

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
УНИВЕРСИТЕТ ИТМО

О.А. Степанов

**МЕТОДЫ ОБРАБОТКИ НАВИГАЦИОННОЙ
ИЗМЕРИТЕЛЬНОЙ ИНФОРМАЦИИ**

Учебное пособие

Санкт-Петербург

 **УНИВЕРСИТЕТ ИТМО**

2017

УДК 629.783
ББК 39.62
С79

Рецензенты:

А.М. Боронахин – д-р техн.наук, профессор Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В.И. Ульянова (Ленина);

А.В. Небылов – д-р техн.наук, профессор Санкт-Петербургского государственного университета аэрокосмического приборостроения.

Рекомендовано к печати Ученым советом факультета Системы управления и робототехники, протокол № 7 от 22.11.2016.

Степанов О.А.

С79 Методы обработки навигационной измерительной информации. – СПб: Университет ИТМО, 2017. – 196 с.

Излагаются общие принципы, используемые при обработке измерительной информации с целью оценивания неизвестных параметров и случайных последовательностей. Рассматриваются различные подходы и методы построения алгоритмов оценивания. Значительное внимание уделяется рекуррентному дискретному фильтру Калмана и его аналогу для непрерывного времени. В прикладном плане изложение ориентировано на решение задач, связанных с обработкой навигационной информации. Материал представлен в виде восьми разделов, содержащих теоретическую часть и значительное количество примеров и задач для самостоятельного решения.

Учебное пособие подготовлено в помощь студентам, обучающимся по специальностям, связанным с обработкой избыточной измерительной информации, в частности, для направления подготовки бакалавров 24.03.02 «Системы управления движением и навигация» и направления подготовки магистров 24.04.02 «Системы управления движением и навигация» в рамках дисциплин «Математические основы теории оценивания», «Методы и задачи обработки навигационной информации».

ISBN 978-5-7577-0554-5

© Университет ИТМО, 2017

© О.А. Степанов, 2017

Содержание

Введение.....	6
Условные сокращения	7
Основные определения.....	8
1. Примеры и постановки задач оценивания постоянных параметров при обработке навигационной информации.....	10
1.1. Оценивание коэффициентов полинома	10
1.2. Задача начальной выставки инерциальной навигационной системы. Простейший случай	12
1.3. Постановка линейной задачи оценивания.....	13
1.4. Определение относительного сдвига реализаций	14
1.5. Определение координат по измерениям дальностей до точечных ориентиров	16
1.6. Определение координат и скорости по спутниковым данным	18
1.7. Постановка нелинейной задачи оценивания и ее линеаризация.....	19
Задачи к разделу	22
Контрольные вопросы	24
2. Решение задач оценивания на основе детерминированного подхода. Метод наименьших квадратов	25
2.1. Основные положения и постановка задачи в методе наименьших квадратов.....	25
2.2. Общее решение задачи синтеза алгоритмов на основе метода наименьших квадратов в линейном случае.....	30
2.3. Анализ точности в методе наименьших квадратов в линейном случае.....	34
2.4. Решение нелинейных задач оценивания. Линеаризованные и итерационные алгоритмы.....	39
Задачи к разделу	47
Контрольные вопросы	48
3. Решение задач оценивания на основе небайесовского подхода. Метод максимума правдоподобия.....	49

3.1. Основные положения и постановка задачи.....	49
3.2. Метод максимума правдоподобия	52
3.3. Общее решение линейной гауссовской задачи. Взаимосвязь с алгоритмами метода наименьших квадратов.....	55
3.4. Решение нелинейной гауссовской задачи. Взаимосвязь с алгоритмами метода наименьших квадратов.....	57
Задачи к разделу	60
Контрольные вопросы	61
4. Байесовский подход. Оптимальные оценки	63
4.1. Формула Байеса и апостериорная плотность.....	63
4.2. Постановка и общее решение байесовской задачи оценивания	65
4.3. Свойства оптимальных байесовских оценок	67
4.4. Решение линейной гауссовской задачи. Взаимосвязь с алгоритмами метода наименьших квадратов	69
4.5. Линейные оптимальные оценки. Постановка задачи и ее общее решение	71
4.6. Повышение точности оценивания при использовании нелинейных алгоритмов	73
Задачи к разделу	77
Контрольные вопросы	78
5. Алгоритмы комплексной обработки измерений.....	79
5.1. Примеры и постановка задачи комплексной обработки.....	79
5.2. Комплементарный фильтр. Инвариантная схема обработки	82
5.3. Неинвариантная схема обработки.....	84
5.4. Централизованная и децентрализованная схемы обработки	87
5.5. Рекуррентная схема обработки.....	90
Задачи к разделу	91
Контрольные вопросы	94
6. Случайные последовательности	95
6.1. Определение случайной последовательности и ее описание.....	95
6.2. Стационарные случайные последовательности. Дискретный белый шум	97
6.3. Марковские последовательности	100
6.4. Формирующий фильтр	101
6.5. Динамика изменения матрицы ковариаций марковской последовательности	105
Задачи к разделу	109
Контрольные вопросы	110

7. Фильтрация случайных последовательностей. Дискретный фильтр Калмана	112
7.1. Постановка и решение задачи нерекуррентного оптимального линейного оценивания случайных последовательностей.....	112
7.2. Постановка линейной задачи рекуррентной оптимальной фильтрации случайных последовательностей	117
7.3. Дискретный фильтр Калмана для случайных последовательностей	119
7.4. Уравнения ошибок фильтра Калмана. Инновационная последовательность.....	127
7.5. Динамика изменения матрицы ковариаций и установившийся режим в задаче фильтрации	129
7.6. Постановка и общее решение задачи рекуррентной оптимальной фильтрации случайных последовательностей с нелинейными измерениями	133
Задачи к разделу	136
Контрольные вопросы	137
8. Фильтрация случайных процессов. Фильтр Калмана-Бьюси.....	138
8.1. Случайные процессы и методы их описания	138
8.2. Формирующий фильтр	142
8.3. Постановка и общее решение задачи оптимальной линейной фильтрации. Фильтр Калмана-Бьюси	145
8.4. Связь непрерывных и дискретных алгоритмов фильтрации.....	150
Задачи к разделу	154
Контрольные вопросы	155
Приложение 1	157
П1. Основные матричные операции.....	157
Приложение 2	169
П2. Случайные величины и векторы.....	169
П2.1. Случайная величина и ее описание.....	169
П2.2. Гауссовские случайные величины и их характеристики.....	171
П2.3. Случайные векторы и их описание	173
П2.4. Гауссовские случайные векторы и их характеристики.....	175
Приложение 3	179
П3. Обыкновенные дифференциальные уравнения	179
Литература	182
Предметный указатель.....	185
История кафедры Информационно-навигационных систем	193

Введение

Предлагаемое учебное пособие подготовлено на основании многолетнего опыта, накопленного автором при чтении курса лекций «Математические основы теории оценивания» для студентов 3-его курса кафедры информационно-навигационных систем Университета ИТМО, и опыта прочтения лекций, посвященных методам теории фильтрации при решении задач обработки навигационной информации, читаемых студентам старших курсов трех ведущих вузов Санкт-Петербурга: Университета ИТМО, Государственного электротехнического университета (ЛЭТИ) им. В.И. Ульянова (Ленина) и Государственного университета аэрокосмического приборостроения. Значительную роль при формировании содержания учебного пособия сыграл и опыт, накопленный при чтении аналогичного курса лекций для студентов Технологического института г. Карлсруэ (Германия).

В учебном пособии излагаются общие принципы, используемые при обработке измерительной информации с целью оценивания неизвестных параметров. Рассматриваются различные подходы к построению алгоритмов оценивания в зависимости от уровня априорной информации статистического характера об оцениваемых параметрах и ошибках измерения, в том числе такие, как метод наименьших квадратов, метод максимума правдоподобия и байесовский метод. Анализируется взаимосвязь алгоритмов, получаемых в рамках рассматриваемых подходов. Методы и алгоритмы, полученные для постоянного вектора, обобщаются применительно к оцениванию случайных последовательностей, наиболее важным из которых является рекуррентный дискретный фильтр Калмана. Кроме того, рассматривается фильтр Калмана для непрерывного времени и обсуждается переход от непрерывных алгоритмов к их дискретным аналогам. Отметим, что при изложении материала обсуждается не только проблема синтеза алгоритмов, но и весьма важная для практических приложений проблема анализа точности. В этой связи предпочтение отдается алгоритмам, вытекающим из байесовского подхода, которые, помимо самой оценки, обеспечивают

возможность выработки текущих характеристик точности в виде соответствующей матрицы ковариаций ошибок оценивания.

В прикладном плане пособие ориентировано на решение задач, связанных с обработкой навигационной информации. Именно в этой области активно используются излагаемые в пособии алгоритмы оценивания. Материал представлен в виде восьми разделов (лекций). Каждый раздел включает теоретическую часть и значительное количество иллюстрирующих примеров, в которых излагается постановка рассматриваемых задач и конкретизируются обсуждаемые алгоритмы. В конце разделов приведены задачи для самостоятельного решения и перечень контрольных вопросов.

Пособие предназначено для студентов, обучающихся по специальностям, связанным с обработкой избыточной измерительной информации. При его составлении за основу взят материал книг [25, 26], адаптированный для изучения студентами соответствующих специальностей и ориентированный на ознакомление с основами теории и принципами построения алгоритмов оценивания неизвестных параметров и случайных последовательностей.

При изложении материала активно используется математический аппарат в объеме, традиционно изучаемом на первых курсах технических вузов. Для понимания содержания учебного пособия требуется хорошее знание основ теории вероятностей, матричного исчисления и теории дифференциальных уравнений. В связи с этим, с целью удобства читателя, в приложениях к учебному пособию приведены основные сведения по каждой из перечисленных тем.

Автор выражает благодарность всем сотрудникам и аспирантам кафедры информационно-навигационных систем Университета ИТМО и научно-образовательного центра АО «Концерн» ЦНИИ «Электроприбор», помогавшим отрабатывать предлагаемую рукопись. Особая благодарность Уткиной Т.П. и Богомолу В.В., выполнившим значительный объем работ по окончательному редактированию текста.

Условные сокращения

ИНС	– инерциальная навигационная система
МНК	– метод наименьших квадратов
ММНК	– модифицированный метод наименьших квадратов
МФП	– максимум функции правдоподобия
ОМНК	– обобщенный метод наименьших квадратов
ОФК	– обобщенный фильтр Калмана
СКО	– среднеквадратическое отклонение
СНС	– спутниковая навигационная система

- ФК – фильтр Калмана
- ф.р.в. – функция распределения вероятностей
- ф.п.р.в. – функция плотности распределения вероятностей

Основные определения

Измерения – набор величин, вырабатываемых соответствующими датчиками. Применительно к навигационным задачам в качестве таких датчиков могут выступать акселерометры, гироскопы, барометры, дальнометры и т.п.

Под измерениями в настоящем учебном пособии понимаются также значения координат и/или составляющих скорости подвижного объекта и/или параметров его ориентации, вырабатываемые инерциальными навигационными системами, системами счисления и аппаратурой потребителей спутниковых навигационных систем.

Для измерений, как правило, используется обозначение y_i , $i = 1, 2, \dots$ – для случая дискретного времени или $y(t)$ – для случая непрерывного времени.

Измерения, как и все определяемые далее величины, могут быть как скалярными, так и векторными.

Ошибки измерения – разность между истинными и измеренными значениями величин – ускорений, скоростей, координат и т.д.

Для ошибок измерений, как правило, используются обозначение v_i , $i = 1, 2, \dots$ – для случая дискретного времени или $v(t)$ – для случая непрерывного времени.

Оцениваемый параметр – параметр, который отыскивается в результате обработки накопленных измерений.

Для оцениваемых параметров, как правило, используется обозначение x_i , $i = 1, 2, \dots$ – для случая дискретного времени или $x(t)$ – для случая непрерывного времени.

Модель измерений – математические соотношения $y_i = s_i(x_i, v_i)$ или $y(t) = s(t, x(t), v(t))$, определяющие, каким образом связаны измерения с оцениваемыми параметрами и ошибками измерения.

Далее используются линейные и нелинейные модели или функции $\varphi(x)$, т.е. такие, которые удовлетворяют либо не удовлетворяют принципу суперпозиции:

$$\varphi(ax_1 + bx_2) = a\varphi(x_1) + b\varphi(x_2),$$

где a и b – произвольные постоянные величины.

Модель оцениваемых параметров – математические соотношения в виде разностных или дифференциальных уравнений, определяющих, ка-

ким образом эти параметры изменяются во времени. Далее, как правило, используются линейные модели, задаваемые линейными уравнениями.

Задача оценивания – задача, целью которой является нахождение неизвестных параметров (получение оценки) с использованием измерений и информации о моделях для измерений и самих оцениваемых параметров.

Оценка параметра – значение параметра, получаемое в результате решения задачи оценивания.

Для полученных оценок параметров, как правило, используется обозначение \hat{x}_i , $i=1,2,\dots$ – для случая дискретного времени или $\hat{x}(t)$ – для случая непрерывного времени.

Алгоритм оценивания – последовательность вычислений и логических операций, выполняемых с измерениями с целью получения оценок отыскиваемых параметров.

Ошибки оценивания – разность между истинными значениями параметров и их оценками.

Для ошибок оценок параметров, как правило, используется обозначение $\varepsilon_i = x_i - \hat{x}_i$, $i=1,2,\dots$ – для случая дискретного времени или $\varepsilon(t) = x(t) - \hat{x}(t)$ – для случая непрерывного времени.

Критерий оценивания – функция или функционал, который подлежит минимизации или максимизации при нахождении оцениваемых параметров

РАЗДЕЛ 1

Примеры и постановки задач оценивания постоянных параметров при обработке навигационной информации

Рассмотрим примеры типичных линейных и нелинейных задач оценивания постоянных параметров, с которыми приходится сталкиваться при обработке измерительной информации [1, 2, 4, 5, 8, 10-12, 15, 17, 19, 20, 22–34].

1.1. Оценивание коэффициентов полинома

Предположим, что в моменты времени t_i , $i = \overline{1..m}$ с использованием имеющегося на борту летательного аппарата датчика проведены измерения высоты $h(t)$, и при этом можно полагать, что в течение интервала проведения измерений эта высота не меняется. Тогда, вводя обозначение $h = x$ и полагая, что измерения содержат ошибки v_i , $i = \overline{1..m}$, задачу определения h можно свести к оцениванию неизвестной постоянной величины x по набору зашумленных измерений

$$y_i = x + v_i, \quad i = \overline{1..m}. \quad (1.1)$$

Вводя составленный из единичек столбец H , $H^T = [1, 1, \dots, 1]$, m -мерные векторы $y = (y_1, \dots, y_m)^T$ и $v = (v_1, \dots, v_m)^T$, измерения (1.1) можно записать как

$$y = Hx + v. \quad (1.2)$$

Отличительная особенность этого выражения заключается в том, что зависимость измерений от оцениваемого параметра – линейная. В этой связи можно говорить о линейном характере измерений. Можно ввести более сложную модель для описания изменения высоты за время проведения измерений. В частности, если считать, что высота меняется по линейному закону (представляет собой **линейный тренд**), то измерения, проводимые с интервалом Δt , можно представить в виде

$$y_i = x_0 + Vt_i + v_i, \quad i = \overline{1..m}, \quad (1.3)$$

где x_0, V – подлежащие определению неизвестные начальная высота и вертикальная скорость, полагаемая постоянной; $t_i = (i-1)\Delta t$ – моменты времени от начала наблюдения.

Измерения (1.3) и в этом случае могут быть представлены с помощью соотношения (1.2), для чего достаточно ввести оцениваемый вектор и матрицу H в следующем виде:

$$x = (x_1, x_2)^T = (x_0, V)^T; \quad H^T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_m \end{bmatrix}. \quad (1.4)$$

Можно получить представление (1.2) и для более общего случая, когда изменение высоты описывается полиномом $n-1$ -первого порядка, и тогда задача в математическом плане может быть сведена к оцениванию коэффициентов этого полинома по измерениям

$$y_i = x_1 + x_2 t_i + x_3 t_i^2 + \dots + x_n t_i^{n-1} + v_i, \quad i = \overline{1, m}. \quad (1.5)$$

К такой постановке нередко сводится задача предварительной обработки измерений используемых датчиков для снижения уровня их шумов.

С необходимостью оценивания коэффициентов полинома приходится также сталкиваться и при решении так называемой задачи калибровки приборов. Ее суть заключается в том, что показания прибора сравниваются либо с эталонным значением измеряемого параметра, либо с показаниями другого, более точного измерителя. При этом требуется построить модель изменения ошибок во времени с тем, чтобы в дальнейшем при отсутствии эталона учесть эти изменения и тем самым повысить точность измерения в штатном режиме. Пример реализаций ошибок измерений, явно содержащих постоянную составляющую в одном случае и квадратичный тренд (полином второго порядка) в другом, приведен на рис. 1.1.

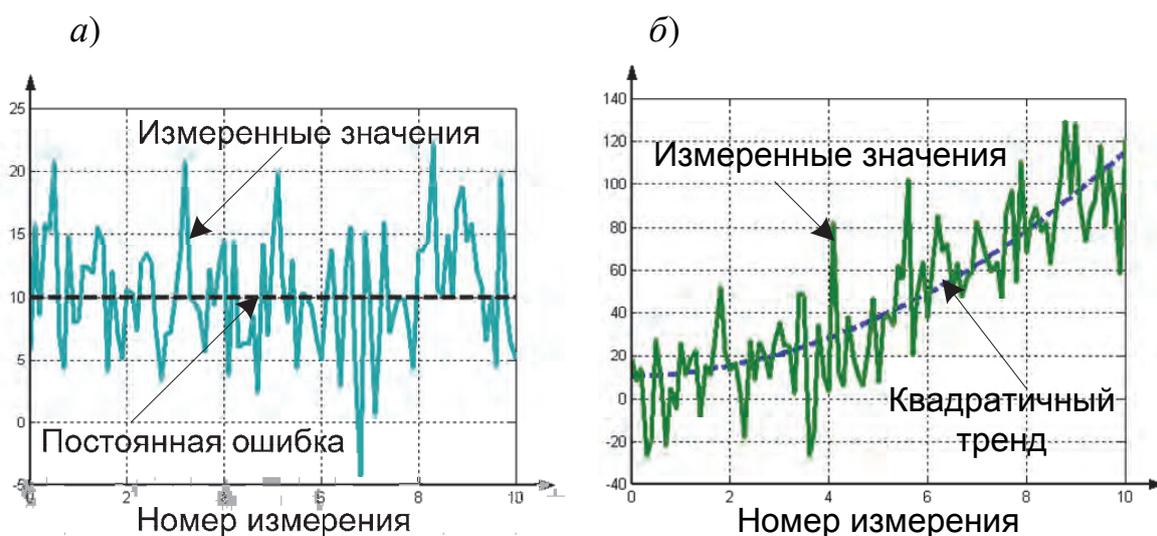


Рис. 1.1. Примеры реализации ошибок измерений, содержащих постоянную величину (а) и квадратичный тренд (б), $t = 10$ с, $\Delta t = 0,1$ с

В качестве аргумента в выражении (1.5) не обязательно должно фигурировать время. Это могут быть и другие физические величины, например температура. Известно, что ошибки высокоточных измерителей существенным образом зависят от температуры. Для снижения этих ошибок предусматривается специальная система термостабилизации. Если предварительно описать, а затем учитывать зависимость изменения погрешностей прибора от температуры, например с помощью полиномиальной модели, то в ряде случаев удастся заметно снизить требования к дорогостоящей системе термостабилизации.

1.2. Задача начальной выставки инерциальной навигационной системы. Простейший случай

Известно, что, если пренебречь ошибками чувствительных элементов инерциальной навигационной системы (ИНС), погрешность выработки скорости для одного из каналов ИНС в простейшем случае может быть приближенно описана с помощью следующего выражения [1, 12]:

$$\Delta V_E(t_i) = -\alpha(0)\sqrt{gR} \sin \omega_{uu} t_i + \Delta V_E(0) \cos \omega_{uu} t_i, \quad i = \overline{1, m},$$

где $\alpha(0)$, $\Delta V_E(0)$ – начальная ошибка построения вертикали (угла между истинной и вырабатываемой в ИНС плоскостями горизонта) и начальная ошибка скорости; g – ускорение силы тяжести; R – радиус Земли;

$\omega_{uu} = \sqrt{\frac{g}{R}}$ – частота, соответствующая периоду Шулера: $T_{uu} = 2\pi \sqrt{\frac{R}{g}} \approx 84$

мин.

Предположим, что объект, на котором установлена ИНС, неподвижен, т.е. известно, что скорость нулевая. В этом случае вырабатываемые инерциальной системой показания скорости фактически будут представлять собой ошибку ИНС. Снимая эти показания в дискретные моменты времени, можем записать:

$$y(t_i) = -\alpha(0)\sqrt{gR} \sin \omega_{uu} t_i + \Delta V_E(0) \cos \omega_{uu} t_i + v_i, \quad (1.6)$$

где v_i – ошибка съема показаний.

Используя набор измерений, можно оценить начальную ошибку построения вертикали и ошибку выработки скорости. По сути это есть задача начальной выставки ИНС в ее простейшем рассмотрении. Вводя вектор оцениваемых параметров и матрицу H с помощью соотношений

$$x^T = (\alpha(0), \Delta V_E(0)); \quad (1.7)$$

$$H^T = \begin{bmatrix} -\sqrt{gR} \sin \omega_{uu} t_1, & -\sqrt{gR} \sin \omega_{uu} t_2, & \cdot & \cdot & -\sqrt{gR} \sin \omega_{uu} t_m \\ \cos \omega_{uu} t_1, & \cos \omega_{uu} t_2, & \cdot & \cdot & \cos \omega_{uu} t_m \end{bmatrix}, \quad (1.8)$$

нетрудно и эти измерения представить в виде, аналогичном (1.2).

Если предположить, что задача решается на малом, по сравнению с периодом Шулера, интервале времени, то, раскладывая в ряд значения функций $\sin \omega_u t_i$ и $\cos \omega_u t_i$, эту задачу легко свести к задаче оценивания коэффициентов полинома. В частности, если считать, что $\sin \omega_u t_i \approx \frac{2\pi}{T_u} t_i$, $\cos \omega_u t_i \approx 1$ при $t_i \ll T_u$, то измерения могут быть приближенно записаны с помощью полинома 1-го порядка

$$y(t_i) \approx -\alpha(0)\sqrt{gR} \frac{2\pi}{T_u} t_i + \Delta V_E(0) + v_i = \Delta V_E(0) - g\alpha(0)t_i + v_i. \quad (1.9)$$

С оцениванием коэффициентов полиномов второго и третьего порядка приходится иметь дело и в задаче выработки так называемых динамических параметров, при решении которой погрешности ИНС по составляющим скорости и перемещения на интервале времени ΔT приближенно описываются с помощью полиномиальных моделей [1].

1.3. Постановка линейной задачи оценивания

Все перечисленные выше задачи могут быть сведены к следующей **общей постановке линейной задачи оценивания**.

Задан постоянный n -мерный вектор $x = (x_1, \dots, x_n)^T$

$$\dot{x} = 0, \quad (1.10)$$

и имеется m -мерный вектор измерений $y = (y_1, \dots, y_m)^T$

$$y = Hx + v, \quad (1.11)$$

где H – $m \times n$ -мерная матрица, а $v = (v_1, \dots, v_m)^T$ – m -мерный вектор, описывающий ошибки измерения.

Требуется, располагая измерением (1.11), найти в каком-то смысле «хорошую» оценку $\hat{x}(y)$ неизвестного вектора x . Этот вектор в задачах оценивания называется **вектором состояния (state vector)**. Запись $\dot{x} = \frac{dx}{dt} = 0$

означает, что вектор x постоянный, поскольку решение этого простейшего дифференциального уравнения $x = \text{const}$. Такая запись используется в целях установления преемственности рассматриваемой в этой главе задачи оценивания постоянного вектора с более общими задачами оценивания изменяющегося вектора, поведение которого во времени может быть описано с помощью дифференциальных или разностных уравнений.

При решении сформулированной задачи важным представляется не только сам алгоритм вычисления оценки по имеющимся измерениям, но и умение количественно охарактеризовать уровень **ошибки оценки**, определяемой как

$$\varepsilon(y) = x - \hat{x}(y), \quad (1.12)$$

т.е. количественно оценить точность вырабатываемой предлагаемым алгоритмом оценки. Последнее особенно актуально при решении задач обработки навигационной информации. Таким образом, в рассматриваемой проблеме можно выделить две важные задачи: **задачу синтеза алгоритма**, т.е. получения конкретной процедуры вычисления оценок $\hat{x}(y)$, и **задачу анализа точности**, суть которой заключается в выявлении свойств ошибок (1.12), в частности, в определении уровня этих ошибок.

1.4. Определение относительного сдвига реализаций

Рассмотрим далее примеры нелинейных задач оценивания. Нередко при решении прикладных задач приходится сталкиваться с задачей определения относительного сдвига двух реализаций. Поясним ее суть. Пусть задана (известна) в общем случае нелинейная функция $s(t)$ скалярного аргумента и проведены измерения типа

$$y_i = s(t_i + \tau) + v_i, \quad i = \overline{1..m}, \quad (1.13)$$

где $v_i, i = \overline{1..m}$ – ошибки измерения функции в точках $t_i + \tau$, значения которых заданы с точностью до неизвестной постоянной величины τ .

Требуется, зная функцию $s(t)$ и располагая значениями $y_i, t_i, i = \overline{1..m}$, оценить τ , т. е. необходимо определить временное запаздывание (сдвиг) одной реализации относительно другой. В случае, например, когда $s(t)$ является гармоническим колебанием, это выражение можно конкретизировать как

$$y_i = A \sin(\omega t_i + \phi_0) + v_i, \quad i = \overline{1..m}, \quad (1.14)$$

где A – амплитуда; $\omega = 2\pi f$ – круговая частота; ϕ_0 – фаза.

При известных значениях амплитуды и частоты получаем **задачу определения фазы**. Отличительная особенность этого выражения заключается в том, что зависимость измерений от оцениваемого параметра – нелинейная. В этой связи можно говорить о нелинейном характере измерений. В общем виде в **нелинейных задачах оценивания** измерения можно записать как

$$y = s(x) + v, \quad (1.15)$$

где x, v – n - и m -мерные векторы; $s(x) = (s_1(x), \dots, s_m(x))^T$ – m -мерная вектор-функция.

В рассмотренном примере при известных значениях амплитуды и частоты (1.14) сводится к (1.15), если $x = \phi_0$ и $s(x) = (A \sin(\omega t_1 + x), \dots, A \sin(\omega t_m + x))^T$. Геометрический смысл этой задачи ясен из рис. 1.2.

Если требуется решить **задачу определения частоты** $x = \omega$, то при известной фазе и амплитуде $s(\bullet) = (A \sin(xt_1 + \phi_0), \dots, A \sin(xt_m + \phi_0))^T$. $s(\cdot) = (A \sin(xt_1 + \phi_0), \dots, A \sin(xt_m + \phi_0))^T$. Если требуется определить не только фазу, но и частоту сигнала, то в (1.15) следует принять $x^T = (\phi_0, \omega)$, а $s(\bullet) = (A \sin(x_2 t_1 + x_1), \dots, A \sin(x_2 t_m + x_1))^T$. В этом случае получаем **задачу фазовой автоподстройки частоты**, которую нередко приходится решать при построении различных измерительных приборов. В общем случае в **задаче определения параметров гармонического сигнала** могут быть неизвестны все три параметра – амплитуда, частота и фаза, тогда вектор оцениваемых параметров будет трехмерным $x^T = (A, \phi_0, \omega)$. Нетрудно заметить, что при известных значениях частоты и фазы зависимость измерений от амплитуды линейная (см. задачу 1.2).

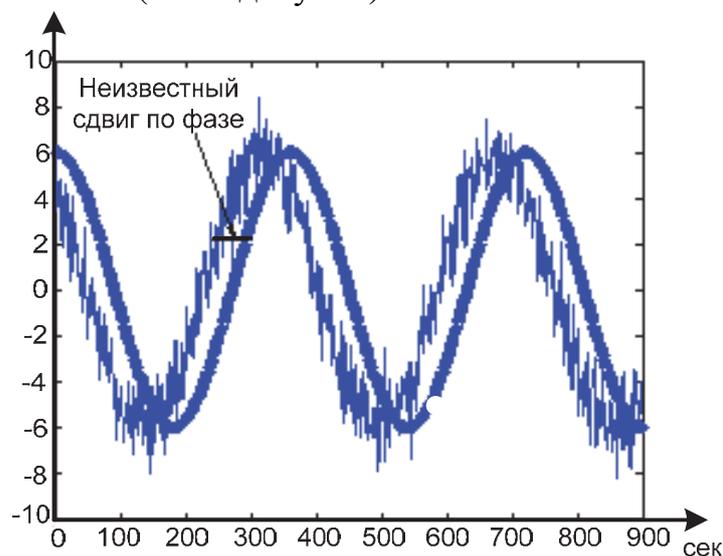


Рис. 1.2. К задаче определения фазы

Большинство радионавигационных систем также основано на определении сдвига одной измеренной с ошибками реализации функции типа (1.13) относительно другой (эталонной) – её точно известной копии [29, 30]. В частности, возможность нахождения координат в современных спутниковых навигационных системах (СНС) обеспечивается за счет одновременного измерения нескольких задержек, обусловленных конечным временем распространения радиоволн между спутниками и потребителем. Эти задержки измеряются путем сопоставления огибающей, выделяемой из принимаемых от спутников сигналов, с их копиями, формируемыми в приемной аппаратуре потребителя [5, 23].

К аналогичной постановке сводится и задача так называемой **корреляционно-экстремальной навигации**, или **навигации по геофизическим полям** [4, 11, 15, 25, 28]. Идея этого метода навигации заключается в уточнении координат подвижного объекта на основе сопоставления реализации измеренных значений некоторого геофизического параметра, например

рельефа местности, с реализацией, вычисленной с использованием заранее снятой карты. В этом случае измерения (1.13) будут соответствовать одномерному варианту этой задачи, если временной аргумент заменить на пространственный, а функция $s(\bullet)$ будет соответствовать снимаемому с карты изменению рельефа местности от этой пространственной координаты (рис. 1.3). Здесь также требуется найти пространственный сдвиг измеренной реализации относительно карты.

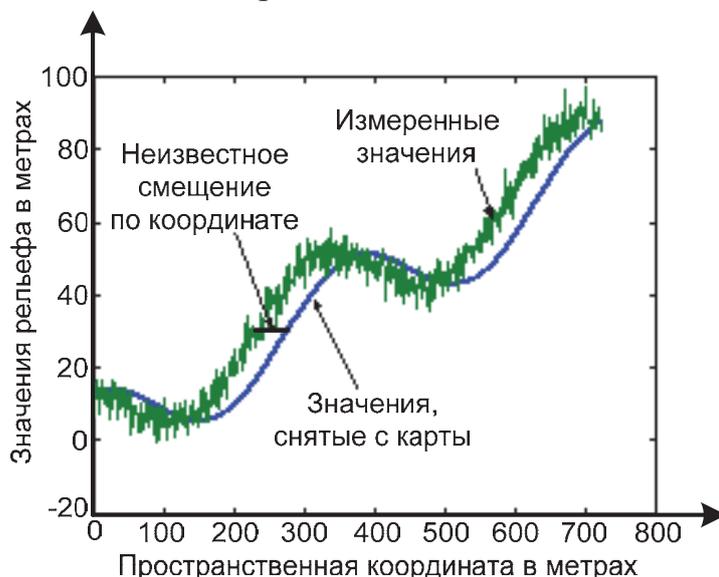


Рис. 1.3. К одномерной задаче уточнения координат места по рельефу местности

1.5. Определение координат по измерениям дальностей до точечных ориентиров

В качестве примера еще одной нелинейной задачи оценивания можно привести задачу определения координат объекта на плоскости с использованием измерений дальностей до ориентиров, координаты которых предполагаются известными (рис. 1.4). В этом случае $x = (x_1, x_2)^T$, и измерения принимают вид

$$y_i = s_i(x) + v_i = \sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2} + v_i, \quad i = \overline{1, m}, \quad (1.16)$$

где x_1, x_2 ; x_1^i, x_2^i , $i = \overline{1, m}$ — координаты объекта и ориентиров соответственно, а

$$s_i(x) = D_i(x) = \sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2}.$$

Идея определения местоположения с использованием измерений дальностей достаточно простая. При наличии безошибочного измерения дальности до одного точечного ориентира координаты объекта становятся известными с точностью до его расположения на окружности, радиус которой совпадет с измеренным значением дальности. Линии постоянного зна-

чения измеряемых параметров в навигации называются **изолиниями положения** [10]. В данном случае они представляют собой окружности. Располагая двумя измеренными без ошибок значениями дальностей, координаты объекта можно получить как одну из двух возможных точек пересечения этих изолиний. Поскольку измерения содержат ошибки, то вместо линий будут формироваться полосы, заключенные между окружностями, равными максимально и минимально возможным значениям дальностей при заданном уровне ошибок измерения. При наличии двух измерений возможные координаты объекта будут с высокой степенью вероятности располагаться внутри фигуры, формируемой в результате пересечения двух полос. При наличии большего количества измерений возникает проблема определения координат с максимальной точностью.

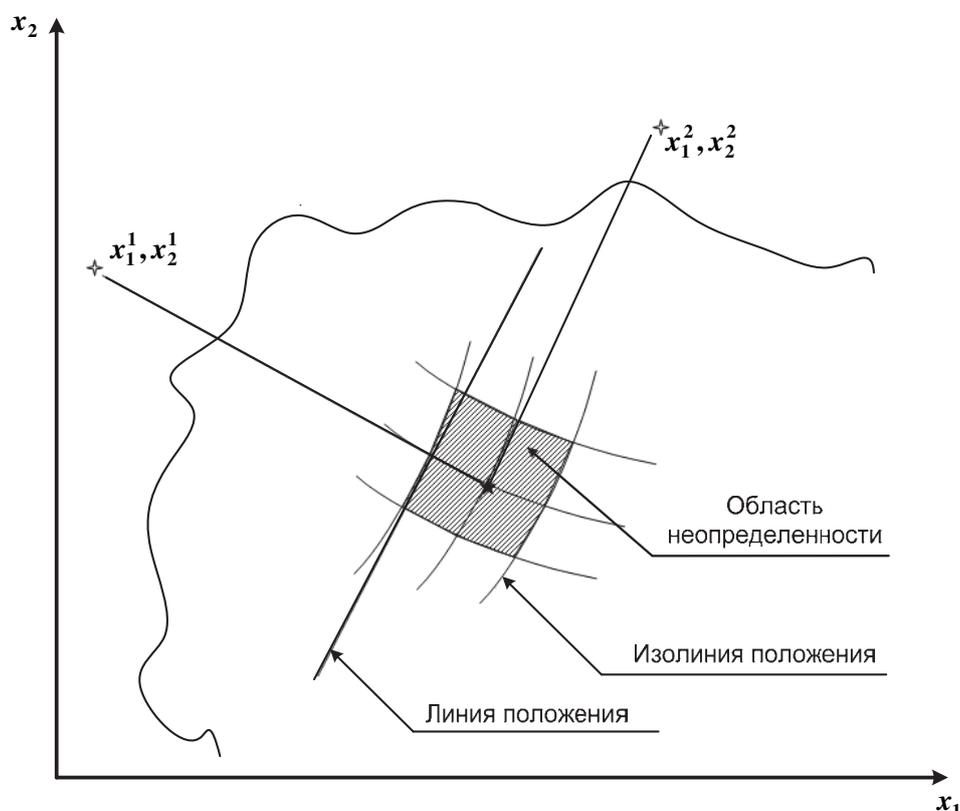


Рис. 1.4. Определение координат объекта на плоскости с использованием измерений дальности до двух ориентиров с известными координатами

Если все измерения содержат одинаковую постоянную составляющую ошибки, то, обозначая ее как x_3 , можем записать

$$y_i = s_i(x) + v_i = \sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2} + x_3 + v_i, \quad i = \overline{1, m}. \quad (1.17)$$

Такую составляющую будем называть в дальнейшем систематической составляющей ошибки или просто **систематической ошибкой**. Понятно, что для определения координат по этим измерениям требуется оценить вектор $x = (x_1, x_2, x_3)^T$.

В то же время цель решения задачи заключается в нахождении координат объекта x_1, x_2 , а неизвестная постоянная составляющая ошибки измерения x_3 наряду с ошибками v_i лишь «мешает» ее решению. В связи с этим параметры типа x_3 нередко называют **мешающими параметрами**. Процедура включения неизвестных параметров в оцениваемый вектор называется **расширением оцениваемого вектора состояния (augmentation)**.

1.6. Определение координат и скорости по спутниковым данным

Легко обобщить рассмотренную в предыдущем примере задачу на случай определения координат в трехмерном пространстве. Именно такого типа измерения используются при определении координат с помощью упоминавшихся уже СНС. Измерив задержки между принимаемыми и формируемыми в аппаратуре потребителя реализациями и располагая известной скоростью распространения радиоволн, измеренные значения дальностей до спутников можно представить как

$$\rho_i = \sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2 + (x_3^i - x_3)^2} + c\Delta t + \varepsilon_i, \quad (1.18)$$

где x_1, x_2, x_3 – неизвестные координаты потребителя на момент приема сигналов, заданные в прямоугольной геоцентрической системе координат; $x_j^i, j=1,2,3$ – координаты i -го спутника в той же системе координат, передаваемые потребителю в навигационном сообщении; Δt – ошибка часов потребителя; ε_i – суммарная ошибка измерения; c – скорость света [5, 19, 23].

Заметим, что орбиты и расположение на них спутников выбираются таким образом, чтобы практически в любой точке земного шара в любое время можно было бы получить измерения не менее чем от четырех спутников. В состав СНС помимо спутников входят наземные контрольные станции, задача которых заключается в уточнении параметров движения (координат и скоростей) спутников (рис. 1.5).

Компонентами оцениваемого вектора $x = (x_1, x_2, x_3, \Delta t)^T$ являются координаты потребителя и ошибка его часов.

Помимо координат, с использованием СНС могут быть определены также и составляющие скорости потребителя. С этой целью используются измерения доплеровских смещений несущей частоты ($f_j^{\text{доп}}$), возникающих за счет взаимного перемещения спутника и потребителя.

Используя эти измерения, можно записать

$$\dot{\rho}_i = \frac{(x_1^i - x_1)(\dot{x}_1^i - \dot{x}_1) + (x_2^i - x_2)(\dot{x}_2^i - \dot{x}_2) + (x_3^i - x_3)(\dot{x}_3^i - \dot{x}_3)}{\sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2 + (x_3^i - x_3)^2}} + c\dot{\Delta t} + \tilde{\varepsilon}_i, \quad (1.19)$$

где $\dot{x}_j; \dot{x}_j^i, j = 1, 2, 3$ – составляющие скорости потребителя и i -го спутника соответственно; Δt^i – ошибка из-за дрейфа часов потребителя; $\tilde{\epsilon}_i$ – суммарная ошибка доплеровских измерений. Здесь оцениванию подлежит вектор, включающий составляющие скорости потребителя $\dot{x}_1, \dot{x}_2, \dot{x}_3$ и ошибку из-за дрейфа его часов.

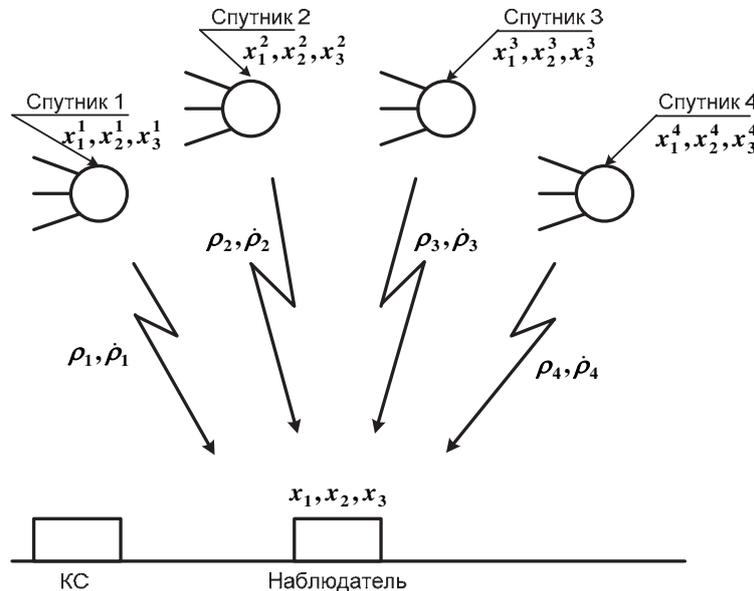


Рис. 1.5. Определение координат и скорости по данным СНС:

КС – контрольные станции, определяющие параметры движения спутников

Как видно из представленных выражений, ошибки измерения дальности и скорости ее изменения, обусловленные ошибками часов потребителя и дрейфом их ухода, носят систематический характер. Их наличие и определило используемые для измеряемых параметров (1.18), (1.19) термины – псевдодальность и псевдоскорость, а также тот факт, что для решения навигационной задачи одновременно требуются измерения не менее чем от четырех спутников.

1.7. Постановка нелинейной задачи оценивания и ее линеаризация

Все перечисленные выше задачи могут быть сведены к следующей **общей постановке нелинейной задачи оценивания**.

Задан постоянный неизвестный n -мерный вектор $x = (x_1, \dots, x_n)^T$

$$\dot{x} = 0, \quad (1.20)$$

и имеется m -мерный вектор измерений $y = (y_1, \dots, y_m)^T$

$$y = s(x) + v, \quad (1.21)$$

где $s(\bullet) = (s_1(x), \dots, s_m(x))^T$ – m -мерная в общем случае нелинейная функция, а $v = (v_1, \dots, v_m)^T$ – m -мерный вектор, характеризующий ошибки измерения.

Требуется, располагая измерением (1.21), найти в каком-то смысле «хорошую» оценку $\hat{x}(y)$ неизвестного вектора x .

Здесь, как и в линейном случае, необходимо выделять задачу синтеза алгоритмов и задачу анализа точности, суть которой заключается в определении свойств ошибок оценивания (1.12). Ясно, что такая постановка включает в себя и линейный случай, поскольку, заменив в (1.21) $s(x)$ на Hx , приходим к задаче (1.10), (1.11).

Далее при обсуждении возможных путей построения алгоритмов оценивания и анализа их точности будут рассматриваться оба класса задач: линейные и нелинейные. Однако основное внимание, тем не менее, будет уделяться линейным задачам. Объясняется это не только простотой решения таких задач, но и тем фактом, что многие нелинейные задачи в ряде случаев могут без значительных потерь в точности быть сведены к линейным постановкам. Осуществляется это на основе **линеаризации функции** $s(\bullet) = s(\cdot) = (s_1(x), \dots, s_m(x))^T$, т.е. приближенного представления ее в виде ряда Тейлора, в котором сохраняются только члены первого порядка:

$$s(x) \approx s(x^n) + \left. \frac{ds}{dx^T} \right|_{x=x^n} (x - x^n) = s(x^n) + H(x^n)(x - x^n), \quad (1.22)$$

где x^n – точка линеаризации;

$$H(x^n) = \left. \frac{ds}{dx^T} \right|_{x=x^n}. \quad (1.23)$$

При построении алгоритмов оценивания, основанных на линеаризации, в качестве оцениваемого вектора состояния удобно использовать вектор отклонений в виде

$$\delta x = (x - x^n). \quad (1.24)$$

Перенесем известные слагаемые в левую часть и введем обозначение

$$\tilde{y}(x^n) \stackrel{\Delta}{=} y - s(x^n). \quad (1.25)$$

Тогда можно записать следующее приближенное соотношение:

$$\tilde{y}(x^n) \approx \left. \frac{ds}{dx^T} \right|_{x=x^n} (x^n) \delta x + v = H(x^n) \delta x + v, \quad (1.26)$$

в котором $\tilde{y}(x^n)$ представляет собой измерение, полученное из исходного путем вычитания известных величин.

Таким образом, исходная нелинейная задача приближенно сведена к линейной задаче оценивания δx по измерениям (1.26).

Ясно, что линеаризованное описание будет допустимым лишь в ограниченной окрестности точки линеаризации. Точность такого представления ориентировочно можно оценить для скалярного случая величиной

$$\delta = \frac{d^2 s}{dx^2} (x - x^n)^2, \quad (1.27)$$

определяющей уровень слагаемых второго порядка в разложении Тейлора и зависящей, с одной стороны, от величины второй производной $\frac{d^2 s}{dx^2}$, а с другой – от ожидаемых возможных отклонений действительных неизвестных значений оцениваемого параметра от точки линеаризации, т.е. от разности $(x - x^n)$.

О допустимости линеаризованного описания при решении задач оценивания можно судить, сопоставляя уровень ожидаемых ошибок измерения v с ожидаемыми значениями

$$\delta = \frac{d^2 s}{dx^2} (x - x^n)^2.$$

Проиллюстрируем сказанное на примере рассмотренной в 1.5 задачи определения координат объекта на плоскости по измерениям дальностей до точечных ориентиров по измерениям (1.16). Используя описанную процедуру линеаризации, можем записать:

$$\tilde{y}(x^n) \approx H(x^n)\delta x + v, \quad (1.28)$$

где
$$\tilde{y}_i(x^n) = y_i - D_i(x^n), \quad i = \overline{1, m}; \quad (1.29)$$

$$D_i(\hat{x}) = \sqrt{(x_1^i - x_1^n)^2 + (x_2^i - x_2^n)^2}; \quad (1.30)$$

$$H(\hat{x}) = \begin{bmatrix} (x_1^n - x_1^1) / D_1(\hat{x}) & (x_2^n - x_2^1) / D_1(\hat{x}) \\ (x_1^n - x_1^2) / D_2(\hat{x}) & (x_2^n - x_2^2) / D_2(\hat{x}) \\ \cdot & \cdot \\ (x_1^n - x_1^m) / D_m(\hat{x}) & (x_2^n - x_2^m) / D_m(\hat{x}) \end{bmatrix} =$$

$$= - \begin{bmatrix} \sin \Pi_1(x^n) & \cos \Pi_1(x^n) \\ \sin \Pi_2(x^n) & \cos \Pi_2(x^n) \\ \cdot & \cdot \\ \sin \Pi_m(x^n) & \cos \Pi_m(x^n) \end{bmatrix} \quad (1.31)$$

– матрица размерности $m \times 2$.

В этих выражениях $\Pi_i(x^n)$ – угол, отсчитываемый от оси ox_2 – определяет ориентацию в пространстве вектора в направлении от $x^n = (x_1^n, x_2^n)^T$ к точечному ориентиру. Сам этот вектор задает направление градиента для измеряемого навигационного параметра (в данном случае дальности) – направление наибольшего его изменения. Компоненты $H_i^T(x^n)$ соответствуют производным по направлениям вдоль координат ox_1 и ox_2 . Суть линеаризации в этом примере заключается в том, что изолинии, представляющие собой, как отмечалось выше, окружности, заменяются (см. рис. 1.4) прямыми, которые называются **линиями положения**

Для величины (1.27) в этом примере справедливо следующее выражение:

$$\delta = \frac{d^2 D}{dx_1^2} (x_1 - x_1^n)^2 + \frac{d^2 D}{dx_2^2} (x_2 - x_2^n)^2 + \frac{d^2 D}{dx_1 dx_2} (x_1 - x_1^n)(x_2 - x_2^n),$$

которое при $x_1 - x_1^n \approx x_2 - x_2^n \approx \Delta$ можно записать в виде

$$\delta = \left(\frac{d^2 D}{dx_1^2} + \frac{d^2 D}{dx_2^2} + \frac{d^2 D}{dx_1 dx_2} \right) \Delta^2.$$

Считая для простоты, что одна из координат известна, а направление прямой, соединяющей координаты точечного ориентира с точкой линеаризации, совпадает с направлением неизвестной координаты, для величины δ получаем

$$\delta = \frac{d^2 D}{dx_1^2} \Delta^2 = \frac{\Delta^2}{D} = \frac{\Delta}{D} \Delta. \quad (1.32)$$

Отсюда следует, что погрешность линеаризованного представления зависит от отношения уровня возможной ошибки в координате к величине дальности до точечного ориентира. Само значение δ при этом следует сравнивать с уровнем ошибки измерения дальности. Оценивая возможную погрешность линеаризованного представления при решении задачи определения координат по спутниковым данным с учетом высоты орбиты в 20000 км, можно констатировать, что при неточном задании точки линеаризации с ошибкой в 1–10 км значение $\delta \approx (0,5 - 5)$ м. Эта величина соизмерима с уровнем ошибки измерения дальности в современных спутниковых системах.

Задачи к разделу

Задача 1.1. Пусть имеется набор измерений

$$y_i = x_1 + x_2 + v_i, \quad i = \overline{1, m}. \quad (1)$$

Запишите постановку задачи в виде (1.11) в случае, когда по измерениям (1) необходимо найти только x_1 , а сумма $\varepsilon_i = x_2 + v_i$ трактуется как ошибка. Сделайте то же самое, полагая, что оцениванию подлежит x_1 и x_2 .

Задача 1.2. Пусть в дискретные моменты $t_i, i = \overline{1, m}$ заданы измерения

$$y_i = A \sin(\omega t_i + \phi_0) + v_i, \quad i = \overline{1, m},$$

где A – амплитуда; $\omega = 2\pi f$ – круговая частота; ϕ_0 – фаза.

Считая частоту и фазу известными, запишите постановку задачи оценивания амплитуды в виде (1.11).

Задача 1.3. Пусть в дискретные моменты времени t_i , $i = \overline{1,m}$ заданы измерения $y_i = x_1 + A \sin \omega t_i + v_i$, в которых круговая частота $\omega = 2\pi f$ считается известной, а x_1 и A – неизвестные величины.

Запишите постановку задачи оценивания в виде (1.11) в случае, когда по этим измерениям ставится задача найти только амплитуду, а сумма $\varepsilon_i = x_1 + v_i$ трактуется как ошибка. Сделайте то же самое, полагая, что оцениванию подлежит x_1 и A .

Задача 1.4. Пусть заданы измерения, как в задаче 1.2, и требуется найти фазу при известных значениях частоты и амплитуды. Запишите линеаризованную постановку задачи, принимая в качестве точки линеаризации $\bar{\phi}_0$. Сделайте то же самое в случае, если стоит задача нахождения частоты при известных значениях фазы и амплитуды.

Задача 1.5. Задача слежения за подвижным объектом – (задача траекторного измерения) (рис.1.6). Пусть на плоскости известны координаты точки x_1^*, x_2^* , из которой до подвижного объекта, осуществляющего прямолинейное движение с постоянной скоростью, в дискретные моменты времени могут быть проведены измерения дальностей ρ_i и пеленгов θ_i , с ошибками $\delta\rho_i$ и $\delta\theta_i$, $i = \overline{1,m}$. Сформулируйте задачу оценивания составляющих скорости и координат подвижного объекта в начальный момент времени по измерениям (1.20), (1.21), полагая, что имеются измерения дальностей (а); пеленгов (б); дальностей и пеленгов – задача комплексной обработки (в).

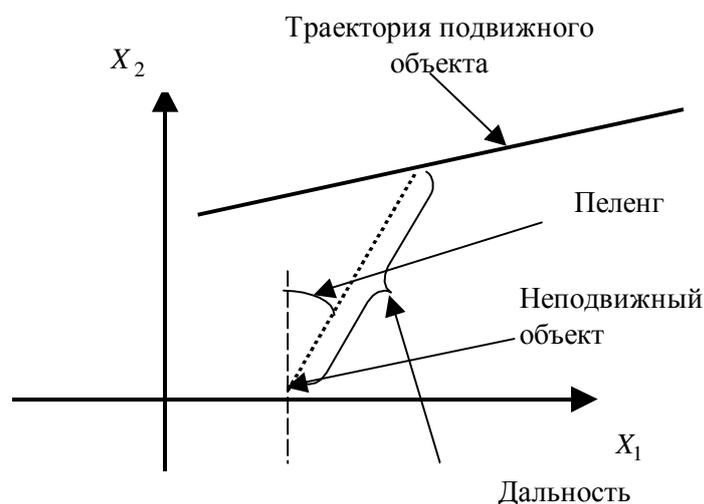


Рис. 1.6. К задаче определения координат подвижного объекта по измерениям дальностей и пеленгов

Задача 1.6. Сформулируйте задачу слежения за подвижным объектом, полагая, что выполнены условия задачи 1.5, а ошибки измерения дальности

стей и пеленгов имеют, кроме того, подлежащие определению постоянные составляющие.

Задача 1.7. Сформулируйте в линеаризованном виде две предыдущие задачи как одну задачу комплексной обработки измерений дальностей и пеленгов в целях оценивания составляющих скорости и координат подвижного объекта.

Контрольные вопросы

1. Поясните смысл и приведите соответствующие постановки линейных задач оценивания постоянной скалярной величины, коэффициентов линейного и квадратичного трендов, одномерных координат и скорости при равномерном движении объекта, а также задачи оценивания коэффициентов полинома.
2. Поясните смысл и приведите математическую постановку задачи оценивания при выставке ИНС в ее простейшем варианте.
3. Поясните смысл и приведите математическую постановку задачи оценивания сдвига реализаций. Проиллюстрируйте ее на примере решения задачи оценивания фазы гармонического сигнала. Почему эта задача является нелинейной?
4. Поясните смысл и приведите математические постановки задач оценивания частоты гармонического сигнала и фазовой автоподстройки частоты.
5. Поясните смысл и приведите соответствующие математические постановки нелинейных задач оценивания координат на плоскости и в пространстве по измерениям дальностей до точечных ориентиров при наличии и отсутствии постоянных составляющих ошибок измерения. Что такое мешающие параметры?
6. Сформулируйте в общем виде постановку задачи оценивания постоянного вектора по зашумленным измерениям и проиллюстрируйте ее на примерах. Поясните, в чем особенности задач синтеза алгоритмов и анализа их точности.
7. Поясните смысл процедуры линеаризации и сформулируйте соответствующую ей линеаризованную постановку задачи на примере определения координат по точечным ориентирам и слежения за подвижным объектом по измерениям дальностей на плоскости.

РАЗДЕЛ 2

Решение задач оценивания на основе детерминированного подхода. Метод наименьших квадратов

В настоящем разделе рассмотрим так называемый **детерминированный (нестохастический) подход**, в котором при решении задач оценивания не вводятся предположения о случайности искомого вектора x и ошибок измерений v , и, следовательно, отсутствует необходимость в привлечении какой-либо априорной информации статистического характера [6, 21, 25, 27, 31-34].

2.1. Основные положения и постановка задачи в методе наименьших квадратов

Отличительная особенность этого подхода заключается в том, что задача синтеза алгоритма, т.е. получения процедуры вычисления оценки неизвестного вектора x по измерениям y , основана на выборе таких ее значений, которые обеспечивают минимизацию некоторого выбранного критерия, характеризующего меру близости между измеренными и вычисленными значениями $s(x)$ или Hx . В качестве простейшего варианта такого критерия может быть выбрана функция

$$J^{\text{МНК}}(x) = (y - s(x))^T (y - s(x)) = \sum_{i=1}^m (y_i - s_i(x))^2. \quad (2.1)$$

Разности $\mu_i = y_i - s_i(x)$ обычно называют **невязками измерений**.

Подход, основанный на минимизации критерия типа (2.1), получил наименование **метода наименьших квадратов (МНК) (least squares method (LSM))** [22-23, 30-33]. В дальнейшем алгоритмы, основанные на минимизации других критериев, аналогичных (2.1), также будем называть алгоритмами МНК.

Критерий (2.1) и соответствующая ему оценка

$$\hat{x}^{\text{МНК}}(y) = \arg \min_x (y - s(x))^T (y - s(x)) \quad (2.2)$$

имеют вполне понятный смысл – выбрать такое значение искомого параметра, при котором минимизируется сумма квадратов отклонений вычисляемых величин от их измеренных значений, т.е. минимизируется сумма квадратов невязок измерений.

Определяя производную для $J^{\text{МНК}}(x)$ в соответствии с приведенными в приложении правилами (П1.63) и учитывая необходимое условие минимума, можем записать так называемую **систему нормальных уравнений**:

$$\frac{dJ^{\text{МНК}}(x)}{dx} = -2 \frac{ds^T(x)}{dx} (y - s(x)) = 0. \quad (2.3)$$

Следует напомнить, что (2.3) является лишь необходимым условием, и для обеспечения локального минимума требуется проверять справедливость выполнения достаточного условия

$$\frac{\partial^2}{\partial x \partial x^T} J^{\text{МНК}}(x) \Big|_{\hat{x}^{\text{МНК}}(y)} \geq 0. \quad (2.4)$$

Вместо (2.1) в качестве критерия может быть использована функция

$$J^{\text{ОМНК}}(x) = (y - s(x))^T Q (y - s(x)), \quad (2.5)$$

в которой Q – некоторая симметричная неотрицательно определенная матрица. Так, если считать Q диагональной матрицей с элементами q_i , $i = \overline{1, m}$, то вместо (2.1) будем иметь

$$J^{\text{ОМНК}}(x) = \sum_{i=1}^m q_i (y_i - s_i(x))^2.$$

Смысл введения весовой матрицы Q заключается в том, чтобы обеспечить возможность по-разному учитывать вклад отличий измеренных и вычисленных значений, соответствующих различным компонентам вектора измерений. В этом случае говорят об **обобщенном методе наименьших квадратов (ОМНК)**. Его иногда называют **методом взвешенных наименьших квадратов (weighted least squares method)**.

В наиболее общем случае, вводя симметричную неотрицательно определенную матрицу D и некоторый известный вектор $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)^T$, в качестве минимизируемого критерия можно использовать функцию

$$J^{\text{ММНК}}(x) = (y - s(x))^T Q (y - s(x)) + (x - \bar{x})^T D (x - \bar{x}), \quad (2.6)$$

которая при диагональном характере матриц Q и D примет вид

$$J^{\text{ММНК}}(x) = \sum_{i=1}^m q_i (y_i - s_i(x))^2 + \sum_{j=1}^n d_j (x_j - \bar{x}_j)^2.$$

Будем называть далее метод, основанный на минимизации критерия типа (2.6), **модифицированным методом наименьших квадратов (ММНК)**. Здесь \bar{x}_j , $j = \overline{1, n}$ представляют собой некоторые заранее известные возможные значения оцениваемых величин. Таким образом, смысл

второго дополнительного слагаемого заключается в том, что при отличии получаемых оценок $\hat{x}_j(y)$ от \bar{x}_j устанавливается определенный штраф, уровень которого задается коэффициентами d_j , $j = \overline{1..n}$. Эти коэффициенты определяют уровень «доверия» к априорной информации о возможных значениях x_j , заданных \bar{x}_j , $j = \overline{1..n}$.

Из сказанного следует, что нахождение оценок, соответствующих методу наименьших квадратов или одному из его обобщенных вариантов, сводится к отысканию минимума функций (2.1), (2.5) или (2.6).

◆ **Пример 2.1.** Пусть требуется оценить амплитуду A гармонического колебания по измерениям (1.14) в предположении, что фаза и частота известны.

Рассмотрим, к чему сводится решение этой задачи с использованием МНК. Критерий (2.1) при $x = A$ в этой задаче запишется как

$$J^{\text{МНК}}(x) = \sum_{i=1}^m (y_i - x \sin(\omega t_i + \phi_0))^2.$$

Оценка, соответствующая данному критерию, отыскивается достаточно просто. Действительно,

$$\frac{dJ^{\text{МНК}}(x)}{dx} = 2 \sum_{i=1}^m (y_i - x \sin(\omega t_i + \phi_0)) \sin(\omega t_i + \phi_0) = 0,$$

откуда

$$\hat{x}^{\text{МНК}} = \frac{1}{\sum_{i=1}^m \sin^2(\omega t_i + \phi_0)} \sum_{i=1}^m \sin(\omega t_i + \phi_0) y_i = \sum_{i=1}^m \tilde{q}_i y_i,$$

где \tilde{q}_i – коэффициенты, определяемые как

$$\tilde{q}_i = \frac{\sin(\omega t_i + \phi_0)}{\sum_{i=1}^m \sin^2(\omega t_i + \phi_0)}.$$

Ясно, что выполнено и достаточное условие (2.4), поскольку

$$\frac{d^2 J^{\text{МНК}}(x)}{dx^2} = 2 \sum_{i=1}^m \sin^2(\omega t_i + \phi_0) \geq 0.$$

На рис. 2.1 представлены реализация измеренных значений гармонического колебания с частотой 1 рад/с, фазой $\frac{\pi}{2}$ и единичной амплитудой на интервале 2 с с шагом 0,02 с и реализации значений гармонического колебания без ошибок, вычисленные при тех же параметрах и трех значениях амплитуды $A_1 = 0.5, 0.5$, $A_2 = 1.0, 1.0$ и $A_3 = 1.5, 1.5$.

