}_{n+1} - \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T \left[\left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1} \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right)^T + Q_{\lambda,n+1} \right]^{-1} \times \\
&\quad \times \left(\frac{\partial S_{0,n+1}}{\partial \tilde{x}_{n+1}} \right) \tilde{D}_{n+1}, \quad D_0 = D(0), \\
\frac{\partial f_{0n}^T}{\partial x_n^*} K_n + K_n \frac{\partial f_{0n}}{\partial x_n^*} &= -L_n, \\
\tilde{x}_{n+1} &= x_n^* + \Delta t f_{0n} - V_n W_n^{-1} V_n^T K_n (x_n^* - x_n^{np}), \\
\tilde{D}_{n+1} &= D_n + \Delta t \left(\frac{\partial f_{0n}}{\partial x_n^*} \right) D_n + \Delta t D_n \left(\frac{\partial f_{0n}}{\partial x_n^*} \right)^T + \Gamma_n Q_{\eta n} \Gamma_n^T, \\
\text{где } f_{0n} &= f_{0n}(x_n^*, D_n), \quad S_{0,n+1} = S_{0,n+1}(\tilde{x}_{n+1}, D_{n+1}).
\end{aligned}$$

ПРИЛОЖЕНИЕ

1. Дельта-функция

К понятию дельта-функция можно прийти путем предельного перехода в обычных функциях. Одной из таких функций является единичный прямоугольный импульс в начале координат с основанием равным $2l$

$$\delta_l(x) = \begin{cases} 1/2l, & |x| < l, \\ 0, & |x| > l. \end{cases}$$

Для этой функции и любого $\varepsilon > l$ можно записать

$$\int_{-\varepsilon}^{\varepsilon} \delta_l(x) dx = 1, \quad \int_0^{\varepsilon} \delta_l(x) dx = \int_{-\varepsilon}^0 \delta_l(x) dx = 1/2, \quad (\text{П1})$$

$$\int_{x_0-\varepsilon}^{x_0+\varepsilon} \delta_l(x-x_0)dx = 1, \quad \int_{x_0}^{x_0+\varepsilon} \delta_l(x-x_0)dx = \int_{x_0-\varepsilon}^{x_0} \delta_l(x-x_0)dx = 1/2. \quad (\text{П2})$$

Очевидно, при предельном переходе $l \rightarrow 0$ функция $\delta_l(x)$ не имеет предела в обычном смысле. Однако интегралы (П1) и (П2) сохраняют свои значения. Рассуждая формально, приходим к определению дельта-функции $\delta(x-x_0) = \lim_{l \rightarrow 0} \delta_l(x-x_0)$

$$\delta(x-x_0) = \begin{cases} \infty, & x = x_0, \\ 0, & x \neq x_0, \end{cases} \quad \int_{x_0-\varepsilon}^{x_0+\varepsilon} \delta(x-x_0)dx = 1 \text{ при любом } \varepsilon > 0.$$

Причем площадь «всплеска» в бесконечность дельта-функции в начала координат или особой точке x_0 равна единице. Дельта-функция не существует как обычная функция. Она относится к классу обобщенных функций.

Путем предельного перехода можно получить следующие соотношения

$$\int_a^b f(x)\delta(x_0-x)dx = \int_a^b f(x)\delta(x-x_0)dx = f(x_0), \quad x_0 \in (a,b), \quad (\text{П3})$$

$$\int_a^b f(x)\delta(a-x)dx = \int_a^b f(x)\delta(x-a)dx = \frac{1}{2}f(a), \quad (\text{П4})$$

$$\int_a^b f(x)\delta(b-x)dx = \int_a^b f(x)\delta(x-b)dx = \frac{1}{2}f(b), \quad (\text{П5})$$

отражающие фильтрующие свойства дельта-функции.

При $x_0 > b$ и $x_0 < a$ интеграл (П3) равен нулю.

Если x_0 является точкой разрыва первого рода функции $f(x)$, то

$$\int_a^b f(x)\delta(x-x_0)dx = \frac{1}{2}[f(x_0^+) + f(x_0^-)], \quad a < x_0 < b,$$

где $f(x_0^+)$ и $f(x_0^-)$ - значения функции $f(x)$ справа и слева от точки разрыва.

Из соотношений (П1) и (П2) следует, что интеграл от дельта-функции в пределах от $-\infty$ до u равен единичной ступенчатой функции

$$\int_{-\infty}^u \delta(x) dx = 1(u) = \begin{cases} 0, & u < 0, \\ 1/2, & u = 0, \\ 1, & u > 0. \end{cases}$$

Таким образом, дельта-функция представляет собой производную единичной ступенчатой функции. Применяя операцию интегрирования по частям, можно убедиться, что для любой функции, имеющей непрерывные производные до n -го порядка включительно в точке x_0 , имеет место

$$\int_{x_0 - \varepsilon}^{x_0 + \varepsilon} f(x) \delta^{(n)}(x - x_0) dx = (-1)^n f^{(n)}(x_0).$$

Выразим дельта-функцию интегралом Фурье. Используя для непрерывной функции известное из теории интеграла Фурье соотношение

$$f(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} f(x) \exp\{-j\omega(x-u)\} dx,$$

можно записать, изменив порядок интегрирования,

$$f(u) = \int_{-\infty}^{\infty} f(x) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{-j\omega(x-u)\} d\omega \right] dx.$$

Сравнив полученное выражение с (П3), получаем при $a=-\infty$, $b=\infty$, что $\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\{j\omega x\} d\omega$.