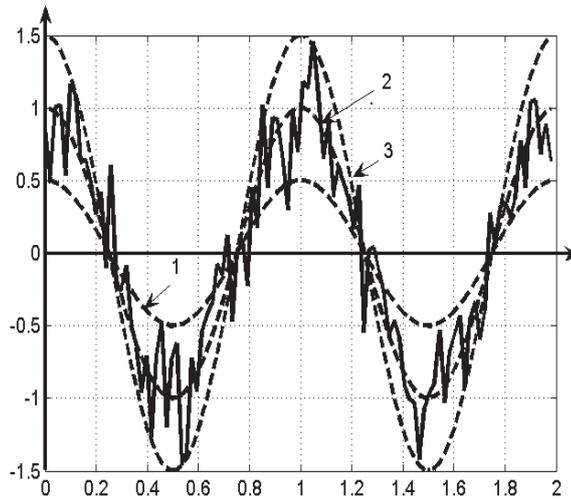


Рис. 2.1. Измеренные и вычисленные значения гармонического колебания при различных значениях амплитуды: 1 – 0,5; 2 – 1,0; 3 – 1,5

Из этого рисунка становится прозрачным геометрический смысл задачи нахождения оценки с помощью МНК – подобрать (найти) такое значение амплитуды, при котором будет достигаться наилучшее совпадение измеренных и вычисленных реализаций гармонического колебания.

Минимизируемый критерий $J^{\text{МНК}}(x) = \sum_{i=1}^m (y_i - x \sin(\omega t_i + \phi_0))^2$ для рассматриваемого примера изображен на рис. 2.2, из которого видно, что функция $J^{\text{МНК}}(x)$ представляет собой параболу, имеющую один экстремум в точке, которая и определяет значение оценки, соответствующей МНК.

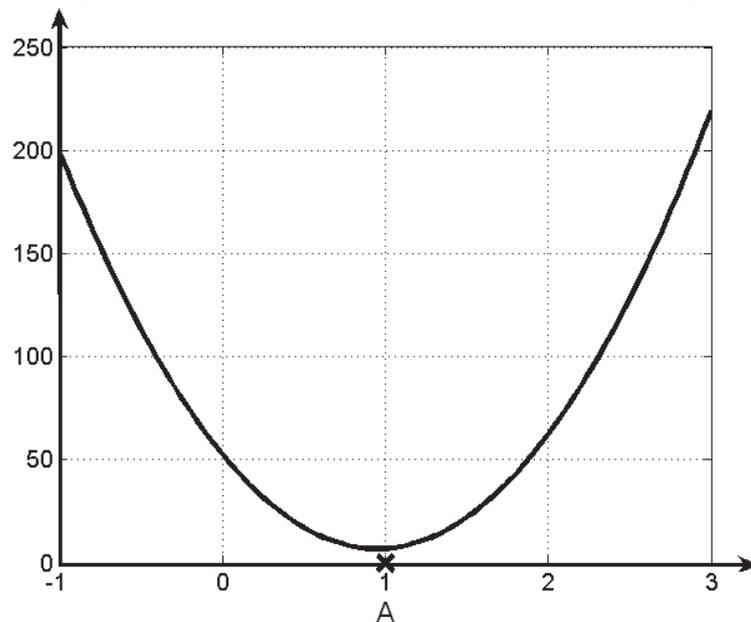


Рис. 2.2. Вид минимизируемого критерия $J^{\text{МНК}}(x)$

Значение самого минимизируемого критерия для трех значений амплитуды $A_1 = 0.5, 0.5$, $A_2 = 1.0, 1.0$ и $A_3 = 1.5, 1.5$ и минимальное значение критерия, достигаемое в рассматриваемом примере при $\hat{x} = 0.9448, 0.9448$, приведены в табл. 2.1.

Т а б л и ц а 2.1

Значения минимизируемого критерия при разных значениях амплитуды

$x = A$	0.5	1.0	1.5	$\hat{x} = \hat{A} = 0.9448$
$J^{\text{МНК}}(x)$	18.8	9.0	24.4	8.86

Из таблицы следует, что значение критерия достигает своей минимальной величины в точке, не совпадающей с истинным значением амплитуды. Это является следствием наличия ошибок измерения, порождающих отличие оценки от истинного значения отыскиваемой величины, что вполне объяснимо, поскольку выражение для оценки может быть записано как

$$\hat{x}^{\text{МНК}} = \frac{\sum_{i=1}^m \sin(\omega t_i + \phi_0) (x \sin(\omega t_i + \phi_0) + v_i)}{\sum_{i=1}^m \sin^2(\omega t_i + \phi_0)} = x + \sum_{i=1}^m \tilde{q}_i v_i,$$

и, таким образом, для ошибки оценки справедливо выражение

$$\hat{x}^{\text{МНК}} - x = \sum_{i=1}^m \tilde{q}_i v_i.$$

Нетрудно получить решение рассматриваемой задачи и для случая ОМНК. Так, при диагональном виде матрицы Q и $x = A$ критерий (2.5) запишется как

$$J^{\text{ОМНК}}(x) = \sum_{i=1}^m q_i (y_i - x \sin(\omega t_i + \phi_0))^2.$$

Оценка, соответствующая этому критерию, находится достаточно просто. Действительно,

$$\frac{dJ^{\text{ОМНК}}(x)}{dx} = 2 \sum_{i=1}^m q_i (y_i - x \sin(\omega t_i + \phi_0)) \sin(\omega t_i + \phi_0) = 0,$$

откуда

$$\hat{x}^{\text{ОМНК}} = \frac{1}{\sum_{i=1}^m q_i \sin^2(\omega t_i + \phi_0)} \sum_{i=1}^m q_i \sin(\omega t_i + \phi_0) y_i = \sum_{i=1}^m \tilde{q}_i y_i,$$

где \tilde{q}_i – коэффициенты, отыскиваемые в виде

$$\tilde{q}_i = \frac{q_i \sin(\omega t_i + \phi_0)}{\sum_{i=1}^m q_i \sin^2(\omega t_i + \phi_0)}.$$

Не представляет особого труда найти оценку и для ММНК, поскольку и в этом случае решение задачи также сведется к нахождению расположения точки экстремума параболы

$$J^{\text{ММНК}}(x) = J^{\text{ОМНК}}(x) + d(x - \bar{x})^2. \quad \blacklozenge$$

2.2. Общее решение задачи синтеза алгоритмов на основе метода наименьших квадратов в линейном случае

Простота получения алгоритма вычисления оценки, т.е. задачи синтеза, в рассмотренном выше примере является следствием того факта, что измерения линейным образом зависят от неизвестной амплитуды и ошибок измерения. Рассмотрим, к чему сводится алгоритм решения задачи оценивания в линейной постановке в общем случае, т.е. тогда, когда измерения могут быть представлены в виде (1.11). Для обычного МНК при таких измерениях минимизируемый критерий (2.1) запишется как

$$J^{\text{МНК}}(x) = (y - Hx)^T (y - Hx). \quad (2.7)$$

Нетрудно заметить, что относительно x функция (2.7) представляет собой квадратичную форму

$$J^{\text{МНК}}(x) = x^T H^T H x - 2x^T H^T y + y^T y, \quad (2.8)$$

которая при невырожденной $H^T H$ имеет единственный экстремум, и при этом достаточное условие (2.4) выполнено.

Принимая во внимание соотношение (П1.61), систему нормальных уравнений, соответствующую критерию (2.7), можно записать в виде

$$H^T (y - Hx) = 0.$$

Условие невырожденности $H^T H$ будем называть **условием наблюдаемости**. Использование этого термина вполне оправданно, поскольку в этом случае имеем:

$$\hat{x}^{\text{МНК}}(y) = (H^T H)^{-1} H^T y, \quad (2.9)$$

или

$$\hat{x}^{\text{МНК}}(y) = K^{\text{МНК}} y, \quad (2.10)$$

где

$$K^{\text{МНК}} = (H^T H)^{-1} H^T. \quad (2.11)$$

Отсюда следует, что при отсутствии ошибок измерений и выполнении условия наблюдаемости

$$\hat{x}^{\text{МНК}}(y) = (H^T H)^{-1} H^T H x \equiv x, \quad (2.12)$$

т.е. оценка совпадает с истинным значением вектора.

Поступая аналогично при решении линейной задачи в случае ОМНК и считая невырожденной $H^T Q H$, запишем:

$$J^{\text{ОМНК}}(x) = (y - Hx)^T Q (y - Hx); \quad (2.13)$$

$$\hat{x}^{\text{ОМНК}}(y) = K^{\text{ОМНК}} y, \quad (2.14)$$

где

$$K^{\text{ОМНК}} = (H^T Q H)^{-1} H^T Q. \quad (2.15)$$

Заметим, что, как правило, матрица Q не вырождена, и, таким образом, при выполнении условия наблюдаемости матрица $H^T Q H$ также является невырожденной.

Для ММНК (см. задачу 2.1) получим следующий набор соотношений:

$$J^{\text{ММНК}}(x) = (y - Hx)^T Q (y - Hx) + (x - \bar{x})^T D (x - \bar{x}); \quad (2.16)$$

$$\hat{x}^{\text{ММНК}}(y) = \bar{x} + K^{\text{ММНК}}(y - H\bar{x}), \quad (2.17)$$

где

$$K^{\text{ММНК}} = (D + H^T Q H)^{-1} H^T Q. \quad (2.18)$$

Представленные соотношения сведены в табл. 2.2.

Т а б л и ц а 2.2

Наблюдаемые критерии в линейной задаче оценивания

Метод	Критерий	Алгоритм
МНК	$J^{\text{МНК}}(x) = (y - Hx)^T (y - Hx)$	$\hat{x}^{\text{МНК}}(y) = K^{\text{МНК}} y,$ $K^{\text{МНК}} = (H^T H)^{-1} H^T$
ОМНК	$J^{\text{ОМНК}}(x) = (y - Hx)^T Q (y - Hx)$	$\hat{x}^{\text{ОМНК}}(y) = K^{\text{ОМНК}} y,$ $K^{\text{ОМНК}} = (H^T Q H)^{-1} H^T Q$
ММНК	$J^{\text{ММНК}}(x) = (y - Hx)^T Q (y - Hx) + (x - \bar{x})^T D (x - \bar{x})$	$\hat{x}^{\text{ММНК}}(y) = \bar{x} + K^{\text{ММНК}}(y - H\bar{x}),$ $K^{\text{ММНК}} = (D + H^T Q H)^{-1} H^T Q$

Обращаем внимание на одно весьма важное обстоятельство – все получающиеся оценки линейным образом зависят от измерений. Это есть следствие линейности измерений (1.11) и квадратичного характера минимизируемых критериев (2.7), (2.13), (2.16).

Понятно, что с использованием приведенных соотношений для рассмотренного выше примера 2.1 (стр. 30) легко получить представленные в нем выражения для оценок. Для этого достаточно учесть вид матрицы H , которая, как нетрудно понять, записывается как $H^T = [\sin(\omega t_1), \sin(\omega t_2), \dots, \sin(\omega t_m)]$.

Рассмотрим еще два примера.

◆ **Пример 2.2.** Конкретизируем выражения для оценок в задаче оценивания неизвестной скалярной величины по скалярным измерениям

(1.1), выбирая в критерии (2.13) для простоты матрицу Q диагональной с элементами $q_i > 0, i = \overline{1, m}$, а в критерии (2.16) полагая $D = d$.

Очевидно, что минимизируемый критерий и в этом случае имеет вид параболы, и, поскольку матрица $H^T = [1, 1, \dots, 1]$, выражения для оценок будут определяться в соответствии с соотношениями, приведенными в табл. 2.3.

Из выражений в табл. 2.3 следует, что в МНК оценка представляет собой среднеарифметическое от измерений y_i . Для ОМНК оценка определяется в результате «взвешивания» измерений с нормированными коэффициентами:

$$\sum_{i=1}^m \tilde{q}_i^{\text{ОМНК}} = 1; \quad \tilde{q}_i^{\text{ОМНК}} = \frac{q_i}{\sum_{i=1}^m q_i}.$$

Для ММНК изменяется нормировка, что обусловлено наличием априорной информации о возможном значении оцениваемой величины x , равном \bar{x} , и добавляется дополнительное слагаемое.

Т а б л и ц а 2.3

Алгоритмы оценивания для трех вариантов МНК

для простейшего примера нахождения x по измерениям $y_i = x + v_i, i = \overline{1, m}$

Метод	Критерий	Алгоритм
МНК	$J^{\text{МНК}}(x) = \sum_{i=1}^m (y_i - x)^2$	$\hat{x}^{\text{МНК}} = \frac{1}{m} \sum_{i=1}^m y_i$
ОМНК	$J^{\text{ОМНК}}(x) = \sum_{i=1}^m q_i (y_i - x)^2$	$\hat{x}^{\text{ОМНК}} = \sum_{i=1}^m \tilde{q}_i^{\text{ОМНК}} y_i, \quad \tilde{q}_i^{\text{ОМНК}} = \frac{q_i}{\sum_{i=1}^m q_i}$
ММНК	$J^{\text{ММНК}}(x) = d(x - \bar{x})^2 + \sum_{i=1}^m q_i (y_i - x)^2$	$\hat{x}^{\text{ММНК}} = \beta \bar{x} + \sum_{i=1}^m \tilde{q}_i^{\text{ММНК}} y_i, \quad \beta = \frac{d}{d + \sum_{i=1}^m q_i}, \quad \tilde{q}_i^{\text{ММНК}} = \frac{q_i}{d + \sum_{i=1}^m q_i} \blacklozenge$

◆ **Пример 2.3.** Будем полагать, что требуется оценить начальное значения координаты x_0 и значение скорости V по измерениям типа (1.3).

Критерий МНК в этой задаче будет определяться как

$$J^{\text{МНК}}(x_0, V) = \sum_{i=1}^m (y_i - x_0 - V t_i)^2.$$

Принимая во внимание обозначения (1.4), легко убедиться в том, что этот критерий представляет собой квадратичную форму (2.8), которая может быть записана в виде

$$J^{\text{МНК}}(x) = (x_1, x_2)^T \begin{pmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 2 \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m t_i y_i \end{pmatrix} + \sum_{i=1}^m y_i^2$$

или

$$J^{\text{МНК}}(x) = mx_1^2 + x_2^2 \sum_{i=1}^m t_i^2 + 2x_1x_2 \left(\sum_{i=1}^m t_i \right) - 2 \left(x_1 \sum_{i=1}^m y_i + x_2 \sum_{i=1}^m t_i y_i \right) + \sum_{i=1}^m y_i^2.$$

Используя соотношение (2.9), получим следующее выражение:

$$\begin{bmatrix} \hat{x}_1^{\text{МНК}} \\ \hat{x}_2^{\text{МНК}} \end{bmatrix} = \begin{bmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m t_i y_i \end{bmatrix}. \quad (2.19)$$

Для ОМНК при диагональной матрице Q в критерии появятся сомножители q_i , и выражения для оценки преобразуются к виду

$$\hat{x}^{\text{ОМНК}} = \begin{bmatrix} \sum_{i=1}^m q_i & \sum_{i=1}^m q_i t_i \\ \sum_{i=1}^m q_i t_i & \sum_{i=1}^m q_i t_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^m q_i y_i \\ \sum_{i=1}^m q_i t_i y_i \end{bmatrix}.$$

Не представляет труда конкретизировать в этом примере и выражение для критерия и оценки применительно к ММНК.

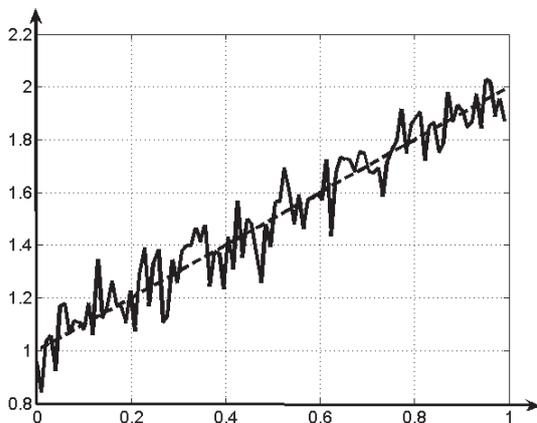


Рис. 2.3. Реализация измерений $y_{1i} = x_0 + Vt_i + v_i, i = \overline{1..m}$

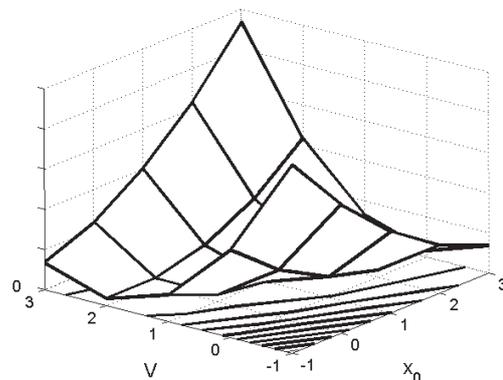


Рис. 2.4. Минимизируемый критерий $J^{\text{МНК}}(x_0, V)$

На рис. 2.3 представлена реализация измеренных значений координат на фоне графика, соответствующего их действительным значениям для случая, когда измерения проводятся на интервале 10 с с шагом в 1 с при $x_0 = 1$ м и $V = 1$ м/с. На рис. 2.4 приведена функция, описывающая поведение критерия $J^{\text{МНК}}(x)$ и представляющая собой вытянутый вдоль одной из осей параболоид, соответствующий квадратичной форме. ♦

2.3. Анализ точности в методе наименьших квадратов в линейном случае

При постановке задачи в разделе 1.1 отмечалось, что, помимо решения задачи синтеза алгоритмов, весьма важной является задача анализа точности, т.е. исследования свойств ошибок получаемых оценок. Для линейного случая с использованием соотношений (2.10), (2.14) для МНК и ОМНК могут быть получены явные выражения для ошибок оценок:

$$\begin{aligned}\varepsilon^{\text{МНК}}(y) &= x - \hat{x}^{\text{МНК}}(y) = x - K^{\text{МНК}} y = (E - K^{\text{МНК}} H)x - K^{\text{МНК}} v; \\ \varepsilon^{\text{ОМНК}}(y) &= x - \hat{x}^{\text{ОМНК}}(y) = x - K^{\text{ОМНК}} y = (E - K^{\text{ОМНК}} H)x - K^{\text{ОМНК}} v.\end{aligned}$$

Поскольку для МНК и ОМНК справедливы соотношения

$$E - K^{\text{МНК}} H = E - (H^T H)^{-1} H^T H = 0; \quad (2.20)$$

$$E - K^{\text{ОМНК}} H = E - (H^T QH)^{-1} H^T QH = 0, \quad (2.21)$$

ошибки, соответствующие этим двум методам, могут быть записаны так:

$$\varepsilon^{\text{МНК}}(y) = -K^{\text{МНК}} v; \quad (2.22)$$

$$\varepsilon^{\text{ОМНК}}(y) = -K^{\text{ОМНК}} v. \quad (2.23)$$

Для ММНК в силу того, что $(E - K^{\text{ММНК}} H) \neq 0$, уравнение для ошибок оценок примет вид

$$\varepsilon^{\text{ММНК}}(y) = x - \hat{x}^{\text{ММНК}}(y) = (E - K^{\text{ММНК}} H)(x - \bar{x}) - K^{\text{ММНК}} v. \quad (2.24)$$

Из полученных соотношений вытекает следующий вывод. Ошибки оценок, соответствующие МНК и ОМНК в линейном случае, не содержат слагаемых, зависящих от вектора оцениваемых параметров, а зависят только от ошибок измерения. В этих условиях можно говорить **об инвариантности (независимости) ошибок оценивания** относительно искомого вектора. Что касается ММНК, то для него ошибка оценивания зависит еще и от значения самого оцениваемого вектора, и, таким образом, она не обладает свойством инвариантности.

Достоинство рассмотренных методов заключается в том, что на этапе синтеза алгоритмов какая-либо априорная информация статистического характера не привлекается. Однако ее отсутствие затрудняет решение задачи анализа точности. Такая возможность открывается, если ввести пред-

положение о случайном характере как ошибок измерения (в случае МНК и ОМНК), так и оцениваемого вектора (в случае ММНК).

Введем такие предположения, полагая, в частности, что ошибки измерения v_i , $i = \overline{1..m}$ являются центрированными случайными величинами, для которых определена матрица ковариаций R . В этом случае из соотношений (2.22), (2.23) вытекает, что ошибки оценок для МНК и ОМНК также будут центрированными с матрицами ковариаций:

$$P^{\text{МНК}} = M \left\{ (K^{\text{МНК}} v)(K^{\text{МНК}} v)^T \right\} = (H^T H)^{-1} H^T R H (H^T H)^{-1}; \quad (2.25)$$

$$P^{\text{ОМНК}} = (H^T Q H)^{-1} H^T Q R Q H (H^T Q H)^{-1}. \quad (2.26)$$

Для того чтобы вычислить матрицу ковариаций ошибок оценивания для ММНК, необходимо не только дополнительно ввести предположение о случайном характере оцениваемого вектора и задать статистические свойства для векторов v и x по отдельности, но и определить их взаимные статистические свойства. Если, к примеру, считать, что x и v – не коррелированные между собой случайные векторы с математическими ожиданиями \bar{x} и $\bar{v} = 0$ и матрицами ковариаций R , P^x , то с использованием представления (2.24) нетрудно убедиться в том, что ошибки оценивания также являются центрированными, и для соответствующей им матрицы ковариаций можем получить соотношение

$$P^{\text{ММНК}} = (E - K^{\text{ММНК}} H) P^x (E - K^{\text{ММНК}} H)^T + K^{\text{ММНК}} R (K^{\text{ММНК}})^T, \quad (2.27)$$

где матрица $K^{\text{ММНК}}$ задается выражением (2.18).

Если в критерии (2.13) принять $Q = R^{-1}$, то оценки ОМНК и матрица ковариаций их ошибок могут быть записаны как

$$\hat{x}^{\text{ОМНК}}(y) = (H^T R^{-1} H)^{-1} H^T R^{-1} y, \quad (2.28)$$

$$P^{\text{ОМНК}} = (H^T R^{-1} H)^{-1}. \quad (2.29)$$

Если, кроме того, считать, что $D = (P^x)^{-1}$, то $K^{\text{ММНК}} = \left((P^x)^{-1} + H^T R^{-1} H \right)^{-1} H^T R^{-1}$, и выражения для оценки (2.17) и соответствующей ей матрице ковариаций (см. задачу 2.3) примут вид:

$$\hat{x}^{\text{ММНК}}(y) = \bar{x} + \left((P^x)^{-1} + H^T R^{-1} H \right)^{-1} H^T R^{-1} (y - H\bar{x}); \quad (2.30)$$

$$P^{\text{ММНК}} = \left((P^x)^{-1} + H^T R^{-1} H \right)^{-1}. \quad (2.31)$$

Возможность вычисления матриц ковариаций ошибок оценивания обеспечивает существенное продвижение в решении задачи анализа точности, поскольку удастся охарактеризовать ее на количественном уровне. В частности, диагональные элементы полученных матриц ковариаций ошибок

оценивания (2.25)–(2.27) представляют собой дисперсии ошибок оценивания компонент искомого вектора x .

Значения диагональных элементов матрицы P^x обычно называют **априорными дисперсиями ошибки оценивания**. Такое название вполне обосновано, если считать, что до проведения измерений в качестве априорной оценки целесообразно использовать значение математического ожидания \bar{x} . Сопоставление априорных дисперсий с соответствующими **апостериорными дисперсиями** (диагональными элементами матриц ковариаций (2.25)–(2.27)), т. е. с теми, которые получены после проведения измерений с использованием различных алгоритмов, позволяет оценить эффективность этих алгоритмов. Соответственно, матрицу ковариаций P^x называют **априорной матрицей**, а матрицы (2.25)–(2.27) – **апостериорными матрицами ковариаций ошибок оценивания**.

Если дополнительно, наряду с предположением об известных значениях математических ожиданий и матриц ковариаций, считать, что ошибки измерения и оцениваемый вектор – гауссовские, то в силу линейного характера преобразований (2.22)–(2.24) это обеспечивает гауссовский характер и ошибок оценивания для рассматриваемых методов. Иными словами, становится известной ф.п.р.в. для вектора ошибок оценивания. Наличие ф.п.р.в. обеспечивает полное описание их статистических свойств. В частности, для них могут быть рассчитаны такие характеристики, как вероятная ошибка, предельная ошибка, квантили и т.д. (см. Приложение 2).

◆ **Пример 2.4.** Будем считать, что в примере 2.2 (стр. 35) x является случайной величиной с математическим ожиданием \bar{x} и дисперсией σ_0^2 , а ошибки измерения – центрированные, не коррелированные между собой и с x случайные величины с дисперсиями r_i^2 , $i=\overline{1,m}$ (неравноточные измерения), а в частном случае $r_i^2 = r^2$, $i=\overline{1,m}$ (равноточные измерения). Получим выражения для ошибок оценивания и соответствующих им дисперсий для трех вариантов МНК, применительно к задаче оценивания скалярной величины x .

При сделанных предположениях R – диагональная матрица с элементами r_i^2 , $i=\overline{1,m}$; матрица $H^T = [1, 1, \dots, 1]$; Q – диагональная матрица с элементами $q_i = 1/r_i^2$ в первом случае, и $q_i = 1/r^2$, $i=\overline{1,m}$ – во втором. Таким образом, нетрудно получить представленные в табл. 2.4 выражения.

В табл. 2.5 приведены результаты расчета СКО ошибок оценивания в зависимости от числа измерений для ОМНК и ММНК при одинаковых дисперсиях ошибок измерения $r^2=1$, $m = \overline{1,10}$.

Таблица 2.4

Уравнения ошибок оценивания и значения их дисперсий для трех вариантов МНК в простейшем случае

Метод	Ошибки оценки	Дисперсии	
		Неравноточные измерения	Равноточные измерения
МНК	$\varepsilon_{\text{МНК}} = \frac{1}{m} \sum_{i=1}^m \nu_i$	$D_{\text{МНК}} = \frac{\sum_{i=1}^m r_i^2}{m^2}$	$r_i^2 = r^2$
ОМНК	$\varepsilon_{\text{ОМНК}} = \sum_{i=1}^m \tilde{q}_i \text{ОМНК} \nu_i, \tilde{q}_i \text{ОМНК} = \frac{q_i}{\sum_{i=1}^m q_i}$	$D_{\text{ОМНК}} = \left(\sum_{i=1}^m \frac{1}{r_i^2} \right)^{-1}$	$D_{\text{ОМНК}} = \frac{r^2}{m}$
ММНК	$\varepsilon_{\text{ММНК}}(y) = \beta(x - \bar{x}) + \sum_{i=1}^m \tilde{q}_i \text{ММНК} \nu_i,$ $\tilde{q}_i \text{ММНК} = \frac{q_i}{d + \sum_{i=1}^m q_i}, \beta = \frac{d}{d + \sum_{i=1}^m q_i}$	$D_{\text{ММНК}} = \left[\frac{1}{\sigma_0^2} + \sum_{i=1}^m \frac{1}{r_i^2} \right]^{-1}$	$D_{\text{ММНК}} = \frac{\sigma_0^2 r^2}{r^2 + \sigma_0^2 m}$

**Значения СКО ошибок оценивания в зависимости от числа измерений
для ОМНК и ММНК при разных значениях σ_0**

Метод	Номер измерения									
	1	2	3	4	5	6	7	8	9	10
ОМНК	1	0,7	0,58	0,5	0,45	0,41	0,38	0,35	0,33	0,32
ММНК $\sigma_0 \geq 10, \bar{x} = 0$	1	0,7	0,58	0,5	0,45	0,41	0,38	0,35	0,33	0,32
ММНК $\sigma_0 = 1, \bar{x} = 0$	0,7	0,5 8	0,5	0,45	0,41	0,38	0,35	0,33	0,32	0,3

Из таблицы следует, что для ОМНК и ММНК при $\sigma_0 \geq 10$ отличия в точности не проявляются. Влияние априорной информации в данном случае незначительно, поскольку $\sigma_0 \gg r$. При $\sigma_0 = r = 1$ это влияние существенно лишь при малом числе измерений. Фактически СКО для ОМНК совпадает с СКО для ММНК на предыдущем шаге. ♦

Конкретизируем теперь соотношения (2.25), (2.26), (2.31) для задачи оценивания двумерного вектора, рассмотренной в примере 2.3. (стр. 32).

♦ **Пример 2.5.** Получим выражения для матриц ковариаций ошибок оценивания коэффициентов полинома первой степени по измерениям $y_i = x_0 + Vt_i + v_i, i = \overline{1,m}$, считая, что ошибки измерения являются некоррелированными между собой центрированными случайными величинами с одинаковыми дисперсиями, равными r^2 , а оцениваемые коэффициенты представляют собой центрированные ($\bar{x} = 0$), не коррелированные между собой и с ошибками измерений случайные величины с матрицей ковариаций

$$P^x = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}.$$

С использованием соотношений (1.4), (2.25), (2.26) при диагональной $Q = \{q_i\}$ в критерии для ОМНК легко получить следующие выражения:

$$P^{\text{ММНК}} = r^2 \begin{bmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{bmatrix}^{-1}; \quad (2.32)$$

$$P^{\text{ОМНК}} = \begin{bmatrix} \sum_{i=1}^m q_i & \sum_{i=1}^m q_i t_i \\ \sum_{i=1}^m q_i t_i & \sum_{i=1}^m q_i t_i^2 \end{bmatrix}^{-1}. \quad (2.33)$$

Принимая $Q = R^{-1}$, а для ММНК дополнительно полагая $\bar{x} = 0$, $D = (P^x)^{-1}$ и учитывая тот факт, что $R = r^2 E$, E – единичная матрица, легко убедиться в том, что матрицы ковариаций ошибок оценивания в МНК и ОМНК в этом случае будут между собой совпадать, т.е. $P^{\text{ОМНК}} = P^{\text{МНК}}$, а для ММНК

$$P^{\text{ММНК}} = \begin{bmatrix} \frac{1}{\sigma_0^2} + \frac{m}{r^2} & \frac{1}{r^2} \sum_{i=1}^m t_i \\ \frac{1}{r^2} \sum_{i=1}^m t_i & \frac{1}{\sigma_1^2} + \frac{1}{r^2} \sum_{i=1}^m t_i^2 \end{bmatrix}^{-1} \quad \blacklozenge \quad (2.34)$$

2.4. Решение нелинейных задач оценивания.

Линеаризованные и итерационные алгоритмы

Выше были получены простые алгоритмы вычисления оценок, соответствующих МНК и его модификациям. При введении дополнительных предположений о случайном характере оцениваемого вектора и ошибок измерения также легко вычисляются и матрицы ковариаций ошибок оценивания, с помощью которых анализируется точность оценивания. Простота процедур вычисления оценок и их характеристик точности в данном случае является следствием линейного характера рассматриваемых задач. Решение проблемы оценивания существенно усложняется, если зависимость измерений от оцениваемых параметров нелинейная. В то же время, как уже отмечалось, достаточно широкий круг нелинейных задач, связанных с обработкой навигационной информации, может эффективно решаться с использованием полученных выше алгоритмов, применение которых в нелинейном случае основано на описанной в подразделе 1.7 процедуре линеаризации. Обсудим этот вопрос более подробно.

Проведя линеаризацию, исходную нелинейную задачу легко свести к линейной постановке, в которой в качестве измерения фигурируют значения (1.25), представляемые в виде (1.26). С использованием этого соотношения легко получить линеаризованные варианты МНК и его модификации. Эффективность создаваемых при этом алгоритмов в немалой степени зависит от того, насколько удачно выбрана точка линеаризации. Чем ближе эта точка к истинному значению оцениваемого параметра, тем точнее будет линеаризованное представление и тем точнее будет оценка, полученная на основе линеаризованных алгоритмов. Таким образом, для повышения эффективности алгоритмов, основанных на линеаризации, точку линеаризации целесообразно выбирать как можно ближе к истинному неизвестному значению оцениваемого параметра. Из сказанного следует, что для повышения эффективности линеаризованных алгоритмов разумно использовать достаточно очевидный прием. Суть его заключается в много-

кратной повторной обработке измерений и использовании получаемых результатов для уточнения расположения точки линеаризации. Поясним смысл этой процедуры.

Выбрав начальную точку линеаризации x^n и используя приближение

$$s(x) \approx s(x^n) + \left. \frac{ds}{dx^T} \right|_{x=x^n} (x - x^n) = s(x^n) + H(x^n)(x - x^n),$$

сформируем первоначальную оценку искомого вектора с помощью соотношения

$$\hat{x}^{(1)} = \bar{x} + K(x^n) \left[y - s(x^n) - H^{(1)}(x^n)(\bar{x} - x^n) \right], \quad (2.35)$$

в котором $H^{(1)}(x^n) = \left. \frac{ds}{dx^T} \right|_{x=x^n}$, а $K(x^n)$ вычисляется в соответствии с пра-

вилами, соответствующими выбранному варианту МНК. Повторим эту операцию до тех пор, пока значение оценки не перестанет сколько-нибудь существенно изменяться. В общем виде данный алгоритм для случая ММНК имеет вид:

$$\hat{x}^{(\gamma+1)} = \bar{x} + K(\hat{x}^{(\gamma)}) \left[y - s(\hat{x}^{(\gamma)}) - H^{(\gamma)}(\hat{x}^{(\gamma)})(\bar{x} - \hat{x}^{(\gamma)}) \right]; \quad (2.36)$$

$$K(\hat{x}^{(\gamma)}) = P(\hat{x}^{(\gamma)}) H^T(\hat{x}^{(\gamma)}) R^{-1}; \quad (2.37)$$

$$P(\hat{x}^{(\gamma)}) = \left((P^x)^{-1} + H^T(\hat{x}^{(\gamma)}) R^{-1} H(\hat{x}^{(\gamma)}) \right)^{-1}; \quad (2.38)$$

$$\gamma = 0, 1, 2, \dots, \quad \hat{x}^{(0)} = \bar{x}.$$

Для случая ОМНК в этих выражениях следует положить $(P^x)^{-1} = 0$ и $\bar{x} = 0$ а для МНК, кроме того, $R^{-1} = E$. Такие алгоритмы в теории оценивания получили наименование **итерационных алгоритмов (iterated algorithm)** или **алгоритмов с локальными итерациями** [20, 21, 25, 32, 34].

Заметим, что в силу выполнения равенства типа $KH = E$ при использовании МНК или ОМНК выражение для оценки (2.36) может быть представлено в следующей, удобной при практической реализации, рекуррентной относительно номера итерации форме: $\hat{x}^{(\gamma+1)} = \hat{x}^{(\gamma)} + \delta \hat{x}^{(\gamma+1)}$, где $\delta \hat{x}^{(\gamma+1)} = K(\hat{x}^{(\gamma)}) \left[y - s(\hat{x}^{(\gamma)}) \right]$.

Если полагать, что ошибки измерения представляют собой центрированный случайный вектор с матрицей ковариаций R , а оцениваемый вектор является случайным вектором с математическим ожиданием \bar{x} и матрицей ковариаций P^x , то выражение (2.38) с точностью до предположения о справедливости линеаризованного описания будет определять **расчетную матрицу ковариаций ошибок оценивания**. Этот термин используется в связи с тем, что матрица ковариаций, получаемая в предположении о справедливости линеаризованного представления функции $s(x)$, в общем

случае отличается от **действительной матрицы ковариаций**, которая для произвольной оценки $\tilde{x}(y)$, в том числе и для оценки, вырабатываемой с помощью линеаризованного или итерационного алгоритмов, задается как

$$\tilde{P} = M \left\{ (x - \tilde{x}(y))(x - \tilde{x}(y))^T \right\}. \quad (2.39)$$

В этой связи встает **проблема адекватности**, т.е. соответствия расчетных матриц ковариаций их действительным значениям.

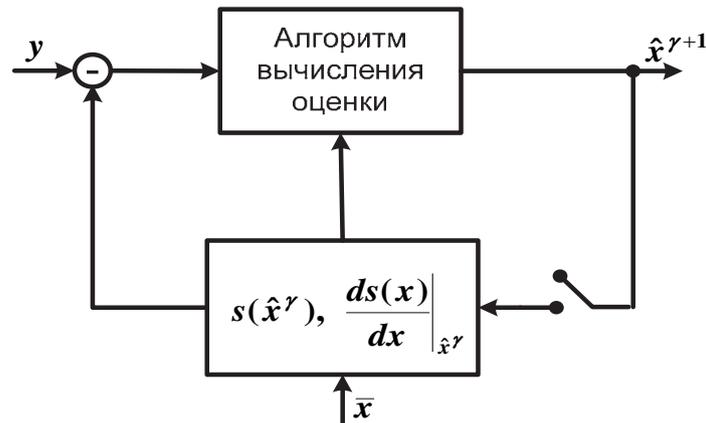


Рис. 2.5. Блок-схема итерационного алгоритма оценивания

Важно также подчеркнуть, что получаемые в результате алгоритмы перестают быть линейными относительно измерений, поскольку появляется нелинейная зависимость от измерений, в силу того, что от измерений зависят матрицы $K(\hat{x}^{(\gamma)}) = K(\hat{x}^{(\gamma)}(y))$. Зависимость от измерений появляется также и для расчетной матрицы ковариаций $P(\hat{x}^{(\gamma)}) = P(\hat{x}^{(\gamma)}(y))$.

Блок-схема, соответствующая итерационному алгоритму, представлена на рис. 2.5.

◆ **Пример 2.6.** Получим основанные на линеаризации обычный и итерационный алгоритмы, соответствующие МНК, в задаче оценивания фазы гармонического колебания в предположении, что амплитуда и частота известны. Для простоты считаем амплитуду единичной.

Критерий (2.1) для МНК при $x = \phi_0$ в этом случае запишется как

$$J^{\text{МНК}}(x) = \sum_{i=1}^m (y_i - \sin(\omega t_i + x))^2. \quad (2.40)$$

Для нахождения оценки, соответствующей критерию (2.40), следует найти точку расположения на оси абсцисс его минимального значения. Для этого необходимо решить нормальное уравнение, соответствующее необходимому условию существования экстремума, которое в данном случае примет вид

$$\frac{dJ^{\text{МНК}}(x)}{dx} = 2 \sum_{i=1}^m (y_i - \sin(\omega t_i + x)) \cos(\omega t_i + x) = 0.$$

Ясно, что сведение задачи минимизации критерия (2.40) к задаче нахождения корней этого нелинейного уравнения не упрощает ее решения. Используем линейризованное описание функций $\sin(\omega t_i + x)$:

$$\sin(\omega t_i + x) \approx \sin(\omega t_i + x^n) + (x - x^n) A \cos(\omega t_i + x^n), \quad i = \overline{1, m},$$

где x^n – выбранная точка линейризации. Тогда критерий в окрестности точки линейризации будет представлять собой параболу:

$$J^{\text{мнк}}(x) \approx \sum_{i=1}^m (\tilde{y}_i(x^n) - (x - x^n) \cos(\omega t_i + x^n))^2, \quad (2.41)$$

где $\tilde{y}_i(x^n) = y_i - \sin(\omega t_i + x^n)$.

Примеры графиков для минимизируемого критерия, соответствующего исходной нелинейной функции и ее линейризованному описанию при истинном значении фазы $\phi_0 = \frac{\pi}{2}$ и $x^n = \bar{\phi}_0 = \frac{3\pi}{4}$, изображены на рис. 2.6.

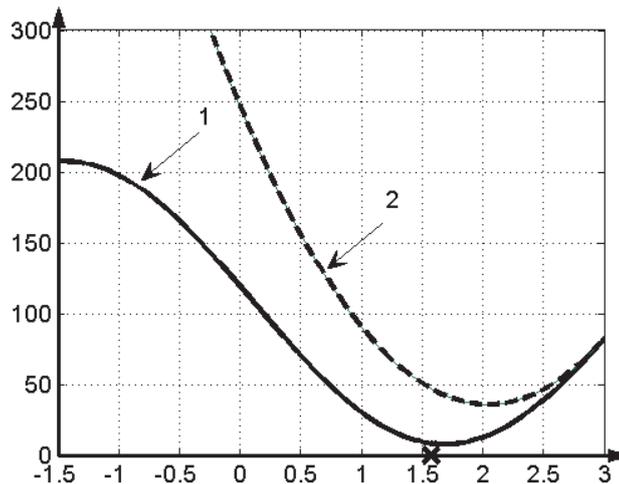


Рис. 2.6. График критерия $J^{\text{мнк}}(\phi_0)$ (1) и его приближенного описания (2) в задаче оценивания фазы

Эта оценка соответствует точке минимума параболы (2.41). Полагая дополнительно, что ошибки измерения являются некоррелированными между собой случайными величинами с одинаковой дисперсией r^2 , нетрудно найти выражение и для соответствующей этой оценке расчетной дисперсии

$$(\sigma^{\text{мнк}}(\bar{\phi}_0))^2 = \frac{r^2}{\sum_{i=1}^m \cos^2(\omega t_i + \bar{\phi}_0)}.$$

Принимая в качестве следующей точки линейризации $x^n = \hat{\phi}_0^{(1)}$ и проводя повторную обработку измерения, получаем более точное описание поведения минимизируемого критерия в окрестности точки экстремума, а, следо-

вательно, и более точное значение оценки и соответствующей ей расчетной дисперсии. Общие выражения для оценки (2.36) и ее расчетной дисперсии, соответствующие итерационному алгоритму в рассматриваемой задаче, конкретизируются как

$$\hat{\phi}_0^{(\gamma+1)} = \bar{\phi}_0 + \frac{\sum_{i=1}^m \cos(\omega t_i + \hat{\phi}_0^{(\gamma)}) \left(y_i - \sin(\omega t_i + \hat{\phi}_0^{(\gamma)}) - \cos(\omega t_i + \hat{\phi}_0^{(\gamma)}) (\bar{\phi}_0 - \hat{\phi}_0^{(\gamma)}) \right)}{\sum_{i=1}^m \cos^2(\omega t_i + \hat{\phi}_0^{(\gamma)})}, \quad (2.42)$$

$$(\sigma^{\text{МНК}}(\hat{\phi}_0^{(\gamma)}))^2 = \frac{r^2}{\sum_{i=1}^m \cos^2(\omega t_i + \hat{\phi}_0^{(\gamma)})}.$$

Используя такую процедуру вычисления оценки до тех пор, пока она не перестает существенным образом изменяться, получаем оценку со значительно меньшей ошибкой, чем для одной итерации, т.е. при использовании линеаризованного алгоритма.

Описанная ситуация поясняется с помощью приводимых ниже рисунков и результатов вычисления оценки, ее ошибки и расчетной дисперсии для нескольких итераций, представленных в табл. 2.6. Предполагалось, что истинное значение фазы $\phi_0 = \frac{\pi}{2}$, а точка линеаризации $\phi_0^{\text{л}} = \frac{3\pi}{4}$.

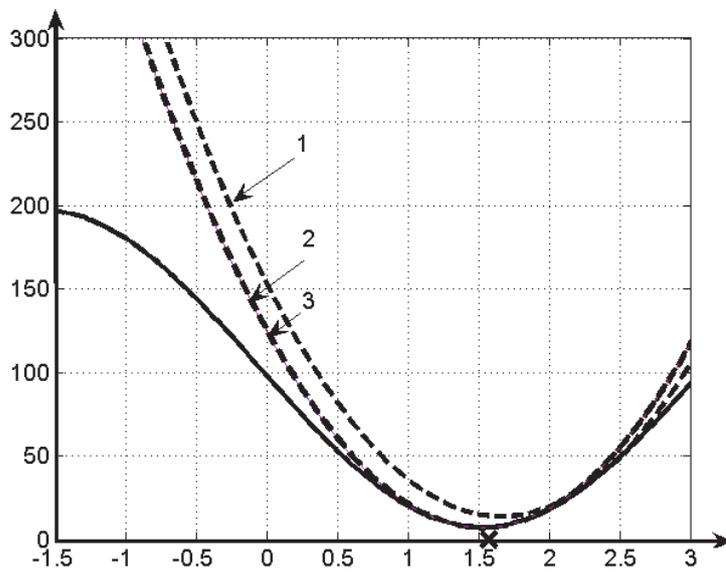


Рис. 2.7. Графики критерия $J^{\text{МНК}}(\phi_0)$ и его приближенного описания при разном числе итераций

Заметим, что выражение для расчетной дисперсии практически не меняется в силу малого изменения значения производной функции $\sin(\omega t_i + x)$ при изменении точки линеаризации.

Значения оценок $\hat{\phi}_0^{(\gamma)}$, их ошибок $\varepsilon^{(\gamma)}$ и расчетных СКО $\sigma^{\text{МНК}}(\hat{\phi}_0^{(\gamma)})$ при использовании итерационного алгоритма в задаче оценивания фазы

$\phi_0^{\text{л}} = \frac{3\pi}{4}$ истинное значение $\phi_0 = \frac{\pi}{2}$			
Номер итерации	$\hat{\phi}_0^{(\gamma)}$	$\varepsilon^{(\gamma)}$	$\sigma^{\text{МНК}}(\hat{\phi}_0^{(\gamma)})$
1	1,6499	0,0791	0,0422
2	1,5233	-0,0475	0,0422
3	1,5164	-0,0544	0,0422
4	1,5161	-0,0547	0,0422
5	1,5161	-0,0547	0,0422

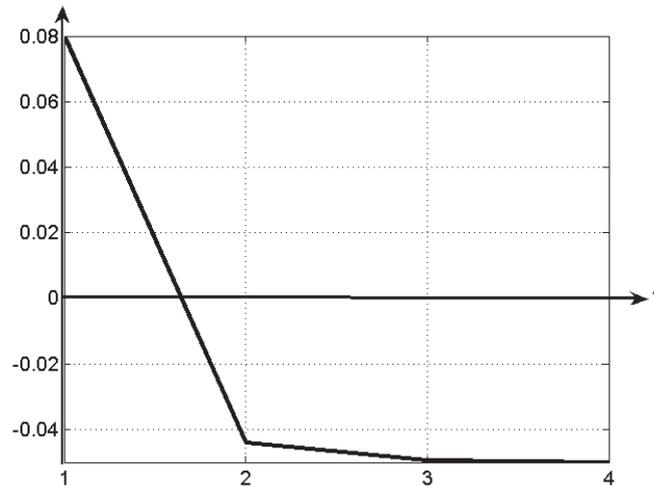


Рис. 2.8. Графики зависимости ошибок оценок фазы от номера итераций

Используя рекуррентную относительно номера итерации процедуру, можем записать

$$\hat{\phi}_0^{(\gamma+1)} = \hat{\phi}_0^{(\gamma)} + \delta\hat{\phi}_0^{(\gamma+1)},$$

где

$$\delta\hat{\phi}_0^{(\gamma+1)} = \frac{\sum_{i=1}^m \cos(\omega t_i + \hat{\phi}_0^{(\gamma)}) (y_i - \sin(\omega t_i + \hat{\phi}_0^{(\gamma)}))}{\sum_{i=1}^m \cos^2(\omega t_i + \hat{\phi}_0^{(\gamma)})}. \quad \blacklozenge$$

Линеаризованный и итерационный алгоритмы широко используются при решении задачи определения координат по точечным ориентирам. В частности, конкретизируем эти алгоритмы для примера решения такой задачи на плоскости.

◆ **Пример 2.7.** Будем полагать, что есть измерения (1.16) до двух точечных ориентиров, для простоты считая, что один из них расположен на оси ox_1 , а второй – на оси ox_2 . Для начала запишем алгоритм МНК, основанный на линеаризации.

Предварительно заметим, что минимизируемый критерий имеет вид

$$J^{\text{МНК}}(x_1, x_2) = \sum_{i=1}^2 \left(y_i - \sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2} \right)^2, \quad (2.43)$$

а его приближение, соответствующее линеаризованному описанию функции $s_i(x) = D_i(x) = \sqrt{(x_1^i - x_1)^2 + (x_2^i - x_2)^2}$ в окрестности точки линеаризации, представляет собой параболоид

$$J^{\text{МНК}}(x_1, x_2) \approx \sum_{i=1}^2 \left(\tilde{y}_i - H_{i1}(x)(x_1 - x_1^n) - H_{i2}(x)(x_2 - x_2^n) \right)^2, \quad (2.44)$$

где $H_{i1}(x^n) = -\sin \Pi_i(x^n)$, $H_{i2}(x^n) = -\cos \Pi_i(x^n)$,

$$\tilde{y}_i(x^n) \stackrel{\Delta}{=} y_i - D_i(x_n) = y_i - \sqrt{(x_1^i - x_1^n)^2 + (x_2^i - x_2^n)^2}, \quad i = 1, 2.$$

Вводя

$$H = - \begin{bmatrix} \sin \Pi_1(x^n), \cos \Pi_1(x^n) \\ \sin \Pi_2(x^n), \cos \Pi_2(x^n) \end{bmatrix} = \begin{bmatrix} H_1(x^n) \\ H_2(x^n) \end{bmatrix}; \quad \tilde{y} = \begin{bmatrix} \tilde{y}_1(x_n) \\ \tilde{y}_2(x_n) \end{bmatrix}; \quad \delta x = x - x^n$$

и привлекая (2.9), запишем

$$\delta \hat{x}^{\text{МНК}} = (H_1^T(x_n)H_1(x_n) + H_2^T(x_n)H_2(x_n))^{-1} (H_1^T(x_n)\tilde{y}_1(x_n) + H_2^T(x_n)\tilde{y}_2(x_n)),$$

где $H_i(x^n) = -(\sin \Pi_i(x^n), \cos \Pi_i(x^n))$, $i = 1, 2$.

Предполагая дополнительно, что ошибки измерения – не коррелированные между собой случайные величины с одинаковыми дисперсиями, равными r^2 , конкретизируем также выражение (2.25) для расчетной матрицы ковариаций ошибок оценивания

$$P^{\text{МНК}}(x_n) = r^2 \left(H_1^T(x_n)H_1(x_n) + H_2^T(x_n)H_2(x_n) \right)^{-1}.$$

Будем считать, что точка линеаризации выбрана в начале координат, т.е. $x^n = 0$. В этом случае $\Pi_1(x^n) = 90^\circ$, $\Pi_2(x^n) = 0$ и, следовательно, $H = -E$. Принимая во внимание этот факт, получаем:

$$\delta \hat{x}^{\text{МНК}} = - \begin{bmatrix} \tilde{y}_1(x_n) \\ \tilde{y}_2(x_n) \end{bmatrix}; \quad P^{\text{МНК}} = \begin{bmatrix} r^2 & 0 \\ 0 & r^2 \end{bmatrix}.$$

Таким образом, уточненные координаты объекта, вычисляемые с помощью МНК, будут иметь вид

$$\hat{x}_i^{\text{МНК}} = x_i^n + \delta \hat{x}_i^{\text{МНК}} = x_i^n + D_i(x_n) - y_i$$

или с учетом того, что точка линеаризации выбрана в начале координат,

$$\hat{x}_i^{\text{МНК}} = \delta \hat{x}_i^{\text{МНК}} = D_i(0) - y_i, \quad i = 1, 2.$$

Выбирая теперь в качестве $x^n = \delta \hat{x}_i^{\text{МНК}}$, повторим вычисления еще раз, при этом заметим, что при $\delta \hat{x}_i^{\text{МНК}} \neq 0$ матрица H уже не будет единичной, так как $\Pi_1(x^n)$ и $\Pi_2(x^n)$ будут отличны от 90° и нуля соответственно. Для реализации итерационного алгоритма следует повторить описанные действия до тех пор, пока $\delta \hat{x}_i^{\text{МНК}}$ на очередном шаге не станет пренебрежимо малой.

Ясно, что полученные оценки будут совпадать с оценками ОМНК, если весовая матрица Q выбрана диагональной с элементами $1/r^2$, $\bar{x} = 0$, а матрица D нулевая.

Легко обобщить представленный алгоритм для случая, когда проведено m измерений. В частности, считая, что $x \equiv \delta x = (x - x^n)$ – центрированный вектор с матрицей ковариаций P^x , а ошибки измерения v_i – не коррелированные между собой и с вектором x центрированные с.в. с заданными дисперсиями r_i^2 , $i = \overline{1, m}$, для расчетных матриц ковариаций ошибок оценок, соответствующих различным вариантам МНК, можно записать:

$$P^{\text{МНК}}(x^n) = \left(\sum_{i=1}^m M_i(x^n) \right)^{-1} \left(\sum_{i=1}^m r_i^2 M_i(x^n) \right) \left(\sum_{i=1}^m M_i(x^n) \right)^{-1}; \quad (2.45)$$

$$P^{\text{ОМНК}}(x^n) = \left(\sum_{i=1}^m \frac{1}{r_i^2} M_i(x^n) \right)^{-1}; \quad (2.46)$$

$$P^{\text{ММНК}}(x^n) = \left((P^x)^{-1} + \sum_{i=1}^m \frac{1}{r_i^2} M_i(x^n) \right)^{-1}, \quad (2.47)$$

где

$$M_i(x^n) = \begin{bmatrix} \sin^2 \Pi_i(x^n) & 0.5 \sin 2\Pi_i(x^n) \\ 0.5 \sin 2\Pi_i(x^n) & \cos^2 \Pi_i(x^n) \end{bmatrix}. \quad (2.48)$$

Если считать измерения равноточными, т.е. полагать $r_i^2 = r^2$, $i = \overline{1, m}$, то приведенные выражения упрощаются к виду:

$$P^{\text{МНК}}(x^n) = P^{\text{ОМНК}}(x^n) = r^2 \left(\sum_{i=1}^m M_i(x^n) \right)^{-1}; \quad (2.49)$$

$$P^{\text{ММНК}}(x^n) = \left((P^x)^{-1} + \frac{1}{r^2} \sum_{i=1}^m M_i(x^n) \right)^{-1}. \quad (2.50)$$

При использовании итерационного алгоритма для получения расчетной матрицы ковариаций вместо x^n следует принять значение, полученное для последней итерации.

Заметим, что число измерений и количество точечных ориентиров могут между собой не совпадать. ♦

Из полученных в примере 2.7 соотношений следует, что расчетная точность оценивания в задаче определения координат по точечным ориентирам в значительной степени определяется взаимным расположением ориентиров. Действительно, вводя радиальную среднеквадратическую погрешность определения координат места σ_m , можем записать

$$\sigma_m \equiv DRMS = \sqrt{P^{\text{ОМНК}}(1,1) + P^{\text{ОМНК}}(2,2)} = r \sqrt{Sp\left(\sum_{i=1}^m M_i(x^i)\right)^{-1}}. \quad (2.51)$$

Отсюда следует, что σ_m при одинаковой величине r в пренебрежении априорной информации определяется произведением среднеквадратической ошибки измерения дальностей на некоторый коэффициент

$$PDOP = \sqrt{Sp\left(\sum_{i=1}^m M_i(x^i)\right)^{-1}}, \quad (2.52)$$

зависящий от взаимного расположения точечных ориентиров. Этот коэффициент называется **геометрическим фактором (Position Dilution of Precision)**. Например, в частном случае двух точечных ориентиров при $\Pi_2 = \Pi_1 + 90^\circ$ нетрудно убедиться в том, что $\sigma_m \equiv \sqrt{2}r$ (см. задачу 2.2.5). Очевидно, что при увеличении числа используемых ориентиров m расчетная радиальная среднеквадратическая погрешность будет уменьшаться.

Задачи к разделу

Задача 2.1. Задан критерий вида

$$J^{\text{МНК}}(x) = (y - Hx)^T R^{-1} (y - Hx) + (x - \bar{x})^T (P^x)^{-1} (x - \bar{x}).$$

Покажите, что значение x , при котором этот критерий достигает минимального значения, определяется как

$$\hat{x}(y) = \bar{x} + ((P^x)^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} (y - H\bar{x}).$$

Примечание. Задачу решите, используя:

- а) систему нормальных уравнений;
- б) процедуру выделения полного квадрата

Задача 2.2. Покажите, что при решении задачи оценивания вектора x по измерениям (2.1) с помощью МНК и ОМНК будут справедливы следующие соотношения [6, 25]:

$$\begin{aligned} (y - \hat{y}^{\text{МНК}})^T \hat{y}^{\text{МНК}} &= 0, ; \\ (y - \hat{y}^{\text{ОМНК}})^T Q \hat{y}^{\text{ОМНК}} &= 0, \end{aligned}$$

где $\hat{y}^{\text{МНК}} = H\hat{x}^{\text{МНК}} = HK^{\text{МНК}}y$; $\hat{y}^{\text{ОМНК}} = H\hat{x}^{\text{ОМНК}} = HK^{\text{ОМНК}}y$.

Задача 2.3. Пусть в критерии (2.16) $Q = R^{-1}$, $D = (P^x)^{-1}$, и, таким образом, в выражении для оценки $\hat{x}^{\text{МНК}}(y) = \bar{x} + K^{\text{МНК}}(y - H\bar{x})$, со-

ответствующей ММНК, матрица $K^{\text{ММНК}}$ определяется как $K^{\text{ММНК}} = ((P^x)^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1}$. Полагая, что x и v – некоррелированные случайные векторы с матрицами ковариаций P^x и R , а математическое ожидание вектора x равно \bar{x} , покажите, что матрица ковариаций ошибок оценок ММНК определяется как

$$P^{\text{ММНК}} = \left((P^x)^{-1} + H^T R^{-1} H \right)^{-1}.$$

Задача 2.4. Рассчитайте величину геометрического фактора в задаче оценивания координат объекта на плоскости по измерениям (1.16) для случая двух точечных ориентиров, считая, что выполнены следующие предположения: точка линеаризации расположена в начале координат; один из точечных ориентиров расположен на оси ox_1 ; второй – на оси ox_2 ; ошибки измерения – не коррелированные между собой центрированные случайные величины с одинаковыми дисперсиями, равными r^2 .

Контрольные вопросы

1. Сформулируйте постановку задачи оценивания с использованием МНК, ОМНК и ММНК и поясните их особенности.
2. Что такое невязка измерений и система нормальных уравнений? Приведите пример системы нормальных уравнений для задачи оценивания фазы.
3. Решите задачу оценивания амплитуды гармонического сигнала с использованием МНК.
4. Получите алгоритмы вычисления оценок МНК и его модификаций для линейной задачи в общем виде. Проиллюстрируйте применение этих алгоритмов на примере решения простейшей задачи оценивания постоянной величины.
5. Приведите уравнения для ошибок оценок МНК и его модификаций и выражения для матриц ковариаций ошибок оценок для линейной задачи оценивания. Какие дополнительные предположения о свойствах ошибок измерения необходимо ввести для вычисления матриц ковариаций ошибок оценивания при решении задачи с помощью МНК?
6. Поясните, как с использованием измерений (1.21) могут быть получены алгоритмы оценивания, основанные на линеаризации функции. Проиллюстрируйте это на примерах оценивания фазы гармонического сигнала и оценивания координат на плоскости по измерениям дальностей до точечных ориентиров.
7. В чем особенность итерационного алгоритма? При каких условиях этот алгоритм может обеспечить нахождение оценки, соответствующей МНК?

РАЗДЕЛ 3

Решение задач оценивания на основе небайесовского подхода. Метод максимума правдоподобия

В настоящем разделе рассмотрим **небайесовские алгоритмы**, которые строятся в предположении о случайном характере только ошибок измерения, и при этом считаются полностью известными их статистические свойства. Последнее означает, что задана ф.п.р.в. $f_v(v)$. Неизвестный оцениваемый вектор, как и в методе наименьших квадратов и его модификациях, предполагается неслучайным (детерминированным) вектором. Подход к получению оценок при таких предположениях получил название **небайесовского, или классического подхода (метода)** [7, 16, 25, 27].

3.1. Основные положения и постановка задачи

Введение предположения о случайном характере ошибок измерения с известной ф.п.р.в. $f_v(v)$ уже на этапе проектирования алгоритмов дает возможность при фиксированных значениях x рассматривать измерения как случайный вектор, свойства которого определены с помощью условной к x ф.п.р.в. $f(y/x)$. В полной мере это относится и к оценке $\tilde{x}(y)$, и к ее ошибке $\varepsilon(y) = x - \tilde{x}(y)$, представляющим собой преобразования измерений y . Принимая во внимание соотношение $y = s(x) + v$ и фиксируя x , для $f(y/x)$ можно записать выражение

$$f(y/x) = f_v(y - s(x)), \quad (3.1)$$

в котором $f_v(\bullet)$ – ф.п.р.в. ошибок измерения.

Так, полагая, что в выражении $y = s(x) + v$ ошибки измерения представляют собой гауссовский центрированный вектор с известной матрицей ковариаций R , $f(y/x)$ может быть конкретизирована в виде

$$f(y/x) = \frac{1}{(2\pi)^{m/2} \sqrt{\det R}} \exp\left(-\frac{1}{2}(y - s(x))^T R^{-1}(y - s(x))\right). \quad (3.2)$$

Если, кроме того, считать v_i , $i = \overline{1, m}$ не зависимыми между собой случайными величинами с дисперсиями r_i^2 , $i = \overline{1, m}$, то $f(y/x)$ представляется как

$$f(y/x) = \frac{1}{(2\pi)^{\frac{m}{2}} \sqrt{\prod_{i=1}^m r_i^2}} \exp\left(-\frac{1}{2} \sum_{i=1}^m \frac{(y_i - s_i(x))^2}{r_i^2}\right). \quad (3.3)$$

Количественная характеристика качества оценивания x по измерениям y может быть введена с помощью скалярной функции $L(x - \tilde{x}(y))$, устанавливающей определенный штраф за отличие оценки от истинного значения оцениваемого параметра и называемой **функцией потерь**. Наибольшее распространение при анализе качества оценок в задачах обработки навигационной информации получила **квадратичная функция потерь**

$$L(x - \tilde{x}(y)) = \sum_{i=1}^n (x_i - \tilde{x}_i(y))^2 = (x - \tilde{x}(y))^T (x - \tilde{x}(y)) = Sp\{(x - \tilde{x}(y))(x - \tilde{x}(y))^T\}.$$

Введем связанный с ней критерий в виде математического ожидания от этой функции

$$J(x) = M_{y/x}\{L(x - \tilde{x}(y))\} = M_{y/x}\{(x - \tilde{x}(y))^T (x - \tilde{x}(y))\}, \quad (3.4)$$

который в силу справедливости перестановки местами операций взятия математического ожидания и вычисления следа матриц также может быть записан в виде

$$J(x) = M_{y/x}\{Sp(x - \tilde{x}(y))(x - \tilde{x}(y))^T\} = Sp\tilde{P}(x),$$

где

$$\tilde{P}(x) = M_{y/x}\{(x - \tilde{x}(y))(x - \tilde{x}(y))^T\} \quad (3.5)$$

– матрица ковариаций ошибок оценивания.

Важно заметить, что знак математического ожидания соответствует функции $f(y/x)$. Для того чтобы подчеркнуть этот факт, у этой матрицы и критерия введен аргумент x .

Таким образом, **задача синтеза алгоритма в рамках небайесовского подхода** может быть конкретизирована следующим образом.

Найти алгоритм вычисления оценки неизвестного детерминированного вектора x по измерениям $y = s(x) + v$, в которых v – m -мерный случайный вектор ошибок измерения с заданной $f_v(v)$ исходя из минимизации критерия (3.4), т.е.

$$\hat{x}(y) = \arg \min_{\tilde{x}(y)} M_{y/x} \left\{ \sum_{i=1}^n (x - \tilde{x}_i(y))^2 \right\}.$$

Задача **анализа точности в рамках небайесовского подхода** сводится к вычислению матрицы ковариаций ошибок оценивания (3.5).

Критерий (3.4) получил наименование **среднеквадратического критерия**, а оценка, обеспечивающая его минимум – **оптимальной в среднеквадратическом смысле небайесовской оценкой**. Заметим, что критерий

(3.4) принципиально отличается от критериев, рассмотренных в предыдущем разделе, поскольку цель решения задачи заключается в обеспечении определенных требований к ошибкам оценки искомого вектора, а не к вычисленным значениям измеряемых параметров.

К сожалению, для сформулированной задачи не удастся установить общего правила нахождения оценок, минимизирующих критерий (3.4). В связи с этим выбор того или иного алгоритма оценивания в предположении о случайном характере ошибок измерения и детерминированном характере неизвестного вектора проводят путем сравнения соответствующих им значений критерия и сопоставления свойств получающихся оценок. Важную роль при таком сопоставлении играют такие понятия, как несмещенность, состоятельность и эффективность. Приведем определения этих понятий и поясним их смысл.

В рамках небайесовского подхода оценка $\tilde{x}(y)$ называется несмещенной, если ее математическое ожидание совпадает с истинным значением параметра x , т.е.

$$M_{y/x}\{\tilde{x}(y)\} = x, \quad (3.6)$$

или, записывая более подробно,

$$M_{y/x}\{\tilde{x}(y)\} = \int \tilde{x}(y) f(y/x) dy = x.$$

Как правило, при минимизации критерия (3.3) накладывают дополнительное требование к несмещенности оценки. Оценка, обеспечивающая минимум этого критерия при выполнении требования (3.6), называется **небайесовской несмещенной оценкой с минимальной дисперсией**.

Для объяснения понятия состоятельности предположим, что имеется последовательность измерений скалярной величины $y_i = x + v_i$, $i = \overline{1, m}$, и с их использованием вырабатывается оценка \tilde{x}_m . Эта оценка называется **состоятельной**, если она сходится по вероятности к истинному значению оцениваемой величины при увеличении числа измерений m , т.е.

$$\lim_{m \rightarrow \infty} \Pr(x - e < \tilde{x}_m < x + e) = 1, \quad (3.7)$$

где e – сколь угодно малое положительное число.

Аналогично можно определить состоятельность оценки и в векторном случае.

Понятие эффективности оценки связано с **неравенством Рао–Крамера**. Для несмещенных оценок $\tilde{x}(y)$ **неравенство Рао–Крамера в рамках небайесовского подхода** формулируется следующим образом [7, 25, 27]:

$$\tilde{P}(x) \geq I^{-1}(x), \quad (3.8)$$

где

$$I(x) = M_{y/x} \left\{ \frac{\partial \ln f(y/x)}{\partial x} \left(\frac{\partial \ln f(y/x)}{\partial x} \right)^T \right\}. \quad (3.9)$$

Из этого неравенства следует, что можно указать такую матрицу $I(x)$ при фиксированном значении x , которая будет всегда меньше либо равна матрице ковариаций для любой несмещенной оценки. Напомним, что знак неравенства для матриц означает справедливость неравенства для соответствующих квадратичных форм, см. соотношение (П1.35). В этом смысле матрица $I^{-1}(x)$ определяет **предельно достижимую точность решения задачи в рамках небайесовского подхода**. Будем называть эту матрицу матрицей, характеризующей нижнюю границу точности, или просто **матрицей нижней границы**. Для справедливости приведенного неравенства требуется, чтобы $f(y/x)$ удовлетворяла условиям регулярности, суть которых сводится к абсолютной интегрируемости и существованию первых и вторых производных по x . Оценка, для которой в (3.8) достигается знак равенства, называется **эффективной небайесовской оценкой**. Стоящая справа в выражении (3.9) матрица называется **информационной матрицей Фишера**.

Если найден алгоритм вычисления эффективной оценки $\hat{x}(y)$, то для нее, как следует из (3.8), справедливо следующее неравенство:

$$M_{y/x} \{ (x - \hat{x}(y))(x - \hat{x}(y))^T \} \leq M_{y/x} \{ (x - \tilde{x}(y))(x - \tilde{x}(y))^T \}, \quad (3.10)$$

означающее, что какой бы ни выбирался любой другой алгоритм вычисления несмещенной оценки $\tilde{x}(y)$, соответствующая ему матрица ковариаций ошибок будет всегда меньше или равна матрице, обратной информационной матрице Фишера.

Использование неравенства Рао–Крамера оказывается весьма полезным при анализе точности, поскольку с его помощью удастся оценить потенциально достижимую точность без построения самой процедуры оценивания.

Введенные выше определения, касающиеся свойств получаемых оценок, позволяют более предметно проводить их сопоставление.

3.2. Метод максимума правдоподобия

Наибольшее распространение в рамках небайесовского подхода получил алгоритм вычисления оценок, основанный на максимизации $f(y/x)$ как функции x при фиксированных значениях измерений y . Эта функция в теории оценивания получила название **функции правдоподобия**, а метод вычисления, основанный на ее максимизации – **метод максимума функции правдоподобия (МФП)** [7, 16, 21, 25]. Заметим, что значение $f(y^*/x)$, умноженное на величину малого приращения измерений Δy , т.е. $f(y^*/x) \Delta y$, в скалярном случае приближенно определяет вероятность для измерения попасть в интервал $(y^*, y^* + \Delta y)$, т.е. $\text{Pr}(y^* < y < y^* + \Delta y) \approx f(y^*/x) \Delta y$. Таким образом, смысл процедуры максимизации функции правдоподобия заключается в том, чтобы при фикси-

рованных значениях измерений подобрать такое значение искомого параметра, при котором эта вероятность достигает наибольшего значения. Часто вместо функции правдоподобия имеют дело с ее логарифмом или **логарифмической функцией правдоподобия** $\ln f(y/x)$. Обычно эти функции определяются с точностью до произвольного постоянного множителя.

Из теории оценивания известно, что оценка, соответствующая максимуму функции правдоподобия, обладает важными свойствами: она состоятельна; в асимптотике при неограниченном увеличении числа измерений $m \rightarrow \infty$ не смещена и нормальна (имеет гауссовское распределение) [7]. **Кроме того, установлен следующий факт: если существует эффективная небайесовская оценка, то она является оценкой, максимизирующей функцию правдоподобия** [7]. Перечисленные свойства оценки, соответствующей максимуму функции правдоподобия, и объясняют ее широкое использование в рамках небайесовского подхода. Однако следует иметь в виду, что эта оценка не является общим решением задачи минимизации критерия нахождения несмещенных оценок с минимальной дисперсией и даже не всегда при ограниченном объеме измерений является просто несмещенной [16].

Итак, **оценка максимального правдоподобия (maximum likelihood estimate)** отыскивается путем выбора значения x , которое максимизирует $f(y/x)$, т.е.

$$\hat{x}^{\text{мп}}(y) = \arg \max_x f(y/x), \quad (3.11)$$

либо

$$\hat{x}^{\text{мп}}(y) = \arg \max_x \ln f(y/x).$$

Для обеспечения максимума функции правдоподобия требуется, чтобы соответствующая оценка удовлетворяла необходимому условию максимума

$$\left. \frac{d}{dx} f(y/x) \right|_{\hat{x}^{\text{мп}}(y)} = 0,$$

или

$$\left. \frac{d}{dx} \ln f(y/x) \right|_{\hat{x}^{\text{мп}}(y)} = 0. \quad (3.12)$$

Эти уравнения получили названия **уравнений правдоподобия**.

Как и в случае МНК, приведенные условия являются лишь необходимыми для обеспечения максимума, и каждое из полученных решений должно проверяться на достаточное условие типа

$$\left. \frac{d^2}{dx^2} \ln f(y/x) \right|_{\hat{x}^{\text{мп}}(y)} \leq 0. \quad (3.13)$$

Рассмотрим два примера.

◆ **Пример 3.1.** Конкретизируем алгоритм вычисления оценки, максимизирующей функцию правдоподобия, для задачи оценивания скалярной величины x по скалярным измерениям

$$y_i = x + v_i, i = \overline{1, m}, \quad (3.14)$$

в которых $v_i, i = \overline{1, m}$ – не зависящие между собой гауссовские случайные величинами с одинаковыми дисперсиями r^2 , т.е. $R = r^2 E$.

Так как функция правдоподобия для этого примера имеет вид (3.3), а $s(x) = x$, алгоритм нахождения оценки сведется к минимизации критерия

$$J^{\text{мфп}}(x) = -\frac{1}{2r^2} \sum_{i=1}^m (y_i - x)^2,$$

откуда

$$\hat{x}^{\text{мфп}}(y) = \frac{1}{m} \sum_{i=1}^m y_i.$$

Таким образом, оценка, соответствующая максимуму функции правдоподобия, в данном примере представляет собой среднеарифметическое по всем измерениям.

Проанализируем свойства этой оценки.

Поскольку $M_{y/x} \{ \hat{x}^{\text{мфп}}(y) \} = \frac{1}{m} M_{y/x} \left\{ \sum_{i=1}^m (x + v_i) \right\} = x$, оценка является несмещенной, а дисперсия ее ошибки вычисляется как

$$M_{y/x} \left\{ \left(\hat{x}^{\text{мфп}}(y) - x \right)^2 \right\} = \frac{1}{m^2} M_{y/x} \left\{ \left(\sum_{i=1}^m v_i \right)^2 \right\} = \frac{r^2}{m}.$$

В силу того, что дисперсия ошибки оценки стремится к нулю при увеличении числа измерений m , оценка $\hat{x}^{\text{мфп}}(y)$ является состоятельной.

Рассчитаем предельно достижимую точность для этого примера. В данном случае имеем

$$\frac{\partial \ln f(y/x)}{\partial x} = H^T \frac{1}{r^2} (y - Hx),$$

где H^T – строка, состоящая из единиц.

Принимая во внимание тот факт, что

$$M_{y/x} \left\{ (y - Hx)(y - Hx)^T \right\} = r^2 E_m,$$

получаем

$$\tilde{P}(x) \geq \frac{r^2}{m}.$$

Таким образом, в задаче оценивания скалярной величины по измерениям (3.14) в условиях, когда ошибки измерения – независимые между собой гауссовские случайные величины с одинаковыми дисперсиями r^2 , диспер-

сия ошибки несмещенной оценки не может быть меньше $\frac{r^2}{m}$. Отсюда следует, что оценка, соответствующая максимуму функции правдоподобия, в этом примере является несмещенной эффективной оценкой, а следовательно, и несмещенной оценкой с минимальной дисперсией. \blacklozenge

3.3. Общее решение линейной гауссовской задачи.

Взаимосвязь с алгоритмами метода наименьших квадратов

Решение задачи нахождения несмещенной эффективной оценки в примере 3.1 оказалось достаточно простым. Более того, нетрудно было заметить, что полученный алгоритм совпал с алгоритмом МНК. Это является следствием того, что в примере решалась линейная гауссовская задача. Именно в этом случае алгоритм нахождения оценки, соответствующей максимуму функции правдоподобия, реализуется наиболее просто. Покажем это.

Будем считать, что рассматривается линейная задача оценивания постоянного n -мерного вектора x по измерениям m -мерного вектора измерений $y = Hx + v$, в предположении, что ошибки измерения v представляют собой центрированный случайный гауссовский вектор с матрицей ковариаций R . Получим оценку, соответствующую максимуму правдоподобия, и проанализируем ее свойства.

При сделанных предположениях функция правдоподобия совпадает с выражением (3.2), в котором $s(x) = Hx$, т.е.

$$f(y/x) = \frac{1}{(2\pi)^{m/2} \sqrt{\det R}} \exp\left(-\frac{1}{2}(y - Hx)^T R^{-1}(y - Hx)\right),$$

а логарифмическая функция правдоподобия может быть записана в виде

$$J^{\text{мфп}}(x) = (y - Hx)^T R^{-1}(y - Hx). \quad (3.15)$$

Таким образом,

$$\hat{x}^{\text{мфп}}(y) = \arg \max_x N(y, Hx, R) = \arg \min_x J^{\text{мфп}}(x). \quad (3.16)$$

Заметим, что полученный критерий (3.15) совпадает с критерием (2.13) для ОМНК при $Q = R^{-1}$, что позволяет получить следующие выражения:

$$\hat{x}^{\text{мфп}}(y) = K^{\text{мфп}} y; \quad (3.17)$$

$$P^{\text{мфп}} = K^{\text{мфп}} R (K^{\text{мфп}})^T = (H^T R^{-1} H)^{-1}, \quad (3.18)$$

где
$$K^{\text{мфп}} = (H^T R^{-1} H)^{-1} H^T R^{-1}. \quad (3.19)$$

Оценка (3.17) является несмещенной.

Действительно, поскольку $M_{y/x}\{y\} = Hx$, а $E - K^{\text{мфп}} H = 0$, то $M_{y/x}\{(x - K^{\text{мфп}} y)\} = 0$.

В силу того, что ошибка оценки $\varepsilon^{\text{мфп}}(y) = x - \hat{x}^{\text{мфп}}(y)$ может быть представлена в виде, аналогичном (2.23), т.е.

$$\varepsilon^{\text{мфп}}(y) = -K^{\text{мфп}}v, \quad (3.20)$$

эта ошибка не зависит от оцениваемого вектора x , а зависит только от ошибок измерения. Т.е. ошибка оценки в методе максимума правдоподобия в рассматриваемой задаче инвариантна относительно оцениваемого вектора. Существенно, что и матрица ковариаций здесь также не зависит от оцениваемого вектора.

Поскольку в рассматриваемой задаче достаточно просто отыскивается матрица, характеризующая нижнюю границу точности, нетрудно также убедиться в том, что оценка (3.17) является эффективной. Действительно, поскольку $s(x) = Hx$, то, принимая во внимание соотношение

$$\frac{\partial \ln f(y/x)}{\partial x} = H^T R^{-1}(y - Hx)$$

и тот факт, что выполнены условия регулярности, легко вычислить математическое ожидание в (3.9) и для информационной матрицы Фишера получить

$$I(x) = H^T R^{-1} H. \quad (3.21)$$

Сопоставляя $I^{-1}(x)$ с матрицей ковариаций $P^{\text{мфп}}(x)$, задаваемой выражением (3.18), убеждаемся в их совпадении.

Таким образом, в линейной гауссовской задаче оценка максимального правдоподобия (3.17) является несмещенной эффективной небайесовской оценкой с матрицей ковариаций (3.18) или, что то же самое, несмещенной небайесовской оценкой с минимальной дисперсией.

Обсудим теперь связь с МНК. Замечаем, что полученные выражения (3.17), (3.18) идентичны выражениям (2.28), (2.29), соответствующим ОМНК при $Q = R^{-1}$. Это вполне логично, поскольку в рассматриваемом случае задача минимизации критерия ОМНК совпадает с задачей максимизации функции правдоподобия, так как критерий (3.15) совпадает с критерием ОМНК.

Из сказанного следует, что оценка, соответствующая методу максимума функции правдоподобия в задаче оценивания x по измерениям (1.11) при гауссовском характере ошибок измерения, совпадает с оценкой обобщенного метода наименьших квадратов, если весовая матрица в его критерии $Q = R^{-1}$.

Поскольку оценка МФП, как отмечалось выше, совпадет с оценкой ОМНК, то для оценок ОМНК при соответствующем выборе матрицы будет справедлив вывод.

В линейной гауссовской задаче оценка, соответствующая ОМНК, при выборе весовой матрицы в критерии такой, что $Q = R^{-1}$, является

несмещенной эффективной небайесовской оценкой с матрицей ковариаций (3.18).

В силу отмеченного совпадения результаты, рассмотренные в примерах 2.2 -2.5 и касающиеся ОМНК, можно трактовать как результаты, соответствующие методу максимума функции правдоподобия, если дополнительно предположить гауссовский характер ошибок измерения и принять $Q = R^{-1}$.

3.4. Решение нелинейной гауссовской задачи.

Взаимосвязь с алгоритмами метода наименьших квадратов

Обсудим теперь, к чему сводится алгоритм нахождения оценки, соответствующей максимуму функции правдоподобия в нелинейной задаче оценивания постоянного неизвестного n -мерного вектора x по m -мерным измерениям $y = s(x) + v$, где $s(\bullet) = (s_1(x), \dots, s_m(x))^T$ – m -мерная в общем случае нелинейная функция, а v – m -мерный центрированный случайный гауссовский вектор ошибок измерения с матрицей ковариаций R . В этом случае логарифмическая функция правдоподобия может быть записана в виде

$$J^{\text{мфп}}(x) = \ln f(y/x) = -\frac{1}{2}(y - s(x))^T R^{-1}(y - s(x)). \quad (3.22)$$

Отсюда следует, что для нахождения оценки необходимо либо отыскивать максимум этого критерия, либо решать систему нелинейных уравнений

$$\frac{\partial \ln f(y/x)}{\partial x} = \frac{ds^T(x)}{dx} R^{-1}(y - s(x)) = 0 \quad (3.23)$$

с последующей проверкой условия (3.13).

Нетрудно заметить, что вид критерия (3.22), соответствующий рассматриваемому гауссовскому случаю, с точностью до постоянного множителя совпадает с критерием (2.5) обобщенного метода наименьших квадратов, если в ОМНК в качестве весовой матрицы выбрать матрицу $Q = R^{-1}$. Отсюда вытекает следующий вывод, аналогичный выводу для линейной задачи. **Оценка, соответствующая методу максимума функции правдоподобия, в нелинейной задаче оценивания x по измерениям (1.21) при гауссовском характере ошибок измерения совпадает с оценкой обобщенного метода наименьших квадратов, если весовая матрица в его критерии $Q = R^{-1}$.**

Из сказанного следует, что рассмотренные в подразделах 2.3, 2.4 пути построения алгоритмов и анализ их точности можно трактовать как методы решения задач оценивания на основе максимизации функции правдоподобия в предположении гауссовского характера ошибок измерения. В частности, здесь могут быть использованы линеаризованный и итерационный алгоритмы. Следует заметить при этом, что при анализе точности

должна быть использована матрица, подобная матрице (2.39). При этом при вычислении математического ожидания необходимо использовать условную ф.п.р.в. $f(y/x)$.

Ясно, что при нахождении этой матрицы, например в соответствии с **методом статистических испытаний**, необходимо выполнять значительный объем вычислений. И здесь весьма полезным оказывается привлечение неравенства Рао–Крамера, поскольку в нелинейной гауссовской задаче по аналогии с тем, как это было сделано для линейного случая, достаточно просто найти матрицу, характеризующую нижнюю границу точности. Действительно, полагая, что функция $s(x)$ обеспечивает выполнение условий регулярности, и принимая во внимание соотношение

$$\frac{\partial \ln f(y/x)}{\partial x} = \frac{ds^T(x)}{dx} R^{-1}(y - s(x)), \quad (3.24)$$

получаем

$$I(x) = \frac{ds^T(x)}{dx} R^{-1} \frac{ds(x)}{dx^T}. \quad (3.25)$$

Нетрудно заметить, что если применительно к рассматриваемой задаче считать допустимым линеаризованное описание для $s(x)$ и приближенно решать задачу методом максимума правдоподобия с использованием (3.17), то матрица ковариаций, вычисляемая согласно (3.18), совпадет с $I^{-1}(x)$, поскольку

$$I^{-1}(x) \approx I^{-1}(x_n) = \left(H^T(x_n) R^{-1} H(x_n) \right)^{-1} = P^{\text{мфп}}.$$

Конкретизируем полученные соотношения применительно к двум примерам.

◆ **Пример 3.2.** Получим выражение для нахождения нижней границы точности в задаче оценивания скалярной величины x по измерениям

$$y_i = s_i(x) + v_i, i = \overline{1, m}, \quad (3.26)$$

в которых $v_i, i = \overline{1, m}$ предполагаются независимыми между собой гауссовскими случайными величинами с одинаковыми дисперсиями r^2 , т. е. $R = r^2 E$.

Используя выражение (3.25), запишем

$$I(x) = \frac{1}{r^2} \sum_{i=1}^m \left(\frac{ds_i(x)}{dx} \right)^2.$$

Если ввести величину, характеризующую значения производной

$$\bar{g}(x) = \sqrt{\frac{1}{m} \sum_{i=1}^m \left(\frac{ds_i(x)}{dx} \right)^2},$$

то выражение для нижней границы точности можно записать как

$$P(x) \geq I^{-1}(x) = \frac{r^2}{\bar{g}^2(x)m}. \quad (3.27)$$

Заметим, что $I^{-1}(x) = (\sigma^{\text{МНК}}(x))^2$, где $(\sigma^{\text{МНК}}(x))^2$ – значение расчетной дисперсии ошибки оценки, соответствующей линеаризованному МНК с точкой линеаризации $x = x^n$.

Нетрудно конкретизировать полученные соотношения для примера 2.6. В частности, величина, характеризующая значение производной, будет определяться как

$$\bar{g}^2(x) = \frac{\sum_{i=1}^m \cos^2(\omega t_i + x)}{m}. \quad \blacklozenge$$

◆ Пример 3.3. Получим выражение для матрицы, характеризующей нижнюю границу точности, в нелинейной задаче определения координат места по измерениям дальностей (1.16) до m точечных ориентиров, полагая, что ошибки измерения являются независимыми между собой гауссовскими центрированными случайными величинами с дисперсиями r_i^2 , $i = \overline{1.m}$.

Принимая во внимание (3.25), для матрицы нижней границы можем записать выражение: $I^{-1}(x) = \left(\sum_{i=1}^m \frac{1}{r_i^2} M_i(x) \right)^{-1}$, в котором $M_i(x)$ определяются соотношением (2.64), т. е.

$$M_i(x) \triangleq \begin{bmatrix} \sin^2 \Pi_i(x) & 0.5 \sin 2\Pi_i(x) \\ 0.5 \sin 2\Pi_i(x) & \cos^2 \Pi_i(x) \end{bmatrix}.$$

Если предположить, что линеаризованное описание в области априорной неопределенности справедливо, то увидим, что матрица ковариаций для ОМНК при замене значения x^n на истинное значение x , как следует из примера 2.7, будет совпадать с расчетной матрицей, характеризующей нижнюю границу точности. ◆

В общем случае при использовании неравенства Рао–Крамера для анализа точности в нелинейных задачах необходимо помнить, что действительная матрица ковариаций ошибок оценивания для исследуемых алгоритмов задается соотношением

$$\tilde{P} = M \left\{ (x - \tilde{x}(y))(x - \tilde{x}(y))^T \right\}$$

и может существенно отличаться от $I^{-1}(x)$. В частности, такая ситуация может иметь место для линеаризованного алгоритма в силу приближенности описания (1.22). Вместе с тем неравенство (3.8), которое в данном случае конкретизируется как

$$\tilde{P}(x) \geq \left(\frac{ds^T(x)}{dx} R^{-1} \frac{ds(x)}{dx^T} \right)^{-1},$$

справедливо всегда, если $s(x)$ обеспечивает выполнение условий регулярности для функции правдоподобия (3.2). Это означает, что какой бы алгоритм ни был выбран для решения рассматриваемой задачи, матрица ковариаций его ошибок будет всегда превышать матрицу, стоящую в правой части неравенства и вычисляемую при фиксированном значении неизвестного вектора.

Задачи к разделу

Задача 3.1. Получите выражение для функции правдоподобия и дисперсии, характеризующей нижнюю границу точности, в задаче оценивания скалярной величины x по измерениям

$$y_i = s_i(x) + \varepsilon_i,$$

где ошибки измерения

$$\varepsilon_i = d + v_i$$

представляют собой сумму центрированной гауссовской случайной величины с дисперсией σ_d^2 , описывающей систематическую составляющую, и независимых между собой и от d центрированных гауссовских случайных величин с одинаковыми дисперсиями $r_i^2 = r^2$, $i = \overline{1.m}$.

Покажите также, что при получении оценки соответствующей максимуму функции правдоподобия можно использовать критерий

$$J^{\text{мфп}}(x) = \frac{1}{2r^2} \left(\sum_{i=1}^m (y_i - s_i(x))^2 - \frac{\sigma_d^2}{m\sigma_d^2 + r^2} \left(\sum_{i=1}^m (y_i - s_i(x)) \right)^2 \right).$$

Задача 3.2. Покажите, что выражение

$$J^{\text{мфп}}(x) = \frac{1}{2r^2} \left(\sum_{i=1}^m (y_i - s_i(x))^2 - \frac{\sigma_d^2}{m\sigma_d^2 + r^2} \left(\sum_{i=1}^m (y_i - s_i(x)) \right)^2 \right)$$

из задачи 3.1 может быть представлено в виде

$$J^{\text{мфп}} = \sum_{i=1}^m \frac{(y_i - s_i(x) - \hat{d}_{i-1}(x))^2}{r^2 + \tilde{\sigma}_{i-1}^2}, \quad (1)$$

где

$$\hat{d}_i(x) = \hat{d}_{i-1}(x) + \frac{\tilde{\sigma}_{i-1}^2}{\tilde{\sigma}_{i-1}^2 + r^2} (y_i - s_i(x) - \hat{d}_{i-1}(x)), \quad ; \quad (2)$$

$$\tilde{\sigma}_i^2 = \frac{\tilde{\sigma}_{i-1}^2 r^2}{\tilde{\sigma}_{i-1}^2 + r^2} \quad i = \overline{1.m}; \quad \tilde{\sigma}_0^2 = \sigma_d^2, \quad \hat{d}_0(x) = 0. \quad (3)$$

Задача 3.3. Пусть имеются два измерителя, вырабатывающие показания

$$y_1 = x + v_1;$$

$$y_2 = x + v_2,$$

в которых $x = (x_1, x_2)^T$ представляет собой двухмерный вектор, задающий координаты объекта на плоскости. Считая, что этот вектор является детерминированным вектором, а двухмерные векторы ошибок измерения являются центрированными гауссовскими векторами с матрицами ковариаций R_1, R_2 , получите алгоритм вычисления эффективной оценки и соответствующей ему матрицы ковариаций. Сопоставьте результаты с решением для ОМНК, полагая матрицу Q блочно-диагональной с блоками R_1^{-1} и R_2^{-1} .

Задача 3.4. Получите основанный на линеаризации и итерационный алгоритмы оценивания, соответствующие методу максимума правдоподобия в задаче оценивания фазы гармонического колебания по измерениям (1.14), т.е.

$$y_i = A \sin(\omega t_i + \phi_0) + v_i, i = \overline{1, m},$$

в предположении, что амплитуда A и частота $\omega = 2\pi f$ известны, а ошибки измерения являются некоррелированными между собой гауссовскими случайными величинами с одинаковой дисперсией r^2 . Сопоставьте результаты с решением для МНК

Задача 3.5. Покажите, что оценки, соответствующие методу максимума функции правдоподобия, не зависят от линейных невырожденных преобразований по отношению к используемым измерениям.

Задача 3.6. Полагая, что решению подлежит линейная задача оценивания вектора x по измерениям $y = Hx + v$, а оценки отыскиваются в виде $\hat{x}(y) = Ky$, найдите выражение для матрицы K , удовлетворяющей условию $E - KH = 0$ и обеспечивающей минимум критерия (3.4), т.е. $J = M_{y/x} \left\{ (x - \hat{x}(y))^T (x - \hat{x}(y)) \right\}$.

Предварительно убедитесь в том, что при выполнении оговоренного условия при вычислении выбранного критерия J достаточно ввести предположение только о случайном характере вектора ошибок измерения v и задаться лишь его первыми двумя моментами.

Полагая далее, что v – центрированный случайный вектор с известной матрицей ковариаций R , получите выражение для соответствующей матрицы ковариаций ошибок таких оценок.

Контрольные вопросы

1. Сформулируйте задачу оценивания в рамках небайесовского подхода. В чем особенности такой постановки по сравнению с постановкой задачи с использованием МНК и его модификаций?

2. Дайте определение несмещенной и состоятельной оценок, а также несмещенной небайесовской оценки с минимальной дисперсией.
3. Приведите формулировку неравенства Рао–Крамера и поясните его смысл. Что такое эффективная оценка?
4. Поясните суть метода максимума правдоподобия, что такое функция правдоподобия и уравнение правдоподобия?
5. Получите выражение для информационной матрицы Фишера для линейной и нелинейной гауссовской задач оценивания.
6. Получите решение линейной гауссовской задачи оценивания методом максимума правдоподобия. Является ли полученная оценка эффективной? Какова взаимосвязь с алгоритмами метода наименьших квадратов?
7. Как соотносится матрица, характеризующая нижнюю границу точности, с действительной матрицей ковариаций ошибок оценивания в нелинейной гауссовской задаче и расчетной матрицей, соответствующей линеаризованному алгоритму?

РАЗДЕЛ 4

Байесовский подход. Оптимальные оценки

Рассмотрим теперь задачу оценивания (1.21) в рамках **байесовского подхода**, т.е. полагаем, что оцениванию подлежит n -мерный постоянный вектор x по m -мерному вектору измерений $y = s(x) + v$. Основная особенность этого подхода заключается в предположении о случайном характере оцениваемого вектора x и ошибки измерения v . Это дает возможность ввести совместную ф.п.р.в. $f(x, v)$, которая, в свою очередь, позволяет перейти к ф.п.р.в. $f(x, y)$ и затем ввести условную (апостериорную) ф.п.р.в. $f(x/y)$ [6, 7, 16, 25, 27, 37]. Ранее при обсуждении метода максимума правдоподобия уже вводилась условная плотность $f(y/x)$. Учитывая важность понятия апостериорной плотности в задаче оценивания, решаемой в рамках байесовского подхода, обсудим это понятие более подробно.

4.1. Формула Байеса и апостериорная плотность

Итак, будем полагать, что совместная ф.п.р.в. $f(x, y)$ определена. Для этой плотности справедлива **формула умножения плотностей вероятности**

$$f(x, y) = f(x/y)f(y) = f(y/x)f(x). \quad (4.1)$$

Входящие в эти соотношения плотности $f(x/y)$ и $f(y/x)$ являются условными плотностями и определяют статистические свойства векторов x и y при условии, что вектор, стоящий справа от черты, фиксирован. При решении задач оценивания условные плотности $f(x/y)$ и $f(y/x)$ также называют **апостериорными плотностями**, тем самым подчеркивается тот факт, что эти плотности соответствуют апостериорной ситуации, т.е. такой, при которой один из векторов, связанный с оцениваемым вектором, фиксируется. В этом случае исходные плотности $f(x)$, $f(y)$ принято называть **априорными**. Из соотношения (4.1) следует, что

$$f(x/y) = \frac{f(x, y)}{f(y)} = \frac{f(x, y)}{\int f(x, y) dx}. \quad (4.2)$$

Зависящая от измерений y функция $f(y) = \int f(x, y) dx$, по сути, обеспечивает нормировку, т.е равенство единице для выражения, стоящего в правой части (4.2).

Соотношение (4.2) известно как **формула Байеса**. Выражения (4.2) обеспечивают возможность нахождения условных плотностей по известной совместной плотности $f(x, y)$. Из них, в частности, следует, что при независимости векторов x и y условные и априорные ф.п.р.в. между собой совпадают.

Запишем соотношения, позволяющие находить параметры **условной гауссовской плотности распределения** [18]. Предположим, что совместная ф.п.р.в. двух гауссовских векторов x и y размерности n и m имеет вид

$$f(x, y) = N\left(\begin{pmatrix} x^T \\ y^T \end{pmatrix}; \begin{pmatrix} \bar{x}^T \\ \bar{y}^T \end{pmatrix}, P\right),$$

где
$$P = \begin{bmatrix} P^x & P^{xy} \\ (P^{xy})^T & P^y \end{bmatrix}.$$

Ясно, что для каждого вектора по отдельности для ф.п.р.в. справедливо следующее представление:

$$f(x) = N(x; \bar{x}, P^x);$$

$$f(y) = N(y; \bar{y}, P^y).$$

Используя соотношение (4.2) и правила обращения блочных матриц, можно показать, что условная плотность $f(x/y)$ также является гауссовской, т.е. $f(x/y) = N(x; \hat{x}(y), P^{x/y})$, а ее параметры определяются с помощью следующих соотношений:

$$\hat{x}(y) = \bar{x} + P^{xy} (P^y)^{-1} (y - \bar{y}); \quad (4.3)$$

$$P^{x/y} = P^x - P^{xy} (P^y)^{-1} (P^{xy})^T. \quad (4.4)$$

Эти соотношения определяют правило нахождения параметров условной гауссовской плотности для двух совместно гауссовских векторов.

◆ **Пример 4.1.** Предположим, что задана скалярная гауссовская случайная величина x с плотностью $f(x) = N(x; \bar{x}, \sigma_0^2)$ и набор величин

$$y_i = s(x) + v_i, \quad i = \overline{1, m},$$

в которых $s(x)$ – известная в общем случае нелинейная функция, например, $s(x) = x^3$. Считаем, что $v_i, i = \overline{1, m}$ – независимые между собой и от x гауссовские случайные величины с ф.п.р. $f(v_i) = N(v_i; 0, r^2)$. Запишем выражение для апостериорной плотности. Используя (4.2) и сделанные предположения, получим

$$f(x/y) = \frac{1}{c(y)} \exp - \frac{1}{2} \left(\frac{x^2}{P^x} + \frac{1}{r^2} \sum (y_i - s(x))^2 \right),$$

где нормирующий множитель определяется как

$$c(y) = \int_{-\infty}^{+\infty} \exp - \frac{1}{2} \left(\frac{x^2}{P^x} + \frac{1}{r^2} \sum (y_i - s(x))^2 \right) dx.$$

Функция $c(y)$ с точностью до коэффициентов совпадает с $f(y)$. Она обычно называется **нормирующим множителем**. Для частного случая, когда $s(x) = x$, в этих выражениях $s(x)$ надо заменить на x . Для этого частного случая нетрудно показать, что апостериорная плотность является гауссовской $f(x/y) = N(x; \hat{x}(y), P^{x/y})$, а ее параметры определяются в виде

$$\hat{x}(y) = \bar{x} + \frac{\sigma_0^2}{r^2 + \sigma_0^2 m} \left(\sum_{i=1}^m (y_i - \bar{x}) \right);$$

$$P = \left[\frac{1}{\sigma_0^2} + \frac{m}{r^2} \right]^{-1} = \frac{\sigma_0^2 r^2}{r^2 + \sigma_0^2 m}.$$

Для того чтобы убедиться в этом, необходимо сначала записать параметры гауссовской плотности составного вектора $z = (x, v_1, v_2, \dots, v_m)^T$, затем найти параметры гауссовской плотности для составного вектора $\tilde{z} = (x, y_1, y_2, \dots, y_m)^T = (x, y^T)^T$. Это сделать нетрудно, если учесть, что $\tilde{z} = Tz$, где T – матрица, легко определяемая из соотношения $y_i = x + v_i$, $i = \overline{1, m}$. Далее необходимо воспользоваться соотношениями (4.3.), (4.4.). Нетрудно убедиться в том, что в рассматриваемом примере при $s(x) = x$ функция $f(y)$ будет определяться как

$$f(y) = N(y; \bar{x} I_{m \times 1}; \sigma_0^2 I_{m \times m} + r^2 E_{m \times m}),$$

где $I_{m \times m}$, $E_{m \times m}$ – $m \times m$ матрица, состоящая из единиц и единичная матрица. ◆

4.2. Постановка и общее решение байесовской задачи оценивания

Сформулируем теперь постановку задачи оценивания, решаемую в рамках байесовского подхода, и приведем для нее общее решение.

Как и в предыдущем разделе, введем квадратичную функцию потерь

$$L(x - \tilde{x}(y)) = \sum_{i=1}^n (x_i - \tilde{x}_i(y))^2 = (x - \tilde{x}(y))^T (x - \tilde{x}(y))$$

и связанный с ней критерий в виде математического ожидания квадратичной функции потерь

$$J^B = M_{x,y} \{L(x - \tilde{x}(y))\} = M_{x,y} \left\langle Sp \left\{ (x - \tilde{x}(y))(x - \tilde{x}(y))^T \right\} \right\rangle = Sp \{ \tilde{P} \}.$$

Этот критерий в рамках байесовского подхода называют **байесовским риском или средними байесовскими потерями**.

Заметим, что он отличается от аналогичного критерия, введенного в предыдущем разделе, поскольку здесь математическое ожидание вычисляется с учетом случайного характера как x , так и y .

Задача оценивания вектора x по измерениям y в рамках байесовского подхода формулируется так: найти такую оценку, которая обеспечит минимум математического ожидания для квадратичной функции потерь, т.е.

$$\hat{x}(y) = \arg \min_{\tilde{x}(y)} M_{x,y} \left\{ \sum_{i=1}^n (x - \tilde{x}_i(y))^2 \right\} = \arg \min_{\tilde{x}(y)} Sp \{ \tilde{P} \}.$$

Такая оптимальная в среднеквадратическом смысле оценка в дальнейшем называется **оптимальной байесовской оценкой или, как правило, просто оптимальной оценкой**.

В теории оценивания весьма важным является следующее утверждение.

Оптимальной байесовской оценкой является математическое ожидание, соответствующее апостериорной плотности $f(x/y)$, т.е.

$$\hat{x}(y) = \int x f(x/y) dx. \quad (4.5)$$

Действительно, записывая критерий в виде

$$J^B = \int \int (x - \tilde{x}(y))^T (x - \tilde{x}(y)) f(x/y) dx f(y) dy$$

и дифференцируя его по оценке, можем записать

$$\frac{d}{d\tilde{x}(y)} \int (x - \tilde{x}(y))^T (x - \tilde{x}(y)) f(x/y) dx = -2 \int (x - \tilde{x}(y))^T f(x/y) dx = 0.$$

Отсюда следует, что

$$\int x^T f(x/y) dx = \tilde{x}^T(y) \int f(x/y) dx.$$

Принимая во внимание условие нормировки, получаем (4.5). Таким образом, для нахождения оптимальной оценки в рамках **байесовского подхода** требуется вычисление многократного интеграла (4.5).

При анализе точности в рамках байесовского подхода используются **условная и безусловная апостериорные матрицы ковариаций ошибок оценивания**:

$$P(y) = M_{x|y} \{ [x - \hat{x}(y)][x - \hat{x}(y)]^T \} = \int [x - \hat{x}(y)][x - \hat{x}(y)]^T f(x/y) dx, \quad (4.6)$$

$$P = M_{x,y} \{ [x - \hat{x}(y)][x - \hat{x}(y)]^T \} = \iint [x - \hat{x}(y)][x - \hat{x}(y)]^T f(x,y) dx dy. \quad (4.7)$$

Матрица $P(y)$ характеризует точность оценивания при их конкретных значениях измерений, а P – точность оценивания в среднем по всем измерениям.

Важно обратить внимание на следующее обстоятельство. В отличие от алгоритмов, получаемых с использованием МНК или небайесовского подхода, в рамках байесовского подхода указывается не только алгоритм вычисления самой оценки, но и процедура вычисления соответствующей ей

матрицы ковариаций $P(y)$, характеризующей расчетную точность оценивания для конкретной реализации измерений, используемой при нахождении оценок. Это имеет принципиальное значение при решении задач оценивания, связанных с обработкой навигационной информации.

С учетом сказанного **под задачей синтеза оптимального алгоритма в рамках байесовского подхода** будем понимать нахождение процедуры, обеспечивающей не только вычисление самой оценки (4.5), но и соответствующей ей текущей характеристики точности в виде условной апостериорной матрицы ковариаций (4.6).

Под задачей анализа точности в рамках байесовского подхода будем понимать получение безусловной апостериорной матрицы ковариаций P (4.7).

4.3. Свойства оптимальных байесовских оценок

Определение 1.

Байесовская оценка $\tilde{x}(y)$ называется **несмещенной**, если выполнено следующее равенство

$$M_y \tilde{x}(y) = \bar{x}.$$

Заметим, что это определение отличается от того, которое было введено при рассмотрении небайесовского подхода. В то же время оно представляется вполне логичным, поскольку здесь оцениваемый вектор считается случайным.

Свойство 1. Оценка (4.5) является несмещенной. В справедливости этого нетрудно убедиться, поскольку

$$\begin{aligned} M_y \{\hat{x}(y)\} &= \int \hat{x}(y) f_y(y) dy = \iint x f(x|y) f_y(y) dx dy = \\ &= \iint x f(y|x) f_x(x) dy dx = \int x \int f(y|x) dy f_x(x) dx = M_x \{x\}. \end{aligned}$$

Свойство 2 (свойство ортогональности). Ошибка оптимальной оценки не коррелирована (ортогональна) с измерениями, т.е.

$$M_{x,y} \{(x - \hat{x}(y)) y^T\} = 0. \quad (4.8)$$

Это равенство легко доказать, если вычислять математическое ожидание последовательно, сначала по $f(x|y)$, а затем по $f(y)$.

Аналогично можно показать, что **ошибка оценки не коррелирована (ортогональна)** с оптимальной оценкой, т.е.

$$M_{x,y} \{(x - \hat{x}(y)) \hat{x}^T(y)\} = 0.$$

Свойство 3. Матрицы ковариаций ошибки оценки (4.6), (4.7) удовлетворяют следующим неравенствам:

$$\tilde{P} - P \geq 0; \quad \tilde{P}(y) - P(y) \geq 0. \quad (4.9)$$

где $\tilde{P}(y), \tilde{P}$ – условная и безусловная матрицы, задаваемые выражениями, аналогичными (4.6), (4.7), и характеризующие точность произвольной

оценки $\tilde{x}(y)$. Напомним, что выражение $P \geq 0$ для матрицы означает неотрицательную определенность соответствующей ей квадратичной формы (см. соотношение (П1.34)). В связи с приведенным неравенством можно говорить о том, что оптимальная в среднеквадратическом смысле оценка минимизирует матрицу ковариаций ошибок оценивания, а сама апостериорная матрица ковариаций P характеризует **потенциальную точность оптимального оценивания в рамках байесовского подхода**. Диагональные элементы этих матриц соответственно определяют потенциальную точность оценивания компонент x_j , $j = \overline{1..n}$.

Свойство 4. Оптимальная оценка (4.5) минимизирует определители матриц P и $P(y)$.

Свойство 5. Оптимальная оценка m -мерного вектора \tilde{x} , связанного с n -мерным вектором x линейным преобразованием $\tilde{x} = Tx$, где T – произвольная $m \times n$ матрица, определяется как $\hat{\tilde{x}}(y) = T\hat{x}(y)$, где $\hat{x}(y)$ – оптимальная оценка вектора x .

В этом нетрудно убедиться, используя (4.5).

Определение 2.

В рамках байесовского подхода также вводится понятие эффективности оценок, которое, как и при небайесовском подходе, вытекает из **неравенства Рао–Крамера**

$$P = M_{x,y} \{ (x - \hat{x}(y))(x - \hat{x}(y))^T \} \geq (I^B)^{-1}, \quad (4.10)$$

где I^B определяется как

$$I^B = M_{x,y} \left\{ \frac{\partial \ln f_{x,y}(x,y)}{\partial x} \left(\frac{\partial \ln f_{x,y}(x,y)}{\partial x} \right)^T \right\}. \quad (4.11)$$

Справедлива следующая цепочка неравенств:

$$\tilde{P} \geq P \geq (I^B)^{-1}. \quad (4.12)$$

Оценку $\hat{x}(y)$ называют **эффективной байесовской оценкой**, если в (4.12) знак неравенства переходит в равенство. Знак «больше либо равно» в неравенстве $P \geq (I^B)^{-1}$ означает, что оптимальная в среднеквадратическом смысле оценка не всегда обязана быть эффективной. В этом смысле матрица I^B определяет **предельно достижимую точность оценивания в рамках байесовского подхода**. В связи с этим матрицу $(I^B)^{-1}$, так же, как в небайесовском подходе, называют **матрицей нижней границы точности**. Следует обратить внимание на тот факт, что в рамках байесовского подхода устанавливается граница для безусловной матрицы ковариаций. Соотношение между $(I^B)^{-1}$ и матрицей $P(y)$, которая для нелинейных задач зависит от измерений, не устанавливается.

Свойство 6. Если эффективная оценка существует, то она представляет собой оптимальную в среднеквадратическом смысле оценку, определяемую в виде (4.5). [7]

4.4. Решение линейной гауссовской задачи.

Взаимосвязь с алгоритмами метода наименьших квадратов

Наиболее просто в рамках байесовского подхода решается задача оценивания в линейном гауссовском случае.

Пусть требуется найти оптимальную в среднеквадратическом смысле байесовскую оценку вектора x по измерениям (1.2), т.е. $y = Hx + v$, в которых x и v предполагаются совместно гауссовскими векторами. В силу гауссовского характера x , v и линейности измерений совместная плотность $f(x, y)$, а следовательно, и апостериорная плотность $f(x/y)$ также будут гауссовскими. Для нахождения этой плотности следует действовать так же, как и в примере, рассмотренном в п.4.1: записать совместную плотность для вектора $z = (x^T, v^T)^T$, затем найти параметры гауссовской плотности вектора $\tilde{z} = (x^T, y^T)^T$, принимая во внимание, что $\tilde{z} = Tz$, а затем найти параметры условной плотности $f(x/y)$. Вычисляя значения $\bar{y} = H\bar{x}$, P^{xy} , P^y , нетрудно конкретизировать выражения для оценки и матрицы ковариаций. В частном случае, когда векторы x и v между собой независимы, т.е. $f(z) = f(x) \cdot f(v)$, учитывая используемые обозначения $P^v = R$, $P^{x/y} = P$, получаем:

$$\bar{y} = H\bar{x}; \quad P^{xy} = P^x H^T; \quad P^y = HP^x H^T + R.$$

Отсюда следует:

$$\hat{x}(y) = \bar{x} + K(y - H\bar{x}); \quad (4.13)$$

$$K = P^x H^T (HP^x H^T + R)^{-1}; \quad (4.14)$$

$$P = P^x - P^x H^T (HP^x H^T + R)^{-1} HP^x = (E - KH)P^x. \quad (4.15)$$

С использованием леммы об обращении матриц [6]

$$\left[P^{-1} + H^T R^{-1} H \right]^{-1} = P - PH^T (HPH^T + R)^{-1} HP,$$

матрицы P и K могут быть также представлены в виде

$$P = \left((P^x)^{-1} + H^T R^{-1} H \right)^{-1}; \quad (4.16)$$

$$K = PH^T R^{-1}. \quad (4.17)$$

В линейной гауссовской задаче оптимальная в среднеквадратическом смысле оценка является эффективной, поскольку

$$\frac{\partial \ln f_{x,y}(x, y)}{\partial x} = - \left((P^x)^{-1} (x - \bar{x}) - H^T R^{-1} (y - s(x)) \right),$$

то, подставляя это выражение в (4.11), можем записать

$$P = P(y) = (I^B)^{-1} = \left((P^x)^{-1} + H^T R^{-1} H \right)^{-1}. \quad (4.18)$$

Важно подчеркнуть, что апостериорная плотность в рассматриваемой задаче является гауссовской, а алгоритм вычисления оптимальной в среднеквадратическом смысле оценки – линейным относительно измерений y . Существенно также, что условная матрица ковариаций $P(y)$ от измерений не зависит и, следовательно, она совпадет с безусловной матрицей ковариаций P , что в общем нелинейном случае не выполняется.

Установим связь оптимальных оценок с оценками, отыскиваемыми методом наименьших квадратов. Убедимся в справедливости следующего утверждения.

В линейной гауссовской задаче при отсутствии корреляции между оцениваемым вектором и вектором ошибок измерения и соответствующем выборе критерия в модифицированном методе наименьших квадратов (ММНК), т.е. когда в качестве \bar{x} принимается значение, совпадающее с математическим ожиданием вектора x , а в качестве матриц D и Q – матрицы $D = (P^x)^{-1}$, $Q = R^{-1}$, оценки ММНК совпадают с оптимальными байесовскими оценками.

Как отмечалось выше, апостериорная плотность в линейной гауссовской задаче также является гауссовской, а следовательно, она представляет собой симметричную функцию, и соответствующее ей математическое ожидание $\hat{x}(y)$, т.е. оптимальная оценка (4.5), совпадает со значением x , при котором эта плотность достигает максимального значения.

Запишем следующее выражение для совместной ф.п.р.в.:

$$f(x, y) = f(y/x) f(x) = f_v(y - Hx) f(x).$$

Таким образом, для апостериорной плотности будем иметь

$$f(x/y) = c \exp \left\{ -\frac{1}{2} \left((y - Hx)^T R^{-1} (y - Hx) + (x - \bar{x})^T (P^x)^{-1} (x - \bar{x}) \right) \right\}, \quad (4.19)$$

где c – нормирующий множитель, не зависящий от x .

Принимая во внимание такое представление для $f(x/y)$, нетрудно понять, что максимизация апостериорной плотности равносильна минимизации критерия

$$J(x) = (y - Hx)^T R^{-1} (y - Hx) + (x - \bar{x})^T (P^x)^{-1} (x - \bar{x}), \quad (4.20)$$

совпадающего с критерием, используемым в ММНК, при соответствующем выборе матриц Q и D . Этот факт и объясняет совпадение оценки ММНК при независимых x и v с оптимальной байесовской оценкой, которая в линейной гауссовской задаче, в свою очередь, совпадает с линейной оптимальной оценкой. Соответственно совпадают и матрицы ковариаций их ошибок.

4.5. Линейные оптимальные оценки.

Постановка задачи и ее общее решение

Зачастую для простоты получаемого алгоритма оценивания вводятся ограничения на класс оценок. Например, нередко считают, что отыскиваемая байесовская оценка линейным образом зависит от измерений, т.е. определяется как

$$\tilde{x}(y) = \bar{x} + K(y - \bar{y}). \quad (4.21)$$

Принимая во внимание сказанное, задачу нахождения оценки вектора x по измерениям y можно сформулировать следующим образом: найти оценку, минимизирующую математическое ожидание квадратичной функции потерь в классе линейных оценок вида (4.21). Нетрудно убедиться, что оценка, определяемая в виде (4.21), является несмещенной.

Таким образом, стоит задача нахождения **линейных несмещенных байесовских оценок с минимальной дисперсией** или, что то же самое, задача нахождения **оптимальных в среднеквадратическом смысле байесовских линейных оценок**. Далее в целях сокращения будем говорить просто о **линейных оптимальных оценках**. Алгоритм, обеспечивающий нахождение таких оценок, назовем **линейным оптимальным алгоритмом**.

Подставляя (4.21) в выражение для критерия J^B , можем записать

$$\begin{aligned} J^B &= M_{x,y} \left\{ (x - \bar{x} - K(y - \bar{y}))^T (x - \bar{x} - K(y - \bar{y})) \right\} = \\ &= M_{x,y} \left\{ Sp[(x - \bar{x} + K(y - \bar{y}))(x - \bar{x} + K(y - \bar{y}))^T] \right\} = \\ &= Sp M_{x,y} \left\{ (x - \bar{x})(x - \bar{x})^T + K(y - \bar{y})(x - \bar{x})^T + \right. \\ &\quad \left. + (x - \bar{x})(y - \bar{y})^T K^T + K(y - \bar{y})(y - \bar{y})^T K^T \right\} = \\ &= Sp \left\{ P^x + KP^{yx} + P^{xy} K^T + KP^y K^T \right\}. \end{aligned} \quad (4.22)$$

Таким образом, алгоритм вычисления оценки сводится к задаче параметрической оптимизации критерия J^B относительно матрицы K . Иными словами – не требуется нахождение самой апостериорной плотности.

Справедливо следующее утверждение.

Для того чтобы линейная оценка (4.21) при решении задачи оценивания вектора x с использованием известных фиксированных значений вектора измерений y обеспечивала минимум критерия (4.22), необходимо и достаточно, чтобы матрица K^{lin} , используемая при вычислении этой оценки, удовлетворяла уравнению (доказательство – в задаче 4.1)

$$K^{lin} P^y = P^{xy}. \quad (4.23)$$

Если матрица P^y не вырождена, то тогда из (4.23) следует

$$K^{lin} = P^{xy} (P^y)^{-1} \quad (4.24)$$

и, таким образом,

$$\hat{x}(y) = \bar{x} + K^{lin}(y - \bar{y}) = \bar{x} + P^{xy}(P^y)^{-1}(y - \bar{y}). \quad (4.25)$$

Используя (4.25), нетрудно записать выражение для матрицы ковариаций:

$$\begin{aligned} P^{lin} &= M_{x,y} \{ (x - \hat{x}(y))(x - \hat{x}(y))^T \} = \\ &= P^x + K^{lin} P^y (K^{lin})^T - K^{lin} P^{yx} - P^{xy} (K^{lin})^T. \end{aligned}$$

Преобразуя это выражение с привлечением (4.24), получаем

$$P^{lin} = P^x - P^{xy}(P^y)^{-1}P^{yx} = P^x - K^{lin}P^{yx}. \quad (4.26)$$

Таким образом, процедура вычисления оптимальных в среднеквадратическом смысле линейных оценок задается соотношениями (4.24)–(4.26) и для синтеза алгоритма необходимо располагать априорными математическими ожиданиями \bar{x} и \bar{y} и матрицами P^x , P^y , P^{xy} . Знание этих матриц обеспечивает также решение задачи анализа точности, поскольку с их использованием удастся найти апостериорную матрицу ковариаций ошибок P^{lin} , соответствующих оптимальному линейному алгоритму.

В частном случае, когда в линейной задаче в измерениях $y = Hx + v$ векторы x и v не коррелированы, а v – центрирован, нетрудно убедиться в том, что (см. задачу 4.2) соотношения (4.25), (4.26) для оптимальной линейной оценки и матрицы ковариаций ее ошибок сводятся к выражениям (4.13) – (4.17). Напомним, что эти соотношения определяют алгоритм нахождения оптимальной байесовской оценки, отыскиваемой без ограничения на класс используемых оценок в виде условного математического ожидания (4.5).

Таким образом, справедлив следующий важный вывод.

В линейной гауссовской задаче оптимальная линейная байесовская оценка и просто оптимальная байесовская оценка между собой совпадают.

Рассмотрим пример.

◆ **Пример 4.2.** Найдем выражение для оптимальной линейной оценки и соответствующей ей дисперсии скалярной случайной величины x , равномерно распределенной в интервале $\left[-\frac{b}{2}, \frac{b}{2}\right]$, по измерениям $y_i = x + v_i$, $i = \overline{1, m}$, в которых v_i – независимые между собой и от x случайные величины, равномерно распределенные в интервале $\left[-\frac{a}{2}, \frac{a}{2}\right]$.

Применительно к рассматриваемой задаче для оптимальной линейной оценки и соответствующей ей дисперсии получим:

$$\hat{x}^{lin}(y) = \bar{x} + \frac{\sigma_0^2}{r^2 + \sigma_0^2 m} \sum_{i=1}^m (y_i - \bar{x} - \bar{v}), ;$$

$$P^{lin} = \frac{\sigma_0^2 r^2}{r^2 + \sigma_0^2 m},$$

которые после подстановки значений математических ожиданий $\bar{x} = 0$, $\bar{v} = 0$ и дисперсий $\sigma_0^2 = b^2 / 12$, $r^2 = a^2 / 12$ запишем в следующем виде:

$$\hat{x}^{lin}(y) = \frac{b^2}{(a^2 + b^2 m)} \sum_{i=1}^m y_i ; \quad P^{lin} = \frac{a^2 b^2}{12(a^2 + b^2 m)}.$$

В частности, при $a \ll b$ имеем: $\hat{x}^{lin}(y) \approx \frac{1}{m} \sum_{i=1}^m y_i$; $P^{lin} \approx \frac{b^2}{12m}$.

Приведенный алгоритм будет оптимальным линейным алгоритмом не только для случайных величин с равномерным распределением, но и для любых центрированных случайных величин, имеющих произвольное распределение и заданные дисперсии $\sigma_0^2 = b^2 / 12$, $r^2 = a^2 / 12$. В частности, этот алгоритм будет оптимальным для центрированных гауссовских случайных величин с такими же дисперсиями. \blacklozenge

С учетом сказанного можно сформулировать следующий вывод.

Алгоритм вычисления линейных оптимальных оценок в линейной задаче полностью определяется первыми двумя моментами для составного вектора $z = (x^T, v^T)^T$ и не зависит от вида ф.п.р.в. $f(x, v)$. Иными словами, линейный оптимальный алгоритм в линейной задаче не зависит от характера распределений оцениваемого вектора и ошибок измерения. Заметим, что в нелинейной задаче это утверждение несправедливо.

4.6. Повышение точности оценивания при использовании нелинейных алгоритмов

В общем случае точность оценивания, достигаемая при использовании нелинейного алгоритма, может быть повышена по сравнению с точностью линейного оптимального алгоритма. Проиллюстрируем это на примере 4.2, рассмотренном в предыдущем пункте. С этой целью получим оптимальную байесовскую оценку для этого примера.

Запишем сначала выражения для $f(x, y) = f(y/x)f(x)$ и $f_y(y)$. С учетом специфики рассматриваемой задачи эти функции можно представить в виде

$$f(x, y) = \begin{cases} \frac{1}{b} f(y/x), & x \in [0, b] \\ 0, & x \notin [0, b] \end{cases} ;$$

$$f(y) = \frac{1}{b} \int_0^b f(y/x) dx.$$

Для получения значений $f(y)$ при фиксированных значениях измерений требуется располагать $f(y/x)$. Можно записать

$$f(y/x) = \prod_{i=1}^m f_v(y_i - x),$$

где
$$f(y_i/x) = f_v(y_i - x) = \begin{cases} \frac{1}{a}, & x \in [y_i - a, y_i], \\ 0, & x \notin [y_i - a, y_i] \end{cases}, \quad i = \overline{1, m},$$

и таким образом

$$f(y/x) = \begin{cases} c^*, & x \in \Omega, \\ 0, & x \notin \Omega. \end{cases} \quad (4.27)$$

В этом соотношении c^* – некоторая константа, а область Ω представляет собой отрезок, формируемый в результате пересечения всех интервалов $[y_i - a, y_i]$, $i = \overline{1, m}$, т. е.

$$\Omega \equiv \bigcap_{i=1}^m [y_i - a, y_i] = [d_1, d_2] = [y_{\max} - a, y_{\min}]. \quad (4.28)$$

Границы этого интервала $d_1 = y_{\max} - a$ и $d_2 = y_{\min}$ определяются максимальным y_{\max} и минимальным y_{\min} измеренными значениями. Учитывая (4.27), (4.28) и тот факт, что

$$\hat{x}(y) = \frac{\int_0^b xf(y/x)dx}{\int_0^b f(y/x)dx},$$

для оптимальной байесовской оценки можем записать выражение

$$\hat{x}(y) = \frac{1}{c_2 - c_1} \int_{c_1}^{c_2} x dx = \frac{1}{c_2 - c_1} \left(\frac{x^2}{2} \Big|_{c_1}^{c_2} \right) = \frac{(c_2 + c_1)}{2}, \quad (4.29)$$

в котором $[c_1, c_2]$ – отрезок, представляющий собой пересечение априорной области $[0, b]$ и области Ω , так что

$$c_1 = \max\{0, d_1\}; \quad c_2 = \min\{b, d_2\}.$$

Важно подчеркнуть при этом, что оценка (4.29) нелинейным образом зависит от измерений. Выражение (4.29) есть следствие того факта, что апостериорная плотность в рассматриваемой задаче соответствует равномерному распределению на отрезке $[c_1, c_2]$.