Кроме того, из прямого преобразования Фурье имеем $\int_{-\infty}^{\infty} \delta(x) \exp\{-j\omega x\} dx = 1$, т.е. спектр дельта-функции является равномерным на всех частотах с интенсивностью равной единице.

2. Сведения из алгебры матриц

Прямоугольная таблица mn чисел, составленная из m строк и n столбцов называется прямоугольной матрицей размера $(m \times n)$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{nm} \end{pmatrix} = \| a_{ij} \|, \quad i = \overline{1, m}, \quad j = \overline{1, n}.$$

Элементы a_{ij} называются компонентами матрицы, первый индекс указывает строку, второй - столбец матрицы. Если $m=n$ матрица называется квадратной.

Матрица размера $(m \times 1)$ является вектором-столбцом

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Матрица размера $(1 \times n)$ называется вектором-строкой

$$x = \| x_1 \quad x_2 \quad \cdots \quad x_n \|.$$

Матрицы и векторы, имеющие одну компоненту, называются скалярной величиной. Матрица, у которой все компоненты равны нулю, обозначается 0 . Квадратная матрица, у которой все компоненты, расположенные вне главной диагонали, равны нулю, называется диагональной. Диагональная матрица, у которой все компоненты, стоящие на главной диагонали, равны единице, называется единичной матрицей и обозначается I .

Сумма двух прямоугольных матриц $A = \| a_{ij} \|$ и $B = \| b_{ij} \|$ одинакового размера $(m \times n)$ называется матрица $C = \| c_{ij} \|$ того же размера, компоненты которой равны сумме соответствующих компонент матриц A и B : $c_{ij} = a_{ij} + b_{ij}$, $C = A + B$.

Произведение матрицы A на скалярную величину β называется матрица C , соответствующие компоненты которой равны $c_{ij} = \beta a_{ij}$, $C = \beta A$.

Произведением двух прямоугольных матриц $A=||a_{ij}||$ размера $(m \times n)$ и $B=||b_{ij}||$ размера $(n \times r)$ является матрица $C=||c_{ij}||$ размера $(m \times r)$, компонент которой $c_{ij} = \sum_{q=1}^n a_{iq} b_{qi}$, $i = \overline{1, m}$, $j = \overline{1, r}$,

$C=AB$, т.е. когда число столбцов в первом сомножителе равно числу строк во втором сомножителе.

Операции умножения матриц обладают следующими свойствами: $(AB)C=A(BC)$, $(A+B)C=AC+BC$, $A(B+C)=AB+AC$. Операция умножения не обладает переместительным свойством, т.е. в общем случае $AB \neq BA$. Если $AB=BA$, то матрицы A и B называются перестановочными (коммутативными).

Скалярное произведение (обозначаемое точкой) двух векторов одинаковой размерности можно представить в виде

$$x \cdot y = x^T y = y^T x = y \cdot x.$$

Если скалярное произведение двух ненулевых векторов равно нулю, то эти два вектора ортогональны.

Определителем (детерминантом) квадратной матрицы размера $n \times n$, $n > 1$ называется число

$$\det A = |A| = \sum_{q=1}^n (-1)^{q+1} a_{1q} M_{1q},$$

где M_{1q} - определитель квадратной матрицы размера $(n-1) \times (n-1)$, полученный из A вычеркиванием первой строки и q -го столбца. Квадратная матрица называется особенной (вырожденной), если ее определитель равен нулю. В противном случае ее называют неособенной (невырожденной). Для каждой неособенной матрицы A существует обратная матрица $A^{-1} = ||a_{ij}^{-1}||$, компоненты которой равны

$$a_{ij}^{-1} = \frac{A_{ij}}{|A|}, \text{ где алгебраическое дополнение } A_{ij} = (-1)^{i+j} M_{ij}.$$

Имеют место следующие соотношения для неособенных матриц $(AB)^{-1} = A^{-1} B^{-1}$, $AA^{-1} = A^{-1}A = I$.

Матрица A^T называется транспонированной относительно матрицы A , если $a_{ij}^T = a_{ji}$, т.е. строки матрицы A становятся столбцами матрицы A^T . Для операции транспонирования

выполняются следующие соотношения $(A+B)^T=A^T+B^T$, $(AB)^T=B^T A^T$, $(A^{-1})^T=(A^T)^{-1}$, $(A^T)^T=A$.

Если $A=A^T$, то матрица называется симметричной.

Для функций от матриц сохраняются свойства скалярных функций в следующих разложениях

$$\exp\{A\} = \sum_{q=0}^{\infty} \frac{A^q}{q!}, \quad \ln\{A\} = \sum_{q=1}^{\infty} \frac{(-1)^{q-1}}{q!} (A-I)^q, \quad (1-A)^{-1} = \sum_{q=0}^{\infty} A^q,$$

если собственные числа матрицы A лежат внутри круга сходимости.