С учетом сказанного нетрудно показать, что

$$P(y) = \int_0^b (x - \hat{x}(y))^2 f(x/y) dx = \frac{(c_2 - c_1)^2}{12}. \quad (4.30)$$

В частности, если $d_1 = y_{\max} - a$ и $d_2 = y_{\min}$ не выходят за пределы априорной области $[0, b]$, то для оптимальной оценки и условной дисперсии будут справедливы следующие соотношения:

$$\hat{x}(y) = \frac{y_{\min} + y_{\max}}{2} - \frac{a}{2}; \quad (4.31)$$

$$P(y) = \int_0^b (x - \hat{x}(y))^2 f(x/y) dx = \frac{(y_{\min} + a - y_{\max})^2}{12}.$$

Из выражения (4.31) фактически следует, что оптимальная оценка представляет собой среднее арифметическое максимального и минимального значений измерений, подсчитанное с учетом известного математического ожидания для ошибок измерений.

В целях сопоставления точности с помощью метода статистических испытаний были рассчитаны значения среднеквадратических ошибок $\sigma^{opt}(m)$ и $\sigma^{lin}(m)$, характеризующие точность оптимальных нелинейной и линейной оценок.

В табл. 4.1 и на рис. 4.1 приведены результаты, полученные при $b=1$, $a=0,1$ и разном количестве измерений $m=1, 2, \dots, 100$, при этом число реализаций $L=1000$.

Т а б л и ц а 4.1

Значения среднеквадратических ошибок оптимальных линейных и нелинейных оценок

Число измерений	10	20	100
Линейная оптимальная оценка	0,0091	0,0065	0,0029
Нелинейная оптимальная оценка	0,0064	0,0034	0,00075

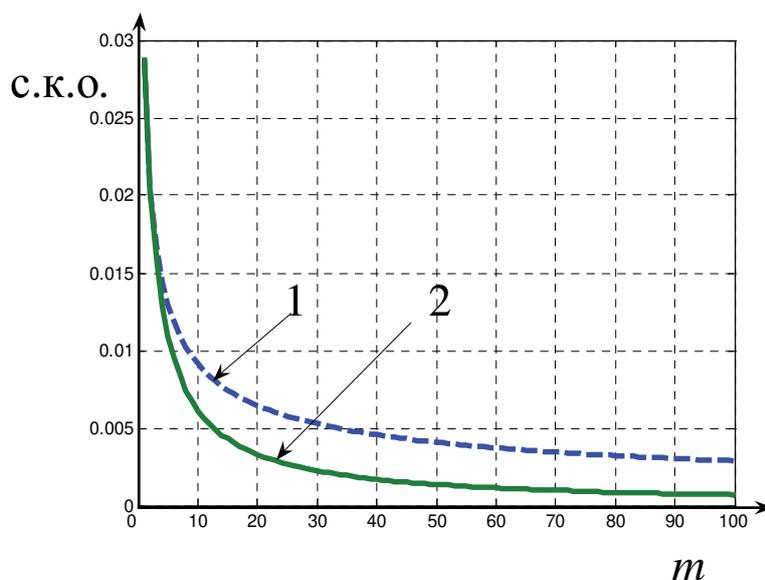


Рис. 4.1. Значения с.к.о. для оптимального линейного (1) и нелинейного (2) алгоритмов при разном числе измерений

Из представленных результатов следует, что точность нелинейного оптимального алгоритма заметно выше точности алгоритма, оптимального в классе линейных алгоритмов, что вполне согласуется с приведенным в начале подраздела утверждением. Формально такое отличие объясняется тем, что вид апостериорной функции плотности распределения вероятностей (рис. 4.2) заметно отличается от гауссовского.

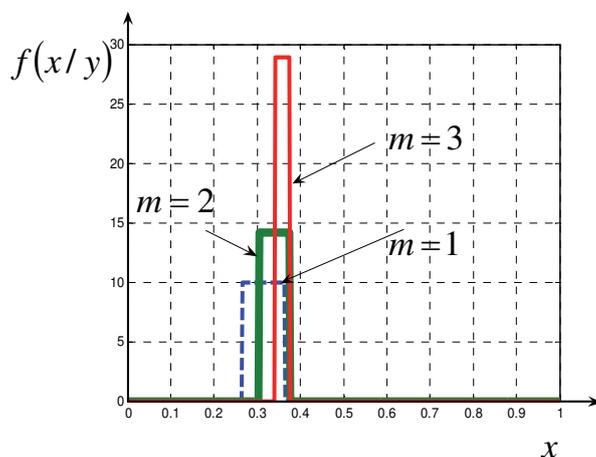


Рис. 4.2. Апостериорные ф.п.р.в. при $m=1, 2, 3$

На качественном уровне можно дать следующее объяснение такого отличия. Замечаем, что при выполнении условия $a \ll b^2 m$ алгоритм вычисления оптимальной линейной оценки фактически сведется к нахождению среднего арифметического по всем значениям $y_i - a/2$, $i = \overline{1, m}$, принадлежащим отрезку $[y_{\min} - a/2, y_{\max} - a/2]$. В то же время в оптимальном нелинейном алгоритме апостериорная область значений $[c_1, c_2]$ определяется пересечением априорной области $[0, b]$ и отрезка $[y_{\max} - a, y_{\min}]$.

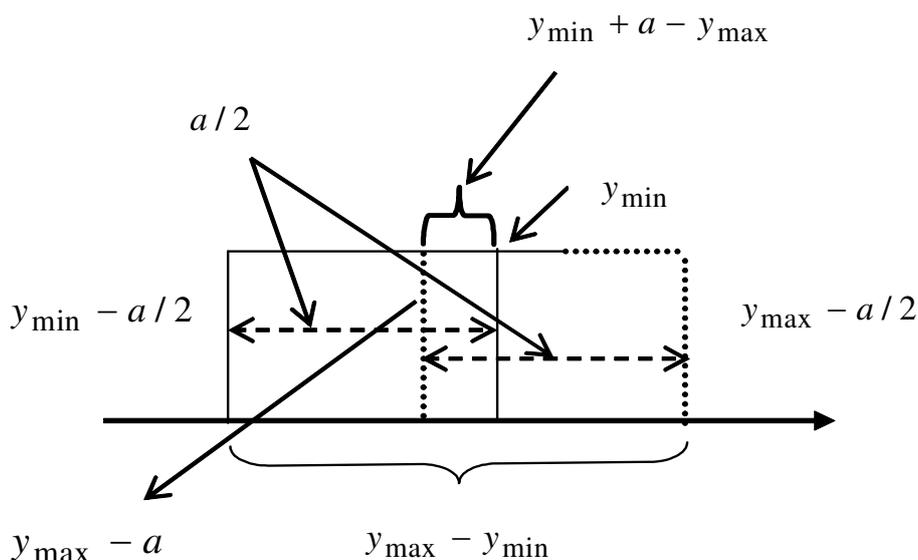


Рис. 4.3. Формирование апостериорной области для оптимальной нелинейной оценки

Очевидно, что при увеличении числа измерений повышается вероятность появления измерения с минимальным (близким к нулю) и максимальным (близким к a) значениями ошибки, т. е. $y_{\min} \rightarrow x$, а $y_{\max} \rightarrow x + a$. Отсюда следует, что длина отрезка $y_{\max} - y_{\min}$, определяющего область значений всех измерений, которые используются при вычислении среднего арифметического, стремится к величине a , задающей размер априорной области. В то же время длина отрезка $y_{\min} + a - y_{\max}$, определяющего размер апостериорной области для оптимальной нелинейной оценки, будет стремиться к нулю (рис. 4.3).

Задачи к разделу

Задача 4.1. Запишите выражение для оптимальной оценки и ее дисперсии для задачи нахождения скалярной с.в. x по измерениям $y_i = x + v_i$ при независимых x и v_i , $i = \overline{1..m}$, для которых заданы $f(x) = N(x; \bar{x}, \sigma_0^2)$, $f(v) = N(v; 0, R)$, где R – диагональная матрица с элементами r_i^2 , $i = \overline{1..m}$. Упростите полученные выражения для случая, когда $r_i^2 = r^2$, $i = \overline{1..m}$, и сопоставьте их с аналогичными выражениями, соответствующими ММНК.

Задача 4.2. Найдите такие ф.п.р.в. для ошибок измерения и оцениваемого скаляра x в условиях примера 4.1, при которых полученный там алгоритм будет оптимальным байесовским алгоритмом.

Задача 4.3. Покажите, что для оптимальной байесовской оценки (4.5) центрированного вектора x справедливо следующее соотношение:

$$M_{x,y} \hat{x}(y) \hat{x}^T(y) = P^x - P,$$

где P^x и P – априорная и апостериорная безусловная матрицы ковариаций.

Задача 4.4. Решая задачу нахождения оптимальной линейной несмещенной оценки так, как она сформулирована в подразделе 4.4, и полагая, что векторы x и y – центрированные, т.е. $\bar{x} = 0$ и $\bar{y} = 0$, докажите следующее утверждение.

Для того чтобы линейная оценка $\tilde{x}(y) = Ky$ вектора x по измерениям (1.21) обеспечивала минимум критерия (4.24), необходимо и достаточно, чтобы матрица K^{lin} , используемая при вычислении этой оценки, удовлетворяла уравнению

$$K^{lin} P^y = P^{xy}. \quad (1)$$

Задача 4.5. Запишите выражения для оптимальной в среднеквадратическом смысле линейной оценки и соответствующей ей апостериорной матрицы ковариаций ошибок, полагая, что решению подлежит линейная задача оценивания по измерениям (1.2), т.е.

$y = Hx + v$, в условиях, когда вектор оцениваемых параметров x и вектор ошибок измерения v считаются случайными коррелированными между собой векторами с нулевыми математическими ожиданиями и для них

задана матрица ковариаций в виде $P^{x,v} = \begin{bmatrix} P^x & B \\ B^T & R \end{bmatrix}$.

Важным здесь представляется тот факт, что вид плотности $f(x,v)$ не предполагается гауссовским, более того считается, что вид этой плотности не задается.

Задача 4.6. Получите выражение для оптимальной в среднеквадратическом смысле линейной оценки и соответствующей ей матрицы ковариаций, полагая, что решению подлежит предыдущая задача 4.5 в условиях, когда ошибки измерения отсутствуют, а $B = 0$.

Задача 4.7. Получите выражение для оптимальной в среднеквадратическом смысле линейной оценки и соответствующей ей матрицы ковариаций, полагая, что решению подлежит задача 4.5 при $B = 0$, а измерения имеют вид: $y = Hx + v + u$, где u – известный m -мерный вектор.

Контрольные вопросы

1. Что такое апостериорная плотность и как она может быть найдена в задаче оценивания?
2. Сформулируйте постановку задачи получения оптимальной в среднеквадратическом смысле оценки вектора x по измерениям $y = s(x) + v$ без введения ограничений на класс используемых оценок и приведите ее общее решение.
3. Перечислите свойства оптимальных байесовских оценок.
4. Как модифицируется постановка задачи получения линейной оптимальной в среднеквадратическом смысле оценки вектора x по измерениям y по сравнению с постановкой задачи, в которой не вводится ограничение на класс используемых оценок?
5. Приведите необходимые и достаточные условия оптимальности при нахождении линейной оценки вектора x по измерениям y . Опишите алгоритм нахождения линейной оптимальной оценки.
6. При каких предположениях оптимальная в среднеквадратическом смысле оценка вектора x по линейным измерениям $y = Hx + v$, совпадает с оценками, соответствующими методу наименьших квадратов?
7. Приведите пример задачи, в которой дисперсия ошибки оценивания для оптимального алгоритма существенно ниже дисперсии ошибки оценивания для линейного оптимального алгоритма.

РАЗДЕЛ 5

Алгоритмы комплексной обработки измерений

Задача комплексной (совместной) обработки измерений заключается в получении оценки искомого вектора с использованием измерений, поступающих от всех имеющихся датчиков и систем.

Цель этого раздела – с одной стороны, показать, что задача комплексной обработки по сути представляет собой обычную задачу оценивания, а с другой – обсудить специфику применения рассмотренных ранее методов для такой задачи в зависимости от уровня имеющейся априорной информации [25, 27, 31].

5.1. Примеры и постановка задачи комплексной обработки

В простейшем варианте с задачей комплексной обработки приходится иметь дело при нахождении n -мерного вектора неизвестных параметров x по данным двух датчиков (систем), показания которых могут быть представлены в следующем виде:

$$y_1 = x + v_1;$$

$$y_2 = x + v_2.$$

Здесь v_1, v_2 – ошибки каждого из датчиков. Предполагается, что датчики или системы обеспечивают измерения одних и тех же параметров. В частности, при определении координат и скорости это могут быть показания различных навигационных систем, таких, например, как инерциальные и спутниковые системы и т.п.

Нетрудно заметить, что эта задача представляет собой частный случай задачи (1.10), (1.11), т.е. оценивания x по измерениям $y = Hx + v$. В этом

легко убедиться, положив $y = (y_1^T, y_2^T)^T$, $v = (v_1^T, v_2^T)^T$, $H = \begin{bmatrix} E_n \\ E_n \end{bmatrix}$, где E_n – $n \times n$

единичная матрица.

Отсюда с очевидностью следует, что, задача комплексной обработки измерений двух датчиков (систем) представляет собой стандартную линейную задачу оценивания.

Предположим, что требуется найти оценку x по двум наборам измерений:

$$\begin{aligned}y_1 &= x + v_1, \\y_2 &= \tilde{s}(x) + v_2,\end{aligned}$$

где y_2 – вектор размерности l ; $\tilde{s}(x) = (\tilde{s}_1(x), \dots, \tilde{s}_l(x))^T$ – l -мерная нелинейная вектор-функция, зависящая от n -мерного неизвестного вектора x ; v_1 и v_2 – векторы соответствующих размерностей.

Особенность этих измерений заключается в том, что одно из них обеспечивает непосредственное измерение всех компонент искомого вектора, а другое – некоторую функцию от этого вектора x . Важно подчеркнуть, что размерность измерения y_2 в общем случае не совпадает с размерностью вектора оцениваемых параметров и в этом смысле может быть произвольной.

Вводя $l + n$ -мерную вектор-функцию $s(x)$ и вектор v

$$s(x) = \begin{bmatrix} x \\ \tilde{s}(x) \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

приходим к нелинейной задаче оценивания (1.20), (1.21), сформулированной в разделе 1.7, т.е. к оцениванию x по измерениям $y = s(x) + v$.

Таким образом, задача комплексной обработки в рассматриваемом случае представляет собой стандартную, но уже нелинейную задачу оценивания.

Как и прежде, основное внимание в дальнейшем будем уделять линейным задачам или задачам, которые могут быть сведены к линейным. Проведя линеаризацию $\tilde{s}(x)$, например в точке $x^0 = 0$, можем записать

$$\begin{aligned}y_1 &= x + v_1; \\ \tilde{y}_2 &\approx y_2 - s(0) = \tilde{H}x + v_2,\end{aligned}$$

где $\tilde{H} = \left. \frac{d\tilde{s}(x)}{dx^T} \right|_{x=0}$ – $l \times n$ матрица. После линеаризации задачу комплексной

обработки опять же нетрудно свести к стандартной линейной задаче оценивания x по измерениям $y = Hx + v$. Действительно, это легко сделать, если а в качестве H принять матрицу размерности $(l + n) \times n$, т.е.

$$H = \begin{bmatrix} E_{n \times n} \\ \left. \frac{d\tilde{s}(x)}{dx^T} \right|_{x=0} \end{bmatrix},$$

считая при этом, что $y = (y_1^T, y_2^T)^T$, $v = (v_1^T, v_2^T)^T$.

С задачей рассмотренного типа приходится сталкиваться, например, при уточнении (коррекции) показаний навигационной системы о координатах объекта с использованием некоторых корректирующих измерений. В качестве таких измерений могут быть использованы, например, дальности до точечных ориентиров. В частном случае, полагая, что $x = (x_1, x_2)^T$ – двухмерный вектор координат объекта на плоскости, и име-

ются навигационная система, измеряющая $y_1 = x + v_1$, где $y_1 = (y_1^{(1)}, y_1^{(2)})^T$, $v_1 = (v_1^{(1)}, v_1^{(2)})^T$, и корректирующее измерение ($l = 1$) дальности до точечного ориентира, можем записать:

$$\begin{aligned} y_1^{(1)} &= x_1 + v_1^{(1)}; \\ y_1^{(2)} &= x_2 + v_1^{(2)}; \\ y_2 &= \sqrt{(x_1^o - x_1)^2 + (x_2^o - x_2)^2} + v_2, \end{aligned}$$

где $x = (x_1^o, x_2^o)^T$ – известные координаты точечного ориентира.

После линеаризации функции $s(x)$, например, в точке $x^* = 0$, корректирующее измерение y_2 можно представить как

$$\tilde{y}_2(x^*) = y_2 - s(0) = \tilde{H}x + v_2 = -x_1 \sin \Pi - x_2 \cos \Pi + v_2.$$

Таким образом, вводя $v = (v_1^{(1)}, v_1^{(2)}, v_2)^T$, $H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\sin \Pi & -\cos \Pi \end{bmatrix}$, нетрудно

все измерения $y = (y_1^{(1)}, y_1^{(2)}, y_2)^T$ записать как $y = Hx + v$. Здесь Π – угол, определяющей ориентацию единичного вектора $(\sin \Pi, \cos \Pi)^T$ относительно оси Ox_2 .

В наиболее общем случае задача комплексной обработки в линейной постановке может быть сформулирована следующим образом:

оценить постоянный n -мерный вектор x , используя набор из m векторных измерений вида

$$y_j = H_j x + v_j, \quad j = \overline{1, m},$$

в которых H_j – $m_j \times n$ матрицы; v_j – m_j -мерные векторы ошибок измерения; j – номер измерения.

Введя $m_\Sigma \times n$ матрицу H , где $m_\Sigma = \sum_{j=1}^m m_j$, составные векторы измерений

y и их ошибок v в виде

$$H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_m \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix},$$

замечаем, что и в этом общем случае задача сводится к обычной постановке – оценить x по $y = Hx + v$.

Обсудим в следующих разделах различные варианты построения алгоритмов комплексной обработки.

5.2. Комплементарный фильтр. Инвариантная схема обработки

При наличии данных от двух и более измерителей при решении задачи комплексной обработки нередко используется так называемый комплементарный фильтр. Такой фильтр используется в условиях, когда априорная информация о векторе x отсутствует и предполагается лишь наличие априорной статистической информации об ошибках измерения. **Суть комплементарного (инвариантного) фильтра** заключается в использовании разностных измерений Δy , не содержащих искомого вектора x , с целью получения оценки ошибок одного измерителя v_1 на фоне ошибок другого измерителя v_2 (рис. 5.1).

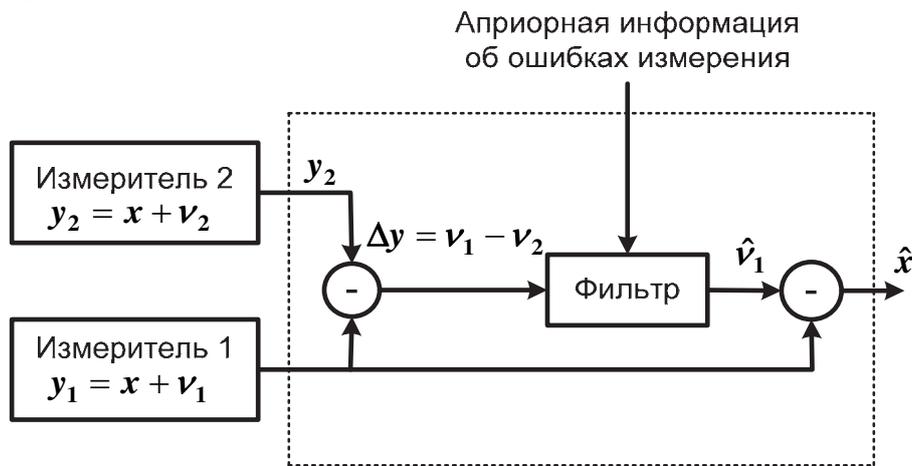


Рис. 5.1. Комплементарный фильтр, инвариантная схема комплексной обработки.

Проиллюстрируем применение этого фильтра и проанализируем свойства получающихся при этом оценок на примере простейшей задачи комплексной обработки измерений

$$y_1 = x + v_1;$$

$$y_2 = x + v_2,$$

рассмотренной в предыдущем подразделе.

Будем полагать, что ошибки измерения v_1 , v_2 являются не коррелированными между собой центрированными векторами с известными матрицами ковариаций $R_j > 0$, $j = 1, 2$.

Для построения комплементарного фильтра сформируем разностные измерения [31]

$$\Delta y = y_1 - y_2 = v_1 - v_2, \quad (5.1)$$

из которых исключается отыскиваемый вектор x , и используем их для оценивания ошибок одного измерителя на фоне ошибок другого измерителя. Для определенности будем считать, что в качестве оцениваемого вектора выступает v_1 , а в качестве ошибок измерения – вектор v_2 .

Наличие информации о математических ожиданиях и матрицах ковариаций ошибок v_1, v_2 позволяет, опираясь на результаты подраздела 4.5, сформулировать задачу получения для v_1 оптимальной в среднеквадратическом смысле линейной оценки и соответствующей ей матрицы ковариаций по измерениям Δy , которые, как нетрудно показать, могут быть определены в виде

$$\hat{v}_1 = (R_1^{-1} + R_2^{-1})^{-1} R_2^{-1} (y_1 - y_2); \quad (5.2)$$

$$P^{v_1} = (R_1^{-1} + R_2^{-1})^{-1}. \quad (5.3)$$

Оценку искомого вектора x с использованием данных от комплементарного фильтра сформируем как

$$\hat{x} = y_1 - \hat{v}_1 = x + v_1 - \hat{v}_1 = (R_1^{-1} + R_2^{-1})^{-1} [(R_1^{-1} + R_2^{-1})y_1 - R_2^{-1}(y_1 - y_2)].$$

Это выражение легко преобразовать к виду

$$\hat{x} = (R_1^{-1} + R_2^{-1})^{-1} [R_1^{-1}y_1 + R_2^{-1}y_2]. \quad (5.4)$$

Очевидно, что ошибка оценки $\varepsilon = x - \hat{x} = v_1 - \hat{v}_1$ не зависит (**инвариантна**) от оцениваемого вектора x и, таким образом, обладает относительно него свойством инвариантности. Именно это и определило название для такой схемы обработки – **инвариантная схема**. Соответствующие алгоритмы получили также наименование **инвариантных алгоритмов обработки**. Иногда говорят об инвариантном характере оценок. Матрица ковариаций для полученной таким образом ошибки оценки x будет совпадать с матрицей (5.3) для ошибок оценки v_1 .

Запишем измерения y_1, y_2 как $y = Hx + v$, где $y = (y_1^T, y_2^T)^T, v = (v_1^T, v_2^T)^T, H = \begin{bmatrix} E_n \\ E_n \end{bmatrix}$ и $R^v = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$. Для решения задачи получения оценки x при сделанных предположениях может быть использован ОМНК, описанный в разделе 2. Полагая в критерии (2.13) $Q = R^{-1}$, нетрудно убедиться, что оценка (5.4) совпадает с оценкой ОМНК (задача 5.1).

Из сказанного следует, что комплементарный фильтр (**инвариантную схему**) в рассматриваемом случае можно трактовать как **специальным образом организованную процедуру получения оценок, соответствующих ОМНК**. Ее особенность заключается в том, что не требуется вводить каких-либо предположений о статистических свойствах искомого вектора x при получении алгоритма оценивания.

Заметим, что, если дополнительно предположить, что v_1 и v_2 – гауссовские, то оценка (5.2) для ошибки v_1 первого измерения становится оптимальной байесовской оценкой, а оценка (5.4) будет совпадать, в свою

очередь, с оценкой, максимизирующей функцию правдоподобия. Иными словами, оценка, полученная с использованием инвариантного алгоритма, будет соответствовать оценке МФП. В этих условиях можно говорить о том, что комплементарный фильтр (инвариантная схема) обеспечивает получение оценок, соответствующих МФП, ошибка которой также не зависит от вектора оцениваемых параметров.

Описанный выше прием может быть использован и при обработке измерений, заданных в виде

$$y_1 = x + v_1; \quad (5.5)$$

$$y_2 = Hx + v_2. \quad (5.6)$$

Для получения комплементарного фильтра (инвариантного алгоритма) в этом случае следует сформировать разностное измерение

$$\Delta y = y_2 - Hy_1 = -Hv_1 + v_2.$$

Далее, считая ошибки v_1, v_2 не коррелированными между собой центрированными векторами с известными матрицами ковариаций $R_j > 0, j = 1, 2$, нетрудно в рамках байесовского подхода получить следующие выражения для линейной оптимальной оценки и соответствующей ей ошибки вектора v_1 по измерениям $\Delta y = y_2 - Hy_1$:

$$\hat{v}_1 = \left(R_1^{-1} + H^T R_2^{-1} H \right)^{-1} H^T R_2^{-1} (y_2 - Hy_1),$$

$$P^{v_1} = \left(R_1^{-1} + H^T R_2^{-1} H \right)^{-1}.$$

Вычитая полученное значение оценки ошибки первого измерителя \hat{v}_1 из его показаний y_1 , для оценки x получаем

$$\hat{x} = y_1 + \left(R_1^{-1} + H^T R_2^{-1} H \right)^{-1} H^T R_2^{-1} (y_2 - Hy_1). \quad (5.7)$$

Нетрудно убедиться (задача 5.2) в том, что и в этом случае полученная таким образом оценка и соответствующая ей матрица ковариаций будут соответствовать оценкам ОМНК или МФП, если предположить гауссовский характер ошибок измерений.

Важно подчеркнуть, что для построения инвариантной схемы среди используемых измерений можно было бы выделить хотя бы одно измерение типа $y = x + v$, обеспечивающее непосредственное измерение искомого вектора x .

5.3. Неинвариантная схема обработки

Следует обратить внимание на тот факт, что описанная в подразделе 5.1 инвариантная схема обеспечивает нахождение оптимальных в среднеквадратическом смысле байесовских оценок по разностным измерениям Δy лишь для ошибок измерений, а не для искомого параметра. Для получения оптимальной байесовской оценки искомого вектора x требуется ввести

предположение о его случайном характере. В этом случае задачу комплексной обработки целесообразно сводить к традиционной постановке задачи оценивания вектора x по измерениям $y=Hx$ и использовать для решения байесовские алгоритмы, описанные в подразделах 4.4. или 4.5. Поскольку ошибка байесовской оценки $x - \hat{x}(y) = (E - KH)(x - \bar{x}) + Kv$ зависит не только от ошибок измерения, но и от самого оцениваемого параметра, алгоритм, с помощью которого отыскивается эта оценка, называют иногда **неинвариантным алгоритмом**. Соответствующая этому алгоритму блок-схема приведена на рис. 5.2. Здесь, в отличие от предыдущего случая, оба измерения равноправно используются в алгоритме, в котором привлекается априорная информация как об ошибках измерения, так и о самом оцениваемом векторе.

Заметим, что инвариантная схема также может быть представлена в виде блок-схемы, в которой каждое измерение используется равноправно, если под алгоритмом понимать часть, выделенную пунктиром (см. рис. 5.1).

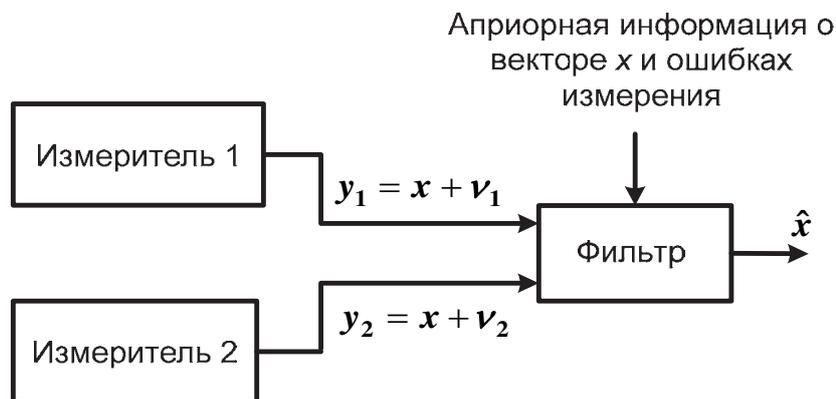


Рис. 5.2. Неинвариантная (байесовская) схема построения алгоритма комплексной обработки

Принципиальное отличие инвариантного и неинвариантного алгоритмов заключается в том, что в первом случае используется априорная информация только о стохастических свойствах ошибках измерения, а во втором – привлекается еще информация и о векторе оцениваемых параметров.

Достоинство инвариантной схемы заключается в том, что не требуется вводить каких-либо предположений о статистических свойствах относительно оцениваемого вектора. В ряде случаев это оказывается вполне оправданным, так как зачастую ввести адекватное описание для x бывает затруднительно. Вместе с тем следует помнить, что при наличии информации о векторе оцениваемых параметров отказ от нее может привести к существенной потере в точности. Рассмотрим пример, иллюстрирующий возможный проигрыш в точности в случае неучета априорной информации о свойствах оцениваемого вектора.

◆ **Пример 5.1.** Предположим, что в моменты времени t_i ($i = \overline{1..m}$) проведены измерения высоты летательного аппарата не только с использованием спутниковой системы, но и с данными от баровысотомера. Запишем эти измерения в виде

$$y_i^{CHC} = h_i + v_i^{CHC}, \quad i = \overline{1..m}; \quad (5.8)$$

$$y_i^{BB} = h_i + v_i^{BB}, \quad i = \overline{1..m}. \quad (5.9)$$

Введем векторы $x = (h_1, \dots, h_m)^T$, $v^{CHC} = (v_1^{CHC}, v_2^{CHC}, \dots, v_m^{CHC})^T$, $v^{BB} = (v_1^{BB}, v_2^{BB}, \dots, v_m^{BB})^T$ и матрицу $H = E_m$. Ясно, что, используя такие обозначения, эти измерения могут быть представлены в виде $y = Hx + v$.

Будем для начала считать, что информация о статистических свойствах оцениваемого вектора отсутствует и такая информация имеется лишь для векторов ошибок различных измерителей. При этом полагаем, что они гауссовские, центрированы и не коррелированы между собой, а матрицы ковариаций для каждого из них заданы как $R_j = r_j^2 E_m$, $j = CHC, BB$.

Решим задачу с использованием комплементарного фильтра (инвариантной схемы) и разностных измерений

$$\Delta y_i = y_i^{CHC} - y_i^{BB} = v_i^{CHC} - v_i^{BB}.$$

Используя (4.13), нетрудно получить выражение для оценки (задача 5.3) и убедиться в том, что матрица ковариаций ошибок оценивания для вектора x будет диагональной, и при $r_j^2 = r$, $j = CHC, BB$ для нее будет справедли-

во выражение $P^{мфп} = \frac{r^2}{2} E_m$. Отсюда следует, что дисперсия ошибки измерения высоты в каждый дискретный момент снизится лишь в два раза.

Будем теперь полагать, что имеется априорная информация о векторе. В частности, предположим, что на интервале наблюдения высота постоянна и представляет собой центрированную гауссовскую случайную величину с дисперсией σ_0^2 . В этом случае, используя материалы подраздела 4.4, можно получить оптимальную байесовскую оценку и соответствующую ей дисперсию ошибки оптимальной оценки (задача 5.4). При сделанных предположениях можем записать матрицу ковариаций в виде $P^x = \sigma_0^2 I_{m \times m}$, где $I_{m \times m}$ – матрица, составленная из единичек. Поскольку в силу вырожденности $P^x = \sigma_0^2 I_{m \times m}$ применить выражение для вычисления матрицы ковариаций ошибок оценивания в виде (4.16) не представляется возможным, используем выражение (4.15). Учитывая специфику матрицы $H = E_m$, получаем

$$P = \sigma_0^2 I_{m \times m} \left(E_m - (\sigma_0^2 I_{m \times m} + 2r^2 E_m)^{-1} \sigma_0^2 I_{m \times m} \right).$$

Принимая во внимание вытекающее из (П1.59) соотношение

$$(\sigma_0^2 I_{m \times m} + 2r^2 E_m)^{-1} = \frac{1}{2r^2} \left(E_m - \frac{2\sigma_0^2}{2m\sigma_0^2 + r^2} I_{m \times m} \right) I_{m \times m},$$

можно показать, что диагональные элементы матрицы ковариаций для оцениваемого вектора, собственно говоря, определяющие точность оценивания высоты, будут задаваться соотношением

$$\sigma_h^2 = \frac{\sigma_0^2 r^2}{2m\sigma_0^2 + r^2}. \quad (5.10)$$

Смысл этого выражения вполне понятен, так как при сделанных предположениях, в сущности, решается задача оценивания постоянной скалярной величины по $2m$ измерениям с независимыми ошибками измерения с дисперсиями r^2 .

Пренебрегая вкладом априорной информации и полагая $r^2 \ll \sigma_0^2$, получаем, что дисперсия ошибки определения высоты снизится в $2m$ раз, т.е. $\sigma_h^2 \approx \frac{r^2}{2m}$, что существенно выше, чем в предыдущем случае. ♦

5.4. Централизованная и децентрализованная схемы обработки

Рассмотрим сформулированную выше задачу комплексной обработки измерений

$$y_j = H_j x + v_j, \quad j = \overline{1, i}, \quad (5.11)$$

в которых H_j – $m_j \times n$ матрицы; v_j – m_j -мерные векторы ошибок измерения; j – номер измерения.

Введем следующие обозначения:

$$Y_i = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \end{bmatrix}, \quad V_i = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \end{bmatrix}, \quad \mathbf{H}_i = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_i \end{bmatrix},$$

где векторы измерений Y_i и их ошибок V_i – размерности $m_\Sigma = m_\Sigma = \sum_{j=1}^i m_j$, а

матрица \mathbf{H}_i – размерности $m_\Sigma \times n$.

Запишем измерения (5.11) как

$$Y_i = \mathbf{H}_i x + V_i.$$

Обсудим решение этой задачи в предположении, что x и v_j – случайные центрированные векторы с заданными матрицами ковариаций P^x , R_j , $j = \overline{1, i}$. В целях простоты считаем, что все эти векторы не коррелирова-

ны между собой. В этом случае матрица ковариаций \mathbf{R}_i для составного вектора V_i будет иметь следующий вид:

$$\mathbf{R}_i = \begin{bmatrix} R_1 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & R_i \end{bmatrix}.$$

При сделанных предположениях для решения этой задачи можно использовать оптимальный байесовский алгоритм и получить следующие выражения для оптимальной в среднеквадратическом смысле оценки и соответствующей ей матрицы ковариаций:

$$\hat{x}_i(Y_i) = P_i \mathbf{H}_i^T \mathbf{R}_i^{-1} Y_i, \quad (5.12)$$

$$P_i = \left((P^x)^{-1} + \sum_{j=1}^i H_j^T R_j^{-1} H_j \right)^{-1}. \quad (5.13)$$

Отсюда имеем

$$\hat{x}_i(Y_i) = \left((P^x)^{-1} + \sum_{j=1}^i H_j^T R_j^{-1} H_j \right)^{-1} \left(\sum_{j=1}^i H_j^T R_j^{-1} y_j \right). \quad (5.14)$$

При получении оптимальной оценки по всему имеющемуся набору измерений могут быть использованы две схемы обработки. Одна из них получила название **централизованной схемы обработки** (рис. 5.3). Такое название объясняется тем, что выполнение всех вычислений предполагается осуществлять централизованно в одном алгоритме.

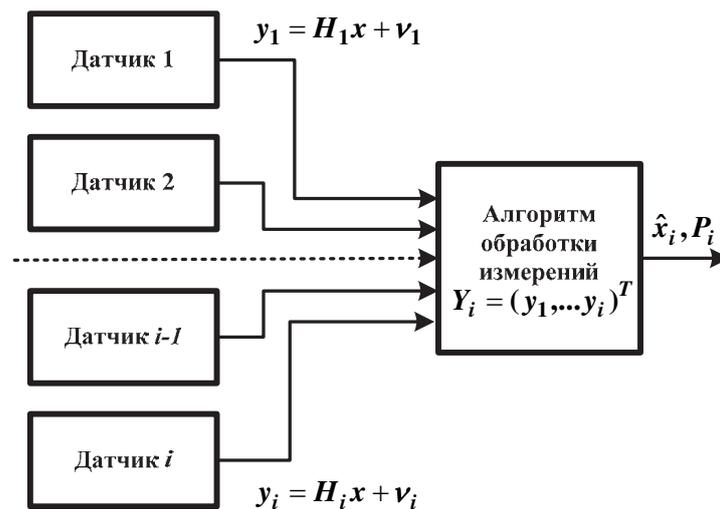


Рис. 5.3. Централизованная схема вычисления оптимальной оценки

Вместе с тем при вычислении искомой оценки может быть использована и другая схема.

Предположим, что справедливо неравенство

$$(P^x)^{-1} \ll H_j^T R_j^{-1} H_j, \quad j = \overline{1, i}, \quad (5.15)$$

означающее, что априорной информацией о векторе x можно пренебречь. Введем частные оптимальные оценки

$$\hat{x}^{(j)} \cong (P^{(j)}) H_j^T R_j^{-1} y_j, \quad (5.16)$$

где

$$P^{(j)} = (H_j^T R_j^{-1} H_j)^{-1}, \quad (5.17)$$

т. е. оценки искомого вектора, полученные с использованием данных только частных j -х измерений.

Принимая во внимание (5.14) (5.16), легко убедиться в том, что выражение для оптимальной оценки, полученной по всему набору измерений, может быть преобразовано к виду

$$\hat{x}_i \cong \left(\sum_{j=1}^i (P^{(j)})^{-1} \right)^{-1} \left(\sum_{j=1}^i (P^{(j)})^{-1} \hat{x}^{(j)} \right). \quad (5.18)$$

Иными словами, оптимальная оценка, полученная с использованием всего набора измерений, представляет собой «взвешенную» сумму частных оптимальных оценок.

Предположение (5.15) можно и не вводить, если при получении частных оценок вместо (5.17) использовать соотношения

$$P^{(j)} = \left(\frac{1}{i} (P^x)^{-1} + H_j^T R_j^{-1} H_j \right)^{-1}. \quad (5.19)$$

Блок-схема, соответствующая этому варианту, приведена на рис. 5.4. Эта схема называется **децентрализованной** в силу того, что вычисления частных оценок и соответствующих им матриц ковариаций может осуществляться в отдельных блоках, а искомая оценка затем определяется путем взвешивания этих частных оценок.

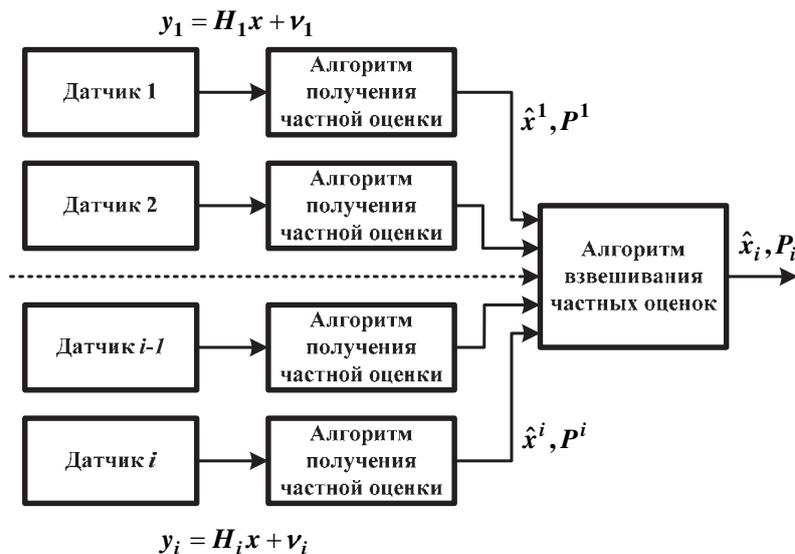


Рис. 5.4. Децентрализованная схема обработки

Заметим, что, если дополнительно предположить, что x и v_j являются гауссовскими векторами, то полученные выше оценки будут оптимальными байесовскими оценками, минимизирующими среднеквадратический критерий без ограничения на класс используемых алгоритмов.

5.5. Рекуррентная схема обработки

Наибольшее применение при решении задач обработки навигационной информации получила рекуррентная схема. Суть этой схемы заключается в том, что искомая оценка получается не в результате обработки всего набора измерений, а формируется путем последовательной обработки каждого имеющегося измерения и результатов, полученных на предыдущем шаге обработки. Соответствующие алгоритмы получили название **рекуррентных алгоритмов**.

Идею получения таких алгоритмов легко пояснить на простейшем примере оценивания скалярной величины по скалярным измерениям

$$y_i = x + v_i.$$

Если считать, что ошибки измерения представляют собой некоррелированные случайные величины с одинаковыми дисперсиями, то в этом случае оценку, соответствующую, например, ОМНК, можно определить в виде среднего арифметического накопленных измерений, т.е. $\hat{x}_i = \frac{1}{i} \sum_{j=1}^i y_j$. За-

писывая цепочку равенств

$$\hat{x}_i = \frac{1}{i} \sum_{j=1}^i y_j = \frac{y_i + \sum_{j=1}^{i-1} y_j}{i} = \frac{y_i}{i} + \frac{i-1}{i} \hat{x}_{i-1} = \hat{x}_{i-1} + \frac{1}{i} (y_i - \hat{x}_{i-1}),$$

получаем

$$\hat{x}_i = \hat{x}_{i-1} + \frac{1}{i} (y_i - \hat{x}_{i-1}). \quad (5.20)$$

По аналогии получим рекуррентный алгоритм применительно к рассмотренной в предыдущем подразделе задаче, полагая, что измерения (5.11) поступают в схему обработки последовательно. Введем оценку и матрицу ковариаций ее ошибок, соответствующую обработке $(i-1)$ -го измерения в виде

$$\hat{x}_{i-1}(Y_{i-1}) = \left((P^x)^{-1} + \sum_{j=1}^{i-1} H_j^T R_j^{-1} H_j \right)^{-1} \left(\sum_{j=1}^{i-1} H_j^T R_j^{-1} y_j \right), \quad (5.21)$$

$$P_{i-1} = \left((P^x)^{-1} + \sum_{j=1}^{i-1} H_j^T R_j^{-1} H_j \right)^{-1}. \quad (5.22)$$

Тогда можно записать следующую цепочку равенств:

$$\begin{aligned}
\hat{x}_i(Y_i) &= P_i \left(\sum_{j=1}^i H_j^T R_j^{-1} y_j \right) = P_i \left(H_i^T R_i^{-1} y_i + \sum_{j=1}^{i-1} H_j^T R_j^{-1} y_j \right) = \\
&= P_i \left(H_i^T R_i^{-1} y_i + P_{i-1}^{-1} P_{i-1} \left(\sum_{j=1}^{i-1} H_j^T R_j^{-1} y_j \right) \right) = P_i \left(H_i^T R_i^{-1} y_i + P_{i-1}^{-1} \hat{x}_{i-1}(Y_{i-1}) \right) = \\
&= P_i \left(H_i^T R_i^{-1} y_i + P_{i-1}^{-1} \hat{x}_{i-1}(Y_{i-1}) \right) + P_i H_i^T R_i^{-1} \left(H_i \hat{x}_{i-1}(Y_{i-1}) - H_i \hat{x}_{i-1}(Y_{i-1}) \right) = \\
&= P_i H_i^T R_i^{-1} (y_i - H_i \hat{x}_{i-1}(Y_{i-1})) + P_i \left(P_{i-1}^{-1} + H_i^T R_i^{-1} H_i \right) \hat{x}_{i-1}(Y_{i-1}) = \\
&= \hat{x}_{i-1}(Y_{i-1}) + P_i H_i^T R_i^{-1} (y_i - H_i \hat{x}_{i-1}(Y_{i-1})); \\
P_i &= \left((P^x)^{-1} + \sum_{j=1}^i H_j^T R_j^{-1} H_j \right)^{-1} = \\
&= \left((P^x)^{-1} + \sum_{j=1}^{i-1} H_j^T R_j^{-1} H_j + H_i^T R_i^{-1} H_i \right)^{-1} = \left(P_{i-1}^{-1} + H_i^T R_i^{-1} H_i \right)^{-1},
\end{aligned}$$

из которых с очевидностью вытекают искомые рекуррентные соотношения:

$$\hat{x}_i(Y_i) = \hat{x}_{i-1}(Y_{i-1}) + P_i H_i^T R_i^{-1} (y_i - H_i \hat{x}_{i-1}(Y_{i-1})); \quad (5.23)$$

$$P_i = \left(P_{i-1}^{-1} + H_i^T R_i^{-1} H_i \right)^{-1}. \quad (5.24)$$

Важно подчеркнуть, что эти соотношения носят рекуррентный характер как для оценки, так и для матрицы ковариаций. Если ввести матрицу

$$K_i = P_i H_i^T R_i^{-1}, \quad (5.25)$$

то выражение для оценки можно представить как

$$\hat{x}_i = \hat{x}_{i-1}(Y_{i-1}) + K_i (y_i - H_i \hat{x}_{i-1}(Y_{i-1})). \quad (5.26)$$

Принимая во внимание выражения (4.14), (4.15), матрицу ковариаций ошибок оценивания P_i и матрицу K_i можно вычислять как

$$P_i = P_{i-1} - P_{i-1} H_i^T (H_i P_{i-1} H_i^T + R_i)^{-1} H_i P_{i-1} = (E - K_i H_i) P_{i-1}, \quad (5.27)$$

$$K_i = P_{i-1} H_i^T (H_i P_{i-1} H_i^T + R_i)^{-1}. \quad (5.28)$$

Как будет показано в следующих разделах, фактически это есть соотношения фильтра Калмана для рассматриваемой здесь задачи оценивания постоянного вектора.

Задачи к разделу

Задача 5.1.

Решите задачу оценивания вектора по измерениям

$$y_1 = x + v_1;$$

$$y_2 = x + v_2,$$

с использованием метода обобщенных наименьших квадратов, полагая, что в критерии (2.13) $Q = \begin{bmatrix} R_1^{-1} & 0 \\ 0 & R_2^{-1} \end{bmatrix}$.

Полагая, что ошибки измерения представляют не коррелированные между собой центрированные векторы с матрицами ковариаций R_1 и R_2 , запишите, используя (2.29), матрицу ковариаций для ошибок полученной оценки

Сопоставьте полученные соотношения с результатами, соответствующими комплементарному фильтру для рассматриваемой задачи.

Задача 5.2. Используя соотношения, соответствующие ОМНК в задаче нахождения неизвестного n -мерного вектора x по измерениям типа (5.5), (5.6), записываемым в виде

$$\begin{aligned} y_1 &= x + v_1; \\ y_2 &= Hx + v_2, \end{aligned}$$

где y_2 – вектор размерности l ; H – матрица $l \times n$, конкретизируйте выражения для оценки этого вектора, считая, что $Q = R^{-1}$.

Дополнительно полагая, что v_1 и v_2 – некоррелированные между собой центрированные векторы с известными матрицами ковариаций $R_j > 0$, $j = 1, 2$, а R – матрица ковариаций вектора $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, получите с использованием (2.29) выражение для матрицы ковариаций ошибок. Убедитесь в том, что полученные оценки совпадают с оценками, соответствующими комплементарному фильтру. При каких условиях эта оценка будет соответствовать оценке МФП?

Задача 5.3. В примере 5.1 получите выражение для оценки вектора x по измерениям (5.8.), (5.9) и матрицы ковариаций ошибок в случае использования для решения задачи комплементарного фильтра.

Задача 5.4. В примере 5.1 получите выражение для оценки вектора x по измерениям (5.8.), (5.9) и матрицы ковариаций ошибок в случае использования для решения задачи оптимального байесовского алгоритма.

Задача 5.5. В моменты времени $t_i = \Delta t(i-1)$, $i = \overline{1, m}$ с равным интервалом Δt проведены измерения высоты летательного аппарата с использованием спутниковой системы и данных от баровысотомера. Ошибки измерения представляют не зависимые между собой для разных моментов времени случайные величины с дисперсиями r_{CHC}^2 и r_{BB}^2 . Высота объекта описывается в виде полинома первой степени $h_i = x_0 + Vt_i$.

Сформулируйте задачу комплексной обработки этих измерений в целях получения оптимальных оценок вектора $x = (x_1, x_2)^T = (x_0, V)^T$, полагая, что его компоненты являются не зависимыми между собой и от ошибок измерений гауссовскими с.в. с математическим ожиданием $(\bar{x}_0, 0)^T$ и дисперсиями σ_0^2, σ_V^2 . Запишите выражение для матрицы ковариаций ошибок оптимальных оценок.

Задача 5.6. Рассмотрим обсуждавшуюся выше задачу коррекции показаний навигационной системы о двухмерных координатах объекта по измерению дальности до одного точечного ориентира, полагая, что она может быть решена в линеаризованной постановке и, таким образом, измерения могут быть записаны в следующем виде:

$$y_1^{(1)} = x_1 + v_1^{(1)};$$

$$y_1^{(2)} = x_2 + v_1^{(2)};$$

$$y_2 = \tilde{H}x + v_2 = -x_1 \sin \Pi - x_2 \cos \Pi + v_2,$$

где $x = (x_1, x_2)^T$ – двухмерный вектор, задающий координаты объекта на плоскости, а угол Π задает ориентацию единичного вектора $-(\sin \Pi, \cos \Pi)^T$ относительно оси Ox_2 .

Пусть $v_1 = (v_1^{(1)}, v_1^{(2)})^T$ – центрированный вектор ошибок измерений с матрицей ковариаций R , которой соответствуют параметры эллипса ошибок a, b, τ , а ошибка v_2 – не коррелированная с этим вектором центрированная случайная величина с дисперсией, совпадающей с b^2 .

Конкретизируйте комплементарный (инвариантный) фильтр для оценки координат и убедитесь в том, что он соответствует ОМНК при $Q = R^{-1}$.

Найдите такое значение угла Π , при котором величина радиальной среднеквадратической ошибки будет наименьшей.

Задача 5.7. Имеются две системы, вырабатывающие измерения

$$y_1 = x + v_1;$$

$$y_2 = x + v_2,$$

в которых $x = (x_1, x_2)^T$ представляет собой двухмерный вектор, задающий координаты объекта на плоскости. Предполагается, что двухмерные векторы ошибок измерения являются некоррелированными центрированными случайными векторами с матрицами ковариаций $R_1 > 0, R_2 > 0$.

Получите выражение для матрицы ковариаций ошибок оценок, соответствующих ОМНК, полагая матрицу Q блочно-диагональной с блоками R_1^{-1} и R_2^{-1} .

Используйте предположение о том, что размеры малой и большой полуосей эллипсов, соответствующих этим матрицам, одинаковые, т.е.

$a_1 = a_2 = a$, и $b_1 = b_2 = b$, и определите, при какой взаимной ориентации этих эллипсов радиальная среднеквадратическая ошибка будет наименьшей (наибольшей).

Контрольные вопросы

1. Сформулируйте задачу комплексной обработки и проиллюстрируйте ее на примерах. Что такое комплементарный фильтр (инвариантный алгоритм)?
2. На примере обработки данных от двух измерителей поясните суть инвариантного алгоритма. Почему используется такое название? Приведите соответствующую ему схему обработки измерений. Какова его связь с алгоритмом ОМНК и МФП?
3. Поясните, в чем особенность неинвариантного алгоритма. Приведите соответствующую ему схему обработки измерений. Обсудите достоинства и недостатки этого алгоритма по сравнению с инвариантным алгоритмом.
4. Поясните, в чем особенность централизованной и децентрализованной схем обработки данных от нескольких измерителей.
5. Поясните, каким образом могут быть получены рекуррентные алгоритмы оценивания в задаче оценивания вектора постоянных параметров.

РАЗДЕЛ 6

Случайные последовательности

В этом разделе приводятся основные определения, связанные с понятием случайных последовательностей, и обсуждаются их свойства и методы описания [18, 21, 25, 27, 31-34].

6.1. Определение случайной последовательности и ее описание

Рассмотрим последовательность скалярных величин $x_1, x_2, x_3, \dots, x_i, \dots$. **Последовательность называется случайной**, если каждая из входящих в нее величин x_i является случайной. На рис. 6.1 приведен пример реализации случайной последовательности. Если полагать, что нижний индекс соответствует значениям времени t_i , то в этом случае будем говорить о временной случайной последовательности. В дальнейшем речь будет идти о временных случайных последовательностях, и нижний индекс будет трактоваться как значение дискретного времени.

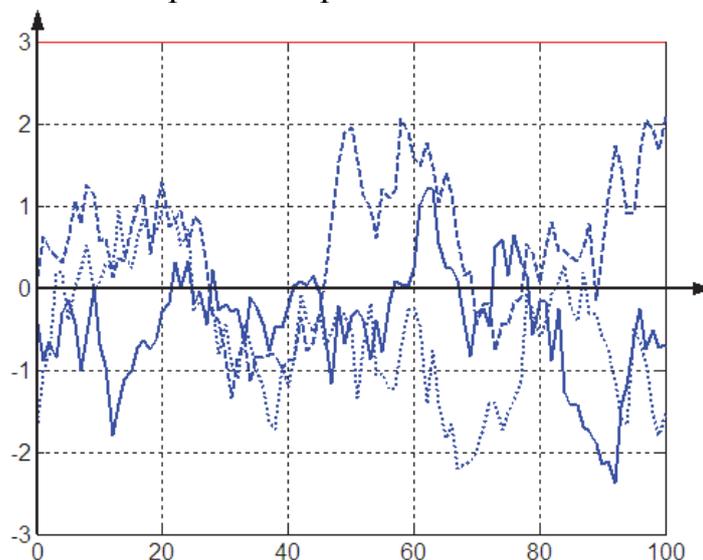


Рис. 6.1. Пример набора реализаций случайной последовательности

Пусть $x_i, i=1,2,\dots$ – случайная последовательность. Если зафиксировать, например, два произвольных момента времени t_i и t_j , то в результате можно сформировать двухмерный случайный вектор, компонентами кото-

рого являются две случайные величины x_i и x_j . Совместные статистические свойства этого вектора можно определить либо с помощью ф.р.в., либо с помощью ф.п.р.в. Аналогичные функции можно ввести и для вектора, включающего большее число значений последовательности. Ясно, что описание случайной последовательности будет полностью заданным, если будут определены функции плотности распределения вероятности $f(x_1, x_2, \dots, x_i)$ для любого конечного набора значений последовательности в произвольно выбранные моменты времени t_1, t_2, \dots, t_i .

При решении прикладных задач нередко ограничиваются заданием некоторых характеристик случайной последовательности, наиболее важными из которых являются **математическое ожидание и дисперсия случайной последовательности**. Располагая для произвольного i -го момента времени ф.п.р.в. $f(x_i)$, эти характеристики можно определить:

$$\bar{x}_i = M \{x_i\} = \int x_i f(x_i) dx_i ; \quad (6.1)$$

$$\sigma_i^2 = M \{(x_i - \bar{x}_i)^2\} = \int (x_i - \bar{x}_i)^2 f(x_i) dx_i . \quad (6.2)$$

Если считать известной совместную ф.п.р.в. $f(x_i, x_j)$ для значений последовательности в произвольные моменты времени, то можно ввести еще одну весьма важную характеристику

$$k(i, j) = M \{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\} = \iint (x_i - \bar{x}_i)(x_j - \bar{x}_j) f(x_i, x_j) dx_i dx_j , \quad (6.3)$$

называемую **корреляционной функцией случайной последовательности**.

Аналогичная характеристика для двух разных последовательностей называется **взаимно корреляционной функцией**.

Функция (6.3) задает значения коэффициента корреляции между случайными величинами, соответствующими i -му и j -му моментам времени. В силу соотношения (4.1)

$$f(x_i, x_j) = f(x_i / x_j) f(x_j) ,$$

где $f(x_i / x_j)$ – условная плотность для значений последовательности в i -й момент при условии, что значение в j -й момент зафиксировано.

Поскольку при $i = j$, $f(x_i / x_j) = \delta(x_i - x_j)$, т.е. переходная плотность представляет собой дельта-функцию, ясно, что

$$k(i, i) = \sigma_i^2 , \quad (6.4)$$

т.е. дисперсия последовательности в фиксированный момент времени совпадает с величиной корреляционной функции при совпадающих значениях аргумента $i = j$.

Случайные последовательности могут быть не только скалярными, но и векторными. Под **случайной n -мерной векторной последовательностью**

понимают последовательность, значения которой представляют собой n -мерный случайный вектор.

Пусть $x_i, i=1,2,..- n$ -мерная случайная последовательность. Заметим, что здесь и далее для указания дискретного момента времени, которому соответствует вектор, так же, как и для обозначения его компонент, используется нижний индекс.

Если требуется одновременное указание момента времени и номера компоненты, то используется двойная индексация, т. е. $x_i = (x_{i1}, \dots, x_{in})^T$, причем на первом месте всегда будет стоять индекс, соответствующий времени, а на втором – индекс, соответствующий номеру компоненты.

Все приведенные выше определения могут быть обобщены на случай векторной последовательности. В частности, корреляционная функция будет представлять собой **корреляционную матрицу**, определяемую как

$$k(i, j) = \int \int (x_i - \bar{x}_i)(x_j - \bar{x}_j)^T f(x_i, x_j) dx_i dx_j. \quad (6.5)$$

При одинаковых значениях аргумента $i = j$ корреляционная матрица совпадает с ковариационной матрицей, т.е.

$$k(i, i) = P_i = \int (x_i - \bar{x}_i)(x_i - \bar{x}_i)^T f(x_i) dx_i. \quad (6.6)$$

6.2. Стационарные случайные последовательности.

Дискретный белый шум

Рассмотрим некоторые наиболее распространенные типы случайных последовательностей. Зафиксируем моменты времени в порядке их возрастания и поставим им в соответствие значения последовательности x_1, x_2, \dots, x_k . Важным классом случайных последовательностей являются **стационарные случайные последовательности.**

Стационарной в широком смысле называется такая последовательность, у которой математическое ожидание от времени не зависит, а корреляционная функция зависит от разности $(i - j)$, т.е.

$$\bar{x}_i = \bar{x};$$

$$k(i - j) = M\{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\}.$$

Из последнего соотношения также следует независимость от времени дисперсии стационарной последовательности, поскольку

$$\sigma^2 = k(0).$$

Если перечисленные свойства не выполняются, то последовательность называется **нестационарной.**

Стационарной в узком смысле является такая последовательность, для которой плотность $f(x_1, \dots, x_k)$, введенная для любого конечного набора значений последовательности, при одновременном изменении всех момен-

тов времени на некоторую величину μ не изменяется, т.е. $f(x_1, \dots, x_k) = f(x_{1+\mu}, \dots, x_{k+\mu})$.

Последовательность, у которой математическое ожидание – нулевое ($\bar{x} = 0$), называется **центрированной**. Реализации, представленные на рис. 6.1, соответствуют центрированной последовательности. Пример реализации нецентрированной случайной последовательности приведен на рис. 6.2.

Понятно, что нецентрированная случайная последовательность, у которой математическое ожидание непостоянно, является нестационарной, так что реализация на рис. 6.2 представляет собой пример нестационарной последовательности.

Рассмотрим последовательность, значения которой в разные моменты времени между собой независимы:

$$f(x_1, x_2, \dots, x_k) = \prod_{j=1}^k f(x_j). \quad (6.7)$$

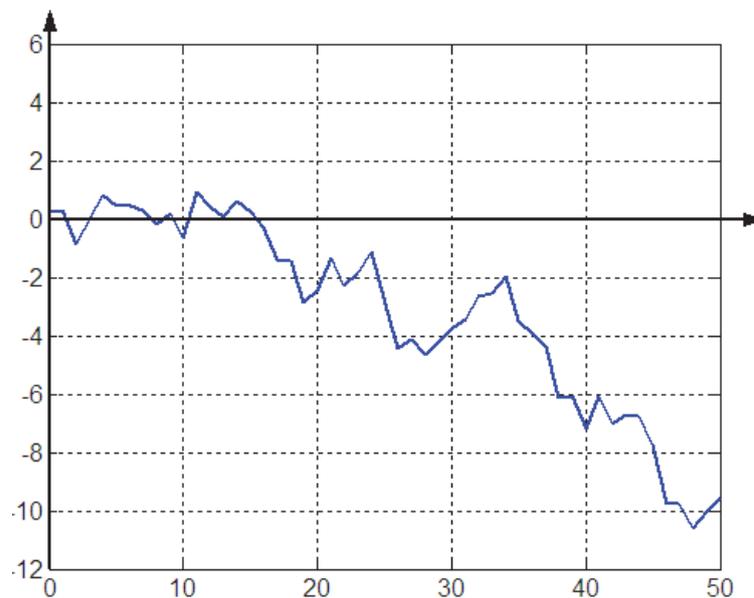


Рис. 6.2. Пример реализации нецентрированной случайной последовательности

Поскольку из условия независимости следует некоррелированность значений случайной последовательности, нетрудно понять, что корреляционная функция такой последовательности будет иметь следующий вид:

$$k(i, j) = \sigma_i^2 \delta_{ij}, \quad (6.8)$$

где δ_{ij} – символ Кронекера.

Последовательность, имеющая корреляционную функцию вида (6.8) называется **дискретным белым шумом**. Иными словами, дискретный белый шум – это такая последовательность, значения которой в различные моменты времени между собой не коррелированы. Если белый шум – центрированный, а дисперсии для всех моментов времени одинаковы, то шум будет стационарным. Пример реализации такого шума представлен на рис. 6.3.