Следом квадратной матрицы A называется сумма ее компонент, расположенных на главной диагонали

$$SpA = \sum_i a_{ii}.$$

Для следа матриц имеют место следующие соотношения

$$SpA + SpB = Sp(A+B), \quad Sp(BC) = Sp(CB).$$

Операция дифференцирования вектора x размера $(m \times 1)$ или матрицы A размера $(m \times n)$ по скалярному аргументу t определяется выражениями

$$\frac{dx}{dt} = \left\| \frac{dx_i}{dt} \right\|, \quad i = \overline{1, m},$$

$$\frac{dA}{dt} = \left\| \frac{da_{ij}}{dt} \right\|, \quad i = \overline{1, m}, \quad j = \overline{1, n}.$$

Производные от матричных функций скалярного аргумента t имеют вид

$$\frac{d(AB)}{dt} = \frac{dA}{dt} B + A \frac{dB}{dt}, \quad \frac{d(A)^{-1}}{dt} = -A^{-1} \left(\frac{dA}{dt} \right) A^{-1}.$$

Операция интегрирования матричной функции скалярного аргумента сводится к соотношению

$$\int_0^t A_\tau d\tau = \left\| \int_0^t a_{ij}(\tau) d\tau \right\|, \quad i = \overline{1, m}, \quad j = \overline{1, n}.$$

Производная скалярной функции $f(x)$ по вектору x размера $(n \times 1)$ есть вектор-строка

$$\frac{df(x)}{dx} = \left\| \frac{\partial f(x)}{\partial x_j} \right\|, \quad j = \overline{1, n}.$$

Производная вектор-функции $y(x)$ размера $(m \times 1)$ по вектору x размера $n \times 1$ называется матрицей Якоби

$$\frac{dy}{dx} = \left\| \frac{\partial y_i}{\partial x_j} \right\|, \quad i = \overline{1, m}, \quad j = \overline{1, n}.$$

Производная скалярной функции $f(A)$ по матрице A определяется формулой

$$\frac{d}{dA} f(A) = \left\| \frac{\partial}{\partial a_{ij}} f(A) \right\|, \quad i = \overline{1, m}, \quad j = \overline{1, n}.$$

Если $y = Ax$, где y и x - векторы, A - постоянная матрица, то

$$\frac{dy}{dx} = A.$$

Если $y = x^T C x$, C - симметричная постоянная матрица, то

$$\frac{dy}{dx} = C^T x + Cx = 2Cx.$$

3. Связь вариационного исчисления и принципа максимума (минимума) с понятиями теоретической механики

В теории оптимального управления множителям Лангража $\lambda(t)$ вариационного исчисления отвечают в методе принципа максимума вспомогательная функция $\psi(t) = -\lambda(t)$ (в принципе минимума $\psi(t) = \lambda(t)$), а в теоретической механике обобщенный импульс $p(t)$. С

помощью векторов обобщенных импульсов $p(t)$ и координат $q(t)$ можно описать поступательное и вращательное движение сложных динамических систем. Рассматриваются консервативные системы без рассеяния энергии под действием стационарных потенциальных сил в декартовой системе координат: $x=x(q)$, $p=p(\dot{x})=m\dot{x}$, m - диагональная матрица постоянных масс системы^{*)}. Введем важную при определении экстремумов функционалов качества функцию Лагранжа, как разность кинетической и потенциальной энергии системы

$$L(x, \dot{x}) = T(x, \dot{x}) - U(x).$$

Потенциальная энергия зависит только от вектора координат, а кинетическая энергия может зависеть как от вектора координат, так и от вектора скоростей координат. По определению обобщенным импульсом называют частную производную кинетической энергии по скорости координаты, которую можно также записать через функцию Лагранжа

$$p = \frac{\partial T(x, \dot{x})}{\partial \dot{x}} = \frac{\partial L(x, \dot{x})}{\partial \dot{x}}. \quad (\text{П6})$$

Если кинетическая энергия системы не зависит от координат компонент вектора x , то $T(\dot{x}) = \frac{1}{2} \dot{x}^T m \dot{x}$.

Перейдем к рассмотрению одного из основных соотношений вариационного исчисления - уравнению Эйлера. Функция Лагранжа удовлетворяет уравнению Эйлера

$$\frac{d}{dt} \left[\frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right] - \frac{\partial L(x, \dot{x})}{\partial x} = 0, \quad (\text{П7})$$

которое дает необходимые условия экстремума функционала качества

$$J_t = \int_0^t L(x, \dot{x}) dt = \int_0^t [T(x, \dot{x}) - U(x)] dt.$$

^{*)} Системы с рассеянием энергии, например, из-за трения, сопротивления воздуха и т.д., называется диссипативными системами.

Этому необходимому условию соответствует равенство нулю первой вариации функционала (принцип Гамильтона)

$$\delta J_t = \delta \int_0^t L(x, \dot{x}) dt = 0.$$

Вторая вариация позволяет различать минимумы и максимумы: для минимума функционала необходимо $\delta^2 J_t > 0$, для максимума $\delta^2 J_t < 0$.

Всякому функционалу соответствует свое дифференциальное уравнение Эйлера и наоборот, дифференциальное уравнение равновесия сил, действующих на систему, можно рассматривать как уравнение Эйлера для некоторого функционала. И действительное поведение системы будет доставлять этому функционалу экстремальное значение. Таким образом, описание законов природы на языке вариационного исчисления полностью эквивалентно их описанию на языке дифференциальных уравнений.