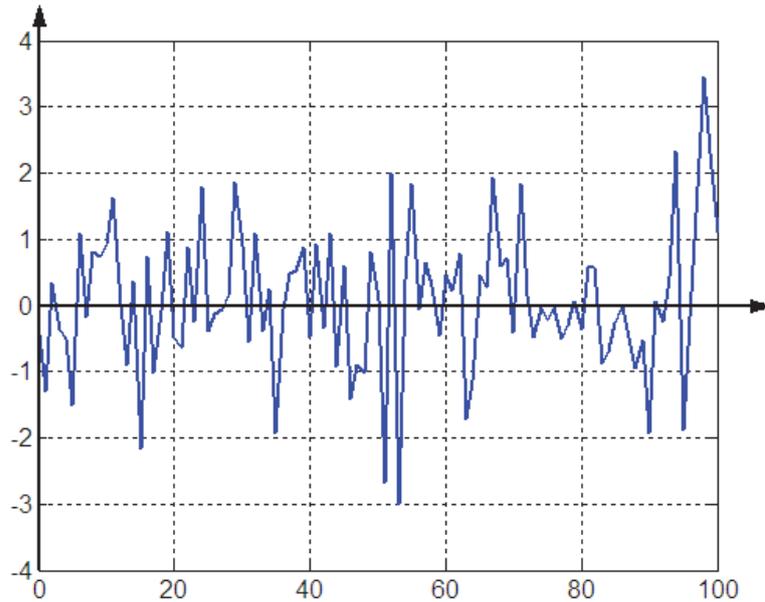


Рис. 6.3. Пример реализации стационарного белого шума

Заметим, что корреляционная функция характеризует временные свойства последовательности, определяя количественно уровень статистической зависимости ее значений в разные моменты времени. Вместе с тем существенным является вид функций плотностей распределения $f(x_1)$, $f(x_1, x_2)$, $f(x_1, x_2, x_3)$ и т.д. К примеру, можно ввести гауссовские последовательности.

Случайная последовательность называется **гауссовской**, если плотности распределения вероятности $f(x_1, \dots, x_k)$ для любого набора ее значений в произвольные моменты времени являются гауссовскими. В частности, если это требование выполнено для белого шума, то такой шум называется **гауссовским дискретным белым шумом**. Важно подчеркнуть, что для гауссовского белого шума для любого набора значений справедливо соотношение (6.7). Таким образом, наряду с (6.8), для любых $i \neq j$ можно записать

$$f(x_i, x_j) = f(x_i)f(x_j).$$

Гауссовские последовательности обладают весьма важным свойством, заключающимся в том, что, располагая математическим ожиданием (6.1) и корреляционной функцией (6.3), можно определить ф.п.р.в. $f(x_1, \dots, x_k)$ для любого набора значений последовательности в произвольные моменты времени. Иными словами, задание (6.1) и (6.3) обеспечивает полное описание его статистических свойств. Это есть следствие того факта, что гауссовская плотность полностью определяется двумя первыми моментами.

◆ **Пример 6.1.** Пусть для скалярной стационарной гауссовской последовательности заданы математическое ожидание \bar{x}_i и корреляционная

функция $k(i)$. Требуется записать плотность распределения $f(x_1, x_2, x_3)$ для вектора, составленного из значений последовательности $x = (x_1, x_2, x_3)^T$.

Поскольку вектор $x = (x_1, x_2, x_3)^T$ – гауссовский, то для отыскания соответствующей ему ф.п.р.в. достаточно найти математическое ожидание $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)^T$ и ковариационную матрицу

$$P^x = \begin{bmatrix} k(0) & k(t_1 - t_2) & k(t_1 - t_3) \\ k(t_2 - t_1) & k(0) & k(t_2 - t_3) \\ k(t_3 - t_1) & k(t_2 - t_3) & k(0) \end{bmatrix}.$$

Отсюда следует, что

$$f(x_1, x_2, x_3) = N(x_1, x_2, x_3; \bar{x}_1, \bar{x}_2, \bar{x}_3, P^x). \quad \blacklozenge$$

6.3. Марковские последовательности

Рассмотрим скалярную последовательность x_i . Зафиксируем моменты времени в порядке возрастания и сформируем вектор x_1, x_2, \dots, x_k . Для этого вектора может быть записана совместная плотность распределения вероятности $f(x_1, \dots, x_k)$, для которой (см. задачу 6.1), как следует из формулы умножения плотностей вероятности (4.1), в общем случае справедливо представление

$$f(x_1, \dots, x_k) = f(x_k / x_{k-1}, x_{k-2}, \dots, x_1) f(x_{k-1} / x_{k-2}, x_{k-3}, \dots, x_1) \dots f(x_1). \quad (6.9)$$

При исследовании временных свойств последовательностей существенным является вопрос о зависимости статистических свойств ее значений в текущий момент времени x_k от значений в предыдущие моменты времени x_1, x_2, \dots, x_{k-1} . Эта зависимость определяется условными плотностями вероятности вида $f(x_k / x_1, \dots, x_{k-1})$. Важным классом случайных последовательностей являются марковские последовательности. Отличительная особенность марковских последовательностей заключается в том, что статистические свойства ее значений в текущий момент определяются лишь значениями в ближайший предшествующий момент времени. С учетом сказанного можно дать следующее определение марковской последовательности.

Марковской последовательностью называется такая случайная последовательность, для которой свойства ее значений x_k в момент t_k при известном значении в ближайший момент t_{k-1} зависят только от значений последовательности в этот момент и не зависят от значений последовательности в моменты $t < t_{k-1}$, т.е.

$$f(x_k / x_{k-1}, x_{k-2}, \dots, x_1) = f(x_k / x_{k-1}). \quad (6.10)$$

Условная плотность вероятности марковской последовательности $f(x_k / x_{k-1})$ называется **плотностью вероятности перехода** или **переходной плотностью**.

Для марковской последовательности выражение (6.9) может быть преобразовано к виду

$$f(x_k, \dots, x_1) = f(x_1) \prod_{j=2}^k f(x_j / x_{j-1}).$$

Таким образом, задание плотности вероятности перехода и ф.п.р.в. для начального значения $f(x_1)$ достаточно для нахождения совместной функции плотности распределения вероятностей для набора значений марковской последовательности в произвольные моменты времени.

Рассмотренные выше типы случайных последовательностей легко обобщаются на векторный случай.

6.4. Формирующий фильтр

Важную роль при решении прикладных задач играют последовательности, задаваемые с помощью рекуррентных разностных уравнений вида

$$x_i = \Phi_i x_{i-1} + \Gamma_i w_i, i = 1, 2, \dots, \quad (6.11)$$

в которых x_i – n -мерный вектор; w_i – некоррелированный с x_0 , центрированный p -мерный дискретный белый шум с корреляционной функцией вида

$$M\{w_i w_j^T\} = \delta_{ij} Q_i, \quad (6.12)$$

где Q_i – $p \times p$ -матрица ковариаций; Φ_i , Γ_i – известные $n \times n$ и $n \times p$ -матрицы.

Вектор x_0 , задающий значение последовательности в начальный момент времени, также предполагается случайным с известным математическим ожиданием \bar{x}_0 и матрицей ковариаций P_0 .

Нетрудно заметить, что для такой последовательности справедливо соотношение (6.10), т.е. последовательность, определяемая (6.11), является марковской. Если значение последовательности в начальный момент времени x_0 и белый шум w_i являются гауссовскими, т.е.

$$f(x_0) = N(x_0; \bar{x}_0, P_0); \quad (6.13)$$

$$f(w_i) = N(w_i; 0, Q_i), \quad (6.14)$$

то последовательность x_i будет **гауссовской марковской последовательностью**.

Этот факт есть следствие того, что линейные преобразования гауссовских векторов порождают гауссовский вектор, а векторы x_0 и w_1, \dots, w_i являются совместно гауссовскими (см. задачу 6.2).

Уравнение (6.11) называется **формирующим фильтром (shaping filter) случайной последовательности**. Это название обусловлено тем, что случайная последовательность формируется в результате преобразования дискретного белого шума с помощью разностного уравнения (6.11). Входной

белый шум w_i , получил наименование **порождающего шума (generating или forcing noise)**. Матрица Φ_i называется **матрицей динамики**, матрица Γ_i – **матрицей порождающих шумов**, а вектор x_i – **вектором состояния**.

Принимая во внимание результаты задачи 6.3, нетрудно убедиться в том, что эволюция во времени математического ожидания и матрицы ковариаций для последовательности (6.11) определяется с помощью рекуррентных соотношений:

$$\bar{x}_i = \Phi_i \bar{x}_{i-1}; \quad (6.15)$$

$$P_i = M \left\{ (x_i - \bar{x}_i)(x_i - \bar{x}_i)^T \right\} = \Phi_i P_{i-1} \Phi_i^T + \Gamma_i Q_i \Gamma_i^T, \quad (6.16)$$

позволяющих вычислять искомые значения на текущем шаге с использованием известных значений на предыдущем шаге.

Из (6.15) с очевидностью следует, что последовательность, формируемая с помощью соотношения (6.11), при центрированном характере порождающих шумов будет центрированной, когда центрированным является вектор начальных условий.

При необходимости нетрудно, используя правила преобразования случайных векторов и соотношения (6.11), (6.12), получить матрицу ковариаций P_i^X для составного вектора $X_i = (x_0^T, x_1^T, \dots, x_i^T)^T$, со значениями последовательности во все предшествующие моменты времени, включая начальное значение (см. задачу 6.6).

Достоинство описания последовательностей с использованием формирующего фильтра (6.11) и соотношений (6.15), (6.16) заключается в том, что с их помощью рекуррентно может решаться как задача формирования самой последовательности, так и задача вычисления соответствующих ей расчетных статистических характеристик в виде математических ожиданий и матриц ковариаций. В частности, диагональные элементы матрицы ковариаций определяют расчетные дисперсии $\sigma_i^2(j) = P_i[j, j]$, $j = \overline{1, n}$ для каждой компоненты вектора состояния. С учетом сказанного формирующие фильтры (6.11) широко используются при решении прикладных задач для моделирования случайных последовательностей. С этой целью необходимо, с использованием датчика случайных чисел сформировать случайный вектор x_0 , векторы w_i и, привлекая рекуррентные соотношения (6.11), получать необходимые значения реализаций, формируемых в общем случае с использованием соотношения

$$z_i = H_i x_i. \quad (6.17)$$

Далее, если это особо не оговаривается, при рассмотрении примеров получения реализаций с помощью формирующих фильтров для определенности предполагается, что используются датчики гауссовских случайных величин.

◆ **Пример 6.2.** Частным случаем последовательности (6.11) является последовательность, в которой отсутствуют порождающие шумы, т.е. последовательность вида

$$x_i = \Phi_i x_{i-1}. \quad (6.18)$$

Для определения ее математического ожидания необходимо использовать (6.15), а матрица ковариаций в этом случае может быть найдена, если задаться матрицей ковариаций в начальный момент времени P_0 и выражением $P_i = \Phi_i P_{i-1} \Phi_i^T$.

Если матрица Φ_i – единичная, т.е. $\Phi_i = E$, то в результате получаем последовательность

$$x_i = x_{i-1}, \quad (6.19)$$

представляющую собой постоянный вектор с математическим ожиданием \bar{x}_0 и матрицей ковариаций P_0 . ◆

◆ **Пример 6.3.** Если в уравнении формирующего фильтра (6.11) $\Phi_i = E$ и имеются порождающие шумы, причем $\Gamma_i = E$, то тогда

$$x_i = x_{i-1} + w_i. \quad (6.20)$$

Ясно, что такая последовательность может быть представлена в виде суммы двух некоррелированных между собой слагаемых, одним из которых является постоянный случайный вектор, а другим – сумма некоррелированных между собой значений дискретного белого шума от первого до текущего моментов времени, т.е.

$$x_i = x_0 + \sum_{j=1}^i w_j. \quad (6.21)$$

Поскольку порождающий шум считается центрированным, то $\bar{x}_i = \bar{x}_0$, для любого i . Соотношение (6.16) в данном случае конкретизируется как

$$P_i = P_{i-1} + Q_i.$$

Если матрицы ковариаций для порождающего шума постоянные, т.е. $Q_i = Q$, то это позволяет записать

$$P_i = P_0 + iQ. \quad (6.22)$$

Из последних двух соотношений следует, что, несмотря на постоянство матриц, определяющих формирующий фильтр, получающаяся последовательность не является стационарной, поскольку соответствующая ей матрица ковариаций при увеличении времени также увеличивается.

Последовательность (6.20) называется **последовательностью с некоррелированными приращениями**. Это определение есть следствие того, что для нее приращения $(x_i - x_j)$ на непересекающихся интервалах времени между собой не коррелированы. В этом легко убедиться, поскольку на таких интервалах приращения определяются значениями белого шума, соответствующими различным моментам времени. Если все векторы – гаус-

совские, то такие приращения будут между собой независимыми, а последовательность (6.20) будет называться **последовательностью с независимыми приращениями**.

В скалярном случае последовательность (6.20) называется **винеровской последовательностью**. В англоязычной литературе для нее используется термин **random walk** (случайное блуждание). ♦

♦ **Пример 6.4.** В разделе 1 приводился пример описания изменения высоты за время проведения измерений в виде

$$h_i = x_0 + Vt_i, \quad i = \overline{1, m}, \quad (6.23)$$

где x_0, V – начальная высота и вертикальная скорость, полагаемая постоянной; $t_i = (i-1)\Delta t$ – моменты времени от начала наблюдения.

В разделе 1 такая модель изменения высоты сводилась к заданию двух неизвестных величин. Покажем, что эту модель можно получить и с использованием формирующего фильтра. Введем вектор состояния

$$x_i = (x_{i1}, x_{i2})^T = (h_i, V)^T. \quad (6.24)$$

Поскольку

$$\begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{i-1,1} \\ x_{i-1,2} \end{bmatrix},$$

где Δt – интервал проведения измерений, то при $\Gamma_i = 0$, $H_i = [1, 0]$,

$$\Phi_i = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (6.25)$$

получим

$$z_i = H_i x_i = [1, 0] x_i = h_i. \quad (6.26)$$

Если задаться матрицей ковариаций для начального момента времени в виде

$$P_0 = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_V^2 \end{bmatrix},$$

то, используя (6.16), нетрудно убедиться в том, что

$$\sigma_{h_i}^2 = P_i[1, 1] = \sigma_0^2 + \sigma_V^2 [(i-1)\Delta t]^2.$$

Если добавить порождающий шум с дисперсией q_1^2 в первое уравнение

$$\begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{i-1,1} \\ x_{i-1,2} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} w_i,$$

то вместо (6.23) получим следующую более общую модель для описания изменений высоты:

$$h_i = x_0 + Vt_i + \sum_{j=1}^i w_j.$$

Поскольку матрица динамики определяется выражением (6.25), а матрица при порождающих шумах $\Gamma_i = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, то в данном случае с помощью (6.16) получим

$$\sigma_{h_i}^2 = P_i[1,1] = \sigma_0^2 + \sigma_v^2 [(i-1)\Delta t]^2 + q^2(i-1).$$

На рис. 6.4 приведены примеры реализаций изменения высоты, полученные путем моделирования с помощью формирующего фильтра и датчика случайных чисел, для двух указанных случаев при следующих исходных данных: время 100 с, дискретность проведения измерений 1 с, $\sigma_0 = 10$ м, $\sigma_v = 0,1$ м/с при (а) $q_1 = 0$ и $q_1 = 1$ м (б). Здесь же приведены величины $\pm 3\sigma_{h_i}$, задающие расчетные утроенные среднеквадратические отклонения и в гауссовском случае определяющие область наиболее вероятных значений такой последовательности в каждый момент времени.

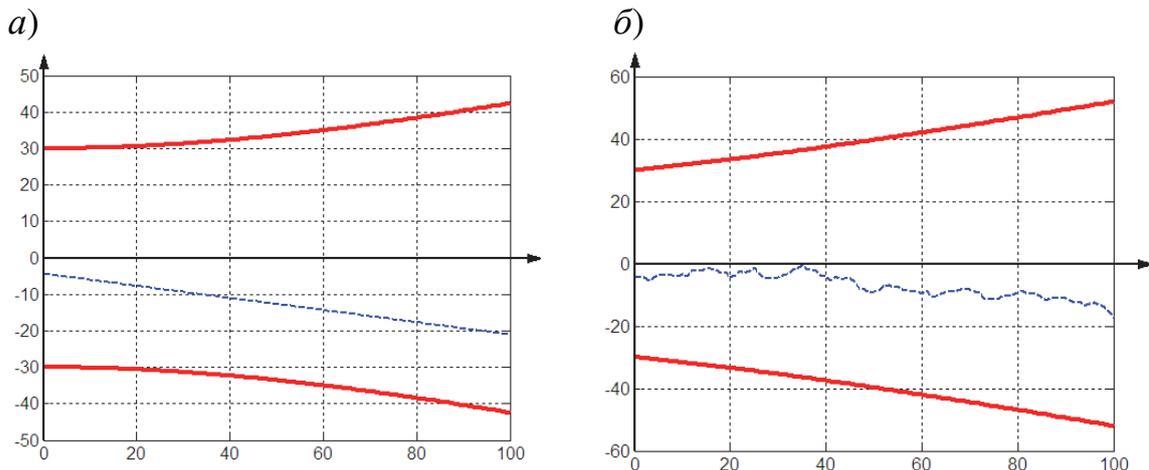


Рис. 6.4. Графики реализаций изменения высоты и соответствующих утроенных значений расчетных СКО при отсутствии и наличии порождающего шума

6.5. Динамика изменения матрицы ковариаций марковской последовательности

Из (6.16) следует, что в общем случае матрица ковариаций марковской последовательности, описываемой с помощью формирующего фильтра, во времени изменяется, причем P_i может как «возрастать», т.е. $P_i \geq P_{i-1}$, так и «убывать»: $P_i \leq P_{i-1}$.

Поясним это на скалярном примере, полагая

$$x_i = \Phi x_{i-1} + w_i \quad (6.27)$$

и считая $\Gamma = 1$ и $Q_i = q^2$ для всех i . В этом случае, вводя обозначение $P_i = \sigma_i^2$, выражение (6.16) запишем как

$$\sigma_i^2 = \sigma_{i-1}^2 \Phi^2 + q^2.$$

Если соотношение величин Φ^2 , q^2 и σ_0^2 таковы, что $\sigma_0^2 \Phi^2 + q^2 > \sigma_0^2$, т.е. $\sigma_0^2 > \frac{q^2}{1-\Phi^2}$, то дисперсия при увеличении i возрастает, т.е. $\sigma_i^2 > \sigma_{i-1}^2$. Если

же $\sigma_0^2 \Phi^2 + q^2 < \sigma_0^2$, $\left(\sigma_0^2 < \frac{q^2}{1-\Phi^2} \right)$, то дисперсия уменьшается. Когда суще-

ствует решение уравнения $\sigma_\infty^2 \Phi^2 + q^2 = \sigma_\infty^2$, дисперсия последовательности (6.27) с увеличением i перестает меняться и принимает значение $\sigma_\infty^2 = \frac{q^2}{1-\Phi^2}$. Это возможно при $\Phi^2 < 1$. Если принять $\sigma_0^2 = \frac{q^2}{1-\Phi^2}$, то дис-

персия будет постоянной для всех i . Ясно также, что при $\Phi > 1$ дисперсия может только возрастать, и поэтому не существует положительного решения уравнения $\sigma_0^2 \Phi^2 + q^2 = \sigma_0^2$; при $\Phi = 1$ дисперсия также будет возрастать.

В векторном случае, для того чтобы матрица ковариаций (6.16) при $\Phi_i = \Phi$, $\Gamma_i = \Gamma$, $Q_i = Q$ сохранялась постоянной, необходимо, чтобы существовало решение P_∞ для следующего уравнения:

$$P_\infty = \Phi P_\infty \Phi^T + \Gamma Q \Gamma^T. \quad (6.28)$$

В этой ситуации последовательность (6.11) при центрированном характере порождающего шума будет **стационарной марковской последовательностью**.

◆ **Пример 6.5.** Пусть задана винеровская последовательность (6.20), и при этом $P_0 = \sigma_0^2$, $Q = q^2$.

Приведем пример реализации этой последовательности и соответствующих ей в каждый момент времени расчетных значений СКО: $\sigma_i = \sqrt{P_i} = \sqrt{\sigma_0^2 + q^2 i}$, при разных соотношениях дисперсии начального значения и дисперсий порождающих шумов. Для определенности полагаем, что x_i описывает изменение вертикальной координаты.

Как отмечалось выше, значения такой последовательности в каждый момент времени представляют собой сумму двух слагаемых: случайной величины и суммы значений белозумной последовательности, накопленных к текущему моменту.

При $\sigma_0^2 = 0$ слагаемое в виде постоянной величины отсутствует. Такая реализация и соответствующие ей расчетные СКО в виде $\pm \sigma_i$ представлены на рис. 6.5, а, при $q = 1 \text{ м/с}$, $\sigma_0 = 0$. Здесь и в следующем примере для определенности считается, что значения последовательности формируются на интервале 100 с с дискретностью 1 с. При выполнении условия $q^2 i = \sigma_0^2$

($q = 1,0 \text{ м/с} = 1,0 \text{ м/с}$, $\sigma_0 = 10 \text{ м} = 10 \text{ м}$) для последнего момента вклад двух слагаемых в результирующее значение дисперсии одинаков (рис. 6.5, б).

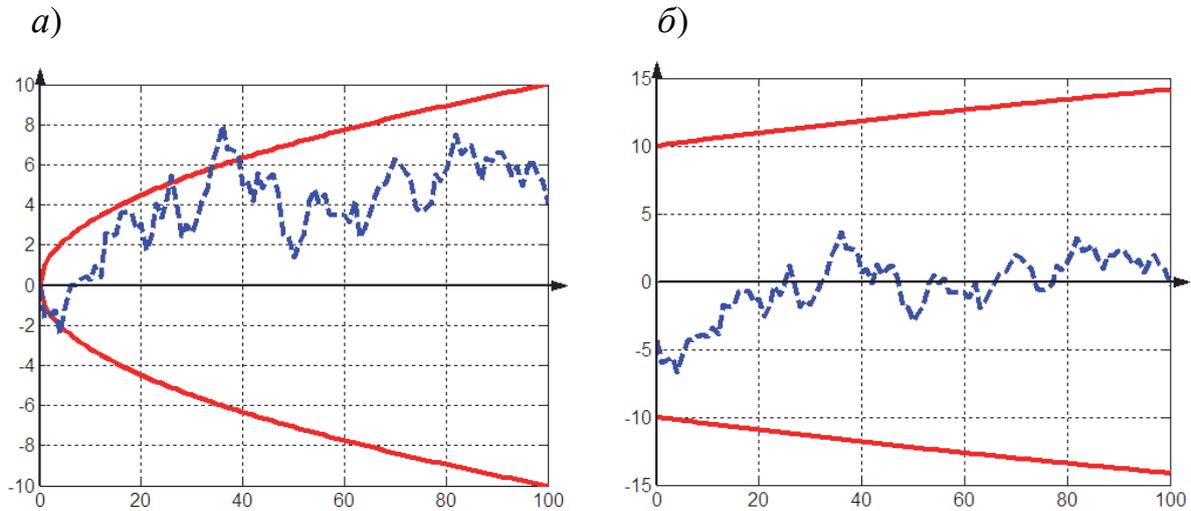


Рис. 6.5. Реализация винеровской последовательности и соответствующие ей расчетные значения СКО при нулевом и ненулевом начальных значениях

При выполнении условия $q^2 i \ll \sigma_0^2$ ($i = 100$, $q = 1,0 \text{ м/с} = 1,0 \text{ м/с}$, $\sigma_0 = 10 \text{ м} = 100 \text{ м}$) реализации будут мало отличаться от реализаций для постоянного вектора (рис. 6.6).

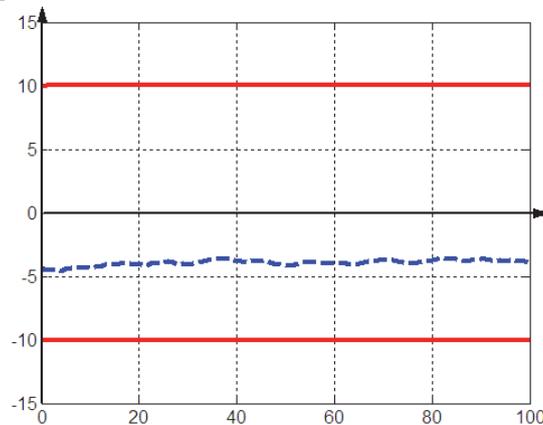


Рис. 6.6. Реализация винеровской последовательности и соответствующие ей расчетные значения СКО при выполнении условия $q^2 i \ll \sigma_0^2$

Для всех трех вариантов формируемые последовательности в рассмотренном примере являются нестационарными, хотя в последнем случае эта нестационарность в сравнительном плане проявляется весьма незначительно, т.е. СКО изменчивой части существенно меньше СКО постоянной составляющей. ◆

◆ **Пример 6.6.** Пусть требуется сформировать реализации центрированной случайной последовательности с корреляционной функцией

$k(i-j) = \sigma^2 e^{-\alpha|i-j|}$. Такая последовательность обычно называется **экспоненциально-коррелированной последовательностью**.

Нетрудно убедиться в том, что это может быть сделано с помощью формирующего фильтра вида (6.11), в котором $\Phi = e^{-\alpha}$, $\Gamma = \sigma\sqrt{1-e^{-2\alpha}}$; x_0 – центрированная случайная величина с дисперсией $P_0 = \sigma_0^2$, причем $\sigma_0 = \sigma$, а w_i – центрированный, дискретный белый шум с единичной дисперсией, независимый от x_0 [21, 25].

Это действительно так, поскольку при таких параметрах уравнение (6.28) преобразуется к тождеству: $\sigma^2 = \sigma^2 e^{-2\alpha} + \sigma^2(1 - e^{-2\alpha})$, означающему, что дисперсия этой последовательности постоянна. Кроме того, используя определение корреляционной функции, можем также записать следующее равенство:

$$k(i, i-1) = k(i - (i-1)) = k(1) = M\{x_i x_{i-1}\} = M\{\Phi x_{i-1} x_{i-1} + \Gamma w_i x_{i-1}\} = \sigma^2 e^{-\alpha},$$

которое нетрудно обобщить на случай $i - j > 1$.

Ниже приведены графики реализации такой последовательности при $\alpha = 0,1$, $\sigma = 1\text{ м}$, для трех случаев: стационарный случай, при выполнении условий $\sigma_0^2 \Phi^2 + \Gamma^2 = \sigma_0^2$, $\sigma = \sigma_0 = 1\text{ м} = 1$ м (рис. 6.7); нестационарный случай с возрастающей дисперсией при выполнении условий $\sigma_0^2 \Phi^2 + \Gamma^2 > \sigma_0^2$, $\sigma > \sigma_0 = 0$ (рис. 6.8, а), нестационарный случай с убывающей дисперсией при выполнении условий $\sigma_0^2 \Phi^2 + \Gamma^2 < \sigma_0^2$, $\sigma < \sigma_0 = 3\text{ м}$ (рис. 6.8, б). На графиках представлены как сами значения реализаций, так и соответствующие им утроенные расчетные значения СКО в виде $\pm 3(\sigma_i = \sqrt{P_i})$.

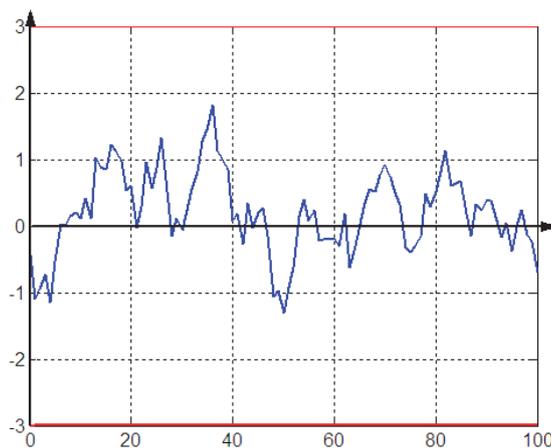


Рис. 6.7. Реализация экспоненциально коррелированной последовательности (стационарный случай)

Во всех случаях, как отмечалось выше, поскольку выполнено условие $\Phi^2 = e^{-2\alpha} < 1$, дисперсия выходит на установившийся режим $\sigma_\infty^2 = \frac{\sigma^2(1 - e^{-2\alpha})}{1 - e^{-2\alpha}} = \sigma^2 = 1$.

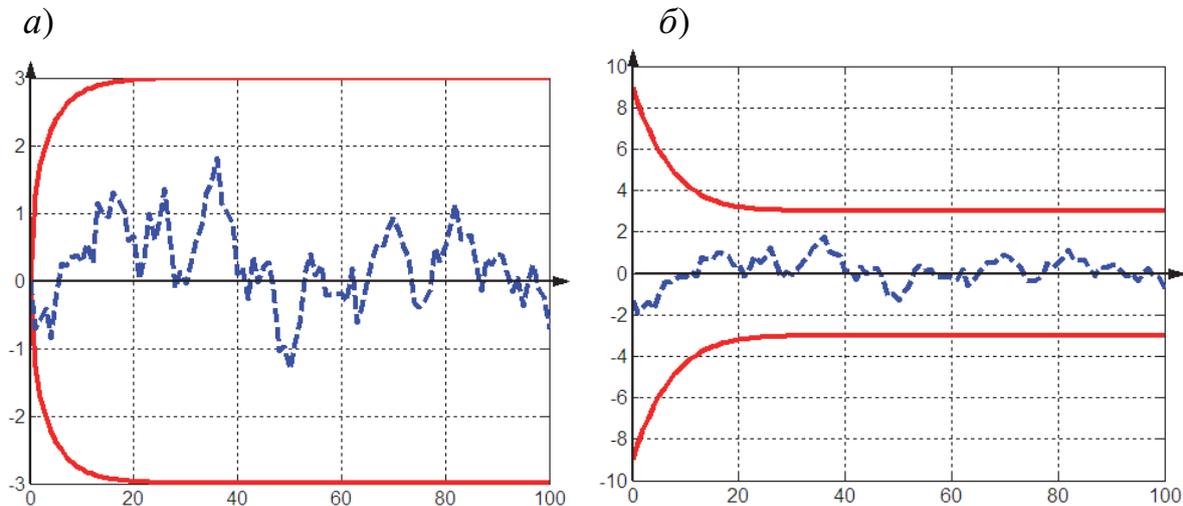


Рис. 6.8. Реализации экспоненциально коррелированной последовательности и соответствующие утроенные расчетные значения СКО (нестационарный случай)

Из представленных соотношений следует, что, если требуется сформировать последовательность с заданными σ^2 и α при единичной дисперсии дискретного шума, то для получения такой последовательности необходимо выполнение условия $\Gamma^2 = \sigma^2(1 - \Phi^2) = \sigma^2(1 - e^{-2\alpha})$. Это условие при $\alpha \ll 1$ может быть записано как $\Gamma^2 \approx 2\sigma^2\alpha$. ♦

Задачи к разделу

Задача 6.1. Задана совместная плотность распределения вероятности $f(x_1, \dots, x_k)$, покажите, что для нее справедливо следующее представление: $f(x_1, \dots, x_k) = f(x_k / x_{k-1}, x_{k-2}, \dots, x_1) f(x_{k-1} / x_{k-2}, x_{k-3}, \dots, x_1) \dots f(x_1)$.

Задача 6.2. Поясните, почему векторы x_0 и w_1, \dots, w_i , имеющие гауссовские ф.п.р.в. (6.13), (6.14), причем x_0 не коррелирован с w_l , а w_l, w_j не коррелированы между собой при $l \neq j$, $l, j = \overline{1, i}$, являются совместно гауссовскими.

Задача 6.3. Покажите, что эволюция во времени математического ожидания и матрицы ковариаций для последовательности (6.11) определяются соотношениями (6.15), (6.16):

$$\bar{x}_i = \Phi_i \bar{x}_{i-1};$$

$$P_i = M \left\{ (x_i - \bar{x}_i)(x_i - \bar{x}_i)^T \right\} = \Phi_i P_{i-1} \Phi_i^T + \Gamma_i Q_i \Gamma_i^T.$$

Задача 6.4. Рассчитайте матрицу ковариаций для случайного вектора, формируемого согласно выражению $x_i = \sum_{j=1}^i w_j$, в котором w_j , $j = \overline{1, i}$ – не коррелированные между собой центрированные случайные векторы с одинаковой матрицей ковариаций Q для каждого момента времени.

Задача 6.5. Рассчитайте матрицу ковариаций для случайного вектора, формируемого согласно выражению $x_i = \sum_{j=1}^i w_j$, в котором для всех $j = \overline{1, i}$, $w_j \equiv w$ – один и тот же центрированный случайный вектор с матрицей ковариаций Q .

Задача 6.6. Получите правило нахождения матрицы ковариаций P^{X_i} составного вектора $X_i = (x_0^T, x_1^T, \dots, x_i^T)^T$, компоненты которого описываются соотношениями (6.11), (6.12). Найдите эту матрицу для случая $i = 1$.

Задача 6.7. Получите выражение для ф.п.р.в. составного вектора $X_i = (x_0^T, x_1^T, \dots, x_i^T)^T$ для гауссовской марковской последовательности, задаваемой (6.11)–(6.14).

Задача 6.8. Пусть в предыдущей задаче $i = 1$, и $\Gamma_1 = E$, т.е. $x_1 = \Phi_1 x_0 + w_1$. Учитывая, что ф.п.р.в. для составного вектора $X_1 = (x_0^T, x_1^T)^T$ может быть представлена в виде (см. задачу 6.6)

$$f(X_1) = N(X_1; 0, P_1), \quad (1)$$

где

$$P_1 = \begin{bmatrix} P_0 & P_0 \Phi_1^T \\ \Phi_1 P_0 & \Phi_1 P_0 \Phi_1^T + Q_1 \end{bmatrix}, \quad (2)$$

убедитесь, что представление (1) для ф.п.р.в. идентично тому, которое получено в предыдущей задаче.

Пояснение. Раскройте выражение (1) и при вычислении P_1^{-1} воспользуйтесь правилом обращения блочных матриц (выражение (П1.47) Приложения 1).

Задача 6.9. Получите выражение для корреляционной функции центрированной винеровской последовательности, для которой начальное условие нулевое, а порождающий шум имеет одинаковую дисперсию q^2 .

Контрольные вопросы

1. Дайте определение случайной последовательности, поясните, что такое математическое ожидание, дисперсия и корреляционная функция для

случайной последовательности. Приведите примеры скалярных случайных последовательностей.

2. Поясните, что такое центрированная, стационарная и гауссовская случайные последовательности.
3. Дайте определение дискретного белого шума. Может ли дискретный белый шум быть нестационарным или негауссовским?
4. Введите понятие марковской последовательности. Поясните, что такое формирующий фильтр, приведите примеры.
5. Получите рекуррентные соотношения для нахождения математического ожидания и матрицы ковариаций марковской последовательности, заданной с помощью формирующего фильтра.
6. Что такое винеровская последовательность, каковы ее свойства? Является ли винеровская последовательность марковской?
7. Поясните, почему при постоянных матрицах Φ , Γ , Q в соотношениях (6.11), (6.12) и невыполнении равенства (6.28) последовательность не является стационарной?
Проиллюстрируйте это на примере.

РАЗДЕЛ 7

Фильтрация случайных последовательностей. Дискретный фильтр Калмана

В разделах 2–4 применительно к задаче оценивания вектора постоянных параметров были подробно рассмотрены возможные варианты построения алгоритмов в зависимости от уровня предполагаемой априорной информации об оцениваемом векторе и ошибках измерения. В этой главе мы рассмотрим задачу оценивания последовательностей, описывающих изменяющиеся во времени параметры, и при этом будем исходить из предположения о том, что они носят случайный характер. Такое предположение позволяет сформулировать задачу в рамках байесовского подхода как задачу нахождения оптимальных в среднеквадратическом смысле оценок. Предполагается, что случайная последовательность задается с помощью линейного формирующего фильтра и имеются измерения, линейным образом зависящие от оцениваемой последовательности. Описывается знаменитый рекуррентный дискретный фильтр Калмана, позволяющий элегантно решить обсуждаемую задачу [7, 18,21, 25, 27, 30-35].

7.1. Постановка и решение задачи нерекуррентного оптимального линейного оценивания случайных последовательностей

Несмотря на то, что наша цель заключается в получении рекуррентного алгоритма, для установления преемственности рассматриваемой в настоящем разделе задачи оценивания последовательности с задачей оценивания постоянного вектора рассмотрим сначала постановку и алгоритм решения задачи в нерекуррентной форме.

Рассмотрим следующую простейшую задачу оценивания случайных последовательностей. Предположим, что с помощью корреляционных и взаимно корреляционных функций заданы статистические свойства двух скалярных случайных последовательностей x_i и y_i , $i = 1, 2, \dots$:

$$k_x(v, \mu) = M \{ (x_v - \bar{x}_v)(x_\mu - \bar{x}_\mu) \}; \quad (7.1)$$

$$k_y(v, \mu) = M \{ (y_v - \bar{y}_v)(y_\mu - \bar{y}_\mu) \}; \quad (7.2)$$

$$k_{xy}(v, \mu) = M \{ (x_v - \bar{x}_v)(y_\mu - \bar{y}_\mu) \}, \quad v, \mu = 1, 2, \dots \quad (7.3)$$

и известны их математические ожидания \bar{x}_j и \bar{y}_j , $j = \overline{1, i}$.

Считаем также, что в результате проведения измерений зафиксированы все накопленные к текущему моменту i значения последовательности y_j , $j = \overline{1, i}$, т.е. известен вектор $Y_i = (y_1, \dots, y_i)^T$.

Требуется, располагая измеренными значениями последовательности y_i , полученными на текущий момент, найти оценку значений последовательности x_j для некоторого момента времени j .

Из представленной постановки следует, что необходимо оценить значения одной случайной последовательности по известным значениям другой, коррелированной с ней последовательности.

Различное соотношение между i и j в приведенной постановке соответствует различным задачам оценивания [18, 21]:

$j = i$ – **задача фильтрации (filtering problem)**, когда дискретный момент времени, для которого отыскивается оценка, совпадает с текущим моментом – моментом поступления очередного измерения;

$j < i$ – **задача интерполяции или сглаживания (smoothing problem)**, когда оценка отыскивается для более раннего момента времени по сравнению с текущим моментом;

$j > i$ – **задача прогноза или экстраполяции (forecasting problem)**, когда оценка отыскивается для будущего момента времени.

Специфика этих задач поясняется на рис. 7.1. Задачи фильтрации и прогноза обычно решают в **режиме реального времени (on-line mode)**, когда оценку требуется получить на текущий момент времени или в текущий момент времени предсказать значение оцениваемого параметра. Задачи же сглаживания решаются в так называемом **камеральном режиме обработки (off-line mode)**, когда оценки формируются после того, как получены все измерения

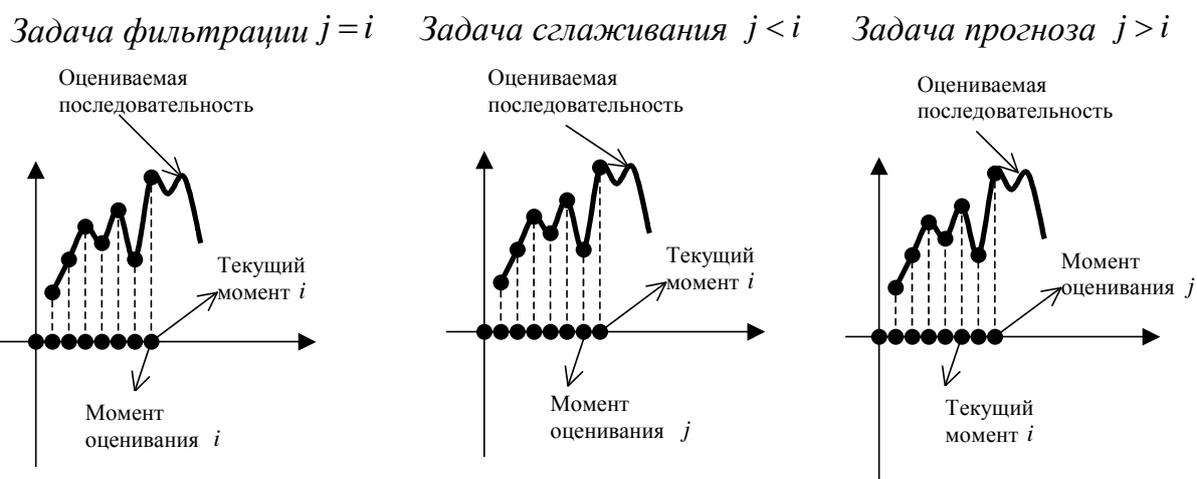


Рис. 7.1. Различные типы задач оценивания

В целях выявления взаимосвязи задач оценивания случайной последовательности с задачей оценивания постоянного вектора введем составной вектор $(x_j^T, Y_i^T)^T$ и рассмотрим сначала, к чему сводятся нерекуррентные алгоритмы оценивания, т.е. такие, в которых при вычислении оценок на текущем шаге используется весь набор измерений $Y_i = (y_1, \dots, y_i)^T$. Введем также обозначение $\hat{x}_{j|i}(Y_i)$, означающее, что оценка x_j на момент j формируется по измерениям, накопленным к текущему моменту i , т.е. полагаем, что в общем случае момент времени, для которого вычисляются оценки, может не совпадать с текущим моментом времени.

Конкретизация рассмотренных в разделах 2–4 постановок задачи оценивания постоянного вектора зависела от объема используемой априорной статистической информации. В рамках рассматриваемой постановки предполагаются известными два первых момента для вектора $(x_j^T, Y_i^T)^T$, что позволяет по аналогии с тем, как это было сделано в подразделе 4.5, сформулировать задачу нахождения несмещенных оценок $\hat{x}_{j|i}(Y_i)$, минимизирующих среднеквадратический критерий

$$J_{j|i}^{\bar{0}} = M \left\{ \left(x_j - \hat{x}_{j|i}(Y_i) \right)^2 \right\} \quad (7.4)$$

в классе оценок, зависящих от измерений линейным образом.

Такие оценки будем называть **оптимальными в среднеквадратическом смысле линейными оценками случайной последовательности или линейными несмещенными оценками с минимальной дисперсией.**

В сущности, это есть байесовская постановка задачи оценивания случайных последовательностей с ограничением на класс используемых оценок при минимизации выбранного критерия. Опираясь на результаты подраздела 4.5, в котором приведен алгоритм решения такой задачи, с учетом соотношений (4.24), (4.25) для искомой линейной оценки можно записать выражение

$$\hat{x}_j(Y_i) = \bar{x}_j + K_{j|i}(Y_i - \bar{Y}_i), \quad (7.5)$$

где $K_{j|i}$ – матрица-строка размерности $1 \times i$, удовлетворяющая **уравнению**

Винера–Хопфа для дискретного времени

$$K_{j|i} P^{Y_i} = P^{x_j Y_i}, \quad (7.6)$$

в котором

$$P^{x_j Y_i} = M \left\{ (x_j - \bar{x}_j)(y_\mu - \bar{y}_\mu) \right\}, \quad \mu = \bar{1..i} \quad (7.7)$$

– матрица-строка, определяющая корреляцию x_j со значениями y_1, \dots, y_i ;

$$P^{Y_i} = M \left\{ (y_\nu - \bar{y}_\nu)(y_\mu - \bar{y}_\mu) \right\}, \quad \nu, \mu = \bar{1..i} \quad (7.8)$$

– матрица ковариаций для всего набора измерений размерности $m \times m$.

Апостериорная дисперсия для ошибки оценки (7.5) задается соотношением

$$P_{j/i}^{lin} = P^{x_j} - K_{j/i} P^{Y_i x_i}, \quad (7.9)$$

в котором

$$P^{x_j} = M \left\{ (x_j - \bar{x}_j)^2 \right\} \quad (7.10)$$

– дисперсия оцениваемой последовательности.

Элементы входящих в приведенные выражения матриц $P^{x_j Y_i}$, P^{Y_i} , а также величина P^{x_j} легко отыскиваются с использованием (7.1) – (7.3).

Если матрица P^{Y_i} не вырождена, то можно записать явное выражение для оценок и соответствующих им дисперсий:

$$K_{j/i} = P^{x_j Y_i} (P^{Y_i})^{-1}; \quad (7.11)$$

$$P_{j/i}^{lin} = P^{x_j} - P^{x_j Y_i} (P^{Y_i})^{-1} P^{Y_i x_j}. \quad (7.12)$$

В подразделе 4.3 (соотношение (4.8)) отмечалось, что ошибка оценки, минимизирующей среднеквадратический критерий типа (7.4) в классе линейных оценок, ортогональна ко всему набору измерений или произвольной их комбинации. В данном случае это соотношение может быть записано в одной из следующих эквивалентных форм:

$$M \left\{ (x_j - \hat{x}_j(Y_i)) Y_i^T \right\} = 0, \quad (7.13)$$

т.е.

$$M \left\{ x_j Y_i^T \right\} = M \left\{ \hat{x}_j(Y_i) Y_i^T \right\}, \quad (7.14)$$

или

$$M \left\{ x_j \hat{x}_j^T(Y_i) \right\} = M \left\{ \hat{x}_j(Y_i) \hat{x}_j^T(Y_i) \right\}, \quad (7.15)$$

т.е.

$$M \left\{ x_j \hat{x}_j(Y_i) \right\} = M \left\{ \hat{x}_j(Y_i) \hat{x}_j(Y_i) \right\}. \quad (7.16)$$

Обращаем внимание на тот факт, что приведенное решение на основе уравнения Винера–Хопфа справедливо для любого типа оценок: фильтрации, сглаживания и прогноза.

В полной мере приведенные результаты будут справедливыми и в случае, если последовательности являются векторными, и при этом в качестве критерия вместо (7.4) используется критерий

$$J_{j/i}^{\delta} = M \left\{ (x_j - \hat{x}_{j/i}(Y_i))^T (x_j - \hat{x}_{j/i}(Y_i)) \right\}.$$

Полагая, что x_i и y_i ($i=1, 2..$) представляют собой n - и m -мерные последовательности с математическими ожиданиями \bar{x}_j, \bar{y}_j ($j = \overline{1, i}$) для линейных оптимальных оценок, минимизирующих такой критерий, будут справедливы приведенные выше выражения. При этом матрицы $P^{x_j Y_i}$, P^{Y_i} и P^{x_j} бу-

дуг иметь размерности $n \times [i \times m]$, $n \times [i \times m]$, $[i \times m] \times [i \times m]$ и $n \times n$ соответственно.

Нетрудно заметить, что после того, как введен вектор $(x_j^T, Y_i^T)^T$, задачи оценивания случайной последовательности для каждого момента времени в представленной постановке ничем не отличаются от задачи оценивания постоянного вектора, рассмотренной в разделах 2–4. Особенность заключается лишь в том, что необходимые для их решения матрицы ковариаций формируются с использованием корреляционных и взаимно корреляционных функций (7.1)–(7.3).

◆ **Пример 7.1.** Предположим, что x_i представляет собой стационарную центрированную случайную последовательность с корреляционной функцией вида $k(i-j) = \sigma^2 e^{-\alpha|i-j|}$ и на интервале $\mu = \overline{1.i}$ проводятся измерения этой последовательности, которые могут быть представлены как

$$y_i = x_i + v_i, \quad (7.17)$$

где v_i – независимая от x_i последовательность, представляющая собой центрированный дискретный белый шум с дисперсией r^2 . Требуется найти оценку последовательности в некоторый момент времени j .

Применительно к рассматриваемому примеру имеем:

$$P^{x_j} = \sigma^2; \quad (7.18)$$

$$P^{Y_i} = \{M(x_\nu x_\mu) + M(v_\nu v_\mu)\} = \{\sigma^2 e^{-\alpha|\nu-\mu|} + r^2 \delta_{\nu\mu}\}, \quad \nu, \mu = \overline{1.i}, \quad (7.19)$$

$$P^{x_j Y_i} = \{M(x_j x_\mu)\} = \{\sigma^2 e^{-\alpha|j-\mu|}\}, \quad \mu = \overline{1.i}. \quad (7.20)$$

Принимая во внимание эти выражение, можно записать следующее соотношение для матрицы $K_{j/i}$:

$$K_{j/i} = P^{x_j Y_i} (P^{Y_i})^{-1} = \sigma^2 \begin{pmatrix} e^{-\alpha|j-1|} & e^{-\alpha|j-2|} & \dots & e^{-\alpha|j-i|} \end{pmatrix} \times$$

$$\times \begin{bmatrix} \sigma^2 + r^2 & \sigma^2 e^{-\alpha|1|} & & & \sigma^2 e^{-\alpha|1-i|} \\ \sigma^2 e^{-\alpha|1|} & \sigma^2 + r^2 & & & \\ & & & & \\ & & & \sigma^2 + r^2 & \sigma^2 e^{-\alpha|i|} \\ \sigma^2 e^{-\alpha|i-1|} & & & \sigma^2 e^{-\alpha|i|} & \sigma^2 + r^2 \end{bmatrix}^{-1}.$$

Используя (7.5), (7.9) и последнее соотношение, можно конкретизировать выражения для интересующей нас оценки и соответствующей ей апостериорной дисперсии ошибки. ◆

Важно подчеркнуть, что в рассмотренной постановке не предполагается задание функциональной зависимости между измеряемой и оцениваемой

последовательностями, а исходная информация задается в виде математических ожиданий, корреляционных и взаимно корреляционных функций. В сущности, это есть обобщение задачи линейной регрессии для случайных последовательностей, решаемой при наличии информации лишь об их первых и вторых моментах. Причем полученное решение соответствует так называемой *нерекуррентной* процедуре, при которой для нахождения оценок (7.5) каждый раз привлекается весь набор имеющихся измерений. Хотя приведенные выражения и позволяют определять оценки и соответствующие им характеристики точности, использование их при решении прикладных задач затруднительно. В частности, это связано с тем, что при увеличении числа измерений размерность матрицы P^Y все время возрастает, что осложняет процедуру ее обращения. Далее основное внимание уделяется задачам рекуррентной фильтрации, при решении которых удается построить удобные для реализации алгоритмы.

7.2. Постановка линейной задачи рекуррентной оптимальной фильтрации случайных последовательностей

Сформулируем задачу построения **рекуррентных алгоритмов нахождения оптимальных в среднеквадратическом смысле линейных оценок случайных последовательностей**, т.е. задачу построения **линейных рекуррентных оптимальных алгоритмов**. Применительно к задаче оценивания постоянного вектора рекуррентные алгоритмы рассматривались в подразделе 5.5. Их суть заключается в том, что искомая оценка для заданного набора измерений i формировалась путем последовательной обработки очередного i -го измерения и оценки, полученной на предыдущем шаге обработки по набору измерений $Y_{i-1} = (y_1, \dots, y_{i-1})^T$. Соответствующая текущей оценке матрица ковариаций также вычислялась с использованием матрицы ковариаций для предыдущего шага. Применительно к рассматриваемой задаче рекуррентные алгоритмы могут быть получены в случае, когда решается задача фильтрации, т.е. $i = j$. При этом предполагается, что случайная последовательность задана с помощью линейного формирующего фильтра вида (6.11), а измерения линейным образом зависят от оцениваемой последовательности.

Постановка такой задачи в общем случае может быть сформулирована следующим образом.

Задана n -мерная случайная последовательность в виде формирующего фильтра

$$x_i = \Phi_i x_{i-1} + \Gamma_i w_i, \quad (7.21)$$

и имеются m -мерные измерения, связанные с этой последовательностью соотношением вида

$$y_i = H_i x_i + v_i, \quad (7.22)$$

где w_i – p -мерный вектор порождающих шумов; v_i – m -мерный вектор ошибок измерения; Φ_i , H_i , Γ_i – известные матрицы размерности $n \times n$, $m \times n$, $n \times p$ соответственно. При этом w_i и v_i представляют собой дискретные, центрированные белые шумы:

$$M\{w_i w_j^T\} = \delta_{ij} Q_i; \quad M\{v_i v_j^T\} = \delta_{ij} R_i. \quad (7.23)$$

Вектор начальных условий x_0 считается центрированным вектором с матрицей ковариаций P_0 , а векторы x_0 , w_i , v_i – не коррелированными между собой, т.е.

$$M\{w_i v_j^T\} = 0, \quad M\{x_0 v_i^T\} = 0, \quad M\{x_0 w_i^T\} = 0. \quad (7.24)$$

Центрированный характер случайных векторов предполагается лишь для упрощения получаемых выражений.

Требуется, располагая накопленными к текущему моменту времени i измерениями $Y_i = (y_1^T, y_2^T, \dots, y_i^T)^T$, найти рекуррентный алгоритм вычисления оптимальных в среднеквадратическом смысле несмещенных линейных оценок последовательности (7.21), минимизирующих критерий

$$J_i^0 = M\{(x_i - \hat{x}_{i/i}(Y_i))^T (x_i - \hat{x}_{i/i}(Y_i))\}, \quad (7.25)$$

и рекуррентный алгоритм вычисления матриц ковариаций ошибок оценивания

$$\varepsilon_{i/i}(Y_i) = x_i - \hat{x}_{i/i}(Y_i), \quad (7.26)$$

определяемых как

$$P_i = M\{(x_i - \hat{x}_{i/i}(Y_i))(x_i - \hat{x}_{i/i}(Y_i))^T\}. \quad (7.27)$$

В соответствии с терминологией, введенной в подразделе 7.1, можно также говорить о **задаче фильтрации вектора состояния**, описываемого с помощью соотношения (7.21), по измерениям (7.22) или о **задаче фильтрации, сформулированной на языке пространства состояний**.

Далее для оценки $\hat{x}_{i/i}(Y_i)$ и ее ошибки будем использовать обозначения $\hat{x}_{i/i}(Y_i) \equiv \hat{x}_i(Y_i) \equiv \hat{x}_i$, $\varepsilon_{i/i}(Y_i) \equiv \varepsilon_i(Y_i) = \varepsilon_i$.

Нетрудно заметить, что рассмотренная в разделах 2-4 задача оценивания постоянного вектора представляет собой частный случай сформулированной здесь задачи фильтрации последовательности. Действительно, эта задача получается, если в уравнениях (7.21) принять $w_i \equiv 0$ и $\Gamma_i \equiv 0$, а $\Phi_i \equiv E$.

Можно выделить следующие отличительные особенности приведенной постановки по сравнению с той, которая рассматривалась ранее:

- решается задача оценивания марковской последовательности;
- формирующий фильтр для марковской последовательности – линейный;

- введены ошибки измерения и определена функциональная зависимость измерений от оцениваемой последовательности и ошибок измерений, причем эта зависимость также линейная.

Эти особенности создают предпосылки успешного решения задачи получения удобных в вычислительном отношении рекуррентных линейных алгоритмов в виде дискретного фильтра Калмана, соотношения для которого и рассмотрим далее.

Важно подчеркнуть, что в представленной постановке, как и в предыдущем подразделе, не вводится информация о законах распределения случайных последовательностей и решение задачи предполагается находить в классе линейных оценок.

7.3. Дискретный фильтр Калмана для случайных последовательностей

При построении оптимального рекуррентного алгоритма сформулированной выше задачи предполагается, что оценка на текущем шаге \hat{x}_i вычисляется с использованием очередного измерения y_i , оценки \hat{x}_{i-1} и соответствующей ей матрицы ковариаций ошибок P_{i-1} для предыдущего шага (рис. 7.2).

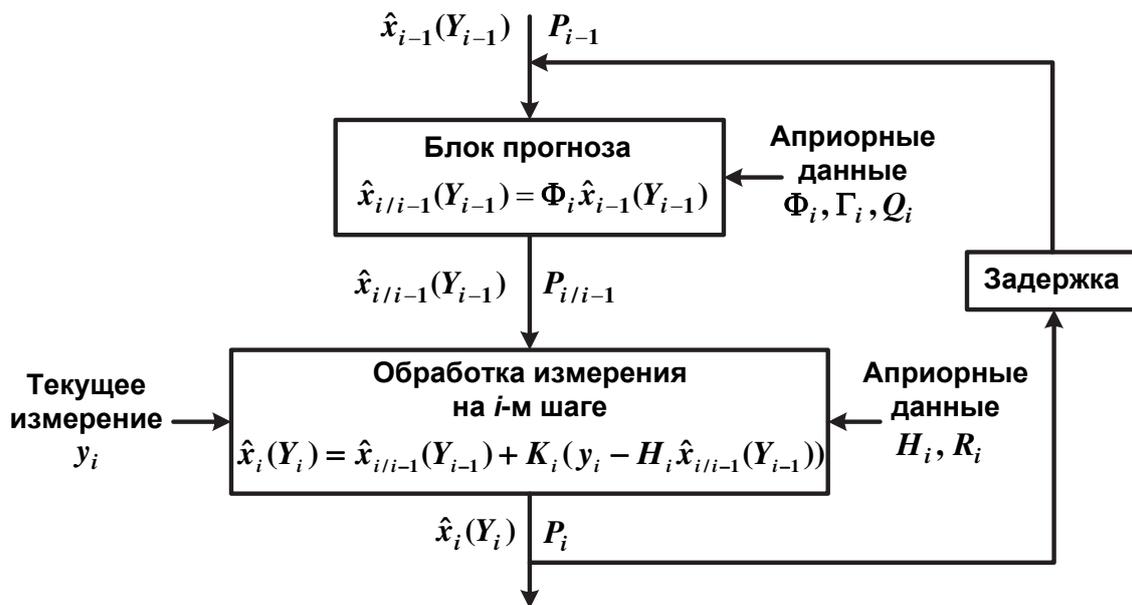


Рис. 7.2. Блок-схема фильтра Калмана для дискретного времени

В подразделе 5.5 был получен рекуррентный алгоритм применительно к задаче оценивания постоянного вектора. Там подчеркивалось, что на каждом шаге фактически происходит уточнение оценки, полученной с использованием предшествующих измерений, путем добавления корректирующего слагаемого, представляющего собой произведение коэффициента усиления на разность между текущим измерением и его прогнозируемым значением. При получении рекуррентного алгоритма для случайной последо-

вательности, задаваемой с помощью формирующего фильтра (7.21), логично предположить, что его структура сохранится прежней. Особенность применительно к изменяющейся последовательности должна проявиться в том, каким образом находить оценку прогноза значений последовательности и прогноза значений текущего измерения на очередном шаге обработки, располагая значением оптимальной оценки \hat{x}_{i-1} , соответствующей набору измерений Y_{i-1} . Ясно, что это следует делать с учетом уравнения динамики (7.21) для оцениваемой последовательности.

Принимая во внимание сказанное выше, можно показать, что решение задачи получения рекуррентного алгоритма вычисления оптимальных в среднеквадратическом смысле линейных несмещенных оценок, минимизирующих среднеквадратический критерий, задается в виде набора соотношений, в котором выделяют два блока (см. рис. 7.2) [6, 18, 21, 25, 31-35].

В первом блоке (блоке прогноза) вычисляется линейная оптимальная оценка значений вектора состояний в i -й момент по измерениям Y_{i-1} , т.е. **оценка прогноза** на один шаг вперед $\hat{x}_{i/i-1}$ (**state estimate extrapolation** [32]) и соответствующая ей **матрица ковариаций ошибок прогноза** (**error covariance extrapolation** [32]), которые отыскиваются с использованием оптимальной оценки \hat{x}_{i-1} и матрицы ковариаций P_{i-1} на предыдущем шаге:

$$\hat{x}_{i/i-1} = \Phi_i \hat{x}_{i-1}; \quad (7.28)$$

$$P_{i/i-1} = \Phi_i P_{i-1} \Phi_i^T + \Gamma_i Q_i \Gamma_i^T, \quad (7.29)$$

где

$$P_{i/i-1} = M \left\{ (x_i - \hat{x}_{i/i-1}(Y_{i-1}))(x_i - \hat{x}_{i/i-1}(Y_{i-1}))^T \right\} = M \left\{ (\varepsilon_{i/i-1})(\varepsilon_{i/i-1})^T \right\};$$

$$\varepsilon_{i/i-1} = x_i - \hat{x}_{i/i-1}(Y_{i-1}).$$

Во втором блоке осуществляется вычисление искомой текущей оценки \hat{x}_i и расчетной матрицы ковариаций ее ошибок P_i с использованием текущего измерения y_i и результатов, полученных в блоке прогноза:

$$\hat{x}_i = \hat{x}_{i/i-1} + K_i (y_i - H_i \hat{x}_{i/i-1}); \quad (7.30)$$

$$K_i = P_{i/i-1} H_i^T (H_i P_{i/i-1} H_i^T + R_i)^{-1}; \quad (7.31)$$

$$P_i = P_{i/i-1} - P_{i/i-1} H_i^T (H_i P_{i/i-1} H_i^T + R_i)^{-1} H_i P_{i/i-1} = (E - K_i H_i) P_{i/i-1}. \quad (7.32)$$

Эти соотношения и определяют знаменитый **фильтр Калмана (ФК) для дискретного времени**. Рекуррентный характер этого алгоритма очевиден, поскольку при нахождении прогноза и соответствующей ему матрицы ковариаций, помимо априорной информации в виде матриц Φ_i , Γ_i и Q_i , используется только оценка и матрица ковариаций ее ошибок, полученные на предыдущем шаге, а при нахождении очередной оценки и соответствующей ей матрицы ковариаций – лишь результаты, полученные в блоке прогноза, само текущее измерение и априорная информация в виде матриц H_i

и R_i . Несмотря на использование только очередного измерения, получаемая в результате оценка является оптимальной (минимизируют критерий (7.25)) по всему набору измерений, и в блок-схеме сохранена зависимость оценок от соответствующих наборов измерений.

В англоязычной литературе для соотношений (7.30) обычно используются термины **state estimate observational update** или просто **state estimate update**, а для (7.32) – **error covariance update** [32]. Матрица K_i называется **матрицей коэффициентов усиления ФК (Kalman gain matrix)** или просто **коэффициентом усиления ФК**.

Как и в подразделе 4.4, можно показать, что для матриц K_i , P_i будут также справедливы следующие, удобные в некоторых практических приложениях представления:

$$K_i = P_i H_i^T R_i^{-1}; \quad (7.33)$$

$$P_i = (P_{i|i-1}^{-1} + H_i^T R_i^{-1} H_i)^{-1}. \quad (7.34)$$

Вычисление коэффициента усиления и матрицы ковариаций в соответствии с выражениями (7.31), (7.32) предпочтительнее в тех случаях, когда $m \ll n$, в то время как при $n \ll m$ и диагональной матрице R_i удобнее использовать выражения (7.33), (7.34).

Следует обратить внимание на тот факт, что матрицы ковариаций ошибок оценок P_i не зависят от измерений, а определяются только матрицами ковариаций Q_i и R_i , характеризующими свойства порождающих шумов и ошибок измерения, и матрицами наблюдения H_i .

Таким образом, все вычисления, связанные с нахождением матриц ковариаций и коэффициента усиления, могут быть в принципе выполнены заранее. Иногда блок, реализующий эти вычисления, называют **ковариационным каналом**, а блок, реализующий вычисления оценок (7.30) – **оценочным каналом**. Весьма существенно, что алгоритм ФК является линейным относительно измерений, а коэффициент усиления также зависит лишь от матриц Q_i , R_i и H_i и, следовательно, не зависит от измерений.

Впервые соотношения для ФК были приведены и доказаны в знаменитой работе Калмана [35]. Доказательство проведено на основе использования свойства ортогональности.

После выхода в свет этой публикации появились различные варианты доказательства оптимальности оценок, получаемых с помощью ФК.

Важно еще раз обратить внимание на тот факт, что здесь рассматривается задача получения оценок, оптимальных в классе линейных, при этом каких-либо предположений о законах распределения случайных последовательностей не вводится. Если предположить гауссовский характер распределений, то оказывается, что приведенный выше алгоритм будет обеспечивать получение оптимальных байесовских оценок, т.е. оценок, которые

минимизируют среднеквадратический критерий типа (7.4) без введения ограничений на класс используемых оценок.

Т а б л и ц а 7.1

Соотношения для алгоритма дискретного фильтра Калмана

Постановка задача фильтрации	
Уравнение для вектора состояния	$x_i = \Phi_i x_{i-1} + \Gamma_i w_i$
Измерения	$y_i = H_i x_i + v_i$
Начальные условия	$\bar{x}_0 = 0, P_0$
Порождающие шумы	$\bar{w}_i = 0, M\{w_i w_j^T\} = \delta_{ij} Q_i;$
Шумы измерения	$\bar{v}_i = 0, M\{v_i v_j^T\} = \delta_{ij} R_i;$
Взаимная корреляция	$M\{x_0 w_i^T\} = 0; M\{w_i v_i^T\} = 0; M\{x_0 v_i^T\} = 0$
Матрицы	$\Phi_i, - n \times n, \Gamma_i - n \times p, Q_i - p \times p$ $H_i - m \times n, R_i - m \times m$
Минимизируемый критерий $J_i^{\circ} = M_{x_i, Y_i} \left\{ (x_i - \hat{x}_i(Y_i))^T (x_i - \hat{x}_i(Y_i)) \right\}$	
Решение задачи фильтрации	
Прогноз	$\hat{x}_{i/i-1} = \Phi_i \hat{x}_{i-1}$
Матрица ковариаций ошибок прогноза	$P_{i/i-1} = \Phi_i P_{i-1} \Phi_i^T + \Gamma_i Q_i \Gamma_i^T$
Оценка	$\hat{x}_i = \hat{x}_{i/i-1} + K_i (y_i - H_i \hat{x}_{i/i-1})$
Коэффициент усиления	$K_i = P_{i/i-1} H_i^T (H_i P_{i/i-1} H_i^T + R_i)^{-1}$ вариант 1 $K_i = P_i H_i^T R_i^{-1}$ вариант 2
Матрица ковариаций ошибок оценивания	$P_i = (E - K_i H_i) P_{i/i-1}$ вариант 1 $P_i = (P_{i/i-1}^{-1} + H_i^T R_i^{-1} H_i)^{-1}$ вариант 2

Подчеркнем, что алгоритм ФК определяет не только удобную процедуру вычисления самих оценок, что обеспечивает решение **задачи синтеза алгоритма оценивания случайной последовательности**, но и процедуру

вычисления расчетной матрицы ковариаций, характеризующей текущую точность алгоритма оценивания, что важно при решении **задачи анализа точности оценивания случайной последовательности**.

В частности, диагональные элементы определяют расчетные дисперсии ошибок оценивания, которые, в свою очередь определяют расчетные СКО ошибок оценивания для всех компонент вектора состояния.

Сформулированная задача и представленные соотношения сведены в табл. 7.1.

Поскольку с помощью ФК в рамках рассмотренной постановки отыскиваются оценки, оптимальные в классе линейных, можно говорить о том, что с помощью ФК рассчитывается **потенциальная точность оценивания случайной последовательности (7.21) по измерениям (7.22) с использованием линейных оценок**.

Конкретизируем приведенные соотношения алгоритма ФК на простейших примерах.

◆ **Пример 7.2.** Получим выражения для фильтра Калмана в задаче оценивания скалярного параметра $x_i = x_{i-1} = x$ по скалярным измерениям $y_i = x + v_i$, в которых $x \equiv x_0$ – центрированная случайная величина с дисперсией σ_0^2 ; v_i – не зависящий от x центрированный дискретный белый шум, т.е. $M\{v_i v_j\} = \delta_{ij} r_i^2$.

В данном случае $\Phi = H = 1$, $\Gamma = Q \equiv 0$. Поскольку оцениваемый параметр не меняется, блок прогноза существенно упрощается:

$$\begin{aligned}\hat{x}_{i|i-1} &= \hat{x}_{i-1}; \\ P_{i|i-1} &= P_{i-1}.\end{aligned}$$

Таким образом, выражения для оценки примут вид

$$\hat{x}_i = \hat{x}_{i-1} + K_i (y_i - \hat{x}_{i-1}).$$

Для вычисления коэффициента усиления и дисперсии ошибки оптимальной оценки здесь удобно использовать выражения (7.33), (7.34):

$$\begin{aligned}K_i &= \frac{P_{i-1}}{P_{i-1} + r_i^2}; \\ P_i &= (P_{i-1}^{-1} + R_i^{-1})^{-1} = \frac{P_{i-1} r_i^2}{P_{i-1} + r_i^2}, P_0 = \sigma_0^2.\end{aligned}$$

Если предположить, что дисперсии всех измерений одинаковы, т.е. $r_i^2 = r^2$, то легко убедиться в том, что

$$P_i = \frac{P_0 r^2}{i P_0 + r^2}.$$

Отсюда следует, что в условиях, когда априорная дисперсия существенно больше дисперсии измерений, т.е. при $P_0 \gg r^2$, выражения для коэф-

коэффициента усиления и апостериорной дисперсии могут быть определены как $K_i = \frac{1}{i}$, $P_i = \frac{r^2}{i}$, и, таким образом, $\hat{x}_i = \hat{x}_{i-1} + \frac{1}{i}(y_i - \hat{x}_{i-1})$.

Как и следовало ожидать, последнее выражение в этом случае представляет собой рекуррентную формулу вычисления среднего арифметического

$$\hat{x}_i = \frac{1}{i} \sum_{j=1}^i y_j.$$

На рис. 7.3 приведены реализации ошибок фильтрации и соответствующие утроенные расчетные значения ее СКО в виде $\pm 3(\sigma_i = \sqrt{P_i})$ для случая $P_0 = \sigma_0^2 = 1$, $r^2 = 1$.

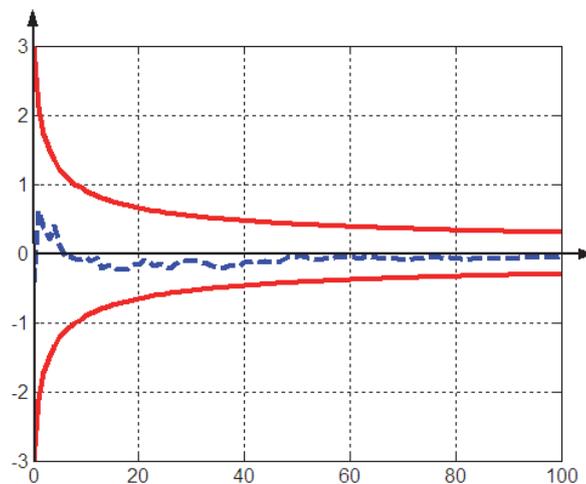


Рис. 7.3. Реализации ошибок фильтрации и соответствующие утроенные значения СКО при оценивании постоянной величины

Как следует из графика, величина СКО ошибки оценивания неограниченно уменьшается, в данном случае на рассматриваемом интервале времени, достигая уровня примерно 0,1. ♦

♦ **Пример 7.3.** Пусть требуется решить задачу фильтрации винеровской последовательности $x_i = x_{i-1} + w_i$ по измерениям $y_i = x_i + v_i$, в которых x_0 — центрированная случайная величина с дисперсией σ_0^2 ; w_i , v_i — не зависящие друг от друга и от x_0 центрированные дискретные белые шумы такие, что $M\{v_i v_j\} = \delta_{ij} r^2$, $M\{w_i w_j\} = \delta_{ij} q^2$ [21].

В данном случае $\Phi = H = 1$, $\Gamma = 1$. Несмотря на то, что оцениваемый параметр здесь изменяется, выражения для оценок будут иметь тот же вид, что и в предыдущем примере, т.е.

$$\hat{x}_{i/i-1} = \hat{x}_{i-1}; \quad \hat{x}_i = \hat{x}_{i-1} + K_i(y_i - \hat{x}_{i-1}).$$

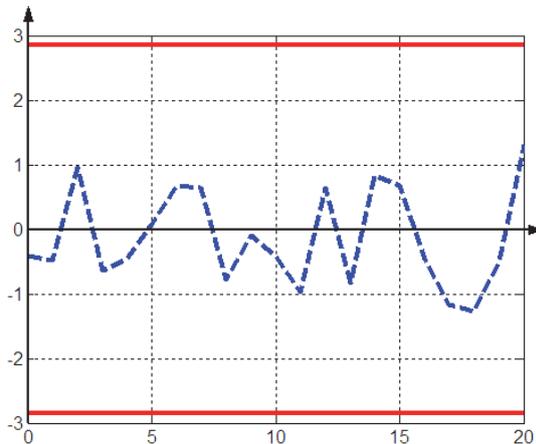
Для дисперсий ошибок прогноза и дисперсий ошибок оценок будем иметь следующие выражения:

$$P_{i/i-1} = P_{i-1} + q^2;$$

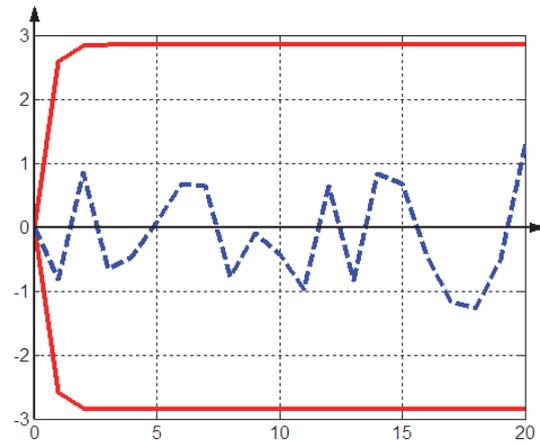
$$P_i = P_{i/i-1} \left(1 - \frac{P_{i/i-1}}{P_{i/i-1} + r^2} \right) \text{ или } P_i = (P_{i-1} + q^2) \left(\frac{r^2}{P_{i-1} + q^2 + r^2} \right).$$

На рис. 7.4 представлены результаты решения задачи фильтрации винеровской последовательности на интервале 20 с при $q = 1$, $r^2 = 4$ в виде реализации ошибок фильтрации и соответствующих утроенных значений СКО ($\pm 3(\sigma_i = \sqrt{P_i})$) при разных значениях $P_0 = \sigma_0^2$.

а) $P_0 = \sigma_0^2 = P_\infty = 1$



б) $P_\infty > P_0 = \sigma_0^2 = 0$



в) $(P_0 = 3^2) > P_\infty = 1$

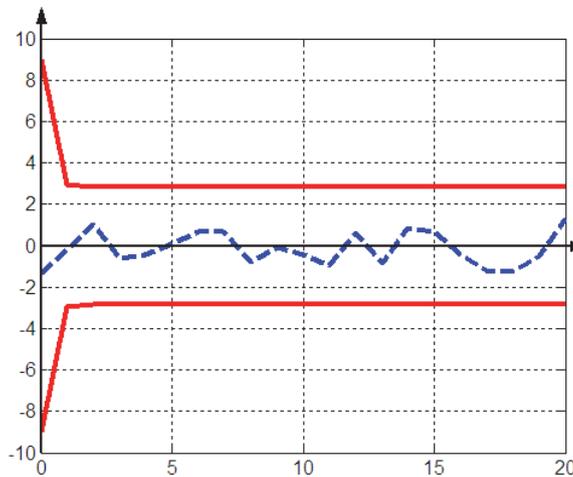


Рис. 7.4. Реализация ошибок фильтрации винеровской последовательности и соответствующие утроенные значения СКО:

а – стационарный случай, дисперсия ошибки постоянна; б – дисперсия ошибки увеличивается и выходит на установившееся значение; в – дисперсия ошибки уменьшается и выходит на установившееся значение ◆

◆ **Пример 7.4.** Получим соотношения ФК для рассмотренной в примере 7.1 задачи фильтрации скалярной стационарной экспоненциально-

коррелированной последовательности, имеющей корреляционную функцию $k(i-j) = \sigma^2 e^{-\alpha|i-j|}$ [6].

С учетом примера 6.6 уравнения (7.21), (7.22) в данном случае конкретизируются следующим образом: $x_i = \Phi x_{i-1} + w_i$; $y_i = x_i + v_i$, где $\Phi = e^{-\alpha}$, x_0 – центрированная случайная величина с дисперсией $P_0 = \sigma^2$, а w_i , v_i – не зависящие от x_0 центрированные дискретные белые шумы с дисперсиями $q^2 = \sigma^2(1 - e^{-2\alpha})$ и r^2 соответственно.

Уравнения для блока прогноза здесь могут быть записаны как

$$\hat{x}_{i/i-1} = \Phi \hat{x}_{i-1}; \quad P_{i/i-1} = \Phi^2 P_{i-1} + q^2,$$

а уравнения для оценок и дисперсий их ошибок – в виде

$$\hat{x}_i = \Phi \hat{x}_{i-1} + K_i (y_i - \Phi \hat{x}_{i-1})$$

$$P_i = \left(\frac{1}{P_{i/i-1}} + \frac{1}{r^2} \right)^{-1} = \left(\frac{1}{\Phi^2 P_{i-1} + q^2} + \frac{1}{r^2} \right)^{-1}, \text{ где } K_i = \frac{P_i}{r^2}.$$

На рис. 7.5 представлен пример реализаций ошибок фильтрации экспоненциально-коррелированной последовательности и соответствующих им утроенных расчетных значений СКО в виде $\pm 3(\sigma_i = \sqrt{P_i})$, при $\alpha = 0,1$, $\sigma = 1$ м, $q^2 = 1(1 - e^{-0,2}) = 0,2$ м² и двух значениях дисперсии ошибок измерения: $r^2 = 1$ м² (а) и $r^2 = (0,1)^2$ м² (б)

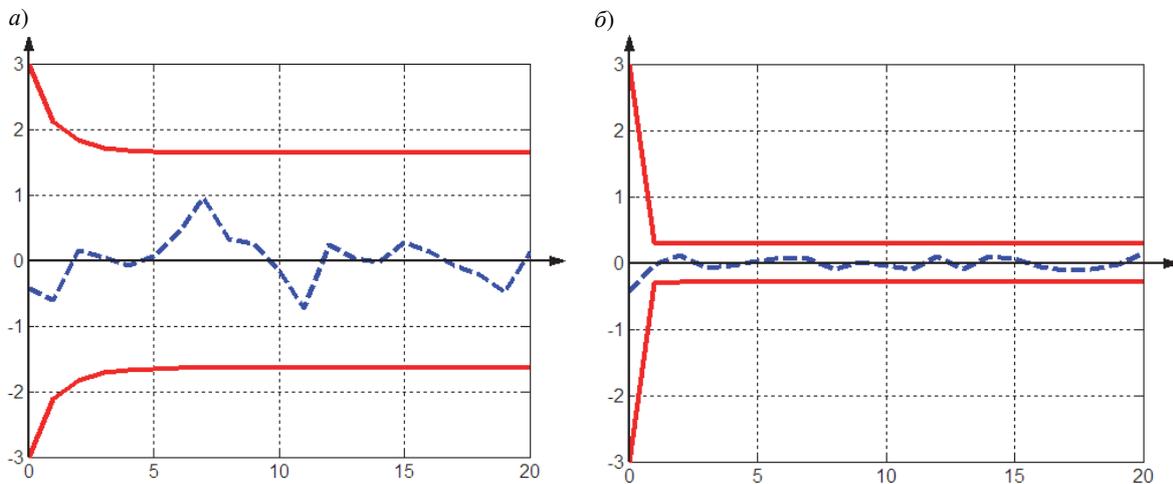


Рис. 7.5. Реализации ошибок фильтрации экспоненциально-коррелированной последовательности и соответствующих им утроенных расчетных значений СКО при разных уровнях ошибок измерения ♦

Как видно из графиков, дисперсии достигают некоторых установившихся значений, величина которых существенным образом зависит от уровня ошибок измерения и порождающего шума. Причины и возможность наличия установившегося значения обсуждаются далее.

7.4. Уравнения ошибок фильтра Калмана.

Инновационная последовательность

Вычитая из уравнения формирующего фильтра (7.21) значения оценок (7.28), (7.30), нетрудно получить уравнения для **ошибок прогноза** и **ошибок фильтрации**, вырабатываемых ФК:

$$\varepsilon_{i|i-1} = \Phi_i x_{i-1} + \Gamma_i w_i - \Phi_i \hat{x}_{i-1} = \Phi_i \varepsilon_{i-1} + \Gamma_i w_i; \quad (7.35)$$

$$\begin{aligned} \varepsilon_i &= \Phi_i x_{i-1} + \Gamma_i w_i - \Phi_i \hat{x}_{i-1} - K_i (H_i x_i - H_i \hat{x}_{i|i-1}) + K_i v_i = \\ &= \Phi_i \varepsilon_{i-1} + \Gamma_i w_i - K_i H_i (\Phi_i \varepsilon_{i-1} + \Gamma_i w_i) + K_i v_i = \\ &= (E - K_i H_i) \Phi_i \varepsilon_{i-1} + (E - K_i H_i) \Gamma_i w_i + K_i v_i. \end{aligned} \quad (7.36)$$

Из последнего уравнения, в частности, следует, что ошибка линейной оптимальной оценки является марковской последовательностью и для нее справедливо следующее рекуррентное соотношение:

$$\varepsilon_i = (E - K_i H_i) \Phi_i \varepsilon_{i-1} + (\Gamma_i - K_i H_i) w_i + K_i v_i, \quad (7.37)$$

в котором в правой части присутствуют как порождающие шумы, так и шумы измерения.

В том, что для матриц ковариаций ошибок (7.35), (7.36) будут справедливы соотношения (7.29), (7.32) или, что то же самое, (7.34), нетрудно убедиться непосредственно (см. задачу 7.1), используя выражение для коэффициента усиления (7.31).

Ясно также, что для оценок, отыскиваемых с помощью ФК, будет выполняться условие ортогональности (7.15), т.е.

$$M \{ (x_i - \hat{x}_i(Y_i)) \hat{x}_j^T(Y_j) \} = 0. \quad (7.38)$$

Любопытно отметить, что (см. задачу 7.3)

$$M \{ \varepsilon_i(Y_i) x_i^T \} = M_{x_i/Y_i} \{ (x_i - \hat{x}_i^T(Y_i)) x_i^T \} = P_i. \quad (7.39)$$

Важную роль при анализе свойств оценок ФК играет **невязка измерения**, которая в рассматриваемом случае определяется как

$$\mu_i = y_i - H_i \hat{x}_{i|i-1}(Y_{i-1}). \quad (7.40)$$

Эта последовательность обладает весьма важным свойством, заключающимся в том, что она представляет собой дискретный центрированный белый шум, т.е.

$$M \{ \mu_i \mu_j^T \} = \delta_{ij} L_i, \quad (7.41)$$

где

$$L_i = H_i P_{i|i-1} H_i^T + R_i. \quad (7.42)$$

Тот факт, что матрица ковариаций невязки определяется соотношением (7.42), доказывается весьма просто (см. задачу 7.2). Убедимся в справедливости (7.41). Поскольку $M_{y_i/Y_{i-1}} \{ y_i \} = H_i \hat{x}_{i|i-1}(Y_{i-1})$, то при $j < i$

$$M_{Y_i} \left\{ (y_i - H_i \hat{x}_{i/i-1}(Y_{i-1})) y_j^T \right\} = M_{Y_{i-1}} M_{y_i/Y_{i-1}} \left\{ \mu_i y_j^T \right\} = 0,$$

т.е.

$$M \left\{ \mu_i y_j^T \right\} = 0, \text{ для } j < i. \quad (7.43)$$

Аналогично $M \left\{ \mu_i \mu_j^T \right\} = 0$, для $j < i$. Если $j > i$, то такое равенство легко получить, убеждаясь в том, что $M \left\{ \mu_i y_j^T \right\} = 0$. Таким образом, последнее равенство справедливо для любых несовпадающих j и i , а следовательно, справедливо и (7.41). Из доказанного утверждения, в сущности, вытекает, что случайная последовательность y_i с помощью соотношения (7.40) преобразуется в белый шум. В связи с этим такая операция называется иногда **выбеливанием**.

Из (7.40), (7.43) следует, что невязка, соответствующая текущему моменту времени, ортогональна набору предыдущих измерений $Y_{i-1} = (y_1, \dots, y_{i-1})^T$, а текущее измерение может быть представлено в виде суммы

$$y_i = H_i \hat{x}_{i/i-1}(Y_{i-1}) + \mu_i.$$

Заметим, что выражение для оценки, вырабатываемой в фильтре Калмана, может быть записано как сумма двух слагаемых:

$$\hat{x}_i = \hat{x}_{i/i-1} + K_i \mu_i(Y_{i-1}), \quad (7.44)$$

одно из которых представляет собой линейную комбинацию набора предыдущих измерений $Y_{i-1} = (y_1, \dots, y_{i-1})^T$, а второе – линейную комбинацию компонент вектора невязки $\mu_i(Y_{i-1})$, ортогонального к Y_{i-1} . Так как в этом векторе содержится новая информация, которой нет в Y_{i-1} , именно за счет использования μ_i и происходит обновление информации, что и объясняет используемый для невязки (7.40) термин – **инновационная**, или **обновляющаяся последовательность (innovations sequences [32])**.

◆ **Пример 7.5.** Конкретизируем приведенные уравнения для ошибок оценивания применительно к задаче фильтрации постоянной величины и винеровской последовательности.

Для постоянной величины:

$$\begin{aligned} \varepsilon_{i/i-1} &= \varepsilon_{i-1}; \\ \varepsilon_i &= (1 - K_i) \varepsilon_{i-1} + K_i v_i; \\ \mu_i &= y_i - \hat{x}_{i/i-1}(Y_{i-1}) = \varepsilon_{i-1} + v_i; \\ L_i &= P_{i-1} + r^2. \end{aligned}$$

Для винеровской последовательности:

$$\varepsilon_{i/i-1} = \varepsilon_{i-1} + w_i;$$

$$\begin{aligned}\varepsilon_i &= (1 - K_i)\varepsilon_{i-1} + (1 - K_i)w_i + K_i v_i; \\ \mu_i &= y_i - \hat{x}_{i/i-1}(Y_{i-1}) = \varepsilon_{i/i-1} + v_i = \varepsilon_{i-1} + w_i + v_i; \\ L_i &= P_{i-1} + q^2 + r^2.\end{aligned}$$

Любопытно отметить, что невязка измерения, формируемая как разность между измеренными значениями и текущей оценкой

$$\tilde{\mu}_i = y_i - \hat{x}_i(Y_i) = \varepsilon_i + v_i = (1 - K_i)\varepsilon_{i-1} + (1 - K_i)w_i + (1 + K_i)v_i,$$

в отличие от невязки, формируемой как разность между измеренными значениями и оценкой прогноза, уже не является белым шумом. ♦

7.5. Динамика изменения матрицы ковариаций и установившийся режим в задаче фильтрации

Немаловажным при исследовании поведения ошибок ФК является вопрос о характере изменения соответствующих матриц ковариаций во времени. Выше при рассмотрении скалярных примеров уже отмечалось, что дисперсия ошибки фильтрации может достигать некоторого установившегося значения. Обсудим этот вопрос более подробно.

Анализ уравнения (7.32) для матрицы ковариаций показывает, что в ней имеются два слагаемых. Первое слагаемое

$$P_{i/i-1} = \Phi_i P_{i-1} \Phi_i^T + \Gamma_i Q_i \Gamma_i^T \quad (7.45)$$

отражает факт изменения матрицы ковариации ошибок оценивания при вычислении оценки прогноза, обусловленный, в частности, наличием порождающих шумов. Чтобы проанализировать характер этого изменения, введем матрицу $\Delta P_i^{(1)} = P_{i/i-1} - P_{i-1}$ и представим ее в виде $P_{i/i-1} = P_{i-1} + \Delta P_i^{(1)}$. Из результатов подраздела (6.5) и уравнения (7.45) следует, что матрица ковариаций ошибок прогноза $P_{i/i-1}$ может как «возрастать» по сравнению с матрицей ковариаций ошибок P_{i-1} на предыдущем шаге, так и «убывать». Убывание в принципе возможно, если матрица ковариаций (дисперсии) самой оцениваемой последовательности уменьшается. Однако при решении прикладных задач, как правило, происходит возрастание матрицы ковариаций ошибок оцениваемой последовательности. Это, в частности, означает, что дисперсии ошибок оценок каждой компоненты, определяемые диагональными элементами, могут лишь увеличиваться или, в крайнем случае, сохраняться на прежнем уровне. Такое поведение обусловлено тем, что значение вектора состояния изменяется в соответствии с уравнением формирующего фильтра (6.11), в правой части которого присутствует порождающий шум, увеличивающий неопределенность в знании вектора состояния на каждом шаге. Даже в случае, если значение вектора состояния в какой-то момент времени было известно точно, например $P_{i-1} = 0$, при прогнозе соответствующая матрица ковариаций $P_{i/i-1} = \Gamma_i Q_i \Gamma_i^T$ возрастает

и становится отличной от нуля, поскольку $\Gamma_i Q_i \Gamma_i^T \geq 0$ при $Q > 0$. Более подробно это обсуждается далее в примере 7.6.

Второе слагаемое в выражении для матрицы ковариаций, отражающее влияние очередного (текущего) измерения, может быть представлено в виде

$$\Delta P_i^{(2)} = P_{i|i-1} - P_i = P_{i|i-1} H_i^T (H_i P_{i|i-1} H_i^T + R_i)^{-1} H_i P_{i|i-1}.$$

Так как справа стоит неотрицательно определенная матрица, то $P_{i|i-1} - P_i \geq 0$ и, следовательно, $P_i \leq P_{i|i-1}$, т.е. матрица ковариаций ошибок оценок при обработке очередного измерения может лишь уменьшиться или сохраниться на прежнем уровне по сравнению с матрицей ковариаций ошибок прогноза. Это соотношение представляется вполне логичным, поскольку использование новой измерительной информации не должно снижать точность оценивания.

Таким образом, для матрицы ковариаций можно записать

$$P_i = P_{i-1} + \Delta P_i^{(1)} - \Delta P_i^{(2)}.$$

Ясно, что при равенстве между собой слагаемых $\Delta P_i^{(1)}$ и $\Delta P_i^{(2)}$ матрица ковариаций ошибок фильтрации не будет зависеть от времени, т.е. будет постоянной. В этом случае говорят об **установившемся режиме задачи фильтрации**. В частности, такой режим может иметь место при решении задач фильтрации, в которых все используемые в соотношениях фильтра Калмана матрицы (Φ , Γ , H , Q , R) являются постоянными. На существование установившегося решения уже обращалось внимание, когда приводились результаты моделирования задач фильтрации винеровской и экспоненциально-коррелированной последовательностей по их измеренным на фоне белого шума значениям в примерах 7.3, 7.4. Очевидно, что для того чтобы имел место установившийся режим, необходимо существование решения следующего уравнения:

$$\begin{aligned} P_\infty &= \Phi P_\infty^\phi \Phi^T + \Gamma Q \Gamma^T - (\Phi P_\infty^\phi \Phi^T + \Gamma Q \Gamma^T) H^T \times \\ &\times \left(H (\Phi P_\infty^\phi \Phi^T + \Gamma Q \Gamma^T) H^T + R \right)^{-1} H (\Phi P_\infty^\phi \Phi^T + \Gamma Q \Gamma^T). \end{aligned} \quad (7.46)$$

Здесь P_∞^ϕ определяет матрицу ковариаций ошибок фильтрации для установившегося режима. При выполнении условия (7.46) матрица ковариаций ошибок прогноза и коэффициент усиления также сохраняются постоянными:

$$\begin{aligned} P_\infty^{np} &\equiv \Phi P_\infty \Phi^T + \Gamma Q \Gamma^T; \\ K_\infty &= P_\infty^\phi H^T R^{-1}. \end{aligned}$$

Таким образом, уравнение фильтра Калмана превращается в стационарное разностное уравнение

$$\hat{x}_i = \Phi \hat{x}_{i-1} + K_\infty (y_i - H \Phi \hat{x}_{i-1}) = (E - K_\infty H) \Phi \hat{x}_{i-1} + K_\infty y_i,$$

которое соответствует так называемому **фильтру Винера для дискретного времени**.

◆ **Пример 7.6.** Проанализируем возможную динамику изменения дисперсии ошибок и получим фильтр Винера применительно к задаче, рассмотренной в примере 7.3, т.е. к задаче оценивания винеровской последовательности по ее измеренным значениям на фоне белого шума.

Поскольку $\Phi = 1$, то $P_{i/i-1} = P_{i-1} + q^2$, и таким образом на каждом шаге при прогнозе дисперсия ошибки увеличивается на величину, совпадающую с дисперсией порождающего шума q^2 , т.е. $\Delta P_i^{(1)} = q^2$ (рис. 7.6).

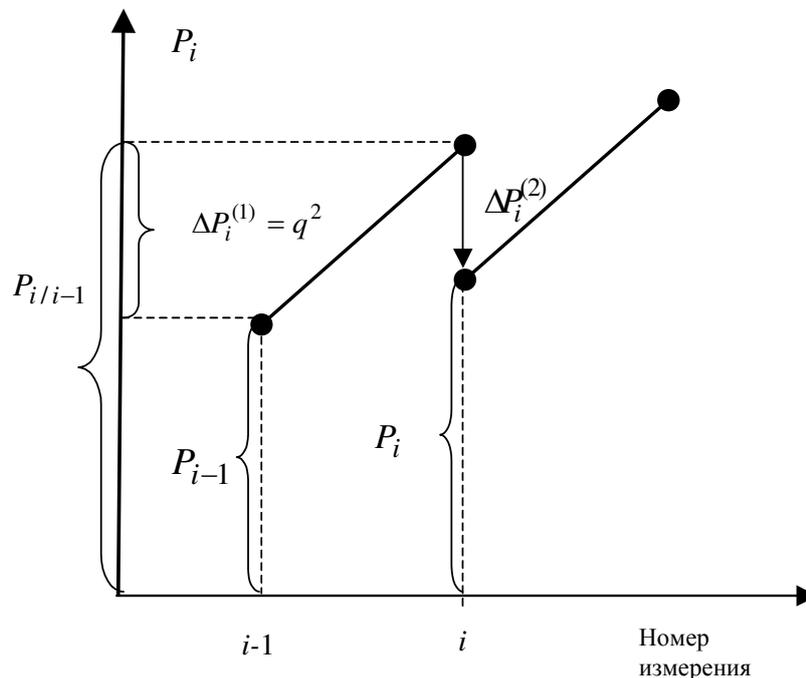


Рис. 7.6. Динамика изменения дисперсии ошибок фильтрации и прогноза в задаче оценивания винеровской последовательности

При обработке же очередного измерения, как показано на рис. 7.6, дисперсия ошибки уменьшается, так как

$$\Delta P_i^{(2)} = P_{i/i-1} - P_i = P_{i/i-1} \left(\frac{P_{i/i-1}}{P_{i/i-1} + r^2} \right) \geq 0.$$

Если $\Delta P_i^{(2)} < q^2$, то дисперсия ошибки на очередном шаге будет больше, чем на предыдущем, т.е. $P_i > P_{i-1}$, и при увеличении числа измерений дисперсия ошибки будет возрастать. Если $\Delta P_i^{(2)} > q^2$, то дисперсия ошибки при увеличении числа измерений будет уменьшаться: $P_i < P_{i-1}$.

При $\Delta P_i^{(1)} = \Delta P_i^{(2)}$ дисперсия на текущем и предыдущем шагах будет одинаковой, и ее значение при заданных q^2 и r^2 можно найти путем решения уравнения

$$\frac{(P_{i-1} + q^2)^2}{P_{i-1} + q^2 + r^2} = q^2.$$

Нетрудно убедиться в том, что решение этого уравнения будет совпадать с решением уравнения вида (7.46), которое в данном случае записывается как

$$(P_\infty^\phi + q^2) \left(\frac{r^2}{P_\infty^\phi + q^2 + r^2} \right) - P_\infty^\phi = 0$$

или

$$(P_\infty^\phi)^2 + P_\infty^\phi q^2 - q^2 r^2 = 0.$$

Решив уравнение, для установившейся дисперсии получим

$$P_\infty^\phi = -\frac{q^2}{2} \pm \sqrt{\frac{q^4}{4} + q^2 r^2} = -\frac{q^2}{2} \pm \frac{q^2}{2} \sqrt{1 + \frac{4r^2}{q^2}}.$$

Заметим, что рассматриваемое уравнение имеет два решения. В принципе необходимо проанализировать оба эти решения, если они положительны. Если одно из них отрицательно, то, поскольку дисперсия отрицательной быть не может, естественно, используется только одно положительное решение. Если значение дисперсии в начальный момент времени принять равным $\sigma_0^2 = P_\infty$, то дисперсия будет неизменной, начиная с первого измерения; если $\sigma_0^2 > P_\infty$, то дисперсия будет уменьшаться, стремясь сверху к установившемуся значению $\sigma_i^2 \xrightarrow{i \rightarrow \infty} P_\infty$. Если $\sigma_0^2 < P_\infty$, то дисперсия будет увеличиваться, приближаясь к тому же самому значению снизу. При наличии установившегося значения у дисперсии ошибки фильтрации коэффициент усиления и дисперсия ошибки прогноза будут также стремиться к установившимся значениям:

$$K_\infty = \frac{P_\infty^\phi}{r^2}; \quad P_\infty^{np} = P_\infty^\phi + q^2.$$

Так, к примеру, полагая $q^2 = 1$, $r^2 = 2$, получаем $P_\infty = 1$ и $K_\infty = \frac{1}{2}$. Таким образом, фильтр Винера будет иметь вид

$$\hat{x}_i = 0,5\hat{x}_{i-1} + 0,5y_i = \frac{\hat{x}_{i-1} + y_i}{2},$$

т.е. оценка на текущем шаге представляет собой среднее арифметическое между оценкой на предыдущем шаге и текущим измерением.

Как показано в задаче 7.6, при выполнении условия $q \ll r$ можно записать: $P_\infty \approx rq$; $P_\infty^{np} \approx q(r+q)$; $K_\infty \approx \frac{q}{r}$. В случае же, когда $q \gg r$, $P_\infty \approx r^2$,

$P_\infty^{np} \approx q^2$, $K_\infty \approx 1$. ♦

Нетрудно конкретизировать уравнение (7.46) (задача 7.7) и для установившегося режима задачи фильтрации экспоненциально-коррелированной последовательности по измерениям на фоне белого шума.

7.6. Постановка и общее решение задачи рекуррентной оптимальной фильтрации случайных последовательностей с нелинейными измерениями

Рассмотрим более общую, чем в предыдущем разделе, постановку рекуррентной оптимальной байесовской задачи фильтрации случайных последовательностей. Во-первых, снимем ограничение на линейный характер оценок, используемых при минимизации критерия (7.25), а во-вторых, будем считать, что зависимость измерений от оцениваемых параметров может быть нелинейной.

Постановка такой задачи в общем случае может быть сформулирована следующим образом [24, 25, 30, 37].

Задана n -мерная случайная последовательность в виде формирующего фильтра

$$x_i = \Phi_i x_{i-1} + \Gamma_i w_i, \quad (7.47)$$

и имеются m -мерные измерения

$$y_i = s_i(x_i) + v_i, \quad (7.48)$$

где w_i – p -мерный вектор порождающих шумов; v_i – m -мерный вектор ошибок измерения; Φ_i , Γ_i – известные матрицы размерности $m \times n$, $n \times p$ соответственно; $s_i(x_i) = (s_{i1}(x_i), \dots, s_{im}(x_i))^T$ – известная m -мерная функция, в общем случае нелинейная относительно аргумента x .

Последовательности w_i и v_i представляют собой дискретные, центрированные белые шумы, для которых значения в несовпадающие моменты времени, кроме того, независимы между собой, а вектор начальных условий x_0 считается центрированным вектором с матрицей ковариаций P_0 . Векторы x_0 , w_i , v_i также считаются независимыми между собой.

Предполагаются также известными законы распределения для случайного вектора x_0 и последовательностей w_i и v_i , задаваемые с помощью соответствующих ф.п.р.в. $f_{x_0}(x_0)$, $f_{w_i}(w_i)$ и $f_{v_i}(v_i)$.

Требуется, располагая измерениями $Y_i = (y_1^T, y_2^T, \dots, y_i^T)^T$ и не вводя ограничений на класс используемых оценок, **найти** рекуррентный алгоритм вычисления оптимальных в среднеквадратическом смысле оценок последовательности (7.1), минимизирующих критерий

$$J_i^{\circ} = M \left\{ (x_i - \hat{x}_i(Y_i))^T (x_i - \hat{x}_i(Y_i)) \right\}, \quad (7.49)$$

и соответствующих им характеристик точности в виде матриц ковариаций ошибок оценивания (7.26).

Когда измерения линейны и при минимизации (7.49) используются только линейные оценки, решение задачи определяется соотношениями (7.28)–(7.32). Их обоснование проведено в 7.3 исходя из соотношений для линейной задачи оценивания постоянного вектора. В рассматриваемом случае для получения искомых алгоритмов будем опираться на подраздел 4.2, согласно которому искомая оценка определяется как

$$\hat{x}_i(Y_i) = \int x_i f(x_i / Y_i) dx_i, \quad (7.50)$$

где $f(x_i / Y_i)$ – апостериорная ф.п.р.в. или просто апостериорная плотность.

Характеристики точности в виде **условной и безусловной апостериорных матриц ковариаций ошибок оптимальных оценок случайной последовательности** будут задаваться следующими соотношениями:

$$P_i(Y_i) = \int (x_i - \hat{x}_i(Y_i))(x_i - \hat{x}_i(Y_i))^T f(x_i / Y_i) dx_i; \quad (7.51)$$

$$P_i = \iint (x_i - \hat{x}_i(Y_i))(x_i - \hat{x}_i(Y_i))^T f(x_i, Y_i) dx_i dY_i. \quad (7.52)$$

Здесь безусловная апостериорная матрица ковариаций P_i характеризует **потенциальную точность решения сформулированной задачи оптимальной фильтрации** в среднем по всем измерениям, а матрица $P_i(Y_i)$ характеризует расчетную точность для данного текущего набора измерений. Именно для этой матрицы и предполагается получить рекуррентный алгоритм ее вычисления.

Оценку (7.50) будем называть **оптимальной в среднеквадратическом смысле байесовской или просто оптимальной оценкой случайной последовательности**. Она будет обладать всеми свойствами оптимальных оценок, подробно рассмотренных в подразделе 4.3.

Здесь так же, как это было сделано в 4.2, уместно обратить внимание на то, что помимо указания на правило вычисления оптимальной оценки (7.50) задается и правило (7.51) вычисления соответствующей ей **текущей апостериорной матрицы ковариаций ошибок $P_i(Y_i)$** , характеризующей расчетную точность оценивания для конкретного набора измерений. Таким образом, как и в случае оценивания постоянного вектора, под **задачей синтеза оптимальных алгоритмов оценивания случайных последовательностей** будем понимать задачу построения процедуры, обеспечивающей вычисление как самой оценки (7.50), так и соответствующей ей условной апостериорной матрицы ковариаций (7.51). Саму эту процедуру назовем **оптимальным алгоритмом**. Под **задачей же анализа точности оценивания случайных последовательностей** будем понимать задачу вычисления и анализа безусловной апостериорной матрицы ковариаций P_i .

Отметим основные отличительные особенности приведенной постановки по сравнению с постановкой в подразделе 7.2.

1. Не вводятся ограничения на класс используемых оценок при минимизации выбранного критерия.
2. Предполагается нелинейным характер зависимости измерений от оцениваемых параметров.
3. Считаются известными функции плотности распределения для вектора начальных условий, порождающих и измерительных шумов, а не только их первые два момента.
4. Вместо условий некоррелированности (7.23), (7.24) значений случайных векторов в различные моменты времени предполагаются выполненными условия их независимости.

Из соотношений (7.50), (7.51) следует, что, как и в случае оценивания вектора постоянных параметров, для решения сформулированной задачи необходимо располагать апостериорной плотностью $f(x_i / Y_i)$, нахождение которой и составляет основное содержание и основную сложность задачи нелинейной фильтрации [24, 37].

Решение задачи рекуррентной оптимальной фильтрации существенным образом упрощается при гауссовском характере апостериорной плотности. Нетрудно понять, что апостериорная плотность будет гауссовской, если зависимость измерений от оцениваемой последовательности линейная, т.е. $s_i(x_i) = H_i x_i$, а векторы начальных условий, порождающих и измерительных шумов являются гауссовскими, т.е.

$$f(x_0) = N(x_0; 0, P_0); f(w_i) = N(w_i; 0, Q_i); f(v_i) = N(v_i; 0, R_i).$$

Таким образом, ясно, что, если в задаче, рассматриваемой в подразделе 7.2, дополнительно предположить гауссовский характер начальных условий, порождающих и измерительных шумов, то апостериорная плотность будет тоже гауссовской. Более того, можно показать, что эта задача решается с использованием соотношений (7.28)-(7.32) для ФК.

Отсюда следует, что приведенные в подразделе 7.3 соотношения ФК в гауссовском случае обеспечивают получение оптимальных в среднеквадратическом смысле байесовских оценок случайной последовательности (7.21) по измерениям (7.22), т.е. оценок, минимизирующих среднеквадратический критерий без ограничений на класс оценок.

Из приведенных утверждений следует, что для рассматриваемого частного случая оценки, вырабатываемые ФК, обладают всеми свойствами оптимальных оценок (см. подраздел 4.3).

В частности, условная и безусловная матрицы ковариаций ошибок оптимальных оценок будут совпадать между собой, т.е. $P_i = P_i(Y_i)$, и если определить произвольный, не обязательно линейный, алгоритм вычисления оценок $\tilde{x}(Y_i)$ и найти соответствующую безусловную матрицу ковариаций их ошибок \tilde{P}_i , то в соответствии со свойствами 3, 4, приведенными в подразделе 4.3, всегда будут выполняться следующие соотношения:

$$\begin{aligned}\tilde{P}_i - P_i &\geq 0, \\ \det(\tilde{P}_i) &\geq \det(P_i).\end{aligned}$$

Важным также является свойство 5, означающее, что получение оптимальной оценки полного вектора состояний обеспечивает получение оптимальной оценки вектора, представляющего собой произвольное линейное преобразование оцениваемого вектора.

Таким образом, матрица ковариаций ошибок оценок (7.32), вычисляемая в фильтре Калмана, характеризует **потенциальную точность оценивания случайной гауссовской последовательности, формируемой с помощью уравнений (7.21).**

Следует заметить, что соотношения (7.50), (7.51) могут быть использованы и в более общем случае, когда нелинейными являются не только измерения, но и уравнения формирующего фильтра. Для получения конкретных алгоритмов используются различные методы аппроксимации апостериорной плотности, как это описано в многочисленных работах, в частности в [24, 37].

Задачи к разделу

Задача 7.1. Пусть в задаче фильтрации, сформулированной в п. 7.2, известна матрица ковариаций ошибок фильтрации на предыдущем шаге $P_{i-1} = M \{ \varepsilon_{i-1} \varepsilon_{i-1}^T \}$, а на текущем шаге прогноз и оценка вычисляются в виде $\hat{x}_{i/i-1} = \Phi_i \hat{x}_{i-1}$; $\hat{x}_i = \hat{x}_{i/i-1} + K_i (y_i - H_i \hat{x}_{i/i-1})$, где матрица K_i задается выражением (7.31). Покажите, что для матриц ковариаций ошибок прогноза и ошибок фильтрации справедливы соотношения (7.29), (7.34).

Задача 7.2. Покажите, что матрица ковариаций для невязки $\mu_i = y_i - H_i \hat{x}_{i/i-1}(Y_{i-1})$ определяется выражением $L_i = H_i P_{i/i-1} H_i^T + R_i$.

Задача 7.3. Докажите справедливость соотношения $M \{ \varepsilon_i(Y_i) x_i^T \} = P_i$.

Задача 7.4. Запишите уравнение ошибок для установившегося режима задачи из примера 7.6.

Задача 7.5. Найдите значение дисперсии ошибки фильтрации, ошибки прогноза и коэффициента усиления в установившемся режиме в задаче оценивания винеровской последовательности из примера 7.6 для случая, когда $q \ll r$ и $q \gg r$.

Задача 7.6. Конкретизируйте уравнение (7.46) для установившегося значения дисперсии ошибки фильтрации скалярной экспоненциально-коррелированной последовательности по измерениям на фоне белого шума. Найдите его приближенное решение и выражения для дисперсии ошибки прогноза, коэффициента усиления и фильтр Винера при выполне-

нии условия $q \gg r$. Установите связь уравнения для установившегося значения дисперсии ошибки фильтрации с аналогичным уравнением в примере 7.6.

Контрольные вопросы

1. Сформулируйте постановку задачи нахождения оптимальных в среднеквадратическом смысле линейных оценок одной случайной последовательности по измерениям другой, коррелированной с ней последовательности. Запишите ее решение с использованием дискретного варианта уравнений Винера–Хопфа.
Поясните, в чем специфика задач фильтрации, сглаживания и прогноза.
2. Сформулируйте постановку линейной задачи рекуррентного оценивания случайной последовательности, заданной с помощью формирующего фильтра.
3. Какие основные блоки содержит алгоритм дискретного фильтра Калмана? Поясните, почему этот алгоритм является рекуррентным. Конкретизируйте эти блоки на примере решения задачи оценивания постоянной скалярной величины.
4. Поясните, что такое ошибки прогноза и ошибки фильтрации. Запишите уравнение для этих ошибок. Почему ошибки фильтрации являются марковской последовательностью?
5. Как соотносятся матрицы ковариаций для ошибок прогноза и фильтрации при наличии установившегося режима?
6. При каких условиях существует установившийся режим в задаче фильтрации?
7. Что такое фильтр Винера? Поясните его связь с фильтром Калмана. Приведите пример.
8. Приведите постановку и общее решение задачи рекуррентной оптимальной байесовской фильтрации случайных последовательностей.
9. В чем отличие постановки задачи рекуррентной оптимальной байесовской фильтрации от задачи рекуррентной оптимальной линейной фильтрации случайных последовательностей?
10. При каких дополнительных условиях по сравнению с условиями задачи п.7.2 фильтр Калмана будет оптимальным без введения ограничения на класс используемых оценок?

РАЗДЕЛ 8

Фильтрация случайных процессов. Фильтр Калмана–Бьюси

Как правило, на практике постановка прикладных задач фильтрации формулируется в непрерывной форме с использованием дифференциальных уравнений. В этом случае приходится иметь дело со случайными процессами. Вопросам фильтрации случайных процессов с позиций подхода, предложенного Р. Калманом, и посвящен настоящий раздел [11, 18, 20, 21, 26, 30-34, 35, 37].

8.1. Случайные процессы и методы их описания

Случайным процессом $x(t)$ в скалярном случае называется такая функция времени t , значение которой при любом фиксированном значении t является случайной величиной.

Далее по аналогии с тем, как это было сделано применительно к случайным последовательностям, для случайных процессов $x(t)$ и их возможных реализаций $x(t)$ будем использовать одинаковые обозначения $x(t)$.

Основные характеристики и классы случайных процессов вводятся по аналогии со случайными последовательностями. В частности, корреляционная функция для случайного процесса определяется как

$$k_x(t_1, t_2) = \iint (x(t_1) - \bar{x}(t_1))(x(t_2) - \bar{x}(t_2))^T f(x_1, t_1, x_2, t_2) dx_1 dx_2, \quad (8.1)$$

а для стационарных процессов, как следует из подраздела 6.2, корреляционная функция будет зависеть от разности аргументов.

◆ **Пример 8.1.** В качестве примера корреляционной функции стационарного процесса можно привести функцию вида

$$k_x(\tau) = \sigma_x^2 e^{-\alpha|\tau|}. \quad (8.2)$$

Процесс с такой корреляционной функцией называется **экспоненциально-коррелированным** и является обобщением экспоненциально-коррелированной последовательности, введенной в примере 6.6 [25]. Здесь $k_x(0) = \sigma_x^2$ – дисперсия процесса, а $\tau_k = 1/\alpha$ – **интервал корреляции**.

Графики функции (8.2) при единичной дисперсии и двух значениях интервалов корреляции $\tau_k = 1$ с и $\tau_k = 0.1$ с приведены на рис. 8.1.

Для описания свойств стационарных процессов, помимо корреляционной функции, используют его **спектральную плотность**, которая вводится как преобразование Фурье от корреляционной функции, т.е.

$$S_x(\omega) = \int_{-\infty}^{\infty} k_x(\tau) \exp(-j\omega\tau) d\tau. \quad (8.3)$$

Справедливо и обратное представление:

$$k_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) \exp(j\omega\tau) d\omega. \quad (8.4)$$

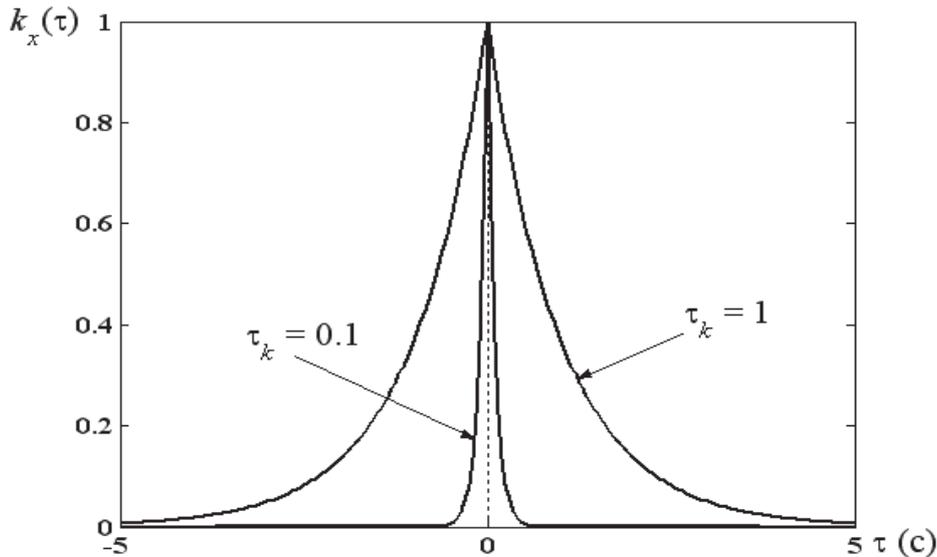


Рис. 8.1. Корреляционная функция $k_x(\tau) = \sigma_x^2 e^{-\alpha|\tau|}$ при двух значениях интервалов корреляции ♦

Функция (8.3) иногда называется **энергетическим спектром** стационарного процесса [16] или **спектральной плотностью средней мощности**. В англоязычной литературе для (8.3) используется термин **power spectral density (PSD)** [32]. В силу четного характера функций $S_x(\omega)$, $k_x(\tau)$ приведенные соотношения можно записать как

$$S_x(\omega) = \int_{-\infty}^{\infty} k_x(\tau) \cos \omega\tau d\tau = 2 \int_0^{\infty} k_x(\tau) \cos \omega\tau d\tau;$$

$$k_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) \cos \omega\tau d\omega = \frac{1}{\pi} \int_0^{\infty} S_x(\omega) \cos \omega\tau d\omega.$$

Из последнего выражения с очевидностью вытекает следующее соотношение:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) d\omega = k_x(0) = \sigma_x^2, \quad (8.5)$$

означающее, что площадь, ограниченная $S_x(\omega)$ и осью абсцисс, с точностью до постоянного коэффициента определяет величину дисперсии процесса σ_x^2 .

◆ **Пример 8.2** Получим выражение для спектральной плотности экспоненциально-коррелированного процесса с корреляционной функцией (8.2) и проанализируем ее поведение. Используя таблицы преобразования Фурье, можно показать, что спектральная плотность, соответствующая функции (8.2), будет определяться как [18, 26]

$$S_x(\omega) = \frac{2\sigma_x^2\alpha}{\alpha^2 + \omega^2}. \quad (8.6)$$

Графики спектральной плотности при двух значениях интервала корреляции $\tau_k = 1/\alpha$ приведены на рис. 8.2.

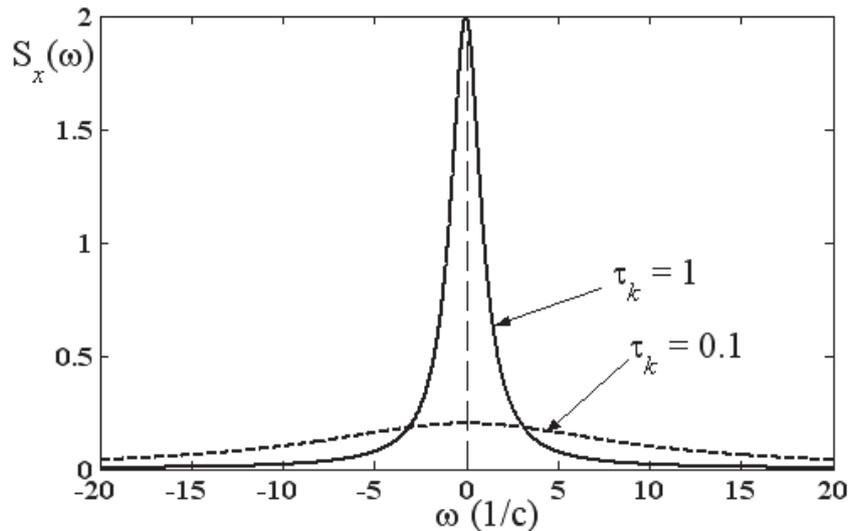


Рис. 8.2. Графики спектральной плотности экспоненциально-коррелированного процесса при разных значениях интервала корреляции

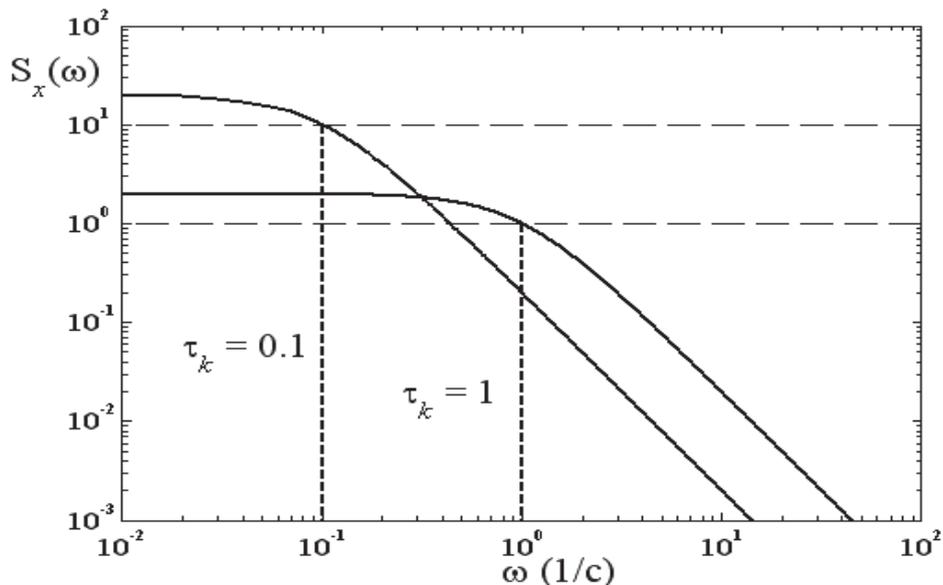


Рис. 8.3. Графики спектральной плотности экспоненциально-коррелированного процесса при разных значениях интервала корреляции в логарифмическом масштабе ◆

Полоса частот, в которой эта спектральная плотность отлична от нуля, может быть оценена величиной α , поскольку при изменении ω в пределах от 0 до α значения спектральной плотности уменьшаются лишь в два раза – от $S_x(0) = \frac{2\sigma_x^2}{\alpha}$ до $S_x(\alpha) = \frac{\sigma_x^2}{\alpha}$, существенно снижаясь при $\omega > \alpha$.

Наиболее отчетливо это проявляется, если построить график спектральной плотности для положительных значений частот в логарифмическом масштабе. Пример таких графиков при $\sigma = 1$ и двух значениях $\alpha = 1 \cdot \text{с}^{-1}$ и $\alpha = 10 \cdot \text{с}^{-1}$ представлен на рис. 8.3.

Процесс с постоянной спектральной плотностью во всем диапазоне частот, для которого $S_x(\omega) = Q$, называется **белым шумом**. Величина Q называется **интенсивностью белого шума**. Корреляционная функция для белого шума имеет вид

$$k_x(\tau) = Q\delta(\tau), \quad (8.7)$$

который вытекает из следующего представления для дельта-функции $\delta(\tau)$ [14]:

$$\frac{1}{2\pi} \int e^{j\omega\tau} d\omega = \delta(\tau).$$

Из выражения (8.7) следует, что непрерывный белый шум, в отличие от дискретного, имеет бесконечную дисперсию. Значения же в любые несопадающие моменты времени, как и в случае дискретного времени, не коррелированы.

Обращаем внимание на то, что размерность интенсивности белого шума равна размерности его дисперсии, умноженной на единицу времени, т.е.

$$[Q] = [\text{единицы}^2][\text{время}].$$

Сказанное вполне согласуется с выражением $k(\tau) = Q\delta(\tau)$, поскольку здесь слева стоит величина, имеющая размерность $[\text{единицы}^2]$, а справа – величина, размерность которой определяется как $[Q][\text{время}]^{-1}$. Это есть следствие того, что дельта-функция, у которой аргументом является время, имеет размерность, обратную размерности времени $[\delta(t)] = [\text{время}]^{-1}$, так как из свойств дельта-функции следует, что

$$\int_{-\infty}^{\infty} \delta(\tau) d\tau = 1.$$

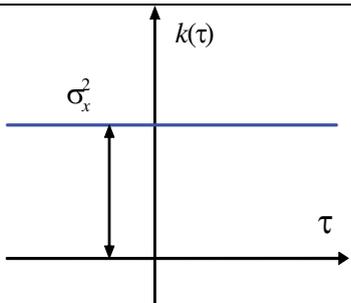
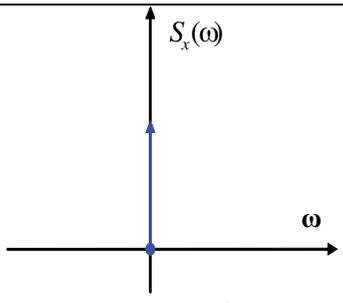
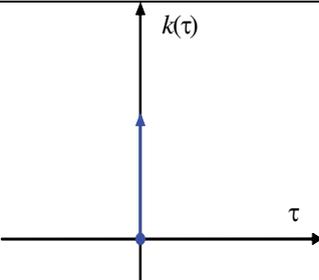
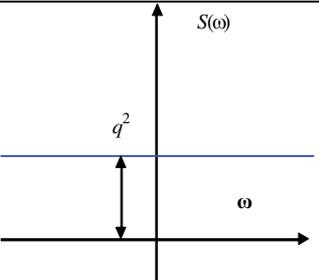
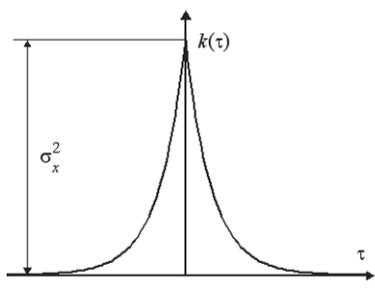
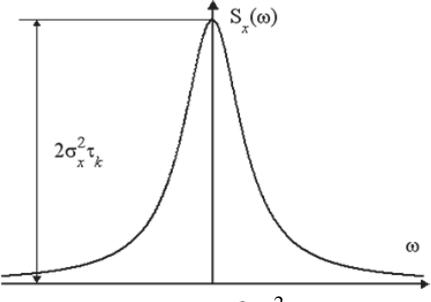
Можно отметить, что белый шум обладает в некотором смысле противоположными свойствами по сравнению с частным случаем процесса, представляющего собой неизменную во времени случайную величину, для которого корреляционная функция постоянна, а спектральная плотность – дельта-функция:

$$k_x(\tau) = \sigma_x^2, ; \quad S_x(\omega) = 2\pi\sigma_x^2\delta(\omega).$$

Примеры графиков корреляционных функций и спектральных плотностей случайной константы, белого шума и экспоненциально-коррелированного процесса сведены в табл. 8.1.