Рассмотрим движение материального тела массой m в однородном поле тяготения по координате x . Функция Лагранжа имеет вид

$$L(x, \dot{x}) = T(\dot{x}) - U(x) = \frac{1}{2} m(\dot{x})^2 - mgx, \quad (\text{П8})$$

где g - ускорение силы тяжести.

После подстановки функции (П8) в уравнение (П7) и вычислений получим уравнение Эйлера в форме второго закона Ньютона

$$P = mg.$$

В качестве еще одного примера рассмотрим движение в однородном поле тяготения маятника, представляющего собой подвешанное на жесткой невесомой проволоке длиной l материальное тело массой m . Рассматривается линейное приближение при малых угловых отклонениях Θ . Кинетическая и потенциальная энергии соответственно равны

$$T = \frac{1}{2} ml^2 [\dot{\Theta}]^2, \quad U = mg(1 - \cos \Theta)l.$$

Полная энергия

$$E = T + U = \frac{1}{2} ml^2 [\dot{\Theta}]^2 + mg(1 - \cos \Theta)l.$$

Для консервативной системы $E = \text{const}$ и дифференциальное уравнение движения маятника можно получить из условия $\partial E / \partial t = 0$.

Отсюда

$$\frac{d^2 \Theta}{dt^2} + \frac{g}{l} \sin \Theta = 0. \quad (\text{П9})$$

Получим это уравнение с использованием методов вариационного исчисления. Функция Лагранжа равна

$$L = T - U = \frac{1}{2} ml^2 [\dot{\Theta}]^2 - mg(1 - \cos \Theta)l. \quad (\text{П10})$$

После подстановки функции (П10) в уравнение (П7) и вычислений получаем уравнение Эйлера в форме дифференциального уравнения известного вида (П9).

Функция Гамильтона (гамильтониан) в механике по определению является суммой кинетической и потенциальной энергией

$$H_p = T + U = -(T - U) + 2T = -L + \sum_{i=1}^r \dot{x}_i p_i = -L + (\dot{x})^T p. \quad (\text{П11})$$

Выражение (П11) совпадает с соотношением (5.14) с точностью до обозначения вспомогательной переменной, если функционал J принять за дополнительную координату x_{r+1} , а функцию Лагранжа обозначить производной по этой координате \dot{x}_{r+1} , $p_{r+1} = -1$ ($\psi_{r+1} = -1$).

На основании полученного выражения (П11) определим дифференциал гамильтониана

$$dH_p = -\frac{\partial L}{\partial t} dt - \sum_{i=1}^r \frac{\partial L}{\partial x_i} dx_i - \sum_{i=1}^r \frac{\partial L}{\partial \dot{x}_i} d\dot{x}_i + \sum_{i=1}^r \dot{x}_i dp_i + \sum_{i=1}^r p_i d\dot{x}_i. \quad (\text{П12})$$

Согласно соотношению (П6) третий и пятый члены (П12) сокращаются. Выпишем полный дифференциал гамильтониана

$$dH_p = \frac{\partial H_p}{\partial t} dt + \sum_{i=1}^r \frac{\partial H_p}{\partial x_i} dx_i + \sum_{i=1}^r \frac{\partial H_p}{\partial p_i} dp_i. \quad (\text{П13})$$

Первый член (П13), как будет показано ниже, равен нулю. В результате сравнения остальных членов соотношений (П12) и (П13) получаем $\frac{\partial H_p}{\partial x_i} = -\frac{\partial L}{\partial x_i} = -\frac{\partial}{\partial t} \left[\frac{\partial L}{\partial \dot{x}} \right]$. Но согласно (П6) $\frac{\partial L}{\partial \dot{x}} = p$, поэтому

$$\frac{\partial H_p}{\partial x_i} = -\dot{p}_i, \quad (\text{П14})$$

а также

$$\frac{\partial H_p}{\partial p_i} = \dot{x}_i. \quad (\text{П15})$$

Уравнения (П14) и (П15) называются каноническими уравнениями Гамильтона. Из этих уравнений следует

$$\frac{\partial H_p}{\partial t} = \sum_{i=1}^r \frac{\partial H_p}{\partial x_i} \frac{\partial x_i}{\partial t} + \sum_{i=1}^r \frac{\partial H_p}{\partial p_i} \frac{\partial p_i}{\partial t} = -\sum_{i=1}^r \dot{p}_i \dot{x}_i + \sum_{i=1}^r \dot{x}_i \dot{p}_i = 0.$$

т.е. во время движения консервативной системы функция Гамильтона явно не зависит от времени и полная механическая энергия сохраняет свое значение.

Из физических соображений следует, что оптимальное управление в динамической системе может быть получено, если обеспечить максимум мощности или энергии (гамильтониана) в ней с учетом ограничений на управляющие сигналы.

СПИСОК ЛИТЕРАТУРЫ

1. Левин Б.Р. Теоретические основы статистической радиотехники. Кн.1. М.: Сов.радио, 1974.
2. Стратонович Р.Л. Избранные вопросы теории флуктуаций в радиотехнике. М.: Сов.радио, 1961.
3. Пугачев В.С. Теория случайных функций и ее применение к задачам автоматического управления. М.: Физматгиз, 1962.
4. Вентцель А.Д. Курс теории случайных процессов. М.: Наука, 1975.
5. Сосулин Ю.Г. Теория обнаружения и оценивания стохастических сигналов. М.: Сов.радио, 1978.
6. Тихонов В.И., Миронов М.А. Марковские процессы. М.: Сов.радио, 1977.
7. Дынкин Е.Б. Марковские процессы. М.: Физметгиз, 1963.
8. Казаков В.А. Введение в теорию марковских процессов и некоторые радиотехнические задачи. М.: Сов.радио, 1973.
9. Дуб Дж.Л. Вероятностные процессы. М.: Изд-во иностр.лит., 1956.
10. Никитин Н.Н., Разевиг В.Д. Методы цифрового моделирования стохастических дифференциальных уравнений и оценка их погрешностей // Журнал вычислительной математики и матем.физики. 1978. Т18, № 1.
11. Большаков И.А., Ракошиц В.С. Прикладная теория случайных потоков. М.: Сов.радио, 1978.
12. Гнеденко Б.Ф., Колмогоров А.Н. Предельные распределения для сумм независимых случайных величин. М.: Гостехиздат, 1949.
13. Городецкий А.Я. Статистический анализ и синтез фотонных систем. СПб.: Изд-во СПбГТУ, 1996.
14. Mandelbrot B.B. The fractal Geometry of Nature. – Freeman, New York, 1983.
15. Фракталы в физике /пер. с англ.под ред. Л.Пьетронеро, Э.Тозатти. – М.: Мир, 1988.
16. Федер Енс. Фракталы/пер.с англ. М.: Мир, 1991.
17. Gefen Y., Ahrony A., Alexander S. Anomalous diffusion on percolatung clisters // Phys.Rev.Lett. 1983
18. Homsy G.M. Viscous findering in porous media // Ann.Rev. Fluid Mech., 19, 1987.
19. Бибермак Л.М., Воробьев В.С., Якубов И.Т. Кинетика неравновесной низкотемпературной плазмы. М.: Наука, 1982.
20. Городецкий А.Я., Заборовский В.С. Информатика. Фрактальные процессы в компьютерных сетях. СПб.: Изд-во СПбГТУ, 2000.
21. Mandelbrot B.B., Van Ness J.W.Fractional Brownian motions, fractional noises and applications.// STAM Review. 10, 1968.
22. Барбашин Е.А. Функции Ляпунова. М.: Наука, 1970.
23. Малкин И.Г. Теория устойчивости движения. М.: Наука, 1966.

24. Математические основы теории автоматического регулирования. Под ред.Чемоданова Б.К. М.: Высшая школа, Т1, 1977.
25. Ван Трис Г. Синтез оптимальных нелинейных систем управления. М.: Мир, 1964.
26. Пугачев В.С., Казаков И.Е., Евланов Л.Г. Основы статистической теории автоматических систем. М.: Машиностроение, 1974.
27. Свешников А.А. Прикладные методы теории случайных функций. М.: Наука, 1969.
28. Первозванский А.А. Случайные процессы в нелинейных автоматических системах. М.: Гос.изд-во физ.-мат.лит., 1962.
29. Лившиц И.А., Пугачев В.И. Вероятностный анализ систем автоматического управления. Кн.2. М.: Сов.радио, 1963.
30. Казаков И.Е. Статистическая теория систем управления в пространстве состояний. М.: Наука, 1975.
31. Коловский М.З. Об оценке точности решений, получаемых методом статистической линеаризации. // Автоматика и телемеханика, 10, 1966.
32. Bunke H. Statistische Linearisierung // Zeitschr.angew. Math.und Mech. 2, 1972.
33. Стратонович Р.Л., Сосулин Ю.Г. Оптимальное обнаружение марковского процесса в шуме.// Техническая кибернетика. № 6, 1964.
34. Сосулин Ю.Г. Оптимальное обнаружение марковских сигналов на фоне марковских помех при дискретном времени // Техническая кибернетика // Техническая кибернетика. № 4, 1966.
35. Стратонович Р.Л., Сосулин Ю.Г. Оптимальный прием сигналов на фоне негауссовых помех. // Радиотехника и электроника. Т11, № 4, 1966.
36. Городецкий А.Я. Структура оптимальных фильтров полезного сигнала, передаваемого случайным импульсным потокам // Радиотехника и электроника. Т21, № 4, 1976.
37. Городецкий А.Я. Восстановление марковского сообщения из дискретной случайной последовательности // Техническая кибернетика. № 5, 1977.
38. Городецкий А.Я. Субоптимальный алгоритм оценивания в задачах фотонной техники.// Робототехника и техническая кибернетика. Сб.тр. СПбГТУ, 1993.
39. Левин Б.Р. Теоретические основы статистической радиотехники. Кн.2. М.: Сов.радио, 1975.
40. Тихонов В.И. Статистическая радиотехника. М.: Сов.радио, 1976.
41. Колмогоров А.И. Интерполирование и экстраполирование стационарных случайных последовательностей. // Изв.АН СССР. Сер.мат.Т5, № 1, 1941.
42. Wiener N. The Extrapolation, Interpolation and Smoothing of Stationary Tume Series. Yohn. Wiley and Sons. Inc. New York, 1949.
43. Kalman R.E., Busy R. New Results in Limar Filtering and Prediction Theory.// ASMEY. Basic Eng, V83,-1961.
44. Стратонович Р.Л. К теории оптимальной нелинейной фильтрации // Теория вероятностей и ее применение. 2, 1959.

45. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. М.: Радио и связь, 1991.
46. Тихонов В.И., Кульман Н.К. Нелинейная фильтрация и квазикогерентный прием. М.: Сов.радио, 1975.
47. Медич Дж. Статистически оптимальные линейные оценки и управление / пер. с англ. М.: Энергия, 1973.
48. Сосулин Ю.Г. Интерполяция и экстраполяция марковских сигналов при наличии помех. // Изв.АН СССР. Техн.кибернетика № 5, 1969.
49. Хелстром К. Квантовая теория проверки гипотез и оценивания. М.: Мир, 1979.
50. Клаудер Дж., Сударшан Э. Основы квантовой оптики. М.: Мир, 1970.
51. Городецкий А.Я. К квантовой теории оценивания марковского сообщения. // Проблемы передачи информации № 2, 1983.
52. Райбман Н.С., Чадеев В.М. Адаптивные модели в системах управления. М.: Сов.радио, 1966.
53. Ли Р. Оптимальные оценки, определение характеристик, управление. М.: Наука, 1966.
54. Райбман Н.С. Что такое идентификация. М.: Наука, 1970.
55. Norros I. A storage model with self-similar input // Queueing System. v.9, 1994.
56. Leland W.E., Taggu M.S., Willinger W. and Wilson D.V. On the self-similar nature of Internet traffic (Extended Versuon). // IEEE/ACM Transations on Networking. v.2, № 2, 1994.
57. Ryn B. and Lowen S Point process models for self-semilar Network Traffic, with applications.// Stochastic Models. № 14. 1998.
58. Qiong Li, David L., Mills. On the long-range dependence of packet round-trip delays in Internet.// Processings of IEEE ICC 98. v2, 1998.
59. Леннинг Д.Х., Беттин Р.Г. Случайные процессы в задачах автоматического управления. М.: ИЛ, 1958.
60. Пугачев В.С. Теория случайных функций и ее применение к задачам автоматического управления. М.: Физматиздат, 1962.
61. Растригин Л.А. Статистические методы поиска. М.: Наука, 1968.
62. Гладков Д.И. Негradientные методы случайного поиска. Труды по статистическим проблемам оптимизации, 1968.
63. Цыпкин Я.З. Адаптация и обучение в автоматических системах. М.: Наука, 1968.
64. Robins H., Monro S. A stochastic approximation method. // Math. Stat. № 22, 1951.
65. Зубов В.И. Теория оптимального управления. М.: Судостроение, 1988.
66. Летов А.М. Динамика полета и управление. М.: Наука, 1969.
67. Болтянский В.Г., Гамкрелидзе Р.В., Понтрягин Л.С. К теории оптимальных процессов // ДАН СССР. Т 110, № 1, 1956.
68. Беллман Р., Калаба Р. Динамическое программирование и современная теория управления. М.: Наука, 1969.

69. Козлов В.Н., Куприянов В.Е., Шашихин В.Н. Вычислительная математика и теория управления. Под ред. Козлова В.Н. СПб.: Изд-во СПбГТУ, 1996.
70. Тихонов А.Н. О методах регуляризации задач оптимального управления.// ДАН СССР. Т 162, № 4, 1965.
71. Солодовников В.В. Статистическая динамика линейных систем автоматического управления. М.: Физматгиз, 1960.
72. Фельдбаум А.А. Основы теории оптимальных автоматических систем. М.: Наука, 1966.
73. Куликовский Р. Оптимальные и адаптивные процессы в системах автоматического регулирования. М.: Наука, 1967.
74. Kalman R.E., Koercke R.W. Optimal synthesis of linear sampling control systems using generalized performance indexes // Trans.ASME, v.80, 1960.
75. Кунцевич В.М. Системы экстремального управления. М.: Физматгиз, 1961.
76. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М.: Наука, 1969.
77. Батков А.М., Щукин Б.А. Оптимальное управление, основанное на методе статистической линеаризации. Современные методы проектирования САУ. Машиностроение, 1967.
78. Kushner H.I. On the stochastic maximum principle: fixed time of control.// Journ. Math. Anal. and Appl. v11, 1965.
79. Городецкий А.Я. Оптимизация управления в системах с фотонным каналом связи.// Робототехника и техническая кибернетика. Сб.тр. СПбГТУ, 1999.