Т а б л и ц а 8.1

Корреляционные функции и спектральные плотности простейших процессов

Вид процесса	Корреляционная функция	Спектральная плотность
Случайная константа с дисперсией σ_x^2	 $k(\tau) = \sigma_x^2$	 $S_x(\omega) = 2\pi\sigma_x^2\delta(\omega)$
Белый шум интенсивности q^2	 $k_x(\tau) = \sigma_x^2\delta(\tau)$	 $S_x(\omega) = q^2$
Экспоненциально-коррелированный процесс	 $k_x(\tau) = \sigma_x^2 e^{-\alpha \tau }$	 $S_x(\omega) = \frac{2\sigma_x^2\alpha}{\alpha^2 + \omega^2}$

8.2. Формирующий фильтр

Рассмотрим n -мерный случайный процесс $x(t)$, задаваемый в виде

$$\dot{x}(t) = F(t)x(t) + G(t)w(t). \quad (8.8)$$

где $F(t)$, $G(t)$, $Q(t)$ – известные $n \times n$ -, $n \times p$ - и $p \times p$ -матрицы. Полагаем, что в начальный момент времени $x(0)$ представляет собой случайный вектор с известным математическим ожиданием $\bar{x}(0)$ и матрицей ковариаций $P(0)$, $w(t)$ – не зависящий от $x(0)$ центрированный p -мерный белый шум

$$M(w(t+\tau)w^T(t)) = Q(t)\delta(t-\tau). \quad (8.9)$$

Как и в случае последовательностей, уравнение (8.8), используемое для описания процессов, называется **формирующим (shaping) фильтром**; белый шум в правой части – **порождающим (generating, forcing) белым шумом** или **шумом системы**; матрица $F(t)$ – **матрицей динамики**; матрица $G(t)$ – **матрицей порождающих шумов**, а вектор $x(t)$ – **вектором состояния**.

Используем выражение (ПЗ.8) из приложения 3 и запишем решение уравнения (8.8) в виде

$$x(t) = \Phi(t, t_1)x(t_1) + \int_{t_1}^t \Phi(t, \tau)G(\tau)w(\tau)d\tau, \quad (8.9)$$

где $\Phi(t, t_1)$ – фундаментальная матрица для уравнения $\dot{x}(t) = F(t)x(t)$.

Можно показать, что математическое ожидание, матрица ковариаций и корреляционная функция определяются следующими соотношениями (задача 8.1):

$$\bar{x}(t) = \Phi(t, t_0)\bar{x}(t_0); \quad (8.10)$$

$$P(t) = \Phi(t, t_0)P(t_0)\Phi^T(t, t_0) + \int_{t_0}^t \Phi(t, \tau)G(\tau)Q(\tau)G^T(\tau)\Phi^T(t, \tau)d\tau; \quad (8.11)$$

$$k(t_2, t_1) = \begin{cases} \Phi(t_2, t_1)P(t_1), & t_2 > t_1, \\ P(t_2)\Phi^T(t_1, t_2), & t_2 \leq t_1. \end{cases} \quad (8.12)$$

Принимая во внимание выражения (ПЗ.9), (ПЗ.10), нетрудно также убедиться, что матрица ковариаций (8.11) является решением дифференциального уравнения

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t). \quad (8.13)$$

Частным случаем уравнений (8.8), (8.13) являются стационарные уравнения, в которых матрицы F , Q и G постоянны:

$$\dot{x}(t) = Fx(t) + Gw(t); \quad (8.14)$$

$$\dot{P} = FP + PF^T + GQG^T. \quad (8.15)$$

Обсудим условия, при которых на выходе стационарной системы (8.14) можно получить стационарный процесс. Очевидно, что в первую очередь необходимо, чтобы математическое ожидание процесса в начальный момент времени равнялось нулю, т.е. $\bar{x}(t_0) = 0$, в противном случае математическое ожидание процесса $x(t)$, как следует из (8.10), будет зависеть от времени. Для того чтобы от времени не зависела матрица ковариаций, определяемая уравнением (8.13), необходимо выполнение условий, при которых существует его установившееся решение. Это означает, что существует матрица P_∞ , такая что при $P = P_\infty$

$$\dot{P}_\infty = FP_\infty + P_\infty F^T + GQG^T = 0. \quad (8.16)$$

Если матрицу ковариаций $P(0)$ для вектора $x(0)$ выбрать совпадающей с решением этого уравнения $P(0) = P_\infty$, то процесс $x(t)$, порожденный уравнением (8.14), становится стационарным, поскольку $P(t) \equiv P(0)$, кроме того, как вытекает из (8.12), корреляционная функция

$$k(\tau) = \Phi(\tau)P_\infty, k(-\tau) = P_\infty\Phi^T(\tau) \quad (8.17)$$

будет зависеть только от τ . Из (8.17) также следует, что

$$k(\tau) = k^T(-\tau) = \Phi(\tau)P_\infty. \quad (8.18)$$

Таким образом, **условиями стационарности процесса на выходе стационарной системы** при поступлении на ее вход белого шума являются центрированность значений процесса в начальный момент времени, наличие решения уравнения (8.16) и выбор начальной матрицы ковариаций, совпадающей с этим решением.

Если установившееся решение (8.16) существует, но начальная матрица ковариаций не совпадает с P_∞ , то, поскольку $P(t) \rightarrow P_\infty$ при увеличении времени, процесс после завершения переходного режима при $t \rightarrow \infty$ можно считать стационарным.

З а м е ч а н и е 1. Если дополнительно предположить, что $x(0)$ и порождающий шум – гауссовские, т.е.

$$f(x(0)) = N(x(0); \bar{x}(0), P(0)); \quad (8.19)$$

$$f(w(t)) = N(w(t); 0, Q(t)), \quad (8.20)$$

то и процесс $x(t)$ также будет гауссовским. Это есть следствие того, что при линейных преобразованиях гауссовских векторов в результате получается гауссовский вектор.

З а м е ч а н и е 2. Используя выражение (8.9), нетрудно убедиться в том, что процесс $x(t)$ является марковским. Действительно, если зафиксировать моменты времени $t_1 > t_2 > t_3$, то, как следует из (8.9), значение процесса в момент t_3 при фиксированных его значениях в моменты t_1 и t_2 зависит только от момента t_2 и не зависит от t_1 . При этом существенным является тот факт, что $w(t)$ – белый шум, не зависящий в статистическом смысле от начальных условий $x(0)$.

◆ **П р и м е р 8.3.** Рассмотрим формирующий фильтр

$$\dot{x}(t) = -\alpha x(t) + \sqrt{2\sigma_x^2\alpha}w(t). \quad (8.21)$$

Получим корреляционную функцию, соответствующую этому процессу, и конкретизируем уравнение (8.15).

Поскольку $F = -\alpha$, $q_w = \sqrt{2\sigma_x^2\alpha}$, уравнение (8.14) примет вид

$$\dot{P} = -\alpha P - \alpha P + q_w^2.$$

В силу того что $\Phi(t, t_0) = e^{-\alpha(t-t_0)}$, решение этого уравнения можно представить как

$$P(t) = P(t_0)e^{-2\alpha(t-t_0)} + 2\sigma_x^2\alpha \int_{t_0}^t e^{-2\alpha(t-\tau)} d\tau.$$

Понятно, что (8.16) сводится к уравнению $2\alpha P_\infty = 2\sigma_x^2\alpha$, имеющему решение $P_\infty = \sigma_x^2$. Таким образом, при $P(0) = \sigma_x^2$ процесс будет стационарным, а соответствующая ему корреляционная функция примет вид

$$k(\tau) = k(-\tau) = \Phi(\tau)P = \sigma_x^2 e^{-\alpha\tau}. \quad \blacklozenge$$

8.3. Постановка и общее решение задачи оптимальной линейной фильтрации. Фильтр Калмана-Бьюси

Приведем математическую постановку и общее решение задачи оптимальной линейной фильтрации, рассматриваемой в рамках концепции пространства состояний. Напомним, что такое решение задачи фильтрации было получено Р. Калманом сначала для дискретного времени, а уже затем совместно с Бьюси – для случая непрерывного времени [36]. Это и определило название для этого фильтра – **фильтр Калмана-Бьюси**.

Математическую постановку калмановской задачи непрерывной оптимальной линейной фильтрации сформулируем по аналогии с тем, как это было сделано для дискретного случая.

Предположим, что задан n -мерный марковский процесс

$$\dot{x}(t) = F(t)x(t) + G(t)w(t), \quad x(t_0) = x_0 \quad (8.22)$$

и имеются m -мерные измерения

$$y(t) = H(t)x(t) + v(t), \quad (8.23)$$

где $F(t)$, $G(t)$, $H(t)$ – известные в общем случае, зависящие от времени матрицы размерности $n \times n$, $n \times p$, $m \times n$ соответственно; x_0 – центрированный вектор начальных условий с матрицей ковариаций P_0 ; $w(t)$, $v(t)$ – не коррелированные между собой и с начальными условиями x_0 центрированные белые шумы заданной интенсивности, т.е.

$$M \{x_0 w^T(t)\} = 0; \quad M \{w(t) v^T(t)\} = 0; \quad M \{x_0 v^T(t)\} = 0; \quad (8.24)$$

$$M \{w(t) w^T(\tau)\} = Q(t) \delta(t - \tau), \quad Q(t) \geq 0; \quad (8.25)$$

$$M \{v(t) v^T(\tau)\} = R(t) \delta(t - \tau), \quad R(t) > 0. \quad (8.26)$$

Задача фильтрации в пространстве состояния для **непрерывного времени** формулируется так. С использованием измерений (8.23) $Y(t) = \{y(\tau) : \tau \in [0, t]\}$, накопленных на интервале $[0, t]$, в момент времени t необходимо получить линейную оптимальную в среднеквадратическом смысле оценку вектора $x(t)$, т.е. такую, которая обеспечивает минимизацию критерия

$$r^{\hat{}}(t) = M \left\{ (x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t)) \right\}. \quad (8.27)$$

Можно показать, что оценка и соответствующая ей матрица ковариаций ошибок определяются с помощью следующих соотношений [18, 36]:

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + K(t)(y(t) - H(t)\hat{x}(t)); \quad (8.28)$$

$$K(t) = P(t)H(t)^T R^{-1}(t); \quad (8.29)$$

$$\dot{P}(t) = P(t)F(t)^T + F(t)P(t) - P(t)H(t)^T R^{-1}(t)H(t)P(t) + G(t)Q(t)G^T(t). \quad (8.30)$$

Эти уравнения и определяют знаменитый **непрерывный фильтр Калмана–Бьюси (ФКБ)**.

Важно заметить, что невязка измерений $y(t) - H(t)\hat{x}(t)$, называемая, как и в дискретном случае, инновацией, представляет собой белый шум с матрицей интенсивности, равной R .

Располагая (8.8), (8.28), нетрудно (задача 8.3) получить уравнение и для ошибок фильтрации.

Матрица ковариации называется апостериорной матрицей ковариации.

Из (8.28) следует, что выражение для оценки может быть записано как

$$\dot{\hat{x}}(t) = (F(t) - K(t)H(t))\hat{x}(t) + K(t)y(t), \quad (8.31)$$

т.е. фильтр Калмана представляет собой линейную в общем случае нестационарную динамическую систему, свойства которой определяются матрицей $(F(t) - K(t)H(t))$, а на ее вход поступают измерения, «взвешенные» с помощью матрицы $K(t)$.

На рис. 8.3, 8.4 представлены блок-схемы формирования оцениваемого процесса и его измерений и блок-схема фильтра Калмана–Бьюси.

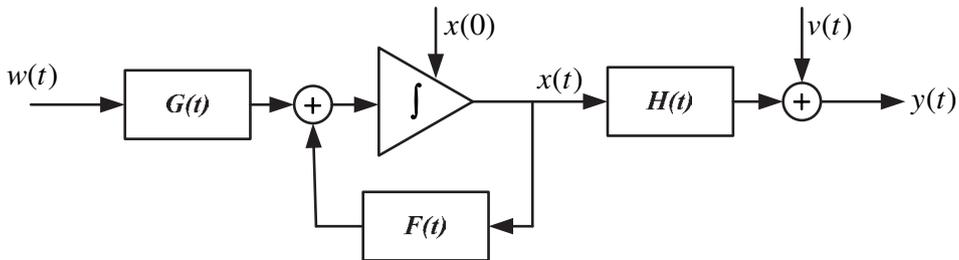


Рис. 8.3. Блок-схема формирования оцениваемого процесса и его измерений

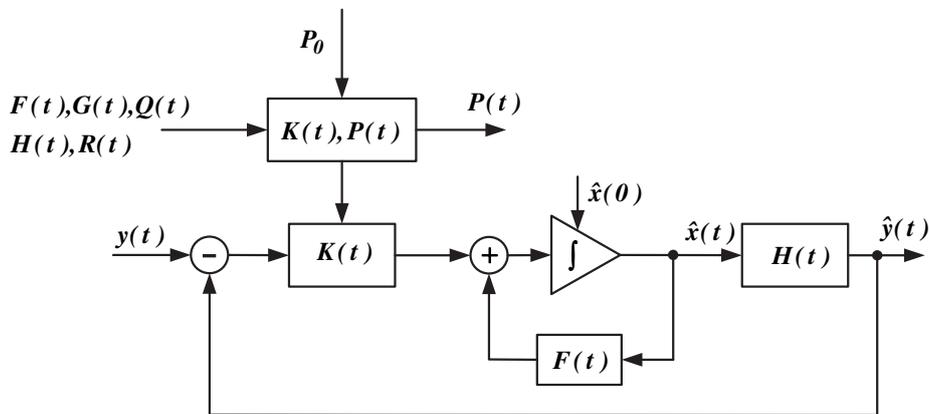


Рис. 8.4. Блок-схема фильтра Калмана–Бьюси

Сопоставление этих схем показывает, что фильтр Калмана–Бьюси в своей структуре содержит блок, аналогичный формирующему фильтру, в который добавляется единичная, отрицательная обратная связь, а вместо порождающего белого шума на вход поступают невязки измерений $\mu(t) = y(t) - H(t)\hat{x}(t)$, «взвешиваемые» с матрицей $K(t)$.

Как и в дискретном случае, в приведенном алгоритме можно условно выделить два «канала»: **оценочный «канал»**, которому соответствует линейное дифференциальное уравнение для оптимальных оценок (8.28), и **ковариационный «канал»**, которому соответствует **нелинейное дифференциальное уравнение Рикатти** (8.30). Причем это уравнение, так же как и матрица коэффициента усиления $K(t)$, не зависит от измерений, а зависит только от матриц $F(t), G(t), Q(t)$ и $H(t), R(t)$, определяющих свойства оцениваемого и измеряемого процессов. Оба эти уравнения необходимо решать при отыскании оценок, причем уравнение для матрицы ковариаций можно решить заранее, поскольку оно не зависит от поступающих измерений.

Приведенные выше соотношения вместе с условиями, при которых решается задача, сведены в табл. 8.2 [18, 21, 26, 36].

Т а б л и ц а 8.2

**Постановка задачи фильтрации
и алгоритм непрерывного фильтра Калмана–Бьюси**

Исходные данные	
Уравнение для вектора состояния	$\dot{x} = F(t)x + G(t)w(t)$
Измерения	$y = H(t)x + v(t)$
Начальные условия	x_0 –центрированный случайный вектор с матрицей ковариаций P_0
Порождающие шумы	$M\{w(t)w^T(\tau)\} = Q(t)\delta(t - \tau), Q(t) \geq 0$
Шумы измерения	$M\{v(t)v^T(\tau)\} = R(t)\delta(t - \tau), R(t) > 0$
Взаимная корреляция	$M\{x_0 w^T(t)\} = 0; M\{w(t)v^T(t)\} = 0 \quad M\{x_0 v^T(t)\} = 0$
Матрицы	$F(t) - n \times n, G(t) - n \times p \quad H(t) - m \times n$
Постановка задача фильтрации и минимизируемые критерии	
С использованием измерений $Y(t) = \{y(\tau) : \tau \in [0, t]\}$ найти оценку $\hat{x}(t)$, минимизирующую критерий $r^\delta(t) = M\{(x(t) - \hat{x}(t))^T(x(t) - \hat{x}(t))\}$, и матрицу ковариаций ошибок оценивания $P(t) = M\{(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^T\}, \quad r^\delta(t) = SpP(t)$	
Решение задачи фильтрации	
Начальные условия	$\hat{x}(0) = 0, P(0) = P_0$
Оценка	$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + K(t)(y(t) - H(t)\hat{x}(t))$

Коэффициент усиления	$K(t) = P(t)H^T(t)R^{-1}(t)$
Матрица ковариаций ошибок оценивания	$\dot{P}(t) = P(t)F^T(t) + F(t)P(t) - P(t)H^T(t)R^{-1}(t)H(t)P(t) + G(t)Q(t)G^T(t)$

Важно обратить внимание на требование невырожденности матрицы, определяющей интенсивность белозумных ошибок измерений, т.е. $R(t) > 0$. Это обусловлено тем фактом, что при вычислении матрицы ковариаций требуется матрица, обратная $R(t)$.

♦ **Пример 8.4.** Решим задачу фильтрации постоянной скалярной величины

$$\dot{x} = 0$$

по измерениям на фоне белого шума, т.е.

$$y(t) = x + v(t),$$

в которых x – центрированная случайная величина с дисперсией σ_x^2 , $v(t)$ – не коррелированный с x центрированный белый шум с корреляционной функцией

$$M\{v(t)v^T(\tau)\} = r^2\delta(t - \tau).$$

В этом примере $F = 0$, $G = 0$, $H(t) = 1$, $P_0 = \sigma_0^2$, и уравнения (8.28)–(8.30) сводятся к следующим:

$$\dot{\hat{x}}(t) = K(t)(y(t) - \hat{x}(t));$$

$$K(t) = P(t) / r^2; \dot{P}(t) = P^2(t) / r^2.$$

Путем непосредственной подстановки нетрудно убедиться в том, что

$$P(t) = \frac{\sigma_x^2 r^2}{r^2 + \sigma_x^2 t} \text{ и } K(t) = \frac{\sigma_x^2}{r^2 + \sigma_x^2 t}.$$

Если выполнено неравенство $\frac{r^2}{t} \ll \sigma_x^2$, означающее малое влияние априорной информации, можно записать следующие приближенные выражения: $P(t) = \frac{r^2}{t}$, $K(t) = \frac{1}{t}$, из которых для оценки получаем уравнение в виде

$$\dot{\hat{x}}(t) = \frac{1}{t}(y(t) - \hat{x}(t)).$$

Учитывая, что фундаментальная матрица в данном случае определяется как

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

общее решение линейного нестационарного уравнения для оценки применительно к рассматриваемой задаче можно записать

$$\hat{x}(t) = \frac{t_0}{t} x_0 + \int_0^t \frac{\tau}{t} y(\tau) \frac{1}{\tau} d\tau \rightarrow \frac{1}{t} \int_0^t y(\tau) d\tau.$$

Отсюда следует, что при сделанных предположениях в рассматриваемом примере, так же, как и в аналогичной дискретной задаче, алгоритм получения оптимальной оценки сводится к вычислению среднего значения от всех поступивших измерений. ♦

Замечание 1. Помимо самих оценок, в фильтре Калмана–Бьюси вырабатывается матрица ковариаций $P(t)$, диагональные элементы которой характеризуют текущую точность оценивания, что весьма важно при решении задач обработки навигационной информации.

Замечание 2. Поскольку общее решение нестационарного дифференциального уравнения (8.31) определяется с помощью соотношения (ПЗ.5), оценки, вырабатываемые ФКБ, линейным образом зависят от измерений. Так что оценки ФКБ являются оптимальными в среднеквадратическом смысле линейными оценками.

Заметим, что в приведенной постановке задачи, как и в предыдущем разделе, не оговаривался вид функций плотности распределения для порождающих шумов, белозумных ошибок измерений и начальных условий. Если дополнительно предположить их гауссовский характер, то оценки, определяемые согласно приведенным соотношениям, будут оптимальными в среднеквадратическом смысле без добавления слова «линейные», иными словами, они обеспечат минимизацию критерия (8.27) без предположений о линейном характере алгоритма оценивания.

Замечание 3. В случае, когда решается задача фильтрации для стационарной системы, для которой матрицы F, G, Q, R постоянны, уравнение (8.22) может иметь установившееся решение, т.е.

$$P_{\infty} F^T + F P_{\infty} - P_{\infty} H^T R^{-1} H P_{\infty} + G Q G^T = 0.$$

В этом случае, поскольку $K = P_{\infty} H^T R^{-1}$, уравнение (8.28) для оценки также становится стационарным, т.е.

$$\dot{\hat{x}} = F \hat{x} + K(y - H \hat{x}).$$

Можно показать, что этому уравнению соответствует передаточная функция

$$W(p) = (pE - (F - KH))^{-1} K.$$

Именно эта передаточная функция будет получена при решении задачи фильтрации в установившемся режиме при ее решении, которое было предложено в свое время Н. Винером. Таким образом, при сделанных предположениях фильтр Калмана будет совпадать с фильтром Винера.

Замечание 4. В ряде случаев при получении решения уравнения (8.30) при отсутствии порождающих шумов, т.е. при

$$\dot{P}(t) = P(t)F^T(t) + F(t)P(t) - P(t)H^T(t)R(t)^{-1}H(t)P(t),$$

удобно перейти к уравнению для обратной матрицы, используя соотношение $P(t)P^{-1}(t) = E$. Нетрудно убедиться, что в этом случае

$$\frac{dP^{-1}(t)}{dt} = -F^T(t)P^{-1}(t) - P^{-1}(t)F(t) + H^T(t)R(t)^{-1}H(t),$$

или, принимая во внимание (ПЗ.10)

$$P^{-1}(t) = \Phi^T(t_0, t)P_0^{-1}\Phi(t_0, t) + \int_{t_0}^t \Phi^T(\tau, t_0)H^T R^{-1}H\Phi(\tau, t_0)d\tau.$$

8.4. Связь непрерывных и дискретных алгоритмов фильтрации

Несмотря на то, что постановка прикладных задач формулируется, как правило, в непрерывной форме, их решение осуществляется с использованием средств вычислительной техники, что требует дискретного представления задачи. Для этого в первую очередь необходимо, задавшись интервалом дискретизации Δt , найти такую n -мерную случайную последовательность

$$x_i = \Phi_i x_{i-1} + \Gamma_i w_i \quad (8.32)$$

с дискретным p -мерным центрированным белым шумом w_i , математическое ожидание и матрица ковариаций которой в дискретные моменты времени совпадали бы с аналогичными параметрами для непрерывного процесса (8.8), т.е.

$$m_i \equiv Mx_i = Mx(t_i) \equiv \bar{x}(t_i), \quad (8.33)$$

$$P_i \equiv M \left\{ (x_i - m_i)(x_i - m_i)^T \right\} = M \left\{ (x(t_i) - \bar{x}(t_i))(x(t_i) - \bar{x}(t_i))^T \right\}. \quad (8.34)$$

Эти равенства будем называть **условиями стохастической** (статистической) **эквивалентности**, под которыми понимается совпадение статистических характеристик (8.33), (8.34) для непрерывного случайного процесса и его дискретного аналога для одних и тех же моментов времени [26, 32].

Из сказанного следует, что требуется найти такие матрицы Φ_i , Γ_i и матрицу ковариаций дискретного белого шума Q_i , которые обеспечат выполнение этих условий. Получающуюся в результате последовательность будем называть **последовательностью, эквивалентной в стохастическом смысле** процессу (8.8).

Принимая во внимание соотношения (8.10), заметим, что при выборе матрицы Φ_i в (8.32) в виде

$$\Phi_i = \Phi(t_{i-1} + \Delta t, t_{i-1}), \quad (8.35)$$

условие (8.32) будет выполнено.

Используя (8.11), запишем

$$P(t_i) = \Phi(t_i, t_{i-1})P(t_{i-1})\Phi^T(t_i, t_{i-1}) + \int_{t_{i-1}}^{t_i} \Phi(t_i, \tau)G(t_{i-1} + \tau)Q(t_{i-1} + \tau)G^T(t_{i-1} + \tau)\Phi^T(t_i, \tau)d\tau. \quad (8.36)$$

Принимая во внимание выражение (6.16), нетрудно заметить, что при вычислении матрицы Φ_i с помощью (8.35) для выполнения равенства $P_i = P(t_i)$, обеспечивающего совпадение матриц ковариаций непрерывного процесса и его дискретного аналога, матрицы Γ_i и Q_i следует выбирать такими, что

$$\Gamma_i Q_i \Gamma_i^T = Q_{di}^*, \quad (8.37)$$

где

$$Q_{di}^* = \int_{t_{i-1}}^{t_{i-1} + \Delta t} \Phi(t_{i-1} + \Delta t, \tau)G(\tau)Q(\tau)G^T(\tau)\Phi^T(t_{i-1} + \Delta t, \tau)d\tau. \quad (8.38)$$

Поскольку размерность матрицы Q_i соответствует размерности вектора w_i и равна p , размерность матрицы Γ_i должна быть $n \times p$.

Таким образом, для получения дискретного представления процесса, задаваемого уравнением (8.8), необходимо воспользоваться соотношениями (8.35), (8.37), (8.38).

Предположим, что непрерывная система стационарна, т.е. матрицы F, G, Q не зависят от времени. В этом случае матрица динамики определяется как матричная экспонента

$$\Phi_i = e^{F(\Delta t)}\Phi_i \approx e^{F\Delta t} \approx \sum_{v=0}^k F^v \Delta t^v / v!, \quad (8.39)$$

а выражения (8.37), (8.38) примут вид

$$\Gamma Q_d \Gamma^T = Q_d^*, \quad (8.40)$$

где

$$Q_d^* = \int_0^{\Delta t} \Phi(\tau)GQG^T\Phi^T(\tau)d\tau. \quad (8.41)$$

Полагая в $\Phi(\tau) \approx \Phi(\tau^*) = \Phi^*$, где $\tau^* \in [0, \Delta t]$, по теореме о среднем можем записать

$$Q_d^* = \Phi^* G Q G^T (\Phi^*)^T \Delta t. \quad (8.42)$$

Если, к примеру, в силу малости Δt принять $\Phi_i^* \approx E$, то

$$Q_d^* \approx G Q G^T \Delta t. \quad (8.43)$$

Используя (8.37), (8.38), запишем

$$\Gamma Q_d \Gamma^T \approx G Q G^T \Delta t. \quad (8.44)$$

Полагая

$$\Gamma = G \Delta t, \quad (8.45)$$

для матрицы ковариаций дискретного белого шума, соответственно, получаем следующее выражение:

$$Q_d = \frac{1}{\Delta t} Q. \quad (8.46)$$

Если ввести дискретный шум как среднее значение непрерывного белого шума на интервале дискретизации, т.е.

$$w_i = \frac{1}{\Delta t} \int_0^{\Delta t} w(t) dt, \quad (8.47)$$

то соответствующая ему матрица ковариаций будет определяться как $M(w_i w_i^T) = \frac{1}{\Delta t} Q$ (см. задачу 8.2), т.е. при выборе матриц с помощью соотношений (8.45), (8.46) матрица ковариаций дискретного шума соответствует матрице ковариаций непрерывного шума, осредненного на интервале дискретизации.

Следует иметь в виду приближенный характер приведенных соотношений. В особой степени это касается матриц Q_d и Γ . Для более точного их нахождения требуется вычисление матрицы Q_d^* согласно (8.38) с последующим использованием соотношения (8.37).

Чтобы завершить получение постановки задачи фильтрации в дискретной форме, необходимо найти матрицы H_i и R_i , с помощью которых можно записать соответствующие измерения. При определении этих матриц исходят из того, что набор измерений на интервале $t \in [t_{i-1}, t_{i-1} + \Delta t]$ заменяется одним измерением, формируемым путем осреднения, т.е.

$$y_i = \frac{1}{\Delta t} \int_{t_{i-1}}^{t_i} (H(t)x(t) + v(t)) dt. \quad (8.48)$$

В этом случае матрицы H_i и R_i определяются как

$$H(t) \approx H(t_i) = H_i; \quad (8.49)$$

$$R(t) \approx R(t_i), \quad t \in [t_{i-1}, t_{i-1} + \Delta t]; \quad (8.50)$$

$$R_i \approx \frac{1}{\Delta t} R(t_i). \quad (8.51)$$

Последнее соотношение следует из того факта, что

$$v_i = \frac{1}{\Delta t} \int_{t_{i-1}}^{t_i} v(t) dt.$$

Ясно, что для получения дискретного аналога непрерывной задачи важным представляется правильный выбор интервала дискретизации Δt . Этот интервал, в первую очередь, необходимо выбирать из условий малой изменчивости на интервале дискретизации $t \in [t_{i-1}, t_{i-1} + \Delta t]$ матриц $F(t)$, $H(t)$ и матриц, входящих в (8.38). Кроме этого, необходимо учитывать возмож-

ную изменчивость процесса, обусловленную внутренней динамикой, определяемой матрицей $F(t) \approx F(t_i) = F_i$, которая предполагается постоянной для $t \in [t_{i-1}, t_{i-1} + \Delta t]$. Известно, что эта изменчивость определяется собственными частотами, которые зависят от собственных чисел матрицы. В этой связи при выборе интервала дискретизации стремятся обеспечить справедливость соотношения

$$\Delta t \leq k / |\lambda|_{\max}, \quad (8.52)$$

где k принимается равным 0,1–0,2, а $|\lambda|_{\max}$ – максимальный модуль собственных чисел матрицы динамики.

♦ **Пример 8.5.** Получим дискретный вариант задачи фильтрации экспоненциально-коррелированного процесса, описываемого уравнением

$$\dot{x}(t) = -\alpha x(t) + \sqrt{2\sigma_x^2 \alpha} w(t),$$

по измерениям

$$y(t) = x(t) + v(t),$$

где σ_x^2 – дисперсия процесса; $\tau_k = 1/\alpha$ – интервал корреляции; $w(t)$ – белый шум единичной интенсивности; $v(t)$ – независимый от $w(t)$ и $x(0)$ белый шум измерения с интенсивностью r^2 .

Во-первых, заметим, что здесь можно непосредственно воспользоваться выражениями (8.41), не прибегая к его приближенному представлению. Действительно, поскольку $e^{F\tau} = e^{-\alpha\tau}$, $G = \sqrt{2\sigma_x^2 \alpha}$, $Q = 1$, можем записать

$$Q_d^* = 2\sigma_x^2 \alpha \int_0^{\Delta t} e^{-2\alpha\tau} d\tau = \sigma_x^2 (1 - e^{-2\alpha\Delta t}).$$

Последнее выражение есть следствие равенства

$$\int_0^{\Delta t} e^{-2\alpha\tau} d\tau = \frac{1}{2\alpha} \int_{-2\alpha\Delta t}^0 e^x d\tau = \frac{1}{2\alpha} (1 - e^{-2\alpha\Delta t}).$$

Таким образом, принимая $Q_d = \frac{1}{\Delta t}$ и используя (8.40), т.е. $\Gamma Q_d \Gamma^T = Q_d^*$, получаем следующее выражение для Γ :

$$\Gamma = \sqrt{\sigma_x^2 (1 - e^{-2\alpha\Delta t}) \Delta t}.$$

Полагая $e^{-2\alpha\Delta t} \approx 1 - 2\alpha\Delta t$, можем также записать приближенное выражение

$$\Gamma = \left(\sqrt{2\sigma_x^2 \alpha} \right) \Delta t,$$

очевидно совпадающее с (8.45).

Таким образом, дискретную задачу можно сформулировать так: оценить последовательность

$$x_i = e^{-\alpha\Delta t} x_{i-1} + \sqrt{\sigma_x^2 (1 - e^{-2\alpha\Delta t}) \Delta t} w_i, \quad (8.53)$$

по измерениям

$$y_i = x + v_i,$$

где $y_i = y(t_i) = \int_{t_i}^{t_i+\Delta t} y(\tau) d\tau$, $v_i = v(t_i) = \int_{t_i}^{t_i+\Delta t} v(\tau) d\tau$, полагая, что w_i , v_i есть дискретные независимые между собой белозумные последовательности с дисперсиями $Q_d = \frac{1}{\Delta t}$ и $r_d^2 = \frac{r^2}{\Delta t}$.

В первом приближении вместо (8.53) можно использовать

$$x_i = e^{-\alpha\Delta t} x_{i-1} + \left(\sqrt{2\sigma_x^2\alpha}\right) \Delta t w_i,$$

либо

$$x_i = x_{i-1} - \alpha\Delta t x_{i-1} + \left(\sqrt{2\sigma_x^2\alpha}\right) \Delta t w_i,$$

при $Q_d = \frac{1}{\Delta t}$.

Заметим, что, если принять дисперсию порождающего шума единичной, т.е. $Q_d = 1$, то тогда коэффициент при порождающем шуме в последних двух уравнениях должен определяться как $\sqrt{2\sigma_x^2\alpha\Delta t}$.

С учетом соотношения (8.52) и того факта, что $\lambda = \tau_k$, при выборе интервала дискретизации рекомендуется учитывать условие $\frac{\Delta t}{\tau_k} \leq 0.1$. ♦

Задачи к разделу

Задача 8.1. Покажите, что математическое ожидание, матрица ковариаций и корреляционная функция процесса, задаваемого формирующим фильтром (8.8), будут определяться соотношениями (8.10)–(8.12).

Задача 8.2. Пусть $w(t)$ представляет собой центрированный белый шум с интенсивностью Q . Введем случайную последовательность, определяемую как

$$w_i = \frac{1}{\Delta t} \int_{t_i}^{t_i+\Delta t} w(t) dt,$$

где Δt – интервал дискретизации.

Найдите корреляционную функцию для этой последовательности.

Задача 8.3. Располагая уравнениями для оцениваемого процесса

$$\dot{x}(t) = F(t)x(t) + G(t)w(t)$$

и уравнениями для оценки

$$\dot{\hat{x}}(t) = F(t)\hat{x}(t) + K(t)(y(t) - H(t)\hat{x}(t)),$$

запишите уравнения для ошибок фильтрации. Является ли процесс, описывающий ошибки фильтрации, стационарным и марковским?

Задача 8.4. Запишите уравнения для оценки и апостериорной дисперсии в фильтре Калмана в задаче фильтрации винеровского процесса, описываемого уравнением $\dot{x}(t) = \sqrt{2\sigma_x^2} \alpha w(t)$, при использовании измерений $y(t) = x(t) + v(t)$, в которых $v(t)$ – независимый от $w(t)$ и $x(0)$ белый шум измерения с интенсивностью r^2 . Получите решения этих уравнений.

Задача 8.5. Запишите уравнения для оценки и апостериорной дисперсии в фильтре Калмана для задачи фильтрации марковского процесса (8.21) по его измерениям $y(t) = x(t) + v(t)$, в которых $v(t)$ – независимый от $w(t)$ и $x(0)$ белый шум измерения с интенсивностью r^2 . Получите решения этих уравнений.

Задача 8.6. Получите выражения для апостериорной дисперсии для непрерывной задачи оценивания центрированной случайной величины с дисперсией σ_0^2 по измерениям $y(t) = x + v(t)$, в которых $v(t)$ – белый шум интенсивности r^2 , независимый от x . Получите аналогичное выражение при переходе от непрерывной задачи к ее дискретному аналогу и сопоставьте эти выражения.

Задача 8.7. Запишите алгоритм ФКБ в задаче оценивания координат и скорости объекта, которые удовлетворяют уравнению

$$\dot{X} = V;$$

$$\dot{V} = 0,$$

при наличии измерений только координат, т.е.

$$y = X + v,$$

где $v(t)$ – центрированный некоррелированный с X белый шум интенсивности r^2 .

Получите выражения для дисперсии ошибки оценивания координат и скорости.

Примечание. Для получения решения для матрицы ковариаций рекомендуется перейти к уравнению для обратной матрицы – см. замечание 4 в конце подраздела 8.3

Задача 8.8. Покажите, что рассмотренная в задаче 8.7 задача может быть сформулирована как задача оценивания двух постоянных величин. Получите выражение для матрицы ковариаций для такой постановки.

Контрольные вопросы

1. Дайте определение математического ожидания, дисперсии и корреляционной функции для случайного процесса.

2. В чем отличие случайной последовательности от случайного процесса? Перечислите их общие характеристики.
3. Какой процесс является стационарным? Что такое спектральная плотность процесса? Как найти дисперсию процесса, зная его спектральную плотность?
4. Дайте определения формирующего фильтра для случайного процесса. Приведите соотношения для математического ожидания и матрицы ковариаций для случайного процесса, задаваемого формирующим фильтром.
5. Сформулируйте постановку задачи оптимальной линейной фильтрации случайного процесса.
6. Приведите и прокомментируйте соотношения для непрерывного ФКБ. Сопоставьте блок-схемы формирующего фильтра и ФК.
7. Поясните процедуру перехода от непрерывной задачи фильтрации к ее дискретному аналогу.

ПРИЛОЖЕНИЯ

Приложение 1

П1. Основные матричные операции

В этом подразделе приведены основные понятия из теории матричного исчисления. При его формировании использованы материалы книг [3, 9].

Матрицей называется таблица, содержащая в общем случае n строк и m столбцов, элементами которой могут быть числа, символы и выражения различного вида. Матрицу A можно представить как

$$A = \begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & a_{1m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i1} & \cdot & a_{ij} & \cdot & a_{im} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \cdot & \cdot & \cdot & a_{nm} \end{bmatrix}. \quad (\text{П1.1})$$

Для матрицы могут быть использованы также следующие обозначения:

$$A = \{a_{ij}\}; A = \{A(i, j)\}; A = \{A[i, j]\}, i = \overline{1, n}, j = \overline{1, m}.$$

Вектор

$$a = \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ a_n \end{bmatrix} \quad (\text{П1.2})$$

и строка

$$a = [a_1, \dots, a_n] \quad (\text{П1.3})$$

представляют собой частные случаи матриц размерности $n \times 1$ и $1 \times n$ соответственно.

Квадратная матрица. Квадратной называется матрица, у которой число строк и столбцов одинаково, т.е. матрица размерности $n \times n$.

Диагональная матрица. Квадратная матрица A размерности n называется диагональной, если все ее элементы, кроме тех, которые стоят на диагонали, равны нулю.

$$D = \begin{bmatrix} d_{11} & 0 & \dots & 0 & 0 \\ 0 & d_{22} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{n-1n-1} & 0 \\ 0 & 0 & \dots & 0 & d_{nn} \end{bmatrix}. \quad (\text{П1.4})$$

Для диагональной матрицы нередко используется следующее обозначение: $D = \text{diag} \{d_{ii}\}$, $i = \overline{1..n}$.

Единичная матрица. Единичной матрицей называется диагональная матрица с единицами на диагонали. Обычно для этой матрицы используется обозначение E , т.е.

$$E = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}. \quad (\text{П1.5})$$

Транспонированная матрица. Пусть задана матрица $A = \{a_{ij}\}$, $i = \overline{1..n}$, $j = \overline{1..m}$ размерности $n \times m$. Транспонированной матрицей A^T называется такая матрица размерности $m \times n$, которая определяется как

$$A^T = \{a_{ji}\}, \quad j = \overline{1..m}, \quad i = \overline{1..n}. \quad (\text{П1.6})$$

Очевидно, что строки такой матрицы являются столбцами матрицы A , а строки A – столбцами A^T .

Верхней треугольной матрицей называется матрица, у которой отличны от нуля элементы, стоящие на диагонали и выше:

$$A = \begin{bmatrix} * & * & \dots & * & * \\ 0 & * & \dots & * & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & * & * \\ 0 & 0 & \dots & 0 & * \end{bmatrix}. \quad (\text{П1.7})$$

У **нижней треугольной матрицы** от нуля отличны элементы, стоящие на диагонали и ниже.

Симметричная (симметрическая) матрица. Квадратная матрица размерности n называется симметричной, если

$$A = A^T. \quad (\text{П1.8})$$

Блочная матрица. Матрица, имеющая вид

$$A = \begin{bmatrix} A_{11} & \cdot & \cdot & \cdot & \cdot & A_{1m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{i1} & \cdot & A_{ij} & \cdot & \cdot & A_{im} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{n1} & \cdot & \cdot & \cdot & \cdot & A_{nm} \end{bmatrix}, \quad (\text{П1.9})$$

где A_{ij} , $i = \overline{1, n}$, $j = \overline{1, m}$ представляют собой матрицы, т.е. матрица, составленная из блоков, называется блочной матрицей.

Блочно-диагональная матрица – такая, у которой отличны от нуля только блоки, стоящие на диагонали.

След матрицы. Следом квадратной матрицы A называется величина, равная сумме диагональных элементов, т.е.

$$\text{Tr}A = \sum_{i=1}^n a_{ii}. \quad (\text{П1.10})$$

Такое обозначение происходит от английского слова **trace**, одно из значений которого – след.

Иногда эту величину обозначают как

$$\text{Sp}A = \sum_{i=1}^n a_{ii}. \quad (\text{П1.11})$$

Это обозначение соответствует немецкому варианту этого слова, которое произносится как «шпур».

Сложение матриц. Пусть заданы три матрицы A, B, C размерности $n \times m$. Элементы матрицы C , равной сумме двух матриц A и B

$$C = A + B, \quad (\text{П1.12})$$

определяются с помощью следующего выражения

$$c_{ij} = a_{ij} + b_{ij}, \quad i = \overline{1, n}, \quad j = \overline{1, m}. \quad (\text{П1.13})$$

Умножение матрицы на число. При умножении матрицы на число на это число умножается каждый элемент матрицы, т.е.

$$\alpha A = \{\alpha a_{ij}\}. \quad (\text{П1.14})$$

Умножение матриц. Пусть заданы матрицы A и B размерности $n \times m$ и $m \times l$ соответственно. Элементы $n \times l$ матрицы C , равной произведению двух матриц A и B ,

$$C = AB, \quad (\text{П1.15})$$

определяются как $c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}$, $i = \overline{1, n}$, $j = \overline{1, l}$.

Из приведенного определения следует, что операция умножения корректна только в случае, когда число столбцов матрицы, стоящей в произведении на первом месте, равно числу строк матрицы, стоящей на втором месте, т.е. когда

$$C = AB, \quad (\text{П1.16})$$

и имеет следующую структуру:

$$\begin{array}{c}
 \underbrace{\left[\begin{array}{ccc} c_{11} & \cdot & c_{1m} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ c_{n1} & \cdot & c_{nm} \end{array} \right]}_m \\
 \left. \begin{array}{c} n \\ \\ \\ \\ \\ \end{array} \right\} = n \left\{ \begin{array}{c} \underbrace{\left[\begin{array}{ccc} a_{11} & \cdot & a_{1l} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{n1} & \cdot & a_{nl} \end{array} \right]}_l \\ \\ \\ \\ \\ \end{array} \right. \left\{ \begin{array}{c} \underbrace{\left[\begin{array}{ccc} b_{11} & \cdot & b_{1m} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ b_{l1} & \cdot & b_{lm} \end{array} \right]}_m \\ \\ \\ \\ \\ \end{array} \right. \cdot
 \end{array}$$

Будем говорить, что в этом случае **размерности матриц согласованы**. В результате умножения $1 \times n$ строки a^T на $n \times 1$ столбец b получаем скаляр, в то время как умножение столбца b на строку a дает квадратную матрицу, т.е.

$$ba^T = \begin{bmatrix} b_1 a_1 & \cdot & b_1 a_n \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ b_n a_1 & \cdot & b_n a_n \end{bmatrix}. \quad (\text{П1.17})$$

Величину

$$\|a\| = \sqrt{\sum_{i=1}^n a_i^2} = \sqrt{a^T a} \quad (\text{П1.18})$$

будем называть **нормой вектора** a или его **модулем**. С другими определениями нормы вектора, а также нормы матрицы можно ознакомиться в работе [3].

В общем случае операция умножения матриц некоммукативна, т.е.

$$AB \neq BA.$$

Если операция перестановки справедлива, то матрицы называют **коммукативными**.

Из сказанного с очевидностью следует, что $a^T a = Sp(aa^T)$.

При выполнении операций **умножения блочных матриц** можно использовать правила для обычных матриц, т.е.

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Здесь важно иметь в виду, что при выполнении этих операций размерности перемножаемых блоков должны быть согласованы.

Определитель или **детерминант** квадратной матрицы $A = \{a_{ij}\}$ размерности n может быть вычислен с использованием определителей для матриц размерности $n - 1$ с помощью следующего выражения:

$$\det(A) = \sum_{j=1}^n a_{1j}A_{1j}, \quad (\text{П1.19})$$

в котором A_{1j} – **алгебраические дополнения**, т.е. определители для матриц порядка $n - 1$, получаемых вычеркиванием 1-й строки и j -го столбца у матрицы A , умноженные на $(-1)^{1+j}$.

Аналогичное представление может быть получено для произвольной i -й строки

$$\det(A) = \sum_{j=1}^n a_{ij}A_{ij},$$

в котором A_{ij} – определители порядка $n - 1$, получаемые вычеркиванием i -й строки и j -го столбца и умножением на $(-1)^{i+j}$.

К примеру, для двухмерной матрицы легко получить

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Для определителя также используется обозначение

$$\det(A) = |A|.$$

Квадратная матрица называется **невырожденной** или **несингулярной**, если ее определитель отличен от нуля, т.е.

$$\det(A) = |A| \neq 0.$$

В противном случае матрица называется **вырожденной** или **сингулярной**.

Введем сокращенное обозначение для определителей, составленных из элементов прямоугольной $n \times m$ матрицы

$$A = \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{vmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \dots & a_{i_1 k_p} \\ a_{i_2 k_1} & a_{i_2 k_2} & \dots & a_{i_2 k_p} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i_p k_1} & a_{i_p k_2} & \dots & a_{i_p k_p} \end{vmatrix}.$$

Такой определитель называется **минором** p -го порядка, если $1 \leq i_1 < i_2 < \dots < i_p \leq n$ и $1 \leq k_1 < k_2 < \dots < k_p \leq m$. Миноры, у которых $i_1 = k_1, i_2 = k_2, \dots, i_p = k_p$, называются **главными**.

Наибольший из порядков, отличных от нуля миноров, порождаемых матрицей, называется **рангом матрицы**. Если r – ранг прямоугольной $n \times m$ матрицы, то очевидно, что $r < n, m$.

Обратная матрица. Пусть задана невырожденная $n \times n$ матрица $A = \{a_{ij}\}$. Обратной матрицей A^{-1} называется такая, для которой справедливо следующее соотношение:

$$AA^{-1} = E. \quad (\text{П1.20})$$

Обозначим обратную матрицу как $A^{-1} = B = \{b_{ij}\}$.

Для элементов обратной матрицы справедливо соотношение [3, с.345]

$$b_{ij} = \frac{A_{ji}}{\det(A)}, \quad \det(A) = a_{11}A_{11} + \dots + a_{1n}A_{1n}, \quad (\text{П1.21})$$

в котором A_{ji} – алгебраические дополнения.

К примеру, для двухмерной матрицы легко получить

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Заметим, что для обратной матрицы справедливо равенство $A^{-1}A = E$, т.е. прямая и обратная матрицы коммутативны.

Ортогональная матрица. Квадратная матрица A называется ортогональной при условии, что

$$AA^T = E. \quad (\text{П1.22})$$

Характеристическим многочленом квадратной матрицы A называется многочлен

$$p(\lambda) = \det(A - \lambda E). \quad (\text{П1.23})$$

К примеру, для двухмерной матрицы этот многочлен имеет вид

$$p(\lambda) = \det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}.$$

Характеристическое уравнение определяется как

$$p(\lambda) = \det(A - \lambda E) = 0.$$

Собственные значения матрицы. Собственными значениями квадратной матрицы A размерности n называются n корней ее характеристического уравнения

$$p(\lambda) = \det(A - \lambda E) = 0. \quad (\text{П1.24})$$

Совокупность всех n собственных чисел называется **спектром матрицы**.

Собственным вектором квадратной матрицы A называется такой вектор x , для которого выполняется следующее равенство:

$$Ax = \lambda x, \quad (\text{П1.25})$$

т.е. $(A - \lambda E)x = 0$, где λ – собственное число.

Это равенство означает, что умножение матрицы на такой вектор не меняет его направления. Поскольку характеристическое уравнение матрицы размерности n имеет n корней, соответственно для каждого корня (собственного числа) будет свой собственный вектор. Таким образом, матрица имеет n собственных чисел и собственных векторов. Существенно, что некоторые из собственных чисел могут между собой совпадать, а одному и тому же собственному значению может соответствовать несколько различных собственных векторов. Задача их нахождения известна как **проблема собственных чисел**.

Преобразование подобия. Пусть заданы две квадратные $n \times n$ матрицы A и C , причем матрица C не вырождена. Определим квадратную матрицу B как

$$B = CAC^{-1}. \quad (\text{П1.26})$$

Матрицы B и A называются **подобными матрицами**, а матрица C **преобразованием подобия**. Нетрудно убедиться в том, что подобные матрицы имеют одинаковые характеристические многочлены. Действительно,

$$\begin{aligned} \det(B - \lambda E) &= \det(CAC^{-1} - \lambda CC^{-1}) = \det(A - \lambda E) \det(C) \det(C^{-1}) = \\ &= \det(A - \lambda E) \det(CC^{-1}) = \det(A - \lambda E). \end{aligned}$$

Отсюда вытекает, что подобные матрицы имеют одинаковые собственные числа.

Легко убедиться в том, что определитель и след квадратной матрицы A могут быть представлены с использованием собственных чисел, т.е.

$$\text{Tr}A = \sum_{i=1}^n \lambda_i, \quad (\text{П1.27})$$

$$\det(A) = \prod_{i=1}^n \lambda_i. \quad (\text{П1.28})$$

Отсюда следует, что определители и следы подобных матриц одинаковы.

Диагонализация симметричных матриц. Симметричная матрица A всегда с помощью ортогонального преобразования подобия может быть приведена к диагональному виду, т.е. всегда можно найти такую ортогональную матрицу

$$T^T T = E, \quad T^T = [t_1 \quad \dots \quad t_j \quad \dots \quad t_n], \quad (\text{П1.29})$$

для которой

$$TAT^T = \{\lambda_j \delta_{ij}\}, \quad j = \overline{1, n}, \quad (\text{П1.30})$$

где $\lambda_j, t_j, j = \overline{1, n}$ – собственные числа и собственные векторы матрицы A , т.е.

$$A t_j = \lambda_j t_j, \quad (\text{П1.31})$$

причем $t_i^T t_j = \delta_{ij}$. Здесь величина δ_{ij} представляет собой **символ Кронекера**

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (\text{П1.32})$$

Последнее условие означает дополнительное требование, накладываемое на значение модулей собственных векторов, т.е. их модули должны быть равны единице.

Важно заметить, что возможность преобразовать матрицу к диагональному виду в общем случае, когда матрица не является симметричной, существует не всегда. Однако произвольную квадратную матрицу можно преобразовать к другим матрицам специального (говорят еще – нормального или канонического) вида, таким, например, как матрица Жордана, матрица Фробениуса и т.д. [9].

Квадратичная форма. Пусть заданы квадратная матрица A и вектор x размерности n . Квадратичной формой называется выражение

$$y = x^T A x. \quad (\text{П1.33})$$

Квадратичная форма называется **положительно определенной** в случае, если при любых отличных от нуля значениях вектора x величина квадратичной формы положительна, т.е.

$$y = x^T A x > 0, \text{ при } x \neq 0. \quad (\text{П1.34})$$

Матрицу A в этом случае называют **положительно определенной**.

Если знак строгого неравенства заменить на \geq , то тогда квадратичная форма и соответствующая ей матрица будут называться **неотрицательно определенными**.

Если знак в неравенстве заменить на противоположный, то получим **отрицательно определенную** квадратичную форму и матрицу.

Матричные неравенства. Говорят, что квадратная матрица A больше другой квадратной матрицы B при условии, если неравенства

$$x^T A x > x^T B x > 0 \quad (\text{П1.35})$$

или $x^T (A - B)x > 0$ справедливы при отличных от нуля значениях вектора x . Аналогично можно ввести и другие типы неравенств.

Можно показать, что в случае, когда выполняется какое-то неравенство для матриц, то такое же неравенство будет справедливо и для их собственных чисел. Так, если все собственные числа квадратной матрицы A положительны или неотрицательны (отрицательны или не положительны), т.е. $\lambda_j > 0$ или $\lambda_j \geq 0$ ($\lambda_j < 0$ или $\lambda_j \leq 0$), $j = \overline{1, n}$, то соответствующие нера-

венства будут справедливы и для самой матрицы, т.е. $A > 0$ или $A \geq 0$ ($A < 0$ или $A \leq 0$).

Функции от матриц. Некоторые функции от квадратных матриц могут быть введены с помощью степенных рядов. Матрица A^m степени m определяется как

$$A^m = \underbrace{A \cdot A \dots \cdot A}_m. \quad (\text{П1.36})$$

Используя это определение, можно ввести, например, **матричную экспоненту**, задаваемую в виде

$$e^A = \exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!}. \quad (\text{П1.37})$$

Функции $f(A)$ от симметричных матриц могут быть введены и иначе. Для этого сначала матрицу A с помощью ортогонального преобразования сводят к диагональному виду, т.е. получают представление

$$TAT^T = \Lambda = \text{diag}\{\lambda_j\}, \quad j = \overline{1, n},$$

из которого следует, что $A = T^T \Lambda T$.

Далее $f(A)$ определяют как

$$f(A) = T^T f(\Lambda) T, \quad (\text{П1.38})$$

где $f(\Lambda) = \text{diag}\{f(\lambda_j)\}$, $j = \overline{1, n}$.

Можно показать, что введенное таким образом определение будет совпадать с определением с помощью разложения в ряд [9].

Производная и интеграл матрицы, элементы которой зависят, например, от времени, т.е. $A(t) = \{a_{ij}(t)\}$, $i = \overline{1, n}$, $j = \overline{1, m}$, представляют собой матрицу

$$\dot{A}(t) = \{\dot{a}_{ij}(t)\}, \quad i = \overline{1, n}, \quad j = \overline{1, m} \quad \text{либо}$$

$$\int_0^t A(\tau) d\tau = \left\{ \int_0^t a_{ij}(\tau) d\tau \right\}, \quad i = \overline{1, n}, \quad j = \overline{1, m} \quad \int_0^t A(\tau) d\tau = \left\{ \int_0^t a_{ij}(\tau) d\tau \right\}, \quad i = \overline{1, n}, \quad j = \overline{1, m}$$

с элементами в виде производной или интеграла от элементов исходной матрицы.

Можно также ввести **производную скалярной функции $s(x)$ по векторному аргументу**

$$\frac{ds(x)}{dx^T} = \left[\frac{\partial s(x)}{\partial x_1} \dots \frac{\partial s(x)}{\partial x_n} \right]. \quad (\text{П1.39})$$

Запись $\frac{ds(x)}{dx}$ будет означать

$$\frac{ds(x)}{dx} = \begin{bmatrix} \frac{\partial s(x)}{\partial x_1} \\ \dots \\ \frac{\partial s(x)}{\partial x_n} \end{bmatrix}, \quad (\text{П1.40})$$

таким образом,

$$\frac{ds(x)}{dx} = \left[\frac{ds(x)}{dx^T} \right]^T. \quad (\text{П1.41})$$

Отсюда следует, что

$$\frac{d^2s(x)}{dx dx^T} = \frac{d}{dx} \left[\frac{ds(x)}{dx^T} \right] = \begin{bmatrix} \frac{\partial^2 s(x)}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 s(x)}{\partial x_1 \partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial^2 s(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 s(x)}{\partial x_n \partial x_n} \end{bmatrix}, \quad \frac{d^2s(x)}{dx dx^T} = \left[\frac{d^2s(x)}{dx^T dx} \right]^T. \quad (\text{П1.42})$$

Вектор-функцию векторного аргумента и ее производную по векторному аргументу $x = (x_1, \dots, x_n)^T$ определим как

$$s(x) = \begin{bmatrix} s_1(x_1, \dots, x_n) \\ \vdots \\ s_m(x_1, \dots, x_n) \end{bmatrix}. \quad (\text{П1.43})$$

и

$$\frac{ds(x)}{dx^T} = \begin{bmatrix} \frac{\partial s_1(x)}{\partial x_1} & \dots & \frac{\partial s_1(x)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial s_m(x)}{\partial x_1} & \dots & \frac{\partial s_m(x)}{\partial x_n} \end{bmatrix}, \quad \text{а} \quad \frac{ds^T(x)}{dx} = \begin{bmatrix} \frac{\partial s_1(x)}{\partial x_1} & \dots & \frac{\partial s_m(x)}{\partial x_1} \\ \dots & \dots & \dots \\ \frac{\partial s_1(x)}{\partial x_n} & \dots & \frac{\partial s_m(x)}{\partial x_n} \end{bmatrix}, \quad (\text{П1.44})$$

т.е. $\frac{ds^T(x)}{dx} = \left[\frac{ds(x)}{dx^T} \right]^T.$

В общем случае, если A – квадратная симметричная матрица, то $\frac{d}{dx} [g^T(x) A g(x)] = 2 \frac{dg^T(x)}{dx} A g(x).$

Производную скалярной функции $s(A)$ по матрице (матричный градиент) введем как

$$\frac{ds(A)}{dA} = \left\{ \frac{ds}{da_{ij}} \right\}.$$

К примеру, для квадратной матрицы A справедливо [32, с. 23]:

$$\frac{ds(Sp(A))}{dA} = E, \quad \frac{ds(Sp(BAC))}{dA} = B^T C^T;$$

$$\frac{ds(Sp(ABA^T))}{dA} = A(B + B^T). \quad (\text{П1.45})$$

Формула обращения блочных матриц [18, с. 107]. Пусть

$$P = \begin{bmatrix} P^x & P^{xy} \\ (P^{xy})^T & P^y \end{bmatrix}, \quad (\text{П1.46})$$

где P^x , P^y , P^{xy} – $n \times n$, $m \times m$ и $n \times m$ матрицы, причем для P^x , P^y существуют обратные матрицы. В этом случае обратная матрица P^{-1} определяется как

$$P^{-1} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \quad (\text{П1.47})$$

где

$$A = \left[P^x - P^{xy} (P^y)^{-1} P^{yx} \right]^{-1},$$

$$B = - \left[P^x - P^{xy} (P^y)^{-1} P^{yx} \right]^{-1} P^{xy} (P^y)^{-1},$$

$$C = \left[P^y - P^{yx} (P^x)^{-1} P^{xy} \right]^{-1}.$$

Справедливы также следующие соотношения:

$$A = (P^x)^{-1} + (P^x)^{-1} P^{xy} C P^{yx} (P^x)^{-1},$$

$$B = -A P^{xy} (P^y)^{-1} = -(P^x)^{-1} P^{xy} C,$$

$$C = (P^y)^{-1} + (P^y)^{-1} P^{yx} A P^{xy} (P^y)^{-1}.$$

В табл. П1.1 представлены некоторые полезные матричные соотношения, в частности, так называемая **лемма об обращении матриц** [18, с. 216]

$$\left[P^{-1} + H^T R^{-1} H \right]^{-1} = P - P H^T (H P H^T + R)^{-1} H P, \quad (\text{П1.48})$$

где P , R – квадратные невырожденные матрицы размерности n и m ; H – $m \times n$ матрица.

Из этой леммы следует, что

$$\left[P^{-1} + R^{-1} \right]^{-1} = P - P(P + R)^{-1} P, \quad (\text{П1.49})$$

а также

$$\left[r^2 E + q^2 I \right]^{-1} = \frac{1}{r^2} \left(E - \frac{q^2}{nq^2 + r^2} I \right), \quad (\text{П1.50})$$

Последнее соотношение легко получается, если принять

$$P^{-1} = r^2 E, \quad H = [1, 1 \dots 1], \quad R^{-1} = q^2.$$

Т а б л и ц а П1.1

Некоторые полезные матричные соотношения

Соотношение	Примечание
$(AB)^T = B^T A^T$ (П1.51)	A, B – произвольные матрицы согласованной размерности
$Tr(ABC) = Tr(BCA) = Tr(CAB)$ (П1.52)	A, B, C – произвольные матрицы согласованной размерности
$Tr(A) = \sum_{i=1}^n \lambda_i$ (П1.53)	A – квадратная матрица, λ_i – собственные числа A
$\det(AB) = \det(A) \det(B)$ (П1.54)	A, B – квадратные матрицы одинаковой размерности
$\det(A) = \prod_{i=1}^n \lambda_i$ (П1.55)	A – квадратная матрица, λ_i – собственные числа A
$(AB)^{-1} = B^{-1} A^{-1}$ (П1.56)	A, B – квадратные невырожденные матрицы одинаковой размерности
$\left[P^{-1} + H^T R^{-1} H \right]^{-1} =$ $= P - PH^T (HPH^T + R)^{-1} HP$ (П1.57)	P, R – квадратные невырожденные матрицы размерности n и m , H – $m \times n$ матрица
$\left[P^{-1} + R^{-1} \right]^{-1} = P - P(P + R)^{-1} P$ (П1.58)	P, R – квадратные невырожденные матрицы одинаковой размерности
$\left[r^2 E + q^2 I \right]^{-1} = \frac{1}{r^2} \left(E - \frac{q^2}{nq^2 + r^2} I \right)$ (П1.59)	I – n -мерная квадратная матрица, составленная из единиц; r^2, q^2 – положительные величины
$\frac{\partial}{\partial x^T} (Ax) = A, \quad \frac{\partial}{\partial x} (x^T A^T) = A^T$ (П1.60)	A – квадратная симметричная матрица,

$\frac{\partial}{\partial x}(x^T Ax) = 2Ax,$ $\frac{\partial}{\partial x^T}(x^T Ax) = 2x^T A \quad (\text{П1.61})$	$x = (x_1, \dots, x_n)^T$ – n -мерный вектор
$\frac{\partial^2}{\partial x \partial x^T}(x^T Ax) = 2A \quad (\text{П1.62})$	
$\frac{\partial}{\partial x}[g^T(x)Ag(x)] = 2\frac{\partial g^T(x)}{\partial x}Ag(x),$ <p>в частности,</p> $\frac{\partial}{\partial x}[(y - Hx)^T A(y - Hx)] =$ $= 2H^T A(y - Hx) \quad (\text{П1.63})$	A – квадратная симметричная матрица, $g(x) = (g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n))^T$ – m -мерная вектор-функция, x, y – n - и m -мерные векторы, H – $m \times n$ -матрица
$x^T Ax - 2x^T z =$ $= (x - A^{-1}z)A(x - A^{-1}z) - z^T A^{-1}z \quad (\text{П1.64})$	A – квадратная невырожденная симметричная матрица, x, z – n -мерные векторы
<p>Задана $P = \begin{bmatrix} P^x & P^{xy} \\ (P^{xy})^T & P^y \end{bmatrix}$</p> <p>Найти $P^{-1} = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad (\text{П1.65})$</p>	$C = \left[P^y - P^{yx} (P^x)^{-1} P^{xy} \right]^{-1}$ $B = - \left[P^x - P^{xy} (P^y)^{-1} P^{yx} \right]^{-1} P^{xy} (P^y)^{-1}$ $A = \left[P^x - P^{xy} (P^y)^{-1} P^{yx} \right]^{-1}$

Приложение 2

П2. Случайные величины и векторы

П2.1. Случайная величина и ее описание

Случайной называется такая величина, значение которой заранее неизвестно, и можно лишь указать некую числовую меру (вероятность) того, что она будет принадлежать той или иной заранее определенной области значений [13, 14, 16, 25, 32].

Случайную величину будем считать заданной, если определена функция, позволяющая вычислять вероятность того, что случайная величина будет принадлежать тому или иному интервалу или их произвольному набору. Такой функцией является **функция распределения вероятностей**, представляющая собой скалярную функцию $F_x(x)$ и определяющую вероятность того, что случайная величина \mathbf{x} принадлежит открытому интервалу $(-\infty, x)$, т.е. вероятность того, что $\mathbf{x} < x$. Таким образом,

$$F_x(x) = \mathbf{Pr}(\mathbf{x} : \mathbf{x} < x). \quad (\text{П2.1})$$

Функция (П2.1) является неотрицательной, неубывающей, непрерывной слева функцией, удовлетворяющей условиям:

$$F_x(-\infty) = \Pr(\mathbf{x} : \mathbf{x} < -\infty) = 0; \quad (\text{П2.2})$$

$$F_x(\infty) = \Pr(\mathbf{x} : \mathbf{x} < \infty) = 1. \quad (\text{П2.3})$$

Помимо функции распределения вероятности, для описания свойств случайных величин используют также **функцию плотности распределения вероятности (ф.п.р.в.)**, определяемую как

$$f_x(x) = \frac{dF_x(x)}{dx}. \quad (\text{П2.4})$$

Ф.п.р.в. вероятностей обладает следующими свойствами: эта функция является неотрицательной, $f_x(x) \geq 0$);

$$F_x(x) = \int_{-\infty}^x f_x(u) du, \quad (\text{П2.5})$$

$$\int_{-\infty}^{\infty} f_x(u) du = 1, \quad (\text{П2.6})$$

$$\Pr(x_1 \leq \mathbf{x} < x_2) = \Pr(\mathbf{x} : \mathbf{x} < x_2) - \Pr(\mathbf{x} : \mathbf{x} < x_1) = F_x(x_2) - F_x(x_1) \quad (\text{П2.7})$$

или
$$\Pr(x_1 \leq \mathbf{x} < x_2) = F_x(x_2) - F_x(x_1) = \int_{x_1}^{x_2} f_x(u) du. \quad (\text{П2.8})$$

Принимая во внимание (П2.4), можно записать

$$f_x(x) = \lim_{dx \rightarrow 0} \frac{F_x(x+dx) - F_x(x)}{dx} = \lim_{dx \rightarrow 0} \frac{\Pr(x \leq \mathbf{x} < x+dx)}{dx};$$

из этого соотношения следует, что при малых dx

$$\Pr(x \leq \mathbf{x} < x+dx) = F_x(x+dx) - F_x(x) \approx f_x(x)dx. \quad (\text{П2.9})$$

Помимо ф.р.в. и ф.п.р.в., для описания статистических свойств с.в. используется набор ее числовых характеристик. Основными из них являются **математическое ожидание, моменты, дисперсия, среднеквадратическое отклонение (СКО)**; для СКО иногда используют термин **стандартное отклонение** или **стандарт**. Если речь идет об ошибке, то используется термин **среднеквадратическая ошибка**, для которой будем использовать аббревиатуру **с.к.о.** В англоязычной литературе для среднеквадратической ошибки используют термин **Root Mean Square Error (RMSE)**. Связь перечисленных характеристик с ф.п.р.в. определяется соотношениями, представленными в табл. П2.1.

Случайная величина, имеющая нулевое математическое ожидание, называется **центрированной случайной величиной**.

Предположим, что с помощью некоторой в общем случае нелинейной функции $g(\bullet)$ в результате преобразования случайной величины \mathbf{x} с известной ф.п.р.в. $f_x(x)$ сформирована новая с.в. $y = g(\mathbf{x})$. Справедливы следующие соотношения:

$$\bar{y} = M_y \{y\} = M_x \{g(x)\} = \int g(x) f_x(x) dx;$$

$$M_y \{y^n\} = M_x \{g^n(x)\} = \int g^n(x) f_x(x) dx.$$

Т а б л и ц а П2.1

Основные числовые характеристики случайной величины

Характеристика	Определение
Математическое ожидание	$M_x(x) = \bar{x} = \int x f_x(x) dx$ (П2.10)
Момент n -го порядка	$M_x(x^n) = \int x^n f_x(x) dx$ (П2.11)
Центральный момент n -го порядка	$M_x(x - \bar{x})^n = \int (x - \bar{x})^n f_x(x) dx$ (П2.12)
Дисперсия	$D = \int (x - \bar{x})^2 f_x(x) dx$ (П2.13)
	$D = M_x x^2 - \bar{x}^2$ (П2.14)
Среднеквадратическое отклонение (СКО)	$\sigma = \left(\int (x - \bar{x})^2 f_x(x) dx \right)^{1/2}$ (П2.15)

Примечание. В приведенных соотношениях пределы интегрирования считаются бесконечными. В дальнейшем, когда пределы не указываются, они предполагаются бесконечными.

П2.2. Гауссовские случайные величины и их характеристики

Гауссовской или нормальной случайной величиной называется такая, для которой ф.р.в. и ф.п.р.в. имеют вид

$$F_x(x) = \frac{1}{(2\pi)^{1/2} \sigma} \int_{-\infty}^x \exp\left\{-\frac{(t - \bar{x})^2}{2\sigma^2}\right\} dt; \quad (\text{П2.16})$$

$$f_x(x) = \frac{1}{(2\pi)^{1/2} \sigma} \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma^2}\right\}.$$

Эти функции называют соответственно **гауссовским (нормальным) распределением вероятности** и **гауссовской (нормальной) функцией плотности распределения вероятностей**.

Для гауссовской ф.п.р.в. используется следующее обозначение:

$$f_x(x) = \frac{1}{(2\pi)^{1/2} \sigma} \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma^2}\right\} = N(x; \bar{x}, \sigma^2). \quad (\text{П2.17})$$

Вид гауссовских ф.р.в. и ф.п.р.в. и их зависимость от математического ожидания и СКО показан на рис. П2.1.

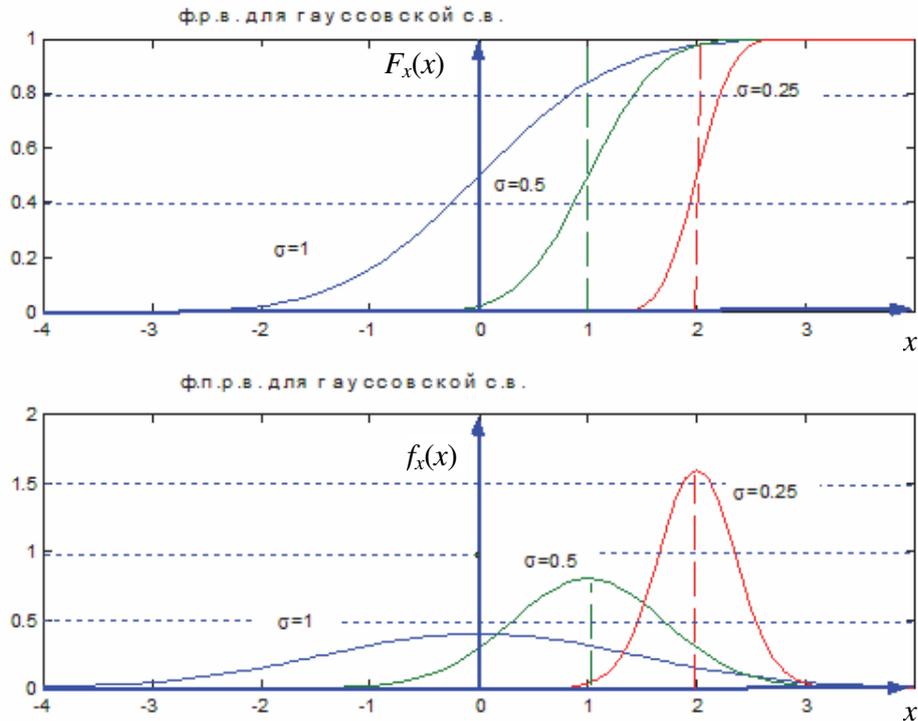


Рис. П2.1. Графики ф.р.в. и ф.п.р.в. гауссовской случайной величины при различных значениях математического ожидания ($\bar{x} = 0, \bar{x} = 1, \bar{x} = 2$) и СКО $\sigma = 1; \sigma = 0,5; \sigma = 0,25$

Нетрудно также заметить, что для гауссовской ф.п.р.в. медиана, математическое ожидание и мода между собой совпадают. Понятно, что такое распределение является унимодальным.

Нечетные центральные моменты гауссовской случайной величины равны нулю, т.е. [14, с.576]

$$\int (x - \bar{x})^{2k-1} f_x(x) dx = 0,$$

а для четных моментов справедливо выражение

$$\int (x - \bar{x})^{2k} f_x(x) dx = 1 \times 3 \times \dots \times (2k - 1) \sigma^{2k}, k = 1, 2, \dots$$

В табл. П2.2 представлены значения вероятности $\Pr[\bar{x} - k\sigma \leq x < \bar{x} + k\sigma] = \Pr[|x - \bar{x}| \leq k\sigma]$ для гауссовской с.в. при различных значениях k .

Т а б л и ц а П 2.2

Значения вероятности $\Pr[|x - \bar{x}| \leq k\sigma]$ для гауссовского распределения при различных k

k	1	2	3	4
$\Pr[x - \bar{x} \leq k\sigma]$	0,6827	0,9545	0,9973	0,9999

Обычно величину, равную 3σ , называют **предельным значением** или **предельной ошибкой**, если с.в. описывает погрешности тех или иных измерений. Тот факт, что для гауссовской случайной величины $\Pr\left[|x - \bar{x}| \leq 3\sigma\right] = 0,997$, называют **правилом трех сигм**.

П2.3. Случайные векторы и их описание

Случайным называется вектор, каждая компонента которого является случайной величиной. Для случайного вектора $\mathbf{x} = (x_1, \dots, x_n)^T$ его свойства в полном объеме задаются **совместной ф.р.в.** или **совместной ф.п.р.в.**, определяемыми в следующем виде:

$$F_{\mathbf{x}}(x) = \Pr(\mathbf{x}_1 < x_1, \dots, \mathbf{x}_n < x_n); \quad (\text{П2.18})$$

$$f_{\mathbf{x}}(x) = \frac{\partial^n F_{\mathbf{x}}(x)}{\partial x_1 \dots \partial x_n}; \quad (\text{П2.19})$$

$$F_{\mathbf{x}}(x) = \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_1} f_{\mathbf{x}}(u) du_1 \dots du_n. \quad (\text{П2.20})$$

Выражение (П2.18) задает вероятность события, при котором для каждой компоненты выполняется неравенство $\mathbf{x}_j < x_j$, $j = \overline{1, n}$.

Совместная ф.п.р.в., так же, как и в одномерном случае, является неотрицательной функцией и удовлетворяет **условию нормировки**

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) dx_1 \dots dx_n = 1. \quad (\text{П2.21})$$

Кроме того, совместная ф.п.р.в. удовлетворяет **условиям согласованности**, которые при $m < n$ записываются как

$$f_{x_1, x_2, \dots, x_m}(x_1, x_2, \dots, x_m) = \int \dots \int f_{\mathbf{x}}(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n) dx_{m+1} \dots dx_n, \quad (\text{П2.22})$$

и является симметричной функцией своих аргументов. Последнее означает, что ф.п.р.в. для вектора $\mathbf{x} = (x_1, \dots, x_n)^T$ не зависит от того, в какой последовательности расположены его компоненты, в частности $f_{x_i, x_j}(x_i, x_j) = f_{x_j, x_i}(x_j, x_i)$.

Вероятность попадания случайного вектора в область Ω , его математическое ожидание \bar{x} и **матрица ковариаций** P , которая является обобщением понятия дисперсии на многомерный случай, имеют вид

$$\Pr(\mathbf{x} \in \Omega) = \int_{\Omega} f_{\mathbf{x}}(x) dx; \quad (\text{П2.23})$$

$$\bar{x} = \int x f_{\mathbf{x}}(x) dx = M_{\mathbf{x}}(x); \quad (\text{П2.24})$$

$$P = \int (x - \bar{x})(x - \bar{x})^T f_{\mathbf{x}}(x) dx = M_{\mathbf{x}}(xx^T) - \bar{x}\bar{x}^T. \quad (\text{П2.25})$$

Здесь и в дальнейшем интегралы понимаются как многократные, причем если область интегрирования не указана, как уже отмечалось выше, то пределы по каждой компоненте предполагаются от $-\infty$ до $+\infty$. Диагональные элементы матрицы ковариаций определяют дисперсии соответствующих компонент случайного вектора.

Математическое ожидание $M_{x_i, x_j} \{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\}$ для двух случайных величин x_i и x_j называется **коэффициентом корреляции**. Таким образом, недиагональные элементы

$P_{ij} = M_{x_i, x_j} \{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\} = \iint (x_i - \bar{x}_i)(x_j - \bar{x}_j) f_{x_i, x_j}(x_i, x_j) dx_i dx_j$, $i \neq j$, $i, j = \overline{1, n}$ определяют коэффициенты корреляции между различными компонентами.

В силу свойства симметричности ф.п.р.в. справедливо равенство $P_{ij} = P_{ji}$, означающее симметричность матрицы ковариаций, т.е. $P = P^T$. Важным свойством матрицы ковариаций является тот факт, что она является неотрицательно определенной матрицей, т.е. такой, для которой $x^T P x \geq 0$ при любом $x \neq 0$.

Если математическое ожидание случайного вектора – нулевое, то, как и в скалярном случае, такой вектор называется **центрированным**.

Если $M_{x_i, x_j} \{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\} = 0$, то случайные величины называются **некоррелированными** или **ортогональными**. Отсюда следует, что для случайного вектора, у которого компоненты не коррелированы между собой, матрица ковариаций имеет диагональный вид. Если заданы два случайных вектора, то можно ввести **матрицу взаимной корреляции**, определяемую как

$$B = \int \int (x - \bar{x})(y - \bar{y})^T f_{x, y}(x, y) dx dy,$$

где $f_{x, y}(x, y)$ – совместная ф.п.р.в. Если эта матрица равна нулю, то говорят о том, что **случайные векторы некоррелированы или ортогональны**.

Важным также является понятие **независимости** с.в. Случайные величины x_i , $i = \overline{1, n}$ называются взаимно **независимыми**, если совместная ф.п.р.в. равна произведению ф.п.р.в. для каждой из этих с.в., т.е.

$$f_x(x_1 \dots x_n) = \prod_{i=1}^n f_{x_i}(x_i).$$

Аналогично вводится определение и для **независимых случайных векторов**.

Независимые случайные величины являются некоррелированными, поскольку

$$\begin{aligned} M_{x_i, x_j} \{(x_i - \bar{x}_i)(x_j - \bar{x}_j)\} &= \iint (x_i - \bar{x}_i)(x_j - \bar{x}_j) f_{x_i, x_j}(x_i, x_j) dx_i dx_j = \\ &= \int (x_i - \bar{x}_i) f_{x_i}(x_i) dx_i \int (x_j - \bar{x}_j) f_{x_j}(x_j) dx_j = 0. \end{aligned}$$

Обратное утверждение в общем случае несправедливо.

Приведенные выше понятия, конкретизированные для двухмерного вектора $\mathbf{x} = (x_1, x_2)^T$, представлены в табл. П2.3.

Т а б л и ц а П2.3

Основные определения и соотношения для двухмерного случайного вектора

Соотношение	Определение
Функция распределения вероятностей	$F_{\mathbf{x}}(x) = \mathbf{Pr}(\mathbf{x}_1 < x_1, \mathbf{x}_2 < x_2);$
Связь ф.п.р.в. и ф.р.в.	$f_{\mathbf{x}}(x) = \frac{\partial^2 F(x)}{\partial x_1 \partial x_2}, F_{\mathbf{x}}(x) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{\mathbf{x}}(u) du_2 du_1$
Условие нормировки	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) dx_1 dx_2 = 1$
Симметричность ф.п.р.в.	$f_{\mathbf{x}}(x_1, x_2) = f_{\mathbf{x}}(x_2, x_1)$
Условия согласованности	$f_{x_1}(x_1) = \int f_{\mathbf{x}}(x_1, x_2) dx_2$ $f_{x_2}(x_2) = \int f_{\mathbf{x}}(x_1, x_2) dx_1$
Независимость	$f_{\mathbf{x}}(x_1, x_2) = f_{x_1}(x_1) f_{x_2}(x_2)$
Некоррелированность	$M_{\mathbf{x}}(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) = 0$
Вероятность попадания случайного вектора в область Ω	$\mathbf{Pr}(\mathbf{x} \in \Omega) = \int_{\Omega} f_{\mathbf{x}}(x) dx_1 dx_2$
Математическое ожидание	$\bar{x}_i = \int \int x_i f_{\mathbf{x}}(x_1, x_2) dx_1 dx_2 = \int x_i f_{x_i}(x_i) dx_i,$ $i = 1, 2$
Коэффициенты корреляции	$P_{12} = P_{21} = \iint (x_1 - \bar{x}_1)(x_2 - \bar{x}_2) f_{\mathbf{x}}(x_1, x_2) dx_1 dx_2$
Дисперсии компонент	$\sigma_i^2 = \iint (x_i - \bar{x}_i)^2 f_{x_i}(x_i) dx_i, i = 1, 2$
Матрица ковариаций	$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, P_{ii} = \sigma_i^2, i = 1, 2, P = P^T \geq 0$

П.2.4. Гауссовские случайные векторы и их характеристики

Гауссовским случайным вектором называется такой, для которого ф.п.р.в. определяется в виде

$$f_{\mathbf{x}}(x) = N(x; \bar{x}, P) = \frac{1}{(2\pi)^{n/2} (\det P)^{1/2}} \exp\{-0.5(x - \bar{x})^T P^{-1}(x - \bar{x})\}. \quad (\text{П2.26})$$

В этом выражении \bar{x} и P представляют собой математическое ожидание и матрицу ковариаций, которые, как и в одномерном случае, полностью определяют гауссовскую ф.п.р.в.

Векторы называются **совместно гауссовскими**, если их совместная плотность распределения – гауссовская. Заметим, что возможны ситуации, при которых каждый вектор или с.в. по отдельности – гауссовские, а их совместная плотность таковой не является.

Если совместно гауссовские векторы \mathbf{x} и \mathbf{y} не коррелированы, т.е. $M_{\mathbf{x},\mathbf{y}}\{(x - \bar{x})(y - \bar{y})^T\} = 0$, то они и независимы между собой и, таким образом, $f_{\mathbf{x},\mathbf{y}}(x, y) = f_{\mathbf{y}}(y)f_{\mathbf{x}}(x)$.

В частности, в этом легко убедиться на примере, когда \mathbf{x} и \mathbf{y} – скаляры, поскольку

$$\begin{aligned} f_{\mathbf{x},\mathbf{y}}(x, y) &= N\left(x, y; \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} P_{11} & 0 \\ 0 & P_{22} \end{bmatrix}\right) = \frac{1}{(2\pi) P_{11}^{1/2} P_{22}^{1/2}} \exp\left\{-\frac{(x - \bar{x})^2}{2P_{11}} - \frac{(y - \bar{y})^2}{2P_{22}}\right\} = \\ &= \frac{1}{(2\pi)^{1/2} P_{11}^{1/2}} \exp\left\{-\frac{(x - \bar{x})^2}{2P_{11}}\right\} \frac{1}{(2\pi)^{1/2} P_{22}^{1/2}} \exp\left\{-\frac{(y - \bar{y})^2}{2P_{22}}\right\}. \end{aligned}$$

Как отмечалось ранее, в общем случае такое утверждение несправедливо.

Проанализируем вид двумерной гауссовской ф.п.р.в. Предположим, что матрица ковариаций – недиагональная, т.е.

$$P = \begin{bmatrix} \sigma_1^2 & r^* \\ r^* & \sigma_2^2 \end{bmatrix}. \quad (\text{П2.27})$$

Вводя **нормированный коэффициент корреляции** в виде

$$r = \frac{r^*}{\sigma_1 \sigma_2}, \quad (\text{П2.28})$$

нетрудно убедиться в том, что

$$P^{-1} = \frac{1}{(1 - r^2)} \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{r}{\sigma_1 \sigma_2} \\ -\frac{r}{\sigma_1 \sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}. \quad (\text{П2.29})$$

Таким образом, ф.п.р.в. для двумерного гауссовского вектора может быть записана как

$$N(x; 0, P) = \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - r^2}} \exp\left\{-\frac{1}{2(1 - r^2)} \left(\frac{x_1^2}{\sigma_1^2} - \frac{2rx_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right\}. \quad (\text{П2.30})$$

Уравнение

$$g(x_1, x_2) = \frac{x_1^2}{\sigma_1^2} - \frac{2rx_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2} = c^2 \quad (\text{П2.31})$$

при разных значениях c задает эллипсы. Но при этом их оси развернуты относительно вертикальной оси на некоторый угол. В качестве примера на рис. П2.3 изображены изолинии при $r = 0,75$.

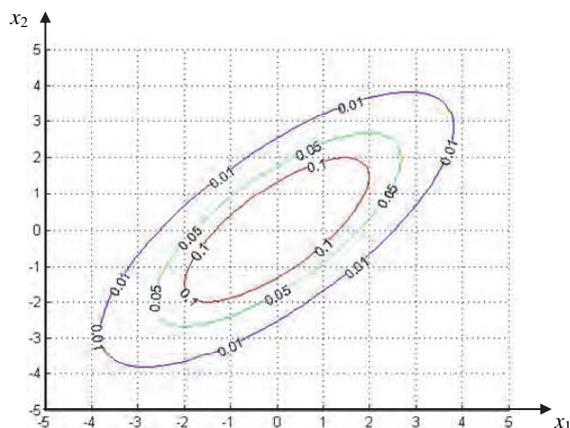


Рис. П2.3. Изолинии ф.п.р.в. для центрированного гауссовского вектора при $\sigma_1 = \sigma_2 = 1$

При решении навигационных задач на плоскости нередко полагают, что координаты объекта представляют собой гауссовский случайный вектор с математическим ожиданием в точке его предполагаемого местонахождения. Для описания неопределенности расположения точки на плоскости используют введенные выше **эллипсы равных вероятностей**, в частности эллипс, соответствующий уравнению (П2.31) при $c = 1$.

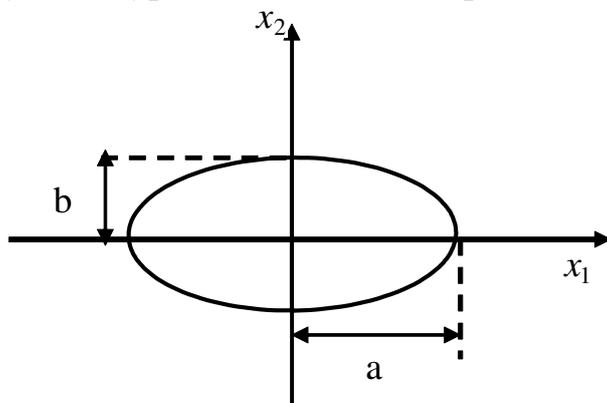


Рис. П2.4. Эллипс ошибок для двумерного гауссовского вектора с независимыми компонентами

Поскольку этот эллипс пересекает оси в точках, совпадающих со значениями соответствующих СКО, т.е. при $x_2 = 0, x_1 = \sigma_1$, а при $x_1 = 0, x_2 = \sigma_2$, он получил наименование **среднеквадратического эллипса ошибок**, или **стандартного эллипса** [10]. В навигационных приложениях для его описания используют **параметры эллипса: большую a и малую b полуоси и дирекционный угол τ** , задающий ориентацию большой полуоси относительно оси x_2 . Эти три параметра полностью определяют

матрицу ковариаций двумерной гауссовской плотности [25]. На рис. П2.4 изображен частный случай, когда $\sigma_2 = b$, $\sigma_1 = a$, $\tau = 90^\circ$, и таким образом

$$P = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}, \quad (\text{П2.32})$$

т.е. размеры полуосей эллипса определяют значения СКО по каждой координате.

При оценивании точности местоположения подвижных объектов весьма важным представляется умение охарактеризовать неопределенность местоположения одним числом. Для этих целей обычно используют значения вероятности попадания точки на плоскости в ту или иную заданную область Ω . Для двумерного центрированного гауссовского вектора с ф.п.р.в. (П2.30) эта вероятность определяется как

$$\Pr(\mathbf{x} \in \Omega) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \iint_{\Omega} \exp\left\{-\frac{1}{2(1-r^2)}g(x_1, x_2)\right\} dx_1 dx_2, \quad (\text{П2.33})$$

где $g(x_1, x_2)$ – эллипс равных вероятностей (П2.31).

Если в качестве Ω выступает область, ограниченная $g(x_1, x_2)$, то, переходя к полярным координатам, можно показать, что [16, с. 68]

$$\Pr(\mathbf{x}: g(x_1, x_2) \leq c^2) = 1 - \exp\left\{-\frac{c^2}{2(1-r^2)}\right\}. \quad (\text{П2.34})$$

Для случая независимых с.в. при $\sigma_1 = \sigma_2 = \sigma$ эллипс превращается в окружность радиусом $R = c\sigma$ и, таким образом, из (П2.34) получаем, что вероятность нахождения случайного вектора в круге с таким радиусом определяется распределением Рэлея

$$\Pr\left(\mathbf{x}: \frac{\sqrt{x_1^2 + x_2^2}}{\sigma} \leq \frac{R}{\sigma}\right) = \Pr(\mathbf{x}: \sqrt{x_1^2 + x_2^2} \leq R) = F(R) = 1 - \exp\left\{-\frac{R^2}{2\sigma^2}\right\}, \quad R > 0. \quad (\text{П2.35})$$

Величина R , соответствующая 50-процентному попаданию гауссовского случайного вектора в круг заданного радиуса, т.е. когда вероятность попадания равна 0,5, называется **круговой вероятной ошибкой (КВО)**, а круг, соответственно, **кругом равных вероятностей**. В случае когда эллипс представляет собой окружность, т.е. при независимых с.в. и равных СКО $\sigma_1 = \sigma_2 = \sigma$, 50-процентное попадание в круг ($\Pr=0,5$) достигается при его радиусе, равном $1,177\sigma$. Для $R = 3,4\sigma$ $\Pr=0,997$. Если это не так, то радиус круга, при котором достигается вероятность попадания в него, равная 0,5, следует отыскивать с помощью соотношения (П2.33).

Иногда используют понятие **радиальной среднеквадратической ошибки (Distance Root Mean Square (DRMS))**

$$DRMS = \sqrt{\sigma_1^2 + \sigma_2^2}. \quad (\text{П2.36})$$

В зависимости от значений матрицы ковариаций или параметров эллипса этой величине соответствует 65–68% попадания в круг такого радиуса. Нередко используют удвоенную радиальную среднеквадратическую ошибку (2DRMS). Ей соответствует вероятность попадания в круг радиуса, равного удвоенной радиальной ошибке. Точное значение вероятности зависит от конкретных соотношений дисперсий и коэффициента корреляции, а примерная ее величина определяется как $\text{Pr}=0,95$.

Приложение 3

П3. Обыкновенные дифференциальные уравнения

Обыкновенным дифференциальным уравнением в форме Коши называется уравнение вида [6, 14, 26, 32]

$$\dot{x}(t) = F(t, x(t), u(t)), \quad (\text{П3.1})$$

где $x(t) = (x_1(t), \dots, x_n(t))^T$ – n -мерный вектор; $u(t) = (u_1(t), \dots, u_p(t))^T$ – p -мерный вектор входного воздействия, под которым может пониматься как управление, так и возмущение; $F(\cdot) = (F_1(\cdot), \dots, F_n(\cdot))^T$ – n -мерная в общем случае нелинейная вектор-функция. Вектор $x(t) = (x_1(t), \dots, x_n(t))^T$ называют **фазовым вектором**.

В тех случаях, когда $F(t, x(t), u(t)) = F(x(t), u(t))$ не зависит явно от времени, уравнение называется **стационарным**. При наличии такой зависимости уравнение называется **нестационарным**.

При $u(t) \equiv 0$ уравнение называют **однородным**, а при $u(t) \neq 0$ – **неоднородным**.

Уравнение

$$\dot{x}(t) = F(t)x(t) + G(t)u(t), \quad (\text{П3.2})$$

в котором $F(t), G(t)$ – матрицы соответствующей размерности, называется **линейным дифференциальным уравнением**.

Если матрицы $F(t), G(t)$ зависят от времени, то уравнение называется **нестационарным линейным дифференциальным уравнением**.

Если матрицы F, G не зависят от времени, т.е.

$$\dot{x}(t) = Fx(t) + Gu(t), \quad (\text{П3.3})$$

уравнение называется **стационарным линейным дифференциальным уравнением**.

Решением дифференциального уравнения называется такая функция времени $x(t)$, $x(t_0) = x_0$, подстановка которой в исходное дифференциальное уравнение превращает его в тождество. Значение функции $x(t_0) = x_0$ в начальный момент называется **начальным условием**.

Фундаментальной, или переходной, матрицей системы уравнений (П3.2) называется такая матрица, которая удовлетворяет уравнению

$$\frac{d\Phi(t, t_0)}{dt} = F(t)\Phi(t, t_0) \quad (\text{ПЗ.4})$$

при начальном условии вида $\Phi(t_0, t_0) = E$.

Общее решение дифференциального уравнения (ПЗ.2), удовлетворяющее начальному условию $x(t_0) = x_0$, имеет вид [6, с. 507; 14, с. 42]

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^t \Phi(t, \tau)G(\tau)u(\tau)d\tau. \quad (\text{ПЗ.5})$$

Первое слагаемое представляет общее решение линейного однородного нестационарного дифференциального уравнения

$$\dot{x}(t) = F(t)x(t).$$

Второе слагаемое является **частным решением** дифференциального уравнения (ПЗ.2) при $x(t_0) = 0$. Частным такое решение называется потому, что оно соответствует нулевым начальным условиям и зависит от конкретного (частного) вида входного воздействия $u(t)$.

Для того чтобы показать, что (ПЗ.5) является решением (ПЗ.3), необходимо воспользоваться следующим правилом [14, с. 114]:

$$\frac{d}{dt} \int_0^t g(t, \tau)d\tau = g(t, t) + \int_0^t \frac{d}{dt} g(t, \tau)d\tau. \quad (\text{ПЗ.6})$$

Дифференцируя (ПЗ.8) и учитывая, что

$$\dot{\Phi}(t, t_0) = F(t)\Phi(t, t_0),$$

получаем

$$\begin{aligned} \dot{x}(t) &= F(t)\Phi(t, t_0)x(t_0) + \Phi(t, t)G(t)u(t) + \int_{t_0}^t F(t)\Phi(t, \tau)G(\tau)u(\tau)d\tau = \\ &= F(t)x(t) + G(t)u(t). \end{aligned}$$

Для стационарных уравнений (ПЗ.3) фундаментальная матрица зависит от разности аргументов и представляет собой **матричную экспоненту** для F

$$\Phi(t - t_0) = e^{F(t-t_0)} = \sum_{i=0}^{\infty} \frac{1}{i!} F^i (t - t_0)^i. \quad (\text{ПЗ.7})$$

Поскольку $F e^{F(t-t_0)} = F \sum_{i=0}^{\infty} \frac{1}{i!} F^i (t - t_0)^i$, легко убедиться в том, что $e^{F(t-t_0)}$

удовлетворяет уравнению типа (ПЗ.4).

Таким образом, общее решение стационарного линейного дифференциального уравнения (ПЗ.3) может быть записано как

$$x(t) = e^{F(t-t_0)}x(t_0) + \int_{t_0}^t e^{F(t-\tau)}Gu(\tau)d\tau. \quad (\text{ПЗ.8})$$

Рассмотрим линейное матричное дифференциальное уравнение

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + Q(t). \quad (\text{ПЗ.9})$$

Нетрудно убедиться, что общим решением этого уравнения является матрица

$$P(t) = \Phi(t, t_0)P(t_0)\Phi^T(t, t_0) + \int_{t_0}^t \Phi(t, \tau)Q(\tau)\Phi^T(t, \tau)d\tau, \quad (\text{ПЗ.10})$$

где $\Phi(t, t_0)$ – фундаментальная матрица для уравнения $\dot{x}(t) = F(t)x(t)$.

Действительно, осуществляя дифференцирование этого соотношения, с учетом (ПЗ.6) получаем

$$\begin{aligned} \dot{P}(t) &= F(t)\Phi(t, t_0)P(t_0)\Phi^T(t, t_0) + \Phi(t, t_0)P(t_0)\Phi^T(t, t_0)F^T(t) + \\ &Q(t) + \int_{t_0}^t F(t)\Phi(t, \tau)Q(\tau)\Phi^T(t, \tau)d\tau + \int_{t_0}^t \Phi(t, \tau)Q(\tau)\Phi^T(t, \tau)F(t)^T d\tau. \end{aligned}$$

Принимая во внимание (ПЗ.10), получаем (ПЗ.9). Для стационарного случая решение (ПЗ.10) имеет вид

$$P(t) = e^{F(t-t_0)}P(t_0)e^{F^T(t-t_0)} + \int_{t_0}^t e^{F(t-\tau)}Qe^{F^T(t-\tau)}d\tau. \quad (\text{ПЗ.11})$$

ЛИТЕРАТУРА

1. Анучин О.Н., Емельянец Г.И. Интегрированные системы ориентации и навигации для морских подвижных объектов: Изд. 2-е, переработ. и дополн. / Под общ. ред. В.Г.Пешехонова. – СПб.: ЦНИИ «Электроприбор», 2003. – 389 с.
2. Бартенев В.А., Гречкосеев А.К., Козорез Д.А., Красильщиков М.Н., Пасынков В.В., Себряков Г.Г., Сыпало К.И. Современные и перспективные информационные ГНСС-технологии в задачах высокоточной навигации. Москва, 2014. – 192 с.
3. Беллман Р. Введение в теорию матриц. – М.: Наука, 1969, 368 с.
4. Белоглазов И.Н., Джанджгава Г.И., Чигин Г.П. Основы навигации по геофизическим полям. – М.: Наука, 1985. – 328 с.
5. Болдин В.А., Zubинский В.И., Зурабов Ю.Г. и др. Глобальная спутниковая радионавигационная система ГЛОНАСС. – М.: ИПРЖР, 1998. – 400 с.
6. Брайсон А.Е., Хо Ю Ши. Прикладная теория оптимального управления. – М.: Мир, 1972. – 544 с.
7. Ван Трис Г. Теория обнаружения оценок и модуляции. Т. 1. Теория обнаружения, оценок и линейной модуляции. – М.: Сов. радио, 1972. – 744 с.
8. Веремеенко К.К., Желтов С.Ю., Ким Н.В., Себряков Г.Г., Красильщиков М.Н. Современные информационные технологии в задачах навигации и наведения беспилотных маневренных летательных аппаратов. – М.: Физико-математическая литература, 2009. – 556 с.
9. Ганмахер Ф.Р. Теория матриц. – М.: Физико-математическая литература, 2004. – 560 с.
10. Груздев Н.М. Оценка точности морского судовождения. – М.: Транспорт, 1989. – 191 с.
11. Дмитриев С.П. Высокоточная морская навигация. – Л.: Судостроение, 1991. – 222 с.

12. **Емельянец Г.И., Степанов А.П.** Интегрированные инерциально-спутниковые системы ориентации и навигации. / Под общей редакцией академика РАН Пешехонова В.Г. – СПб.: ГНЦ РФ АО «Концерн «ЦНИИ «Электроприбор», 2016. – 394 с.
13. **Ивановский Р.И.** Теория вероятностей и математическая статистика. Основы, прикладные аспекты с примерами и задачами в среде Mathcad. СПб.: БХВ-Петербург, 2008. – 528 с.
14. **Корн Г., Корн Т.** Справочник по математике для научных работников. – М.: Наука, 1973. – 832 с.
15. **Красовский А.А., Белоглазов И.Н., Чигин Г.П.** Теория корреляционно-экстремальных навигационных систем.– М.: Наука, 1979.– 448 с.
16. **Левин Б.Р.** Теоретические основы статистической радиотехники. Т 2. – М.: Сов. радио, 1966.
17. **Матвеев В.В., Распопов В.Я.** Основы построения бесплатформенных инерциальных навигационных систем. / По общей редакцией д.т.н Распопова В.Я. – СПб.: ГНЦ РФ ОАО «Концерн «ЦНИИ «Электроприбор», 2009. – 280 с.
18. **Медич Дж.** Статистически оптимальные линейные оценки и управление. – М.: Энергия, 1973. – 440 с.
19. **Михайлов Н.В.** Автономная навигация космических аппаратов при помощи спутниковых радионавигационных систем. – СПб.: Политехника, 2014. – 362 с.
20. **Ривкин С.С., Ивановский Р.И., Костров А.В.** Статистическая оптимизация навигационных систем. – Л.: Судостроение, 1976. – 280 с.
21. **Сейдж Э., Мелс Дж.** Теория оценивания и ее применение в связи и управлении. – М.: Связь, 1976. – 496 с.
22. **Соколов С.В., Погорелов В.А.** Стохастическая оценка, управление и идентификация в высокоточных навигационных системах. – М.: Физико-математическая литература, 2016. – 264 с.
23. **Соловьев Ю.А.** Спутниковая навигация и ее приложения. – М.: ЭКО-ТРЕНЗ 325, 2003. – 325 с.
24. **Степанов О.А.** Применение нелинейной фильтрации в задачах обработки навигационной информации: Изд. 3-е. – СПб.: ЦНИИ «Электроприбор», 2003. – 370 с.
25. **Степанов О.А.** Основы теории оценивания с приложениями к задачам обработки навигационной информации. Ч. 1. Введение в теорию оценивания. – СПб.: ГНЦ РФ ОАО «Концерн «ЦНИИ «Электроприбор», 2010. – 509 с.
26. **Степанов О.А.** Основы теории оценивания с приложениями к задачам обработки навигационной информации. Часть 2 Введение в тео-

- рию фильтрации. – СПб.: ГНЦ РФ ОАО «Концерн «ЦНИИ «Электроприбор», 2012. – 417 с.
- 27. Степанов О.А.** Основные подходы и методы решения прикладных задач обработки измерительной информации. Материалы XVI конференции молодых учёных «Навигация и управление движением», 2014. С.12-35.
 - 28. Щербинин В.В.** Построение инвариантных корреляционно-экстремальных систем навигации и наведения летательных аппаратов. – М.: МГТУ им. Н.Э. Баумана, 2011. – 230 с.
 - 29. Ярлыков М.С.** Статистическая теория радионавигации. - М.: Радио и связь, 1985.
 - 30. Ярлыков М.С.** Марковской теория оценивания в радиотехнике. – М.: Сов. радио, 2004. – 461 с.
 - 31. Brown R.G., Hwang P.Y.S.** Introduction to Random Signal and Applied Kalman Filtering with Matlab Exercises and Solutions. 3rd Ed. John Wiley and Sons, 1997.
 - 32. Gelb A.** Applied Optimal Estimation. M.I.T. Thirteenth printing // Press, Cambridge, MA, 1994.
 - 33. Gibbs, Bruce P.** Advanced Kalman Filtering, Least-Squares and Modeling: A Practical Handbook, John Wiley&Sons, Inc., 2011.
 - 34. Grewal S.M., Andrews A.P.** Kalman Filtering. Theory and Practice Prentice Hall, New Jersey 1993, 382 p.
 - 35. Kalman R.E.** A New Approach to Linear and Filtering Prediction Problems. Trans. ASME, J. Basic Eng., 1960, vol. 82 D.
 - 36. Kalman R.E., Bucy R.S.** New Results in Linear Filtering and Prediction Theory, Transactions of the ASME - Journal of Basic Engineering, Vol. 83: pp. 95-107 (1961). Калман Р.Е., Бьюси Р.С. Новые результаты в теории линейной фильтрации и предсказания // Теоретические основы инженерных расчетов . – 1961. – N 1. – Сер. Д.
 - 37. Simo Särkkä** Bayesian Filtering and Smoothing. Cambridge University Press, 2013. P. 256.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

А

- алгоритм инвариантный 83
 - итерационный 40
 - небайесовский 49
 - неинвариантный 85
 - оптимальный 134
 - - линейный 71, 73
 - - - рекуррентный 117
 - оценивания 9
 - рекуррентный 90
 - с локальными итерациями 40

В

- вектор 157
 - функция 166
 - случайный 173
 - - гауссовский 175
 - центрированный 174
 - собственный 163
 - состояния 13, 102, 143
 - фазовый 179
- векторы случайные независимые 174
 - - некоррелированные 174
 - - ортогональные 174
 - совместно гауссовские 176
- величина случайная 169
 - - гауссовская 171
 - - нормальная 171
 - - центрированная 170
- величины случайные независимые 174
- величины случайные некоррелированные 174
 - - ортогональные 174
- выбеливание 128

Г

градиент матричный 166

Д

детерминант 161

диагонализация матриц 163

дисперсия 96, 170

- апостериорная 36

- априорная 36

дополнение алгебраическое 161

З

задача анализа точности 14, 50, 67, 122, 134

- интерполяции 113

- калибровки 11

- комплексной обработки измерений 81

- корреляционно-экстремальной навигации 15

- навигации по геофизическим полям 15

 - - - спутниковым данным 18

 - - - точечным ориентирам 16

- начальной выставки 12

- определения параметров гармонического сигнала 15

 - - координат 16

 - - фазы 14

 - - частоты 15

- оценивания 9

 - - коэффициентов полинома 10

 - - линейная 13

 - - нелинейная 14, 19

задача прогноза 113

- синтеза алгоритма 14, 50, 67, 122, 134

- сглаживания 113

- слежения за подвижным объектом 23

- траекторного измерения 23

- фазовой автоподстройки частоты 15

- фильтрации 113, 118, 145

 - - для непрерывного времени 145

- экстраполяции 113

значение матрицы собственное 162

значение предельное 173

И

измерения 8

измерения корректирующие 80

изолиния положения 17

инвариантность ошибок оценивания 34

интеграл матрицы 165
интенсивность белого шума 141
интервал корреляции 138

К

канал ковариационный 121, 147
- оценочный 121, 147
коэффициент корреляции 175
- - нормированный 176
- усиления 121

критерий оценивания 9
- среднеквадратический 50
круг равных вероятностей 178

Л

лемма об обращении матриц 167
линеаризация функции 20
линия положения 21

М

матрица 157
- блочная 159
- блочно-диагональная 159
- взаимной корреляции 174
- вырожденная 161
- диагональная 158
- динамики 102, 143
- единичная 158
- информационная Фишера 52
- квадратная 157
- ковариаций 40, 120, 173
- - апостериорная 36
- - - безусловная 66, 134
- - - текущая 134
- - - условная 66, 134
- - априорная 36
- - действительная 41
- - расчетная 40
- корреляционная 97
- коэффициентов усиления 121
- невырожденная 161
- неотрицательно определенная 164
- несингулярная 161
- нижней границы точности 52, 68
- обратная 162
- ортогональная 162

- отрицательно определенная 164
- переходная 179
- положительно определенная 164
- порождающих шумов 102, 143
- симметричная 158
- сингулярная 161
- транспонированная 158
- фундаментальная 179
- матрица треугольная 158
 - - верхняя 158
 - - нижняя 158
- матрицы коммутативные 160
 - подобные 163
- метод классический 49
 - максимума функции правдоподобия 52
 - наименьших квадратов 25
 - - - взвешенных 26
 - - - модифицированный 26
 - - - обобщенный 26
 - небайесовский 49
 - статистических испытаний 58
- минор матрицы 162
 - главный 162
- многочлен характеристический 162
- множитель нормирующий 65
- модель измерений 8
 - оцениваемых параметров 8
- модуль вектора 160
- момент 170
- Н**
- невязки измерений 25, 127
- независимость ошибок оценивания 34
 - случайных величин 174
- неравенства матричные 164
- неравенство Рао–Крамера 51, 68
- норма вектора 160
- О**
- ожидание математическое 96, 170
- определитель 161
- ортогональность ошибки 67
- отклонение среднеквадратическое 170
 - стандартное 170

- оценка байесовская 66, 68, 71, 72, 134
 - максимального правдоподобия 53
 - небайесовская 52, 56
 - несмещенная 51, 56, 67, 71, 114
 - оптимальная 66, 72, 134
 - - в среднеквадратическом смысле 50, 71, 114, 134
 - - линейная 71
 - прогноза 120
 - параметра 9
 - состоятельная 51
 - с минимальной дисперсией 51, 56, 71, 114
 - эффективная 52, 68

ошибка круговая вероятная 178

- оценки 13, 67
- предельная 173
- прогноза 127
- систематическая 17
- среднеквадратическая 170
 - - радиальная 178
- фильтрации 127

ошибки измерения 8

- оценивания 9

П

параметр мешающий 18

- оцениваемый 8

параметры эллипса ошибок 177

плотность апостериорная 63, 66

- априорная 63
- вероятности перехода 100
- переходная 100
- спектральная 139
 - - средней мощности 139
- условная 64
 - - гауссовская 64

подход байесовский 63, 66

- детерминированный 25
- классический 49
- небайесовский 49
- нестохастический 25

полуоси эллипса ошибок 177

последовательность винеровская 104

- инновационная 128
- марковская 100

- - гауссовская 101
- - стационарная 106
- обновляющаяся 128
- с независимыми приращениями 104
- с некоррелированными приращениями 103
- случайная 94
 - - векторная 96
 - - гауссовская 99
 - - нестационарная 97
 - - стационарная 97
 - - - в узком смысле 97
 - - - в широком смысле 97
 - - центрированная 98
- экспоненциально-коррелированная 108
- потери средние байесовские 66
- правило трех сигм 173
- преобразование подобия 163
- проблема адекватности 41
 - собственных чисел 163
- производная вектор-функции 166
 - матрицы 165
 - скалярной функции по матрице 166
 - скалярной функции по векторному аргументу 165
- процесс случайный 138
 - экспоненциально-коррелированный 138

Р

- размерность матриц 160
- ранг матрицы 162
- распределение гауссовское 171
 - нормальное 171
- расширение вектора состояния 18
- режим камеральной обработки 113
 - реального времени 113
 - фильтрации установившийся 130
- решение дифференциального уравнения 179
 - - - общее 180
 - - - частное 180

риск байесовский 66

С

- свойство ортогональности 67
- символ Кронекера 164
- система нормальных уравнений 26
- след матрицы 159

сложение матриц 159
спектр матрицы 163
- энергетический 139
стандарт 170
строка 157
схема обработки децентрализованная 89
- - инвариантная 83
- - неинвариантная 85
- - рекуррентная 90
- - централизованная 88

Т

точность потенциальная 68, 123, 134
- предельно достижимая 52, 68
тренд квадратичный 11
- линейный 10

У

угол дирекционный 177
умножение блочных матриц 160
- матриц 159
- матрицы на число 159
уравнение Винера–Хопфа 114
- дифференциальное линейное 179
- - нелинейное 147
- - обыкновенное в форме Коши 179
- неоднородное 179
- нестационарное 179
- однородное 179
- правдоподобия 53
- Рикатти 147
- стационарное 179
- характеристическое 162
условие наблюдаемости 30
- начальное 179
- нормировки 173
- согласованности 173
- стационарности 144
- стохастической эквивалентности 150

Ф

фактор геометрический 47
фильтр Винера 131
- инвариантный 82
- Калмана 120
- - для дискретного времени 120

- Калмана-Бьюси 145
 - - - непрерывный 146
- комплементарный 82
- формирующий 101, 143
- форма квадратичная 164
 - - неотрицательно определенная 164
 - - отрицательно определенная 164
 - - положительно определенная 164
- формула Байеса 64
 - обращения блочных матриц 167
 - умножения плотностей вероятности 63
- функция взаимно-корреляционная 96
 - корреляционная 96
 - от матрицы 165
 - плотности распределения вероятностей 170
 - - - - гауссовская 171
 - - - - нормальная 171
 - - - - совместная 173
 - потерь 50
 - - квадратичная 50
- функция правдоподобия 52
 - - логарифмическая 53
 - распределения вероятностей 169
 - - - совместная 173

Ш

- шпур 159
- шум дискретный белый 98
 - - - гауссовский 99
 - белый 141
 - порождающий 102, 143
 - системы 143

Э

- экспонента матричная 165, 180
- эллипс ошибок среднеквадратический 177
 - стандартный 177
- эллипсы равных вероятностей 177

Миссия университета – генерация передовых знаний, внедрение инновационных разработок и подготовка элитных кадров, способных действовать в условиях быстро меняющегося мира и обеспечивать опережающее развитие науки, технологий и других областей для содействия решению актуальных задач.

ИСТОРИЯ КАФЕДРЫ ИНФОРМАЦИОННО-НАВИГАЦИОННЫХ СИСТЕМ

Кафедра «Информационно-навигационные системы» является одной из старейших в университете.

В 30-е гг. XX века отечественная промышленность приступила к созданию новых навигационных приборов – гирокомпасов, лагов, эхолотов. В связи с этим в 1934 г. рядом предприятий и организаций было внесено предложение во Всесоюзный совет народного хозяйства (ВСНХ) СССР о подготовке необходимых инженерных кадров. В сентябре 1937 г. была создана кафедра навигационных приборов. С 1937/38 учебного года на кафедре была начата подготовка специалистов.

В разработке профиля кафедры принимали участие видные ученые и специалисты промышленности и флота страны, в том числе академик А.Н. Крылов, профессора Б.И. Кудревич и К.С. Ухов, доцент Д.Н. Иконников, инженеры М.М. Блюк, М.А. Шиф и многие другие.

Первым заведующим кафедрой стал профессор Л.П. Шишелов, занимавший в то время должность декана факультета «А» (точной механики).

В том же 1937 г. на кафедре организована лаборатория, заведующим которой был назначен П.А. Ильин. Он окончил ЛИТМО в 1931 г. в числе первых выпускников и, приобретя опыт работы на производстве, в 1936 г. поступил в аспирантуру. С этого момента вся его деятельность была связана с кафедрой и направлена на совершенствование учебного процесса. Благодаря его активной деятельности и связям с промышленностью лаборатория была оснащена современными на тот период навигационными приборами, а учебный процесс обеспечен необходимыми учебными пособиями. В 1940 г. состоялся первый выпуск инженеров по навигационным приборам.

Во время Великой Отечественной войны в период блокады Ленинграда на время эвакуации ЛИТМО лабораторное оборудование кафедры оставалось в осажденном городе и было сохранено. Поэтому практически сразу после реэвакуации института в 1944 г. на кафедре были возобновлены занятия.

С 1945 г. по 1962 г. кафедру возглавлял крупный ученый в области навигации, гидрографии, штурманских приборов, доктор технических наук, профессор К.С. Ухов, который при поддержке профессора Б.И. Кудревича

основал научно-педагогическую школу «Навигационные приборы». При нем в 1956 г. в связи с изменением номенклатуры специальностей кафедра была переименована в кафедру гироскопических и навигационных приборов. Чтобы удовлетворить потребности промышленности, в эти годы на кафедре обучалось от 4 до 5 групп студентов, включая группы вечернего образования.

С 1963 г. по 1970 г. кафедрой заведует (по совместительству) главный инженер Ленинградского научно-производственного объединения «Азимут», лауреат Ленинской премии, доктор технических наук, профессор С.Ф. Фармаковский. В этот период коллектив кафедры значительно укрепил и расширил связи с промышленностью. В гироскопическом приборостроении получают развитие газовые опоры как в осях подвеса, так и в качестве главных опор ротора гироскопа. На кафедре гироскопических и навигационных приборов разрабатываются теоретические основы расчета и проектирования газодинамических подшипников. Результаты использованы в НПО «Азимут» и других организациях.

Во второй половине 60-х гг. под руководством профессора П.А. Ильина разворачиваются работы по исследованию и разработке лазерных гироскопов. Для проведения работ по договорам с Московским институтом электроавтоматики при кафедре в 1966 г. создается отраслевая лаборатория квантовой гироскопии под научным руководством с.н.с. Г.М. Кузнецова.

В 1969 г. в издательстве «Машиностроение» публикуется монография д.т.н., профессора М.А. Сергеева «Наземные гироскопы». В 1970 г. он избирается по конкурсу заведующим кафедрой и руководит ею до 1982 г. В промышленности ведутся работы по созданию космических летательных аппаратов, сверхзвуковых самолетов, атомных подводных лодок, межконтинентальных ракет. В этих условиях кафедра расширяет профиль подготовки специалистов и в 1973 г. переименовывается в кафедру бортовых приборов управления.

С 1982 г. по 1991 г. кафедру возглавляет профессор В.А. Каракашев, создавший совместно со своими учениками О.Н. Анучиным и Г.И. Емельянцевым (впоследствии доктора наук) научно-педагогическую школу «Инерциальные навигационные системы». На кафедре разрабатываются новые учебные планы, совершенствуется преподавание дисциплин, большое внимание уделяется внедрению в учебный процесс вычислительной техники. Преподаватели кафедры проходят переподготовку и повышение квалификации в области вычислительной техники. На кафедре ведутся научно-исследовательские работы совместно и по заказам навигационно-гидрографического института, ЦНИИ «Электроприбор», завода «Баррикада» в Волгограде.

С 1991 г. кафедру возглавляет директор ГНЦ РФ ЦНИИ «Электроприбор» (в настоящее время АО «Концерн «ЦНИИ «Электроприбор»), лауреат Ленинской и Государственной премий, в то время член-корреспондент, впоследствии академик РАН В.Г. Пешехонов. Под руководством В.Г. Пешехонова кафедра ориентируется на подготовку специалистов в

области создания приборов и систем, измеряющих перемещения, скорости и ускорения морских подвижных объектов и образующих интегрированные навигационно-управляющие комплексы судов и кораблей различного назначения, а также комплексы пространственной ориентации и стабилизации информационно-измерительной аппаратуры на их борту. В соответствии с произошедшими изменениями в научной и педагогической деятельности кафедры в апреле 2003 г. кафедра получила свое современное наименование – она была переименована в кафедру информационно-навигационных систем.

По инициативе академика В.Г. Пешехонова, начиная с 2004/05 учебного года, студенты и преподаватели кафедры активно включаются в межвузовскую программу интенсивной подготовки специалистов, которая реализуется на базе ЦНИИ «Электроприбор». С этой целью между ЦНИИ «Электроприбор» и тремя ведущими университетами города (СПб ГЭТУ «ЛЭТИ», СПбГУ ИТМО и СПб ГУАП) заключается и успешно реализуется соглашение о межвузовской углубленной подготовке инженерных кадров и проведении совместных научных исследований в области навигации на базе ЦНИИ «Электроприбор». В сентябре 2008 г. кафедра преобразуется в базовую при ЦНИИ «Электроприбор».

За работы в области объединения усилий вузов при подготовке специалистов коллективу, в состав которого вошли профессора СПб ГЭТУ «ЛЭТИ» Ю.В. Филатов и Ю.А. Лукомский, а также д.т.н. О.А. Степанов (ЦНИИ «Электроприбор»), в 2010 г. присуждается премия Правительства Санкт-Петербурга «За выдающиеся достижения в области высшего и среднего профессионального образования за 2010 г.» в области интеграции образования, науки и промышленности.

Укрепляется международное сотрудничество, в частности с международной общественной организацией «Академия навигации и управления движением», учрежденной в 1995 г. как общественное объединение ведущих ученых в области навигации и управления движением и в настоящее время насчитывающей более 400 членов, работающих в различных университетах и организациях десяти стран (России, Украины, Германии, США, Франции, Китая, Южной Кореи, Белоруссии, Турции, Индии). Ведущие профессора кафедры читают лекции в Китае и Финляндии, активно выступают с докладами на авторитетных международных и российских конференциях.

В 2011 г. в ЦНИИ «Электроприбор» открывается новый комплекс учебных помещений, включающий лабораторию, компьютерный класс и зону отдыха для студентов. Лаборатория, оснащенная современным оборудованием для проведения лабораторных работ, фактически является центром коллективного пользования для ведущих вузов города, готовящих специалистов в области навигации и управления.

В 2012 г. кафедра информационно-навигационных систем отметила свой 75-летний юбилей.

В настоящее время кафедра активно участвует в реализации Программы повышения конкурентоспособности Университета ИТМО, в частности, в 2014 г. в рамках открытого конкурса создана международная научная лаборатория «Интегрированные системы ориентации и навигации» (руководители: Степанов О.А. – профессор, д.т.н., заместитель заведующего кафедрой; Троммер Г.Ф. – профессор, директор Института оптимизации систем Технологического университета Карлсруэ, Германия). Ее сотрудники проводят исследования по различным направлениям, связанным с научной тематикой кафедры. За последнее время выполнены работы в рамках семи грантов, поддержанных Российским фондом фундаментальных исследований, и работы по крупному проекту, поддержанному Российским научным фондом, на тему «Решение проблемы высокоточных измерений гравитационного поля в труднодоступных районах Земли с подвижных объектов». В 2014–2016 гг. по результатам научных исследований опубликовано значительное количество печатных работ, в том числе более 120 публикаций, рецензируемых в базах Scopus и Web of Science.

В настоящее время на кафедре созданы все необходимые условия для освоения важной профессии, связанной с созданием информационно-навигационных систем, и проведением в этой области научных исследований на современном уровне.

Олег Андреевич Степанов

МЕТОДЫ ОБРАБОТКИ НАВИГАЦИОННОЙ ИЗМЕРИТЕЛЬНОЙ ИНФОРМАЦИИ

Учебное пособие

В авторской редакции
Редакционно-издательский отдел Университета ИТМО

Редактор Т.П. Уткина
Зав. РИО Н.Ф. Гусарова

Подписано к печати 21.12.2016

Заказ № 3801

Тираж 126 экз.

Отпечатано в учреждении «Университетские телекоммуникации»

Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49