СОДЕРЖАНИЕ

| | |
|--|-----------|
| ПРЕДИСЛОВИЕ | 3 |
| ВВЕДЕНИЕ..... | 6 |
| 1. МЕТОДЫ ОПИСАНИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ..... | 11 |
| 1.1. ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ | 11 |
| 1.2. МОМЕНТНЫЕ И КОРРЕЛЯЦИОННЫЕ ФУНКЦИИ | 16 |
| 1.3. ЭРГОДИЧЕСКИЕ ПРОЦЕССЫ | 22 |
| 1.4. СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ..... | 23 |
| 2. МОДЕЛИ ПРОЦЕССОВ В ИНФОРМАЦИОННЫХ СИСТЕМАХ | 26 |
| 2.1. ГАУССОВСКИЙ ПРОЦЕСС | 26 |
| 2.2. ВИНЕРОВСКИЙ ПРОЦЕСС. БЕЛЫЙ ШУМ | 29 |
| 2.2.1. ОБЩИЕ ОПРЕДЕЛЕНИЯ..... | 29 |
| 2.2.2. СТОХАСТИЧЕСКИЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ И ИНТЕГРАЛЫ | 33 |
| 2.3. МАРКОВСКИЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ..... | 37 |
| 2.3.1. ОПРЕДЕЛЕНИЕ МАРКОВСКИХ ПРОЦЕССОВ | 37 |
| 2.3.2. ДИСКРЕТНЫЙ ПРОЦЕСС С ДИСКРЕТНЫМ ВРЕМЕНЕМ..... | 39 |
| 2.3.3. НЕПРЕРЫВНЫЙ ПРОЦЕСС С ДИСКРЕТНЫМ ВРЕМЕНЕМ | 40 |
| 2.3.4. ДИСКРЕТНЫЙ ПРОЦЕСС С НЕПРЕРЫВНЫМ ВРЕМЕНЕМ | 41 |
| 2.3.5. НЕПРЕРЫВНЫЙ МАРКОВСКИЙ ПРОЦЕСС | 44 |
| 2.3.6. МНОГОМЕРНЫЙ МАРКОВСКИЙ ПРОЦЕСС | 48 |
| 2.3.7. АППРОКСИМАЦИЯ РЕАЛЬНЫХ СИГНАЛОВ МАРКОВСКИМИ ПРОЦЕССАМИ | 49 |
| 2.3.8. ВРЕМЕННАЯ ДИСКРЕТИЗАЦИЯ МАРКОВСКОГО ПРОЦЕССА | 57 |
| 2.4. СЛУЧАЙНЫЕ ТОЧЕЧНЫЕ ПРОЦЕССЫ..... | 61 |
| 2.4.1. ИНТЕГРАЛЬНЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ПОТОКОВ | 61 |
| 2.4.2. ФУНКЦИИ ПЛОТНОСТИ И КОРРЕЛЯЦИИ ПЛОТНОСТИ | 64 |
| 2.4.3. ПУАССОНОВСКИЙ ПРОЦЕСС..... | 69 |
| 2.4.4. МОДУЛИРОВАННЫЕ ПОТОКИ | 71 |
| 2.4.5. МАРКОВСКИЙ ПРОЦЕСС, ПОРОЖДЕННЫЙ ПУАССОНОВСКИМ ПРОЦЕССОМ | 77 |
| 2.5. ФРАКТАЛЬНЫЕ ПРОЦЕССЫ..... | 79 |
| 2.5.1. ФРАКТАЛЬНЫЙ ТОЧЕЧНЫЙ ПРОЦЕСС | 81 |
| 2.5.2. ФРАКТАЛЬНЫЙ ВИНЕРОВСКИЙ ПРОЦЕСС | 91 |
| 3. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ ПРОЦЕССОВ | 94 |
| 3.1. ДИНАМИЧЕСКИЕ СИСТЕМЫ И ИХ ХАРАКТЕРИСТИКИ | 94 |
| 3.2. УСТОЙЧИВОСТЬ ДИНАМИЧЕСКИХ СИСТЕМ..... | 103 |
| 3.3. СЛУЧАЙНЫЕ ПРОЦЕССЫ В ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМАХ | 105 |
| 3.3.1. ОБЩИЕ ОПРЕДЕЛЕНИЯ..... | 105 |
| 3.3.2. ПРЕОБРАЗОВАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ В ЛИНЕЙНЫХ СИСТЕМАХ | 108 |
| 3.3.3. СЛУЧАЙНЫЕ ИМПУЛЬСНЫЕ ПРОЦЕССЫ | 114 |
| 3.3.4. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК..... | 120 |
| 3.4. СЛУЧАЙНЫЕ ПРОЦЕССЫ В НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМАХ | 128 |
| 3.4.1. ОБЗОР МЕТОДОВ | 128 |
| 3.4.2. МЕТОДЫ СТАТИСТИЧЕСКОЙ ЛИНЕАРИЗАЦИИ | 132 |
| 3.4.3. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК..... | 136 |

| | |
|---|------------|
| 3.4.4. Погрешность статистической линеаризации | 141 |
| 4. СТАТИСТИЧЕСКИЕ РЕШЕНИЯ И ОПТИМАЛЬНАЯ ОБРАБОТКА СИГНАЛОВ | 144 |
| 4.1. СВЕДЕНИЯ ИЗ ТЕОРИИ СТАТИСТИЧЕСКИХ РЕШЕНИЙ..... | 144 |
| 4.1.1. Основные понятия | 144 |
| 4.1.2. Байесовские правила решения..... | 146 |
| 4.2. ОБСУЖДЕНИЕ ПОДХОДОВ РЕШЕНИЯ ЗАДАЧИ | 153 |
| 4.3. ОБНАРУЖЕНИЕ И РАЗЛИЧЕНИЕ СИГНАЛОВ | 155 |
| 4.4. НЕБАЙЕСОВСКИЕ ПРАВИЛА РЕШЕНИЯ | 159 |
| 4.4.1. Критерий максимального правдоподобия..... | 159 |
| 4.4.2. Критерий Неймана-Пирсона | 160 |
| 4.4.3. Минимаксное решающее правило | 163 |
| 4.4.4. Последовательный критерий оценивания | 164 |
| 4.5. ИНТЕРВАЛЬНЫЕ ОЦЕНКИ | 165 |
| 4.6. ОБЩИЕ МЕТОДЫ СИНТЕЗА СИСТЕМ ОБНАРУЖЕНИЯ И ФИЛЬТРАЦИИ | 166 |
| 4.6.1. Формулирование задачи синтеза | 166 |
| 4.6.2. Рекуррентные уравнения фильтрации и обнаружения | 169 |
| 4.6.3. Уравнение для апостериорных параметров | 176 |
| 4.6.4. Общие алгоритмы синтеза для многомерного марковского процесса | 181 |
| 4.7. ЛИНЕЙНЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ И ФИЛЬТРАЦИИ | 185 |
| 4.7.1. Алгоритмы синтеза для одномерного марковского процесса | 185 |
| 4.7.2. Алгоритмы синтеза для многомерного марковского процесса | 188 |
| 4.7.3. Алгоритмы синтеза для непрерывного времени | 193 |
| 4.8. КВАЗИОПТИМАЛЬНЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ И ФИЛЬТРАЦИИ | 196 |
| 4.8.1. Постановка задачи..... | 196 |
| 4.8.2. Метод локальной гауссовской аппроксимации..... | 198 |
| 4.8.3. Метод статистической линеаризации | 207 |
| 4.9. ЭКСТРАПОЛЯЦИЯ И ИНТЕРПОЛЯЦИЯ | 217 |
| 4.9.1. Общие определения..... | 217 |
| 4.9.2. Алгоритм экстраполяции | 218 |
| 4.9.3. Алгоритм интерполяции | 220 |
| 4.10. АЛГОРИТМЫ СИНТЕЗА ДЛЯ СЛУЧАЙНЫХ ТОЧЕЧНЫХ ПРОЦЕССОВ..... | 222 |
| 4.10.1. Формулирование задачи..... | 222 |
| 4.10.2. Рекуррентные уравнения фильтрации и обнаружения | 225 |
| 4.10.3. Уравнения для апостериорных параметров | 228 |
| 4.11. НИЖНЯЯ ГРАНИЦА ДИСПЕРСИИ ОШИБКИ | 236 |
| 4.11.1. Отношение сигнал/шум..... | 236 |
| 4.11.2. Неравенство Крамера-Рао..... | 238 |
| 4.12. УЧЕТ УПРАВЛЕНИЯ В ЗАДАЧАХ ОЦЕНИВАНИЯ СОСТОЯНИЯ ДИНАМИЧЕСКИХ СИСТЕМ | 247 |
| 4.13. СТАТИСТИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ДИНАМИЧЕСКИХ СИСТЕМ..... | 249 |
| 4.14. ПРОГНОЗИРОВАНИЕ И УПРАВЛЕНИЕ В КОМПЬЮТЕРНЫХ СЕТЯХ..... | 253 |
| 4.14.1. ВВЕДЕНИЕ | 253 |
| 4.14.2. Алгоритм для моделей сетевого трафика типа приращений точечного процесса | 255 |
| 4.14.3. Алгоритм для моделей сетевого трафика типа фрактального броуновского движения | 257 |
| 5. СТАТИСТИЧЕСКИЙ СИНТЕЗ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДИНАМИЧЕСКИХ СИСТЕМ | 262 |
| 5.1. ВЕДЕНИЕ В ПРОБЛЕМУ СИНТЕЗА..... | 262 |

| | |
|---|------------|
| 5.2. КРИТЕРИИ ОПТИМИЗАЦИИ | 265 |
| 5.3. МЕТОДЫ ОПТИМИЗАЦИИ | 272 |
| 5.3.1. Принцип максимума для непрерывного времени | 272 |
| 5.3.2. Принцип максимума (минимума) для дискретного времени | 281 |
| 5.3.3. Динамическое программирование | 293 |
| 5.3.4. Методы приближенного синтеза оптимального управления | 303 |
| ПРИЛОЖЕНИЕ..... | 308 |
| СПИСОК ЛИТЕРАТУРЫ | 319 |

Городецкий Александр Яковлевич

ИНФОРМАЦИОННЫЕ СИСТЕМЫ.
ВЕРОЯТНОСТНЫЕ МОДЕЛИ И
СТАТИСТИЧЕСКИЕ РЕШЕНИЯ
Учебное пособие

Лицензия ЛР № 020593 от 07.08.97

Налоговая льгота – Общероссийский классификатор продукции
ОК 005-93, т.2; 95 3005 – учебная литература

Подписано в печать 22.04.2003 Формат 60×84/16.

Усл. печ. л. 20,5. Уч.-изд. л. 20,5. Тираж 200.

Отпечатано с готового оригинал-макета, предоставленного
автором, в типографии СПбГПУ.

195251, Санкт-Петербург, Политехническая, 29

Отпечатано на ризографе RN-2000FP

Поставщик оборудования – фирма «Р-ПРИНТ»

Телефон: (812) 110-65-09

Факс: (812) 315-23-04